2012

The Dawn of a Critical Transparency Right for the Profiling Era

Mireille Hildebrandt, *Radboud University Nijmegen*

Digital Enlightenment Yearbook 2012
J. Bus et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-057-4-41

The Dawn of a Critical Transparency Right for the Profiling Era

Mireille HILDEBRANDT*

Abstract. Potential consumers are increasingly profiled to detect their habits and preferences in order to provide for targeted services. Both industry and the European Commission are investing huge sums of money into what they call Ambient Intelligence and the creation of an 'Internet of Things'. Such intelligent networked environments will depend on real time monitoring and profiling, resulting in real time adaptations of the environment. In this contribution Mireille Hildebrandt will assess the threats and opportunities of such autonomic profiling in terms of its impact on individual autonomy and refined discrimination and indicate the extent to which traditional data protection is ineffective as regards profiling. She will then highlight the potential of the draft General Data Protection Regulation to provide a more adequate and effective level playing field for both the industry and individual citizens in the profiling era. The most revolutionary change she detects is not the right to be forgotten or the right to data portability but the right to be informed about the potential consequences of being profiled.

Keywords. profiling, KDD, transparency rights, TETs, democracy, the Rule of Law

Introduction: The Dawn of A New Transparency Right?

The draft General Data Protection Regulation that was released on 25th January 2012[1] contains a separate section on 'Right to Object and Profiling'. It builds on the existing right not to be subject of automated decisions, if such decisions have a significant impact and it transforms the associated transparency right to 'the logic of processing' that informs such decisions.[2] Art. 20 (4) of the proposal stipulates that the controller must provide 'information as to the existence of processing for a measure of the kind referred to in paragraph 1 [automated decisions, mh] and *the envisaged effects of such processing on the data subject* [my emphasis, mh].' Though the present 'right to the logic of processing' does not return in art. 20, the right to be informed of the existence of profiling and its consequences is both daring and much needed. It acknowledges the enormous impact of the computational underground that increasingly determines what we get to see and how we cognize. From the smart algorithms of major search engines to those underlying decision-making on creditworthiness, behavioural advertising, high

* Mireille Hildebrandt is Full Professor of Smart Environments, Data Protection and the Rule of Law at Radboud University Nijmegen, Associate Professor of Jurisprudence at the Erasmus School of Law, Rotterdam and senior researcher at the centre for Law Science Technology and Society at Vrije Universiteit Brussel. The article is a revised version of M. Hildebrandt, "Profiling and the Rule of Law," *Identity in the Information Society* (2008) available at http://www.springerlink.com/content/467887wtv826j6p4/fulltext.pdf.

[1] A good overview of the proposal is provided by Kuner [30]. He speaks of 'A Copernican Revolution in European Data Protection Law'.

[2] Art. 15 jo 12 (a) of the Data Protection Directive D 95/46/EC.

frequency trading and forensic expertise, profiling designates how we perceive the world and what we think we know. To some extend profiling has a much deeper and a more far reaching impact on who we become than identification technologies. In this contribution I will argue that the need to anticipate how profiling technologies categorize and pre-empt us is indeed more urgent and more defining of human agency than the need to prevent identification or to remain anonymous.

I will start with a brief indication of how profiling relates to identification (section 1), before discussing how sophisticated machine profiling differs from the kind of profiling we do in our everyday life (section 2). Next I will investigate profiling as *the* enabling technology for smart environments (section 3). Since profiling produces knowledge, rather than just data, section 5 will look into the threats posed by emerging knowledge-asymmetries due to the proliferation of profiling in smart environments. It is these threats that touch some of the fundamental tenets of democracy and rule of law, being the particular mélange of positive and negative freedom that allows citizens to develop their relative autonomy (section 4). To counter such threats the focus of legal scholars and practitioners should be extended from the protection of personal data to the protection against the undesired application of profiles and the creation of transparency rights regarding group profiles (section 6). Section 7 will argue a need for transparency by design and section 8 concludes with an appeal for cross-disciplinary collaboration to turn the proposed transparency obligation into an effective remedy.

1. Identification and Profiling

We live in the age of identification. For different reasons both government and business enterprise strive to develop effective tools for recurrent identification and authentication. Writers like Scott [49] in his '*Seeing Like a State'*, and Torpey [54] in his '*The Invention of the Passport. Surveillance, Citizenship and the State'*, have described and analyzed the insatiable need of the modern state for the registration of its citizens, originally seeking to attribute and implement tax obligations and register for subscription in the national army. The construction of the territorial nation state required the identification of those aligned to the territory and the nation. As a precondition for taxation and subscription it helped in creating the historical artefact of the territorial nation state. Like the introduction of national languages, national clocks and national currencies, the identification of citizens versus non-citizens in fact was a productive process, not merely the recording of a given fact. Building on the 18th century police state the 19th and 20th century welfare states then claim a need to differentiate between those that are entitled to public benefits and those that have no such right. Also, welfare states claim a pressing social need for identification to prevent fraud, crime and unlawful access in general, and in order to attribute liability, whether criminal or tort. E-government and e-health that aim to provide targeted services reiterate this quest for identification, though in this case the identification needed is more sophisticated and resembles what business undertakings seek when they develop targeted servicing and reinvent customer relationship management (CRM).

Business enterprise is less interested in a full proof registration of the inhabitants of a territory. Its focus is on acquiring relevant data about as many (potential) customers as possible as part of their marketing and sales strategies. As customer loyalty can no longer be taken for granted, companies develop CRM in the hope of surviving the competitive arena of neo-liberal market economies. At the same time they

try to establish which consumers may be persuaded to become their new customers and under what conditions. It seems that they are less interested in unique identification of any particular customer then in a refined type of categorisation that allows them to provide targeted servicing at the right time and in the right place. *Context is all* is not just the key message of adherents to cultural theory. In fact, companies are not just after the attributes of predefined classes of (potential) customers, but would rather invest in finding out which classes they should distinguish in the first place. This is where profiling comes in.

Profiling is as old as life itself. Indeed one could say that the difference between living and lifeless material is the fact that living organisms are capable of self-constitution over and against an environment which is constituted as such by the act of self-constitution [33]. In more simple terms: an organism and its environment co-create each other. Profiling is thus a crucial sign of life, because it consists of a reiterated identification of risks and opportunities by an organism in its environment [21]. Profiling is the interplay between *monitoring* and *adaptation*: to survive and to celebrate life any organism must continuously adapt itself to changes in its surroundings, while it may also manage to adapt its surroundings to its own preferences [33]. Monitoring one's context in this sense is a matter of *pattern recognition*, of discriminating noise from information. Not all data are relevant or valid, and whether this is the case will depend on the context and on the moment. Adequate profiling is always dynamic and caught up in the loop of recognizing a pattern (constructing the profile) and testing its salience (applying the profile). Interestingly enough, such organic profiling is not dependent on conscious reflection. One could call it a cognitive capacity of all living organisms, without thereby claiming consciousness for an amoebe. One could also call it a form of intelligence based on the capacity to adapt: monitoring and testing, subsequent adaptation and reiterated checking is what allows the living to flourish. This is what enables organisms to maintain their identity in the course of time, detecting opportunities to grow and proliferate as well as risks that need to be acted upon.

We may conclude that profiling is not typically human though we have developed our own brand of profiling. Cognitive psychologists speak of stereotyping, which allows us a measure of cognitive economy, as Schauer [46] has saliently argued in his *Profiles, Probabilities and Stereotypes*. What is special about humans is our capacity to reflect upon the profiles our brains come up with. This is a rare capacity, closely related to consciousness and language, and I shall not explore this domain much further, leaving it at the nexus of neurosciences and philosophy of mind ([19], [39], [22]). What matters is our capacity for conscious reflection on the profiles that we have unconsciously generated, because this gives us the freedom to deliberate on them, to reject or to reinforce them and to deliberately apply them. As Rouvroy [45] saliently describes this is what allows our self-formation. It is the precondition for our actions to be qualified as stemming from the freedom to act: we can become aware of the patterns that regulate our actions and review them to change our habits. Though most of our interactions are automated, handled autonomically by the habits that are inscribed in our body and brains, we can bring them to mind and scrutinize their relevance, validity, fairness and justice. This is what turns us into autonomous agents, capable of making a conscious choice for a course of action, deciding by which law to live. Autonomous derives from the Greek *auto nomos*: self and law. To some extent we are indeed capable of living by our own law, and this is also why we can be held accountable for our own actions ([21], [22]).

2. What *is* new? Profiling machines

Automated profiling is new in three ways. First, we are not talking about profiling by organisms but about profiling by machines [16]. Basically these machines are software programs 'trained' to recover unexpected correlations in masses of data aggregated in large databases. Second, we are not talking about making queries in databases, summing up the attributes of predefined categories, but about discovering knowledge we did not know to be 'hidden' in the data ([59], [11]). Third, at this point in time we cannot reflect upon the way profiling impacts our actions, because we basically have no access to the way they are produced and used. This last difference suggests that profiling hampers our freedom to act autonomously, a point I will return to below.

Automated profiling can be described as the process of knowledge discovery in databases (KDD, [17]) or machine learning [36]. KDD is generally thought to consist of the following steps:

1. recording of data
2. aggregation and tracking of data
3. identification of patterns in data (data mining)
4. interpretation of the outcome
5. monitoring data to check the outcome (testing)
6. applying the profiles.

Only the third step is what is called data mining in the sense of using computational algorithms to locate correlations, clusters, association rules and other patterns. An example of such profiling, using genetic algorithms, is driver fatigue detection by Jin et al. [27]. This type of profiling is also called behavioural biometric profiling (BBP) and uses a combination of pupil shape, eye movement frequency and yawn frequency to check tiredness in a driver. The data are mined by means of a feed-forward neural network and a back-propagation learning algorithm. To be honest we must note that BBP is still in an early stage of development, even though some results are highly interesting.[3] Both Zarsky [59] and Custers [11] emphasize that the knowledge generated by profiling machines is new. Zarsky speaks of data mining as 'answering questions users did not know to ask' ([59]: 4). He especially focuses on the difference between classification based on predefined classes and data mining techniques, which provoke unexpected clusters. Custers ([11]:56-58) argues that this type of knowledge is new in comparison with traditional social science, which starts with a hypothesis concerning a population that is tested by applying it to a sample. He points out that, first of all, in the case of KDD the hypothesis emerges in the process of data mining and, second, the hypothesis is tested on data that resemble a population rather than a sample. Some provocative scientists have even declared 'the end of theory' for this reason [4], suggesting that correlations are more important than causation in the era of data analytics.[4] Custers also indicates that when trivial information turns out to correlate with sensitive information, an insurance company or an employer may use the trivial information to exclude a person without this being evident as unfair discrimination (called *masking*). His last point is that the recording of data by means of

[3] See e.g. BBP for aggression detection by means of monitoring of sound, at http://www.soundintel.com/index-en.html. For an overview of behavioral biometrics profiling see [58].

[4] This is both farfetched and premature, but I would claim that both science and the humanities are changing more rapidly and more profoundly than some may suspect. See e.g. [29], [7] and [15].

ICT makes it practically impossible to delete records, especially where they are often shared across different contexts.[5] KDD can thus trace and track correlations in a growing mass of retained data and confront us with inferences drawn from past behaviour that would otherwise be lost to oblivion ([50], [34]). This raises a number of questions in relation to privacy and security, especially with regard to the effectiveness of data protection legislation. Before moving into these anticipated threats I will first describe emerging smart environments such as Ambient Intelligence (AmI) and the Internet of Things (IoT), to explain why autonomic machine profiling will make an increasing difference to our lives.

3. Smart Environments

In this chapter the concept of smart environments refers to hybrid online and offline environments that anticipate their inhabitants (usually referred to as users). I will briefly discuss the vision of Ambient Intelligence (AmI) and the Internet of Things (IoT) as typical scenarios of our technological futures. Smart environments are based on machine leaning, which refers to the discipline that seeks to build computer systems that automatically improve their performance with experience [36].

Both the European Commission [25] and, for instance, Philips [1], have invested heavily in the vision of Ambient Intelligence (AmI), vaguely defined by its 'key elements' ([1]:14): (1) embeddedness, meaning that networked devices are integrated into the environment; (2) context-awareness, since these devices can recognize you and your situational context; (3) personalisation, as they can be tailored towards your needs; (4) adaptiveness, meaning that they may change the environment in response to your behaviours; and (5) anticipation, since they should anticipate your preferences without your deliberate input, the environment will always be one step ahead of you. Related aspects that are often mentioned in the context of AmI are its hidden complexity, the absence of keyboards or monitors, the fact that the environment itself becomes the interface, real time monitoring and ubiquitous and proactive computing. The enabling technologies of this smart environment are sensor technologies, RFID systems, nanotechnology and miniaturization. Together they create *The Internet of Things* [26], which is supposed to turn the offline world online. The IoT consists of things that are tagged and permanently observed while communicating their data through the network that connects them. We must keep in mind, though, that most of these technologies only generate an enormous amount of data, which may not reveal any knowledge until profiling technologies are applied. We may conclude that profiling technologies are the crucial link between an *overdose of trivial data* about our movements, temperature, and interaction with other people or things and *applicable knowledge* about our habits, preferences and the state of the environment. Only after running data mining techniques through the interconnected databases can the things in our environment become smart things and start acting like agents in a multi-agent network (MAS). Profiling thus creates added value in the mass of data, deciding what is noise and what is information.

The vision of AmI depends on a seamless adjustment of the environment to our inferred habits and preferences [13]. The idea is that we need *not* provide deliberate

[5] The draft General Data Protection Regulation devotes a new article to the so-called Right to be forgotten and to erasure (art. 17). On this see [44].

input, but are 'read' by the environment that monitors our behaviour. This presumes what Tennenhouse [53] has described as proactive instead of interactive computing, diminishing human intervention as far as possible. The idea was that to seamlessly adapt the environment we cannot afford to wait for a human interpreter but need profiling machines that draw their own conclusions about what we prefer when and where, hoping to thus solve the problem of endless choice and deliberation. In the mean time even the initiators of AmI have acknowledged the importance of moving away from proactive to interactive smart environments, providing users with relevant feed-back and enabling them to steer the inferences on which smart environments are built [2]. This requires novel types of human-machine interfacing or, better still, novel approaches in human-machine interaction [43][6]. It may be interesting to note that the draft General Data Protection Regulation in art. 20 (2a) stipulates that in the case of a contractual relationship a person may only be subjected to a measure based on automated profiling where the data subject's legitimate interests have been safeguarded, 'such as the right to obtain human intervention'. When decisions affect a person's specific opportunities and more generally her capabilities, transparent human-machine interaction will at some point require human intervention. This should clarify how that person has been or will be categorized assessed and targeted, thus enabling that person to adjust her behaviour or to contest the categorization itself.

4. Democracy and the Rule of Law

Before describing the threats afforded by the socio-technical infrastructure of AmI and the Internet of Things,[7] we need to decide on what kind of threats we wish to detect. In this contribution the focus is not only on threats to individual consumers or tax payers, but also on potential threats to the socio-legal and political framework of democracy and the rule of law. This framework is a historical artefact, providing the constitutional instruments for individual citizens to counter threats to their rights and liberties. To make sense of potential threats against democracy and the rule of law, I will first discuss how these terms are to be understood in relation to profiling.

A sustainable democracy presumes and maintains the rule of law. The rule of law is often defined in reference to the protection of human rights and limited government. With regard to the implications of profiling technologies the most relevant achievement of the rule of law seems to be the mix of what Berlin [6] has coined as negative and positive freedom [21]. Positive freedom – *freedom to* – regards the freedom to participate in public decision-making processes or the freedom to achieve one's personal objectives; negative freedom or liberty – *freedom from* – regards the absence of unreasonable constraints imposed on a person. Positive freedom has a long history, while negative freedom – as a value of liberal democracy – is a relatively recent invention. Stalder [51] in fact suggests that privacy, with its emphasis on negative freedom, is an affordance of the era of the printing press. To nourish a sustainable

[6] The term interface is often deemed too technical, indicating that solutions to difficulties in human-machine interaction cannot be solved by only looking into the technical interface.

[7] In speaking of the affordances of the socio-technical infrastructure I refer to the work of ecological psychologist Gibson [18], see also ([43: 101-103). The idea is that specific technologies have specific affordances for different groups of people, depending on how they enlarge and/or restrict interaction. One could link this to Sen's capability theory [42], by demonstrating how different affordances result in different capabilities, potentially generating far reaching consequences for the exercise of human rights.

democracy we need both types of freedom, as embodied in the rule of law [12]. For this reason privacy is not just a private interest but also a public good. The rule of law establishes constitutional protection of citizens' rights and liberties over and against their government, safeguarded by an independent judiciary that shares the authority of the state. This is called the paradox of the *Rechtsstaat*: the state lends its authority to the courts that permit citizens to contest state authority.

Profiling can endanger both negative and positive freedom. Negative freedom is often equated with opacity, retreat to a private space, the right to oblivion and invisibility to the public eye. It refers to a space and time to regain one's strength, to reflect upon one's objectives and opinions. This negative freedom is matched with a need to act, to anticipate and participate in the public space, for which a measure of transparency is needed. Without transparency one cannot anticipate nor take adequate action. In fact I would claim that negative freedom is an illusion as long as transparency is absent; whereas we may assess reality in the privacy of our thoughts, to the extent that we lack access to the knowledge required to probe the 'reality of our reality' the formation of our will power is steered by what we cannot assess. Thus widespread usage of profiling technologies may endanger the intricate combination of negative and positive freedom whenever we (1) think we are alone, but are in fact watched by machines that observe our online and much of our offline behaviour; (2) think we are making private decisions based on a fair idea of what is going on, while in fact we have no clue as to why service providers, insurance companies or government agencies are dealing with us the way they do.

Referring to what has been discussed in section 2 we should admit that most of our interactions take place without conscious reflection; they are a type of *autonomic* behaviour that is the result of individual learning processes that enable us to move smoothly through everyday life. This, in itself, is not a violation of our negative or positive freedom. As a result of learning processes it may even be the result of the way we exercised our freedom in the past ([56], [22]). However, *autonomous* action (other than autonomic behaviour) is related to the possibility of deliberate reflection on our choices of action. For this we need to have access to the knowledge that impacts these choices. Without such access targeted servicing, customisation and filtering of information might provide us with a comfortable, golden cage; allowing us a reflexive life without reflection ([9], [31], [40], [52]).

5. Threats: Knowledge is Power

The potential threats of profiling must not be conflated with those of data collection per se. First, the implications of profiling for the autonomy of individual citizens do not depend on the collection of personal data but on the processing and mining of these data. The resulting profiles, which are applied to a person because her data match the profile, are often generated by data mining of other people's data. What should concern us here is the process of constructing profiles and their application on people whose data were not used to build the relevant profiles (disabling the applicability of the data protection directive that is focused on the protection of personal data).[8] Informational

[8] Art. 2 (a) of D 95/46/EC defines personal data as 'any information related to an identified or identifiable person'. See 'Opinion 4/2007 on the concept of personal data', of the Art. 29 Working Party. The question of

privacy is all too often reduced to a private interest in the hiding of personal data. This reduction misses out on the *knowledge* asymmetry between profilers and profiled, which has far more implications than the *information* asymmetry that focuses on access to personal data. Second, because of the reduction of privacy to non-disclosure of personal data, privacy is often depicted as a private interest, to be traded against other interests. However, in acknowledging that privacy is not only about personal data, we must face the fact that privacy is also a public good that concerns a citizen's 'freedom from unreasonable constraints on the construction of her identity' ([3]:7, cf. also [2], [45]).[9] This freedom is a precondition for democracy and rule of law, as I have argued in the previous section. The hiding of personal data – which seems a Pavlov reaction of lawyers and other privacy advocates - will, however, not protect us from the impact of group profiling on the construction of our identity, while at the same time the hiding of personal data will reduce the quality of the profiles (and the intelligence of the environment). Third, in the discourse on public and private security, we are often called upon to trade part of our privacy (understood as non-disclosure of personal data) for security. However, neither privacy nor security are fit for private trading. While privacy is a public good in as far as it is constitutive of human agency in a constitutional democracy, security is one of the *raisons d'être* of the state. A state that does not provide its citizens with a minimum of security should be qualified as a failed state, incapable of protecting citizens against each other and/or abusive state officials. With regard to the cliché of a trade-off between privacy and security two points must be made. First, a loss of privacy may imply a loss of security, because it exposes the vulnerability of human identity, whereas the privacy of some may be traded against the security of others.[10] Second, trading with personal data is problematic because we may expect a market failure due to the unequal access to information about the consequences of trading one's personal data (especially in the case of profiling [48]).[11] With Waldron [57] we note that the image of the scale that comes with the notion of a trade-off is a metaphor that functions like a Trojan horse; it delivers an entire network of assumptions that may fit the function of a kitchen scale but not that of weighing competing public goods or human rights.

Profiling machines may spy on you, but why should you care about a machine watching your everyday interactions? In AmI, most of the monitoring and adaptation will be a matter of machine to machine communication, while these machines will not be interested in who you are but in what profit can be gained from which category you fit. How does this relate to privacy and security as constitutive (public) goods that enable citizens to develop the kind of agency that is presumed in constitutional democracy [48]? In the case of profiling, the issue of privacy and security raise the question of 'who is in control: citizens or profilers?' Alas, control is often reduced to

---

what data qualify as personal data is controversial and may become thus contextual that legal certainty as to which data fall within the scope of the directive is lost.

[9] In speaking of the construction of one's identity we endorse a relational non-essentialist conception of identity. See e.g. ([21]: 312-315) on the difference between *idem*-identity and *ipse*-identity.

[10] Schneier [47] suggests that security always involves trade-offs and we had better get used to it. He argues that we should demystify security and start making sensible security trade-offs. My point is that even if both security and privacy involve trade-offs, we should get over the idea that trading privacy for security will indeed provide long term security. The interrelationship between the two is far too complex and we had better invest our energy in win-win solutions (cf. [11]).

[11] Schwartz ([48]: 745) refers to the Calabresis-Melamed analysis that states that property rules work well in the case of few parties, difficult valuations, low transaction costs, while liability rules work well in the case of many parties, monopoly, strategic bargaining and high transaction costs.

hiding or disclosing personal data and this does not cover privacy and security as constitutive public goods. To come to terms with potential threats we need to look deeper into the asymmetries between citizens on the one hand and large organisations that have access to their profiles on the other. I am not referring to the asymmetry of effective access to personal data but to the asymmetry of effective access to knowledge inferred from data aggregates. Especially insofar as this knowledge is protected as part of a trade secret or intellectual property the citizens to whom this knowledge may be applied have no access whatsoever. Zarsky [59] has demonstrated – by analysing a set of examples – how this lack of access can lead to what he calls the 'autonomy trap'. Precisely because a person is not aware of the profiles that are applied to her, she may be seduced to act in ways she would not have chosen otherwise. Imagine that my online behaviour is profiled and matched with a group profile that predicts that the chance that I am a smoker who is on the verge of quitting is 67%. A second profile predicts that if I am offered free cigarettes together with my online groceries and receive news items about the reduction of dementia in the case of smoking I have a 80% chance of not quitting. This knowledge may be sold to the tobacco industry, which may use it to influence my behaviour. In a way, this kind of impact resembles Pavlov's stimulus-response training: it does not appeal to reason but aims to discipline or seduce me into profitable behaviour. My autonomy is circumvented and my intention pre-empted ([35]:3) as long as I am not aware of the knowledge that is used. Zarsky [59] also warns about unfair discrimination, based on refined profiling technologies that allow sophisticated market segmentation. Price discrimination may be a good thing in a free market economy, but the fairness again depends on the awareness of consumers of the way they are categorised [38]. In order to have a fair and free market-economy some rules of the game must be established to prevent unequal bargaining positions, or else we have another market failure. In short the threats can be summarised as concerning (1) privacy, which, however, must not be reduced to hiding one's personal data; (2) security, which, however cannot be traded with privacy since a loss of the one may cause the loss of the other; (3) unfair discrimination, meaning that power relations must be balanced to provide equal bargaining positions; and (4) autonomy, meaning that our negative and positive freedom to act must be established and maintained, since manipulation on the basis of knowledge that one is not aware of violates one's autonomy.

6. The legal framework around profiling

Data have a legal status. They are protected, at least personal data are. Europe tends to understand this protection as a personality right, which opens the possibility to declare certain data to be inalienable. In practice, however, the leaking of personal data is taken to imply consent for storing and using them. Whatever the written safeguards we find in the data protection directive, in practice most people most of the time do not have a hunch of what is happening to which data resulting in the application of what profiles. Some US scholars, notably Lessig [31], favour commoditization in order to facilitate the trading of one’s personal data. In their eyes this should provide at least some kind of citizen's control. However, as discussed above with reference to [48], one may expect a market failure in the sense that due to grotesque knowledge asymmetries the implied consent will be based on ignorance – just like it is today. In both cases one of the problems is that we have no access to the group profiles that have been inferred

from the mass of data that is being aggregated and have not the faintest idea how these profiles impact our chances in life. It may be time to reconsider the legal focus on the protection of personal data, as well as the focus of the privacy advocates who invest in privacy enhancing technologies. What we need is a complementary focus on the dynamically inferred group profiles that need not be derived from one's personal data at all, but may nevertheless contain knowledge about the probability of one's (un)healthy habits, earning capacity, risk-taking, life style preferences, spending habits, political associations etc.

Profiles have no clear legal status. That is, they may be *protected from* access via intellectual property rights by the profiler or be considered part of a company's trade secrets.[12] Protection against, or at least access to profiles is very limited. In data protection legislation one can locate two ways to claim access to a profile. First one can argue that once a profile has been applied to an individual person it becomes a personal data, e.g. in the case of credit scoring practices. This however, does not concern the relevant group profile or its relation to other group profiles, nor the way the profile was generated (by use of which algorithm etc.).[13] Second autonomic application of profiles may fall within the scope of art. 15 of the Data Protection Directive (D 46/95 EC). Paragraph 1 of this article 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

The proposed art. 20 (1) of the Draft General Data Protection Regulation is equivalent in its wording. In short, the article seems to grant European citizens a right not be subjected to an automated decision in the case that this decision makes a significant difference to their life. The existing safeguard has four pitfalls. First, like Bygrave [8] suggests, it may be that if I don't exercise the right, the automated decision is not a violation of the directive. The draft Regulation, however, stipulates that a person may *only* be subjected to automated decisions under specified conditions, implying that this right is not merely a right to object. Such a stipulation will provide for legal certainty, creating a level playing field for all stakeholders involved in autonomic profiling. Second, the 2nd paragraph of art. 15 reads:

> Subject to the other Articles of this Directive, Member States shall provide that a person may be subjected to a decision of the kind referred to in paragraph 1 if that decision:
>
> (a) is taken in the course of the entering into or performance of a contract, provided the request for the entering into or the performance of the contract, lodged by the data subject, has been satisfied or that there are suitable measures to safeguard his legitimate interests, such as arrangements allowing him to put his point of view; or
>
> (b) is authorized by a law which also lays down measures to safeguard the data subject's legitimate interests.

This seems to create many ways out for automated application of profiles. The draft Regulation adds under (a) – as mentioned above – that these suitable measures

[12] Cf. section 41 of the preamble of D46/95/EC. See [14].

[13] The relevant group profile may determine the credit score, which is then a personal data. The profile would indicate that all people with a specific mix of attributes (concerning income, neighbourhood, credit history, gender, profession, educational background) entails a specific credit-risk. This group profile applies to a number of people, it is the result of data mining and not a personal data in the sense of the directive.

may include the right to obtain human intervention. This implies that the person affected by automated profiling can require the data controller to involve a human decision maker. Coupled with the obligation for data controllers to provide 'information as to the existence of processing' for automated decision-making (art. 20 (4)) this will provide inhabitants of smart environments with a strong protection against subliminal manipulations. The third pitfall concerns the fact that as soon as the decision is not automated due to a (routine) human intervention the article no longer applies. However, in the case of autonomic profiling in an AmI environment this would not be an option, because the seamless real time adjustment of the environment rules out such human intervention. This brings us to the fourth and last pitfall of the present data protection regime: as long as one is not aware of being subject to such decisions one cannot exercise this right. The fact that art. 12 grants the right to know 'the logic involved in any automatic processing of data concerning him at least in the case of the automated decisions referred to in art. 15' does not really help if one doesn't know about the automated decisions in the first place. As indicated, the draft Regulation solves this problem, at least in theory, by creating an obligation for data controllers to provide information about the existence of this type of profiling. To the extent that this obligation does not hinge on a person requesting it, this would make a radical difference with the present level of protection. If the obligation to provide this information depends on individual requests, however, we will be back to square one. Obviously, if people are not aware of the computational background of automated decisions they will not request the relevant information and they will not benefit from obligations to provide it. The draft Regulation speaks of 'the controller shall provide the data subject with at least the following information' (art. 20 (4,2) jo 14). This suggests that the data controller should not await a request, but in fact this wording is similar to that under the existing Directive, so we cannot be sure about who should take the initiative here.

As described in the introduction, the draft Regulation provides for a revolutionary novel legal requirement: the obligation for data controllers to provide 'information about the envisaged effects of such processing on the data subject'. Next to an obligation to provide information about the existence of profiling, the data subject must be informed about the envisaged consequences. The problem with autonomic profiling – as argued above – regards the subliminal influences on the process of identity building. The threats to our privacy and autonomy derive from not knowing how our data will cluster together with other data, not knowing how they will form patterns that could disclose future behaviours, inclinations, health or other risks. By forcing data controllers to notify us what risk we are taking by leaking our data the draft Regulation may achieve two important goals: first, the purpose specification principle is reinstated as an important legal rule, because envisaging effects requires ex ante specification of the targeted effects; second, effects that are not intended but can be envisaged due to the generative nature of profiling must be assessed and communicated. This last goal is particularly important in smart environments that thrive on function creep [32]. As indicated above, profiling enables the construction of unexpected knowledge, or unknown unknowns. To some extent the smartness of the environment depends on pattern recognition that cannot be detected by either the naked human eye or algorithmic determination. Having a legal obligation to probe the future capabilities of the smart environment will require data controllers to develop some kind of hermeneutic of their profiling systems to anticipate how they will affect their end-users. To call this a revolutionary but critical obligation would be an understatement.

It seems that we have a double challenge here. First, the existing legal framework lacks adequate protection with regard to the application of group profiles. The existing right of access to the logic of processing is restricted to very specific circumstances that may not apply. Apart from that, the fact that such profiles are generally protected by means of trade secret or intellectual property turns the legal right of access to the logic of processing into an empty shell. Second, insofar as the draft Regulation pays more apt attention to decision-making based on autonomic profiling, the industry will have to invest in a technological and organisational infrastructure that enables compliance with the relevant obligations. Without such a socio-technical infrastructure the Regulation cannot provide us with an effective remedy. This would entail comprehensible, contestable and reliable information for individual citizens about profiles that may be applied to them, including the potential consequences: art. 20 (4) of the draft Regulation seems an excellent starting point for levelling the playing field in this direction. Only if such an infrastructure is in place the rule of law, especially the particular mélange of positive and negative freedom discussed above, can be sustained.

7. Legal Transparency By Design

The idea that legal protection requires articulation into the technological infrastructure against which protection is warranted,[14] seems to gain currency.

The draft Regulation stipulates in art. 23 that data controllers must implement data protection by design and by default. 'By design' refers to appropriate technical and organizational measures and procedures that ensure compliance with the Regulation, 'by default' refers to mechanisms that ensure that 'only those personal data are processed which are necessary for each specific purpose of the processing and are especially not be collected or retained beyond the minimum necessary for those purposes, both in terms of the amount of the data and the time of their storage. In particular, those mechanisms shall ensure that by default personal data are not made accessible to an indefinite number of individuals.' It seems that privacy by design thus refers to socio-technical articulation of all the relevant rights and obligations of the Regulation, whereas privacy by default specifically targets the data minimisation principle.

This is an important milestone, if the draft Regulation is enacted as such. It refers to what has been called constructive technology assessment (CTA), initiated by e.g. Rip et al. [41], building a case for 'upstream' involvement in technological design, i.e. not appointing ethical commissions after a technology is brought to the market but getting involved at the earliest possible stage of technological design. However, as we may guess, designers' good intentions do not determine the actual affordances of a technology, due to the multistability of technological artefacts [24]. Multistability refers to the fact that one and the same technology often affords different behaviours, while it is not always easy to anticipate which behaviour will emerge once the technology is integrated in the socio-technical context of its users. Nevertheless, multistability does not imply that anything goes, or that it makes no sense to anticipate the affordances of technologies under construction. On the contrary, multistability

[14] Modern law is articulated in the technology of the script and can seen as an 'affordance' of the printing press. About the idea that legal norms need articulation in the digital infrastructure that is emerging today, see [21] and [23].

means that upstream involvement of potential end-users and others who may be affected by the technologies, will broaden the scope of technical design and increase the opportunities to construct a socio-technical infrastructure that does not obstruct the flourishing and autonomy of individual citizens.

With regard to smart environments this is a serious challenge. To the extent that environments are smart due to the mining of a mass of behavioural data, minimisation of data processing could actually stifle the feedback mechanisms that make the environment smart. If we want to have our cake and eat it too, we will require effective transparency enhancing tools (TETs) that allow citizens to anticipate how they will be profiled and which consequences this may entail.[15] This will entail a rethinking of the legal-technological infrastructure of smart environments. Evidently, commercial enterprise has an interest in protecting its trade secrets or its intellectual property rights in databases or software programmes. The tension between rights of access and corporate property rights has been acknowledged in section 41 of the preamble of the present Directive. It seems that the draft Regulation resolves this tension by no longer stipulating a right of access to the relevant algorithms and instead requiring data controllers to 'tell' their users how profiling may affect them.[16]

Developing transparency tools would be the first step in the process of providing transparency by design. One could, for instance, think of inference machines capable of calculating the types of consequences that may be envisaged in the case of profiling. However, the complexity of the concerned profiling processes and the growth of information they engender [28], generate various novel challenges. First, the complexity as well as the quantity of information produced by transparency enhancing technologies could overwhelm an individual person, if this information were provided in the form of text, requiring her conscious attention. It seems that our cognitive economy would be flooded without a chance of making sense of the information that is presented. TETs will only succeed in empowering citizens if the human machine interfaces (HMIs) that mediate to provide transparency are as seamless and ubiquitous as the AmI infrastructure they aim to tame. TETs should allow a person to play around with the environment in order to guess how her behaviours trigger proactive interventions of the environment [37]; they should not inundate a person with detailed technical information that requires her scrutiny in a way that nullifies all the 'advantages' of ubiquitous and seamless computing. The HMIs will have to communicate the relevant information in a way that allows one to have 'a feel' of the environment's interpretation of one's behaviour, rather than merely adding more text or graphs to the equation. This, however, does not mean that a more precise access to the technical details must not be available, for instance to enable a person subjected to unfair decision-making on the basis of autonomic profiling, to contest the application of profiles in a court of law. This brings us to a second major challenge, which concerns the fact that at some point such technical detail cannot be provided, due to the fact that autonomic profiling will be self-repairing, self-healing and self-managing to an extent that turns the whole process into a black box that even the designer of the

[15] Cp. [12] about the fact that data protection legislation is mainly a transparency tool, while privacy is considered to be an opacity tool. My point is that the transparency aimed for by the present generation of data protection regimes concerns personal data, without taking note of the results of data processing. The results, consisting of highly sophisticated group profiles, urgently need effective transparency tools.

[16] Note that recital 51 of the preamble of the draft Regulation still speaks of 'the logic of the data undergoing the processing'. It is not clear what this means in terms of access to software codes, since they are protected by copyrights, as the recital notes.

process cannot open. And, to further complicate the issue, even if the technical detail could be disclosed, the human mind could not possibly follow – let alone explain – what happens inside the profiling machines. To follow and check this we would need another machine with similar or even more computing capacities [55].

These challenges should not paralyse us. They should, instead, be a wake-up call for lawyers, politicians and computer engineers to join forces during the construction of the smart infrastructures in which so much capital is presently invested. These new digital infrastructures will match the printing press in terms of their impact on the structure of our societies and this warrants speculative even if rigorous investigation as well as sustained interdisciplinary dialogue.

8. Closing remarks

Advanced profiling technologies answer questions we did not raise. They generate knowledge we did not anticipate, but are eager to apply. As knowledge is power, profiling changes the power relationships between profilers and those being profiled. These asymmetries challenge the relative autonomy of individual citizens and allow an unprecedented dynamic segmentation of society, especially if the vision of Ambient Intelligence is realised: based on refined real time monitoring, followed by proactive adaptation of our smart environment. As long as we lack the legal and technological infrastructure to counter the emerging asymmetry we may find ourselves in a golden cage: an environment that anticipates our preferences before we become aware of them.

This contribution argues that we urgently need to develop legal and technological transparency enhancing tools (TETs) to match the proactive dimension of our smart environments. The draft General Data Protection Regulation acknowledges this need by attributing an obligation to data controllers to supply users with information about the consequences of automated decision-making. To comply with such an obligation the industry will have to involve cognitive scientists, computer engineers, lawyers, designers of interfaces and experts in human-computer interaction with a clear understanding of what is at stake in terms of democracy and the rule of law. The ensuing cross-disciplinary cooperation should allow us to sustain the legal-political framework that safeguards the right to be free from unreasonable constraints on the construction of identity in information society.

References

[1] E. Aarts and S. Marzano (eds.), *The New Everyday. Views on Ambient Intelligence*. Rotterdam, 010, 2003.

[2] E. Aarts and F. Grotenhuis, Ambient Intelligence 2.0: Towards Synergetic Prosperity, in *AmI 2009*, ed. Manfred Tscheligi et al., Berlin Heidelberg: Springer, 2009, 1–13.

[3] P.E. Agre and M. Rotenberg, *Technology and Privacy: The New Landscape*. Cambridge, Massachusetts, MIT, 2001.

[4] Chr. Anderson, The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, *Wired Magazine* **16** (7) (2008).

[5] Art. 29 Working Party (2007), *Working paper 136, Opinion 4/2007 on the concept of personal data*. Brussels, available at: http://ec.europa.eu/justice_home/fsj/privacy/workinggroup/index_en.htm .

[6] I. Berlin, Two concepts of liberty, idem, *Four essays on liberty*, Oxford New York, Oxford University Press, 1969**,** 118-173.

[7] D.M. Berry, *Understanding Digital Humanities: The Computational Turn and New Technology*, London: Palgrave Macmillian, 2012.
[8] L. Bygrave, Minding the Machine, Art.15 and the EC Data Protection Directive and automated profiling, *Computer Law & Security Report* **17** (2001)**,** 17-24.
[9] N. Carr, *The Shallows: What the Internet is Doing to Our Brains,* New York: W.W. Norton, 2010.
[10] A. Cavoukian and T. Hamilton, *The Privacy Payoff. How Successful Businesses Build Consumer Trust*, McGraw-Hill Ryerson, 2002.
[11] B. Custers, *The Power of Knowledge. Ethical, Legal, and Technological Aspects of Data Mining and Group Profiling in Epidemiology*, Nijmegen, Wolf Legal Publishers, 2004.
[12] P. De Hert and S. Gutwirth, Privacy, Data Protection and Law Enforcement. Opacity of the Individual and Transparency of Power, in *Privacy and the Criminal Law*, ed. Erik Claes, Antony Duff, and S. Gutwirth, Antwerp: Intersentia, 2006.
[13] C. Dwyer, The Inference Problem and Pervasive Computing, *Proceedings of Internet Research* 10.0, Milwaukee, WI, 2009.
[14] N. Van Dijk, The Legal Status of Profiles, J.C. Augusto (ed.), *Intelligent Environments 2009, Ambient Intelligence and Smart Environments* 2, Barcelona, IOS, 2009, 510–516.
[15] W.H. Dutton and P.W. Jeffreys, *World Wide Research*, MIT Press, 2010.
[16] G. Elmer, *Profiling Machines. Mapping the Personal Information Economy*. Cambridge, Mass., MIT Press, 2004.
[17] U.M. Fayyad et al. (eds.), *Advances in Knowledge Discovery and Data Mining*, Meno Park, California - Cambridge, Mass. - London England, AAAI Press / MIT Press, 1996.
[18] J. Gibson, *The Ecological Approach to Visual Perception*, Lawrence Erlbaum Associates, New Jersey, 1986.
[19] P. Haggard and B. Libet, Conscious intention and brain activity, *Journal of Consciousness Studies* **8** (11) (2001), 47–63.
[20] M. Hildebrandt, Privacy and Identity, in E. Claes, A. Duff and S. Gutwirth (eds.), *Privacy and the Criminal Law*, Antwerp: Intersentia, 2006, 43-58.
[21] M. Hildebrandt, Defining Profiling: A New Type of Knowledge and Profiling and The Identity of the European Citizen, in *Profiling the European Citizen. A Cross-disciplinary Perspective*. M. Hildebrandt and S. Gutwirth, Springer: Dordrecht, 2008, 17-30 and 303-326.
[22] M. Hildebrandt, *Autonomic* and Autonomous 'Thinking': Preconditions for Criminal Accountability, in M. Hildebrandt and A. Rouvroy (eds.), *Law, Human Agency and Autonomic Computing. The Philosophy of Law Meets the Philosophy of Technology*, Abingdon: Routledge, 2011.
[23] M. Hildebrandt and B.J. Koops, The challenges of Ambient Law and legal protection in the profiling era, *Modern Law Review* **73** (3) (2010), 428–460.
[24] D. Ihde, *Technology and the Lifeworld. From Garden to Earth,* Bloomington and Indianapolis, Indiana University Press, 1990.
[25] ISTAG, *Scenarios for Ambient Intelligence in 2010*, Information Society Technology Advisory Group, 2001**,** available at: http://www.cordis.lu/ist/istag-reports.htm.
[26] ITU, *The Internet of Things*. Geneva, International Telecommunications Union (ITU), 2005.
[27] S. Jin et al., *Driver Fatigue Detection Using a Genetic Algorithm*, Artificial Life and Robotics **11** (1) (2007), 87-90.
[28] J. Kallinikos, *The Consequences of Information. Institutional Implications of Technological Change*, Cheltenham, UK Northampton, MA, USA, Edward Elgar, 2006.
[29] R. D. King et al., The Automation of Science, *Science* **324** (5923) (2009), 85–89.
[30] Ch. Kuner, Ch., The European Commission's Proposed Data Protection Regulation: A Copernican Revolution in European Data Protection Law, *Privacy & Security Law Report* **11** (6) (2012), available at http://www.huntonprivacyblog.com/tag/christopher-kuner/.
[31] L. Lessig, *Code and other laws of cyberspace*. New York, Basic Books, 1999.
[32] B. Masiello and A. Whitten, Engineering Privacy in a Age of Information Abundance, *Intelligent Information Privacy Management*, *AAAI* (2010), 119–124.
[33] H.R. Maturana and F. J. Varela, *Autopoiesis and Cognition: The Realization of the Living*, Dordrecht, Reidel, 1991.
[34] V. Mayer-Schönberger, *Delete: The Virtue of Forgetting in the Digital Age,* Princeton University Press, 2009.
[35] A. McStay, *The Mood of Information: a Critique of Online Behavioural Advertising,* New York, Continuum, 2011.
[36] T. M. Mitchell, *The Discipline of Machine Learning*, Carnegie Mellon University, School of Computer Science, 2006, available at http://www-cgi.cs.cmu.edu/~tom/pubs/MachineLearningTR.pdf.
[37] D. Nguyen, E. Mynatt, *Privacy mirrors: understanding and shaping socio-technical ubiquitous computing systems*, Technical Report GIT-GVU-02-16, Georgia Institute of Technology, USA, 2002.

[38] A. M. Odlyzko, Privacy, economics, and price discrimination on the Internet, ICEC2003 - *Fifth International Conference on Electronic Commerce*, N. Sadeh, (ed.), ACM, 2003, 355-366.
[39] M. Overgaard, The Role of Phenomenological Reports in Experiments on Consciousness, *Psycoloquy* **12**(029) Consciousness Report (1), 2001.
[40] E. Pariser, *The Filter Bubble: What the Internet is Hiding from You*, Penguin, 2011.
[41] A. Rip, et al., *Managing Technology in Society: The Approach of Constructive Technology Assessment*, Pinter Publishers, 1995.
[42] I. Robeyns, The Capability Approach: a theoretical survey, *Journal of Human Development* **6** (1) (2005), 93-114.
[43] Y. Rogers, New theoretical approaches for human-computer interaction, *Annual Review of Information Science and Technology* **38** (1) (2004), 87-143.
[44] J. Rosen, The Right to Be Forgotten, *Stanford Law Review Online* **64** (13) (2012), 88.
[45] A. Rouvroy, Privacy, Data Protection, and the Unprecedented Challenges of Ambient Intelligence, *2 Studies in Ethics, Law, and Technology* **1**, Article 3 (2008), at: http://www.bepress.com/selt/vol2/iss1/art3.
[46] F. Schauer, *Profiles Probabilities and Stereotypes*. Cambridge, Massachusetts / London, England, Belknap Press of Harvard University Press, 2003.
[47] B. Schneier, *Beyond Fear. Thinking Sensibly About Security in an Uncertain World*, Springer, 2006.
[48] P.M. Schwartz, Beyond Lessig's Code for Internet Privacy, in Cyberspace Filters, Privacy-Control and Fair Information Practices, *Wisconsin Law Review* (2000), 743-788.
[49] J.C. Scott, *Seeing Like a State. How Certain Schemes to Improve the Human Condition Have Failed*, New Haven and London, Yale University Press, 1998.
[50] D.J. Solove, *The Digital Person. Technology And Privacy In The Information Age*, New York University Press, New York, 2004.
[51] F. Stalder, The Failure of Privacy Enhancing Technologies (PETs) and the Voiding of Privacy, *Sociological Research Online*, **7** (2) (2002), at: http://www.socresonline.org.uk/7/2/stalder.html.
[52] C. Sunstein, *Republic.com*. Princeton and Oxford, Princeton University Press, 2001.
[53] D. Tennenhouse, Proactive Computing, *Communications of the ACM* **43** (5), 2000, 43-50.
[54] J. Torpey, *The Invention of the Passport. Surveillance, Citizenship and the State*. Cambridge, Cambridge University Press, 2000.
[55] J.P. Van Bendegem, Neat Algorithms in Messy Environments, in M. Hildebrandt and S. Gutwirth (eds.), *Profiling the European Citizen. Cross-Disciplinary Perspectives*, Dordrecht, Springer 2008, 80-83.
[56] F.J. Varela, *Ethical Know-how*. Stanford, Stanford University Press, 1992.
[57] J. Waldron, Security and Liberty: The Image of Balance, *Journal of Political Philosophy* **11** (2) (2003), 191–210.
[58] A. Yannopoulos, V. Androniki, and Th. Varvarigou, Behavioural Biometric Profiling and Ambient Intelligence, in M. Hildebrandt and S. Gutwirth (eds.), *Profiling the European Citizen. Cross-Disciplinary Perspectives,* Springer, Dordrecht, 2008, 89-104.
[59] T.Z. Zarsky, Mine Your Own Business!: Making the Case for the Implications of the Data Mining or Personal Information in the Forum of Public Opinion, *Yale Journal of Law & Technology* **5** (4) (2002-2003), 17-47.