

Towards the Ethical Governance of Smart Society

Mark Hartwood, Barbara Grimpe, Marina Jirotko, and Stuart Anderson

1 Introduction

Smart Society¹ is a term coined by an EU funded Integrating Project (IP) of the same name that aims to capture how contemporary techno-social trends can be harnessed towards solving challenges facing modern society. The “Smart” alludes to the enabling capabilities of innovative, social, mobile and sensor based technologies that in various way are envisaged to create more productive alignments between (growing) demand and (constrained) resources across a number of sectors and application domains.² A key example of this is how to meet growing care needs with diminishing resources as the number of elderly people grows as a proportion of the overall population [9]. While the challenges of urban life form the test bed for the Smart Society project, it is likely to become increasingly relevant in other domains such as finance [6].

¹Smart Society (FP7/2007–2013) Grant agreement n. 600854. <http://www.smart-society-project.eu/>

²Although this might be seen as a tale of improved matchmaking, there are also important nuances in the ways that social computing transforms the sorts of demands, goods and services that are in play. Thus, for example, social platforms can make visible the “hidden” care demands of elders, and also the “hidden” skills of neighbors, and create a market place in which these may be traded (e.g. <http://ce.livingitup.org.uk/hidden-talents-scotland/>).

M. Hartwood (✉) • B. Grimpe • M. Jirotko
Department of Computer Science, Oxford University, Oxford, UK
e-mail: mark.hartwood@cs.ox.ac.uk; barbara.grimpe@cs.ox.ac.uk; marina.jirotko@cs.ox.ac.uk

S. Anderson
School of Informatics, University of Edinburgh, Edinburgh, UK
e-mail: stuart.anderson@ed.ac.uk

Smart Society is partly inspired by the idea of the “Smart City”, a multifaceted concept [19] that recognizes the benefits of urban living but also the strains that are developing on existing infrastructures and resources due to urban growth. The vision relates how cities made “smart” will be more productive, more sustainable, and pleasanter places to live. One aspect of Smart Cities concerns augmenting service infrastructures (such as transport, energy, health and so on) with sensor-based digital technologies able to visualize patterns of service delivery and use stretching across space and time and with a high degree of fidelity [28]. The idea is that service operators can utilize this information to make efficiency savings by tailoring provision to match demand, and by shaping demand through use of incentives or other motivating feedback mechanisms. At the same time shared resources can be used more effectively if users are aware of the global state of the resource and able to coordinate between themselves about how the resource might be utilised. For example, road users can chose an alternative route if they are made aware of patterns of congestion, and if given the right tools they can offer each other advice based upon their local perspective and knowledge.

Smart Society extends Smart City thinking in a number of ways, for example, by including the ideas of:

Hybrid computing How people and machines working together create new sorts of problem solving capability, for example, as in the “wisdom of the crowds”—but also stemming from peoples’ everyday use of their mobile connection to data, algorithms and social networks to solve problems.

Adaptivity Bringing to the appropriate sub-collective to bear to solve a given problem; and

Learning Accreting knowledge of how the system responds to different circumstances and using that to drive subsequent rounds of adaptation.

Smart Society is founded on the idea of “collectives”—groups of people linked by a common identity yet having diverse skills, needs and values. On this definition an example of a collective may be “road users”, incorporating several sub-collectives of pedestrians, cyclists, motorists, and bus users whose common identity is established by their dependence on a shared roads infrastructure, but whose needs, values and skills will vary considerably between these different “categories” of use. In Smart Society, collectives are seen as a source of expertise that may be accessed and exploited. At the same time they are consumers of resources whose patterns of consumption can be shaped by appropriate interventions. Diversity within collectives on the one hand provides a resource pool to enable a collective develop a range of responses to a situation, but on the other it is also source of friction and contention. Taking these elements together, the socio-technical entity powering the Smart Society vision is referred to as a Hybrid and Diversity Aware Collective Adaptive System, or HDA-CAS³ for short.

³On the occasions in this chapter when we refer to CASs we are considering Collective Adaptive Systems more generally and not only the Hybrid, Diversity-Aware sort.

Such collectives already exist and are routinely formed on a more or less ad-hoc basis, for instance, through social platforms such as twitter. An example is a collective anchored via the #UKSnow hashtag which collaboratively monitors the impact of winter weather thereby inducing individuals to adapt by adjusting their travel decisions.⁴ Smart Society aims to engineer more powerful CASs that behave in more predictable ways, that penetrate further into critical or economically significant city infrastructures and services, and that implicate multiple and diverse user constituencies. These elements of scale and intentional design bring with them a series of risks, including those of naively fixing a narrow range of values and of overlooking the need to create governance structures able to evolve and also to mediate between diverse and conflicting value systems.

Thus a key aspect of Smart Society is how to govern them in ways that permit conflicting and diverse perspectives to co-exist within a large-scale evolving CAS, and this question broadly frames the work we present in this chapter. Smart Society from its inception aimed for a multidisciplinary approach to engineering HDA-CASs that incorporates social science understandings of collectives, and ethical orientations to research and innovation [8]. Triggered by a series of recent EU initiatives and research projects [30, 35] the role of the authors in this project is to bring Responsible Research and Innovation practices [42] to a range of technologies, including CASs, in order to shape their impact upon privacy and other social values. An important aspect of this has been to work towards a framework for the *ethical governance* of HDA-CASs.

Our approach has been to couple a conceptual exploration of governance to a social science enquiry into domains where CASs are envisaged, or where CAS-like systems already exist. This unpacking of governance has lead us to the view that the ways in which *CAS might be regulated* (to operate in socially acceptable ways) are quite intimately tied to the ways in which *CASs themselves aim to regulate collectives* (for example, through targeted incentives). Another way of saying this is that CASs gain their effect by instantiating particular forms of social regulation, and moderating how this is achieved is key to producing CASs that are sensitive to important social values.

Exerting influence on a CAS's participants is central to Smart Society, as the Description of Work (DOW) for Smart Society makes clear:

“[The aim of the project is] to develop novel incentives, mechanisms and decision-making algorithms able to drive the emergence of desired system-level behaviours in HDA-CASs taking into account the wider information environment and non-incentivised motivations . . . To introduce a programming paradigm and an architecture for the management and control of HDA-CASs in a goal-oriented fashion.” [10].

In a publication giving an overview of Smart Society these sorts of aims are couched in terms of an everyday example using slightly less technical language:

⁴<http://uksnowmap.com/> mashes up #UKSnow tweets and Google Maps to show geographical patterns of reported snowfall, thus providing a sustaining focus for the collective and a mechanism to propagate snow reporting practices through example and a weak obligation to reciprocity.

“From the analysis of sensor data, machines can “understand” (from low level analysis) that a critical traffic situation has arisen. This initiates a hybrid computation that calculates the best incentives to offer different strata in the driving population in order to align driver behaviour with global policy objectives. . . . Incentives will be given to particular target groups depending upon their needs and expectations. People can ignore such suggestions and decide autonomously on what they believe is best for them” [16].

The above description reveals the Smart Society vision to be a complex one that admits diversity and acknowledges conflicting perspectives. The vision allows for autonomy while at the same time seeking to influence with incentives and persuasive technologies. It aims for a degree of self-regulation by giving participants access to information and resources and broadening their capacity to act, whilst at the same time seeking to impart direction and make wider patterns of behaviour align with centrally defined goals. At first inspection, one might doubt the vision’s coherence, fearing that it contains inherent and insurmountable contradictions (e.g. autonomy versus control, centralised versus self-regulation, individual versus public interests), and yet as we explore the notion of governance more thoroughly, we discover that it is common, perhaps inevitable or necessary, for multiple governance regimes to coexist simultaneously.

The vision for Smart Society articulated in these quotations raises deeply significant social and ethical issues, including:

- who will set the incentive structures or algorithm parameters?
- who gets to set the ultimate direction or goals of the Smart Society—and if this is the State, what new forms of democratic conventions will be needed to control this new and powerful way of implementing policy?
- will we be aware of the machinations of the “unseen hand” that filters the information we see, targets us with incentives and chooses which resources we can access?
- how are the conflicting interests and perspectives of multiple user constituencies mediated?
- and crucially who will be accountable for the effects of the HDA-CAS should things go wrong?

These questions resonate with the general and long-term problem of governing the global knowledge society in a “smart” way [43].

The chapter is organised as follows: First we present our methodology (Sect. 2) and then explore several potential ethical consequences of the Smart Society vision (Sect. 3). This leads us to unpack the concept of governance to better understand the different forms of regulation and their relationship to one another. We observe that modes of governance are not mutually exclusive, but are rather blended in different proportions to achieve different sorts of regulatory effects (Sect. 4). Understanding how forms of social regulation work provides a foundation for understanding the sorts of regulatory effects that HDA-CASs can achieve. It also provides a basis for applying them mindfully, with care and forethought. Finally, we draw upon a “worked example” to show how governance design can be pursued in a way that is sensitive to social values and emerging ethical concerns (Sect. 5).

2 Methodology

It is impossible to study HDA-CAS “in the wild” as the sorts of HDA-CAS envisaged by Smart Society do not yet exist. Thus to understand the implications of HDA-CAS for ethical governance we need to adopt a series of more indirect approaches. To achieve this we have explored:

1. *Emergent ethical issues of contemporary trends in networked, social and mobile computing*: This has principally involved exploring the extensive existing literatures on this topic.
2. *Existing systems or programmes that have some properties in common with HDA-CASs, or that are driven by a similar vision*: Here we have conducted a series of “elite” interviews with powerful stakeholders driving the Smart City agenda. “Elite interviews” aim to explore and learn from the experiences of those in positions of power and influence within a particular arena, be it politics, business, academia or the public sector [1]. This approach allows us to access the accumulated learning accrued from implementing real-life Smart City visions and CAS-like systems. To date we have conducted the following interviews: Senior police officers (2); IT consultant developing SmartCare “apps” (1); Smart City academics (2); Smart City consultants and system integrators (4); Manager of a Regional Intelligent Traffic Management system (1); Civil Servant facilitating Smart Cities programme (1).
3. *User perspectives in contexts corresponding to Smart Society scenarios*: Focus groups with tourists (young travellers) (1); Interviews with “Ride Sharing” scheme participants (8).⁵
4. *Reflective discussions within the Smart Society project itself*: Smart Society project members naturally reflect on the ethical potential of the technologies during co-located and virtual meetings across the project and these are valid and valuable forms of insight.

3 Addressing Ethical Issues in Smart Society

Drawing upon the above empirical work we have found it useful to distinguish between *contextual* and *emergent* ethical issues in relation to CASs. Contextual issues refer to pre-existing ethical sensitivities within a given socio-technical system that reflect interactions between cultural values, supportive infrastructures and system goals. Emergent ethical issues are ones that arise, or are amplified or diminished as a consequence both of reengineering an existing system to function more like a HDA-CAS, or by virtue of the CAS’s evolution. This distinction is important because it enables us to take seriously pre-existing ethical concerns,

⁵In collaboration with our Smart Society partners at Ben Gurion University.

whilst at the same time keeping an open mind as to which ethical issues will assume importance in the future. This awareness of emerging system properties and corresponding ethical issues builds on practice theories of socio-technical order [34, 36]. In the case at hand, it entails an ongoing process for identifying and managing ethical concerns that should function continuously as the HDA-CAS is implemented, as it evolves and as it interacts with wider social and socio-technical systems. Given that ethical concerns are often debatable, conflicting or present as dilemmas then we need to avoid the idea that we can, for the most part, solve ethical problems (cf. [22]). Rather we wish to provide a space for them to be surfaced, negotiated and to enable working compromises to be achieved. We take these processes of identifying and managing ethical concerns to constitute the “ethical governance” of HDA-CASs. We develop some preliminary ideas as to what this governance process should look like and how it intersects with other aspects of CAS governance later on in this document. Here we prime that discussion by drawing upon empirical data to focus on categories of ethical issues that appear relevant to HDA-CASs and Smart City application domains. The intention is to create a sensibility towards relevant ethical concerns, including particular sorts of contemporary or domain specific ethical issues, but also to point to categories of issue attached to wider techno-social trends and anticipated HDA-CAS properties.

3.1 Contextual Issues

A preliminary analysis of interviews and focus groups conducted as to inform this chapter has revealed a variety of pre-existing ethical sensitivities in domains such as social care, tourism and transport. One such example revolves around the safety concerns of those participating in schemes that support “Couch Surfing”⁶ as a source of cheap accommodation, and “Ride Sharing”⁷ as a means towards inexpensive travel. Some female travellers in particular were concerned they may be exposed to risk using these services for instance if they accepted a lift at night alone with a man they did not know. Interviews and focus groups revealed variation in degrees of concern and an array of ad-hoc strategies used to reduce risk. These included: avoiding use of the service altogether; preferring a telephone conversation to arrange the ride to help gain an impression of the driver’s character; keeping a personal record of driver “reputation”; becoming less cautious with experience; and by choosing “safer” rides (e.g. a daytime ride with other passengers). Thus a concerted investigation into a setting can provide valuable insight into important social values that need to be accommodated within the design of a CAS, and yet the

⁶“Couch Surfing”—taking advantage of casual services provided by locals such as offers of accommodation in private homes.

⁷Schemes that allows drivers and commuters to offer and accept lifts and share costs by utilising spare capacity in the cars of those already intending to travel.

process of accommodating social values is often not straightforward. One reason for this is because different social values often compete with each other. In the Ride Sharing schemes, for example, it is hard to balance the need to enhance privacy on the one hand with the need to reveal personal details about drivers and passengers to enhance safety on the other. Layering on properties envisioned for HDA-CASs adds further intricacies to these already complicated situations. An example of this is that users of existing Ride Sharing schemes can choose freely from offers of lifts, whereas Smart Society would use incentives to steer that selection, perhaps to encourage optimal journey times or maximum occupancy. This has the effect of shifting some of the responsibility for choosing a ride to the CAS, implying that if someone should come to harm then liability may be attached to the CAS or its designers. With these complexities in mind our approach is not to attempt a fixed design that roughly satisfies constraints of competing social values as they exist at a point in time, but instead to use an enquiry into social values to inform the design of flexible governance structures that can be renegotiated and modified as circumstances change and as the system evolves. We cover this topic in detail in Sect. 4.1 below

One thing to note from this discussion is that contextual and emerging ethical concerns are not entirely separable. Starting from contextual issues, it is quite natural to then consider how a planned implementation may “mangle” those issues into new types of problem [34]. Thus, understanding existing issues forms the basis for anticipating emerging ethical dilemmas.

3.2 Emerging Issues

Emergence is a key feature of HDA-CASs, and new sorts of ethical dilemmas may arise alongside emerging capabilities and impacts. Forecasting future ethical concerns for evolving, complex, open-ended systems seems a hard task. However, practical methods have been developed towards envisioning a range of alternative possible futures to provide traction for design choices made in the present [17]. These fall under the rubric of “anticipatory governance”, defined as the coupling of foresight and policy to achieve earlier responses to the “unexpected” or emergent consequences of non-linear systems [15]. In this context, foresight is not taken as prediction, but rather as a resource for negotiating possible futures that is informed by combining several sources of knowledge, including: hindsight (i.e. awareness of prior “surprises”), awareness of techno-social trends and dynamics, expertise and perspectives from a range of stakeholders and disciplines, domain overviews, and model based forecasts [15, 17]. Our approach throughout this section has in a modest way been to utilise some of the above anticipation and foresight approaches to understand the implications of governing CASs. For instance, in the remainder of this section we draw out lessons for social values from the accumulated experience of large-scale socio-technical systems with properties similar to CASs that is available from the literature and from our own empirical work. In later sections we

seek to understand how CASs may regulate collectives, and anticipate the different propensities attached to alternate governance regimes. Finally we draw these together with an empirically founded “worked example” (i.e. one drawing upon domain expertise) that considers the governance requirements of a HDA-CAS in a care setting detailed in Sect. 5.

Here we return to possible emergent ethical issues for CASs based upon the sorts of social transformations already wrought by existing complex socio-technical systems:

Social Sorting: CAS that are diversity-aware aim to be sensitive to the mix of capabilities and values present within collectives, and able to stratify populations to target incentives and recruit expertise. However, such an approach is open to undesirable forms of social sorting, identified as the ways that surveillance technologies sift populations and thereby regulate entitlement or access to resources [25].

Representation and transparency: Who decides the global goals a CAS should pursue, and to what extent will participants understand that their behaviour is being directed through the use of incentives and persuasive technologies? Although CASs are envisaged as creating societal benefits, various forms of accountability are needed to ensure such ends are not subverted. It may be suspected that CASs really aim to make life more convenient or lucrative for well-off sponsors, thus certain forms of transparency become needed to preserve confidence and trust.

Direction versus autonomy: The metaphor of “herding sheep” has been used to explore how the behaviour of collectives can be directed⁸, raising the question as to who gets to set the system’s direction—or train and influence the sheepdogs? Similarly the “God of the Smart Society” has been proposed⁹ as an evocative metaphor for the unseen hand guiding the collective’s behaviour, raising the question as to whether a Smart Society should be more paternalistic or more democratic in its constitution?

Incentives and their effects: Attempts to influence human behaviour can result in “perverse outcomes” on those occasions when incentives drive unanticipated and undesired behaviours [37]. This raises issues about monitoring CASs to ensure that their emergent properties are positive and intended. This becomes harder to achieve as the system scales because of the increasing diversity of outcomes and the increasing diversity of views over what outcomes are actually desirable. So although noble intentions are envisaged for CASs such as, reducing traffic congestion or pollution, or creating community goods where none existed previously, defining such intentions will in practice depend upon negotiating between contested perspectives.

⁸“In a similar fashion to herding sheep, the goal is to steer a group of living individuals to comply with our goals.” [2]

⁹By a member of the Smart Society project during a project meeting when the conversation turned to types of ethical concern raised by the project.

CAS boundaries are a further site of ethical concern. Will non-participants be disadvantaged? One can imagine that business owners who depend on passing trade will be upset by changing commuting patterns as drivers participate in a CAS that aims to reduce congestion. Will these “indirect” stakeholders be given a say in how those CASs are configured?

Hybridity within a HDA-CAS aims to blend the capabilities of humans and machines to solve problems either would struggle to solve alone. Questions arise here whether participation is fairly rewarded, whether professional roles are displaced, and how to guard against malicious forms of participation [24, 40].

Flows and mobilities: Attention is needed to the wider impacts of CAS across time and space as they alter flows and mobilities within the proposed Smart City setting. This is because CAS aim to influence the movement of traffic, people, material and immaterial goods, patterns of consumption, transform the knowledge, skills and resources needed to participate in markets, access services and engage in political discourse. With all of these effects there are likely to be winners and losers. As the authors of [18] have argued, increasing the mobility of some stakeholders may entail “immobilities” for other groups.

Automation raises a gamut of issues including the degree of control ceded to algorithms, the redistribution of responsibility and liability (discussed above for the Ride Sharing scenario), the performative shaping of participation (e.g. job applicants aligning their behaviour to the matching algorithm in online job markets such as “Elance”¹⁰), the opaqueness of algorithms and their adaptations, and the filtering effect they have on human experience of the world [12, 20, 21, 23].

Personal integrity: Finally there are a series of values that relate aspects of personal integrity and autonomy such as trust, safety, security and privacy, some of which are discussed above, and others come to the fore in discussions of privacy elsewhere in this volume.

4 Governing Smart Society

This section sets out a simple example to help illustrate principles of governance, their interrelationship and how they are relevant to Smart Society. Our conceptual analysis of governance has led us to the view that the ways in which *CASs might be regulated* (to operate in socially acceptable ways) are quite intimately tied to the ways in which *CASs themselves aim to regulate* collectives (through targeted incentives, for example). In other words, a more thorough understanding of how different forms of governance interact to deliver social regulation helps not only with working out how to design CASs effectively to influence how resources are used, but also to see how this can be done in ways that are ethically sensitive to different contexts. The example we present in this section concerns regulation of public

¹⁰www.elance.com



Fig. 1 Example of a “speed bump”

highways to ensure they function effectively as a shared resource. We explore how speed bumps, sometimes known as “sleeping policeman”, are employed to regulate traffic speed. We illustrate how “speed bumps” feature simultaneously in several intersecting governance regimes, and discuss how any HDA-CAS must similarly exist at the intersection of several governance regimes. We then make the case that CAS design is shaped by, and shapes, governance design. Finally, we explore what this implies for ethical governance for HDA-CASs.

4.1 Understanding Governance

Speed bumps, like the one shown in Fig. 1, configure the driving environment and help regulate traffic speeds in sensitive areas. They are a small component of a wider system of traffic regulation, which we explore in detail below.

Speed bumps are an example of “**environmentally embedded regulation**” [39] illustrating the approach of configuring the physical environment to constrain driving practices in certain ways, in this case to regulate speed for reasons (perhaps) of pedestrian safety. At a base level the material features of the roads and their organisation create a balance of affordances and constraints that shape the possibilities for road use (e.g. speed and overtaking are possible on straight sections, but not where the road bends). This potential of the built environment to regulate social practices is actively exploited by town planners who configure urban spaces

in ways that inhibit crime and anti-social behaviour [41]. Analogously, obtaining desired forms of social computation depends upon carefully structuring virtual user-environments to regulate patterns of social behaviour in specific ways [11]. An example of this is how the moves an ESP game¹¹ are carefully arranged to produce game play that is generative of useful metadata tags.

Whilst the “rules of the road” might be given physicality in the form of speed-bumps or other traffic calming measures, there are a huge range of regulatory cues (signs, lines, grids, lights etc) that signal conventions of road use but do not by themselves enforce compliance. These are part of a **hierarchical and centralised** mode of regulation deriving from legal or institutional authority and policed by the state. Drivers are socialised to these rules formally via driving lessons and the driving test, and compliance is in part maintained through the threat of state (or professionally or institutionally) authorised sanctions. Centralised or hierarchical forms of regulation, in common with other forms, do not determine behaviour. Policing is imperfect, people are willing to risk sanction for some other benefit and the interpretation of rules is a matter of social convention, as is the degree to which they are enforced. So although shared norms and conventions amongst drivers take account of legally sanctioned regulations, they are not wholly determined by them. An example of this is the difference between the actual speed limit on UK motorways (70 MPH) and the *de facto* speed limit which is closer to 80 MPH.¹² Moreover, circumstances continually arise as part of road use that require improvisation and negotiation that would be impossible if official regulations were adhered to rigidly. In computing, this type of regulation is perhaps analogous to the *terms and conditions* attached to services that typically include expected standards of behaviour, allowable and prohibited ways that the service might be used, and sanctions that might be applied should the code deem to have been broken.

The calming effect of speed bumps depends on drivers noticing them, anticipating the jolt and adjusting their practices accordingly—a process that can become more automatic over time. Much of the moment-by-moment organisation of road use depends upon a mix of prior socialisation and situated decision-making, including an appraisal of environmental cues, what other drivers are doing or are likely to do, what the local conventions are, and expectations of how certain traffic situations are likely to evolve [7]. This in turn depends upon reading the intentions of other road users, signalling one’s own intentions, continually adapting one’s own approach in response, as well as adjusting to the adaptations of others. This can

¹¹ A serious game used to generate image metadata such as descriptive tags http://en.wikipedia.org/wiki/ESP_game

¹² A concern voiced about raising the official limit to 80 MPH is that the de-facto limit will then become 90 MPH. The difference arises due to cultural expectations about how regulations are policed. In the UK there is an expectation that the police will not enforce the rule rigidly, but instead allow some leeway, which for all practical purposes leads to raised limit. <http://www.independent.co.uk/news/uk/home-news/motorways-not-safe-enough-for-speed-limit-rise-to-80mph-7745678.html>

be seen as a form of **polycentric governance** [32],¹³ often contrasted to more centralised and hierarchical forms of social regulation, whereby communicating agents *collaboratively self-regulate* their use of a shared resource. This has components of mutual accommodation, sanction and reward, and plays into processes of community norm formation. Polycentric governance is seen to underpin the regulation of knowledge creating communities within Wikipedia, where formation and policing of community norms occurs as part of the communicative practices of community members, rather than being imposed externally. It is also visible in the “discussion fora” of sites like “Zooniverse” where a shared understanding and classificatory practices can emerge for what would otherwise be isolated decision-making tasks of individuals classifying astronomical objects.¹⁴ There are a number of attributes that make polycentric governance a possibility—but a principle among these is “cheap talk” [32]—i.e. easily accessible channels of communication between users of a resource. Design of HDA-CASs should orient to the channels of communication available between participants to take advantage of this type of self-regulation.

Speed bumps are a **motivational** form of governance [31]. They threaten discomfort, the chagrin of passengers and damage to the vehicle should a driver maintain an inappropriate speed. (Of course a thrill seeking teenager might find the bumps a motivation for driving faster.) Many types of social regulation seek to influence human actions through rewards and sanctions built around understandings of how peoples’ actions are motivated.¹⁵ Smart Society aims explicitly to regulate the use of resources through motivational mechanisms such as, incentives, persuasive technologies and reputation services. These are also common approaches to Smart City applications and a feature of interviews with Smart City consultants and implementers. Thus programmes towards more effective domestic energy use outlined by interviewees turned upon making energy consumption visible and therefore accountable,¹⁶ either on a household or neighbourhood basis, perhaps with explicit elements of competition and reward. Sometimes motivational aspects were present in stronger or weaker forms. For example, one interviewee in charge of a regional transport information service wanted to encourage network users to use public transport as often as possible and always provided a public transport

¹³Admittedly speed bumps are somewhat peripheral to polycentric modes of governance. But as we argue below, all the forms of governance presented here are interrelated. Thus how the driving environment is organised (including the presence or absence of speed bumps) shapes the sort of polycentric responses that are possible.

¹⁴<https://www.zooniverse.org/>

¹⁵Benkler suggests there are three classes of reward that people are motivated by: Money, Pleasure (“Intrinsic hedonistic rewards”) and Social (“Social-psychological rewards”) [4].

¹⁶There are a whole series of ethical issues attached to playing off accountability arrangements, particularly how they can create pressure that vulnerable people may be particularly susceptible to, shape behavior in unwanted ways and encourage “gaming” of the system. The worked example at the end of this chapter shows some of these properties for a technology of accountability operating in a care domain.

option in query results, but stopped short of using explicit incentives, partly so that responsibility for the choice remained with the user.

Speed bumps are an adaptation. They are typically placed in response to neighbourhood concerns or other evidence of incautious driving. The approach of adjusting governance measures in response to changing circumstances is referred to as **adaptive governance** and comprises of iterative cycles of monitoring, policy formulation and implementation [27, 39]. A key element of adaptive governance as applied to socio-environmental systems is to bring together diverse forms of expertise, particularly “native” expertise of people living within the system as to how complex socio-ecological systems might evolve in response to change (ibid). In the context of HDA-CASs, adaptive governance would involve forms of reflection that would bring together the expertise of smart society participants with a range of aggregated data describing how a HDA-CAS is behaving. Adaptive governance processes correspond to the cycle of sensing and adapting envisioned for HDA-CAS that will enable it respond to changing circumstances. However, an evolving CAS will most likely produce unpredictable and non-uniform responses to change—be they as a result of new regulations, counter-adaptations, new ways of making measurements, or environmental changes—in ways that demand the renewal of governance arrangements.

The “speed bump” sits within a nexus of diverse concerns voiced by many interested parties,—road users (of varying stripes), pedestrians, residents, motoring organisations, emergency services, environmental organisations, safety campaigners and so on. In this respect the roads analogy bears a strong resemblance to the ambition of HDA-CAS that aim to support diverse user groups with conflicting interests, since road users often have diametrically opposed interests (e.g. cyclists and motorist) and yet have to be accommodated within the same network. The mechanisms by which these voices are heard, how influence is wielded and how resources are allocated form the system of **political governance** of the highways, usually handled in a multi-tiered way via local and national governments and their agencies, but also via other forms of political expression such as campaigning activity. Political governance is a way of organising power and influence. It can be configured to respond to the diversity of interests and values that have to be brokered to create a functional network that roughly satisfies the requirements of many different users and user constituencies. In order to help satisfy the requirement of diversity within HDA-CAS, thought has to be given as to how those user constituencies can influence HDA-CAS configurations.

4.2 Governance Mechanisms as Layered and Intersecting

It should be clear from the above illustrations that managing a complex shared resource like a roads network involves a constellation of governance mechanisms operating simultaneously that serve a variety of purposes whilst at the same time continually interacting and influencing each other. For example, polycentric and

embedded regulation do not preclude one another, but instead tend to occur in mutually supportive (or sometimes disruptive) arrangements. Thus, a junction regulated by traffic lights still depends upon the self-coordinating practices of drivers to achieve its effect. When the lights break down, then traffic will typically continue to flow, but its management shifts towards greater polycentric regulation as the drivers themselves now have to coordinate turn taking [3]. Similar sorts of interdependency relationships can be found with motivational regulation. Coexisting governance arrangements are visible in the way that separate studies of Wikipedia alternately highlight either motivational or polycentric governance mechanisms as accounting for peer production in Wikipedia [4, 13]. We argue that these are different perspectives on a composite phenomenon, rather than competing explanations.

Adaptive governance can be seen to intersect with polycentric, motivational, and embedded modes in aiming towards specific regulatory effects by iterative modification of the physical, informational or incentive structures that underpin those regimes. Similarly political regulation operates over a slower time frame (except for some campaigns being enacted as deliberately surprising, quick interventions in public space) and can also appear “layered on” to other mechanisms¹⁷—although experience of the roads network, communication with other users and access to data about the network are all possible occasions or venues for political discourse or action. Figure 2 shows roughly the relationships between different governance regimes and how they may correspond to Smart Society concepts of evolution and operation. Table 1 shows sample governance mechanism and implementation approaches relevant to computer applications.

Building a CAS can be seen analogously as designing and implementing an ecosystem of governance mechanisms that caters for a diversity of users and fosters the emergence of certain patterns of resource use. This is not the same as designing the behaviour itself. Rather, the relationships between these governance elements need to be carefully thought out in order to allow the system as a whole to emerge in a coherent way.

4.3 *Ethical Governance*

Now we turn to the role that ethical governance has in relation to these various governance regimes. To maintain the analogy with a roads network, we can consider how the road builders and maintainers may have parallels with the designers,

¹⁷An article on the history of Speed “Humps” in Berkley on the City Authority’s web page (<http://www.ci.berkeley.ca.us/ContentDisplay.aspx?id=8238>) tells of how speed humps became contentious and how opposition to them led to shaping how humps are used as an adaptive regulatory measure (“speed hump locations chosen must provide clear safety benefits to balance any potential negative impact”).

Governance Regimes

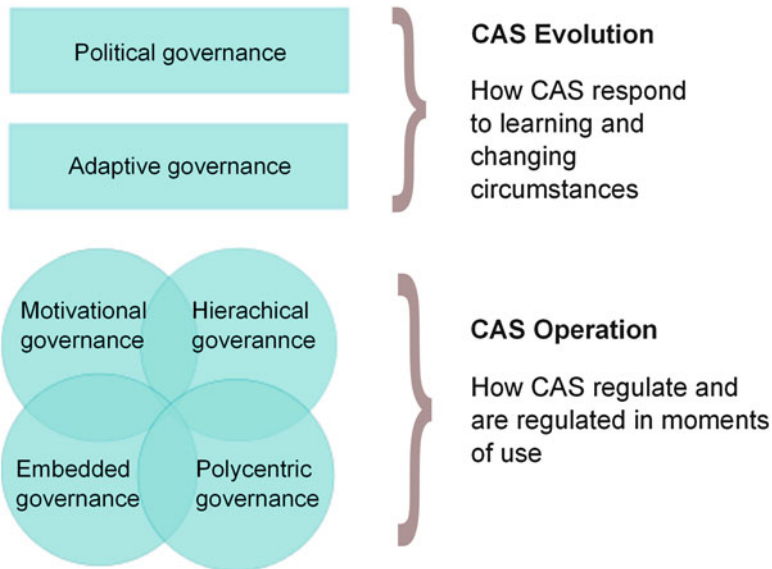


Fig. 2 This figure shows a rough logical arrangement of governance regimes and their relationships to CAS concepts of evolution and operation. This diagram simplifies tremendously the complexity of the relationships between these different aspects of governance

developers and builders of CAS.¹⁸ The road builders wield considerable power over road users in the decisions they make about which roads are built and how the traffic network is regulated—decisions that can affect livelihoods (e.g. where businesses are dependent on passing trade), safety, quality of life (both of drivers and neighbourhoods), the comfort of driving, and impact upon the environment. Designers and implementers of CAS will wield similar powers with respect to a given domain of CAS implementation. Taking care in the production of governance regimes for CAS could include consideration for:

The impact of regulation. Orienting to the practical circumstances in which the activity takes place and considering if the regulation itself poses annoyance, frustration or potential harm to users. The “speed bump” example works well here, because as a mode of regulation it can be potentially very annoying as well as cause damages to vehicles if not noticed. The one in Fig. 1 is painted white to help make speed regulation via bumps less uncomfortable and more palatable.

¹⁸Assuming the Collective Adaptive System doesn’t emerge “spontaneously” as an effect of integrating existing infrastructures and regulatory functions.

Table 1 Sample mechanism and implementation approaches for different forms of governance

| Governance regime | Mechanism | Implementation approach |
|--------------------------|--|--|
| Polycentric | “Cheap talk”—ability to sanction | Discussion boards, chat channels, collaborative filtering, provision of information about the state of the resources and resource users... |
| Motivational | Seeking of monetary, social-psychological or hedonistic reward. Avoidance of sanction. | Policing, monitoring, logging, reputation services, incentives... |
| Environmentally embedded | Structuring physical or virtual environment to achieve regulatory effects. | Visibility arrangements, signs, alerts, workflow organisation, ease or difficulty of interactions... |
| Hierarchical | Laws, regulation, codes of conduct, institutionally backed sanctions and policing. | Terms and conditions, service agreements, codes of conduct, monitoring, penalties, exclusion. |
| Adaptive | Cycles of monitoring, policy formation and implementation. | Sentiment data, sensor data and provenance data, analytics, engagement with users and other experts, discussion fora, AB testing... |
| Political | Representation and decision-making processes. | A constitution, stakeholder representation, discussion fora, executive officers, voting, petitioning... |

Regulating collectives. Adjusting regulatory mechanisms to achieve some new effect has implications at a collective level where understanding the values and social norms associated with the collective, or with communities within the collective, becomes important. An example here lies with the Ride Sharing scheme where interviews with participants reveal a regime of fixed prices between particular destinations based upon communitarian principles of sharing resources and costs. Attempts to raise the price are typically viewed as being “greedy” and resisted. As part of HDA-CAS we might aim to motivate Ride Sharing participants in new ways (perhaps to improve environmental outcomes), but on the basis of existing norms we can see that achieving this via market based principles might be tricky. This might lead us to select a different approach to motivational regulation that relies less on monetary reward for its effect. The Ride Sharing scheme does not have a central constitution or enforcement mechanisms, but it is evident from the interviews that participants orient to a strong set of community norms and standards of behaviour, indicating a strong polycentric aspect to its regulation. Safety has been identified as of key importance to Ride Sharing, and providing for appropriate social regulation to prevent people coming to harm is an important factor to enable a Ride Sharing CAS to gain acceptance beyond single institutional contexts.

Building on existing regulation. A broader principle building upon the above point is to understand, build upon and build out from existing community norms and regulatory mechanisms.

Anticipating the transformatory power of CAS. When CAS are designed to transform how shared resources are managed over existing practices, perhaps by connecting community members in new ways, then one also has to think through what new sorts of regulation might be required in these transformed circumstances. In the Care House scenario described in Sect. 5, potential of CAS to transform accountability regimes, and the danger of losing a qualitative notion of compassion when care tasks are quantified, calls for specific regulatory mechanisms to safe-guard certain core values.

Balancing Governance Regimes. Fashioning an appropriate balance between regimes is important, as each approach contributes important attributes in a mosaic-like way to the overall system of governance. Thus, a builder of a CAS might ask himself which parts of the regulation need to be freer and community directed, and which need to be more rigid and embedded, and which need to be driven by incentives. Failing to think through provision in a particular area could lead to inequity. For instance, a lack of explicit and appropriate structure for political expression could lead to increasing marginalisation of already vulnerable groups.

Understanding values attached to governance. Governance mechanisms themselves are attached to particular values. They can be more or less democratic or participatory in their implementation, for example. Polycentric governance in particular has an important link to autonomy. In writing about digitally augmented mobilities, Buscher et al. propose that people are “served humanely” when representations of the sensed network are used as a resource for “improvised situated action” rather than centralised control [7]. Thus, a system that minimises polycentrism and drives embedded and motivational governance risks being overly controlling and oppressive.

Designing for adaptive governance. At the point of emergence a CAS might carry a lot of intentional design. Once in operation, however, provision should be made for adaptive governance processes to take over the ongoing redesign of the system. This can be kick-started by making the initial design rounds very much like the adaptive governance cycle, with investigations into the prospective domain, participatory policy formation and trial implementation.

Achieving just the right amount of regulation. Governance design should be proportionate to the scale of the system envisaged and the types of communities implicated. Governance of a nationwide traffic network is immensely complicated and intricate, and has evolved to its current form over the entire history of road use. While it serves as a motivating example for this discussion, one should maintain a sense of proportion when bringing the ideas to any real world example.

Adopting governance structures appropriate to the scale of the CAS. As the scale of a CAS changes, it is likely that governance mechanism may become strained and new patterns of governance will be needed to succeed them. For example, issues that can be handled informally between a pair of collaborating colleagues

might need a more formal project management structure to be properly managed within an international research team. An example is how within Wikipedia, governance patterns have changed with changing scale and learning within the wikipedia community [13].

4.4 *Guidance for Governance Design*

This section considers the sort of design procedure one would follow to realise governance mechanisms with the characteristics outlined above. Treating the design of HDA-CAS as if it were a problem in governance design has the helpful property that social values become first class objects for design, as opposed to being “relegated” to informing categories of non-functional requirements which might be addressed late in the day and/or incompletely. That is to say if one wishes to engineer patterns of social behaviour, then one has to understand and work with sociality. Another way to put this is that if we accept that the speed bump’s symbolism is in fact part of its regulatory effect, thinking about how to convey values to influence social orders also becomes an important aspect of design [39].

On this basis, we suggest the following steps for design of HDA-CAS:

1. *Understanding an existing collective, its values and modes of regulation* by characterising the domain in terms of how it functions as a social system—the sort of collective that it corresponds to, the important sub-collectives of which it is composed, how the collective regulates itself, understanding what its core values are and the range of diverse values present.

There are a number of tools that can help surface social values in a concerted way. Perhaps the most prominent of these are Value-Sensitive Design and Reflective Design approaches [14, 38] that depend upon social science modes of enquiry and “disruptive” design practices to probe existing values. An important research issue is to develop these tools to address dimensions of collectivity since current versions focus more on the values of implicated individual stakeholders rather than of communities. A disclosive computer ethics approach can also be used to surface social values that become silently embedded in computer systems [5]. Anticipatory governance too has an important role to play in helping us see the consequences of alternate design choices by generating insights into possible futures. The Care House scenario in Sect. 5 shows how the altering the balance between different governance regimes can have a significant effect on the overall properties of the system, and illustrates entry points for translating knowledge of social values into governance design.

2. *Draw upon existing knowledge and experience* by bringing together diverse forms of information, expertise and interests, including: the “native” domain expertise of CAS participants, sensor and other quantitative data from existing sources, technical expertise, social science expertise and psychosocial understandings of how human practices are influenced by persuasion and incentives.

This reflects the “enquiry” phase of an adaptive governance cycle and implies strong participatory approaches. It also resonates with Responsible Research and Innovation (RRI) maxims of socially embedded and socially responsive innovation [33]. Participatory design approaches can work at scale [29], and it makes sense to implement these by using the Smart Society platform to engage collectives in design-oriented tasks. Finding ways to balance the influence of designers and different constituencies of native participants will provide clues as to the sorts of political governance mechanisms required.

3. *Designing for governance* by drawing on prior steps, the aim would be to identify key regulatory objectives and implement these through a balance of governance mechanisms. These would aim to produce the desired sorts of social organisation and to regulate the system as a whole to behave in ways that are acceptable to the participating collectives.

Working out how to translate from information about a domain (from prior steps) into operational governance regimes presents a real challenge to innovate design approaches that can help deliver Smart Society applications. Some starting points include: using our understanding of governance approaches as outlined above as a way of structuring the design challenges (e.g. as a “checklist” of issues that need to be covered); developing a toolkit of governance structures, such as discussion fora, voting mechanisms, chat channels, incentive mechanism, transparency arrangements, constitutional statements etc (see Table 1) that can be composed into a working application; providing mechanisms that set limits or boundaries on the platform that constrain CAS behaviour along particular dimensions to anticipate and contain certain sorts of unwanted adaptation.

5 A Worked Example of Governance Design for HDA-CAS in a Care Setting

This example derives from an interview with a research consultant working on a project to explore how proximity sensors worn by care home staff and residents can be used as an aid to “reflective practice” [26]. The sensors register each time a carer comes within 1.5 m of a resident. The carer can then view analytics that show those residents they were proximal to, when, and for how long, as well as how overall contact time is shared between residents. A sensor is also located on the care home computer to indicate how much time is spent on administrative tasks. The idea is that staff can interpret this data to rethink their own practice, perhaps prompting consideration of who they spend more time with, who less, and why.

This example has a number of advantages for exemplifying Smart Society concepts:

1. It is a simple case that can be easily extended to incorporate features that give it the properties of a HDA-CAS (an elaborated version is described below).

2. There are evident social values and governance issues attached to the system's use.
3. It falls within the application area of social care, which is seen as an important focus for Smart Society as it moves forward, particularly in relation to use of sensors to assist the delivery of care.

The discussion below attempts to illustrate some of the issues and potential solutions in the governance of a HDA-CAS based upon the principles outlined earlier. The idea is to stimulate a certain way of thinking about CAS and their design, particularly to give attention to the issues, tensions and contradictions that emerge when applied to a real world context. The analysis is not meant to be exhaustive and many of the disciplines within Smart Society would have strong suggestions as to the sorts of mechanism or approaches that might be used to address the different issues that are raised, particularly how incentives can be effectively configured; how reputation and provenance can be factored in; and how social orchestration can be designed to help create the “right” sorts of hybridity. Finally, the example does not reflect in any way the actual intentions of the Mirror project¹⁹ which created the original sensor based app for reflective practice. The projection of an extended system exists only within the context of Smart Society.

5.1 *Smart Society Extensions*

While the computer system is able to *aggregate* the pattern and duration of contacts, these aggregated traces are not particularly meaningful by themselves. As the interviewee has it: “[the sensor] doesn’t tell you the quality of the interaction, it simply tells you an interaction’s occurred”. Interpreting the sensor trace depends on the care staff supplying missing contextual detail: where do the residents usually sit? Which residents prefer attention, which prefer to be left alone? Which registrations are likely to be “artefacts”, and which correspond to “real” interaction? This interpretation of the pattern of contacts by care staff is already a social computation and demonstrates *hybridity* between machine and human capabilities. In particular, it shows how human interpretation can help bridge the *semantic gap* between sense data and meaning.

Of course, in developing this as a Smart Society scenario, the contribution of human-factors colleagues would be to improve *activity recognition* through better sensors and algorithms, although this is unlikely to eliminate the need for human judgment; but perhaps it would alter the sorts of judgment required, with the human needing less to “repair” sensor readings, and able to concentrate more fully on assessing their significance. While human expertise helps bridge the “semantic gap”

¹⁹The EU Mirror project aims to create a series of applications to support reflective professional practice. <http://www.mirror-project.eu/>

between sense data and meaningful interpretations, part of the Smart Society vision is to deliver automated support for sense-making and decision-taking in areas where the computation is easiest for the machine. An extension to the proximity sensor system enabling the discovery of helpful permutations of staff given constraints of duty rotas and shift patterns could be an example of this sort of automation. The work within the project on *lightweight social orchestration* would be concerned with how the blend of automation and human control is realised in practice.

The example has elements of *evolution and adaptation* built-in, since the aim is for the care staff to adjust their practice on the basis of reflecting on sensor data. Simple extensions to the example provide a means to explore *diversity and scale*. Diversity could be present in a number of ways, including: perhaps different types of sensor that vary in the way they provide descriptions of proximity, or to incorporate the different preferences, knowledge and skills of carers and residents (this may enable the system to help determine combinations of carers best able to meet a resident's care needs because of shared interests or values). Diversity becomes an increasingly important consideration when the system is *scaled up* from a single care home to encompass improving care provision across an administrative region. With scale, *governance* issues also come increasingly to the fore, since decision-making and planning would be implicated at multiple levels of organization with each level orienting to different sorts of goals, these are unpicked more fully in a discussion of governance and social values below. Finally, there is scope for building in *reputation mechanisms* and *incentives*, perhaps via resident's rating of the care they receive, through "badges" or other rewards for thoughtful practice.

5.2 Social Values and Governance

The issues presented below represent a value sensitive analysis of the care home example based upon the interview data obtained as part of the empirical component of Smart Society, a conceptual analysis based on our understanding of types of social impact, and an analysis of the technology characteristics. The discussion revolves around design based upon the principles outlined in the governance principles discussed earlier in this chapter.

5.2.1 Embedded Regulation

The following quote is a very good example of how values can be embedded in design, of embedded forms of regulation and how the balance can be struck between different regulatory approaches:

"the original the developers [developer's name] they came up with a kind of dashboard you know - 100% to 0% - critical and colour coded all the way along - Woo Hoo - I said no, no - take off all values - we are not here to tell them what is good or bad, what's critical or what's adequate (...?) not our job."

Table 2 Different models of the sensor based system depending on how far the sensor data circulates

| | Extent of data sharing | Accountability practices |
|---|--|--------------------------------|
| 1 | Only you see your data | Self reflection |
| 2 | The data is shared within the team of carers | Group reflection and oversight |
| 3 | The data is available to the care home manager | Managerial oversight |
| 4 | The data is shared with residents and or their relatives | Customer oversight |

The proposed colour coding pre-configures how “readings” of contact time should be interpreted and as such embeds judgments about what constitutes an appropriate level of contact. These inscribed values imply a regulatory effect similar to that of a thermostat where the aim would be to get the “readings” within an acceptable range. This set-up runs the risk of pushing carers to orient to “getting the reading in the green” as a metric of good care, rather than orienting to quality of interactions and individual need. This points to the more generic danger posed by technologies that *quantify* as framing care in terms of metrics rather than as personal, compassionate, empathic and responsive—characteristics of the quality of interactions. It also shows the power and subtlety of regulatory cues embedded within the user environment and how these should be used mindfully and with sensitivity. In the quote, the IT consultant orients towards a more polycentric mode of regulation that favours greater hybridity by placing a greater emphasis on the discretion and contextual knowledge of the professional carers. We discuss this in further detail in the section on polycentric governance below.

5.2.2 Accountability Regimes

The extent and types of information flows that a technology enables are also implicated in various regulatory effects. An extended version of the sensor system can be configured to create different patterns of disclosure to different audiences and thus, bring different balances of regulatory mechanism into play.

Each of the following patterns of disclosure in Table 2 opens up a different dimension of accountability.

5.2.3 Polycentric Governance

If we think of the care staff as a bounded resource that needs to be allocated effectively to meet the diverse needs of residents then we can also see how, within the context of normal practice, a variety of regulatory structures will play a role in managing the shared resource. One aspect of this will be “centralised” management

practices such as the production of a staff rota to ensure that there is appropriate “cover” at all times. These specifications will not, however, detail precisely who does what and when, which will be a matter partly of routine, partly of negotiation and partly of response to contingency—i.e. regulation of care resources at certain levels have a high degree of polycentrism. That is to say it is the staff and residents collaborate in planning and self-organise their moment-by-moment activities around a negotiated and continually evolving shared sense of what needs doing and what division of labour would best achieve those tasks (which will of course be reflected in more static instruments such as the rota).

The sensor system of this example provides an additional source of information that can feed into reflective practices crucial to polycentric forms of self and mutual regulation. As an aid to self-reflection where a staff member only sees data corresponding to their own activities, this perhaps will prompt them to make adjustments to their own work practices. Sharing everyone’s data between all team members perhaps has a greater potential for insights, ideas and mutual reweaving of priorities, practices and routines. It will also carry greater risks (in extreme cases maybe associated with work place bullying), and will exert subtle pressures toward conforming to the metric of the system.

5.2.4 Motivational Regulation

One way of viewing the sensor system may be like a rather neutral source of information that can be incorporated into reflective practice to optimise use of a constrained care resource. Another is to acknowledge that at the same time, sensor reading can carry very strong moral overtones as to, for example, whether staff are performing as they should, and whether residents are receiving equal and appropriate care, and so on. Hence the high degree of sensitivity that can be attached to how far the sensor traces circulate how easily subjects can be identified. Thus, while in the original example the system is intended as an aid to *reflective practice*, this ostensive purpose is not fixed, and the tool’s strong *evaluative potential* in particular, is something that people can seek to exploit:

“one of the reviewers he clearly cottoned on to it very quickly and said you are really on to something here - you could sell this, it says, as a quality assess- assurance for relatives - so it’s not the carers that get the data it’s the relatives that get the data and you think ‘oh my god’ you know - but that’s exactly your issue now - how far down that road - whose data is it?”

It is a very common experience that people are motivated to adjust their practices if they feel they are being observed or assessed, and it would be easy to behave in a way that gave a “positive” account of resident contact time without actually increasing positive interactions with residents. Thus adaptations motivated by these new types of accountability (from managers or relatives) may be quite negative, and may devalue the sensor systems’ use as an aid to reflection (because the sensor reading can no longer be trusted).

5.2.5 Adaptive Governance

In the section above on embedded regulation, we saw how the IT consultant argued against the use of “colour coding” precisely to remove evaluative connotations. We can see this as a very simple instance of adaptive governance, where the technology is reconfigured to deliver a different regulatory effect by reflecting upon and anticipating its likely or actual effect.

In the above sections we have formulated a problem. The sensor tool threatens to connect residents, managers, relatives and carers in new ways creating new means of surveillance and accountability that contain the possibility for unwanted and unhelpful adaptations, as well as positive ones. In expanding the system to help beyond personal reflective practices, we have to think of the forms of adaptation that might enable these different functions to more happily co-exist.

One strategy might be to use techniques of anonymisation or aggregation, so that data can be examined at a management level or beyond without implicating individuals or individual care homes. This data would still likely be useful, although not ideal, but provide less strong motivations to “game” the system.

Another might be in finding ways of keeping the carers honest such as, enabling residents to annotate data to give some indication of the quality of the interaction in contexts where this may be possible.

There are many further possibilities and combinations of possibilities that have the potential to shape different patterns of practice. These occur at different levels within the system with different implications for the quality of the data that emerges and whether the “real” goals of the system are being met. The point of adaptive governance is that these types of solution should be investigated, trialled and re-evaluated in an ongoing loop of information gathering, discussion and experimentation.

There may be a number of adaptive cycles at different “levels” within the system. Thus the care staff themselves might experiment with different ways of displaying, sharing and interpreting the data locally that helps maintain an emphasis on the “human” elements of care. While at the same time similar processes could be occurring for how data across the region is used to inform care policy, staffing levels and so on.

5.2.6 Political Governance

“But you could imagine - or you could very easily imagine - care home managers deciding that they would want to find these things out and the carers will wear these sensors whether they like it or not and there could be problems without a doubt because - we did come across a couple of carers that didn’t want to wear them. And obviously, you know, we didn’t force them although- . . . I mean it was a small group because we I think there was nine carers in this group and one of them I remember in this test just felt comfortable but peer pressure carried the day and so she says “ok I will do”.”

This quote points us towards the politics of the workplace, and by extension, wider spheres of political involvement that would come to encompass unions, professional bodies, governments, resident and relative care pressure groups, particularly as the scale and scope of the system expands.

One issue that is likely to have political ramifications is how such an expanding system would change the nature and character of care work as a profession. A system that more closely matches need with care expertise across a geographic region could lead to changing shift patterns and demand increasing flexibility or mobility of carers. Such a framework might also enable care increasingly to be delivered remotely or virtually or via robots. It could also alter the sorts of qualifications needed to participate in care and entry into the profession, and how care professionals are remunerated. In the end, it could change or challenge broader social attitudes to care. These issues all raise questions as to who should be setting or shaping and monitoring the overall goals of the system, and the sorts of social and political participation needed to review the values underpinning those goals.

6 Conclusions

This chapter discussed a range of intended and possible empirical features of CAS associated with the vision of a Smart Society, and provided some conceptual elements and empirical illustrations for the ethical governance of such systems. The overall point to be made is that any attempt to construct a framework for ethical governance necessarily remains incomplete and contestable, hence our metaphor to sketch a path towards ethical governance rather than provide a full account. The next, more concrete points to be made relate to this general one, they are derived from our analysis and synthesis of existing forms of governance and CAS features.

First, many CAS are built on, and into, existing forms of governance, that is, different ways of steering society and maintaining social order, and more precisely, different approaches to inherent contradictions like autonomy versus control, centralised versus self-regulation, individual versus public interests. We have distinguished five such different governance regimes (and there may be more): polycentric, motivational, environmentally embedded, adaptive, and political. We have argued that in practice, such different governance modes actually interact with one another and influence each other, thus forming a composite phenomenon rather than competing “juxtaposed” alternatives. This composite phenomenon gets more complicated, and may have ever more emergent properties, as the scale of CAS increases (e.g. national road networks), and as human and technical system “components” change over time, often in numerous feedback-loops (and we strongly advocate such a time-sensitive, historical view of CAS for a more realistic understanding).

Second, CAS may be designed to be as diversity-aware as possible, but many real-life settings will include so many diverging stakeholder needs and interests that ethical problems may not be solved completely, for everybody, and once and for all.

We assume that such disharmony and residual conflict are more the rule than the exception. So we make the case that persisting ethical tensions should be perceived as a structural feature of CAS. The latter may be designed to deal with such tensions, but it may not be realistic to assume that any design can reconcile all possible, and emerging, ethical problems “in the wild”.

However, this precaution does not mean that the whole project needs to be abandoned—quite the contrary. We propose that, third, a basic understanding of ethical governance as being mindful in the production of new, or the reproduction of existing governance regimes; and as continuous processes of identifying and managing recurrent ethical concerns. With this double emphasis on facilitating processes of problem identification, and on a second-order awareness of existing governance regimes, we suggest a procedural understanding of “ethics” here, not (only) a substantial one. It remains to be discussed how such a procedural ethical governance is to be realized in practice.

Finally, given the abstractness of a procedural interpretation of ethical governance and its potential downsides, i.e. second-order ethical problems (e.g. who or what decides about the right procedures; who or what actually takes care of (re)producing governance regimes), more substantial, domain-specific values may need to be considered and designed into a CAS (with the possibility of redesign and recalibration of vision). For instance, a substantial value to be accounted for may be a qualitative notion of care that, in this chapter, has been spelled out in greater detail through a worked example of governance design for HDA-CAS in a care setting. The point to be made here is that there is no abstract or theoretical short cut to the development of such substantial domain specific values, nor are they universally and absolutely true. One has to work through the details of different empirical instances of a given domain to develop a careful, ready-to-be-revised preliminary understanding of important substantial ethical values. This seems also to require methodological innovation to build anticipation more strongly into design processes. We believe that this chapter provides a modest example of how this can be achieved, in this instance, by weaving together the ethics attached to contemporary socio-cultural trends, an understanding of governance design, and elaboration of an empirical case. Although there are still many problems to solve, not least achieving a wider representation of stakeholder engagement when envisioning possible futures, we believe we have the basis of a framework for the ethical governance of CAS that we intend to build upon in Smart Society.

Acknowledgements The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 321480 (GREAT) and no. 600854 (Smart Society). We would also like to thank the Engineering and Physical Sciences Research Council (EPSRC) for funding the Framework for Responsible Research and Innovation project Grant No. EP/J000019/1.

References

1. Aberbach, J.D., Rockman, B.A.: Conducting and coding elite interviews. *Polit. Sci. Polit.* **35**(4), 673–676 (2002)
2. Anderson, S., Bredeche, N., Eiben, A., Kampis, G., van Steen, M.: Adaptive collective systems: herding black sheep. In: *BookSprints for ICT Research* (2013)
3. Baker, L.: Removing roads and traffic lights speeds urban travel. *Scientific American*. Online. <http://www.scientificamerican.com/article.cfm?id=removing-roads-and-traffic-lights> (2009)
4. Benkler, Y.: Coase's penguin, or, linux and the nature of the firm. *The Yale Law J.* **112**, 369–446 (2003)
5. Brey, P.: Disclosive computer ethics: The exposure and evaluation of embedded normativity in computer technology. *Comput. Soc.* **30**(4), 10–16 (2000)
6. Bullock, S.: Prospects for large-scale financial system simulation. The future of computer trading in financial markets. *Foresight driver review DR14* (2011). <http://eprints.soton.ac.uk/272759/> (2011). Accessed 9 Aug 2013
7. Büscher, M., Coulton, P., Efstratiou, C., Gellersen, H., Hemment, D.: Connected, computed, collective: Smart mobilities. In: Grieco, M., Urry, J. (eds.) *Mobilities: New Perspectives on Transport and Society*, pp. 135–158. Burlington: Ashgate (2011)
8. Collective adaptive systems, expert consultation workshop report. ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/fet-proactive/shapfetip-wp2011-12-02_en.pdf (2009). Accessed 9 Aug 2013
9. Demographic change and an ageing population. report of the finance committee. <http://www.scottish.parliament.uk/parliamentarybusiness/CurrentCommittees/59613.aspx> (2013)
10. Description of work. Grant agreement for the smartsociety project, annex i. Internal project documentation (2013)
11. Erickson, T., Kellogg, A.: Social translucence: An approach to designing systems that support social processes. *ACM Trans. Comput. Hum. Interact.* **7**(1), 59–83 (2000)
12. Fleischmann, K., Wallace, W.: Ensuring transparency in computational modeling. *Commun. ACM* **52**(3), 131–134 (2009). <http://dl.acm.org/citation.cfm?id=1467278>
13. Forte, A., Larco, Y.V., Bruckman, A.: Decentralisation in wikipedia governance. *J. Manag. Inf. Syst.* **26**(1), 49–72 (2009)
14. Friedman, B., Kahn, P., Borning, A.: Value sensitive design: theory and methods. Technical report, UW CSE (2002). <http://www.urbansim.org/pub/Research/ResearchPapers/vsd-theory-methods-tr.pdf>
15. Fuerth, L.S.: Foresight and anticipatory governance. *Foresight* **11**(4), 4–32 (2009)
16. Giunchiglia, F.: Hybrid and diversity-aware collective adaptive systems in human computer confluence. In: Ferscha, A. (ed.) *The Next Generation Humans and Computers Research Agenda*, pp. 12–15 (2014). <http://hcsquared.eu/hc2-visions-book>
17. Guston, H.D.: The anticipatory governance of emerging technologies. *J. Kor. Vac. Soc.* **19**(6), 432–441 (2010)
18. Hannam, K., Sheller, M., Urry, J.: Editorial: Mobilities, immobilities and moorings. *Mobilities* **1**(1), 1–22 (2006)
19. Hollands, R.G.: Will the real smart city please stand up? *City Anal. Urban Trends Cult. Theory Policy Action* **12**(3), 303–320 (2008)
20. Introna, L.D., Nissenbaum, H.: Shaping the web: Why the politics of search engines matters. *Inf. Soc.* **16**(3), 169–185 (2000)
21. Johnson, D.G., Mulvey, J.M.: Accountability and computer decision systems. *Commun. ACM* **38**(12), 58–64 (1995)
22. Kjolberg, K.: The notion of “responsible development” in new approaches to governance of nanosciences and nanotechnologies. Ph.D. thesis, University of Bergen (2010)
23. Knobel, C., Bowker, C.G.: Values in design. *Commun. ACM* **54**(7), 26–28 (2011)
24. Lanier, J.: *Who Owns The Future?* Allen Lane, London (2013)
25. Lyon, D.: Surveillance as social sorting. In: Lyon, D. (ed.) *Privacy, Risk and Digital Discrimination*. Routledge, London (2003)

26. Maiden, N., D'Souza, S., Jones, S., Müller, L., Pannese, L., Pitts, K., Prilla, M., Pudney, K., Rose, M., Turner, I., Zachos, K.: Computing technologies for reflective, creative care of people with dementia. *Commun. ACM* **56**(11), 60–67 (2013)
27. McNutt, K., Rayner, J.: Valuing metaphor: A constructivist account of reflexive governance in policy networks. In: 5th Conference on Interpretive Policy Analysis, pp. 23–25, Grenoble (2010)
28. MIT: senseable city laboratory. <http://senseable.mit.edu/>
29. Neuhauser, L., Rothschild, B., Graham, C., Ivey, S.L., Konishi, S.: Participatory design of mass health communication in three languages for seniors and people with disabilities on medicaid. *Am. J. Public Health* **99**(12), 2188–2195 (2009)
30. Options for strengthening responsible research and innovations: Report of the expert group on the state of art in europe on responsible research and innovation. http://ec.europa.eu/research/science-society/document_library/pdf_06/options-for-strengthening_en.pdf Accessed 9 Aug 2013
31. Osterloh, M., Frey, B.S., Frost, J.: Managing motivation, organization and governance. *J. Manag. Govern.* **5**(3), 231–239 (2001)
32. Ostrom, E.: Beyond markets and states: Polycentric governance of complex economic systems. *Am. Econ. Rev.* **100**(3), 641–672 (2010)
33. Owen, R., Macnaghten, P., Stilgoe, J.: Responsible research and innovation: From science in society to science for society, with society. *Sci. Public Policy* **39**, 751–760 (2012)
34. Pickering, A.: The mangle of practice: Agency and emergence in the sociology of science. *Am. J. Sociol.* **99**(3), 559–89 (1993)
35. Progress project. <http://www.progressproject.eu/more-rr-i-resources/>
36. Schatzki, T., Cetina, K., von Savigny, K.: *The Practice Turn in Contemporary Theory*. Routledge, New York (2001)
37. Seddon, J.: *Systems Thinking in the Public Sector*. Triarchy Press, Axminster (2008)
38. Sengers, P., Boehner, K., David, S., Kaye, J.J.: Reflective design. In: *In Proc. Critical Computing*, pp. 49–58. ACM Press, New York (2005)
39. Shah, R.C., Kesan, J.P.: How architecture regulates. *J. Archit. Plann. Res.* **24**(4), 350–359 (2007)
40. Silberman, M.S., Irani, L., Ross, J.: Ethics and tactics of professional crowd work. *XRDS: Crossroads, The ACM Magazine for Students* **17**(2), 39–43 (2010)
41. Slater, J.: Reducing crime through design, supplementary planning document. Tech. rep., Portsmouth City Council. http://www.portsmouth.gov.uk/media/Reducing_Crime_Through_Design_SPD.pdf (2006). Accessed on 24 Mar 2014
42. Stahl, B.C., Eden, G., Jirotko, M.: Responsible research and innovation in information and communication technology - identifying and engaging with the ethical implications of icts. In: Owen, R., Heintz, M., Bessant, J. (eds.) *Responsible Innovation*, pp. 199–218. Wiley, London (2013)
43. Willke, H.: *Smart Governance: Governing the Global Knowledge Society*. Campus, New York (2007)