

# The open agent society: retrospective and prospective views

Jeremy Pitt<sup>1</sup> · Alexander Artikis<sup>2,3</sup>

Published online: 6 October 2015  
© Springer Science+Business Media Dordrecht 2015

**Abstract** It is now more than ten years since the EU FET project ALFEBIITE finished, during which its researchers made original and distinctive contributions to (*inter alia*) formal models of trust, model-checking, and action logics. ALFEBIITE was also a highly inter-disciplinary project, with partners from computer science, philosophy, cognitive science and law. In this paper, we reflect on the interaction between computer scientists and information and IT lawyers on the idea of the ‘open agent society’. This inspired a programme of research whose investigation into conceptual challenges has carried it from the logical specification of agent societies and dynamic norm-governed systems to self-organising electronic institutions, while developing several technologies for agent-based modelling and complex event recognition. The outcomes of this inter-disciplinary collaboration have also influenced current research into using the open agent society as a platform for socio-technical systems, and other collective adaptive systems. We present a number of research challenges, including the ideas of computational justice and polycentric governance, and explore a number of ethical, legal and social implications. We contend that, in order to address these issues and challenges, the continued inter-disciplinary collaboration between computer science and IT lawyers is critical.

---

In Honour of Jon Bing.

---

✉ Jeremy Pitt  
j.pitt@imperial.ac.uk  
Alexander Artikis  
a.artikis@unipi.gr; a.artikis@iit.demokritos.gr

<sup>1</sup> Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, UK

<sup>2</sup> University of Piraeus, 18534 Piraeus, Greece

<sup>3</sup> Institute of Informatics & Telecommunications, NCSR Demokritos, 15310 Athens, Greece

**Keywords** Computers and law · Multi-agent systems · Norm-governed systems · Self-organising systems · Event recognition · Ethical legal and social implications (ELSI)

## 1 Introduction

It is now more than ten years since the EU FET project ALFEBIITE<sup>1</sup> finished: the terms ‘infohabitants’ and ‘universal information ecosystem’ never caught on; its logical framework for ethical behaviour was never fully realised; its domain name was porn-napped; and its end-of-project collected volume, though often referenced, was never actually published.

On the other hand, some of its researchers made significant original and distinctive contributions to formal models of trust (Jones 2002), model-checking (Lomuscio et al. 2009), and logics for action and agency (Sergot 2008). It was also the starting point for, *inter alia*, the development of a model of agent communication that overcame the limitations of the FIPA standard ACL<sup>2</sup> (Jones and Kimbrough 2008), a formal model of forgiveness (Vasalou et al. 2008), and a methodology for the design of socio-technical systems (Jones et al. 2013).

Furthermore, ALFEBIITE was a highly inter-disciplinary project, with partners from computer science, philosophy, cognitive science and law. In this paper, we reflect specifically on the outcomes of the interaction between the computer scientists and the information and IT lawyers on the idea of the ‘open agent society’ (Pitt et al. 2001; Pitt 2005). This inspired a programme of research whose investigation into conceptual challenges has carried it from the logical specification of agent societies and dynamic norm-governed systems to self-organising electronic institutions, while developing several technologies for agent-based modelling and complex event recognition. The outcomes of this inter-disciplinary collaboration have subsequently influenced current research into using the open agent society as a platform for socio-technical systems, and other collective adaptive systems. We present a number of research challenges, including the ideas of computational justice polycentric governance, and explore a number of ethical, legal and social implications. We contend that, in order to address these issues and challenges, the continued inter-disciplinary collaboration between computer science and IT lawyers is critical.

This paper is structured along the lines of this research programme, past and present. Section 2 presents more detail on the background to the ALFEBIITE project, the idea of the open agent society, and surveys a kind of ‘road map’ of past and future development. Given this background, the paper is then effectively divided into two parts: Retrospective and Prospective. In the Retrospective part, various approaches to (computational) logic-based specification of multi-agent

---

<sup>1</sup> Pronounced  $\alpha\beta$ : the acronym stood for “A Logical Framework for Ethical Behaviour between Infohabitants in the Information Trading Economy of the Universal Information Ecosystem”. No-one ever asked twice.

<sup>2</sup> FIPA: the Foundation for Intelligent Physical Agents, see <http://www.fipa.org>; ACL Agent Communication Language, see <http://www.fipa.org/repository/aclspecs.html>.

systems are presented in Sect. 3, and particular technologies for reasoning with and about such specifications are presented in Sect. 4. Having made these developments, we then address the research question: what are the prospects for using the open agent society as the basis for developing socio-technical systems or other collective adaptive systems? Accordingly, in the Prospective part, Sect. 5 presents a quartet of research challenges for socio-technical systems with intelligent components, and Sect. 6 considers a quartet of pressing ethical, legal and social issues which need to be addressed by the ‘digital society’. We conclude in Sect. 7 with some remarks on the value of inter-disciplinary research and the difficulty of evaluating speculative basic research.

## 2 Background: the open agent society

This section has three aims: firstly, to present the background to the ALFEBIITE project; secondly, to discuss the motivation and inspiration for the open agent society; and thirdly to survey a kind of ‘roadmap’ of past developments (to be presented in Sects. 3 and 4) and future research directions (to be presented in Sects. 5 and 6).

The motivation for the open agent society stemmed from two key developments, one foundational, the other strategic. The foundational development was the highly influential paper of Jones and Sergot on *institutionalised power* (Jones and Sergot 1996), which gave the first formal logical characterisation of the concept of ‘counts as’, which was a significant factor in Austin’s (1962) conventional interpretation of speech act theory and its refinement by Searle (1969). The strategic development was the inauguration of FIPA (the Foundation for Intelligent Physical Agents), which sought to define standards for open systems in which intelligent, autonomous and heterogeneous components (i.e. agents) had to interoperate. The confluence of these two developments, unpicked further in Sect. 2.1 resulted in the ALFEBIITE project and the proposal of the open agent society, based on an agent-oriented interpretation of Popper’s prescription of the requirements for an open society (Popper 2002), as described in Sect. 2.2. This leads on to the roadmap surveyed in Sect. 2.3.

### 2.1 Institutionalised power and FIPA

In their paper on *Institutionalised Power* (Jones and Sergot 1996), Jones and Sergot gave the first formal logical characterisation of the notion of *counts as*, a significant factor in speech act theory (Austin 1962; Searle 1969), characterized by (Searle 1969) as constitutive rules of the form “*X counts-as Y in context C*”. The term institutionalised power refers to that characteristic feature of institutions, whereby designated agents, often acting in specific roles, are empowered to create or modify facts of special significance in that institution (*institutional facts*), through the performance of a designated action, often a speech act.

This formal logical characterisation was based on a conditional connective whose intent was to express the consequence relation between the performance by agent

A of some act (e.g. a speech act) bringing about a state of affairs  $X$  which counts as, according to the constraints of an institution, a means of bringing about a state of affairs  $Y$  (for example, in the context of an auction house, the person occupying the role of auctioneer (and no-one else) announcing “sold” counts as, in effect, a contract between buyer and seller).

In the same year that the Jones and Sergot paper was published, 1996, the Foundation for Intelligent Physical Agents (FIPA) held its inaugural meeting. Ostensibly intended to address the issue of interoperability in distributed systems with ‘intelligent’ agents (its unstated purpose was to de-risk a potentially disruptive technology), one of its key technical specifications was on agent communication. Unfortunately, this technical specification over-emphasized the psychological interpretation in Searle’s work with respect to beliefs and intentions. As a result, the FIPA-ACL (agent communication language) semantics was based on an internalised, ‘mentalist’ approach (Pitt and Mamdani 1999). In so doing, the specification overlooked Searle’s other, and more important, contention that speaking a language was to engage in a rule-governed form of behaviour (like playing a game). Consequently, the ACL ‘calculus’ omitted the constitutive aspect of conventional communication, in particular ‘counts as’.

This meant that, for example, in the FIPA standard for a contract-net protocol, it was not clear which action established the institutional fact that a contract exists between two parties. This *does* matter: in the UK, in different legal contexts, when a contract is formally recognized in law can depend on when it was signed, when it was posted, when it was delivered, or when it was opened. Without an explicit representation of which action  $X$  *counted-as*, in specific legal context  $C$ , establishing the institutional fact of a contract  $Y$  (with all its associated terms and conditions), any form of interoperability, or even any use of multi-agent systems as a technology in the emerging ‘Digital Economy’, was going to be extremely problematic. It was, in part, to address this problem that was the motivation for the ALFEBIITE project.

## 2.2 The ALFEBIITE project and the open agent society

In 1999, the EU Future and Emerging Technologies (FET) unit launched a proactive initiative on ‘universal information ecosystems’, which envisaged a complex information environment that scaled, evolved or adapted according to the needs of its ‘infohabitants’ (which, in this context, could be individuals or organizations, virtual entities acting by proxy or on their behalf, smart appliances, intelligent agents, etc.)

This was, of course, a long-term vision with intentionally futuristic terminology to encourage inter-disciplinary rather than incremental research. One of the suitably adventurous projects funded under the initiative was the ALFEBIITE project, which was a highly inter-disciplinary collaboration between computer and cognitive scientists, philosophers and researchers in information law. Its aim was to use empirical psychological and philosophical studies of trust, communication and other social relations to inform the development of a logical framework for characterizing the normative (‘legal’ or ‘ethical’) behaviour of these infohabitants, and the reasoning and decision-making processes underpinning this behaviour.

One of the key ideas to emerge (allowing for the reduction of ‘infohabitant’ to ‘agent’ and ‘ecosystem’ to ‘society’) from this research was the *open agent society*, a flexible network of heterogeneous software processes, each individually aware of the opportunities available to them, capable of autonomous decision-making to take advantage of them, and co-operating to meet transient needs and conditions. This idea picked up on concepts of ‘openness’ from both Hewitt (1986) and Popper (2002) (rather than some of the more mundane notions of openness as being some form of plug-and-play stemming from, for example, the Internet as an open system).

Using Hewitt’s notion of openness, it was possible to assume that the components shared a common language (and so could communicate), but *not* to assume that there was a shared objective, a central controller, or necessarily ideal operation. Allowing for sub-ideal operation, because of competing and conflicting objectives, or even incompetence or malice, is a particular requirement prompted by the unpredictable behaviour manifested by self-interested agents of heterogeneous provenance, exactly as we would find in a human social situation, such as an office.

Critically though, the open agent society was influenced by Popper’s definition of an open (human) society, which included qualities such as accountability and the division of power, a market economy, the rule of law and respect for rights, and the absence of any universal truth. These qualities were reflected in requirements for the open agent society as follows:

- **Accountability:** an information processing component represents a human entity (whether an individual or organization), and the form of that representation needs explicit legal definition, for example in terms of liability, delegation, mandate, ownership and control;
- **Market economy:** in an increasingly instrumented and inter-connected digital environment, ordinary citizens generate vast amounts of data, much of it personal, and its aggregation highly valuable. The generators of the data should be its owners and its beneficiaries, and also entitled to both respect for their privacy and to ‘enjoy their property’ (with the caveats that not all information is subject to intellectual property rights nor data protections rights);
- **Rule of law and respect for rights:** a key characteristic of openness in social organization lies in the empowerment of its agents, that is in establishing how agents may create their own normative relations (permissions, obligations, rights, powers, etc.), given the existence of norm-sanctioning (both allowing and enforcing) institutions; in other words, rights to self-organise are critical;
- **No universal truth:** since any one agent may only have partial knowledge of the environment, and the join of multiple agents’ knowledge bases might be inconsistent, therefore decentralised models of rights management and conflict resolution are required. Moreover, claims to certain knowledge potentially lead to closed systems of thought and expression and to path dependency, and may be a constraint on the adaptation and self-organisation required for robust, resilient and sustainable societies.

To assess our progress in building multi-agent systems that satisfy these requirements, in the next section we present a roadmap outlining the conceptual and technological developments, and the future prospects and challenges.

### 2.3 Road map

To see how the ideas have developed, and in which direction they are further developing, and what needs to be done in order to realise a functional, yet ergonomic version of the open agent society, this section considers a ‘road map’, which also provides a basis for the structure of the rest of the paper. Table 1 illustrates a partial development of the systems specification 2000–2015, from agent societies, to dynamic norm-governed systems, to micro-social systems, and finally to self-organised electronic institutions. These are reviewed in Sect. 3. A number of different action languages have been used in this specification, including the Event Calculus (Kowalski and Sergot 1986), C+ (Giunchiglia et al. 2004) and the Run-Time Event Calculus (RTEC) (Artikis et al. 2015). Furthermore, a number of different protocols have been formalised in these works, including the contract-net (Smith 1980), netbill (Sirbu 1997), floor control (Dommel and Garcia-Luna-Aceves 1995), voting (Robert et al. 2000), argumentation (Brewka 2001), dispute resolution (Vreeswijk 2000) and Paxos (Lamport 1998) protocols.

Table 1 also illustrates some of the technologies that have been developed, including the Society Visualiser, various agent architectures, the MAS simulation environments Presage and Presage2, and the RTEC engine. Some of these are reviewed in Sect. 4. Finally, some of the applications to which these technologies have been applied are shown in the final row.

The structure of the paper is in two parts, Retrospectives and Prospectives, giving rise to the ‘four quartets’ presented in Table 2. The Retrospectives part comprises Sects. 3 and 4 as mentioned above. However, we then ask the question: what happens when we inject the open agent society back into the human society which, in effect, inspired it? This is a (new) type of socio-technical system, and Sect. 5 presents a number of new research questions, while Sect. 6 considers some further ethical, legal and social implications.

**Table 1** The open agent society: a roadmap

Timeline	2000–2004	2005–2008	2009–2011	2012–2015
Model	Agent societies	Dynamic norm-governed MAS	Micro-social systems	Self-organising electronic institutions
Language	Event calculus action language C+		Run-time event calculus	
Technology	Society visualiser causal calculator agent architecture Presage		Run-time event calculus Presage-MS/Presage2	
Application	E-commerce dispute resolution ad hoc networks digital rights management		Resource allocation event recognition	

**Table 2** Structure of the paper (the four quartets)

Retrospectives	<i>MAS system specification</i>	<i>Technologies</i>
	Agent societies	Society visualiser
	Dynamic norm Gov. MAS	Agent architectures
	Micro-social systems	Presage/Presage-2
	Self-organising EI	Run-time event calculus
Prospectives	<i>Challenges</i>	<i>Implications</i>
	Computational justice	Big data
	Collective attention	The “Sharing” economy
	Electronic social capital	Privacy and uberveillance
	Polycentric governance	Design contractualism

### 3 Logic-based specification

This section reviews four different, increasingly refined, logic-based approaches to specification and representation of the open agent society, namely:

- Agent societies (Artikis and Sergot 2010): conditioning the organisation of, and interactions within, multi-agent systems according to a formal specification of institutionalised powers, permissions and obligations;
- Dynamic norm-governed multi-agent systems (Artikis 2012): extended the specification of agent societies with mechanisms for run-time adaptation;
- Micro-social systems (Pitt et al. 2011): extended the specification by the interleaving the rules of social order in agent societies with rules of social choice and rules of social exchange;
- Self-organising electronic institutions (Pitt et al. 2012): extended the specification of run-time adaptation with Elinor Ostrom’s institutional design principles (Ostrom 1990).

The primary requirement of the open agent society being addressed by this work was a characterisation, in some sense, of the “rule of law”. This entails, for example, that the laws should be clear, public, stable and applied consistently; and the processes by which laws are enacted, administered and enforced are accessible, fair and efficient. Therefore, the conceptual challenges addressed by this work include a representation of normative relations and direct execution (administration) of the law. For both of which purposes a rule-oriented logic-based specification is well-suited, with the added advantages of clarity, consistency and machine readability, as presented in Sect. 3.1.

To address issues of enactment which involve adaptation of the rules at run-time, we developed the idea of dynamic norm-governed multi-agent systems (Sect. 3.2). However, underpinning the adaptation or enactment of rules are various processes and protocols for voting, winner determination, information exchange and opinion formation, so that more social relations (besides normative relations) such as

‘gossiping’ and ‘preference’ were required. With such extensions, the agent society was increasingly becoming a microcosm of several aspects of human society, hence micro-social systems (Sect. 3.3). Finally, the requirement also to represent fundamental principles of self-determination, self-organisation and the right to self-organise in ‘fair’ and sustainable systems led to the specification of self-organising electronic institutions, as presented in Sect. 3.4.

### 3.1 Agent societies

In the course of the ALFEBIITE project, an executable specification was developed for a sub-class of computational societies that exhibited the following characteristics (Artikis and Sergot 2010):

- It adopted the perspective of an external observer and viewed societies as instances of normative systems, that is, it described the permissions and obligations of the members of the societies, considering the possibility that the behaviour of the members may deviate from the ideal.
- It explicitly represented the institutionalised powers of the member agents, a standard feature of any norm-governed interaction (see Sect. 2.1). Moreover, it maintained the long established distinction, in the study of social and legal systems, between physical capability, institutionalised power and permission.
- It provided a declarative formalisation of the aforementioned concepts by means of temporal action languages with clear routes to implementation: in particular (but not only) the Event Calculus (Kowalski and Sergot 1986). The specification could be validated and executed at design-time and run-time, for the benefit of the system designers, the agents and their owners.

This sub-class of computational society, as a computational instantiation of the ‘open agent society’, was, over the next few years, the basis for exploring various ideas in virtual organizations and for piecemeal formulation of various different protocols (for voting, dispute resolution, argumentation, e-commerce, etc), until it coalesced in the concept of dynamic norm-governed multi-agent systems.

### 3.2 Dynamic norm-governed multi-agent systems

In some multi-agent systems, environmental, social or other conditions may favor specifications that are modifiable during the system execution. Consider, for instance, the case of a malfunction of a large number of sensors in a sensor network, or the case of manipulation of a voting procedure due to strategic voting. To deal with such issues, we extended our framework for open agent societies to support ‘dynamic’ specifications, that is, specifications that are developed at design-time but may be modified at run-time by the members of a system (Artikis 2012).

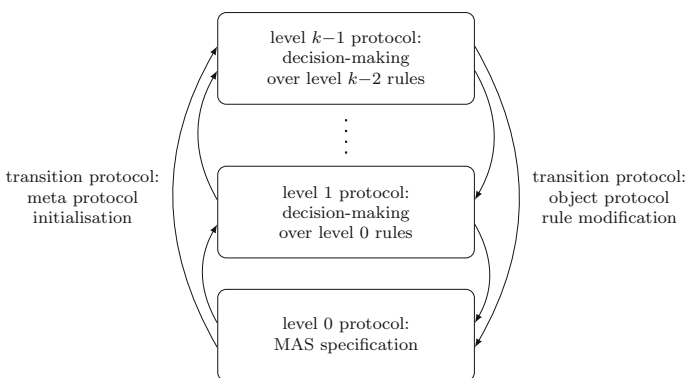
A protocol for conducting business in a multi-agent systems, such as a protocol for e-commerce or resource allocation, is called the ‘object’ protocol. At any point in time during the execution of the object protocol the participants may start a ‘meta’ protocol in order to potentially modify the object protocol rules—for



instance, replace an existing rule-set with a new one. The meta protocol may be any protocol for decision-making over rule modification, such as voting or argumentation. The participants of the meta protocol may initiate a meta-meta protocol to modify the rules of the meta protocol, or they may initiate a meta-meta-meta protocol to modify the rules of the meta-meta protocol, and so on. In general, in a  $k$ -level framework, level 0 corresponds to the main protocol, while a protocol of level  $n$ ,  $0 < n \leq k - 1$  is created, by the protocol participants of a level  $m$ ,  $0 \leq m < n$ , in order to decide whether to modify the protocol rules of level  $n - 1$ . The framework for dynamic specifications is displayed in Fig. 1.

Apart from object and meta protocols, the framework includes ‘transition’ protocols (see Fig. 1); that is, procedures that express, among other things, the conditions in which an agent may successfully initiate a meta protocol (for instance, only the members of the board of directors may successfully initiate a meta protocol in some organizations), the roles that each meta protocol participant will occupy, and the ways in which an object protocol is modified as a result of the meta protocol interactions. Furthermore, an agent’s proposal for specification change is evaluated by taking into consideration the effects of accepting such a proposal on system utility and by modelling a dynamic specification as a *metric space* (Bryant 1985). The enactment of proposals that do not meet the evaluation criteria is constrained.

A protocol specification consists of the ‘core’ rules that are always part of the specification, and the *Degrees of Freedom (DoF)*, that is, the specification components that may be modified at run-time. A protocol specification with  $l$  DoF creates an  $l$ -dimensional specification space where each dimension corresponds to a DoF. A point in the  $l$ -dimensional specification space, or *specification point*, represents a complete protocol specification—a *specification instance*—and is denoted by an  $l$ -tuple where each element of the tuple expresses a ‘value’ of a DoF. One way of evaluating a proposal for specification point change, that is, for changing the current specification of a multi-agent system, is by modelling a dynamic protocol specification as a metric space. More precisely, we compute the ‘distance’ between the proposed specification point and the current point. We constrain the process of specification point change by permitting proposals only if



**Fig. 1** A framework for dynamic specifications

the proposed point is not too ‘far’ from/‘different’ to the current point. The motivation for formalizing such a constraint is to favour gradual changes of a system specification.

Run-time specification change has long been studied in the field of argumentation—see, for example, (Loui 1992; Vreeswijk 2000). Our work was motivated by Brewka’s dynamic argument systems (Brewka 2001). Like Vreeswijk’s work (Vreeswijk 2000), these are argument systems in which the protagonists of a disputation may start a meta level debate, that is, the rules of order become the current point of discussion, with the intention of altering these rules. Run-time specification change has also been the subject of research in distributed systems. Serban and Minsky (2009), for example, presented a framework with which a law change is propagated to the regimentation devices of a multi-agent system, taking into consideration the possibility that during a ‘convergence period’ various regimentation devices operate under different versions of a law, due to the difficulties of achieving synchronized, atomic law update in distributed systems.

In the agents community, various formal models for norm change have been proposed (Broersen 2009). For instance, Boella et al. (2009) replaced the propositional formulas of the Alchourrón et al. (1985) (AGM) framework of theory change with *pairs* of propositional formulas—the latter representing norms—and adopted several principles from input/output logic (Makinson and Torre 2000). The resulting framework includes a set of postulates defining norm change operations, such as norm contraction. Governatori et al. (2007) (Governatori and Rotolo 2008a, b) have presented variants of a Temporal Defeasible Logic (Governatori et al. 2005, 2006) to reason about different aspects of norm change. These researchers have represented meta norms describing norm modifications by referring to a variety of possible time-lines through which the elements of a norm-governed system, and the conclusions that follow from them, can or cannot persist over time.

A key difference between our work and related research is that the framework of dynamic norm-governed multi-agent systems formalises the transition protocol leading from an object protocol to a meta protocol. More precisely, it distinguishes between successful and unsuccessful attempts to initiate a meta protocol, evaluates proposals for specification change by modelling a specification as a metric space, and takes into consideration the effects of accepting a proposal on system utility, by constraining the enactment of proposals that do not meet the evaluation criteria. It also formalises procedures for role-assignment in a meta level.

### 3.3 Micro-social systems

In addition to run-time modification, the specification of agent societies also needs to take into consideration the interplay of the adaptation of these rules with the processes behind the *selection* of the adaptation, based on the following argument. Since, by Popper’s assertion that there should be no universal truth and no centralised control, we are faced with a situation with local information, partial knowledge and (possibly) inconsistent union (of individual knowledge bases). What each network node ‘sees’ is the result of actions by (millions) of actors, some of which are not known, and even those actions which are known, the actor’s motive

may be unknown. Moreover, what a node ‘thinks’ it ‘sees’ may not be consistent with the ‘opinion’ of other nodes. It follows from this that if there is no single central authority that is controlling or coordinating the actions of others, then emphasis is on local decision-making based on locally available information and the perception of locally witnessed events. From this it follows that in the absence of perfect knowledge there is no perfect form of government, therefore the next best thing is a government prepared to modify its policies according to the needs, requirements and expressed opinions of its ‘population’.

In other words, *social organization* is both the requirement for and consequence of any networked computing which impacts on personal, legal or commercial relationships between real-world entities (people or organizations).

Therefore a micro-social system was defined as a distributed computer system or network where the interactions, relationships and dependencies between components is a microcosm of aspects of a human society. The aspects of human society in which we were interested included communication protocols, organizational rules and hierarchies, network structures, inter-personal relationships, and other processes of self-determination and self-organization. In particular, though, we pick out the following three primary set of rules which underpin the social intelligence required to realise a micro-social system which can be applied to a number of issues affecting ad hoc networks (such as dynamism, conflicts, sub-ideality, security and continuity):

- **Rules of Social Order.** Micro-social systems consist of agents whose actions have a conventional significance according to the social rules of an institution; actions are therefore norm-governed.
- **Rules of Social Choice.** Micro-social systems consist of heterogeneous, self-interested agents that can have conflicting preferences in decision-making situations; these agents’ preferences can be aggregated by taking votes over potential outcomes.
- **Rules of Social Exchange.** Micro-social systems, being both open and local, will require agents to gain knowledge over time by exchanging information with each other. Each agent must therefore be capable of reliable opinion formation, based on the opinions gathered from the contacts in their own social networks.

Although the idea of micro-social systems produced some useful insight into the abstract management and operation of multi-agent systems (Pitt et al. 2011), it became clear that, for practical applications such as resource allocation, the interleaving of rules of social order, social choice and social exchange were underpinned by a more fundamental concept: that of *self-organisation*.

### 3.4 Self-organising electronic institutions

Different types of open system, ranging from virtual organizations to cloud and grid computing and to ad hoc, vehicular and sensor networks, all face a similar problem: how to arrange the distribution of resources so that each participating component can successfully complete its tasks—given the constraints that firstly, there may be insufficient resources for every component to successfully complete its tasks;

secondly, that the components may be competing for these resources; and thirdly, that the components may even ‘misbehave’ in the competition for those resources.

This is, in effect, a *common-pool resource* management problem. As it turns out, these problems, and possible solutions, have been well-studied in the social, economic and political sciences. One explanation of a workable solution was offered by the economic and political scientist Elinor Ostrom, whose extensive fieldwork (Ostrom 1990) demonstrated that time and again, throughout human history and geography, communities were able to maintain and sustain common-pool resources, even over generations providing explicit empirical evidence against theoretical ‘proofs’ that collective action on a sufficiently ‘large’ scale will never happen (Olson 1965; Hardin 1968).

In particular, Ostrom observed that when resources were successfully sustained, it was because the individuals in the community voluntarily agreed to regulate and constrain their behaviour according to sets of conventional rules, i.e. rules which were enacted, administered and enforced by the members of the community themselves. These sets of rules, and their self-administration, Ostrom called *self-governing institutions*. However, she further observed that just having a set of rules was not in itself enough: sometimes such institutions were able to sustain the common-pool resource, sometimes not. From her many case studies, though, she was able to identify a set of common features, which if they were all present, enabled successful maintenance of the resource, and failure occurred if one or more features were missing. She then went one step further, and suggested that, when faced with a collective action problem of this kind, if we know the features of a successful self-governing institution, then rather than trust to ‘luck’ or common sense to evolve a solution with the right features, it should be possible to *design* an institution with the requisite features. Accordingly, she specified a set of eight institutional design principles for designing successful self-governing institutions.

In Pitt et al. (2012), it was shown how the framework of dynamic norm-governed systems could be used to axiomatise Ostrom’s institutional design principles in computational logic using the Event Calculus (cf. Sect. 4.2, below). In particular, it was shown that there was a correspondence between the first six design principles and protocols that had been developed both in ALFEBIITE and subsequent projects

**Table 3** Ostrom’s institutional design principles and corresponding protocols

Ostrom principle	Protocol
Clearly defined boundaries	Role-assignment, floor-assignment and access control
Congruence between provision and appropriation rules and the environment	Gossiping and meta-protocols
Those affected by the rules participate in their selection and modification	Voting by enfranchised (empowered) agents
Monitoring by members or agencies appointed by them	Event recognition and reporting
Graduated sanctions	Argumentation
Cheap and efficient dispute resolution protocols	Alternative dispute resolution

(see Table 3). Instances of dynamic norm-governed systems which implemented the principles were defined as *self-organising electronic institutions*.

An experimental multi-agent testbed was implemented, using the PreSage simulation platform (see Sect. 4.1) with a number of agents having to appropriate from a common-pool resource in a series of rounds, with each of the six principles optionally enabled or not in a self-organising electronic institution. The results showed that the more principles that were enabled, the more likely it was that the institution could sustain the resource and maintain ‘high’ levels of membership.

Work on the formalisation of Ostrom’s seventh principle, minimal recognition of the right to self-organise, and the eighth principle, nested enterprises, has been pursued in the context of *holonic institutions* (Pitt and Diaconescu 2015; Diaconescu and Pitt 2015).

## 4 Technologies

The previous section was concerned with the logic-based specification of issues of conceptual significance in the representation of the open agent society. This section is concerned with two technologies used for *reasoning* about, and in, the open agent society, namely Presage2 and the Run-Time Event Calculus (RTEC). Presage2 is a platform for agent animation and multi-agent system simulation, while the Run-Time Event Calculus (RTEC) addressed the performance limitations of the Event Calculus for executing certain run-time tasks, especially event recognition and dealing with temporally long narratives. This is especially important in experimenting with self-organising electronic institutions, where the number of events to process can be ‘large’ (hundreds of thousands).

### 4.1 Agent-based modelling, simulation and animation

Presage2 (Macbeth et al. 2014) is one of many software platforms which would be available for agent-based modelling or multi-agent based simulation. A general survey of the basic features of agent-based modelling tools allows some evaluation of which platforms are suitable for particular tasks (Nikolai and Madey 2009). Alternative platforms which are also suitable for simulation and animation tasks are Netlogo (Wilensky 2011), Repast (North et al. 2006) and MASON (Luke et al. 2004). The feature sets of these platforms are quite similar, although arguably Netlogo is more oriented towards less experienced programmers, with usage being largely graphical user-interface based; while MASON and Presage2 sit at the other end of the scale, with powerful capabilities for more experienced programmers and all functionality being code-based. Repast lies somewhere in between these two points.

Presage2 itself is a second generation agent animation and multi-agent simulation platform, developed from the original Presage platform (Neville et al. 2008) and its variant Presage-*MS* (Carr et al. 2009), which was developed to work with the idea of metric spaces in dynamic norm-governed systems. In this section, we describe only Presage2.

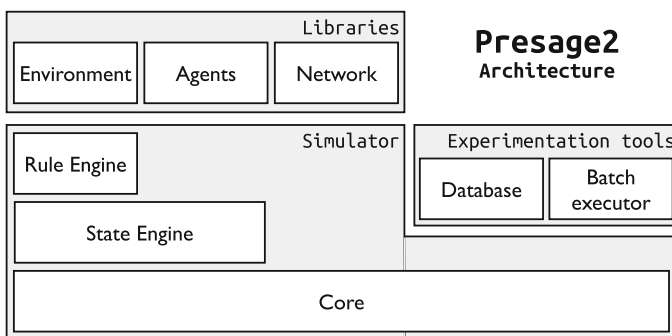
Presage2 is designed as prototyping platform to aid the systematic modelling of systems and generation of simulation results. To this end we require that the platform be able to: simulate computationally intensive agent algorithms with large populations of agents; simulate the networked and physical environment in which the agents interact, including dynamic external events; reason about social relationships between agents including their powers, permissions and obligations; support systematic experimentation; and support aggregation and animation of simulation results.

The platform is composed of packages which work together to control the execution of a simulation as illustrated in Fig. 2:

- Core: controls and executes the main simulation loop and core functions.
- State engine: Stores and updates simulation state.
- Environment and agent libraries: Implementations of common use cases which can be used in the environment and/or agent specifications.
- Communication network simulator: Emulates a dynamic, inter-agent communication network.
- Database: Enables storage of simulation data and results or analysis.
- Batch executor: Tools to automate the execution of batches of simulations.

The Presage2 simulator's core controls the main simulation loop as well as the initialization of a simulation from parameter sets. The simulation uses discrete time, with each loop being a single time step in the simulation. Each time step each agent is given a chance to perform physical and communicative actions, and the simulation state is updated according to these actions as well as to any external events. Agent architecture and computational complexity is not limited (except by the limitations of the computer running the simulation). The platform will wait for every agent's function to terminate before moving to the next time step.

A key concept of multi-agent simulation is that the agents share a common environment. The state engine package simulates this environment as a state space. This package allows the user to control two important functions: The observability of state for each agent, and the effect of an agents action on the state. The former



**Fig. 2** Presage2 agent animator and multi-agent simulator

specifies what state an agent can read given the current environment state, while the latter determines a state change given the current state and the set of all actions performed in the last timestep. The user defines these functions from a set of modules. Each module can be seen as an independent set of rules regarding observability and/or state changes. This method allows for behavioral modules to be slotted into a system without conflict, building up complex system rules from the composition of modules.

An optional Presage2 configuration uses the Drools production rule engine, based on the Rete algorithm (Forgy 1982), as the underlying state storage engine, encapsulating all features of the initial implementation plus additional benefits afforded by Drools. Using Drools offers a much richer state representation. While the initial version uses raw data points with text strings to reference each one, with Drools we can store structured and relational data in the state. Drools also features a forward chaining rule engine, so that declarative rules can be used to modify the state in the engine. Their declarative nature is close to that of action languages such as the Event Calculus (indeed work has been done to allow Event Calculus predicates to be used in Drools directly (Bragaglia et al. 2012), and it has been shown how a dynamic specification in the Event Calculus can be ported into Presage2 via a (currently manual) translation into Drools.

Presage2 is open source and available under the LGPL license from <http://www.presage2.info>. Work on another optional Presage2 configuration is in progress, this one replacing Drools with RTEC, the Run-Time Event Calculus.

## 4.2 The run-time event calculus

The Event Calculus is a logic programming formalism for representing and reasoning about events and their effects (Kowalski and Sergot 1986). It has been frequently used for specifying and reasoning about (open) multi-agent systems due to its simplicity and flexibility. However, the Event Calculus also has a number of well-known limitations. One of these is an issue of *scale*; that is, as the number of agents increases, and/or the number of exchanged messages increases, then the performance and efficiency deteriorate unacceptably. To deal with very large multi-agent systems, we have developed the ‘Event Calculus for Run-Time reasoning’ (RTEC) (Artikis et al. 2015). RTEC includes various optimization techniques for an important class of computational tasks, specifically those in which given a record of what events have occurred (a ‘narrative’) and a set of axioms (expressing the specification of a multi-agent system), we compute the values of various facts (denoting institutionalised powers, permissions, and other normative relations) at specified time points. RTEC thus provides a practical means of informing the decision-making of the agents and their owners, and the system designers. In what follows we briefly discuss the use of RTEC for specifying and executing very large multi-agent systems.

The time model of RTEC is linear and includes integer time-points. Variables start with an upper-case letter, while predicates and constants start with a lower-case letter. *Fluents* express properties that are allowed to have different values at different points in time. The term  $F = V$  denotes that fluent  $F$  has value  $V$ —Boolean

fluents are a special case in which the possible values are **true** and **false**.  $\text{holdsAt}(F = V, T)$  represents that fluent  $F$  has value  $V$  at a particular time-point  $T$ .  $\text{holdsFor}(F = V, I)$  represents that  $I$  is the list of the maximal intervals for which  $F = V$  holds continuously.  $\text{holdsAt}$  and  $\text{holdsFor}$  are defined in such a way that, for any fluent  $F$ ,  $\text{holdsFor}(F = V, T)$  if and only if  $T$  belongs to one of the maximal intervals of  $I$  for which  $\text{holdsFor}(F = V, I)$ .

An *event description* in RTEC includes rules that define the event instances with the use of the  $\text{happensAt}$  predicate, the effects of events with the use of the  $\text{initiatedAt}$  and  $\text{terminatedAt}$  predicates, and the values of the fluents with the use of the  $\text{holdsAt}$  and  $\text{holdsFor}$  predicates, as well as other, possibly a temporal, constraints. Table 4 summarises the RTEC predicates available to the MAS specification developer. We represent the actions of the agents and the environment by means of  $\text{happensAt}$ , while the state of the agents and the environment are represented as fluents. In multi-agent system execution, therefore, the task is to compute the maximal intervals for which a fluent representing an agent or environment variable, such as the institutionalised powers of an agent, has a particular value continuously.

For a fluent  $F$ ,  $F = V$  holds at a particular time-point  $T$  if  $F = V$  has been initiated by an event that has occurred at some time-point earlier than  $T$ , and has not been terminated at some other time-point in the meantime. This is an implementation of the *law of inertia*. To compute the *intervals*  $I$  for which  $F = V$ , that is,  $\text{holdsFor}(F = V, I)$ , we find all time-points  $T_s$  at which  $F = V$  is initiated, and then, for each  $T_s$ , we compute the first time-point  $T_f$  after  $T_s$  at which  $F = V$  is terminated. The time-points at which  $F = V$  is initiated and broken are computed by means of domain-specific  $\text{initiatedAt}$  and  $\text{terminatedAt}$  rules. Consider the following examples from self-organising electronic institutions (see Sect. 3.4):

**Table 4** Main predicates of RTEC

Predicate	Meaning
$\text{happensAt}(E, T)$	Event $E$ occurs at time $T$
$\text{holdsAt}(F = V, T)$	The value of fluent $F$ is $V$ at time $T$
$\text{holdsFor}(F = V, I)$	$I$ is the list of the maximal intervals for which $F = V$ holds continuously
$\text{initiatedAt}(F = V, T)$	At time $T$ a period of time for which $F = V$ is initiated
$\text{terminatedAt}(F = V, T)$	At time $T$ a period of time for which $F = V$ is terminated
$\text{union\_all}(L, I)$	$I$ is the list of maximal intervals produced by the union of the lists of maximal intervals of list $L$
$\text{intersect\_all}(L, I)$	$I$ is the list of maximal intervals produced by the intersection of the lists of maximal intervals of list $L$
$\text{relative\_complement\_all}(I', L, I)$	$I$ is the list of maximal intervals produced by the relative complement of the list of maximal intervals $I'$ with respect to every list of maximal intervals of list $L$



$$\begin{aligned}
&\text{initiatedAt}(\text{allocated}(Ag_R, W, C) = R, T) \text{ if} \\
&\quad \text{happensAt}(\text{allocate}(Ag_H, Ag_R, R, W, C), T), \\
&\quad \text{holdsAt}(\text{pow}(\text{allocate}(Ag_H, Ag_R, R, W, C)) = \text{true}, T)
\end{aligned} \tag{1}$$

$$\begin{aligned}
&\text{initiatedAt}(\text{ifpool}(W, C) = P - R, T) \text{ if} \\
&\quad \text{happensAt}(\text{allocate}(Ag_H, Ag_R, R, W, C), T), \\
&\quad \text{holdsAt}(\text{pow}(\text{allocate}(Ag_H, Ag_R, R, W, C)) = \text{true}, T), \\
&\quad \text{holdsAt}(\text{ifpool}(W, C) = P, T)
\end{aligned} \tag{2}$$

$$\begin{aligned}
&\text{initiatedAt}(\text{leg\_claim\_q}(W, C) = \text{Rest}, T) \text{ if} \\
&\quad \text{happensAt}(\text{allocate}(Ag_H, Ag_R, R, W, C), T), \\
&\quad \text{holdsAt}(\text{pow}(\text{allocate}(Ag_H, Ag_R, R, W, C)) = \text{true}, T) \\
&\quad \text{holdsAt}(\text{leg\_claim\_q}(W, C) = [Ag_R \mid \text{Rest}], T)
\end{aligned} \tag{3}$$

In brief, the exercise of the institutionalised power (*pow*) to perform the *allocate* action initiates values for three fluents: an allocation of resources *R* to an agent *Ag<sub>R</sub>*, a reduction in the amount of pooled resource *ifpool* to allocate, and removing the agent *Ag<sub>R</sub>* from the front of the legitimate claim queue.

The maximal intervals during which an agent *Ag<sub>R</sub>* is allocated resources *R*, for example, are computed using the built-in RTEC predicate `holdsFor` from rule (1) and other similar rules, not shown here, terminating the *allocated* fluent.

RTEC includes various optimisation techniques that allow for very efficient and scalable multi-agent system execution (Artikis et al. 2015). A form of caching stores the results of sub-computations in the computer memory to avoid unnecessary recomputations. A simple indexing mechanism makes RTEC robust to events that are irrelevant to the computations we want to perform and so RTEC can operate without data filtering modules. Moreover, a ‘windowing’ mechanism supports real-time MAS execution. In contrast, no Event Calculus system (e.g. Paschke and Bichler (2008); Chesani et al. 2010, 2013) ‘forgets’ or represents concisely the event history. RTEC is available at <https://github.com/aartikis/RTEC>.

## 5 Some conceptual challenges

In the previous part, we outlined the logic-based specification, and the development of associated technologies, for representing and reasoning with concepts of the open agent society. This completes the Retrospective part of the paper. In the Prospective part, we address the research question: what are the prospects for using the open agent society as the basis for developing socio-technical systems, or other collective adaptive systems? In particular, we are concerned with the potential consequences of embedding the open agent society in the human society that originally inspired it.

In this section, we outline a quartet of research challenges for which the concepts of self-organising electronic institutions and logic-based complex event recognition

are critical for the development of socio-technical systems for/based on the open agent society. These challenges are:

- *computational justice*, as the study of some form of ‘correctness’ in the outcomes from qualitative algorithmic deliberation and decision-making;
- (interoceptive) *collective attention*, as an attribute (internal sense of well-being) of communities that helps solve collective action problems;
- (electronic) *social capital*, as a complexity-reducing short-cut in cooperation dilemmas that have multiple equilibria or are computationally intractable (for example,  $n$ -player games);
- *polycentric self-governance*, reconciling potentially conflicting interests by giving consideration to “common purposes” within multiple centres of decision-making.

Note that the issues of computational justice and polycentric self-governance are related to the open (agent) society requirement for the rule of law. Polycentric governance also impacts on the issues of decentralisation and “no universal truth”, as well as the requirement for a market economy. Collective attention is also a corollary of “no universal truth”: collective action has to proceed on the basis of negotiation, agreement and compromise. Electronic social capital is related to the requirement for a market economy, although from the social capital perspective we are primarily concerned with leveraging positive externalities arising from relational rather than transactional economies.

Finally, in Sect. 6, we consider some ethical, legal and social implications of this research question.

## 5.1 Computational justice

Following on from the original experiments involving resource allocation using self-organising electronic institutions designed according to Ostrom’s principles, a secondary question concerns enduring that the allocation is *fair*. In another experimental testbed implement with Presage2 (Sect. 4.1), we axiomatised Rescher’s theory of distributive justice based on legitimate claims (Rescher 1966). Results showed that, especially in an economy of scarcity, where there was no ‘fair’ resource allocation in any one round, the agents could self-organise the allocation so that it was ‘fair’ over many rounds (Pitt et al. 2014) (note that we used the Gini index, as a commonly used measure of inequality, as the metric for ‘fairness’). This treatment of distributive justice, in conjunction with the formalisation of Ostrom’s principles for self-determination as a form of ‘natural’ justice, coalesced into a wider and deeper programme of research into *computational justice*.

Computational justice (Pitt et al. 2013) is an interdisciplinary investigation at the interface of computer science and philosophy, economics, psychology and jurisprudence, enabling and promoting an exchange of ideas and results in both directions. One of its main goals is to introduce notions borrowed from Social Sciences, such as fairness, equity, transparency and openness into computational

settings. It is also concerned with exporting the developed mechanisms back to Social Sciences, both to better understand their role in social settings, as well as to leverage the knowledge gained in computational settings.

Although ‘justice’ is a concept open to many definitions, the study of computational justice focuses on capturing some element of ‘correctness’ in algorithmic deliberation and/or decision-making. Therefore it focuses on the following ‘qualifiers’, which have been variously used in the social sciences:

- *Natural* or *social* justice, which is concerned with the right of inclusion and participation in the decision making processes affecting oneself.
- *Distributive* justice, which deals with fairly allocating resources amongst a set of agents.
- *Retributive* justice, which addresses the issues of monitoring, reporting, adjudicating on and sanctioning of non-compliant behaviour.
- *Procedural* justice, which considers evaluating whether conventional procedures are ‘fit-for-purpose’.
- *Interactional* justice, which is concerned with the subjective view of the agents and whether they feel they are being treated fairly by the decision makers.

All these qualifiers are important in socio-technical systems, since they address different issues that require ‘conditioning’ in such situations. For example, when distributing collectivized resources, it might be tempting to go for an optimal allocation in terms of maximizing some overall utility. While such an utilitarian view might be the most appropriate in some cases, in socio-technical systems (or any other system where participants can somehow evaluate their satisfaction and act depending on it) it might be better to seek allocations that may be sub-optimal but that take into account the notion of fairness, thereby increasing sustainability in the longer-term at the expense of optimality in the short-term.

Similarly, if a participant violates some norm, it could be subject to some kind of punishment, penalty or sanction. This can be seen either as a direct consequence of the wrong-doing (retributivism view) or as a deterrent for future wrong-doing by the agent being punished as well as by the agents observing the punishment. However, the punishment should be proportional to the offence including both the extent of the violation, as well as recidivism (that is, repeated offences by the same agent). However, note the role of forgiveness (Vasalou et al. 2008) is also an important element of retributive justice.

Natural or social justice is concerned with issues such as the inclusivity of participants in decision-making, for instance to decide how resources are allocated, or how a decision should be made. Inclusive participation of this kind would provide the system with the feature of self-organisation, in the sense of self-governing the resources, and ensure that those who are subject to a set of rules participate in their selection and modification.

Procedural justice is required to provide governance mechanisms which are ‘fit for purpose’ (Pitt et al. 2013), that is, addressing the following sorts of question. Are the rights of members of the institution to participate in collective-choice arrangement adequately represented and protected? Is an institution where decisions

are made by one actor who justifies its decision ‘preferable’ to an institution where the decision is made by a committee that does not offer such justification? Is an institution which expends significant resources on determining the most equitable distribution ‘preferable’ to one that uses a cheaper method to produce a less fair outcome, but has more resources to distribute as a consequence?

Interactional justice allows the participants to make a subjective assessment of whether they are being treated fairly. For instance, in a case of scarcity of resources, if the participants were informed about this scarcity, they would probably better understand not being allocated any resources. Contrarily, if they are not informed about this issue, they might think that they are not being allocated resources because the decision makers are biased towards other participants.

## 5.2 Collective attention

The development of collective awareness has been advocated as enhancing the choice of sustainable strategies by the members of a community and therefore ensuring the adoption of successful strategies (Sestini 2012). However, in communities in which collective awareness is barely present, individual members may experience a diminished appreciation of the global situation and present constrained flexibility in adjustment to change because they do not share the same comprehension of situation with others. They are also less willing to obey the norms and rules set by the community because they do not feel themselves as members of community and are not aware of others seeing them as ones. They understand the situation they are in from a micro-level perception and might additionally recognize the macro-level description of the situation, however, they might not be aware of interactions occurring at the meso level. As a result, individuals make decisions that are sub-optimal from the perspective of the whole system making it less fair, more inefficient and so vulnerable to collapse through instability. Therefore, collective awareness is critical to the formation of socio-technical systems.

It has been argued that collective awareness occurs “when two or more people are aware of the same context and each is aware that the others are aware of it” (Kellogg and Erickson 2002). This awareness of others’ awareness has been indicated as a critical element of collaboration within the communities, especially virtual ones such as computer-mediated communities (Daassi and Favier 2007). In socio-technical systems, an alternative approach might be sought by moving away from mutual knowledge and taking a multi-modal approach (Jones 2002), and define collective awareness of some proposition  $\phi$  as a two part relation: firstly a *belief* that there is a group, and secondly an *expectation* that if someone is a member of that group, then they believe the proposition  $\phi$ .

However, if collective awareness is about being aware of an issue, then something additional is required to do something about it—like the difference between ‘hearing’ and listening. Therefore, we aim to go from collective awareness to *collective attention*.

We define collective attention as “an attribute of communities that helps them solve a collective action problem”, that is, analogous to the way that social capital is defined by (Ostrom and Ahn 2003) as “an attribute of individuals that helps them

solve collective action problems”. Without this community attribute, individuals may take actions that are sub-optimal from a community-wide perspective, leading to diminished utility and sustainability. Individuals may understand the situation they are in from a micro-level perspective (for example, in a power system, reducing individual energy consumption) and might additionally recognize the macro-level requirement (meeting national carbon dioxide emission pledges). However, they might not be aware of interactions occurring at the meso-level which are critical for mapping one to the other.

Therefore, collective attention has a critical role in the formation of electronic institutions, the regulation of behaviour within the context of an institution, and the direction (or selection) of actions intended to achieve a common purpose. If we consider collective attention as being different from mutual knowledge, and base it on expectations for resolving collective action problems instead, then from the human-participant perspective we identify certain requirements as necessary conditions for achieving collective awareness as a precursor to collective action in socio-technical systems. These requirements are:

- Interface cues for collective action, that is, that participants are engaged in a collective action situation;
- Visualisation: appropriate presentation and representation of data, making what is conceptually significant perceptually prominent;
- Social networking: fast, convenient and cheap communication channels to support the propagation of data;
- Feedback: individuals need to know that their (‘small’, individual) action  $X$  contributed to some (‘large’, collective) action  $Y$  which achieved beneficial outcome  $Z$ ;
- Incentives: typically in the form of social capital (Ostrom and Ahn 2003), itself identified as an attribute of individuals that helps to solve collective action problems.

However, preliminary experiments have shown that people have insufficient attention to be sufficiently pro-active in monitoring and responding to all the changes in their environment. Complex event recognition is clearly going to be a critical technology in being able to detect situations that genuinely require user attention, from which the appropriate interface cues and visualisation methods can be drawn.

### 5.3 Electronic social capital

In social systems, it has been observed that *social capital* is an attribute of individuals that enhances their ability to solve collective action problems (Ostrom and Ahn 2003). Social capital takes different forms, for example trustworthiness, social networks and institutions. Each of these forms is a subjective indicator of one individual’s expectations of how another individual will behave in a strategic game: that is, if the individual has a high reputation (trustworthiness) for honoring agreements and commitments, or it is known personally to be reliable (social

network), or a belief there is a set of (institutional) rules (triggering expectations that someone else's behaviour will conform to those rules, and that they will be punished if they do not) [(cf. Jones (2002)].

Therefore, we propose that an electronic form of social capital could be used as an attribute of participants in a socio-technical system, to enhance their ability to solve collective action problems, by reducing the costs and complexity in joint decision-making in repeated pairwise interactions (Petruzzi et al. 2014). Two key features of this framework are firstly, the use of institutions as one social capital attribute, and secondly, the use of the Event Calculus to process events which update all the social capital attributes.

We note, en passant, that as a further direction of research, the role of electronic social capital and its relation to cryptocurrencies such as Bitcoin and Venn, for example in the creation of incentives and alternative market arrangements, needs to be fully explored.

#### 5.4 Polycentric self-governance

It is well-known that managing critical infrastructure, like a national energy generation, transmission and distribution network, will necessarily involve multiple agencies with differing (possibly competing or even conflicting) interests, effectively creating an “overlay” network of relational dynamics which also needs to be resolved. Furthermore, there is some, not always well-understood, inter-connection of public and private ownership that makes the overall system both stable and sustainable.

Therefore, in analysing any such complex system, it is critical to identify the agencies and determine its institutional *common purpose*, what the agency (through its institution) is trying to achieve or maintain, by coordination with other institutions and by the decision-making of its members. Such analysis makes it possible to understand the ‘ecosystem’ of institutions and how they fit together as collaborators or competitors, based on the nature of their purposes and the scope of their influence.

It is then an open question if the ecosystem of institutions can be represented using holonic systems architectures, which are capable of achieving large-scale multi-criteria optimisation (Frey et al. 2013). The key concept here is the idea of *holonic institutions*, whereby each institution is represented as a holon, which can be aggregated or decomposed into supra- and sub-holons respectively [Diaconescu and Pitt (2015), Pitt and Diaconescu (2015)].

The outcome of a positive answer to this question would be twofold. Firstly, that would support polycentric self-governance at all scales of the system, and in particular would support *subsidiarity* (the idea that problems are solved as close to the local source as possible). Secondly, it would encourage the institutions to recognize their role in the “scheme of things” in relation to institutions at the same, higher and lower levels. This is an essential requirement for *adaptive* institutions (RCEP 2010) and this establishment of “systems thinking” as a commonplace practice.

## 6 Some ethical, legal and social issues

Finally, we note that the idea of collective intelligence, including both humans and software agents, and agents as an enabling technology for added-value services, was a key part of the original vision: the open agent society for the user-friendly digital society (Pitt 2005). In this section, we consider some legal and ethical issues that arise when the concept of the open agent society, with its computational instantiations and associated technologies, encounters, and even potentially clashes with, other technological developments like big data, ubiquitous computing, implant devices and the (so-called) sharing economy.

Sections 6.1 and 6.2 require us to reflect more closely on the idea of a market economy, when the transactions are potentially so asymmetric: for example, the positive externalities created by big data when users' information is provided for free or for a relatively rudimentary service; or when sharing economy applications result in work at the edge but value primarily in the middleware. Section 6.3 is also related to this issue, but both this section and Sect. 6.4 on design contractualism are really about accountability: knowing who is doing what with whose data; and knowing what code it is 'safe' to execute.

### 6.1 Big data and knowledge commons

Hess and Ostrom (2006) were concerned with treating knowledge as a shared resource, motivated by the increase in open access science journals, digital libraries, and mass-participation user-generated content management platforms. They then addressed the question of whether it was possible to manage and sustain a knowledge commons, using the same principles used to manage ecological systems with natural resources. A significant challenge in the democratisation of Big Data is the extent to which formal representations of intellectual property rights, access rights, copy-rights, etc. of different stakeholders can be represented in a system of computational justice and encoded in Ostrom's principles for knowledge commons. As observed in Shum et al. (2012), the power of Big Data and associated tools for analytical modelling: "... should not remain the preserve of restricted government, scientific or corporate élites, but be opened up for societal engagement and critique. To democratise such assets as a public good, requires a sustainable ecosystem enabling different kinds of stakeholder in society".

It could be argued that Ostrom's institutional design principles reflect a pre-World Wide Web era of scholarship and content creation, and despite their original insightful work (Ostrom and Hess 2006), these developments make it difficult to apply the principles to non-physical shared sources such as data or knowledge commons, and a further extension of the theory is required to develop applications based on participatory sensing for the information-sharing economy (Macbeth and Pitt 2014). However, this sharing economy is potentially asymmetric in its distribution of risks, investments and rewards.

## 6.2 The “Sharing” economy

Reich (2015) suggests that we are moving towards a economy in which all routine, predictable work is automated or performed by robots, and the profits go to the owners of the robots; and all less predictable work is performed by human beings, and all the profits got to the owners of the middleware (that is, in a peer-to-peer economy, while the intelligence is at the edge, the value is in the network). This model affects not just taxi drivers, plumbers and hotel owners, but will increasingly effect professions such as law, medicine and academia, given the rise of online legal services (such as dispute resolution), health provision, and MOOCs (massive open online courses), and (in the UK) the rise of zero-hour contracts to service those courses.

## 6.3 Privacy and ubersurveillance

The issue here is whether intelligent agents will exacerbate a perceived trend from pervasive computing to persuasive computing and ultimately to coercive computing. The threat is sufficient for the central premise of Dave Eggers’ science fiction novel *The Circle* (McSweeney’s, 2013), that opting out from providing personal data is tantamount to theft, to appear plausible.

It is in this context that programmes for Privacy by Design (Cavoukian 2012), imposing limits on ubersurveillance (Michael et al. 2014) and the ethical issues of wearable, bearable and implant technology in the context of Big Data and the Internet of Things (IoT) (Michael and Michael 2012; Perakslis et al. 2015) are set. Designers should heed “the principles incorporated in the European Union and international treaties as well as the laws of EU member states: the precautionary principle, purpose specification principle, data minimization principle, proportionality principle, and the principle of integrity and inviolability of the body, and dignity” (Perakslis et al. 2015).

## 6.4 Design contractualism

This requirement is based on the observation that affective and pervasive applications are implemented in terms of a sense/respond cycle, called the affective loop (Goulev et al. 2004) or the biocybernetic loop (Serbedzija 2012). The actual responses are determined by decision-making algorithms, which should in turn be grounded within the framework of a mutual agreement, or a social contract (Rawls 1971). This contract should specify how individuals, government, and commercial organizations should interact in a digital, and digitized, world.

In the context of affective computing, this contractual obligation has been called design contractualism by Reynolds and Picard (Reynolds and Picard 2004). Under this principle, the designer makes moral or ethical judgements, and encodes them in the system. In fact, there are already several prototypical examples of this, from the copyleft approach to using and modifying intellectual property, to the IEEE Code of Ethics and the ACM Code of Conduct, to and TrUSTe self-certifying privacy seal.



In some sense, these examples are a reflection of ideas of Lessig that Code is Law (Lessig 2006), or rather in this case, Code is Moral Judgement.

Returning to the security implications of user-generated content and Big Data (see above), design contractualism (Pitt 2012) also underpins the idea of using implicitly-generated data as input streams for Big Data, and treating that data as a knowledge commons (Hess and Ostrom 2006; Macbeth and Pitt 2014). Using the principles of self-governing institutions for managing common pool resources identified by Ostrom (1990), we again advocate managing Big Data from the perspective of a knowledge commons (see above). Design contractualism, from this perspective, effectively defines an analytical framework for collecting and processing user-generated content input to Big Data as a shared resource with both normative, social dimensions, and ecological dimensions.

The normative dimension is the existence of institutional rules embodying the social contract. The social dimension is the belief that there are these rules and that others behaviour will conform to these rules, as a trust shortcut (Jones 2002). The ecological dimension is that the principles offer some protection against poisoning the data well, for example by the “merchants of doubt” identified by Oreskes and Conway (2010).

In short, to the extent that design contractualism relates accountability in the open agent society, or in socio-technical systems with intelligent components operating on behalf of individuals or organizations, it corresponds to an exhortation “don’t be evil”, a worthy corporate motto that seems sadly abandoned.

## 7 Summary and conclusions

In summary, we have charted the development of the ‘open agent society’ from its origins and early formulation in the ALFEBIITE project, as an electronic variation on the themes of Popper’s open society, and identified the requirements of accountability, the rule of law, a market economy and decentralisation. We then reviewed the series of conceptual challenges that have motivated the logical reification of the open agent society from dynamic norm-governed systems for run-time adaptation through to self-organising electronic institutions for fair and sustainable resource allocation. We have also presented two powerful technologies that have been developed in tandem: this includes the multi-agent simulation and animation platform Presage2 and the run-time calculus and reasoning engine for logic-based complex event recognition (RTEC). Both Presage2 and RTEC are available as open source.

Looking forward, we then considered the impact of using the open agent society as a platform for developing socio-technical systems and other collective adaptive systems. We outlined a quartet of research challenges, covering computational justice, polycentric governance, collective attention and electronic social capital, and a quartet of social implications, covering big data, the sharing economy, privacy and design contractualism. These, we believe, raise serious issues for the conceptualisation and formalisation of the open agent society and the development of associated technologies, platforms and application based thereon; and we contend

that the continued inter-disciplinary collaboration of computer scientists and information and IT lawyers will be critical in successfully addressing these issues.

In this paper, we have necessarily focused on the institutional aspects that resulted from the collaboration between the computer scientists and lawyers and philosophers, in particular Jon Bing and Andrew Jones. A separate paper could be written on the outcomes of the collaboration between the computer scientists and the cognitive scientists, in particular Cristiano Castelfranchi, whose original work inspired formal models of trust and economic reasoning, emotions, forgiveness and anticipation.

The discussion of the legal and ethical implications for socio-technical systems has necessarily been limited, but in fact the interaction of technology with society also has ‘political’, ‘cultural’ and even ‘generational’ dimensions that need to be addressed. For example, it could be argued that conflicts of objective interests cannot necessarily be resolved by collective awareness, collective intelligence and collective action. For example, in addressing climate change, there is a complex interaction between (at least) loss aversion bias (people’s preference for avoiding losses over acquiring gains, especially when the loss is incurred by themselves and the gains are accrued by others) and political ‘framing’.

Similarly, the symbiotic partnership of people and technology might not be as benign as might be imagined. The idea of ‘designing’ institutions for ‘fairer’ societies arguably presents serious ideological, political and even moral (Allen et al. 2006) problems. The issue of Artificial Intelligence dominating society has been a staple premise of science fiction, and attracts the attention of leading scientists from other fields. However, with the ever increasing power of data mining and both predictive and prescriptive analytics, it is arguably the case that we should still be more wary of the programmers than the programs.

In conclusion, it was strictly *not* the objective of this paper just to catalogue the authors’ own contributions, but rather to document the retrospective and present the prospective in a way that would acknowledge the implicit and explicit contributions made by our partners on the ALFEBITE project. However, we would end with three final remarks:

- Firstly, to highlight and emphasis the importance of inter-disciplinary research, and acknowledge the beneficial collaborations: it is likely that none of this research would have been so well-found (perhaps even possible) without the insights and understanding that stemmed from working with experts from the fields of law and legal information systems;
- Secondly, that as current research into Artificial Intelligence becomes increasingly dominated by statistics and machine learning, the need for this inter-disciplinary collaboration with legal theorists is actually greater rather then lessened; and
- Thirdly, it is extremely difficult to evaluate the impact of basic research.

On this last point, the current trend of demanding that basic research projects should specify ‘pathways to impact’, or that research papers should be evaluated according to their ‘impact statements’ within some fixed and arbitrary period, should be

treated with some caution and no little skepticism. The fact remains that the pathway to impact is probably even less predictable or manageable than the research programme itself; and as for the ‘impact statement’, this is so time-dependent as to require continual re-assessment, not just at a fixed point. The nature and content of this paper would have been rather different if we had tried writing it five years ago, for example; and might even look very different if it were to be revisited in five years’ time.

**Acknowledgments** This paper is written in the honor and memory of Professor Jon Bing. Jon was one of the key senior personnel working on the ALFEBIITE project at the University of Oslo, who were one of the original partners in the consortium. Jon in particular seemed able to resolve a tension that underlies a sort of Heisenberg’s Principle for computers and law: if one knows the direction in which a technology is heading, then one could not know what the law would decide; but if one asserted a legal decision, then the direction that the technology (or the use of technology) would take could not be known or predicted. Jon seemed to be able to recognize the entanglements and uniquely diagnose a disentanglement. In addition, Jon was a polymath whose knowledge and skill extended far beyond computers and law, but into the realms of music and science fiction, and oddly (but somehow characteristically) an affection for the pachyderm that was as steadfast and resolute as his belief in the proper and appropriate relationship between computer technology and law. But much more than this, Jon was a highly personable collaborator who kindly and strongly supported our faltering first steps in this research programme, and our interactions with him, especially as a result of a joint workshop that he kindly organised in Oslo, which heavily impacted our thinking and research directions (even if a dinner in Oslo had a correspondingly heavy impact on our bank accounts). His hand-written comments (in trademark purple ink) on a manuscript of the Open Agent Society paper is a much (personally) valued relic from the ALFEBIITE project. We would also like to acknowledge the many very helpful comments of the anonymous reviewers, and to acknowledge the contribution of Loretta Anania, the EU Project Officer responsible for the ALFEBIITE project.

## References

- Alchourrón C, Gärdenfors P, Makinson D (1985) On the logic of theory change: partial meet contraction and revision functions. *J Symb Logic* 50(2):510–530
- Allen C, Wallach W, Smit I (2006) Why machine ethics. *IEEE Intell Syst* 21(4):12–17
- Artikis A (2012) Dynamic specification of open agent systems. *J Logic Comput* 22(6):1301–1334
- Artikis A, Sergot M, Paliouras G (2015) An event calculus for event recognition. *IEEE Trans Knowl Data Eng* 27(4):895–908
- Artikis A, Sergot M (2010) Executable specification of open multi-agent systems. *Logic J IGPL* 18(1):31–65
- Austin J (1962) *How to do things with words*. Oxford University Press, Oxford, UK
- Boella G, Pigozzi G, van der Torre L (2009) Normative framework for normative system change. In: *Proceedings of international conference on autonomous agents and multi-agent systems (AAMAS)*. ACM Press, pp 169–176
- Bragaglia S, Chesani F, Mello P, Sottara D (2012) A rule-based calculus and processing of complex events. In: *Rules on the web: research and applications—6th international symposium, RuleML 2012*, Montpellier, France, pp 151–166
- Brewka G (2001) Dynamic argument systems: a formal model of argumentation processes based on situation calculus. *J Logic Comput* 11(2):257–282
- Broersen J (2009) Issues in designing logical models for norm change. In: Vouros G, Artikis A, Stathis K, Pitt J (eds) *Proceedings of international workshop in organised adaptation on multi-agent systems (OAMAS)*, vol. LNCS 5368. Springer, pp 1–17
- Bryant V (1985) *Metric spaces*. Cambridge University Press, Cambridge, UK
- Carr H, Artikis A, Pitt J (2009) Presage-ms: metric spaces in presage. In: *ESAW*, pp 243–246
- Cavoukian A (2012) Privacy by design [leading edge]. *IEEE Technol Soc Mag* 31(4):18–19

- Chesani F, Mello P, Montali M, Torroni P (2010) A logic-based, reactive calculus of events. *Fundam Inform* 105(1–2):135–161
- Chesani F, Mello P, Montali M, Torroni P (2013) Representing and monitoring social commitments using the event calculus. *Auton Agents Multi-agent Syst* 27(1):85–130
- Daassi M, Favier M (2007) Developing a measure of collective awareness in virtual teams. *Int J Bus Inf Syst* 2(4):413–425
- Diaconescu A, Pitt J (2015) Holonic institutions for multi-scale polycentric self-governance. In: Ghose A, Oren N, Telang P, Thangarajah J (eds) COIN 2014, LNAI 9372. Springer, pp 1–17
- Dommel HP, Garcia-Luna-Aceves JJ (1995) Design issues for floor control protocols. In: Proceedings of symposium on electronic imaging: multimedia and networking, vol 2417. IS&T/SPIE, pp 305–316
- Forgy C (1982) Rete: a fast algorithm for the many pattern/many object pattern match problem. *Artif Intell* 19(3597):17–37
- Frey S, Diaconescu A, Menga D, Demeure IM (2013) A holonic control architecture for a heterogeneous multi-objective smart micro-grid. In: 7th IEEE international conference on self-adaptive and self-organizing systems, SASO 2013, Philadelphia, PA, USA, pp 21–30
- Giunchiglia E, Lee J, Lifschitz V, McCain N, Turner H (2004) Nonmonotonic causal theories. *Artif Intell* 153(1–2):49–104
- Goulev P, Stead L, Mamdani A, Evans C (2004) Computer aided emotional fashion. *Comput Graph* 28(5):657–666
- Governatori G, Padmanabhab V, Rotolo A (2006) Rule-based agents in temporalised defeasible logic. In: Proceedings of Pacific Rim International Conference on Artificial Intelligence (PRICAI), LNCS 4099. Springer, pp 31–40
- Governatori G, Palmirani M, Riveret R, Rotolo A, Sartor G (2005) Norm modifications in defeasible logic. In: Proceedings of conference on legal knowledge and information systems (JURIX). IOS Press, pp 13–22
- Governatori G, Rotolo A (2008) Changing legal systems: abrogation and annulment. Part I: Revision of defeasible theories. In: van der Meyden R, van der Torre L (eds) Proceedings of conference on deontic logic in computer science (DEON), LNCS 5076. Springer, pp 3–18
- Governatori G, Rotolo A (2008) Changing legal systems: Abrogation and annulment. Part II: Temporalised defeasible logic. In: Boella G, Pigozzi G, Singh M, Verhagen H (eds) Proceedings of workshop on normative multiagent systems (NORMAS), pp 112–127
- Governatori G, Rotolo A, Riveret R, Palmirani M, Sartor G (2007) Variants of temporal defeasible logics for modelling norm modifications. In: Proceedings of international conference on artificial intelligence and law (ICAIL). ACM
- Hardin G (1968) The tragedy of the commons. *Science* 162(3859):1243–1248
- Hess C, Ostrom E (2006) Understanding knowledge as a commons. MIT Press, Cambridge, MA
- Hewitt C (1986) Offices are open systems. *ACM Trans Inf Syst* 4(3):271–287
- Jones A (2002) On the concept of trust. *Decis Support Syst* 33(3):225–232
- Jones AJ, Artikis A, Pitt J (2013) The design of intelligent socio-technical systems. *Artif Intell Rev* 39(1):5–20
- Jones AJ, Kimbrough SO (2008) The normative aspect of signalling and the distinction between performative and constative. *J Appl Logic* 6(2):218–228
- Jones A, Sergot M (1996) A formal characterisation of institutionalised power. *J IGPL* 4(3):427–443
- Kellogg W, Erickson T (2002) Social translucence, collective awareness, and the emergence of place. A position paper for the role of place in shaping virtual community. In: Proceeding CSCW
- Kowalski R, Sergot M (1986) A logic-based calculus of events. *New Gener Comput* 4:67–95
- Lamport L (1998) The part-time parliament. *ACM Trans Comput Syst* 16(2):133–169
- Lessig L (2006) Code: and other laws of cyberspace, Version 2.0. Basic Books, New York City, NY
- Lomuscio A, Qu H, Raimondi F (2009) MCMAS: a model checker for the verification of multi-agent systems. In: Bouajjani A, Maler O (eds) Computer Aided Verification, Lecture Notes in Computer Science, vol 5643. Springer, pp 682–688
- Loui R (1992) Process and policy: resource-bounded non-demonstrative argument. In: Technical report. Washington University, Department of Computer Science
- Luke S, Cioffi-Revilla C, Panait L, Sullivan K (2004) MASON: a new multi-agent simulation toolkit. In: Proceedings of the 2004 Swarmfest Workshop
- Macbeth S, Busquets D, Pitt JV (2014) System modeling: principled operationalization of social systems using presage2. In: Gianni D, D’Ambrogio A, Tolk A (eds) Modeling and simulation-based systems engineering handbook. CRC Press, Boca Raton, FL, pp 43–66. **(FirstView)**

- Macbeth S, Pitt J (2014) Self-organising management of user-generated data and knowledge. *Knowl Eng Rev* 30(3):237–264
- Makinson D, van der Torre L (2000) Input–output logics. *J Philos Logic* 29:383–408
- Michael K, Michael M, Perakslis C (2014) Be vigilant: there are limits toveillance. In: Pitt J (ed) *The computer after me*, chap 13. IC Press, London, UK
- Michael K, Michael M (2012) Implementing ‘namebers’ using microchip implants: the black box beneath the skin. In: Pitt J (ed) *This pervasive day*, chap 10. IC Press, London, UK
- Neville B, Pitt J (2008) PRESAGE: a programming environment for the simulation of agent societies. In: *Programming multi-agent systems, 6th international workshop, ProMAS 2008, Estoril, Portugal. Revised Invited and Selected Papers*, pp 88–103
- Nikolai C, Madey G (2009) Tools of the trade: a survey of various agent based modeling platforms. *J Artif Soc Soc Simul* 12:2. <http://jasss.soc.surrey.ac.uk/12/2/2.html>
- North MJ, Collier NT, Vos JR (2006) Experiences creating three implementations of the repast agent modeling toolkit. *ACM Trans Model Comp Simul (TOMACS)* 16(1):1–25
- Olson M (1965) *The logic of collective action*. Harvard University Press, Cambridge, MA
- Oreskes N, Conway EM (2010) *Merchants of doubt: how a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. Bloomsbury Press, London, UK
- Ostrom E (1990) *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press, Cambridge, UK
- Ostrom E, Ahn T (2003) *Foundations of social capital*. An Elgar reference collection. Edward Elgar Pub, Cheltenham, UK
- Ostrom E, Hess C (2006) A framework for analyzing the knowledge commons. In: Hess C, Ostrom E (eds) *Understanding knowledge as a commons: from theory to practice*. MIT Press, Cambridge, MA, pp 41–82
- Paschke A, Bichler M (2008) Knowledge representation concepts for automated SLA management. *Decis Support Syst* 46(1):187–205
- Perakslis C, Pitt J, Michael K, Michael M (2015) Pervasive technologies: Principles to consider. *Intern J Med Implant Devices* 5(1):79–93
- Petrucci PE, Busquets D, Pitt JV (2014) Experiments with social capital in multi-agent systems. In: *PRIMA 2014: principles and practice of multi-agent systems*, pp 18–33
- Pitt J, Mamdani A, Charlton P (2001) The open agent society and its enemies: a position statement and research programme. *Telemat Inform* 18(1):67–87
- Pitt J (2005) The open agent society as a platform for the user-friendly information society. *AI Soc* 19(2):123–158
- Pitt J, Ramirez-Cano D, Draief M, Artikis A (2011) Interleaving multi-agent systems and social networks for organized adaptation. *Comput Math Org Theory* 17(4):344–378
- Pitt J (2012) Design contractualism for pervasive/affective computing. *IEEE Technol Soc Mag* 31(4):22–29
- Pitt J, Schaumeier J, Artikis A (2012) Axiomatisation of socio-economic principles for self-organising institutions: concepts, experiments and challenges. *ACM Trans Auton Adapt Syst* 7(4):39:1–39:39
- Pitt J, Busquets D, Macbeth S (2014) Distributive justice for self-organised common-pool resource management. *ACM Trans Auton Adapt Syst* 9(3):14
- Pitt J, Busquets D, Riveret R (2013) Procedural justice and ‘fitness for purpose’ of self-organising electronic institutions. In: *PRIMA 2013: Principles and Practice of Multi-Agent Systems, Lecture Notes in Computer Science*, vol 8291. Springer, Berlin, Heidelberg, pp 260–275
- Pitt J, Busquets D, Riveret R (2013) The pursuit of computational justice in open systems. *AI & SOCIETY* 30(3):359–378
- Pitt J, Diaconescu A (2015) Structure and governance of communities for the digital society. In: *Workshop on Self-Improving System Integration*
- Pitt J, Mamdani A (1999) Some remarks on the semantics of FIPA’s agent communication language. *Auton Agents Multi-Agent Syst* 2(4):333–356
- Popper K (2002) *The open society and its enemies* (New Edition; first published 1945). Routledge
- Rawls J (1971) *A theory of justice*. Harvard University Press, Harvard, MA
- RCEP (2010) 28th report: adapting institutions to climate change. Royal Commission on Environmental Protection, The Stationery Office Limited, UK
- Reich R (2015) The sharing economy is hurtling us backwards. Salon. [http://www.salon.com/2015/02/04/robert\\_reich\\_the\\_sharing\\_economy\\_is\\_hurling\\_us\\_backwards\\_partner/](http://www.salon.com/2015/02/04/robert_reich_the_sharing_economy_is_hurling_us_backwards_partner/)
- Rescher N (1966) *Distributive justice*. Bobbs-Merrill, Indianapolis, IN

- Reynolds C, Picard R (2004) Affective sensors, privacy, and ethical contracts. In: Proceedings CHI 2004 extended abstracts on Human factors in computing systems, pp 1103–1106
- Robert SC, Robert III HM, Evans WJ, Honemann DH, Thomas JB (2000) Robert's rules of order, newly revised, 10th edn. Perseus Publishing, Cambridge, MA
- Searle J (1969) Speech acts. Cambridge University Press, Cambridge, UK
- Serban C, Minsky N (2009) In vivo evolution of policies that govern a distributed system. In: International symposium on policies for distributed systems and networks. IEEE, pp 134–141
- Serbedzija N (2012) Reflective computing—naturally artificial. In: Pitt J (ed) This pervasive day, chap 5. IC Press, London, UK
- Sergot M (2008) Action and agency in norm-governed multi-agent systems. In: Artikis A, O'Hare G, Stathis K, Vouros G (eds) Engineering Societies in the Agents World VIII, Lecture Notes in Computer Science, vol 4995. Springer, pp 1–54
- Sestini F (2012) Collective awareness platforms: engines for sustainability and ethics. IEEE Technol Soc Mag 31(4):54–62
- Shum SB, Aberer K, Schmidt A, Bishop S, Lukowicz P, Anderson S, Charalabidis Y, Domingue J, de Freitas S, Dunwell I, Edmonds B, Grey F, Haklay M, Jelasity M, Karpistenko A, Kohlhammer J, Lewis J, Pitt J, Sumner R, Helbing D (2012) Towards a global participatory platform: democratising open data, complexity science and collective intelligence. Eur Phys J Special Top 214:109–152
- Sirbu M (1997) Credits and debits on the internet. IEEE Spectr 34(2):23–29
- Smith R (1980) The contract net protocol: high-level communication and control in a distributed problem solver. IEEE Trans Comput C-29(12):1104–1113
- Vasalou A, Hopfensitz A, Pitt J (2008) In praise of forgiveness: ways for repairing trust breakdowns in one-off online interactions. Int J Hum Comput Stud 66(6):466–480
- Vreeswijk G (2000) Representation of formal dispute with a standing order. Artif Intell Law 8(2/3):205–231
- Wilensky U (2011) CCL at Northwestern University. Netlogo: a cross-platform multi-agent programmable modeling environment. <http://ccl.northwestern.edu/netlogo/>