The Knowledge Engineering Review, Vol. 22:1, 3–35. © 2007, Cambridge University Press
doi:10.1017/S0269888907001014 Printed in the United Kingdom

On agent technology for e-commerce: trust, security and legal issues

MARIA FASLI

University of Essex, Department of Computer Science, Wivenhoe Park, Colchester CO4 3SQ, UK;
e-mail: mfasli@essex.ac.uk

Abstract

The vision of future electronic marketplaces (e-markets) is that of markets being populated by autonomous intelligent entities—software, trading, e-agents—representing their users or owners and conducting business on their behalf. For this vision to materialize, one fundamental issue that needs to be addressed is that of trust. First, users need to be able to trust that the agents will do what they say they do. Second, they need to be confident that their privacy is protected and that the security risks involved in entrusting agents to perform transactions on their behalf are minimized. Finally, users need to be assured that any legal issues relating to agents trading electronically are fully covered as they are in traditional trading practices. In this paper we consider the barriers for the adoption of agent technology in electronic commerce (e-commerce) which pertain to trust, security and legal issues. We discuss the perceived risks of the use of agents in e-commerce and the fundamental issue of trust in this context. Issues regarding security, and how some of these can be addressed through the use of cryptography, are described. The impact of the use of agent technology on the users' privacy and how it can be both protected as well as hindered by it is also examined. Finally, we discuss the legal issues that arise in agent-mediated e-commerce and discuss the idea of attributing to software agents the status of legal persons or e-persons and the various implications.

1 Introduction

The vision of future electronic marketplaces (e-markets) is that of being populated by autonomous intelligent entities—software, trading, e-agents— representing their users or owners and conducting business on their behalf. E-markets are fully automated from end-to-end: finding potential business partners, negotiating terms and conditions of transactions and contracts, signing and monitoring contracts, payment, and even dispute resolution will all be done automatically by software agents. In theory, the only time humans become involved is in specifying their portfolios or preferences.

For this vision to materialize, one fundamental issue that needs to be addressed is that of trust. Trust becomes critical, in particular, if an agent's actions can cause its user physical, financial or even psychological harm (Bickmore & Cassell, 2001). The prospect of disclosing personal and financial information to an agent and delegating to it the task of conducting business on one's behalf does involve a number of risks. Hence, trust is essential to the uptake and deployment of agent technology.

Concomitant to trust is security, in fact trust can derive from security. Central to electronic commerce (e-commerce) is the ability to perform transactions over the Internet which involve the exchange of private and sensitive information. There are inherent risks in the transmission

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

of information over the Internet which need to be managed, and guarantees need to be provided that transactions and exchanges are secure.

Moreover, agents engage in a number of activities which are significant from a legal perspective: they access computer systems, networks and data, they retrieve and distribute information, they mediate personal and business relations, they negotiate for and buy and sell goods and services. Autonomy implies that agents act without the constant and direct intervention of their users or owners. The introduction of electronic agents that conduct business on behalf of natural or legal persons inevitably raises issues with regards to accountability and liability. The traditional paradigms have all developed along the lines of physical and geographical bounds of space and time. The Internet society, however, is not confined by geographical, legal or corporate boundaries. Policy makers need to consider the repercussions of this new technology and provide safeguards and appropriate legislation that covers their operation.

Despite the excitement about agents and their immense potential, there are serious concerns about the associated trust, security and legal issues.We cannot simply assume that agents will be well designed and implemented, behave correctly and therefore will provide for security and protect the users' privacy; explicit assurances and guarantees must be provided to this effect. For agent technology to fulfil its true potential and for users to be willing to engage with and delegate tasks to agents, a number of challenges need to be overcome. In particular, users need to:

(1) trust that agents do what they say they do;
(2) be confident that their privacy is protected;
(3) be confident that the security risks involved in entrusting agents to perform transactions on their behalf are minimized;
(4) be assured that any legal issues relating to agents trading electronically are fully covered as they are in traditional trading practices.

In this paper we consider the barriers for the adoption of agent technology in e-commerce which pertain to trust, security and legal issues. The paper is organized as follows. In the following section the applications of agent technology in e-commerce are briefly discussed. The perceived risks of the use of agents in e-commerce are presented next. The subsequent sections discuss the complex concept of trust in particular in the context of emarkets. The following section discusses electronic institutions as a means of regulating the agents' behaviour in e-markets. Section 7 discusses another possible solution to trust management as provided by reputation systems. The issue of security is considered next as well as how cryptography can be used to enhance security and trust in agent-mediated e-commerce. The impact of the use of agent technology on the users' privacy and how it can be both protected as well as hindered by it is examined next. The last part of the paper discusses legal issues as they arise as a result of the use of software agents to mediate and negotiate transactions and contracts as well as carry them out. Finally, the issue of attributing legal status to agents is considered. The paper closes with the conclusions.

2 Agents and e-commerce

Despite its name, e-commerce is not fully automated. Typically, online transactions require a significant level of human intervention. Humans undertake the search for products, services, potential vendors and business partners. They evaluate alternatives, decide what goods to buy and when, from which vendor and how much they are willing to pay for them. They engage in potential negotiations, carry out the transactions, and so on. In essence, all the decision-making is done by humans. However, in principle, there is no reason why some aspects of commerce, if not all in some cases, cannot be automated. The concept of an agent as a computational system that can act on a user's behalf seems particularly suitable for e-commerce applications. Agents are different from other software in that they actively seek to satisfy their users' goals, they react to changes in the environment that affect the satisfaction of their goals in a timely fashion, and they have

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

social capabilities, i.e. they are able to interact with each other in a number of ways (communicate, cooperate, negotiate; Wooldridge & Jennings (1995)). In addition, agents can be endowed with the ability to learn and adapt to their environment. It is natural then to consider delegating some of the decisionmaking to such agents.

Indeed, the potential for agents in e-commerce is immense; some researchers and technology analysts consider e-commerce as the potential killer application for agents (Ulfelder, 2000). Agent technology has the potential to revolutionize how individuals as well as businesses and organizations conduct commerce (Maes *et al.*, 1999; He *et al.*, 2003). In particular, agents can:

- find, recommend and compare products, vendors or services;
- participate in e-markets and negotiate the price/terms of a transaction or contract with other participants;
- perform transactions on behalf of their users;
- track the user's interests and offer personalized services;
- monitor conditions and provide notifications;
- retrieve, filter and mine information and knowledge;
- produce and deliver e-services such as information gathering, processing and management.

Although the choices and opportunities for individuals and organizations have increased dramatically due to e-commerce, unless they can actually locate these opportunities—clients, markets, potential partners, openings in the market—they will be unable to take advantage of them. Partially or fully automating some of the processes involved in e-commerce will bring about significant cost savings. Using agents an organization can truly do business 24 hours a day and take advantage of opportunities as they arise globally even if they present themselves in different time zones or continents. As a software agent can search tirelessly for the best matches globally, the costs of searching for partners or products can be minimized and efficiency can thus increase. The searches that these agents can perform on behalf of the user can be intelligent searches involving visiting a number of sites in order to extract the required information and match the user's preferences and find the best possible deal. Such agents can undertake one of the most time-consuming tasks, which is that of negotiating the terms of the transactions and contracts. Agents can make good decisions fast, again minimizing costs and improving overall economic efficiency without necessarily involving humans. Agent technology can also be used to offer personalized services to users.

There has already been some headway towards the direction of automating e-commerce through the use of agent technology. For instance, a number of agents have been developed for finding and comparing products, vendors and services: BargainFinder (Krulwich, 1996), PersonaLogic (Guttman *et al.*, 1999), IntelliShopper (Menczer *et al.*, 2002), Shopbot (Doorenbos *et al.*, 1997)[1]. Agents can also recommend products, vendors and services: MovieLens[2], Amazon[3] (Linden *et al.*, 2003). Agent technology can also be used in complex negotiation situations as has been demonstrated through events such as the Trading Agent Competition[4] (Wellman *et al.*, 2003; Arunachalam & Sadeh, 2004) which features automated trading agents competing in dynamic, complex and stochastic environments where uncertainty is inherent, such as for instance supply chain management.

As the technology matures, more and more complex tasks will be delegated to software agents who will become an indispensable part of an increasingly open, free-market information economy.

1 Although analysts had anticipated that shopbots were going to have a huge impact on vendors and the way business is conducted, due to a number of technical and business-imposed limitations, the technology has not lived up to expectations (Fasli, 2006). Currently comparison sites such as AddALL (`http://www.addall.com/`), DealTime (`http://www.dealtime.com/`), MySimon (`http://www.mysimon.com/`) are not strictly speaking agents. Such sites do not offer impartial advice, but they collate catalogues of products which are provided by the vendors themselves who pay to be listed on the site.

2 `http://movielens.umn.edu/`.

3 `http://www.amazon.com`.

4 `http://www.sics.se/tac/`.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

Agent technology can reduce the costs of trading and thus increase market efficiency and profitability, trading volumes as well as the speed of trading. Agents can enable the move from traditional brick and mortar companies to intelligent and ubiquitous digital business. The vision is that agents will evolve from being simple facilitators to complex and autonomous decision-makers handling incomplete, inconsistent information in real time, and making complex, but good, decisions. For this vision to materialize, we need knowledge-based agents that can reason under uncertainty and with incomplete information, goal-driven agents that can plan, and agents that can learn and adapt to their environment.

Agents play an instrumental role in the realization of a borderless digital 24-hour information economy, but users need to trust them first.

3 Perceived risks

As has often been remarked, there is a close connection between trust and risk (Luhmann, 1979, 1990). Put simply, if there is no risk, the question of trust does not arise. The absence of risk regarding future outcomes implies confidence and certainty. In the presence of risk, trust is required to overcome the inherent uncertainty involved.

Deploying agent technology in e-commerce is certainly not without risk (Brazier *et al.*, 2004b). Agents run on an agent platform, they communicate and interact with other agents, objects and Web services on the Internet and they function autonomously without the user's constant control and intervention. Undoubtedly, mobility increases the risks an agent faces as agents may be hosted by a potentially malicious agent platform, outside the control of their user, and thus they are even more vulnerable (Tschudin, 1999) especially in the context of e-commerce (Claessens *et al.*, 2003).

Most of the time users are unaware of which other agents, objects or Web services their agents are involved and interact with. In an open environment, an agent runs the risk that agent platforms, other agents, humans, viruses, etc., can access, copy or even modify its code or data either by design or by mistake. These attacks can take place while an agent operates, lies in storage or is being transmitted. The risks for the users are considerable and the potential consequences indeed serious. An unauthorized modification of an agent's code or data may influence the agent's behaviour and result, for example, in buying goods for inflated prices, making faulty flight reservations or engaging in malicious behaviour. As an example, while interacting in a marketplace, a buyer agent's reservation price may be modified, thus increasing the likelihood of the agent purchasing goods at higher prices than originally intended. Even accessing data, especially confidential and sensitive information such as bank accounts, credit card numbers and reservation prices, may be sufficient to cause significant damage. For instance, if the reservation price is simply accessed without permission by a seller agent, this gives the latter an advantage during negotiation, as knowing this information, it can adjust its counteroffer in such a way as to achieve maximum profit. Similarly, a seller's reservation price may be increased by another seller so as to steal its potential share of the market, or decreased by a buyer to obtain a better deal.

Significant risks emanate from stealing information, including private information such as addresses, bank accounts, credit card numbers, and an agent's goals and instructions. Such information can be stolen while in transit or in storage. Information on bank accounts and credit card numbers can be misappropriated: money can be transferred out of the user's account or used to purchase goods and services. The agent's identity, and as a consequence that of the user, as well as their authority, can be hijacked by a malicious agent who can then use it and cause damage. Agents or users may be banned from a marketplace if they are found to have acted maliciously, even though a user may not be directly responsible for the agent's behaviour—the misbehaviour may have been the result of the agent's code being tampered with by another entity. Depending on the severity of the damages caused by a misbehaving agent, the user or owner may

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

even have to face litigation. The issue is whether such policies constitute viable business practices and are perceived as fair by users. Apart from the potential financial losses, a user risks their reputation being damaged irreparably. Reputation is in fact an asset which both individuals and companies value as any other asset and has a replacement cost. Reputation is difficult to build, but very easy to tear down.

Risks from non-repudiation are also considerable. An agent or host can deny that a communication exchange or transaction has taken place, thus leaving another agent exposed. For instance, an agent responsible for procurement in a supply chain may order components essential for manufacturing a particular product. However, a supplier may deny ever receiving an order or agreeing on a contract. If such a situation arises, even the mere delay involved in investigating and resolving the dispute may mean significant losses. The buyer agent may not be able to receive the necessary components for manufacturing on time, and as a result it may not be able to satisfy a customer order with subsequent monetary losses as well as damages in its reputation. Logging and auditing of communications and transactions need to be systematic in order to avoid such problems, and be able to trace what has happened in case of possible dispute or litigation.

Another potential risk is that an agent may disappear temporarily or permanently, representing a possible loss of investment in training and acquisition to the user/owner. More significantly, valuable results, trading data and records of transactions can be lost altogether along with the agent. Depending on the situation, tracing and reconstructing transactions may be extremely difficult, if not impossible, as an agent may have interacted with numerous other entities.

Electronic markets are real-time, dynamic and non-deterministic environments. Exceptional situations that are usually characterized by higher risk cannot be ruled out. Users are uncomfortable with the idea of software agents dealing with non-routine or exceptional situations that perhaps the latter are facing for the first time and which they may not necessarily have been designed for or learnt how to deal with. It is a widely held belief that automation is not in a position to deal with exceptional situations (Norman, 1990).

Using agent platforms is also not without risk (Tschudin, 1999). An agent platform hosts an agent and provides services to it. The validity, reliability and trustworthiness of an agent's code and data cannot be easily determined automatically. Agents compete for resources, such as processor cycles, bandwidth, storage space, and services such as communication and mobility. Malicious agents may succeed in migrating to a trustworthy platform and code or data of the system can be modified or stolen, by accident or intentionally. Information on management or access policies can be altered. The platform may exhibit undesirable behaviour, for example, by allowing unknown agents to enter, assigning to or removing resources from agents, killing agents, or allowing third parties to have access to the agents' code and private information, or may simply crash. The platform could also be inundated by denial of service attacks. If an agent platform ceases to function, this means possible financial losses to the agent platform owner as well as the owners and users of the other agents operating on the platform. The reputation of the agent platform and its owners can be irreparably damaged, and, depending on the severity of the misbehaviour, the owner may have to face litigation.

4 Trust

Trust is a fundamental concept in our everyday lives and to a great extent guides our decision-making process.Without trust, virtually all of our social relationships would break down, and it would be impossible to function normally in our everyday lives. If we cannot trust other drivers to stop at a red light, it would become impossible to drive. We depend on others every day, but we are usually comfortable in doing so because we trust that their actions will not be detrimental to us (Rotter, 1980).

Trust can be considered as the generalized expectation that an individual, a group or an artefact (i.e. technology) can be relied upon. Trust is associated with risk and could be viewed

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

as confidence in the face of risk. There are two parties in a trusting relation: the principal who trusts, and the trustee who is the one who is entrusted. In trusting someone, one depends on this person and accepts vulnerability based upon positive expectations of the intentions or behaviour of this person which is not under one's control. The need to trust stems from this vulnerability and dependency on others and the need to delegate tasks to others. Delegation, in particular, presupposes trust, as it inevitably means loss of control. The process of delegation allows for passing responsibility for a task to another person and equipping them with the necessary authority and access to the resources needed to fulfil that task. However, task accountability remains with the principal (Oates, 1993). This means that the principal assumes a risk in that the task may not be completed according to one's standards and one will be personally accountable for the outcome. As a result, it is in the principal's best interest to choose the person who they believe is more capable of accomplishing the task with great care.

Trust can be distinguished into personal and impersonal. Personal trust is subjective and is formed by an individual based on beliefs, observations, reasoning, social stereotypes, communication, past experiences, etc. (Marsh, 1994; Falcone & Castelfranchi, 2004). The interaction between principal and trustee constitutes the basis for the development of trust between the two parties. Thus, trust is built through updating *a priori* expectations about another party by interacting with it. Trust develops following a positive experience from the interaction, even when the initial expectations about the other party's behaviour were negative, and, conversely, trust is reduced following a negative experience. People have different propensities or dispositions towards trust, in other words different baseline levels of trust; some people are more willing to trust than others. This baseline trust depends on an individual's personality and affects decision-making, as well as how new experiences are taken into account in updating one's trust (Marsh, 1994).

Nevertheless, developing personal trust is not always possible due to situational constraints (Shapiro, 1987). In the absence of personal trust, impersonal trust mechanisms can act as a substitute. Intuitively, impersonal trust is trust which can be derived from information or experiences as reported by others. There are a number of trust-inducing mechanisms that fall under this category such as trusted third parties, rule-based trust and word of mouth or reputation mechanisms. In the absence of any personal experience regarding the trustworthiness of someone, third-party intermediaries can be used to mediate between the principal and the trustee. Thus, the principal may not interact directly with the trustee, but through a third party which is mutually trusted by both of them.Alternatively, a principal can adopt the trust of a trusted third party as one's own. Trusted third parties are for instance banks, consumer institutes and governments. Another form of impersonal trust is rule-based trust. Trust can be induced by the knowledge that there are certain performance criteria and sanctions for non-conformance in place, which all members agree to abide by when they first join the marketplace, i.e. institutions (Dellarocas & Klein, 1999). Finally, another form of impersonal trust is word of mouth or reputation. In the absence of personal experience with a person, a principal may seek information from third parties regarding past interactions. The outcomes of these past experiences are aggregated and make up reputation (Resnick *et al.*, 2000). Relying on word of mouth or someone's reputation is in essence reliance on the aggregated experiences of third parties, which are then used to inform one's decision-making. Such third parties can be a community, one's circle of friends, colleagues or the society. This sharing of past interaction histories can take many forms: informal gossip networks or institutionalized review or reputation systems (Kollock, 1999).

5 Trust in e-commerce

Typically, trading in markets involves trusting that the other participants are going to comply with the rules of the protocol and carry out their part of the transaction or honour the terms of the contract agreed, as well as the marketplace as an institution itself. In delegating such tasks

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

to software agents, the need to trust the agents themselves also arises. Thus, trust in emarkets populated by software agents who act on behalf of their users/owners takes two forms:

(1) trust in the software agent as one's representative;
(2) trust in the marketplace infrastructure (that it is fair, non-manipulable, and rules and safeguards are in place) and the other participants (that they will act according to the rules of the protocol, honour contracts and agreements).

5.1 Trust in agent technology

Without doubt a trusting relationship must develop between the user and the agent. There are many valid reasons why users might hesitate to trust personal software agents in particular in e-markets. Generally speaking, there is always reluctance to trust and adopt new technology, and especially technology that one does not understand or comprehend the workings of, or technology that has not been tested extensively. In particular, trust becomes very important if the use of technology is perceived to carry high risks, i.e. can cause its user physical, financial or psychological harm (Bickmore & Cassell, 2001).

In fact, users need to place enormous trust on the agent that represents them and acts on their behalf in a marketplace so they must be confident that agents will do what they ask, and only what they ask. Users also have to entrust agents with private and sensitive information, and agents will have to handle that information in a secure way (Patrick, 2002). Such a software agent needs to have built-in mechanisms to ensure reliability, transparency and make sure than no one can tamper with or manipulate it while it is operating in the market.

The distrust in agents and unwillingness to delegate tasks to them may be attributed to the fact that software agents are perceived as 'faceless strangers' (Shapiro, 1987). Users may often limit their relationships and entrust parties with whom they have previous experience with, have an ongoing relation, or whose performance and reliability has already been established. Moreover, it is often the case that users—individuals, organizations—consider familiarizing themselves and developing relations with their customer base and suppliers as being important in order to improve the level of service provided/received (Child & Linketscher, 2001). However, none of these seems to be applicable to agents.

Another contributing factor to the users' tendency to distrust agents is that of the nature of the tasks being delegated to them. Such tasks are 'twice removed from the interface' (Patrick, 2002). Consider a buyer who is interested in purchasing a particular good. In a traditional shopping trip, the buyer interacts with vendors directly by visiting their shops. A shopping trip nowadays may take place through a computer where the buyer interacts with a Web browser to view information provided by the vendors. The buyer, having information from a number of vendors, is then in a position to compare prices and subsequently make a decision. Hence, the interaction between the buyer and the potential vendor is once-removed from the interface or a disembedded transaction (Riegelsberger & Sasse, 2001). When the shopping task is delegated to a software agent, the buyer interacts with it in order to provide instructions and preferences through some sort of interface. The agent would then search the Internet, gather information provided by one or more vendors and perform comparisons according to the buyer's criteria. As a result, there is no direct connection between the buyer and the search or comparison activities. The interaction between the buyer and the potential vendor is twice-removed (or dis-disembedded). Developing trust can be difficult during once-removed interactions, but even more so in twice-removed interactions. At the same time, as agents facilitate twiceremoved interactions between users and other parties, they are ideal for tasks that require high degrees of privacy or anonymity (see the discussion below).

Finally, the user must be confident that the agent has the mechanisms itself to cope in an uncertain environment where it will have to make decisions on which other agent or service provider to trust in looking after its user's interests. Thus, the user must be sure that the agent

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

will behave in a rational way when faced with new and non-preprogrammed situations. With all of these concerns, developing a trusting relationship between users and their agents is non-trivial.

Allowing agents to act autonomously may be too much to ask from the user right from the start. The user must come to develop faith in the agent's sense of obligation, that is, that it can be trusted to act in the user's best interests to achieve optimal performance (Muir, 1987; Lee & Moray, 1992). In the beginning of the interaction, the agent can perhaps have more of an advisory role, and, as trust builds up, this leads to a more active role in the transactions being carried out, ultimately allowing it to act autonomously in the user's interests. As trust is built gradually through positive experiences, one way forward may be to allow greater degrees of control to users. Control and trust seem to be two opposite sides of the same coin. In essence, a user needs to have high levels of control when they have low levels of trust, whereas when trust is high the level of control can be low (Child & Linketscher, 2001). In the early stages of interaction with a principal, the agent needs to have more built-in mechanisms of control to compensate for the lack of trust. These relate to input, choice and ultimately decision control.

An agent's behaviour can be controlled in three stages: (i) pre-activity, (ii) real time and (iii) post-activity (Jenks & Kelly, 1985). In the first stage, the principal sets and communicates to the agent the criteria for evaluating performance as well as any specific instructions for accomplishing tasks. Obviously a user should give clear instructions to the agent on the task to be performed and the various parameters such as budget and other preferences. However, this instructing phase could potentially fail if (Youll, 2001):

(1) the user does not convey their intentions clearly;
(2) the agent does not fully comprehend the instructions;
(3) the user and the agent interpret the instructions provided differently.

If the agent understands the instructions, the user must be confident that it will execute the instructions correctly, and will only perform the tasks the user intended.

In the second stage, in real time, the principal controls the agent by monitoring its performance. The principal may also choose to intervene if necessary to guide the agent, or correct or even override the agent's decisions. To be able to evaluate that an agent's behaviour is consistent, the user needs to have a clear picture of what the agent does. Updating the user regularly through providing periodic progress reports can facilitate observability and a better understanding of what goes on in the 'black box'. Providing an explanation or justification on decisions and actions reduces the user's uncertainty about what happens inside the agent and provides system process and benchmarking information. Provided that the agent's behaviour is consistent and stable over time, the principal is likely to deem the agent as dependable and eventually as trustworthy of future delegation.

On task completion, the principal reviews the agent's performance and task achievement and may reward good performance or mark and even punish deviation from the user's goals.

As the level of trust increases, the user may relinquish some of the control, thus the agent needs to be able to adapt to the user's ever-changing needs. Hence, in the absence of initial trust, high levels of control and transparency of operations can aid towards the gradual building of it.

5.2 Trust in the marketplace

In traditional as well as e-markets, exchanges are often carried out among participants that have never met before and may have little or no idea about each other's reliability and trustworthiness. Trust in the marketplace is an essential ingredient if participants are to keep negotiating with each other in order to reach a deal and for transactions to go through. This trust reflects the risk that a person undertakes towards the protocol, the marketplace as well as the other participants. One needs to trust the protocol and the market infrastructure that it is fair and non-manipulable.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

E-markets provide a fertile ground for deceitful participants to engage in old as well as new types of fraud. For instance, items may not be delivered at all or may be of different quality than that agreed, the prices charged are higher than those appearing on electronic catalogues or originally agreed, refunds are not given despite being advertised differently, buyers may delay payments, or not pay at all, or post cheques that are not valid.

To minimize risk in e-markets, the issues of trust management and security need to be addressed. In particular, e-markets must address how they intend to provide trust, security, enforce contracts and establish a legal framework. The marketplace itself may provide safeguards and guarantees against fraud and breaches of the protocol, as well as impose sanctions on those who have deviated from the prescribed rules. This can be facilitated through the use of institutions.

In addition, apart from making sure that transactions are secure, there is the need to equip software agents with mechanisms that allow them to reason about and assess trust towards other market participants much in the same way their users would do in traditional market settings. Consequently, we need built-in mechanisms for evaluating the outcomes of interactions with other agents as positive or negative experiences and building a notion of trust through them (Michalakopoulos & Fasli, 2005).

6 Electronic institutions

Designing efficient and robust open e-markets in which heterogeneous selfinterested agents engage in interactions attempting to further their own objectives is an open problem.Traditional mechanism design is concerned with the development of protocols that are stable and individually rational and which lead to socially desirable outcomes. Some of the underlying assumptions of mechanism design are not realistic in the context of heterogeneous software agents in complex environments with limited computational resources and reasoning power (Dash *et al.*, 2003). One of the main assumptions is that if all agents follow the right mechanism, then desirable, efficient and stable outcomes ensue. Moreover, agents are rational and therefore can compute their complete preferences with respect to all possible outcomes. In addition, the society of agents participating in the mechanism is closed, i.e. the agent population is static. These assumptions are problematic, in particular, in the context of e-commerce being conducted by software agents. First, agents do not have unbounded memory and computational power. Second, e-markets are open and dynamic environments in which agents may appear or disappear for a number of reasons. In such open systems agents are heterogeneous and built by different developers with a different understanding of how protocols work, and of course their own agenda. Guiding and regulating the behaviour of the agents in such systems to discourage malicious behaviour, harmful interactions and enhance trust is a difficult problem. Given the inherent heterogeneity, agents cannot be expected to have the same level of sophistication or necessarily follow the rules prescribed by the interaction protocols. Moreover, their interactions may lead to unexpected, dysfunctional behaviour and inadvertently trigger systemic failures (Dellarocas & Klein, 1999).

The counterpart of artificial societies and e-markets, human societies and traditional markets, have dealt with the above issues through developing norms which guide, monitor and regulate the agents' behaviour (Dellarocas & Klein, 1999; Dignum, 2001, 2002; Esteva *et al.*, 2001; Sierra & Noriega, 2002). Norms or social constraints are the essence of institutions (North, 1990). In a similar way, the behaviour of agents within e-markets, virtual organizations or enterprises and in general within multi-agent systems can be regulated via appropriate electronic social institutions. Electronic institutions identify a set of norms that the agents should adhere to when they participate in a particular institution. To this end, interaction protocols (auctions, ContractNet, SNP, etc.) can be augmented with norms to improve their stability, robustness and performance in the presence of heterogeneity, potential malicious behaviour and systemic failures.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

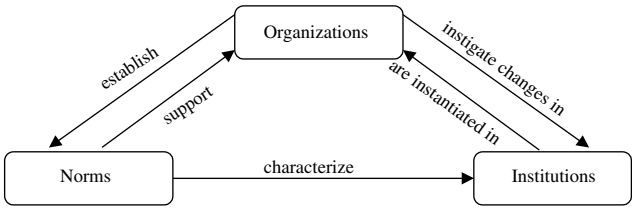


Figure 1 The relationships between norms, institutions and organizations

6.1 Norms, institutions and organizations

Norms are rules, constraints or standards of behaviour which are shared by groups of agents or societies. In general, norms encapsulate abstract values or principles such as 'fairness', 'justice' or 'equality', which may be encoded as specific rules or policies, or conventions. In this respect, norms simplify the interactions among agents and enhance trust. Institutions are amalgamations of interdependent conventions or norms.

Institutions can be formal or informal. The former consist of norms which are encoded as explicit rules or regulations and are enforced within the group or society and prescribe sanctions for deviating from these norms. For instance, a formal norm which is usually explicitly encoded as a statute is that stealing is not permitted behaviour within a society.Any agent deviating from this norm will have to suffer the consequences in the form of sanctions (i.e. prison sentence or fine). Informal institutions are characterized by norms which are not enforced by regulations, but still guide the agents' behaviour, such as taboos, customs, conventions or codes of behaviour. For instance, giving up one's seat on a bus to an elder person is an informal norm, but an agent will not be punished for violating this.

Thus, institutions determine what individual agents are allowed and prohibited from doing, they ascertain deviations from the prescribed behaviour and violation of rules and determine the sanctions to be imposed on the deviating agents. North (1990) studied institutions and their effect on human organizations and postulated that the fundamental role of institutions is to provide a stable framework for interaction, guide the behaviour of human agents and reduce uncertainty.Within an institutional framework, the interaction costs can be potentially reduced while the interactions can grow in complexity and their outcomes can largely be anticipated as participants adhere to the prescribed norms.

At this point we need to clarify the relation between institutions and organizations. Organizations are groups of agents united by some common purpose to achieve a set of objectives. They constitute social entities with structure, resources and authority. They are created with the purpose of taking advantage of opportunities that arise, for example, in markets or societies within the existing institutional frameworks. Organizations, such as for instance corporate organizations and other types of (formal and informal) groups, become instances of specific institutions or a combination of a set of norms (Vazquez-Salceda *et al.*, 2003), i.e. institutions manifest themselves in organizations. Organizations require norms to support their objectives, enhance their functionality and instill confidence and thus, inherently, are governed by institutions. The relationships between norms, institutions and organizations are illustrated in Figure 1. For example, the English auction is a trade institution that prescribes how negotiations can be conducted. A particular auction house X can be an instantiation of the English auction and can conduct auctions using the specific rules as described by the protocol. The auction house may include additional norms that further restrict the behaviour of the agents and which may not typically be part of the English auction protocol[5]. For instance, bidders that wish to participate in an auction may be required to pay a registration fee. The auction house X is an organization—it has structure,

5 McAfee and McMillan (1987) refer to auctions as institutions that include normative aspects.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

members and resources—which instantiates a trade institution and may also impose additional norms.

To summarize, the need for norms stems from the fact that uncertainty is endemic in environments with heterogeneous agents and their interactions need to be structured to promote stability, fairness and trust, and to avoid harmful interactions. Norms characterize institutions. Organizations can be created as instances of specific institutions or a combination of norms. Norms support organizations, and thus organizations establish norms to improve their functionality and interactions. As organizations are not static entities, their norms evolve according to the needs of the constituent agents and the environment in which they operate, and, as a consequence, this may also instigate changes in the institutions. In the following, the terms institution and organization will be used interchangeably where no confusion can arise.

6.2 From norms to institutions

Norms that govern electronic institutions can be distinguished into three categories (Sierra & Noriega, 2002): ontological and communication norms, social interaction norms and norms that impose restrictions on the behaviour of individual agents.

Ontological and communication norms intuitively enable the clear and unambiguous communication among the various participants. Such norms make explicit various aspects that affect communication such as, for instance, the communication language to be used, the performatives that can be exchanged and their meaning, the content language, any underlying ontologies, as well as the roles of the participants themselves and other conventions for exchanging messages. Ontological and communication norms are essential in a system where agents will almost certainly be heterogeneous and they need to be able to communicate with each other in clear unambiguous terms.

Social interaction norms essentially dictate the protocols to be used within the institution and which enable the participants to interact with each other. For instance, in an e-market the Dutch auction is such an interaction protocol which describes how participants can engage in negotiation via this particular protocol. Such norms may also describe the correct sequence of activities within an institution; for example, first an agent needs to register with the auction house, then it can bid in various auctions, payment follows if the agent wins an auction and finally an agent can exit the auction house after successful payment has been completed. These distinct but, in some cases, interdependent activities are usually referred to as scenes (Sierra & Noriega, 2002).

The third class of norms, behaviour norms, are normative rules that dictate permissible and acceptable behaviours as well as prohibitions within the institution, i.e. such norms describe an agent's obligations and rights.

The aim of imposing norms is that they act as deterrents, disincentives and fundamentally are preventative measures against unwanted behaviour (Dignum, 2002). The designer of an auction house, for instance, would like to make it impossible or extremely difficult for the agents to violate any of the prescribed norms. Norms are therefore translated into rules that impose restrictions on the agents' behaviour or sanctions on deviant behaviour. These rules form the policy of a particular organization.

Norms that indicate prohibitions can be translated into regulations. For instance, to prohibit unregistered agents from bidding in auctions, the auction house may require that agents register first. Norms that indicate desirable behaviour are whenever possible translated into restrictions on unwanted behaviour. The issue here is that the institution cannot really enforce such behaviour. For instance, the desired behaviour that we would like the winner of an auction to exhibit, i.e. to pay for the item won, cannot easily be enforced. Alternatively, other rules can be put in place to act as incentives to adhere to the norm or act as deterrents of unwanted behaviour. For example, an auction house may withhold an agent's deposit or entry fee paid, if the agent refuses to pay for a good won in an auction. Furthermore, the auction house can prevent the agent from participating and bidding in any subsequent auctions. Norms that indicate that certain

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

actions are permitted only under specific conditions can be translated into checking that the conditions are satisfied before the agent is allowed to execute the action. As an example, agents may be permitted to pay by credit card, provided that the credit card is valid and their credit limit is not violated. Thus, unless these two conditions are fulfilled the agent is prevented from performing the action of paying by credit card.

Nevertheless, there are other types of unwanted behaviour that cannot be translated from norms into explicit rules that can easily be enforced. For instance, in auctions we would like to prevent participants from colluding and forming rings. Such a norm cannot be easily enforced, as the agents may decide to act outside the institution, and thus the institution has no way of checking their behaviour. In this case, the mechanism or protocol designer might try to deter such behaviour by designing the rules of the game, i.e. the rules of the interaction protocol (e.g. Dutch auction), in such a way that the agents have little incentive to collude. For a ring to be effective, the agreement within the ring needs to be self-enforcing.

Alternatively, instead of attempting to enforce a norm through restricting the agents' behaviour, it may be more efficient to react to violations of norms (Dignum, 2002). For instance, agents that are caught attempting to manipulate the auction can be banned from the auction house, or have their registration fees or deposits withheld as a form of fine. For each norm that the institution would like to enforce a rule is required that specifies either:

(1) the procedure within the institution that enforces the norm; or
(2) the conditions that constitute violations of the norm and the consequences, or sanctions imposed by the institution.

6.3 Agents in electronic institutions

Agents join e-markets in order to interact with other agents and further their own objectives. Marketplaces can be regarded as electronic institutions which are underpinned by a set of norms. By entering a marketplace, an agent assumes a role and commits to adhere to the institution's norms, namely its policy. Every participating agent in such a marketplace should act, and also expect others to act, according to the institution's policy.

Commitments may take the shape of formal contracts between the institution and the agent. Contracts can be private or social (Dellarocas & Klein, 1999). Private contracts are drawn between two or more agents who want to engage in interaction and execute a transaction. The institution's role in this case is to protect the agents throughout the duration of the contract from any potential exceptions, including exceptions that arise as a result of non-conformance or malicious behaviour on behalf of the other agents. A social contract is a commitment on behalf of an agent towards the institution to adhere to its norms. The institution, however, commits to enforce the agent's private contracts as well as protect the agent from the unsolicited behaviour of any other agents or any systemic failures.

Within the context of electronic institutions, agent designers and developers can develop their agents so that they behave according to norms (Dignum, 2001). Agents need not necessarily be aware of the norms that different institutions impose, as long as they are designed to conform to a wide range of principles and not exhibit malicious behaviour. However, although some generic norms may be hardcoded into the agent, such agents cannot adapt to different institutions. Alternatively, agents can be designed so that they are able to reason about norms and make decisions accordingly. In this case, norms need to be represented explicitly alongside the rules for different interaction protocols. For instance, alongside the rules of various auction protocols, an auction house should also include a list of rules that describe desirable behaviour while operating within the auction house. One of these rules can be, for instance, that only registered agents are allowed to place bids in an auction.

Similarly to traditional institutions which facilitate interactions among human participants and regulate and enforce desirable behaviour, electronic institutions can be created to regulate the

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

behaviour of agents within marketplaces and promote fairness and stability and minimize harmful interactions. Inevitably, institutions restrict the agents' autonomy. Agents that want to participate in electronic institutions of any form will lose some of their autonomy, as they have to adhere to the institution's norms. This is the price that agents have to pay in order to be able to interact with others within the context of an institution that provides safeguards against breaches of protocols. However, agents exercise their autonomy in entering into those institutions and committing to comply with their policies in the first place. Nevertheless, an agent being autonomous may decide to violate the rules if it perceives that it can benefit from this, but risks being ostracized and therefore unable to further its objectives.

7 Reputation systems

Trust is an essential prerequisite for establishing and developing open but stable online trading communities. In seeking out and engaging with new partners and in the absence of past interactions with them that would provide an indication of their trustworthiness, one has to resort to alternative means to trust management. One such alternative is the use of reputation reporting systems to encourage trustworthiness and good behaviour in exchanges.

Reputation encapsulates the distributed knowledge of a set of entities about another and is used to predict future behaviour based on past behaviour (Bromley, 1993). When people interact with one another over time, the history of past interactions informs them about each other's dispositions and traits. An interaction such as an economic transaction can be viewed in very simple terms as a prisoners' dilemma situation. The buyer and the seller are faced with the following dilemma: to cooperate with each other, i.e. the buyer pays what is due to the seller and the seller ships the goods as agreed, or to defect, i.e. the buyer does not pay for the goods, and the seller does not ship them or they are of lower quality than that agreed. The socially acceptable outcome of this interaction, which constitutes the Pareto optimal solution, is for both of them to cooperate. However, each has the incentive to defect in order to maximize its own temporary gain. The reason participants fail to cooperate when they do not know each other is the lack of a shadow of the future (Axelrod, 1984). The lack of a common interaction history and the relative ease with which buyers and sellers can change partners from one transaction to the next gives incentives to both parties to provide inferior quality of service or to deviate from the original agreement. Mechanisms for establishing trust attempt to create this shadow of the future. Of course, trust among unknown parties is much more difficult to build. By keeping track of one's past interactions and making the results available to others, reputation systems establish this shadow of the future by creating an expectation that others will consult the aggregated past history. This fear of retaliation in future interactions, if one has not behaved well, creates a powerful incentive to cooperate. For any party deciding whether to initiate a transaction, reputation is a source of information that can reduce uncertainty and guide the decision of whether to trust or not a potential partner.

7.1 Reputation systems in practice

Generally, a reputation system collects, aggregates and distributes feedback about participants' past behaviour (Kollock, 1999; Resnick *et al.*, 2000). Such systems help interested parties to decide whom to trust, encourage trustworthy behaviour and aim to deter deceitful participants. Reputation systems work in a similar way to recommender systems (Montaner *et al.*, 2003). Each participant provides feedback and rates the interaction experience with another participant once a transaction has been completed successfully or otherwise. As in recommender systems, there are different ways of providing feedback, such as rating in terms of a scale for instance from 1 to 5 and/or providing textual comments. Reputation systems can be used to rate individual traders, vendors, sites, software agents or products and services.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

An indicative example of a successful reputation system being used in a large-scale system is that of the Feedback Forum as deployed by eBay. eBay is the largest person-to-person online auction site, with more than four million auctions active at any time. At first glance, the exchanges on eBay may appear highly problematic: the insurance provided is limited and both buyers and sellers accept the associated risks for transacting with unknown parties, which can be significant. Despite this, however, the overall rate of successful transactions on eBay remains remarkably high for a market that opens the possibility to large-scale fraud and deceit. The high rate of successful transactions is attributed by eBay and its users to its reputation system. After a transaction is complete, the buyer and seller have the opportunity to rate each other using 1, 0 and -1 to indicate a positive, neutral and a negative experience, respectively. Participants have running totals of feedback points attached to their user names, which might be pseudonyms. When one is considering whether or not to enter into a transaction with another participant, the reputation rating provides clear guidance as to the participant's trustworthiness in the past. Among the most useful features of the Feedback Forum is that it allows participants to leave textual comments describing in detail their experiences with other buyers and sellers. With such clear reputation indicators, low-quality or unreliable sellers get lower prices and eventually they become isolated or are banned from transacting, leaving the reliable participants to enjoy a healthier market with prices that correspond to the quality of service that they offer. It is very often the case that sellers with excellent reputations may enjoy a premium for their services, as some users may be willing to pay more for the additional security, reliability and higher quality of service. Other auction sites such as Yahoo! Auction and Amazon feature similar reputation systems with variations on the rating scale (5 or 15), several measures (friendliness, prompt response, quality of service) and different ways of aggregating feedback (averaging instead of total feedback score).

Ratings are of course not the only indicators of reputation. Agreeing to be rated when rating is optional is possibly an indication of higher-quality services even before ratings are available. Other ways to promote one's reputation are to use one's real name rather than a pseudonym, and to disclose other information such as a physical store's address or phone number.

There are a number of prerequisites that must be in place in order for an online reputation system to operate effectively and provide incentives for honest and trustworthy behaviour.

(1) Participants must be long-lived entities, human or artificial, that have the potential for future interaction.
(2) The costs for submitting and distributing feedback must be very low.
(3) Feedback information must be aggregated and presented in a way that enables and guides trusting decisions.
(4) Clear guidelines must exist on how the rating system operates and how potential conflicts can be resolved.
(5) The reputation server itself must be reputable and trustworthy.

7.2 Issues in reputation systems

Although reputation systems offer a viable trust management method that emerges from a community to be used by it, there are significant issues to be addressed with regards to eliciting, aggregating and distributing feedback information (Resnick *et al.*, 2000).

Obtaining feedback from users presents three main problems. First, as in recommender systems, users may be reluctant to provide feedback. For instance, after a transaction has been completed, there is little incentive to spend time filling out a form. Nevertheless, a significant number of users choose to do so, at least on eBay, which perhaps is an indication of the participants' willingness to engage with the community, their gratitude or desire to air their frustration. Although some form of economic incentive can be offered for providing feedback, such a scheme introduces additional problems regarding the feedback's reliability and credibility. Second, it is difficult to ensure honest reporting and that the system is non-manipulable. In particular, this relates to

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

two issues: unfair ratings and discriminatory seller behaviour (Dellarocas, 2000a,b). Unfair ratings can be distinguished as follows.

(1) Unfairly high ratings (ballot stuffing): A seller can collude with a group of buyers in order to obtain unfairly high ratings. As a consequence the seller's reputation will be inflated, thus allowing the seller to have a bigger share of the market and perhaps at higher prices than deserved.
(2) Unfairly low ratings (bad-mouthing): Sellers can collude with buyers in order to bad-mouth other sellers that they want to drive out of the market. Sellers can also masquerade as buyers and provide negative ratings to other sellers. Either of these has as an effect that the reputation of the seller being bad-mouthed becomes lower and this may potentially lead to the seller being banned from the marketplace.

A seller's behaviour can be discriminatory towards different buyers, as follows.

(1) Negative discrimination: A seller may provide good service to everyone except a few specific buyers that it does not like. If the number of buyers being discriminated upon is relatively small, the cumulative reputation of the seller will remain good, but an externality will be created against the discriminated buyers.
(2) Positive discrimination: A seller may provide exceptionally good service to a few individuals and average service to the others. This resembles the effects of ballot stuffing. If the number of buyers being favoured is sufficiently large, their high ratings will inflate the reputation of the seller and will create an externality against the rest of the buyers.

Third, it may be particularly difficult to elicit negative feedback. Some users may be afraid of further unpleasant interactions with the other party and so they may refrain from reporting negative experiences.

The next phase, aggregating feedback, poses its own problems. How is feedback aggregated so that it can be useful to participants in their trusting decisions? Are ratings being weighted in some way depending on how recent they are? For instance, eBay displays net feedback which is positives minus negatives whereas others, such as Amazon, display an average. However, these simple numerical ratings may fail to convey important information about online interactions. For instance, did the feedback come from low- or high-value transactions? What were the reputation ratings of the evaluators?

The distributing phase raises additional issues. Who has access to the reputation ratings? Everyone? Only registered users? As each site uses its own way of representing ratings there is a problem with the portability of one's ratings from one system to another. Ideally, for instance, one would like to be able to import one's ratings from eBay into another system, to establish a good name quickly, rather than start from scratch. However, at the time being this is beyond the capabilities of current reputation systems. Finally, name changes present a problem in reputation systems. Users may register with a number of pseudonyms, essentially erasing the effects of prior feedback. One solution might be to assume that new users, i.e. those with no feedback, should always be distrusted until they have somehow proven themselves, either through paying an entry fee (deposit or insurance) or by accepting more risk or worse prices while building their reputations. An alternative is to prevent name changes either by using real names or through requiring users to acquire unique pseudonyms, namely once-in-a-lifetime pseudonyms. Another problem of centralized reputation systems is that of vulnerability to falsified information. In such systems, ratings and reports can be easily falsified. For instance, a seller can easily rate itself several times highly in order to increase its reputation rating and attract more buyers. One solution that has been proposed is the use of decentralized reputation systems (Schneider *et al.*, 2000; Yu & Singh, 2002). In such systems, participants keep their own database of reputation ratings about others with whom they have interacted or they have been told about. Participants

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

can exchange and update their reputation databases when they meet with others, and in this way information about interactions is propagated out to the rest of the participants.

As e-markets on the Internet flourish, reputation systems may constitute a viable solution to the problem of trust management in particular in short-term relationships. Their power stems from their ability to facilitate and induce cooperation in uncertain and often precarious environments without the need for costly and complex enforcement institutions. Reputation systems are being powered by the participants themselves—the community—and work to create a safer community. Nevertheless, there are many challenges that need to be addressed, in particular with regards to fairness and non-manipulation.

8 Security

Security is essential in any computer system and even more so in an open network such as the Internet in which agents, hosts and other entities whose credentials are unknown, or may not be benevolent, have to interact with each other. Hosts, agents and other entities are vulnerable to a number of attacks ranging from eavesdropping communications to stealing sensitive and private data and inflicting damage by design through viruses. It is essential that hosts, agents and their exchanges are appropriately safeguarded against these attacks. Security encompasses mechanisms to ensure confidentiality, integrity, authentication, access control, availability, non-repudiation, and finally logging and auditing (Jansen & Karygiannis, 1999; Rhee, 2003; Stallings, 2003).

8.1 Confidentiality

Typically, confidentiality is the assurance that information about identifiable persons will not be disclosed without consent from those persons, except as allowed and required by law. Identifiable information is any information that will identify or may reasonably lead to the identification of individuals, such as personal, medical and financial information, data on shopping patterns and demographic information. The release of such information would constitute an invasion of privacy for any individual. Hence, confidentiality is closely related to privacy, but the two are not identical. Confidentiality refers to the obligations of individuals and institutions to use information that has been disclosed to them and is under their control appropriately. Thus, rules of confidentiality are observed out of respect for, and to protect and preserve, the privacy of others.

Confidentiality in our specific context is about ensuring that information is only accessible to those only who have authorized access. Confidentiality extends to the agent's code as well as data.

As an inevitable consequence of the nature of tasks being delegated to software agents, namely to conduct business on behalf of their users, they have to carry or exchange data that enable them to do so. Such data can be sensitive financial data, such as credit card numbers or bank account numbers and passwords or pin numbers, or private data that enable the identification of a user such as a name, home address or security number. Protecting these data is of paramount importance in an e-market. Other types of data may also have to be kept confidential. For instance, the terms of a contract with a company may need to be held confidential and protected from competitors. An agent's code may be encoding trading and negotiation strategies or proprietary algorithms; therefore, it is important that these remain confidential as well. Finally, communications among agents, users and hosts that are designated as confidential need to be so, i.e. only an authorized recipient should have access to the contents of the message, otherwise, it should not be possible to obtain any significant information about the message contents. Besides, messages exchanged between hosts may actually involve the exchange of mobile agents.

8.2 Authentication

Authentication is the process of determining whether an agent, a host or a user is, in fact, who or what it claims to be. Thus, a host, an agent or a user attempts to establish that a communication received from another entity is authentic, i.e. it does indeed originate from the latter.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

Authentication is closely related to authorization, but the two are not the same. Authorization is the process of giving someone permission to have access to something. A computer system is supposed to be accessed only by those authorized. Thus, it must attempt to detect and exclude any unauthorized parties. Access to a system is usually controlled by putting in place an authentication procedure to establish with some degree of confidence the identity of the entity requesting access.

8.3 Access control

Access control is about limiting and controlling access to host systems and their resources, applications, agents and data. Thus, each entity trying to gain access must be first identified or authenticated. An agent should be given access rights to information and resources according to the entity that it represents, i.e. an individual or an organization. For instance, an agent representing a user may have access to the user's bank account, but a third agent should not be granted access to this information. Access control is especially important in mobile agent systems as the accessing of resources by mobile agents must be strictly controlled and monitored.

8.4 Integrity

In general, integrity is the assurance that information can only be accessed or modified by those authorized to do so. Integrity has to do with making sure that the agent's code, data or communication exchanges have not been tampered with, altered or destroyed, either intentionally or by accident. Messages should not be duplicated, reordered or replayed. With mobile agents it is vital to ensure that an agent's code cannot be tampered with, and the same applies to the host. In particular, checking the integrity of the agent's code is crucial before allowing it to start execution on a host machine. Checking that no illegal alterations have been made to the agent's state is more difficult.

8.5 Availability

Availability refers to the system being able to provide the prescribed services according to its design specification to authorized parties. Denial of service attacks, for instance, can reduce or interrupt a system's availability or even shut it down altogether. Denial of service attacks are the result of entities usually deliberately attempting to consume an excessive amount of a system's computational resources or services, so that service to other entities (agents, processes, users) is disrupted.

8.6 Non-repudiation

Mechanisms to ensure non-repudiation are vital if e-markets are to be trusted. In essence, non-repudiation prevents either the sender or receiver of a communication from denying that communication. Non-repudiation of origin proves that data have been sent by a particular sender, while nonrepudiation of delivery proves they have been received by the receiver. As software agents close deals and agree on contracts, some means of ensuring that either party cannot deny that a deal or communication has ever taken place is becoming extremely important. Electronic transactions are also potentially subject to fraud. Participants may attempt to use various pretexts to repudiate a transaction. For instance, they may claim that their system has been broken into or has been infected with Trojan horses or viruses to repudiate a transaction.

8.7 Auditing

Some form of logging and auditing must be provided. Records of security related activities including communication exchanges and transactions among agents, users and hosts as well as information on accessing data or resources by agents and hosts need to be maintained for future reference.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

In case of damage caused to a host, agent or data or a potential dispute the audit trail can be checked and actions by individual entities can be verified. Logging transactions and communication exchanges also aids non-repudiation.

9 Cryptography

Modern-day cryptography is considered the science of information security. Cryptography uses linguistic and mathematical techniques for securing information, in particular communication exchanges (Rhee, 2003; Stallings, 2003). In today's Internet-centred world, the field of cryptography has expanded its remit: modern cryptography provides mechanisms for more than just keeping communication exchanges secure, and has a variety of applications including authentication, digital signatures, electronic voting and digital cash. There are two types of modern cryptosystem: symmetric and asymmetric.

Symmetric encryption, also known as conventional or secret key or singlekey encryption, is perhaps the most widely known type of encryption. Symmetric cryptosystems use the same key (the secret key) to encrypt and decrypt a message. The principles are simple. Assume that agent A needs to send a message to B which only B will be able to read. A symmetric encryption algorithm takes as input the plaintext message and by applying a secret key produces a scrambled message, the ciphertext. The ciphertext is then transmitted over the network and on receipt B decrypts it, i.e. applies the algorithm in reverse, using the same key in order to extract the original message. The key is usually a set of bits.

A symmetric cryptosystem is secure provided that the encryption algorithm is hard to break, that both sender and receiver have obtained copies of the secret key in a secure way and that the key remains secret at all times during its use by both parties. The security of symmetric cryptosystems depends on the secrecy of the key and not the algorithm. Provided that an algorithm is strong, the ability of an encryption algorithm to withstand an exhaustion attack increases with respect to the length of the key. Assuming that the key is of sufficient length, usually above 56 bits, the time that it would take to generate all keys, discover the correct one and thus break the algorithm may be prohibitive even for the fastest computers.

The most widely known block[6] encryption algorithms are the data encryption standard (DES), the triple DES (3DES) and the advanced encryption standard (AES). The DES is not considered computationally secure nowadays as its key is only 56 bits long and, with the current performance of computers systems and parallel processing, the key can be potentially discovered within a reasonable amount of time.

However, symmetric cryptosystems present a problem: how to transport the secret key from the sender to the recipient securely and in a tamper-proof way. A number of ways have been devised including physically delivering the key, using a third trusted party or using a public key cryptosystem.

In contrast to symmetric cryptosystems, public key cryptography is a form of cryptography which generally allows users to communicate securely without having prior access to a shared secret key. This type of encryption was proposed by Diffie & Hellman (1976) and constitutes a significant landmark in cryptography.

Public key cryptography depends on the use of mathematical one-way functions, i.e. computations that are easy to do, but hard to reverse. Factoring is one of the most commonly used one-way function. Multiplying two big prime numbers to produce another big number is easy, but factoring this really big number into its two components is very hard. Of course, if one of the big numbers is already known, then factoring is easy. Security is derived from the difficulty in computing the factors p, q if pq is large.

6 A block cipher takes as input the plaintext in fixed-size blocks and produces a block or cipher of equal size for each plaintext block.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

Public key cryptography is asymmetric: it involves the use of two separate keys which are related mathematically, designated as public and private keys. The private key is kept secret, while the public key may be widely distributed. Provided that the keys are sufficiently large, an exhaustive attack would take the world's fastest computers hundreds or thousands of years to obtain the private key from the public key.

Public key cryptography works in the following way.

(1) Each entity (user, host, agent) generates a pair of keys to be used for the encryption and decryption of communications.
(2) Each entity places one of the two keys in a public register or another publicly accessible file, or communicates the key directly to the other party. This now becomes the public key. The second key is kept private.
(3) If B wants to communicate a message to A, B uses A's public key to encrypt the message and subsequently sends it.
(4) When A receives the message, it decrypts it using the private key. If the message is intercepted by a third party, it cannot be deciphered. Only A can decrypt the message, as only A is in possession of the private key which can decrypt it.

Thus, there is no need to distribute any keys. Typically, public key techniques are much more computationally intensive than symmetric algorithms. The two most widely used public key algorithms are the RSA and the Diffie–Hellman.

9.1 How can agents use cryptography in e-commerce?

The most obvious and common application of cryptography is for securing communications, thus ensuring confidentiality and privacy. However, cryptography can also be used to provide for authentication and integrity as well as to ensure non-repudiation of transactions and communication exchanges between agents in e-commerce (Reagle, 1996).

9.1.1 Cryptography for confidentiality and privacy

Encryption can provide for confidentiality—the assurance that no one else can read a message unless they have been entrusted with a private or secret key. When using an asymmetric system, both agents generate their public and private keys, and make available to each other their respective public keys. If agent A wants to send a confidential communication to agent B, then it uses B's public key to encrypt the message, and sends the ciphertext to B. B then uses its private key to decrypt the ciphertext and obtain the original message. What is crucial is that the secret or private keys are secure and remain secret.

A symmetric cryptosystem could be used instead. However, the main advantages of using an asymmetric cryptosystem over a symmetric one are that the two parties can be total strangers with no prior relationships and that they need not agree upon or securely exchange a key.

9.1.2 Cryptography for authentication, data integrity and non-repudiation

Although confidentiality is fundamental to security, another issue that is really important is that of authentication, i.e. being confident that the sender of a communication really is who it claims to be. Public key encryption can be used for authentication purposes. Suppose that an agent A wants to send a message to B. The content of the message may not be that important, but what is really important is that B is in a position to verify that the message has indeed come from A. In this case, A uses its own private key to encrypt the message and sends it to B. When B receives the ciphertext, it can decrypt it using A's public key, thus proving that the message was indeed send by A as this is the only agent that has access to the private key and can encrypt a message that can only be decrypted with A's public key. Thus, authenticity of the message is verified and the source is established. In this case, the entire encrypted message serves as a digital signature. Even if the message is intercepted by a third party, all they would be able to do is read

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

the message by using A's public key to decrypt it. However, the message content is not intended to be confidential in the first place. Nevertheless, it is impossible to alter the message without access to A's private key. Although a third party could change the content of the message, without knowing A's private key, it cannot encrypt it and send it to B as if it was A's original communication. Hence, the content of the message cannot be changed without B realizing it. As a result, data integrity is also ensured. In addition, authentication using public key encryption limits message repudiation: if a message is confirmed as authentic, the sender cannot easily deny sending it. In essence, public key cryptography enables an entity to authenticate messages from complete strangers.

9.2 Digital signatures and certificates

Although digital signatures allow for authentication, in practice, applying the sender's public key is not enough to ensure that a digitally signed document originated from a particular entity (Baker & Hurst, 1998). The recipient must also have confidence that the sender's private key is indeed private and that it actually belongs to the sender. The problem is that any entity X can pretend to be another entity A, and send or broadcast a public key to a number of other entities, who may not necessarily have the means to verify that X is indeed who it claims to be. When there is no prior relationship between sender and recipient, how does the latter ensure that a particular entity is associated with a particular public–private key pair?

One approach that has been developed relies on digital certificates. A digital certificate associates a public key with an entity's identity. In practical terms, a digital certificate is a simple electronic document containing information about an entity including the entity's public key, which is digitally signed by some trusted party. A sender can subsequently attach a copy of its digital certificate, which has been issued by a trusted third party, to any of its messages. On receiving a message accompanied by a digital certificate, the recipient can rely on the public key of the issuing trusted third party to authenticate the message. As a result, a recipient can now link the message to a particular entity. Intuitively, digital certificates are a source of trust: they confirm that a trusted party vouches for the link between an entity and a public key (Baker & Hurst, 1998). The question that arises now is who are the trusted parties and how do we know we can trust them?

A certificate or certification authority (CA) is an entity which issues, publishes, revokes and archives digital certificates that link an entity's identity to its public key.CAsare considered to be trustworthy. Governments, financial and other credible institutions may act as CAs and there are also commercial CAs. A CA will issue a public key certificate confirming that the public key contained in the certificate belongs to the user, organization, agent or other entity noted in it. Intuitively, if one trusts the CA and can verify its signature, then one is also assured that a certain public key does indeed belong to the entity identified in the certificate. The role of the CA is to verify the credentials submitted by applicants, so that the relying parties can trust the information contained in the certificates. To verify the credentials of various entities, CAs often use a combination of procedures and information including enquiring government bureaus, the payment infrastructure, databases and other services.

One of the most common implementations of digital certificates employs the X.509 standard, which defines a framework for the provision of authentication services (Rhee, 2003; Stallings, 2003). X.509 is sometimes referred as the credit card model, as it allows a hierarchy of trusted parties to authenticate other entities in very much the same way as in the credit card industry (Baker & Hurst, 1998). When a consumer purchases goods and pays by credit card, the vendor does not necessarily know the consumer. However, the vendor is willing to accept the credit card as payment. This is because the vendor trusts the credit card, which is a form of digital certificate. The credit card was issued to the consumer by a trusted financial institution, i.e. a bank, whose name is on the card. The vendor may know and trust this bank. However, even if the vendor does not know the bank—it may be a foreign bank—it knows that the bank has

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

been authenticated by the credit card company, i.e. Visa or Mastercard. Hence, if the vendor knows and trusts the credit card company, it can trust the issuing bank and the consumer.

In a similar way, X.509 allows a hierarchical chain of CAs to issue digital certificates that the recipient of the certificate can verify. The original CA in this trust tree is called the root. X.509 has two main advantages. First, many certificates can be related back to a trusted root. Thus, when an entity digitally signs a document, it sends its digital certificate along with all the supported digital certificates associated with its trust hierarchy. Second, it facilitates trusted communication among entities who may not necessarily know each other or have had any prior exchange. The main disadvantage of X.509 is that if the CA root or any member is compromised, then the rest of the tree becomes unreliable and inevitably the security of the system breaks down.

An alternative approach is termed the web of trust model. Similarly to X.509, a web of trust uses digital certificates to link a public–private key pair to an individual. However, in contrast to X.509, the web of trust model is not hierarchical: there are no CAs or trusted third parties. Instead individuals sign each other's keys forming a web of intertwined signatures.An entity can subsequently decide who to trust and how much, based on the confidence in the keys as well as the entities introducing and vouching for a particular participant. This model is better suited for small communities of users who perhaps have regular contact, but it is difficult to implement on a large scale. Pretty good privacy (PGP), one of the best known applications, primarily for encrypting e-mail communications and data files, is based on the web of trust model (Rhee, 2003).

In the same way that CAs can issue certificates to verify the link between an individual or business and a public key, they can issue agents with certificates to verify their link with public keys. Thus, agents can check each other's certificates or credentials before entering into negotiations or completing a transaction. Such infrastructure increases trust in the marketplace and reduces the inherent risks which emanate from dealing with previously unknown participants.

10 Privacy, anonymity and agents

The digital world has made individuals both stronger and weaker: stronger as a result of abundant choice, democratizing access to information, new forms of social interaction and opportunities, and weaker in terms of threats to one's privacy, new forms of electronic fraud, as well as lack of safeguards, clear rules, laws and jurisdiction in a networked borderless world.

10.1 Agents and privacy

Privacy denotes a condition or state in which a natural or legal person (organization, corporation) is more or less inaccessible to others, on the physical, psychological or informational plane (Bygrave, 2001). The ability of natural and legal persons to have control over the release and distribution of information about themselves is crucial to informational privacy. For our discussion, privacy denotes a state of limited accessibility on the informational plane. A breach of privacy may have serious consequences for the user with regards to autonomy, security, integrity and dignity. Agent technology plays a mixed role in relation to one's privacy. On the one hand, agents can render their users and owners more vulnerable to loss of privacy. On the other, they can be used as a means to safeguard and preserve privacy.

As mentioned before, in order for agents to be able to perform transactions on behalf of their owners/users, they may have to carry or exchange sensitive and private information such as personal or financial data. If intercepted by a third party, such data render the user/owner of the agent identifiable, thus breaching their privacy.

An agent encapsulates part of its user's traits and needs through encoding their preferences and other information; in other words, an agent may embody the profile of its user. Data on a user can be collected as an agent performs seemingly innocuous tasks such as searching for information or articles in a repository, searching for a product or comparing prices. For instance, an interest in anarchism may be inferred if a user's agent is searching for articles on anarchist

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

theories, perhaps as part of their research into political science and the relation, if any, between anarchism movements and women. However, one can imagine situations where this is taken one step further, and the user may be suspected of conspiring to overthrow their government based on the articles searched.Although this may seem extreme, in the current climate of continuous alert and fear from terrorism, anything can be taken out of context, sometimes with irreparable consequences.

The monitoring and tapping of agent operations might reveal a considerable amount of information about their users. This information can be used for building or augmenting user profiles, which can be subsequently used by third parties in various ways such as direct marketing, further exacerbating the initial loss of privacy. Of course, the collection of such information and the building of profiles can be facilitated by other agents. Even more worryingly, agent operations can be covert or non-transparent. For instance, an agent can be instructed to monitor someone without the knowledge of the subject. This would compromise the privacy and related interests of the latter, although not those of the agent's owner. Moreover, an agent could pass on information about its user to third parties without the user being aware of the information transfer. For instance, an agent provider (AP) that hires or sells agents to users may collect information on the users and their activities through the agents (Bygrave, 2001). An agent may be recording transactions, sites visited, searches performed and other information and reporting back to the AP, unbeknown to the user. In effect, such an agent would be acting similarly to a Trojan horse.

10.2 Personalization and privacy

E-commerce provides a unique opportunity when compared to traditional commerce. The use of the Internet enables vendors to know the identities of the users who enter their stores. The ability to know exactly who is surfing a vendor's site is invaluable. Having information on the individual users enables merchants and vendors to offer personalized services (Vulkan, 2003).

Agent technology can be used to gather information on the identity of a user and their characteristics in a number of ways: based on an acquired user profile, attributes inferred from observing the user, through learning, i.e. by monitoring the user behaviour and learning their preferences, and correlation with other users (Resnick & Varian, 1997). Even the simple processes of registering with a Website or vendor enables a vendor to start building a profile on the user. Every move that the user makes as it visits different pages can be recorded and be used in order to infer the user's potential interests. Other technology exists to enable vendors to monitors the users' behaviour. Spyware is the general term used for such software, which was coined by American software writer, Steve Gibson. Spyware describes software that is able to track the user's behaviour unbeknown to the user and then transmit this information to the person who is gathering the data, be it a vendor, organization or even a government. Intrinsically, spyware breaches the rights of privacy of the individual user. Naturally, users object to this and there have been a number of instances of users suing the organizations responsible for such breaches of privacy. However, a vendor or organization does not necessarily need sophisticated spyware to be able to track the user's moves and build profiles. Cookies are used to track the movements of users on Websites. A cookie is nothing more than a string of information that is stored from a given Website on the user's machine in order to recognize each user. Cookies can be linked into an organization's or vendor's profiling system and used as a means to first identify a user, and then keep track of its movements, such as pages visited and duration or purchases made. Much information can be mined out of user profiles, such as habits or even spending power. For instance, by tracking the number of holidays and the amounts spent on each, a system can infer information about the spending power of an individual; holidays can subsequently tailored, i.e. the offer luxurious packages only, or only packages with five-star hotels. When the user has

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

given explicit permission for such data to be accumulated and subsequently used for marketing purposes, then there is no problem. The problem arises as the users are often unaware that such data are stored and used in this way. In general, the problem with spyware and cookies is that of regulation. As the Internet has no governing or regulating body, breaches in privacy rights are difficult to regulate. Even if a vendor is not gathering consumer-related data on their own, it is very easy to purchase such data from other retailers and marketing companies.

Vendors can use such information to tailor their offerings to the requirements of the individual user: they can guess which goods or services the user is likely to want to purchase next and offer them in advance. This is one-to-one or push marketing. Thus, vendors have the opportunity to develop personal relationships so that they can better respond to their customers' needs and develop strong customer loyalty. In addition, information on users enables vendors to provide a unique experience to each user, in essence creating a virtual store based on their preferences. This is one way of achieving mass customization (Pine, 1999). Information technology allows vendors to tailor-make the presentation of their stores to the users' requirements and present them only with products and services that may be of interest to them.

However, apart from tailoring their offers and services to individual customers, vendors can use the information held on individuals to measure their desire to acquire a particular product. Accordingly, they can charge different prices to different customers. The tactic of price discrimination that can be facilitated through user profiles is a particularly serious issue and may have negative effects or impact on consumer confidence. For instance, a case involving the online retailer Amazon came to light in September 2000 (Vulkan, 2003). It appeared that Amazon was charging different prices for the same product to different customers. Those that the system could recognize as existing customers were considered as 'loyal' and were offered slightly higher prices than those that the system could not trace or who were occasional buyers. The pricing variations were first noticed when members of the DVDTalk forum realized that they were being charged different prices for the same DVD. The story was reported in the Washington Post. Apparently Amazon had been engaging in price discrimination tactics on various products such as MP3 players and DVDs since the spring of 2000. At the time Amazon claimed that these price differentiations were part of a simple price test. This price discrimination tactic was subsequently discontinued.

10.3 Anonymity

Despite having the potential to compromise the user's privacy, agent technology can also be used to preserve it. For instance, the very nature of the tasks being delegated to agents, i.e. being twice removed from the interface, enables a user to maintain their privacy to some extent. Nonetheless, a third party may actively seek to identify a user.

Another way to safeguard privacy is through anonymity (Brazier *et al.*, 2004a). Anonymity offers a form of privacy and is characterized by the fact that other parties do not know one's identity. There are occasions when anonymity is desirable, for example, to report a crime, search for information, or compare vendors or products. Individuals may not wish to have their searches or purchases recorded because they would want to avoid the junk mail that may result from being identified as the likely purchaser of a particular type of good. Others may be embarrassed to be identified, because of the nature of the material or product. In some cases, having been identified weakens one's negotiation position and bargaining power. If a vendor knows that a user has an interest in science fiction books, because the user has been profiled, the vendor may decide not to offer the user a discount on the next purchase of a science fiction book. To make matters worse, if the vendor practises perfect price discrimination, it may even try to charge the user extra, knowing their interests on the subject (Vulkan, 2003).

In electronic communications, it may often be the case that individuals do not want to reveal their identity. Not revealing one's identity includes concealing one's identity as well as using

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

a pseudonym instead of one's real name. Froomkin (1995, 1996) identified four types of anonymity and pseudo-anonymity in electronic communication, as follows.

(1) Traceable anonymity: the recipient does not have any indication of the sender's identity. Information linking the communication with the sender is in the hands of an intermediary, usually an anonymous remailer.
(2) Untraceable anonymity: the sender of an electronic message cannot be identified.
(3) Untraceable pseudonymity: this is similar to untraceable anonymity, but instead the sender uses a pseudonym to sign a message which cannot be traced back to them. A pseudonym can be used consistently over a period of time to establish a personality and a reputation just like any other online personality.
(4) Traceable pseudonymity: the user signs messages with a pseudonym which can be traced back to them, although not necessarily by the recipient. Traceable pseudonyms can be distinguished into those that have been assigned formally to someone by a third party, and those that have been chosen by the user themselves.

Successful anonymization is technically complex to achieve. Most often users rely on third parties who supply them with the software or the services to facilitate anonymization. Current services facilitating anonymity and pseudonymity that have been designed for human users enable anonymous e-mail and surfing (Martin, 1998). Anonymity in electronic communications can be achieved through anonymous remailer servers. In searching for and comparing the prices of products, the user's anonymity can be protected by passing the requests through one or more anonymizer servers that hide all information about the user's identity. Personas may also be used to protect anonymity as in the IntelliShopper system (Menczer *et al.*, 2002).

However, anonymity is a double-edged sword (Froomkin, 1996). Techniques that facilitate anonymity may incidentally facilitate and support illegal and criminal activities and even encourage dishonest and anti-social behaviour. One can hide behind anonymity to commit online fraud or spread disinformation and hate mail among other things. Anonymity is not prevented by law, but it is not explicitly protected as is the right to one's privacy. The issue is that of finding a balance.

10.4 Protecting privacy

Protecting confidential, sensitive and private information is imperative. An agent carrying or exchanging such information needs to be safeguarded against both passive and active attacks, which are aimed at compromising a user's privacy. This can be achieved in a number of ways, as follows (Brazier *et al.*, 2004a):

(1) placing only minimal confidential information in a software agent;
(2) using appropriate techniques to hide confidential information;
(3) using appropriate protocols when interacting with other agents and agent platforms;
(4) implementing appropriate protective strategies in a software agent;
(5) using appropriate access control policies provided by agent platforms.

Alternatively, one could avoid entrusting a software agent with confidential information altogether, as follows:

(1) use agents without any confidential information, which collect information and report back to the user; upon receipt of information the user can take action;
(2) use electronic cash as a means of payment, which does not require information about the user (Chaum, 1982).

However, an individual or an organization trying to trace the identity of the user could also perform traffic analysis, i.e. track which sites an agent visits or its requests. Traffic analysis could shed light as to who the user of the software agent is. The simplest approach to counter the risk that traffic analysis poses to the user's privacy is through an anonymizing server. An anonymizing

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

server could, for instance, receive requests for information or services from agents and then submit them to third parties after stripping them of all possible identifying information. Provided that there is a high enough traffic, it is difficult for any observer to match agents with specific requests for information or services to third parties. If the requests are sent encrypted with the server's public key, the observer has no way of knowing what the request is all about, and likewise if the answer or result of a service is sent back encrypted (using public key or symmetric encryption), the observer has no way of matching the result or request for a service to a specific user. This preserves the user's privacy to some extent, as it prevents any other third party from identifying the activities of a particular user and building a profile for them through these activities. The volume of requests should be hindering the development of a profile for a particular user.

The traceability of a mobile agent could also be counteracted to some degree if the agent is anonymized, i.e. all identification information is concealed or removed. An anonymizing server would have the task of complicating and ideally blocking any attempt to perform traffic analysis (Brazier *et al.*, 2004a). The server could work by transforming an incoming mobile agent and its identity in such a way that it would be difficult for an observer to relate a software agent that departs the server to an agent that entered it earlier. Again, this solution could work if the volume of incoming and outgoing traffic is sufficiently large. For the destination host, the agent would appear as if it originated from the anonymizing server, and thus the new host would not be able to determine where the agent had been before. Of course, the anonymizing server will have to be regarded as a trusted party in order for a host to accept an incoming agent. An alternative measure would be for the agent to work from varying IP addresses and not visit its home place too often. The advantage of using an anonymizing server is that such an infrastructure makes it much harder and ideally impossible to trace a user by way of traffic analysis: although the agent is known to be affiliated with the server, its real user remains unknown. However, the success of this approach depends on volumes of traffic: too little traffic and it is easy to relate outgoing agents to incoming ones; too much traffic and the server becomes a bottleneck.

Alternative approaches to protecting the user's privacy without using an anonymizing server include the following (Brazier *et al.*, 2004a).

(1) Using a new agent for each new task: a user can have a personal agent who generates task-specific agents, which are subsequently destroyed when they have finished their task. The experiences and information gathered by the task-specific agents can be incorporated into the personal agent.
(2) Hiring agents: agents can be rented or hired from agent providers to perform specific tasks. Although the agent can be traced, the user-client cannot be easily traced, in particular when secure communications are used. It goes without saying that the user needs to trust the agent provider offering this service.

Another alternative is to consider fitting agents with a kind of privacyenhancing technology (PET) such as an identity protector (Senicar *et al.*, 2003). An identity protector can be used in two ways: it can be placed between the user and the agent, or between the agent and the environment. In the former case, the agent cannot collect or exchange any personal data about the user without the prior and explicit approval of the user. Hence, the user has total control of the amount of personal data that is being passed on to the agent. The latter approach gives the agent complete power to obtain and record personal data from its user. The identity protector in this case prevents the agent from distributing the user's personal information unless it is essential to execute a transaction, for instance.

11 Agents and the law

As agents interact on the Internet and participate in e-markets, they engage in a number of activities which are significant from a legal perspective: they access computer systems, networks and

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

data; they retrieve and distribute information; they mediate personal and business relations; they negotiate for and buy and sell goods and services. Autonomy implies that agents act without the constant and direct intervention of their users/owners.

Naturally, the introduction of electronic agents that conduct business on behalf of natural or legal persons raises issues with regards to accountability and liability (Apistola *et al.*, 2002; Andrade *et al.*, 2004). An agent could perform transactions that are very difficult to undo or correct, i.e. buying goods at exorbitant prices or buying multiple items. Depending on the particular situation and the agent's complexity, it might be difficult to determine what exactly went wrong. Thus, through their activities agents may trigger consequences for relevant parties relating to different areas of the law, such as contract, tort, intellectual property and data protection, and even criminal responsibility. The traditional paradigms have all developed along the lines of physical and geographical bounds of space and time. However, the Internet society is not confined to geographical, legal or corporate boundaries. Inevitably, new technologies such as the Internet, and in particular software agents, have an impact on society and consequently on the law too. Policy-makers need to consider the repercussions of this new technology and provide safeguards and appropriate legislation that covers their operation.

There are four parties that are involved in the use of agents and whose interests are affected by any legal considerations (Geurts, 2002).

(1) Agent designer/developer. The agent designer/developer is the person or organization who designs and develops the agent. The designer/developer determines the agent's functionality and sophistication. This is also the person that puts the user manual of the agent together.
(2) Agent supplier or provider. The supplier is the person or organization who sells the agent to the user. The agent designer/developer and supplier may be the same person. The supplier may be able to customize the agent according to the user's preferences. An agent provider is a person or organization that leases the agent to the user; in this case the agent is not sold to the user, instead it is rented.
(3) Agent user. The person or organization who uses the agent. If the agent has been bought from a supplier, then the user is also the legal owner of the agent. The user may configure and customize the agent according to one's personal preferences.
(4) Third parties. Third parties are all entities other than the user with whom the agent interacts, i.e. other agents, users, organizations, hosts. A third party can also be an entity who tampers with the agent.

Providing a framework for guiding and regulating agent-based interactions and exchanges is an interesting problem for lawyers and legal theorists, but also a fundamental issue for the success of agent technology in e-commerce (Sartor, 2002). The question that arises then is who can be held liable, and on what grounds, for infringements made, or for damages caused, by a software agent.

Tort or wrongful acts. The use of an agent can result in damages, which can lead to an action by the counterparty who has suffered the damages in order to obtain compensation. However, who is to be held liable for an agent's misbehaviour? The designer/developer, who perhaps did not test the agent extensively and did not verify its behaviour? What if an agent has been customized by a supplier or provider? Maybe the customization had an adverse effect on the agent's behaviour—it may have been impossible for designers/developers to test all possible configurations. Did the user of the agent do something wrong which caused the agent to misbehave? Or was it a third party that interfered accidentally or by design with the agent which caused this undesirable behaviour? To make matters worse, agents may have the ability to learn and adapt to their environment. As agents learn, their behaviour becomes less predictable, in the sense that the exhibited behaviour is different from the originally designed behaviour. Is it the learning process that is at fault? Was the misbehaviour as a result of an exceptional situation that perhaps the agent did not know how to deal with?

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

Privacy and data protection. Agents may be used to collect information about other users (or even one's own user) unbeknown to the subjects. These data can be personal or financial data which characterize persons and can be used for identifying them, profiling them and subsequently using these profiles for direct marketing and other purposes. Such a collection of data is regarded as unlawful and the person's right to privacy has been breached. Is it the agent's user who is liable? Or the designer/developer who has actually developed a product that infringes individual rights?

Intellectual property rights. Information and search agents can collect and aggregate information from various sources. For instance, articles or other scientific data can be downloaded from repositories. However, information or data may be copyrighted and thus an agent may be infringing copyright laws by downloading information on behalf of the user. So far, there is no standard way of indicating that a piece of information is copyrighted on the Internet. However, this may change in the near future. Nevertheless, this is an important issue, as an agent and its user may find themselves accused of breaking copyright laws.

Product liability. Assuming that an agent is a product, then the liability for a defective agent could potentially be shifted to the designer/developer. The designer/developer may be held responsible for any damages caused by the agent if the exhibited behaviour does not conform to the design specification. However, what if the agent has been customized or configured by the user or another party? Who is liable then? What if the agent learns and adapts to the environment?

Contractual liabilities. Who is to blame if a software agent that has been contracted to provide a service to another software agent does not provide the agreed quality of service or does not provide the service at all? Is it the software agent that is held accountable and, if yes, what sort of sanctions can there be imposed on a software agent? Or is it the legal owner or user that bears the responsibility of ensuring that the agent's behaviour is correct, that it abides by the necessary protocols, and honours commitments and contracts? The former hinges on the possibility of software agents being regarded as legal persons in the eyes of the law. Is the user bound to the terms and conditions of a contract that an agent signed? Are there any situations in which the user is not bound to the terms and conditions?

Criminal responsibility. There may be cases when agents inflict financial losses on their users or other parties. Such agents may very much act like viruses and Trojan horses, stealing private and sensitive information and using it by design to cause damage. Who is to be held responsible then?

Another interesting issue that arises is with regards to the international nature of the transactions that agents may be carrying out on behalf of their users. The physical location of information, processes, services and agents is no longer necessarily the same as the perceived location. This then raises questions regarding which laws apply to a particular situation, such as a transaction that has been conducted by agents that perhaps reside or represent users/owners that reside in different countries governed by different laws. Under which jurisdiction does a transaction fall when it involves a bidding agent based in Sweden and an auction house based in the USA?

11.1 Current legislative provisions

There have been some limited regulatory initiatives with regards to software agents in the USA and Canada. The Uniform Electronics Transaction Act[7] (UETA 1999) has been adopted in several states in the USA and is an effort to establish uniform legal standards for e-commerce transactions. UETA applies to transactions in which parties have agreed to conduct transactions via electronic means. The objective of the Act is to regulate e-commerce and in particular it attempts to establish the legal equivalence of electronic records and signatures with paper writings and traditional signatures. The underlying principle is that electronic transactions should be

7 http://www.law.upenn.edu/bll/ulc/fnact99/1990s/ueta99.htm.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

enforceable at the same level as traditional paper transactions without having to change the rules of law that apply. The Electronic Signatures in Global and National CommerceAct, known as E-Sign[8], came into effect in October 2001. E-Sign, which comprises federal legislation, is designed to bridge the gap between business transactions and online technology. Its primary objective is to facilitate the growth of e-commerce by removing existing legal impediments to the use of electronic contracts. Section 101(a) of the Act establishes that a signature or contract may not be denied legal effect 'solely because it is in electronic form'. The Uniform Computer Information Transactions Act[9] (UCITA) is a more ambitious act in relation to UETA as it attempts to take into account and support recent technological advances, such as for instance software agents. The purpose of the Act is to provide a uniform commercial code applicable to computer information transactions. UCITA refers specifically to electronic agents and provides a legal basis for using agents in e-commerce and some limited provisions on what to do in case something goes wrong. Accordingly, the actions of an agent are attributed to the human who uses the agent. Hence, a person assents to a declaration or a contract of their agent, although they may not be aware of the agent's actions at the time, i.e. responsibility lies with the agent's user or owner. A contract is valid if it is formed by the interaction of electronic agents when they engage in operations that indicate acceptance of an offer, unless the court rules that it is the result of fraud or mistake. The interaction between electronic agents and other natural or legal persons acting on their own, or on behalf of someone else, can also form a valid contract.

Similarly to UCITA, the Uniform E-Commerce Act[10] (UECA) in Canada also defines what an electronic agent is in the context of this statue (Article 19) and also establishes that such agents can form contracts (Article 21). Although these efforts comprise the first step in accepting the use of agent technology as a means to conduct business online, they leave many unanswered questions, in particular with regards to the issues that were raised in the previous section.

Agents do not feature in any European initiatives or current legislation. The EU Directive[11] on e-commerce does not include legislation on electronic agents.

The critical question now is whether the current legislation offers protection for all parties involved when agents are used in e-commerce transactions. For the time being perhaps the answer is yes, as the use of agent technology is neither widespread nor sophisticated. However, in the long run, the issues surrounding agents in e-commerce will have to be carefully considered as intelligent agents that act proactively and autonomously, furthering their design objectives, exemplify qualities that are not covered in the current legislation.

11.2 Agents as legal persons

Legal systems recognize two types of legal person as having will, intention, obligations, rights and contractual capability: natural persons and legal entities. Natural persons are human beings whereas legal entities can be associations, organizations, companies, governments, etc. A contract is a legal transaction comprising declarations of intention of at least two persons, one stated relatively to the other (Weitzenbock, 2001). Contracts are usually enforceable by law. A contract consists of the main obligations of the two (or more) parties coming in agreement and the terms and conditions of the contract. Contractual capability is the ability to perform legal transactions effectively, i.e. actions that imply legal consequences. Having contractual capability involves declaring intentions, but this is only meaningful if the acting persons can understand the consequences of their declarations.

8 http://www.cio.noaa.gov/itmanagement/pl106229.pdf.

9 http://www.law.upenn.edu/bll/ulc/ucita/citam99.htm.

10 http://www.ulcc.ca/en/us/index.cfm?sec=1&sub=1u1.

11 http://www2.dti.gov.uk/industries/ecommunications/ electronic_commerce_directive_0031ec.html.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

For the time being, computer systems and software applications are not attributed any legal status; they are deemed tools to be used by legal or natural persons. However, an agent's nature may be very different to that of a software application. Intelligent software agents are attributed exceptional characteristics such as autonomy, goal-directness and sociality. They become capable of learning from experience and adapting to their environment according to cognitive, reactive and proactive processes quite similar to those used by humans. They are also ascribed mental attitudes such as beliefs, desires, intentions and commitments. These characteristics set agents apart from other pieces of software, rendering them unique.

Agents used in e-commerce applications may negotiate contracts or bid in auctions on behalf of their users. However, such actions imply that a legal transaction takes place between the agent and the other party, which can be another agent, natural person or legal entity (Weitzenbock, 2001; Brazier *et al.*, 2003). For instance, when an agent bids in an electronic auction, it declares its intention to acquire an item for the price of the bid submitted. If it wins the auction, then it has the obligation to provide payment equal to the value of its bid or as otherwise dictated by the rules of the auction. However, this raises some issues with regards to the legal status of agents. In essence, the fundamental question is, can software agents be considered as legal personas? One could argue that agents do not currently have the ability to understand their declarations and their implications, and therefore they cannot engage in any legal transactions—any such transaction is void. Thus, in the eyes of the law, agents do not have a contractual capability, and this presents a problem. If they lack such a capability, then they cannot perform legal transactions such as closing contracts and bidding in auctions. If the agents are to be considered acting as the representatives of users, they still need to have some contractual capability; unless the will or intention to perform an action is distinguished from the action itself. The will or intention is attributed to the natural or legal person, whereas the action is that attributed to the agent. In this case, responsibility lies with the principal, i.e. the agent user. Thus, if the agent was to sign an unauthorized contract, the responsibility would still lie with the user, as it would be assumed that the user has the intention or will and the agent carries out the action. Hence, the risk of transactions would be put entirely on those persons who program, control or otherwise use a software agent. Alternatively, a user could be required to explicitly authorize an agent to sign a contract on their behalf before it is actually finalised. Still, this limits the applicability of agent technology; one of the main advantages of using software agents is that certain processes can be automated as agents can act autonomously in order to further their users' goals.

An alternative and radical proposal is that agents could be potentially attributed a legal persona or they could be granted a special status, that of an e-person (Andrade *et al.*, 2004; Wettig & Zehender, 2004). Such a provision would solve a number of problems. First, the issue regarding the validity of the declarations and contracts being negotiated and concluded by software agents would be resolved. As agents would have some form of legal status, they would be considered as being able to express will and intention. This does not necessarily have an impact on legal theories about consent and declaration, contractual freedom and conclusion of contracts, as legal systems already consider the declarations of other non-natural persons, i.e. legal entities such as organizations. Second, it could limit the owner's liability and responsibility for the agent's behaviour.

Although convenient in some respects, this solution presents a number of difficulties. First of all, what constitutes an agent would have to be answered: is it the software, the hardware or both; and what about mobile agents? More fundamentally, the essential attributes or characteristics that an agent needs to possess or exhibit in order to be granted the status of a legal person or e-person would have to be clearly defined. As natural or legal persons are identified uniquely, the same requirement applies to agents; agents would need to have unique identities. Another requirement for an agent to be rendered a legal person is that it must have a residence or domicile (Andrade *et al.*, 2004). However, mobile agents do not have an established physical location, although in this case the domicile of the agent could be considered to be that of its user.

Agent registers could be set up in a similar fashion to company registers. Such registers could be maintained by trusted third parties such as governments or CAs. Through this registration

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

procedure, an agent could be attributed a unique identity and a domicile. The owner could also grant the agent with a minimum amount of capital which would act as the agent's patrimony (Sartor, 2002). This fund would be deposited in order to ensure that an agent could fulfil its financial obligations and liabilities and it would represent a warranty for the counterparties. As liability safeguarding is very important, this fund could back up claims of the contracting parties in case of problems. Another measure to protect users and other counterparties against possible liabilities would be to establish a compulsory insurance regime for agent-related activities. Hence, a type of agent with limited liability is indeed possible.

When agents are involved in the negotiation of contracts, each contracting party, agent or non-agent can use the agent register to check the soundness of the agents with whom negotiations are being conducted before deciding on the conclusion of the contract. Reputation systems could also be employed to augment this process. However, sometimes negotiations fail. Most often, parties go their separate ways without further obligations. However, in some cases, the negotiations are not fully noncommittal (Weitzenbock, 2002). Costs may have been incurred or expectations may have been raised that have led to breaking off negotiations with other parties, for instance. From a legal point of view, this is the domain of precontractual liability. The agent's patrimony or insurance policy can be used to cover any costs incurred as a result of negotiations being abandoned.

The concept of an agent register raises some privacy issues with regards to the information that would be included in such a register. For instance, would the register contain information connecting the agent to a unique user? Would access to the register be open to everyone? If a user can be identified through its agent, then the user could be profiled through the agent's actions. Access could be limited to only checking that an agent is actually registered with a particular authority as it claims. Access to information regarding the user's identification could be restricted and only granted in exceptional circumstances, i.e. in case of litigation. Moreover, along with the agent's identification, its public key could also be included, which would allow any other entity to communicate and transact with the agent in a secure way.

Of course the crucial question still remains: who is liable? If business is done without any problems or complications, then the legal concerns are irrelevant. However, if problems emerge, legal provisions are crucial. If the agent is regarded as a legal person or an e-person, then in principle *it* is liable. However, as claims practically cannot be realized against the agent, liability falls back again on the owner of the agent. The agent's patrimony, as well as the insurance covering its activities, could be used to cover any liabilities. Of course, an agent should be attributed a way of being represented in case of legal actions against it, as a claim may well be unfounded or malicious.

12 Conclusions

Agent technology for e-commerce is a fascinating area of research and development with the potential to bring about significant benefits to both individuals as well as businesses and organizations. However, it also has the potential to bring about significant benefits to society as a whole.

The Internet, and in particular the use of agent technology to automate financial transactions and other tasks, challenges our traditional views of regulation and responsibility. We may have to rethink and re-evaluate the way that agents operating autonomously can be regulated. Eventually, such smart pieces of software may have to be granted the status of a legal person in order to overcome problems with legal responsibilities in contracting or executing financial transactions. Punishment may involve restricting the agent's rights and access to resources, platforms and marketplaces or even termination.

Agent developers need to know how to design and develop software agents that comply with known legal requirements and the policies of institutions. They also need to adhere to good practices and ethical considerations. Thus, it is crucial that the legal and technical communities embark on interdisciplinary research to deal with these issues. Ultimately such considerations and provisions could help promote trust and confidence in agent technology.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

In time, we may have to rethink certain issues, such as for instance the law and how it applies to the Internet. A possible scenario involves the Internet being considered as a virtual environment, which will be governed by its own laws, regulations and in which different norms may apply. Inevitably, the Internet may be subject to international law, and not that of any single country. How such regulations can be enforced while at the same time the users' rights and privacy are preserved and protected needs to be very carefully considered.

Acknowledgements

This is a revised version of 'Trust, security and legal issues' which appears in Fasli, M. 2007 *Agent Technology for E-commerce*. New York: Wiley. Reproduced with permission.

References

Andrade, F., Novais, P. & Neves, J. 2004 Issues on intelligent electronic agents and legal relations. In *Proceedings of the Law of Electronic Agents Workshop (LEA-04), Rome, Italy*, June 30, pp. 81–94.

Apistola, M., Brazier, F., Kubbe, O., Oskamp, A., Schellekens, M. & Voulon, M. 2002 Legal aspects of agent technology. In *Proceedings of the 17th BILETA Conference, Amsterdam, The Netherlands*, April 5–6.

Arunachalam, R. & Sadeh, N. 2004 The 2003 supply chain management trading agent competition. In *Proceedings of the Trading Agent Design and Analysis Workshop (TADA), Acapulco, Mexico*, August 11.

Axelrod, R. 1984 *The Evolution of Cooperation*. New York, NY: Basic Books.

Baker, S. A. & Hurst, P. R. 1998 *The Limits of Trust*. The Hague: Kluwer Law International.

Bickmore, T. W. & Cassell, J. 2001 Relational agents: a model and implementation of building user trust. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI, Seattle, WA*, March 31–April 5, New York, NY: ACM Press, pp. 396–403.

Brazier, F. M. T., Oskamp, A., Prins, C., Schellekens, M. H. M. & Wijngaards, N. J. E. 2004a Anonymity and software agents: an interdisciplinary challenge. *Artificial Intelligence and Law* **12**(1–2), 137–157.

Brazier, F. M. T., Oskamp, A., Schellekens, M. H. M. & Wijngaards, N. J. E. 2003 Can agents close contracts? In Oskamp, A. & Weitzenboeck, E. (eds.), *Proceedings of the Law of Electronic Agents Workshop (LEA-03), Edinburgh, UK*, June 24, pp. 9–20.

Brazier, F. M. T., Oskamp, A., Prins, J. E. J., Schellekens, M. H. M. & Wijngaards, N. J. E. 2004b Law-abiding and integrity on the Internet: a case for agents. *Artificial Intelligence and Law* **12**(1–2), 5–37.

Bromley, D. B. 1993 *Reputation, Image and Impression Management*. Chichester, UK: Wiley.

Bygrave, L. A. 2001 Electronic agents and privacy: a Cyberspace Odyssey 2001. *International Journal of Law and Information Technology* **9**(3), 275–294.

Chaum, D. 1982 Blind signatures for untraceable payments. In Chaum, D., Rivest, R. L. & Sherman, A. T. (eds.), *Advances in Cryptology: Proceedings of CRYPTO'82 Conference, Santa Barbara, CA*, August, New York: Plenum, pp. 199–203.

Child, M. & Linketscher, N. 2001 Trust issues and user reactions to e-services and electronic marketplaces: a customer survey. Technical Report HPL-2001-32, Hewlett Packard, Trusted E-Services Laboratory, HP Laboratories, Bristol.

Claessens, J., Preneel, B. & Vandewalle, J. 2003 (How) can mobile agents do secure electronic transactions on untrusted hosts? A survey of the security issues and the current solutions. *ACM Transactions on Internet Technology* **3**(1), 28–48.

Dash, R. K., Jennings, N. R. & Parkes, D. C. 2003 Computational-mechanism design: a call to arms. *IEEE Intelligent Systems* **18**(6), 40–47.

Dellarocas, C. 2000a Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In *Proceedings of the 2nd ACM conference on Electronic commerce (EC-00), Minneapolis, MN*, October 17–20, New York, NY: ACM Press, pp. 150–157.

Dellarocas, C. 2000b Mechanisms for coping with unfair ratings and discriminatory behavior in online reputation reporting systems. In *Proceedings of the 21st International Conference on Information Systems (ICIS-00), Brisbane, Australia, December 10–13*. Atlanta, GA: Association for Information Systems Press, pp. 520–525.

Dellarocas, C. & Klein, M. 1999 Civil agent societies: tools for inventing open agent-mediated electronic marketplaces. In Moukas, A., Sierra, C. & Ygge, F. (eds.), *Agent Mediated Electronic Commerce II, Towards Next-Generation Agent-Based Electronic Commerce Systems, IJCAI 1999 Workshop (Lecture Notes in Artificial Intelligence, 1788), Stockholm, Sweden*, July 31. Berlin: Springer, pp. 24–39.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

Diffie, W. & Hellman, M. 1976 Multiuser cryptographic techniques. In *Proceedings of the AFIPS National Computer Conference, New York*, June. Montvale, NJ: AFIPS Press, pp. 109–112.

Dignum, F. 2001 Agents, markets, institutions, and protocols. In Dignum, F. & Sierra, C. (eds.), *Agent Mediated Electronic Commerce, The European AgentLink Perspective (Lecture Notes in Artificial Intelligence, 1991)*. Berlin: Springer, pp. 98–114.

Dignum, F. 2002 Abstract norms and electronic institutions. In Lindemann, G., Moldt, D., Paolucci, M. & Yu, B. (eds.), *Proceedings of the Regulated Agent-based Social Systems: Theory and Applications Workshop (RASTA'02), Bologna, Italy*, July 16, pp. 93–103.

Doorenbos, R. B., Etzioni, O. & Weld, D. S. 1997 A scalable comparison-shopping agent for the World Wide Web. In *Proceedings of the 1st International Conference on Autonomous Agents (Agents 97), Marina del Rey, CA*, February 5–8, New York, NY: ACM Press, pp. 39–48.

Esteva, M., Rodríguez-Aguilar, J. A., Sierra, C., Garcia, P. & Arcos, J. L. 2001 On the formal specification of electronic institutions. In Dignum, F. & Sierra, C. (eds.), *Agent Mediated Electronic Commerce, The European AgentLink Perspective (Lecture Notes in Artificial Intelligence, 1991)*. Berlin: Springer, pp. 126–147.

Falcone, R. & Castelfranchi, C. 2004 Trust dynamics: How trust is influenced by direct experiences and by trust itself. In *Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-04), New York, NY*, July 23–25, IEEE Computer Society, pp. 740–747.

Fasli, M. 2006 Shopbots: A syntactic present, a semantic future. *IEEE Internet Computing* **10**(6), 69–75.

Froomkin, A. M. 1995 Anonymity and its enmities. *Journal of Online Law* Art. 4. Available at http://www.wm.edu/law/publications/jol/95_96/froomkin.html

Froomkin, A. M. 1996 Flood control on the information ocean: living with anonymity, digital cash, and distributed databases. *Journal of Law and Commerce* **15**, 395–479.

Geurts, R. 2002 Legal aspects of software agents. In Prins, J. E. J., Ribbers, P. M. A., van Tilborg, H. C. A., Veth, A. F. L. & van der Wees, J. G. L. (eds.), *Trust in Electronic Commerce: The Role of Trust from a Legal, an Organizational and a Technical Point of View*. The Hague: Kluwer, pp. 231–270.

Guttman, A. G., Moukas, R. H. & Maes, P. 1999 Agent-mediated integrative negotiation for retail electronic commerce. In Noriega, P. & Sierra, C. (eds.), *Agent Mediated Electronic Commerce, 1st International Workshop on Agent Mediated Electronic Trading (AMET'98), Selected Papers (Lecture Notes in Artificial Intelligence, 1571)*. Berlin: Springer, pp. 70–90.

He, M., Jennings, N. R. & Leung, H.-F. 2003 On agent-mediated electronic commerce. *IEEE Transactions on Knowledge and Data Engineering* **15**(4), 985–1003.

Jansen, W. & Karygiannis, T. 1999 Mobile agent security. Technical Report NIST Special Publication 800-19, National Institute of Standards and Technology. Available at `http://csrc.nist.gov/publications/nistpubs/`.

Jenks, J. M. & Kelly, J. M. 1985 *Don't Do, Delegate!* New York: Franklin Watts.

Kollock, P. 1999 The production of trust in online markets. In Lawler, E. J., Macy, M., Thyne, S. & Walker, H. A. (eds.), *Advances in Group Processes*, vol. 16. Greenwich, CT: JAI Press.

Krulwich, B. T. 1996 The BargainFinder agent: comparing price shopping on the Internet. In Williams, J. (ed.) *Bots and Other Internet Beasties*, SAMS.NET. London: Macmillan, pp. 258–263.

Lee, J. & Moray, N. 1992 Trust, control strategies and allocation of function in human–machine systems. *Ergonomics* **35**(10), 1243–1270.

Linden, G., Smith, B. & York, J. 2003 Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing* **7**(1), 76–80.

Luhmann, N. 1979 *Trust and Power: Two Works*. Chichester: Wiley.

Luhmann, N. 1990 Familiarity, confidence, trust: problems and alternatives. In Gambetta, D. (ed.), *Trust: Making and Breaking Cooperative Relations*. Oxford: Blackwell, pp. 94–109.

Maes, P., Guttman, R. & Moukas, A. 1999 Agents that buy and sell: transforming commerce as we know it. *Communications of the ACM* **42**(3), 81–91.

Marsh, S. 1994 Formalising trust as a computational concept. PhD thesis, Department of Mathematics and Computer Science, University of Stirling, UK.

Martin, D. 1998 Internet anonymizing techniques. *login:Magazine*. Available at `http://www.usenix.org/publications/login/1998-5/martin.html`.

McAfee, R. P. & McMillan, J. 1987 Auctions and bidding. *Journal of Economic Literature* **25**(2), 699–738.

Menczer, F., Street, W. N.,Vishwakarma, N., Monge, A. E. & Jakobsson, M. 2002 IntelliShopper: a proactive, personal, private shopping assistant. In *Proceedings of the 1st International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02), Bologna, Italy*, July 15–19, New York, NY: ACM Press, pp. 1001–1008.

Michalakopoulos, M. & Fasli, M. 2005 On deciding to trust. In *Proceedings of the 3rd International Conference on Trust Management (iTrust) (Lecture Notes in Artificial Intelligence, 3477)*. Berlin: Springer, pp. 61–76.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press

Montaner, M., Lopez, B. & Dela, J. 2003 A taxonomy of recommender agents on the Internet. *Artificial Intelligence Review* **19**(4), 285–330.
Muir, B. M. 1987 Trust between humans and machines and the design of decision aids. *International Journal of Man–Machine Studies* **46**, 485–500.
Norman, D. A. 1990 The "problem" of automation: inappropriate feedback and interaction, not "over-automation". In Broadbent, D. E., Baddeley, A. & Reason, J. T. (eds.), *Human Factors in Hazardous Situations*. Oxford: Oxford University Press, pp. 585–593.
North, D. C. 1990 *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
Oates, D. 1993 *Leadership: The Art of Delegation*. London: Century Business.
Patrick, A. S. 2002 Building trustworthy agents. *IEEE Internet Computing* **6**(6), 46–53.
Pine, B. J. 1999 *Mass Customisation*. Boston, MA: Harvard Business School Press.
Reagle, J. M. Jr. 1996 Trust in electronic markets: the convergence of cryptographers and economists. *First Monday*. Available at: `http://www.firstmonday.dk/issues/issue2/markets/`.
Resnick, P., Kuwabara, K., Zeckhauser, R. & Friedman, E. 2000. Reputation systems. *Communications of the ACM* **43**(12), 45–48.
Resnick, P. & Varian, H. R. 1997 Recommender systems. *Communications of the ACM* **40**(3), 56–58.
Rhee, M. Y. 2003 *Internet Security: Cryptographic Principles, Algorithms and Protocols*. Chichester: Wiley.
Riegelsberger, J. & Sasse, A. M. 2001 Trustbuilders and trustbusters: the role of trust in interfaces to e-commerce applications. In *Towards The E-Society: E-Commerce, E-Business, and E-Government, Proceedings of the 1st IFIP Conference on E-Commerce, E-Business, E-Government (I3E 2001), Zurich, Switzerland*, pp. 17–30.
Rotter, J. B. 1980 Interpersonal trust, trustworthiness and gullibility. *American Psychologist* **35**(1), 1–7.
Sartor, G. 2002 Agents in Cyberlaw. In Sartor, G. & Cevenini, C. (eds.), *Proceedings of theWorkshop on the Law of Electronic Agents (LEA 2002), Bologna, Italy, July 13*.
Schneider, J., Kortuem, G., Jager, J., Fickas, S. & Segall, S. 2000 Disseminating trust information in wearable communities. *Personal and Ubiquitous Computing* **4**(4), 245–248.
Senicar, V., Jerman-Blazic, B. & Klobucar, T. 2003 Privacy-enhancing technologies: approaches and development. *Computer Standards and Interfaces* **25**(2), 147–158.
Shapiro, S. P. 1987 The social control of impersonal trust. *American Journal of Sociology* **93**(3), 623–658.
Sierra, C. & Noriega, P. 2002 Agent-mediated interaction: From auctions to negotiation and argumentation. In d'Inverno, W., Luck, M., Fisher, M. & Priest, C. (eds.), *Foundations and Applications of Multi-agent Systems: UKMAS Workshops 1996–2000, Selected Papers (Lecture Notes in Artificial Intelligence, 2403)*. Berlin: Springer, pp. 27–48.
Stallings, W. 2003 *Network Security Essentials: Applications and Standards*. Upper Saddle River, NJ: Prentice-Hall.
Tschudin, C. F. 1999 Mobile agent security. In Klusch, M. (ed.), *Intelligent Information Agents: Agent-Based Information Discovery and Management on the Internet*. Berlin: Springer, ch. 18, pp. 431–446.
Ulfelder, S. June 2000 Undercover Agents. *Computer World*. Available at `http://www.computerworld.com/news/2000/story/0,11280,45452,00.html`.
Vázquez-Salceda, J., Padget, J., Cortés, U., López-Navidad, A. & Caballero, F. 2003 Formalizing an electronic institution for the distribution of human tissues. *Artificial Intelligence in Medicine* **27**(3), 233–258.
Vulkan, N. 2003 *The Economics of E-commerce*. Princeton, NJ: Princeton University Press.
Weitzenbock, E. M. 2001 Electronic agents and the formation of contracts. *International Journal of Law and Information Technology* **9**(3), 204–234.
Weitzenbock, E. M. 2002 Electronic agents and contract performance: good faith and fair dealing. In *Proceedings of the Law of Electronic Agents Workshop (LEA-02): Selected Revised Papers, Bologna, Italy, July 13*, pp. 67–73.
Wellman, M. P., Greenwald, A. R., Stone, P. & Wurman, P. R. 2003 The 2001 Trading Agent Competition. *Electronic Markets* **13**(1).
Wettig, S. & Zehender, E. 2004 A legal analysis of human and electronic agents. *Artificial Intelligence and Law* **12**(1–2), 111–135.
Wooldridge, M. & Jennings, N. R. 1995 Intelligent agents: theory and practice. *Knowledge Engineering Review* **10**(2), 115–152.
Youll, J. 2001 Agent-based electronic commerce: opportunities and challenges. In *Proceedings of the 5th International Symposium on Autonomous Decentralized Systems (ISADS), Dallas, TX*, pp. 146–148.
Yu, B. & Singh, M. P. 2002 An evidential model of distributed reputation management. In *Proceedings of the 1st International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'02), Bologna, Italy, July 15–19*, New York, NY: ACM Press, pp. 294–301.

https://doi.org/10.1017/S0269888907001014 Published online by Cambridge University Press