Counting Sheep: Automated Profiling, Predictions and Control

Martijn van Otterlo
m.vanotterlo@donders.ru.nl
Cognitive Artificial Intelligence
Radboud University Nijmegen
The Netherlands

Paper prepared [*]for, and to be presented at,
the Amsterdam Privacy Conference (7-10 October 2012)
(Working Paper – Please do not cite without permission)

Abstract

Current technologies for big data are increasingly able to automatically gather data, experiment with action strategies, observe results of such strategies, and learn from their effects. When privacy issues are framed as "control over information" then it becomes apparent that some areas in the digital world might be heading to what I call *Walden 3.0*; communities of interest that are controlled by experimentation and reinforcement. Instead of bringing forward Orwell's dystopia *1984* in the privacy domain as is typically done, we sketch how current developments might be better studied in the context of Skinner's utopian novel *Walden II*. We exemplify several issues through a running example from the domain of artificial intelligence, and by mentioning several areas where automated experimentation (may) arise. The article ends with raising questions on how to cope with and study the phenomenon of automated experimentation, and whether Walden 3.0 in the end is dystopian or utopian.

1. Privacy in the Age of Big Data

The adagium [1] of the infamous secret service agency Stasi in former East Germany was *"to know everything"*. Through numerous agents information was covertly gathered about people, their habits and social circles, and put in paper files. In contrast, nowadays, the average *FaceBook* user assembles his personal file i) voluntarily and personally, ii) electronically and digitally, and iii) connecting it automatically to those of people in his social circle. Pressure for doing so does not only come from companies' so-called *free* services, but seems to be mainly of a social nature. The more control people *think* they have about privacy, the more eager they seem to be to disclose even more information (Brandimarte *et al.*, 2009). In addition, we are constantly *measured* and *tracked* by companies and governments. Think of smart energy meters, biometric information on passports, number plate tracking, medical record sharing, etcetera. One may ask whether the concept of *privacy* is still an issue *at all* in the current digital age.

In recent years much has been written about the evolution of privacy (Tene, 2011) in the era of *big data* (Bollier and Firestone, 2010), *the computational turn* (Hildebrandt and Gutwirth, 2008), and *profiling* (van Otterlo, 2012); an era called by some *the petabyte age* [2]. Advances in electronic

[*] This paper was typeset using pdfTeX, Version 3.1415926-2.4-1.40.13 (MiKTeX 2.9), in Utopia font with Math Design).

[1] See the motion picture "Das Leben Der Anderen" for an interesting interpretation (The Lives of Others, http://www.imdb.com/title/tt0405094/)

[2] See the Wired issue: http://www.wired.com/science/discoveries/magazine/16-07/pb_intro

surveillance [3], internet technology and search engines, as well as the rise of social networks, have changed the concept of privacy tremendously. Several recent books elaborate on that, ranging from *Googlization* (Vaidhyanathan, 2011), *social networks* (Andrews, 2011), *filter bubbles* (Pariser, 2011) and *personalization* (Turow, 2011). Not just the amount of data, but also novel ways to analyze that data, changes the playing field of any single individual in the information battle against big companies and governments. Data is becoming a key element for profit and control.

Connecting much of the privacy literature is the idea that hidden underneath the surface all is controlled by *code*. Google's and Facebook's [4] services make heavily use of smart *algorithms* developed in *computer science* and *artificial intelligence* (AI) (Nilsson, 2010). These algorithms learn to predict people's characteristics and intentions (*profiling*), tailor search results to the individual user's needs (*personalization*) or make search results and recommendations more *"social"* (*collaborative filtering*). Previously (van Otterlo, 2012) I have described how algorithmic ideas from *machine learning* form the algorithmic base of profiling. Most privacy papers treat such algorithms as black boxes under a general name such as *data mining* or *analytics*. I would like to argue that opening up these Pandora's boxes at least slightly – that is, developing an *algorithmic literacy* – is vital for understanding what really happens in terms of privacy and control. In order to understand *code* in the privacy context, one needs to go to the level of *models*, which in many cases are *richly structured* combined with *statistics*. Such models are sufficient to understand how code can give us new (and multiple) *algorithmic identities* (Cheney-Lippold, 2011), how they are *exploited* and how new forms of *automated experimentation* are starting to influence our daily lives. Google's biased search results can be mentioned as one example of the latter, but there are many more (van 't Hof *et al.*, 2012) and we need to understand the algorithms responsible for that.

In this paper I consider privacy to be directly related to the influencing of individual's options and perception, or more general to *control*. Privacy refers to the terms of control over information, not the nature of that information. Famous novels such as Orwell's *1984* deal with in-your-face control and archetypical surveillance devices, *Telescreens*. Current control mechanisms in society work much more like other books in the same genre (Claeys, 2010), for example Huxley's *Brave New World* when it comes to soft control using entertainment, sex and drugs, but even more in terms of the glass houses of Zamyatin's *We*. After all, the places we live in in the digital world are fully transparent. However, as I will argue in this paper, technological developments direct us towards a lesser known utopian novel, *Walden II* by Skinner (1948). It describes a society which is trained to behave well using positive reinforcements, based on Skinner's ideas[5] in behaviorism and behavioral engineering.

Understanding profiling, control and experimentation at an algorithmic level may help to understand the evolution of privacy in the digital age. In summary, this paper aims to cover two important things. The first aspect contributes to the algorithmic literacy by providing a simple, yet state-of-the-art, model type from artificial intelligence as running example. This will be the focus of the next section. The second aspect is the main theme; highlighting how models can be exploited for profit and control, and how automated control and even experimentation are within reach in many domains. An understanding of *that* this happens, and *how* it works, may contribute to studies on how privacy evolves, and give clues on how we can counter undesired developments.

[3] See the recent stream of news items and newspaper articles on this topic. An example is the text by Naomi Wolfe, "The new totalitarianism of surveillance technology" in the Guardian, 5th August, at http://www.guardian.co.uk/commentisfree/2012/aug/15/new-totalitarianism-surveillance-technology

[4] Big data companies such as Google and Facebook employ relatively large amounts of PhD graduates from areas such as computer science and artificial intelligence.

[5] The ideas of Skinner have – just like those of Pavlov, Thorndike and Watson – laid a foundation for current algorithmic techniques in artificial intelligence and reinforcement learning (Wiering and van Otterlo, 2012).

2. Profiling, Models and Artificial Intelligence

Models (or *profiles*) are at the heart of profiling, i.e. the art of constructing *predictive theories* about a domain. This amounts to finding causal rules, invariant patterns and statistical regularities in data. Profiles induce *categories*, or sub-groups, to which people can belong to a certain degree. For example whether they are *gay* [6] or a *retirement planner* [7]. The very idea that algorithms (or, code) implement such categories and they can be applied selectively to people, requires a good look at the resulting *algorithmic identities* (de Vries, 2010; Cheney-Lippold, 2011). Algorithms nowadays define how we are seen, by providing a digital lens, tailored by statistics and other biases.

In this section we will take a look at a tiny code example, just enough to exemplify some notions. Note that systems such as Google or Facebook are much more complex; however, conceptually they are very similar in terms of what they do.

2.1 Machine Learning of Models

Artificial intelligence (Nilsson, 2010) is *the science and engineering of making intelligent machines. Machine learning* (Flach, 2012) as a major subfield is *the art and science of algorithms that make sense of data.* It develops algorithms that can *induce* (or, *learn*) theories about a domain from data originating from that domain. For example, based on pictures of apples and pears, a typical machine learning algorithm (e.g. a *neural network*) learns how to *classify* (i.e. discriminate between) apples and pears, thereby *generalizing* this mapping such that previously unseen pictures can be classified too. A very active sub-area, and very much related to profiling, is *activity recognition* (Yang, 2009), which is learning models of human activity. Much research focuses on *video* data, for example to detect suspicious behavior from CCTV data.

An important distinction in machine learning is between *discriminative* and *generative* models [8]. Discriminative models can only *distinguish* between individuals. For example, one could construct a model that would *classify* any given painting into *Mondrian* and *non-Mondrian.* It could do this by looking, for example, at the painting's colors. Alternatively, generative models learn the complete characteristics of the data itself, and are capable of *generating* specific examples (i.e. Mondrian-like paintings) from this model. This class is more complex, but also more suitable for profiling purposes. For Mondrian paintings, these models would represent aspects such as the average line thickness, the distribution of line lengths and so on. Such models can generate *arbitrary* predictions about the data, using *probability* calculations. The most prominent class of models are *probabilistic graphic models* (PGM) (Koller and Friedman, 2009), which include Bayesian networks.

Another distinction in machine learning concerns the *representation* of data models. Many algorithms see the data as consisting of *attributes* [9] with their values. For example, a painting can be described as *colorful* = *yes* and *number-of-lines* = 17. A recent trend is to go beyond that and use *relational* representations (De Raedt, 2008) (combined with probabilistic aspects) to allow for explicit connections between data elements, for example `paintedBy(VictoryBoogieWoogie,Mondrian)` as in a relational database. Such representations are more natural, since most data is fundamentally networked; most particularly when modeling social networks.

[6] See the research on the MIT reality mining dataset at http://reality.media.mit.edu

[7] These are a subclass of *boomer*, see advertising at AOL at http://advertising.aol.com/audiences/boomers

[8] See (Andrzejewksi *et al.*, 2010) for some graphical examples.

[9] See my previous work for examples of attribute-based models (van Otterlo, 2012).

2.2 Models: Representational Parts

Now, let us start with our example model [10], in the context of marketing in social networks, represented [11] using the probabilistic relational language *DT-ProbLog* (Van den Broeck *et al.*, 2010). First there are identifiers (denoted in lowercase, e.g. p1) of individuals in our data:

```
person(martijn).
person(p1).
person(p2).
...
```

In addition, we have information about who is connected in the network to whom (called a trust relation):

```
trusts_directed(martijn,p1).
trusts_directed(p2,p3).
trusts_directed(p1,p2).
...
```

The data here only represents one-way relations for trust. However, using *background knowledge* we can define that this relation is symmetric:

```
trusts(X,Y) IF trusts_directed(X,Y) OR IF trusts_directed(Y,X)
```

Background knowledge goes beyond the data (i.e. beyond specific individuals) and represents *generic* knowledge about the domain itself: the *variables X* and *Y* can stand for any two individuals. A typical example are *family relations* that, for example, define the concepts *parent* or *grandma* (note that one definition uses another here, and generic variables are in uppercase, e.g. P1):

```
parent(P,C) IF mother(P,C) OR father(P,C)
grandma(P1,P2) IF parent(P1,P3) AND parent(P1,P3) and female(P1)
```

Background knowledge also provides us with the means to represent a bit more uncertain (i.e. probabilistic) information. Let us assume we are interested in whether people will buy some unnamed product. In the model we can represent the probability (0.3) that someone will be tempted to buy something if a friend has bought the same product using a rule:

```
0.3 :: buy_trust(Person,Friend).
```

This type of probabilistic knowledge is typically learned from data, by (nontrivial) counting in how many cases someone actually buys a product if one of his friends does. This way, statistical information is contained in the probabilistic facts. Instead of such *implicit* knowledge, one can also put information *explicitly* in the model (Thurman and Schifferes, 2012); the same representation suffices though. In addition, we also specify the conditions for someone to buy the product:

```
buys(Person) IF trusts(Person,Friend) AND buys(Friend)
     WITH PROBABILITY buy_trust(Person,Friend).
```

[10] One can find the complete model at http://dtai.cs.kuleuven.be/problog/tutorial-dtproblog.html

[11] To make the paper accessible to readers of a variety of backgrounds, we modified notation to make it more verbal.

Thus, if some friend buys the product there is a chance of 30 percent (based on the previous rule) that a person will buy the product. Again, note that this is generic knowledge and (without additional rules) applies to all persons appearing in the domain. An interesting aspect is that `buys` represents one of many partial *identities*: it denotes people that will buy the product because a friend does so. Furthermore, membership of this group of people is probabilistic: there is a 30% probability that a random friend belongs to that group (unless we know for sure that this person has bought the product). The rules that define them employ the power of *abstraction*: each rule only mentions a few *relevant* aspects. Note that "real" identities are not necessary; a generic identifier (e.g. `p1`) suffices. Identifiers can help [12] in *connecting* data snippets that are about the same individual.

Summarizing, when we talk about models, we usually mean the generic parts that are not tied to specific individuals, i.e. *rules* denoting under which conditions some individuals have particular properties. Factual information about specific individuals belongs to the data itself. Models represent general patterns using abstraction, thereby *aggregating* knowledge about many individuals.

2.3 Models: Algorithmics, Inference and Learning

We have seen that data and models are two separate things, although they can be represented similarly. Interaction *between* data and models is where *algorithmics* come into play. We can broadly distinguish three different types of algorithms: *deductive*, *abductive* and *inductive*.

Deductive algorithms are about *inferring* information. From our example model, we can (by logical reasoning) infer that `trusts(p3,p1)` using the generic rule for `trust − directed`. Relating to profiling and privacy, deduction can be used to infer *new* knowledge about an individual, for example how likely it is some person will buy the product. Reasoning with probabilistic models can be computationally complex. In our example, the probability that someone buys the product is dependent on the structure of the social network. Namely, for any single person, there could be several persons influencing that person (and that person can, in turn, influence other people) and computing the *right* probabilities is far from trivial since many different situations must be considered.

Where deductive algorithms focus on logical *consequences* of a model, *abductive* inference looks for additional *hypotheses* and *explanations* for facts found in the data. For example, one could observe that someone has bought the product, and *finding out* what could be – according to the model – the cause of that, is an abductive procedure. For example, according to our model, part of the explanation for observing that `p1` has bought the product can be that he has a `trust`-relation with `p2` and `p2` has bought the product. For any complex model, there could be very many possible explanations. Probabilistic models allow for a *ranking* of explanations, enabling to find the *most likely* one.

Inductive algorithms, which include *data mining*, are central to *profiling* and represent the means to get from data to models. Induction takes data and generates a model that *best fits the data*. Two broad categories exist. The first is *parameter learning*, which assumes all the rules mentioned in our example are present, except for the exact probability values. Now, by sophisticated counting, the probabilities can be estimated such that the resulting model "best fits" the statistics observed in the data. A broader inductive class contains algorithms that not only compute the probabilities, but construct *the rules themselves*. Learning then becomes a *search process* in the space of all possible rules. Naturally, many rules could be constructed and the real art of data mining and learning is to find the right set of rules that captures most of the data's variance. Whereas deduction and abduction are generally called *inference* techniques and deal with inferring aspects (of individuals), induction is referred to as *learning* and is targeted at generating or extending the model.

[12] Connecting data to the same identifier will become even easier soon when face recognition is employed at a wide scale.

2.4 The Many Faces of Bias

Machine learning with rich and probabilistic representations is still a form of *statistical reasoning* and therefore all known precautions stay in place (see also (Whyte, 2004) and (Reichmann, 1961)). Aspects such as representiveness of the data, proper data selection, highly improbable events, interpretation of correlations (Tversky and Kahneman, 1981), incomplete data and so on, are to be approached with care. Each right or wrong choice in dealing with data constitutes a *bias* which influences which predictions can be drawn from a model and with what level of validity. Note however, that biases are not necessarily bad; in fact without [13] bias nothing can be learned because all models would seem equally good. Bias in machine learning in the context of big data and profiling has many more facets, cf. (Crawford, 2011; Flach, 2012; van Otterlo, 2012).

In our example model, the most imminent form of bias is *representational*: what to represent, and in which language [14]. For profiling, the choice of which data to track of individuals, which generic relations to consider and how to induce generic patterns from data, are important biases in *generating* the models. However, also in the actual *use* of models, considerable bias is present: what to infer (i.e. induce partial algorithmic identities), from whom, when to do that, and, maybe also, why?

2.5 Counting Sheep: Personalization and Socialization

According to (Thurman and Schifferes, 2012)[p11] *"predicting accurately [...] at an individual level is a considerable technical challenge"*. Even though sheep are quite autonomous and mobile mammals, capable of having a personality, from the perspective of a shepherd and a dog, they are quite *predictable*. In artificial intelligence it is known that very often observable (group) behavior may seem complex, but a prediction model for an individual behavior may very well be simple (Braitenberg, 1984) [15]. This holds especially when these simple-minded individuals interact (Resnick, 1994) [16]. The same may hold for many prediction models used in the area of profiling: it is very probable that simple correlations (as in our example) are enough for the purposes of the model's maker (e.g. predicting whether someone will buy the product). Accurate predictions about an individual may be difficult indeed, but predicting *on average* in a large population is feasible.

In terms of such predictions, two seemingly opposite trends in profiling can be distinguished: *personalization* and *socialization*. The first uses more information about an individual when making predictions, the second uses more information about other people. Both represent a specific form of bias, but from a model's point of view they are very similar.

Personalization is using information about the particular individual to tailor predictions to that individual. Examples are Google's search results based on individual's cookies or GMail contents [17], or the use of geographic location (Kim *et al.*, 2011). Personalization happens more and more, and has received recognition through the *filter bubble* principle by Pariser (2011). From an individual's point of view, this style of *intellectual cocooning* can generate a feeling of information being tuned towards his particular interests, but from a model point of view not much is special. Personalization means that more information about the individual is used to make predictions. This only means that the models are becoming more complex and that more distinctions can be made. The fact that models can be made more personal is paradoxically due to some extent to the fact that there is more data available about people just like you.

The second trend, socialized search and collaborative filtering, is the use of data of *other* people

[13] For example, without bias, Google would not be able to function at all: no search engine can index the whole web, and not all hyperlinks are created equally (Vaidhyanathan, 2011)[p62,63].

[14] The exact semantics of the logical and probabilistic aspects can vary between formalisms cf. (De Raedt, 2008).

[15] For example, a robot following a moving light source.

[16] For example, modeling a realistic school of fish, or a flock of sheep, only three simple rules per individual are needed: i) don't bump into others, ii) try to match your speed with others, and iii) stay close to the group.

[17] See https://www.google.com/experimental/gmailfieldtrial

to color information. A simple example are the book recommendations one gets when shopping at Amazon, movie suggestions at IMDB, and Google's social search efforts. The "if you like this then you might like this too", and the "often bought together" are pieces of information about the book or movie you are currently viewing, based on the models built using opinions and click (and buy) behavior of many other people. Social data can be the target of profiling, but it can also help in the use of models, for example to do group recommendations (e.g. based on implicit consensus between web users). Again, from an individual's point of view, this may extend the variety of algorithmic identities one can be profiled as. Technically, from a model's point of view, not much changes. Rules in the model will now have conditions that mention other individuals (and their properties) in addition to other conditions that depend solely on the individual's data.

3. The Exploitation of Data and Models

So far we established that profiling deals with machine learning, data and models. The selling point towards individuals is that personalization and socialization are very useful (which is often true) However, since model building from data is a demanding task, there must be a good *reason* for doing it; usually either *profit* (*"what are the customer's real wishes?"*) and *surveillence* (e.g. *"spot the terrorist"*). We distinguish four core mechanisms of how technical tools are employed for profiling. Two of them were discussed in the previous sections, and dealt with using data to induce models and employing the model to infer new knowledge about individuals. A third one will be discussed in this section and is about *acting upon* models to optimize a performance criterum (such as profit).Acting upon models is less well studied so far (but see (Turow, 2011)). Section 4 will then discuss the combination of all previous forms in a fully automated fashion. At the end of this paper we explore how the use of models could be studied systematically.

3.1 The Use of Data and Models

Models and profiles are potentially influential when they are invoked. Much has been written about possible dangers for privacy when it comes to inferring new knowledge from profiles. As Anrig *et al.* (2008) put it: *"Profiles discovered by these techniques may not be anticipated, desired or politically correct but modern algorithms do not (yet) have the power to care"*. Imagine that I have a social network containing many individuals who have bought the product. Now, the chance that I am *not* influenced by any one of them (according to the model) becomes smaller as the number of people in my network increases. Due to this phenomenon, I could be placed in the *very-tempted-to-buy* category, without me knowing that, and maybe treated differently based on this categorization. I seem to have no control over how I am perceived (i.e. what my algorithmic identity is according to a particular model) since I do not know that there is such a model, how it looks like, when it is invoked, and when I am being treated differently because of how I am perceived through the model.

The influence an individual has about how and when profiles are being used is very limited. The same holds for control over your own data. Since models are built from data of *many* individuals, and your data is as good as your neighbors', often very few data items are needed to infer new knowledge. This renders many endeavours directed at protection, anonymization and even revocation of data relatively useless. All relevant (statistical) *knowledge* about individuals is already included in the models. One could say that the persons behind the data are forgotten through the model. Debates about opting out, removing personal data or porting it to another provider have therefore might only make sense if people do it on a massively collective scale.

3.2 The Value of Interfering with Models

Models enable attaching a probability to certain attributes, facts or complete situations. A first step into *modifying* this probability distribution to our advantage is by explicitly modeling certain factors we can *influence*. In our example, we assume that we can send an individual *targeted marketing* in order to influence that person's likelihood of buying the product. Let us assume that we learn from observed data that targeted marketing induces a 20 percent chance that a person buys the product, inducing the additional rule:

```
0.2 :: buy_marketing(Person).
```

Now, this creates additional possiblities to express (in the model) when somebody will buy the product. In this case that is the following rule saying that a person will buy the product (with probability 0.2) if he gets a targeted advertisement.

```
buys(Person) IF marketed(Person)
     WITH PROBABILITY buy_marketing(Person).
```

Together with our original model now two factors may influence someone's buying behavior, and we assume one of them is under our control. For that, we need to model explicit *decisions*:

```
IF person(Person) THEN marketed(Person) IS yes OR no.
```

This rule denotes whether we send whoever is filled in for the variable Person advertisements. Note that this single rule represents a possibly huge number of decisions; one for each individual.

Now, having in place the means to express statistical information and a way to influence the behavior of the population through decisions, we need additional elements to express aspects of "good" and "bad" decisions. The general way to do that is by supplying a *reward function* [18] (or, *cost function*). A reward function specifies for any particular state of the world how much (parts of) that state is "worth". In our example model, it is natural to specify that every time a person buys our product, we earn money, i.e. reward in this model is directly related to monetary value. On the other hand, targeted advertisements will *cost* money, and are represented by a negative reward.

```
IF person(P) AND buys(P) THEN REWARD IS 5.
IF person(P) AND marketed(P) THEN REWARD = -3.
```

Note that if all decisions about targeted advertisement are negative (i.e. as in the old setting) then people will still sometimes buy the product if their friends have done so, and we still make some profit (but we do not have any real costs).

For our example domain, it is natural to express costs and rewards in terms of money. However, in any domain some kind of costs may be specified. For example, in surveillence, the occurence of a bomb explosion is reflected as a huge negative cost (if it would occur), and the strip search of an innocent civilian should (at least) count as a small negative cost. In this way, the trade-off between security and inconveniences is explicitly stated in the model. In yet another domain one could trade-off the number of votes against the amount of flyers printed in an election scenario. In this way, privacy aspects can be monetized [19].

[18] See for a more elaborate discussion (van Otterlo, 2009).

[19] Recent work by Spiekerman on "privacy as property" shows how people are forced to value their privacy exactly (http://www.wu.ac.at/ec/faculty/spiek_pres_propertytielburg2012). If we extend the utopian literature connection in this paper even further we might end up in Ayn Rand's novels where literally all aspects of life are monetized.

3.3 Global Optimization: local consequences

Our new model explicitly represents negative and positive effects (rewards). However, because of the probabilistic aspects of the model, one can never say that a certain situation – in our case which specific people buy the product – *will happen*. What we can quantify is what the *probability* is with which a certain situation *can happen*, and in addition, we can quantify what *would be* the costs and profits for that situation. Thus we need to take probability into account and talk about *expected utility* [20]. It is defined in terms of (the total profits of a given situation minus the total costs of that situation) times the probability that this situation might occur. Taking the sum of all possible situations – no matter how unlikely – will produce *the* expected utility.

In our example domain, we can compute the expected value of a *policy*, i.e. a strategy for marketing a subset of the population. Let us assume that our strategy is:

```
[marketed(martijn), marketed(p4), marketed(p27), ...]
```

The decisions, i.e. who gets a targeted advertisement, cause one of many possible situations to occur. For each of them, we can compute the probability. For example, in some of them the person p4 will influence some other person, say p13, with probability 0.3 and in other situations he will not. Overall, the strategy's expected utility expresses the outcome of the tradeoff between costs for targeted advertisements and the expected amount of products sold.

The really interesting thing to do with this model is *finding the globally optimal strategy*, i.e. that strategy with the highest expected utility; a far from trivial task given the many possible strategies. In essence it is *abduction*: find a possible cause (i.e. a marketing strategy) for a situation in which it is very probable that we make a lot of profit. There are many types of bias at work: the best strategy is dependent on the population in the social network, the structure of that network, and all (probabilistic) rules in the model. In fact, we should emphasize that the most important bias in this type of systems is the reward function [21] (see also (van Otterlo, 2009; Wiering and van Otterlo, 2012)). By valuating things differently one gets different solutions of what to do.

An interesting phenomenon from a privacy-as-control point of view is that we have a second issue with the globality of models and the locality of an individual. The very fact that models are optimized at a global scale may very well make the consequences for an individual incomprehensible, meaningless, or even disruptibe. In our example model, the globally optimal strategy is a selection of people who get marketed. Now, depending on the social network relations, it might be optimal to target person p47 too because in the complicated probabilistic interactions that follow from applying the optimal strategy the most (expected) profit will be made. However, from the individual's point of view, it might not make much sense since this person might not have the intention at all to buy the product or to influence his friends in doing so. Whereas advertisements might be ignored, the very fact that decisions are made at a global scale that have consequences for what individuals observe or can do at a local scale is another example of the power imbalance of knowledge and privacy.

4. Towards Walden 3.0 – Automated Profiling

The tools described in the previous sections provide the building blocks for powerful surveillence machines, or advertisement machines, or even politial influence machines. The first tool is *model generation*, which takes observed data about a population and builds a general model. The second tool is *model use*, which takes a model (and an accompanying reward function) and computes the optimal way to exploit the information in the model to *act* upon individuals, for example by targeting

[20] The average number of eyes on a die is (1+2+3+4+5+6)/6 = 3.5 but you will never throw this number actually.

[21] This is equally important in computational work in reinforcement learning. For example; to teach a robot how to escape from a maze, we give it a small negative reward for every step and a big reward if it gets out. This will nudge the robot into learning to go to the exit as quickly as possible.

them with advertisements or by re-ranking their search results. Now, conceptually it is good to separate these tools, but in practice there is a constant dynamic of new data coming in, revised models being built, reward functions being adjusted, and actions being performed based on the new models. One could say that model generation takes data and delivers a model, whereas model use takes a model and causes new data to arise. The resulting *profiling loop* can be applied to any domain where data is digital and ubiquitous, and in the remainder of this section we point to some of them. Interestingly, this feedback loop can be driven using human intervention, but can also function in a fully automated [22] fashion driven purely by real-time algorithms. Intruigingly, Google and Facebook can do this with populations larger than most countries, and they typically do (Olsthoorn, 2012). For example, Google's search query completion is working while you type in what you are looking for.

It is here where the connection to the novel Walden II by B.F. Skinner, introduced at the start of this paper, is the strongest. Walden II depicts a small society in which principles of behaviorism are applied, and where the population is *conditioned* by manipulating how individuals get rewarded (or, to a lesser extent, punished). Good behavior and happy citizens are obtained not by force, but by manipulation. Automated, real-time profiling can form the basis of a very ambient, very covert form of *control* using lots of data and smart algorithms. Quoting Skinner (pp192–193): *"No, the potency of behavioral engineering can scarcely be overestimated. It makes one wonder why the techniques haven't been put to better use long before this. We could teach our children to be satisfied with a very limited and rigorous existence, to despise other forms of society, and to turn from the pleasures of the flesh. We might make such a society last for many years."*

Various forms of this *behavioral engineering* can now be implemented in a digital world. For starters, Google conditions us to accept and believe that a search results list does in fact deliver what we want, and Facebook users are nudged into filling their timeline, even about the period before they actually had an account. In our example model we have refrained from incorporating individual behavior as a consequence of rewarding actions so far, but more detailed models are possible in which individuals get "rewarded" and change their behavior due to that. It all depends on how the reward function of the model is set up and how individuals react to changes in their personal information environments. Again quoting Skinner: *"We aren't satisfied to produce merely a happy people. Our technology is powerful enough to make men happy under many conditions of life."*.

Whereas Walden II was published a long time before any form of digital surveillence was possible, the ideas of controlling a society based on positive reinforcements is very powerful. Based on the technology available, the emergence of a Walden 3.0 seems unavoidable. In the remainder of this section we discuss some technological developments in that direction, and some potential domains.

4.1 Automated Learning and Manipulation

Even though many services on the internet can be studied in terms of automated profiling (or adaptive hypermedia systems, cf.(Steichen *et al.*, 2012)), most are too large to study and inaccessible due to models and algorithms which are kept secret [23]. Our running example stems from recent work (Van den Broeck *et al.*, 2010) which dealt mainly with the abductive part and assumed rules were defined beforehand. The most insightful example of a fully automated system that does model learning *and* use, and is still compact enough to understand fully, is the *robot scientist* developed at Aberystwyth University in the UK (King *et al.*, 2004, 2009). This system uses a language that is very similar to our running example, and it automatically performs scientific experiments concerning the growth of yeast. The system has a current model, does abduction to come up with interesting experiments,

[22] Related phenomena are automated trading systems where most human intervention is removed, mainly due to the speed of acting upon the data. See a story about hypertrading and what can go wrong at BBC: http://www.bbc.co.uk/news/magazine-19214294

[23] Interestingly, the Dutch election system has gone back to the red pencil just because the algorithmics of voting machines was not transparent http://wijvertrouwenstemcomputersniet.nl

and can perform actual experiments on yeast samples. Based on observed behavior the system can learn additional rules in the model and continue with new experiments, incorporating the costs of experiments (i.e. a reward function). A mental step from yeast experimentation to data-driven social science is not too hard to make.

Another example of a general form of *experimentation* using big data comes from Google's *A/B-testing* facility (Christian, 2012) in Google Analytics [24]. It enables a designer of a website or advertisement to *test* possibly hundreds of different versions to a huge audience and to see how different versions trigger different reactions of individuals. For example, a store could experiment with which color website they sell the most products, and learn from their reactions. People are usually unaware that every day they may be placed in several web-based *Skinnerboxes* [25], and that for some part their interaction with information is controlled by covert experiments.

A more general development in automation will come from *probabilistic programming languages* (Poole, 2010) which are being developed in recent years. Unlike traditional programming languages, these new languages support learning and probabilistic inference as built-in capabilities. For example, one could write down parts of a model and the action strategy in a program (similar to our running example model) and probabilities (and rewards) are estimated automatically by the program from data. In other words, a programmer can write down parts of a program (a so-called *partial program*) after which automatically missing parts are filled in by the system itself. One of the interesting features is that such programs could even learn from trial-and-error; i.e. the program could experiment with some strategies and based on the feedback it gets of "how well" things are going according to the reward function definition, it may alter its behavior. This kind of learning is known in artificial intelligence as *reinforcement learning* (van Otterlo, 2009; Wiering and van Otterlo, 2012). Central ideas in this area came from, among others, Skinner himself, and are now used for many intelligent learning tasks, for example a robot making pancakes (Lockerd-Thomaz and Breazeal, 2008). In Skinner (1948)[pp243–244] words: *"It's what the science of behavior calls a 'reinforcement theory'. The things that can happen to us fall into three classes. To some things we are indifferent. Other things we like – we want them to happen, and we take steps to make them happen again. Still other things we don't like – we don't want them to happen and we take steps to get rid of them or keep them from happening again."* If click behavior Facebook *like*'s and other types of *feedback* are taken as *rewards*, by trial-and-error one can experiment with that program strategy that delivers a maximum amount of reward. The opposite is also possible, using *inverse reinforcement learning*: by studying what people do, one could guess *their* implicit reward function they are optimizing at a local scale. Knowing this will give additonal opportunities for global optimization and manipulation.

4.2 Skinnerboxes: Domains for Automated Manipulation

Currently, many digital domains exist in which profiling and experimentation happen, or where they are very likely to happen. Here we briefly point to several of them.

Viral marketing, Advertisements and Shopping Obviously, since our running example deals with viral marketing and targeted advertisements, such settings are ideal for profiling. From a modeler's point of view, the goal (and reward function) is clear, lots of data comes in from buying behavior, and models can be tested quite objectively. For example, Fong (2012) reports on how targeted marketing can influence search activity and willingness to try alternative products. Amazon and other web stores are probably the most visible occasions where we really notice that our data is being used, through recommendations and advertisements. The same goes for loyalty card systems in, for example, supermarkets. One could say that data has become so important that it has changed our view

[24] See http://www.google.com/websiteoptimizer/tutorials.html

[25] Skinner was famous for his experiments in a so-called *Skinnerbox*, a form of puzzle box in which animals were taught some specific skill, see http://en.wikipedia.org/wiki/Operant_conditioning_chamber

on consumer behavior tremendously. (Zwick and Knott, 2009)[p240]: *"We argue that the constant and compounding growth in the volume of data coupled with the rising analytical powers of computers has endowed the customer database with an immediate strategic importance in a company's economic value creation process"*.

Search Engines As Vaidhyanathan (2011)[p84] says: *"Google is a system of almost universal surveillance, yet it operates so quietly that at times it's hard to discern"*. Search engines, in particular Google, are becoming powerful profiling machines. They typically implement all profiling techniques discussed in this paper. Some are visible, such as Google's search query completion, whereas others are much more hidden, such as the influence of Gmail or Google+ circles on search results. An important aspect of search engines is that they function as *gatekeepers* (Introna and Nissenbaum, 2000; Granka, 2010) and can influence and color the information individuals get about the world. Very recently, Google has started [26] censoring so-called *torrent* links in its search results. This is being done under pressure concerning copyright infringements, but this may affect the visibility of other, legitimate content, thereby limiting access to information.

On the Workfloor A company can be seen as a small-sized society, and increasingly workers are being monitored through the use of surveillence cameras, tracking devices (in logistics), or e-mail. Recently, Microsoft applied for a patent [27] about what some call *human resource software from hell* [28]. The basic idea (from the first paragraph of the patent text) is to create *"[a] method to be executed at least in part in a computing device for monitoring, analyzing, and influencing organizational behavior"*. The main goal of the system is *"...and providing analysis results such that desired behaviors are encouraged and undesired behaviors are discouraged."*. How much more "Walden" can it get? Companies lend themselves very well for profiling systems because more control can be obtained in a limited setting where individuals are more dependent on the profiler and consequences of "bad" behavior can be more severe.

Policing and Surveillence Surveillence and security are big business, and governments of Western societies have a strong interest in monitoring and tracking. Using arguments ranging from copyright violations to child porn networks, civil liberties are being sacrificed for more (biometric) data, more cameras and more control. Lately several news items reporting on global surveillence systems [29] have appeared, for example on *Trapwire* [30], the *Domain Awareness System (DAS)* [31] by MicroSoft and New York city but many others exist. A key feature is that they are pervasive and covert.

Reading and News Nick Davies described in his popular bestseller *Flat Earth News* how journalism and reporting were strongly biased by the power of money and corporations. So, we might say that biased news is not new. However, the way in which news and information about the world in a digital world can be biased by personal preferences or by social structures is unprecedented. As an example, Messing and Westwood (2012) describe how social media change what we read and why. Thurman and Schifferes (2012) discusses the increased personalization of news recently and relates it to Pariser's filter bubble. Overall, the consequences of automated profiling on what people read

[26] See http://www.zdnet.com/au/google-takes-small-step-against-online-piracy-7000002490/

[27] The author was interviewed by Jolein de Rooij about this patent. This interview was published in Intermediar-PW, January 2012, at http://www.intermediairpw.nl/

[28] http://www.theregister.co.uk/2011/11/18/microsoft_patent_employee_monitoring/

[29] Hollywood versions of such surveillence systems can be found in (2011) *Person of Interest* (http://www.imdb.com/title/tt1839578/), (2008) *Eagle Eye* (http://www.imdb.com/title/tt1059786/) and (1998) *Enemy of the State* (http://www.imdb.com/title/tt0120660/).

[30] *"TrapWire is a unique, predictive software system designed to detect patterns indicative of terrorist attacks or criminal operations. Utilizing a proprietary, rules-based engine, TrapWire detects, analyzes and alerts on suspicious events as they are collected over periods of time and across multiple locations."*, see http://www.trapwire.com/trapwire.html

[31] See at Privacy S.O.S. http://privacysos.org/2012/august/NYPD-DAS-privacy-ethics

about the world could have tremendous consequences for society, which makes the news domain important to keep an eye on.

And Many More One area profiling machines can and will be employed in are elections, where reward functions optimize vote counts and influence. The influence of data, polls, social networks and search engines could be significant (Conti, 2008). Furthermore, moving to the physical world, it is expected that robots and drones will have a huge impact on data collection, privacy and control in the near future (Sharkey, 2008). These are just two additional examples next to the game industry, the entertainment industry (e.g. media consumption), cloud computing, and many, many more.

5. Discussion, Conclusions and Outlook

In this article we have taken steps towards knowing that, and understanding how, automated profiling and manipulation can happen in many places in our digital society. Through the use of a simple example case we have contributed to the development of *algorithmic literacy* which is needed for a full understanding. The interplay between data and models, and between representation, bias and algorithms are central to the rise of sophisticated prediction machines founded on the algorithmic basis of machine learning. We have identified several important issues, among which the global-local aspects of both the induction of models and the global optimization of models. What happens at the local, individual level in terms of manipulation is incomprehensible without having access to how at a global level information is exploited, and according to which reward function. Furthermore, we have highlighted that these new technologies bring us closer to a Walden 3.0 in which our reality is manipulated through the use of automated and autonomous profiling machines.

Now the question is how good or bad it is that automation, experimentation and manipulation happen in the digital world. In other words, is Walden 3.0 utopian or dystopian, or both? After all, there are huge benefits of personalization, social search and global optimization of search engines. Quoting Vaidhyanathan (2011)[p75]: *"Google never promised to be comfortable and benign: it just promised not to be evil, whatever that means"*. If we see privacy in terms of control over information, then there is growing power imbalance between an individual and prediction machines, and a severe loss in privacy. In the current, digital age, privacy is changing, see also (Schermer, 2011). The fact that the individual's life is determined by how algorithms categorize you and how they act upon that makes you want to know (at least) when that happens, how that happens, and why. Thus, in order to restore (some of) the balance in power, we need to figure out how we get there.

Many previous works have discussed solutions such as *privacy-preserving techniques*, *privacy-by-design*, *aggregation* of data, or various forms of *obfuscation* (Brunton and Nissenbaum, 2011). Whereas that may work in some contexts, they all are vulnerable and do not solve all problems, since data can increasingly be linked to other data. For example, anonymous data in the context of smart energy meters [32] might get deanonimized for people with public agendas (such as scientists); the Googled results can probably be matched against the anonymous living habit of some individual in the data. Other examples include the de-anonymisation of AOL data (Barbaro and Zeller Jr., 2006) and statistical prediction of social security numbers (Acquisti and Gross, 2009).

Among the many other remedies that exist (Schermer, 2011; Tocha *et al.*, 2012), demanding *transparency* (Hildebrandt and Koops, 2010) seems very natural; we would like to know how and when we are profiled. Two problems prevent transparency from being a general solution. The first is that governments and companies *own* the models and are not eager to share this information because of security or commercial reasons. The models are based on statistics *they* gathered, and no general (legal) mechanisms are in place to force them to share that information. The second problem is

[32] See this white paper on some other options for privacy and smart meters by George Danezis http://research.microsoft.com/en-us/projects/privacy_in_metering/privacytechnologyoptionsforsmartmetering.pdf

more severe: let us assume we do get all information that was relevant for generating your search results or selecting your news items for today, *how would we convey this information?*. Any reasonably sized model could produce many hard-to-understand explanations. In addition, the probabilistic information should be accompanied by the underlying statistical data, biases, settings of specific machine learning algorithms and so on. Such information would be impossible to understand for humans. An additional hindering factor would be that this information could have to be supplied to every user, at every webpage and so on, which would make browsing the web impossible.

In this paper we have presented one way of looking at the problem, but we have, by far, not yet even scratched the surface of a solution. Here I would like to argue that the best [33] way to cope with possible privacy erosion by profiling techniques is to study more in detail what happens and hopefully find how can we use other, but equally smart, artificial intelligence techniques to counter the power imbalance that is now developing. An interesting and vital research direction we are going to undertake is to study in more detail what the ramifications are of large-scale automated experimentation on the behavior of people. We need to find ways to determine how and whether it is this happening, and also what happens when people *adapt*, i.e. they change their behavior as a *reaction to* experimentation. I think that studying such data-driven social science settings will need input from behavioral science, from social psychologists and from computational theories, and in general be almost as hard as studying society itself or finding out whether you live inside the Matrix [34]. I do believe that algorithmic literacy and becoming aware of the things happening is the first step.

"You're getting paranoid, Carter. That's a step in the right direction" [35]

References

Acquisti, A. and Gross, R. (2009), Predicting Social Security Numbers from Public Data, *in: Proceedings of the National Academy of Science*, volume 106, pp. 10975–10980.

Andrews, L. (2011), *I know Who You Are and I Saw What You Did*, Free Press, New York.

Andrzejewksi, D., Stork, D. G., Zhu, X. and Spronk, R. (2010), Inferring Compositional Style in the Neo-Plastic Paintings of Piet Mondrian by Machine Learning, *in: Proceedings of Electronic Imaging: Computer Image Analysis in the Study of Art (SPIE)*.

Anrig, B., Browne, W. and Gasson, M. (2008), The Role of Algorithms in Profiling, *in:* Hildebrandt, M. and Gutwirth, S. (eds.), *Profiling the European Citizen: Cross-Disciplinary Perspectives*, chapter 4, Springer, pp. 65–87.

Barbaro, M. and Zeller Jr., T. (2006), A Face is Exposed for AOL Searcher No. 4417749, New York Times August 9.

Bollier, D. and Firestone, C. M. (2010), The Promise and Peril of Big Data, the Aspen Institute, ISBN 0-89843-516-1.

Braitenberg, V. (1984), *Vehicles: Experiments in Synthetic Psychology*, The MIT Press.

Brandimarte, L., Acquisti, A. and Loewenstein, G. (2009), Privacy Concerns and Information Disclosure: An Illusion of Control Hypothesis, *in: INFORMS Annual Meeting*.

Brunton, F. and Nissenbaum, H. (2011), Vernacular Resistance to Data Collection and Analysis: A Political Theory of Obfuscation, *First Monday*, volume 16(5).

Cheney-Lippold, J. (2011), A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control, *Theory, Culture & Society*, volume 28, pp. 164–181.

[33] The really best way is to avoid giving up personal data in the first place, but in our surveillence society this is virtually impossible.

[34] Interesting movies about the reality we live in and the effects of living inside a simulation are The Matrix (1999), The Thirteenth Floor (1999) and eXistenZ (1999).

[35] Quote from the television series *Person of Interest*, season 1, episode 12. Person of interest is about a huge *prediction machine* for surveillence.

Christian, B. (2012), The A/B Test: Inside the Technology that's Changing the Rules of Business, in Wired Magazine, 25th April, available at http://www.wired.com/business/2012/04/ff_abtesting/.

Claeys, G. (ed.) (2010), *The Cambridge Companion to Utopian Literature*, Cambridge University Press, UK.

Conti, G. (2008), Could Googling take down a president?, *Communications of the ACM*, volume 51(1), pp. 71–73.

Crawford, K. (2011), Six Provocations for Big Data, paper presented at Oxford Internet Institute's – A Decade in Internet Time – Symposium on the Dynamics of the Internet and Society (available electronically at http://ssrn.com/abstract=1926431).

De Raedt, L. (2008), *Logical and Relational Learning*, Springer.

de Vries, K. (2010), Identity, profiling algorithms and a world of ambient intelligence, *Ethics and Inf. Technol.*, volume 12(1), pp. 71–85.

Flach, P. (2012), *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*, Cambridge University Press.

Fong, N. M. (2012), Targeted Marketing and Customer Search, working paper series, Temple University Fox School of Business; MIT Sloan School of Management, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2097495.

Granka, L. A. (2010), The Politics of Search: A Decade Retrospective, *The Information Society*, volume 26(5), pp. 364–374.

Hildebrandt, M. and Gutwirth, S. (eds.) (2008), *Profiling the European Citizen: Cross-Disciplinary Perspectives*, Springer.

Hildebrandt, M. and Koops, B. J. (2010), The Challenges of Ambient Law and Legal Protection in the Profiling Era, *The Modern Law Review*, volume 73, pp. 428–460.

Introna, L. D. and Nissenbaum, H. (2000), Shaping The Web: Why the Politics of Search Engines Matters, *The information Society*, volume 15, pp. 169–185.

Kim, B., Ha, J.-Y., Lee, S., Kang, S., Lee, Y., Rhee, Y., Nachman, L. and Song, J. (2011), AdNext: a visit-pattern-aware mobile advertising system for urban commercial complexes, *in: Proceedings of the 12th Workshop on Mobile Computing Systems and Applications*, HotMobile '11, pp. 7–12.

King, R. D., Oliver, J. R. S. G., Young, M., Aubrey, W., Byrne, E., Liakata, M., Markham, M., Pir, P., Soldatova, L. N., Sparkes, A., Whelan, K. E. and Clare, A. (2009), The Automation of Science, *Science*, volume 324(5923), pp. 85–89.

King, R. D., Whelan, K. E., Jones, F. M., Reiser, P. G. K., Bryant, C. H., Muggleton, S. H., Kell, D. B. and Oliver, S. G. (2004), Functional genomic hypothesis generation and experimentation by a robot scientist, *Nature*, volume 427, pp. 247–252.

Koller, D. and Friedman, N. (2009), *Probabilistic Graphical Models*, MIT Press.

Lockerd-Thomaz, A. and Breazeal, C. (2008), Teachable robots: Understanding human teaching behavior to build more effective robot learners, *Artificial Intelligence*, volume 172(6-7), pp. 716–737.

Messing, S. and Westwood, S. J. (2012), How Social Media Introduces Biases in Selecting and Processing News Content, working paper, available at http://www.stanford.edu/ seanjw/papers/SMH.pdf.

Nilsson, N. J. (2010), *The Quest for Artificial Intelligence*, Cambridge University Press.

Olsthoorn, P. (2012), De Almachtige Zoekmachine: Google van Online TomTom naar Personal Coach, *in:* van 't Hof, C., Timmer, J. and van Est, R. (eds.), *Voorgeprogammeerd: Hoe Internet ons Leven Leidt*, Rathenau | Boom Lemma uitgevers, in Dutch.

Pariser, E. (2011), *The Filter Bubble*, Viking (Penguin Books), Great Britain.

Poole, D. (2010), Probabilistic Programming Languages: Independent Choices and Deterministic

Systems, *in:* Dechter, R., Geffner, H. and Halpern, J. (eds.), *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, College Publications, pp. 253–269.

Reichmann, W. J. (1961), *Use and Abuse of Statistics*, Methuen London.

Resnick, M. (1994), *Turtles, Termites, and Traffic Jams*, MIT Press.

Schermer, B. W. (2011), The Limits of Privacy in Automated Profiling and Data Mining, *Computer Law & Security Review*, volume 27, pp. 45–52.

Sharkey, N. (2008), 2084: Big robot is watching you – Report on the future of robots for policing, surveillance and security, available at http://staffwww.dcs.shef.ac.uk/people/N.Sharkey/Future robot policing report Final.doc.

Skinner, B. F. (1948), *Walden Two*, Reprinted 2005, Hacket Publishing Company Inc.

Steichen, B., Ashman, H. and Wade, V. (2012), A Comparative Survey of Personalised Information Retrieval and Adaptive Hypermedia Techniques, *Information Processing and Management*, volume 48(4), pp. 698–724.

Tene, O. (2011), Privacy: The New Generations, *Int. Data Privacy Law*, volume 1(1), pp. 15–27.

Thurman, N. and Schifferes, S. (2012), The Future of Personalization at News Websites, *Journalism Studies*, forthcoming, DOI:10.1080/1461670X.2012.664341.

Tocha, E., Wang, Y. and Cranor, L. F. (2012), Personalization and Privacy: A Survey of Privacy Risks and Remedies in Personalization-based Systems, *User Model User-Adap Inter*, volume 22, pp. 203–220.

Turow, J. (2011), *The Daily You*, Yale University Press, New Haven and London.

Tversky, A. and Kahneman, D. (1981), The Framing of Decisions and the Psychology of Choice, *Science*, volume 211(4481), pp. 453–458.

Vaidhyanathan, S. (2011), *The Googlization of Everything*, University of California Press, Los Angeles.

Van den Broeck, G., Thon, I., van Otterlo, M. and De Raedt, L. (2010), DTProbLog: A Decision-Theoretic Probabilistic Prolog, *in: Proceedings of the National Conference on Artificial Intelligence (AAAI)*.

van Otterlo, M. (2009), *The Logic of Adaptive Behavior*, IOS Press, Amsterdam, The Netherlands.

——— (2012), A Machine Learning View on Profiling, *in:* Hildebrandt, M. and de Vries, K. (eds.), *Privacy, Due process and the Computational Turn – Philosophers of Law meet Philosophers of Technology*, Routledge, in press.

van 't Hof, C., Timmer, J. and van Est, R. (eds.) (2012), *Voorgeprogammeerd: Hoe Internet ons Leven Leidt*, Rathenau | Boom Lemma uitgevers, in Dutch.

Whyte, J. (2004), *Crimes Against Logic*, McGraw-Hill.

Wiering, M. A. and van Otterlo, M. (2012), *Reinforcement Learning: State-of-the-Art*, Springer.

Yang, Q. (2009), Activity Recognition: Linking Low-level Sensors to High-level Intelligence, *in: IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence*.

Zwick, D. and Knott, J. D. (2009), Manufacturing Customers: The Database as New Means of Production, *Journal of Consumer Culture*, volume 9, pp. 221–247.