Electronic Agents and Privacy: A Cyberspace Odyssey 2001

[Published in *International Journal of Law and Information Technology*, 2001, volume 9, no. 3, pp. 275–294]

Lee A. Bygrave[1]

Abstract

In this paper, analysis is carried out of the impact of the use of electronic agents on privacy and related interests. A preliminary examination is made of the various risks to privacy occasioned by agent operations and of the way in which current data protection rules can mitigate these risks. It is suggested that greater thought be given to drafting data protection rules that take full account of electronic agents.

1. Visions of agents

Gareth Morgan has persuasively shown that we tend to view and understand organisations using various metaphorical images that often work at a subconscious, intuitive level.[2] The same point can be made with respect to electronic agents.[3] Concomitantly, how we perceive the impact of electronic agents on privacy will tend to be shaped in part by certain visions of what electronic agents are and are capable of becoming. As Morgan correctly shows in relation to organisations, such visions are not just interpretive constructs; they also provide frameworks for policy and action, including law making. Thus, when analysing the interrelationship of electronic agents and privacy, we must not lose sight of the influence of these images.

Of the various visions of electronic agents which appear to dominate public discourse, two deserve to be singled out for special attention in light of the theme of this paper. The first vision is that of the electronic agent as 'digital secretary' and/or 'digital butler'. This image depicts the agent as subservient and beneficent in relation to its users and indeed wider society. One of the

---

1 B.A.(Hons.), LL.B.(Hons.) (Australian National University); Dr. Juris / LL.D. (Oslo); Senior Research Fellow at the Norwegian Research Centre for Computers and Law, University of Oslo; Barrister of the Supreme Court of New South Wales. Thanks go to Graham Greenleaf for helpful comments on an earlier draft of this paper. This paper has been written in the framework of the ECLIP II project (Electronic Commerce Legal Issues Platform) funded by the European Commission under the specific programme for research, technological development and demonstration on a user-friendly information society (the IST programme). This paper, though, is solely the responsibility of the author and does not represent the opinion of the other contributors to ECLIP or of the European Community, nor is the European Community responsible for any use that might be made of data appearing in this paper. General information on ECLIP is available at http://www.eclip.org/.

2 Morgan, *Images of Organization* (Sage: London 1986). Morgan identifies a range of such images: e.g. organisations as machines, organisms, brains, cultures, political systems, pyschic prisons, processes of flux and transformation, and instruments of domination.

3 By 'electronic agent' is meant a software application which, with some degree of autonomy, mobility and learning capacity, executes specific tasks for a computer user or computer system. See further the paper by Weitzenboeck in this volume. Cf. commentary *infra* n. 9.

most popular and influential sources of this image is Nicholas Negroponte's book, *Being Digital*, which waxes lyrical about the helpful abilities of electronic agents.[4] From a privacy perspective, it is noteworthy that Negroponte's depiction of agents as 'digital butlers' and 'digital sisters-in-law' highlights the fact that such agents will need to carry a great deal of information about their human masters. In the words of Negroponte:

> 'the concept of 'agent' embodied in humans helping humans is often one where expertise is ... mixed with knowledge of you. A good travel agent blends knowledge about hotels and restaurants with knowledge about you (which often is culled from what you thought about other hotels and restaurants).[...] Now imagine a telephone-answering agent, a news agent, or an electronic-mail-managing agent. What they all have in common is the ability to model you'.[5]

Negroponte adds that we are entering into an age of 'true personalization' that is beyond demographics and statistical analysis. This age is characterised by the growing acquaintance by machines as software with human beings as individuals. It is about 'machines understanding individuals with the same degree of subtlety (or more than) we can expect from other human beings, including idiosyncracies (like always wearing a blue-striped shirt) and totally random events, good and bad, in the unfolding narrative of our lives'.[6]

Negroponte characteristically fails to give any detailed consideration to the significant privacy problems that may flow from the acquisition and application of such knowledge. From a privacy perspective, the vision he helps to create of electronic agents is largely sugar-sweet.

The same cannot be said of the other image of electronic agents which is of central importance here. This image is of the electronic agent as a runaway 'big brother': omniscient, omnipotent, malevolent and beyond human control. Probably the most influential depiction of the electronic agent along these lines is in the sci-fi story, '2001: A Space Odyssey', written by Arthur C. Clarke and made into a film by Stanley Kubrick. The story features a computer named HAL (standing for 'Heuristically programmed Algorithmic computer') which controls the operations of a spacecraft in an increasingly paranoid and (from a human perspective) malevolent manner. HAL keeps the spacecraft's human crew under intense surveillance, and manages to knock off most of them before being effectively lobotomised by the last remaining crew member.

I suggest that if and when public concern about the privacy-invasive capabilities of electronic agents becomes widespread, HAL will probably replace the Big Brother of Orwell's *Nineteen Eighty-Four* as the galvanising image of fear. This would be somewhat unfortunate as HAL – like Orwell's Big Brother – represents an extreme that is unlikely to be realised, at least in the near future. Just as unfortunate, though, would be a situation in which public policy about electronic agents is dominated by the relatively Panglossian vision perpetuated by Negroponte. For this vision is equally misleading.

If we are to channel along sensible lines public discourse about the impact of electronic agents on

---

4 Negroponte, *Being Digital* (Hodder & Stoughton: London 1995), especially pp. 149 *et seq*.
5 *Ibid*, p. 155.
6 *Ibid*, p. 165.

privacy and related interests, it is necessary to analyse the various functions and operations of electronic agents relatively independently of the above-described images.[7] The remainder of this paper constitutes a preliminary attempt to provide such an analysis.

2. Agent functions and types

Electronic agents can be differentiated according to the functions they are intended to perform. The chief functional categories of agents are as follows:[8]

- *Filtering agents* – i.e. agents which sift through information directed at their users, letting through only information of interest or relevance to the users or allowing the latter to make an informed choice as to which information is of interest/relevance to them;
- *Search agents* – i.e. agents which seek out information on behalf of their users;
- *User interface agents* – i.e. agents which monitor the interaction of their users with information systems and regulate the interaction in ways that help the users;
- *Broker agents* – i.e. agents which mediate between buyers and vendors of certain products or services;
- *Office/work-flow agents* – i.e. agents which automate basic office tasks;
- *System management agents* – i.e. agents which manage the operations of an information system; and
- *Problem-solving agents* – i.e. agents that function as expert systems for resolving, or helping to resolve, relatively complex issues.

These categories overlap to some extent and agents can combine, in practice, at least several of the above-listed functions.[9] In terms of privacy, the most important point to note here is that the efficient operation of several of these agent types (most notably filtering, search and broker agents) is dependant upon the agent storing information about its user(s) and/or gleaning information about other agents, persons and organisations with which it interacts.

Electronic agents can also be differentiated according to other factors. For instance, they can be differentiated in terms of:

---

[7] I write 'relatively' independently because, as intimated above, it is impossible to be quit such images. In this regard, Curtis Karnow aptly notes the existence of 'a nervous humor in our perceptions of the capabilities and risks of artificial intelligence'. Concomitantly, treatment of these capabilities and risks will tend to 'tread a thin line between anxiety and analysis'. See Karnow, *Future Codes: Essays in Advanced Computer Technology and the Law* (Artech House: Boston/London 1997), p. 108.

[8] For further detail, see e.g. Woolridge & Jennings, 'Intelligent Agents: Theory and Practice' (1995) 10 *The Knowledge Engineering Review*, no. 2, pp. 115–152; Bigus, *Data Mining with Neural Networks: Solving Business Problems from Application Development to Decision Support* (McGraw-Hill: New York 1996), chapter 8. See also the paper by Sonia Gonzalo in this volume.

[9] The degree of agent multi-functionality will depend partly on the level of abstraction and systemic complexity at which the notion of agent is applied. A point of departure of this paper (together with the other papers in this volume) is that an electronic agent is a relatively discrete software module operating within, and partly constitutive of, an information system. The latter, however, might be viewed as an agent too. Indeed, from a practical-regulatory perspective, it might well be unwise to draw hard and fast lines between notions such as 'agent', 'information system' and 'intelligent processing environment' (the latter notion is used in Karnow, *supra* n. 7, pp. 168 *et seq.*). See further section 5 of this paper.

- *mobility* – i.e. the extent to which they can move from one information system to another;
- *interactivity* – i.e. the degree of sophistication with which they can communicate with each other and their users;
- *intelligence* – i.e. their ability to learn and reason;
- *autonomy* – i.e. the extent to which they are able to act independently of their creators, owner or users;
- *agency, control and ownership* – i.e. the persons/organisations who/which own them, control their operations or for whom/which they act; and
- *identifiability* – i.e. the degree to which they can be singularly distinguished from other agents.

The last factor is of obvious significance for privacy. The extent to which agents can be distinguished from other agents tends to encompass the extent to which they can be readily linked to the person(s) or organisation(s) who/which create, own or use them.

As for the factor concerned with agency, ownership and control, an important point is that agents will often act for a multiplicity of parties. For example, an agent may be created and owned by one person/organisation (typically termed an Agent Service Provider or 'ASP') but go into service for other persons/organisations. Also important is whether an agent acts for an organisation or an individual person. As shown further below, this bears on the extent to which data relating to an agent may be safeguarded under data protection law.

3. Privacy risks

3.1 General description

What are the various ways in which use of electronic agents can threaten privacy and related interests?[10] To answer this question, we need first to define what is meant by 'privacy' and 'related interests'. Here, 'privacy' is used to denote a condition or state in which a person (or organisation) is more or less inaccessible to others, either on the spatial, psychological or informational plane. In short, privacy denotes a state of limited accessibility. The notion of 'related interests' denotes interests which privacy helps to promote and which help to promote privacy. The principal such interests are autonomy (i.e. self-determination), integrity (i.e. a person's state of intact, harmonious functionality based on other persons' respect for him/her) and dignity (i.e. a person's intrinsic worth).[11]

At the outset, it should be emphasised that agents play a multifarious role with respect to privacy and related interests. On the one hand, they can help to safeguard the privacy and related interests of their users. Filtering agents, for example, will often play such a role. On the other hand, agents

---

[10] For an alternative though complementary presentation of these threats, see Borking, van Eck & Siepel, *Intelligent Software Agents: Turning a Privacy Threat into a Privacy Protector* (Registratiekamer: The Hague 1999), chapter 4.

[11] For further analysis of privacy and related interests, see Bygrave, *Data Protection Law: Approaching Its Rationale, Logic and Limits* (Kluwer Law International: The Hague/London/New York 2002 – forthcoming), chapters 1 and 7.

can render their users more vulnerable to privacy loss. This can occur when agents are tapped for information about their users. Obviously, too, agents can be actively employed as privacy-invasive instruments – i.e. when they are used to monitor and/or contact other persons/organisations.

The privacy risks associated with agent operations are usefully illuminated if we consider profiling practices. By 'profiling' is meant here the process of inferring a set of characteristics (typically behavioural) about an individual person or collective entity and then treating that person or entity (or other persons or entities) in the light of these characteristics. It will be readily seen from Negroponte's commentary,[12] that agents can embody user profiles; the way in which agents are programmed to operate will often reflect the preferences and needs of the persons/organisations for whom/which they act. Thus, monitoring of agent operations might reveal a considerable amount of information about their users. The tapping of such information can result in even greater privacy loss for those persons/organisations. The reverse side of this privacy loss is the generation or augmentation of user profiles which can subsequently be applied by other persons/organisations in various ways (e.g. direct marketing) that exacerbate the initial privacy loss. The latter tasks can be facilitated, of course, by other agents.

Complicating matters, and adding to the privacy risks of agent operations, is that agent operations can be covert and/or non-transparent. For example, an agent can undertake surveillance tasks (on behalf of its user) without the knowledge of the surveillance subjects. This would compromise the privacy and related interests of the latter though not those of the agent user. At the same time, though, an agent can pass on information about its user to others without the user being aware of the information transfer. For example, an agent can go into service for the user of a particular website or computer but covertly provide information about the user to the agent service provider (ASP).[13] In such a situation, the agent would be acting, in effect, somewhat similarly to a 'Trojan Horse'.

3.2 Privacy risk scenarios

The privacy risks described above may be elucidated more clearly by considering the following five scenarios. Each scenario is illustrated by a diagram. In each diagram, circles represent users of agents while squares represent agents. The arrows represent information flows.

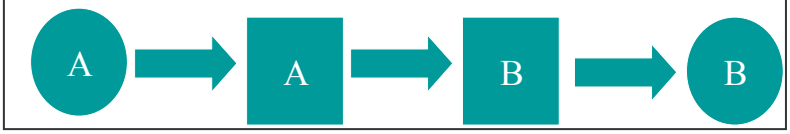
---

[12] See *supra* n. 4 and accompanying text.

[13] As allegedly occurred with the agent, Registration Wizard, which Microsoft included in its Windows 95 operating system: see Karnow, *supra* n. 7, pp. 155–156 and references cited therein.

3.2.1 Scenario 1

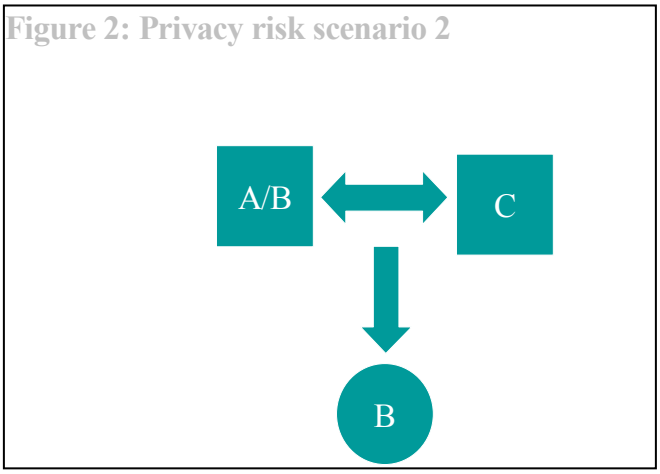
Figure 1: Privacy risk scenario 1



This scenario is the most basic. It involves a person or organisation (A) employing an agent to interact with another person/organisation (B) either directly or via B's agent. The main privacy risk for A is that information about him/her/it flows to B. The extent of damage to A of such information flow will depend on several factors, such as the degree of sensitivity of the information, the extent to which the information flow occurs without A's knowledge and/or consent, and the purposes for which B applies the information.

3.2.2 Scenario 2

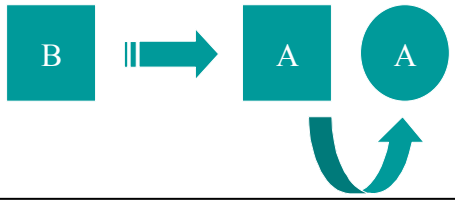
Figure 2: Privacy risk scenario 2



In this scenario, A interacts with C (via C's agent) using an agent that is provided by B. In addition to the risk of information about A flowing to C, there is the risk of information about A flowing to B. Again, the degree of damage to A of such information flow will depend on the same sorts of factors as described in connection with scenario 1.

3.2.3 Scenario 3

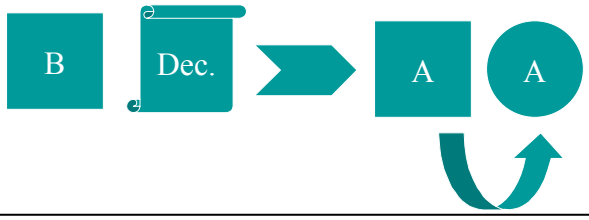
Figure 3: Privacy risk scenario 3



This scenario depicts a situation in which B sends information to A (via A's agent). The information will typically be in the form of advertising material. In other words, the scenario depicts a situation of direct marketing by B with respect to A. Such a situation interferes with A's privacy and autonomy (more particularly A's interest in 'attentional self-determination'; i.e. A's interest in being able to give his/her/its attention to what he/she/it wants).[14] The extent of interference will partly depend on whether the marketing is solicited or unsolicited.

3.2.4 Scenario 4

Figure 4: Privacy risk scenario 4



This scenario does not depict a situation in which A's privacy is directly at risk but it could involve detriment to certain of A's privacy-related interests (and indirectly to A's privacy). The scenario depicts a situation in which agent B makes a decision about agent A which has an effect on A.

[14] See Bygrave, *supra* n. 11, chapter 7 (section 7.2.5).

An example of such a situation occurs in the context of cybermarketing. The advertising banners on Internet websites are frequently programmed to automatically adjust their content and/or format according to the net-browsing data about the site visitor which are stored as 'cookies' on the visitor's computer.[15] This is a form of automated profiling. One possible effect of such practices is unfair discrimination in one or other form of 'weblining' (e.g. A or A's agent – as the website visitor – is offered products or services at a higher price than other, assumedly more valuable consumers have to pay, or A/A's agent is denied an opportunity of purchasing products/services that are made available to others).[16]

3.2.5 Scenario 5

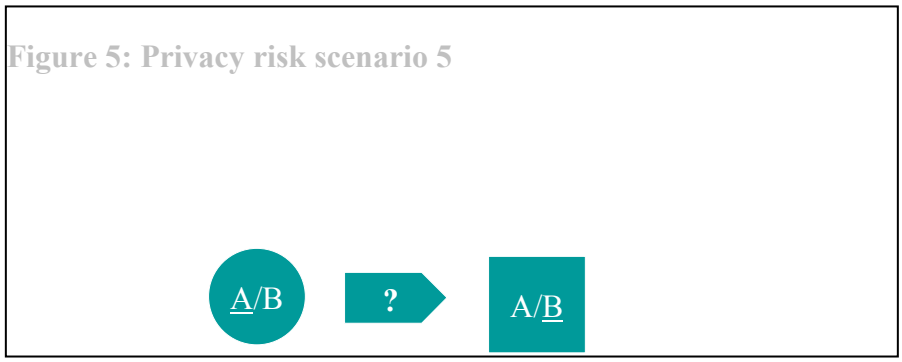


Figure 5: Privacy risk scenario 5

The final scenario concerns the transparency of agent operations: to what extent can A or B discern and comprehend the actions of their own agent together with the actions of the other's agent? Is each agent what it appears to be?

4. Legal issues:

4.1 Data protection provisions of potential relevance to privacy risk scenarios

A relatively large number of legal rules on privacy/data protection may apply to the situations depicted in section 3 so as to ameliorate the privacy risks for the agent user A. Here I shall take the provisions of the 1995 EC Directive on data protection[17] as the legal point of departure for the following discussion. The intention is not to conduct in-depth legal analysis but to flag important issues that can be dealt with in greater detail subsequently.

---

[15] See e.g. US Federal Trade Commission (FTC), *Online Profiling: A Report to Congress*, June 2000, available at http://www.ftc.gov/os/2000/06/onlineprofilingreportjune2000.pdf (last visited 10.8.2001), pp. 2–8.

[16] Further on weblining, see Stepanek, 'Weblining: Companies are using your personal data to limit your choices – and force you to pay more for products', *Business Week Online*, 3.4.2000, available at http://www.businessweek.com/2000/00_14/b3675027.htm (last visited 10.8.2001).

[17] Directive 95/46/EC of the European Parliament and of the Council of 24.10.1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (O.J. L 281, 23.11.1995, pp. 31 *et seq*.) – hereinafter also termed simply 'Directive'.

With respect to privacy risk scenarios 1 and 2 (figures 1 and 2), centrally relevant provisions are Arts. 6–8, 10–12 and 17 of the Directive. These provisions set out basic conditions for the processing of personal data. They place significant restrictions on the non-consensual, covert processing of such data and on the ability to use these data for secondary purposes that are incompatible with the primary purposes for which the data are processed (see Arts. 6–8, 10–12). They also require the taking of security measures for ensuring that personal data are protected from accidental and unlawful destruction, alteration or disclosure (see Art. 17).

The application of these provisions will depend on the information that flows from A to B being classified as personal data pursuant to Art. 2(a) of the Directive. The issue of what constitutes personal data is dealt with further below.

With respect to privacy risk scenario 3 (figure 3), the centrally relevant provision is Art. 14(b) of the Directive which provides persons with a right to object to direct marketing.

As for privacy risk scenario 4 (figure 4), Art. 15(1) is of chief importance. Article 15(1) grants persons a qualified right not to be subject to certain forms of fully automated decision making.[18]

Finally, regarding privacy risk scenario 5 (figure 5), the most pertinent provisions are Arts. 10–12. In summary, Arts. 10 and 11 require data controllers to directly supply data subjects with basic information about the parameters of their data-processing operations, independently of the data subjects' use of own information access rights. Article 12 provides persons a right of access to data kept on them by other persons and organisations. This right of access is not just to data relating directly to them but also to information about the way in which the data are used, including the purposes of the processing, the recipients and sources of the data, and 'the logic involved in any automated processing of data concerning [the data subject] … at least in the case of the automated decisions referred to in Article 15(1)'. The latter component of this access right is of special significance in the context of agent operations.

At the same time, the right of access to logic is qualified by recital 41 in the preamble to the Directive which states that the right 'must not adversely affect trade secrets or intellectual property and in particular the copyright protecting the software'. Yet the recital also states that 'these considerations must not … result in the data subject being refused all information'. In other words, copyright is not permitted to completely pull the blinds on agent transparency. It remains to be seen just how serious the potential for conflict is between the access-to-logic right and intellectual property rights attached to decision-making software. I am inclined to predict that the conflict potential is small, at least in the short term. This is in light of two factors. The first is that information access rights under data protection laws hitherto have often been little used – at least in some jurisdictions.[19] The other factor is that the access-to-logic right will tend to be limited to the type of fully-automated decision making dealt with in Art 15 of the Directive. This type of decision making is narrowly defined; concomitantly, its occurrence is relatively rare.[20]

---

[18] See further section 4.2.1.

[19] With respect to under-usage of access rights in Denmark, see e.g. Blume, 'How to Control Data Protection Rules' (1992) 6 *International Computer Law Adviser*, no. 6, pp. 17, 18. With respect to under-usage of such rights in Norway, see generally the annual reports of the Norwegian Data Inspectorate (*Datatilsynet*).

[20] See further Bygrave, 'Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling' (2000) 7 *Privacy Law & Policy Reporter*, pp. 67–76.

However, it might well become more widespread in the long term.

4.2 Applicability of data protection rules to agent operations

A range of threshold questions arise concerning the applicability of the rules set out in section 4.1 above to agent operations. Again, my point of departure when canvassing these questions is the 1995 EC Directive on data protection.

4.2.1 Agents as users of personal data

The first set of questions concern the extent to which the rules of the Directive may apply to electronic agents in their capacity as processors and users of data on persons other than those for whom they act. Three questions arise in this respect:
(1) can agents be data controllers?;
(2) can agents be data processors?; and
(3) can agents be decision makers for the purposes of Art. 15 of the Directive?

The first of these questions can be dealt with summarily. Although a software agent could fall within the literal ambit of the definition of 'controller' in the Directive,[21] a controller is a body to which certain legal liabilities must directly attach. To my knowledge, electronic agents cannot be sued (or sue) under any European legal system.[22]

As for the second-listed question, again a software agent could fall within the literal ambit of the definition of 'processor' in the Directive.[23] Moreover, since the Directive does not place primary liability for breach of data protection rules on processors but on controllers (see especially Art. 17), agents are not legally hindered from being processors in the same way as they are hindered from being controllers. However, some of the phrasing of the Directive can be interpreted as signalling an intention that processors be more than mere software.[24] Indeed, electronic agents seem to fall more naturally under the category of 'equipment' (see e.g. Art. 4(1)(c)) than of processor.

In any case, whether or not an electronic agent may be a processor appears to have relatively little legal consequence in relation to the Directive – at least for the duties and liabilities of controllers.

---

[21] Article 2(d) of the Directive defines a 'controller' as the 'natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data'. It could be argued that the phrase 'any other body' captures software agents. However, the phrase should probably be read down in light of the list of entities preceding it. Note also recital 12 ('Whereas the protection principles must apply to all processing of personal data by any person whose activities are governed by Community law …') and recital 20 ('Whereas the fact that the processing of personal data is carried out by a person established in a third country …').

[22] See also section 2.1 of the paper by Weitzenboeck in this volume.

[23] A 'processor' is defined in Art. 2(e) as 'a natural or legal person, public authority, agency or any other body which processes personal data on behalf of' a data controller.

[24] See e.g. Art. 16 ('Any *person* acting under the authority of the controller or of the processor, including the processor *himself* …': emphasis added). And, as with the definition of 'controller' in Art. 2(d), the phrase 'any other body' in Art. 2(e) should probably be read down in light of the list of entities preceding it.

If an agent may qualify as a processor, any controller engaging the agent will have to ensure – by way of contract or other legal act (Art. 17(3)) – that the agent provides 'sufficient guarantees in respect of the technical security measures and organizational security measures governing the processing to be carried out' (Art. 17(2)).[25] Yet even if software agents do not qualify as processors, a controller using such an agent will still have to take measures to ensure that the agent operations respect data protection rules, if the controller is to escape liability for damages caused by the agent acting in breach of those rules (see Art. 23).

Turning to the question of whether agents may be decision makers for the purposes of Art. 15, relatively solid grounds exist for a positive answer. Article 15(1) reads:

'Member States shall grant the right to every person not to be subject to a decision which produces legal effects concerning him or significantly affects him and which is based solely on automated processing of data intended to evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, etc.'

Does Art. 15(1) apply to situations in which a human (or corporate) decision maker is apparently absent; i.e. when the process at hand consists primarily of a response on the part of a software agent to particular constellations of data and data input? The issue is actualised by the cybermarketing practices described in section 3.2.4 above.

Supporting a negative answer is that decisions ordinarily involve the adoption of a particular opinion or belief. Hence, the term 'decision' ordinarily connotes a *mental* action. An affirmative answer, though, has stronger foundations. It can be plausibly argued that the term 'decision' should be construed broadly and somewhat loosely, particularly given the rationale for Art. 15(1) and the provision's otherwise detailed qualification of the type of decision it embraces.[26] At the same time, it can be plausibly argued that a human (or corporate) decision maker will still exist even if he/she/it is not directly involved in the process concerned. That decision maker will be the person (or corporation) responsible for applying the software.[27]

4.2.2 Agents in relation to data subjects

The next set of questions concern the extent to which the rules of the Directive may apply to agents in their capacity as data subjects or as entities acting on behalf of data subjects. Four main questions arise here:
(1) can agents be data subjects for the purposes of the Directive?;
(2) to what extent can agent data be personal data?;
(3) can data subject consent be carried out via agents?; and
(4) can controllers' information duties be fulfilled simply through informing agents?

---

[25] The latter requirements are supplemented by Art 16: 'Any person acting under the authority of the controller or ... processor, including the processor himself, who has access to personal data must not process them except on instructions from the controller, unless he is required to do so by law'.

[26] See further Bygrave, *supra* n. 20, pp. 68–71.

[27] While this argument is plausible for current software processes, we should not overlook the future possibility of electronic agents becoming so autonomous in their actions and learning capabilities that it is logically (though not necessarily legally) difficult to link their behaviour with any particular human(s) or corporation(s).

The answer to the first-listed question is negative. Article 2(a) of the Directive defines a 'data subject' as 'an identified or identifiable natural person'.

The second-listed question is more difficult to answer, particularly in the abstract. The starting point for discussion is Article 2(a) of the Directive which defines 'personal data' as

'any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity'.

While this definition appears at first blush simple to construe and apply, it harbours on closer examination numerous ambiguities rendering its interpretation and application less than straightforward. These ambiguities can be summed up in terms of the following questions:

- how easily or practicably must an individual be linked to information in order for the latter to be regarded as 'personal'?
- who is the legally relevant agent of linkage and identification (i.e. who is to carry out linkage and identification)?
- is information 'personal' if it only facilitates the identification of a person in combination with other (auxiliary) information?
- to what extent must data be linkable to just *one* person in order to be 'personal'?

When determining the requisite degree of ease/practicability of linkage, regard is to be had to 'all the means likely reasonably to be used either by the controller or by any other person to identify the said person' (recital 26 in the preamble to the Directive).

Closely related to the question of ease/probability of linkage/identification is the question of who is the legally relevant agent of linkage/identification. Recital 26 of the Directive indicates that *any* person may be the legally relevant agent for identification/linkage. In other words, what is legally decisive is not just the ability of the data controller to link a person to the data, it is any person's ability to do so. This lowers the threshold for determining the circumstances under which data are personal. At the same time, though, the criteria for ease/practicability of linkage/identification exclude from consideration any persons who do not employ means that are reasonably capable of being used for identification. The notion of reasonableness implies, in turn, that account ordinarily should not be taken of persons who are only able to carry out identification by *illegal* means (e.g. computer hacking). Given that the notion of reasonableness also connotes a probability criterion, account should also not be taken of persons who are only able to carry out identification by (objectively) *unexpected* or *unusual* means. In most cases, illegal means will be unexpected or unusual means.

Regarding use of auxiliary information to identify a person, it is fairly clear that the Directive permits this. In other words, when information is indirectly traceable to individual persons (i.e. through combination with other information), it may qualify as 'personal'.

The final issue concerns the extent to which data must allow for individuation; i.e. be linkable to *one* person as opposed to an aggregate of persons. The Directive requires that data be capable of linkage to a *particular individual* person if they are to qualify as personal. Thus, data which are linked to an aggregate of persons (including an organisation) and which do not allow for these persons' individuation will normally fall outside the ambit of the Directive.[28]

In light of the above, the extent to which data relating primarily to an electronic agent may amount to personal data under the Directive is a question of fact that cannot be answered conclusively in the abstract. Nevertheless, there is nothing on the face of the Directive to suggest that data may not qualify as personal merely because they are also linked to software. Moreover, the broad and flexible criterion for linkage/identifiability laid down in recital 26 means that the threshold for what amounts to personal data is relatively low. If readily available means exist for linking agent data to a particular individual person, these data may most likely be regarded as personal for the purposes of the Directive.[29] Concomitantly, agent data that can be linked to an organisation (e.g. corporation) will not qualify as personal under the Directive unless the data can also be linked to a specific individual.[30] The latter possibility might arise with respect to agent data connected to small companies.[31]

Turning to the issue of consent, some of the rules of the Directive stipulate as a basic (albeit alternative) condition for the processing of personal data that the processing be consented to by the data subject (see especially Arts. 7(a) and 8(2)(a)). May this consent be carried out via electronic agents? There is nothing in the Directive to suggest that agents may not ever be used as media for consent. However, such consent will only be valid if it manifests itself in accordance with the Directive's requirements. The basic requirement is that consent is 'any freely given specific and informed indication of … [the data subject's] wishes …' (Art. 2(h)). With respect to the processing of ordinary types of personal data, consent must be 'unambiguous' (Art. 7(a)); with respect to certain types of sensitive personal data, consent must also be 'explicit' (Art. 8(2)(a)).

Obviously, consent provided by electronic agents will be in electronic format. As such, it could run into difficulties in conforming with the requirements of 'unambiguity' or 'explicitness' but it *need* not fall foul of these. The same applies in relation to meeting the requirement that consent be 'informed'.

At the same time, agent operations raise an interesting question with respect to the latter requirement. May the requirement be met when the data subject uses an electronic agent that is programmed to engage in particular transactions but only on the basis that the transactions

---

[28] However, the data protection laws of a minority of European countries – including those of Italy, Austria and Switzerland – expressly cover data on corporations and other legal/juridical persons. For extensive analysis of the issue of data protection for legal persons, see generally Bygrave, *supra* n. 11, Part III.

[29] The issue here is basically the same as the issue of whether data linked primarily to e-mail addresses or machine addresses (i.e. Internet Protocol numbers) may also qualify as personal in particular circumstances. For discussion of the latter issue (and for elaboration of the points raised above in relation to the criteria for identifiability), see Bygrave, *supra* n. 11, chapter 18 (section 18.2); Greenleaf, 'Privacy principles – irrelevant to cyberspace?' (1996) 3 *Privacy Law & Policy Reporter*, pp. 114–115.

[30] Of course, though, agent data linked to a corporation will be covered by those (relatively few) national data protection statutes that expressly safeguard data on corporations.

[31] See further Bygrave, *supra* n. 11, chapter 10 (section 10.3).

conform to a set of privacy standards that have been pre-approved by the data subject?[32] In such a case, the data subject has, as it were, informed him-/herself prior to interaction (via an agent) with potential data controllers. While the drafters of the Directive probably did not envisage this sort of 'self-informing' when they stipulated the requirements on consent, there seems little good reason to treat it as incapable of fulfilling those requirements.

Directly related to the above is the broader duty of information which data controllers have with respect to data subjects, pursuant to Arts. 10–11 of the Directive. May this information duty be fulfilled simply by the controller informing the data subject's agent? Again, the answer is most probably affirmative, at least if the data controller may reasonably expect that the agent will send the information on to its user (the data subject).

A problem might arise when there is a significant time lag between the agent being given the information and the agent passing on the information to its user. Article 11 indicates that the information is, as a point of departure, to be communicated by the controller to the data subject at the time of initiating the data processing. Article 10 is less clear on this point but should probably be read in the same way as Art. 11.

5. A future agenda

How can we reduce the risks to privacy and related interests occasioned by the use of electronic agents? In the following, a somewhat tentative agenda is drawn up for tackling this issue.

The least tentative aspect of the agenda concerns future work on the technological/organisational plane, particularly in the field of systems development. We should develop various privacy-enhancing technologies (PETs) and integrate these in agent operations. Of crucial importance here are mechanisms for ensuring that agent operations are transparent and controllable for agent users. Also crucial are mechanisms for restricting the unintended flow of information about users, via their agents, to others. Thus, we need to build into agent systems a broad range of devices that facilitate, when appropriate, anonymity, pseudonymity, logging, auditing, authentication and access controls.

These proposals echo the important recommendations of the International Working Group on Data Protection in Telecommunications.[33] Fortunately, practical work on this front is already underway. The EU-financed PISA project (standing for 'Privacy Incorporated Software Agents') is an important pioneer in this regard.[34]

---

[32] This sort of transactional set-up is provided for by the Platform for Privacy Preferences (P3P) developed by the World Wide Web Consortium (W3C). See further http://www.w3.org/P3P/ (last visited 12.8.2001).

[33] See the 'Common Position on Intelligent Software Agents' adopted by the Working Group on 29.4.1999; available at http://ig.cs.tu-berlin.de/~dsb/doc/int/iwgdpt/agent_en.htm (last visited 10.8.2001). The recommendations in the Common Position are listed in the appendix to this paper.

[34] For a description of PISA, see Borking, 'Privacy incorporated software agents: a proposal for building a privacy guardian for the electronic age' (2000) 7 *Privacy Law & Policy Reporter*, pp. 91–96. See also the PISA website at http://www.tno.nl/instit/fel/pisa/index.html (last visited 10.8.2001).

Further on the technological/organisational plane, thought should be given to establishing reliable and comprehensive systems for the certification, labelling and registration of electronic agents along with ASPs. Calls for such systems have already been made.[35] I have some reservations, however, about this strategy. Certification systems can easily become cumbersome, expensive and slow. Another problem lies in finding trustworthy bodies to run the systems. Linked to that problem is the risk of multiple certification schemes emerging which apply differing standards and procedures, thus confusing the marketplace. Further, when these schemes are run as business ventures and compete with each other for custom, they might be tempted to cut corners in their setting and/or application of standards.

If we put aside these potential difficulties and embrace the need for certification of agents and ASPs, what would such certification ideally involve? First, in order to be certified, ASPs would need to program their agents so that the latter only transact with other certified agents/ASPs or at least warn their users when they transact with non-certified actors. Certification would also be contingent upon the agents being programmed in accordance with the PET mechanisms listed above. Concomitantly, certification would also be contingent upon the agents being programmed such that they follow the fair information practices laid down in the data protection Directive.

On the legal front, consideration ought to be given to drafting data protection rules which take full account of electronic agents. Without such rules, the above-listed developments in integrating PETs with agent operations might stagnate. The validity of this prognosis, though, is debatable. It could be argued that the present rules in data protection laws already contain sufficient incentives to generate and power the technological/organisational developments called for above. For instance, it is arguable that, taken together, the provisions of the data protection Directive – particularly Arts. 6–8 and 17 – require the establishment of mechanisms for anonymity, pseudonymity, logging and certification in relation to agent operations.

My own view, however, is to treat this argument with scepticism. Fuelling this scepticism is that the overwhelming majority of data protection laws – including the Directive – have been drafted with little, if any, specific consideration being given to electronic agents. Hence, their regulation of agent operations is incidental and/or accidental. Concomitantly, considerable uncertainty surrounds the exact way in which they apply to agents – as shown by the analysis in section 4 above. These characteristics are part and parcel of a larger problem which is that data protection laws still tend to lack sufficiently detailed rules concerning the quality and architecture of information systems.

The only provisions in the Directive which come close to taking account of agent operations are Arts. 15 (right to object to automated profiling) and 12(a) (right of access to logic behind automated profiling). While a relatively large number of countries will soon have enacted similar provisions – mainly as a result of the Directive – most of these countries will be European, at least for the near future. The extent to which non-European countries will enact such provisions remains unclear. The 'safe harbor' agreement which has been concluded recently between the USA and EU,[36] and which stipulates conditions for permitting the flow of personal data from the

---

[35] See e.g. the Common Position of the International Working Group on Data Protection in Telecommunications, *supra* n. 33; Karnow, *supra* n. 7, pp. 178 *et seq.*

[36] See Commission decision of 26 July 2000 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the safe harbour privacy principles and related frequently asked questions issued by the

EU to the USA, omits reference to the principles embodied in Arts. 15 and 12(a). Hence, other non-European countries will probably not be required by the EU to implement the principles either. Unfortunately, legislators in these countries so far have exhibited little willingness to implement the principles of their own accord.[37]

Quite apart from this problem is that the regulatory efficacy of Art. 15 is reduced by its complexity and numerous ambiguities in the way its provisions are formulated. These problems are exacerbated by a paucity of authoritative guidance on the provisions' scope and application. The efficacy of Art. 15 as a regulatory tool is further reduced by the fact that its application is contingent upon a large number of conditions being satisfied.[38]

However, we need to exercise caution before rushing into drafting data protection instruments with a greater number of agent-specific rules. One danger is that of 'legislating on a metaphor' (Bing); i.e. adopting terminology that does not do full justice to the realities of agent operations. This danger is exacerbated in a situation in which the architecture and parameters of agent operations are continuously changing. Accordingly, it is probably most sensible – at least for the present – to draft rules focusing more generally on the architecture of information systems than on the 'agent' concept. This approach is further justified by the fact that, as software, agents are integral components of information systems.

Nevertheless, in the interests of greater regulatory precision and prescriptive guidance, little harm will be done and much will be gained if we begin to think proactively about how best to draft more agent-specific data protection rules. We already have some models upon which to build. The most notable of these is the 'Common Position on Intelligent Software Agents' adopted by the International Working Group on Data Protection in Telecommunications.[39] Also instructive is Germany's Teleservices Data Protection Act of 1997, especially its innovative provisions dealing expressly with electronic consent and transactional anonymity/pseudonymity.[40]

Realistically, though, we can expect that legislators generally will only act to pass data protection rules that take better account of agent technology if:

(i) agents become – in the minds of consumers – indispensable for consumer engagement in electronic commerce; and
(ii) there is evidence suggesting that significant numbers of consumers refrain from engaging in electronic commerce because of perceived privacy risks from agent use.

We are still some way from this situation.

In the longer term, with evermore pervasive and powerful agent technologies, we will need to

---

US Department of Commerce (O.J. L 215, 25.8.2000, pp. 7 *et seq*.).

37 For instance, no specific provision has been made for the principles in the new federal data protection laws for the private sector in Canada and Australia. See further Canada's Personal Information Protection and Electronic Documents Act 2000 and Australia's Privacy Amendment (Private Sector) Act 2000.

38 See further Bygrave, *supra* n. 20.

39 See *supra* n. 33 and the appendix to this paper.

40 For an overview in German, see Engel-Flechsig, 'Teledienstedatenschutz' (1997) 21 *Datenschutz und Datensicherung*, pp. 8–16. For an overview in English, see Bygrave, 'Germany's Teleservices Data Protection Act' (1998) 5 *Privacy Law & Policy Reporter*, pp. 53–54.

consider the extent to which electronic agents themselves should be accorded privacy/data protection rights. This issue parallels the issue of data protection rights for corporate entities, though each issue requires separate analysis. Consideration of the issue of data protection rights for agents is inextricably linked, of course, to the broader question of the legal status of agents and the desirability of conferring upon them legal personality. Some studies have already have been made of these questions.[41] It is noteworthy that the most directly relevant of these studies for the theme of this paper concludes rather provocatively that electronic agents should be given at least three basic rights:[42]

(i) 'privacy' or 'the right to decline to produce information aside from key identification materials';
(ii) 'the right to be free of discrimination, to be able freely to conduct social and economic business'; and
(iii) 'free speech' or 'the right to communicate; to move about in electronic space and to post messages'.

The practical viability of such rights deserves further study.

Appendix

Recommendations from the 'Common Position on Intelligent Software Agents' adopted by the International Working Group on Data Protection in Telecommunications:[43]

> '1.Producers of software agents should reflect in an early stage of design on the implications of the use of intelligent agents for the privacy of individuals. This is necessary to control the consequences that may arise in the near future.
> 2.Developers of agents should ensure that users do not loose control over their systems and information contained therein. They should provide the user with the maximum of transparency on the functioning of the agent. Adding control and feedback mechanisms and safeguards to prevent this will help agent-users to increase trust in using agent technologies.
> 3.Developers of intelligent agents should ensure the proper means by which the privacy of users may be protected and control maintained by data subjects over the uses of their personal data. 4.Technical facilities such as Privacy Enhancing Technologies (PET) are recommended in conjunction with software agents. The following measures are proposed:
> - development of a Trusted Third Party structure for the identification and authentication of all agents;
> - access control mechanisms;
> - tools to give a user control over the actions of third parties' agents that collect personal data;

---

41 See e.g. Solum, 'Legal Personhood for Artificial Intelligences' (1992) 70 *North Carolina Law Review*, pp. 1231–1287; Stone, *Earth and Other Ethics: The Case for Moral Pluralism* (New York 1988), especially pp. 28–30; Karnow, *supra* n. 7, Part V.

42 Karnow, *supra* n. 7, p. 130. See also Karnow, 'The Encrypted Self: Fleshing out the Rights of Electronic Personalities' (1994) 13 *The John Marshall Journal of Computer & Information Law*, pp. 1, 12–13. Karnow intimates that these rights be given not out of concern for agents themselves (or what he terms 'electronic personalities'; 'epers' for short) but as a means of improving the protections for agent users.

43 See *supra* n. 33.

- mechanisms to audit the logged activities;
- integrity mechanisms to control the integrity of stored or exchanged data and to control the integrity of working methods of agents or trusted components, like digital signatures;

These measures can be integrated into the agents. The measures can also be used to build an infrastructure of trusted components.

5.By using a checklist of privacy-compliant design criteria, the designer, supplier, or provider of an agent should design or equip an agent or an agent-environment with proper privacy-enhancing technologies. A framework for certification of the privacy-compliance of software agents is required.'