



# How Information and Communication Technologies Expose Our Deep Personal Privacy

Yair Oppenheim

The Lester and Sally Antin Faculty of Humanities, Tel Aviv University, Tel Aviv, Israel  
Email: yairoppen@gmail.com

**How to cite this paper:** Oppenheim, Y. (2025) How Information and Communication Technologies Expose Our Deep Personal Privacy. *Open Access Library Journal*, 12: e13066.

<https://doi.org/10.4236/oalib.1113066>

**Received:** February 11, 2025

**Accepted:** April 27, 2025

**Published:** April 30, 2025

Copyright © 2025 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Personal privacy, in various aspects, has occupied humanity since the dawn of civilization, evolving and changing throughout human history. For the last two decades, the issue of privacy has constantly been at the center of public debate. Privacy, particularly various aspects of personal privacy<sup>1</sup>, persistently remains the focus of public attention. The growing preoccupation with personal privacy stems from the great changes [1] that information and communication technologies (ICTs) have brought about in all walks of life during the last twenty years. AI (Artificial intelligence) tools have developed as part of a revolution in which ICTs. These tools have capabilities that were not previously known to violate personal privacy by replacing the discussion on personal privacy with discussion on deep and general personal privacy information. I show how AI tool (Decision Tree tool) could be used to expose to the public parts of personal privacy. In this study, I show how it is possible, by analyzing a large database and using Data Mining methods to expose feelings of trust or mistrust. Trust is an emotion that is found in the field of Deep Personal Privacy. Emotions were undiscoverable before the age of ICT and AI. The numerical example presented in this article shows that using an AI tool (Decision Tree) enables the user to expose (prediction) the user's trust (mistrust) in a computerized support system with an accuracy of 68.68%

## Subject Areas

Complex Network Models

<sup>1</sup>Personal privacy is the type of privacy that applies to individual humans, in contrast to general privacy, which also applies to organizations, corporations and agencies.

---

## Keywords

Personal Privacy, Deep and General Personal Privacy, Popper's world, Entropy, ICTs, Decision Tree

---

## 1. Introduction

Personal privacy, in various aspects, has occupied humanity since the dawn of civilization, evolving and changing throughout human history.

Many attempts have been made to redefine personal privacy. However, as Daniel Solove shows [2] [3], none of these definitions is complete and exhaustive in describing the phenomenon of personal privacy [4]-[7]. Several questions remain unclear: What is privacy as a concept? Where and to what does it apply? Where does the private sphere end and the public sphere begin? Where does privacy stand in relation to other concepts such as freedom, autonomy and liberty? How do we balance individual human rights against the public's right to know? Must a state reveal potential threats to its security? And how should it all be regulated in this age of information and communication technologies that follow us wherever we go and whatever we do? The failure to define personal privacy makes it very difficult to formulate social norms or legal regulations that would protect it from being violated by individuals, corporations, social organizations and governments in this era of the information revolution.

The information revolution has made it possible to translate personal privacy into personal privacy information and, consequently, to discuss personal privacy in terms of personal privacy information, too. In this article, I **replace the discussion of personal privacy with the discussion of personal privacy information** and offer a new distinction—between deep personal privacy information and general personal privacy information. **Deep personal privacy information** is information about you that nobody except you knows—in other words, the knowledge you have about yourself minus the knowledge the world (or society) has about you. **General personal privacy information** is information about you that is shared by you and your confidants, but not by the rest of the world (including other individuals, databases, enterprises and organizations). Confidants have an understandable commitment to keep that information in the private sphere and not to disclose it publicly—or they would not be confidants. A confidant may be an individual—e.g., a doctor, a psychotherapist or a family member—or an institution, e.g., an insurance company or a bank.

To help me draw a clear line between the two types of personal privacy and map their place in the universe, I have been using Karl Popper's theory of "three worlds" [8]: world 1—the physical world of natural objects and processes; world 2—the mental or psychological world; and world 3—the realm of the products of human thought.

Their main characteristics are summarized in **Table 1**:

**Table 1.** Popper's world's main characteristics.

	Contents	Status	Examples
World 1	Physical objects	Objective	Table, chair, steel
World 2	Mental experiences	Subjective	Pain, joy
World 3	Knowledge	Objective	Mathematics, biology

Based on this, I place **deep personal privacy** in world 2, and **general personal privacy** in worlds 1 and 3. This distinction allows us to clearly delimit **deep personal privacy** and argue that until the age of ICTs deep personal privacy had been inaccessible, in contrast to general personal privacy, which is culture-dependent and may be negotiated. Moreover, it shows us that while ICTs are affecting both types of privacy, threatening to expose them and make them public, **the greatest change of our time is that ICTs are allowing, for the first time in history, the exposure of deep personal privacy**. In the next sections, I will demonstrate how it can be done.

## 2. How ICTs Contribute to the Exposure of Deep Personal Privacy

The term “information and communication technologies” (“ICTs”) includes big data pools, the Internet of things (IoT), the web, and generative artificial intelligence (AI). The said ICTs aggregate, analyze, fuse and disseminate information on every individual who is active online, using abilities that did not exist until twenty years ago.

**Table 2** maps the various ICTs according to Popper's three worlds theory:

ICTs monitor users' activity and collect their personal information. These capabilities are being exploited all over the world by government agencies [9] as well as by commercial companies using commercial applications. Personal information relating to every aspect of our lives is being collected, stored, analyzed and shared as part of this surveillance process. The collected data undergo “profiling” [10]: data from different sources are combined to create a data pool about an individual, an organization or a physical or Internet entity<sup>2</sup>; all that data are processed using various algorithms, including deep learning, to make a profile of the said individual or entity, which is then used to predict various behavior patterns

**Table 2.** ICT's ingredient mapping to Popper's worlds.

	Contents	Examples related to ICTs and personal privacy
World 1	Physical objects	ICT hardware (servers, communication lines, IoT devices...)
World 2	Mental experiences	Personal information, personal thoughts, personal decisions, personal emotions...
World 3	Knowledge	Computer sciences, ICT software (AI applications, computer programs, social networks, the web, cloud technology, generative AI), personal information

<sup>2</sup>An Internet entity is a virtual entity, such as a Facebook user, while a physical entity is a real-life person or machine.

such as commercial preferences, social habits, even political opinions, with a high level of accuracy. After the information is processed, it becomes a commodity to be traded between various consumers of information, government and private alike. Such processing of our personal information, which is constantly being collected and stored in big data clouds, by means of AI models (mainly deep learning methods) results in increasing exposure of **deep personal privacy** from world 2 and **general personal privacy** from worlds 1 and 3 (by ICTs that reside in worlds 1 and 3). The main types of AI tools used for this purpose are logistic regression, decision trees, and neural networks.

### 3. A Numerical Example of Using AI to Expose Deep Personal Privacy

In this section, I will present a model (example) with real data to demonstrate that it is possible to expose parts of deep personal privacy using AI tools. The example below shows how it is possible to identify an emotion such as trust or distrust in a computerized customer support system. The AI tools I used were **decision trees** and **neural networks**.

A **Decision Tree** is a popular supervised learning algorithm used for classification and regression tasks. It works by recursively splitting the dataset into subsets based on the **most** significant feature that results in the greatest information gain (for classification) or the least variance (for regression). The final output is a tree-like structure where each internal node represents a decision based on a feature, each branch represents the outcome of that decision, and each leaf node represents a class label (for classification) or a predicted value (for regression).

#### Decision Tree Construction [11]

Classification Example:

##### 1. Choose a Feature and Threshold:

- The algorithm evaluates each feature and its possible values (thresholds) to split the dataset. It calculates a score based on **Gini Impurity** or **Information Gain**.

- **Gini Impurity:** Measures how often a randomly chosen element would be incorrectly classified.

(1)  $\text{Gini}(D) = 1 - \sum_i P_i^2$  . where  $p_i$  is the probability of an element belonging to class  $i$ .

- **Information Gain:** Measures the reduction in entropy from a split.

$$(2) \text{IG}(D,A) = H(D) - \sum_{v \in \text{values}} \frac{|D_v|}{|D|} * H(D_v)$$

where  $H(D)$  is the entropy of the dataset  $D$ , and  $D_v$  is the subset of data after splitting on feature  $A$ .

- **Split the Dataset:** After finding the best feature and threshold, split the dataset into two subsets based on this decision.

##### 2. Repeat for Each Subset:

- Apply the same process recursively to each subset until no more splitting is

possible (or until reaching a pre-defined depth or leaf size).

### 3. Assign Class Labels:

- In the leaf nodes, assign a class label based on the majority class within that node.

#### Description of the computerized support system that collects the data

A telecommunications company provides online support services through a computerized support system. The purpose of the system is to resolve technical issues raised by customers: the customer asks questions, and the system responds. The length of the conversation is not limited in time, and there is no restriction on the number of questions that may be asked or on the wording of the questions. The conversation is communicated in natural language.

Every customer request is handled as part of a call (session) that can be completed positively—that is, the customer receives an answer from the system. Either the session ends in a negative way—that is, the client switches to human support or abandons the system without receiving an answer. The system fully records all sessions. (See **Table 3**)

The system has hundreds of users and receives hundreds of calls every day, all of which are fully recorded in real-time. The data discussed here is a dataset of the presented test case.

**Table 3.** An example of a typical session.<sup>3</sup>

2/1/2016	11:33:04	[10540A0DF683000F]	Incoming	query:	How
2/1/2016	11:33:04	[10540A0DF683000F]	Query	Keywords:	1 how
2/1/2016	11:33:10	[10540A0DF683000F]	Incoming	query:	How do
2/1/2016	11:33:10	[10540A0DF683000F]	Query	Keywords:	1 “how 1” do
2/1/2016	11:33:11	[10540A0DF683000F]	Incoming	query	How do
2/1/2016	11:33:11	[10540A0DF683000F]	Query	Keywords:	1 “how 1” “do 1”
2/1/2016	11:33:12	[10540A0DF683000F]	Incoming	query	How do keep
2/1/2016	11:33:12	[10540A0DF683000F]	Query	Keywords:	1 “how 1” “do 1” “i 1” keep
2/1/2016	11:33:17	[10540A0DF683000F]	Incoming	query	How do keep my existing
2/1/2016	11:33:17	[10540A0DF683000F]	Query	Keywords:	1 “how 1” “do 1” “i 1” “keep 1” “my 1” existing exist
2/1/2016	11:33:22	[10540A0DF683000F]	Incoming	query	How do keep my
2/1/2016	11:33:22	[10540A0DF683000F]	Query	Keywords:	1 “how 1” “do 1” “i 1” “keep 1” “my 1”
2/1/2016	11:33:26	[10540A0DF683000F]	Incoming	query	How do keep my old number
2/1/2016	11:33:26	[10540A0DF683000F]	Query	Keywords:	1 “how 1” “do 1” “i 1” “keep 1” “my 1” “old 1” number number (vodafoneUK: English) How do keep my old “number?”—48187427 Can keep my “number?”—38915574 How do transfer my Vodafone number to new SIM?
2/1/2016	11:39:03	[10540A0DF683000F]	Query	Counted answered	

<sup>3</sup>Fles The Original Row Material and the row material after conversion process.

The population of users can be divided into two groups:

{A}—Users who continue in the session until its positive end. {B}—Users who abandon the session and cause it to end negatively.

Our goal in this example was to show how AI tools (Decision Tree) can help us predict which group (A or B) the user belongs to in any given session. Implicitly, it is a measurement of a given user's level of trust in the system in each session.

There are two types of user trust in a support system: the first is general trust that the system will help you resolve technical issues, while the second is trust in the system's ability to help you solve a specific problem. During the session the trust between the user that request technical help and the system that supposed to provide resolutions. If you prefer to continue the conversation until you receive a satisfactory answer, it means you trust in the system. If you abandon the conversation before receiving a satisfactory answer, it means you have lost faith in the system's ability to give you one. The model presented here measured the second type of user trust.

The raw database<sup>4</sup> had about 10 log files in Excel format registered by the system in real-time, each file containing hundreds of sessions. From this raw database, two files containing 2503 sessions in total were selected for the system training stage. A third file containing 1178 sessions was selected for auditing and determining the rate of successful predictions.

#### Infrastructure data used for the decision trees and neural networks

To measure the degree of a user's trust in the system, I defined several attributes which, in my view, reflected at any given moment the degree of the user's belief that the session will end positively (level of trust). These attributes (characteristics) were calculated from the basic information contained in the log files and divided into categories as follows (see **Table 4**):

**Table 4.** Tree decision attributes.

Category	Attribute	Attribute justification
Session IDs	ID	IDs of the specific session
	Date	
	Timestamp	
Session metadata	Number of lines in the session	Assumption that a longer session reflects greater trust in a positive ending
	Session length (time)	
User behavior	Average typing speed	Reflects a user's level of confidence in what they want to say
	Changing the topic during the session	Reflects a lack of user focus and therefore increases chance of failure
	Number of characters in the shortest line	Reflects user focus and a good definition of the problem, and therefore a better chance of the system giving an appropriate answer
	Number of characters in the longest line	Opposite of the previous attribute
	Average number of characters per line	Reflects user focus and a good definition of the problem, and therefore a better chance of the system giving an appropriate answer
Success test	Whether the session ended positively or not	

<sup>4</sup>The data in this example is real data from the telecommunications company.

The attributes were converted to values of {0,1}. **The method conversion for the rest of the attributes is:**

If Session's ended positively assign 1 else, assign 0

If Session's Number of lines in the session  $\geq 10$ , we convert the attribute value to 1 else, convert it to 0.

If the Session's length in milliseconds  $\geq 500,000$ , we convert the attribute value to 1, or else convert to 0.

If Session's Number of topics changes during the session  $\geq 1$ , we convert the attribute value to 1, or else convert it to 0.

If Session's Number of characters in the longest line  $\geq 50$ , we convert the attribute value to 1 else, we convert it to 0.

If Session's Number of characters in the shortest line  $\geq 30$ , we convert the attribute value to 1 else, convert it to 0.

If Session's Typing speed—characters per second  $\geq 1$ , we convert the attribute value to 1, or else convert it to 0.

If Session's Number of characters in the first line  $< 20$ , we convert the attribute value to 1 else, we convert it to 0.

This process enables binary tree searching. (See **Table 5**)

**Table 5.** Summary attributes converting.

Attribute	Positive/high/long value = 1	Negative/low/short value = 0
Session has ended positively or negatively (Positive Y/N)	Y	N
Number of lines in the session	$\geq 10$	$< 10$
Session length in milliseconds	$\geq 500,000$	$< 500,000$
Number of topic changes during the session	$\geq 1$	0
Number of characters in the longest line	$\geq 50$	$< 50$
Number of characters in the shortest line	$\geq 30$	$< 30$
Typing speed—characters per second	$\geq 1$	$< 1$
Number of characters in the first line	$< 20$	$\geq 20$

The decision tree was built according to the Decision Tree Approach using the ID3 [12] algorithm, which uses the entropy formula to calculate the information gain at each stage:

(3)  $IG(Y, X) = H(Y) - H(Y/X) - IG(Y, X)$  is the information gain of concrete stage.

Running an algorithm of the decision tree on the training-providing binary data. Gave values of {0, 1} per leaf in the tree. In accordance with these results, 87 vectors with binary values were built. A characteristic structure of this vector is in **Table 6**.

**Table 6.** Decision rule vectors.

Amount of lines for the session	Session time in milliseconds	The number of times the subject of the session changes	The number of characters in the longest line	The number of characters in the shortest line	The number of characters in the first row of the session	Typing speed	A call ended positively or negatively
0	0	0	1	0	1	1	0

Based on these decision-making vectors, the following algorithm was applied to predict whether a given session will end up in a positive or negative way.

For each session in the control file:

- Compare the row against each of the decision rule vectors
- For all such comparisons, calculate the distance between the decision vector and the session file data. Calculate the distance between the row and the vector as follows:

$$(4) d_n = \left| \sum_{i=1}^7 (x_i - y_i) \right|$$

where  $n$  is the number of the row in the file,  $x_i$  is the value of the characteristic  $i$  in the session line, and  $y_i$  is the value of the characteristic  $i$  in the decision vector.  $d_n$  is the distance between the vectors.

- Select the vector whose distance is the minimum from the row under test. Based on this vector, the system predicts whether the session ends positively or negatively.

The results of the decision tree model used to predict the users' trust are summarized in **Table 7**:

**Table 7.** Trust prediction.

User group	Positive system prediction (true positive (TP) or false positive (FP))	Negative system prediction (true negative (TN) or false negative (FN))	Actual number in group	Precision
A	489	254	743	65.81%
B	320	115	435	73.56%
Total	809	369	1178	68.68%

Based on these results, we get:

$$(5) \text{Recall}^5 = \text{True Positives} / (\text{True Positives} + \text{False Negatives}) = 489 / (489 + 115) = \mathbf{80.96\%}$$

$$(6) \text{F1}^6 = (\text{Precision} * \text{Recall} / (\text{Precision} + \text{Recall})) * 2 = (80.96\% * 68.68\% / (80.96\% + 68.68\%)) * 2 = \mathbf{74.31\%}$$

For supporting the above results, I ran a neural network model with the same data. The neural network characteristics were determined according to empirical

<sup>5</sup>A measure of the performance of a machine learning model, especially in cases of binary classification problems (such as spam detection). Recall describes the ratio between the number of correct cases identified by the model and the total number of correct cases.

<sup>6</sup>F1 score is the harmonic mean of Precision and Recall.

rule:

For the first hidden layer I use the rule:

$$(7) N_h = 2 * (N_i - N_o)$$

$N_h$  = number of neurons in a hidden layer

$N_i$  = number of input neurons—7 in our case

$N_o$  = number of output neurons—1 in our case

The second hidden layer should be **smaller** than the first to create a hierarchical representation, and **7 nodes** are chosen to be smaller than the first layer, ensuring a reduction in dimensionality and preventing overfitting.

These are the reasons for a fully connected network with three layers:

- First hidden layer with 12 nodes
- Second hidden layer with 7 nodes
- An output layer with one node
- Rectified linear activation function for the first two layers and Sigmoid for the output layer.
- Choosing the **right number of epochs and batch size** is crucial for achieving good model performance while avoiding underfitting or overfitting. Since our database is medium size (table of  $1178 \times 8$ ) I choose after trial and error 150 epochs and a batch size of 10.<sup>7</sup>
- For medium and noisy databases, the recommended batch size is  $\sim 16$ . I choose a batch size of 10 after trial and error.

**A general accuracy level of 69.97% was obtained with this parameter.**

The model presented here (the decision tree) measures the level of a given user's trust in the system in each session, with sessions group {A} representing trust in the system's ability to solve a problem, and sessions group {B} representing a lack of trust in the system's ability to solve a problem. The results of both models confirm that AI tools have the ability to detect the level of user trust of the second type in real-time with a general success rate of  $>68\%$ . This is an exposure of an emotion that lies within the realm of deep privacy and has not been reported by the individual to the public or to any confidant. We can also see that it is naturally easier to reveal a feeling of distrust (73.56% probability) than of trust (65.81% probability) because it is easier to destroy trust than to build it, and because negative emotions probably have indirect expressions that are more powerful than in positive emotions<sup>8</sup>.

#### 4. The Current Paradigm of Personal Privacy Has Failed, and There Is Need for a New One

The current paradigm of personal privacy is founded on the basic principles of

<sup>7</sup>**Epoch:** One pass through all the rows in the training dataset.

**Batch size:** The number of samples processed by the model within an epoch before weights are updated.

<sup>8</sup>Polygraph ("lie detector") tests are based on the very similar assumption that when a person is lying, there are noticeable and measurable changes in physiological indicators such as blood pressure, perspiration, etc.

**consent** and **control**. The guiding legal principle is that you should give express consent to the collection of information about you. For this reason, Internet companies are required to obtain customers' advance consent (express or implied) to the collection of their information by means of cookies or otherwise. Such consent should be conscious, *i.e.*, you should be aware that you are giving your consent. [13]. The requirement for consent also includes the second principle of control over information, because by consenting or not consenting, you also choose which categories of information about you may or may not be collected.

According to most applicable laws, consent for information collection is not sweeping, but limited to those categories of personal information that the data subject has agreed to share with the data controller (the entity that collects the information), and to the strict purpose for which the information is being collected. In other words, the use of the information is limited, and there is some element of control over the information. The information must be obtained by fair and reasonable means, with the knowledge and consent of the data subject. The basic principles of consent and control include the element of solitude, which is a manifestation of the right to be let alone; the right to access one's personal information stored in any database and to ask for its rectification or deletion; the element of secrecy and confidentiality; and the accountability of the data controller for complying with these principles and maintaining these rights.

However, the basic assumption that personal information privacy can be protected by regulations and social norms that are based on consent and control is failing for the following reasons:

- AI-based algorithms profile users based on their gender, race and other categories, violating the principle of equal treatment. They are basically constructing and analyzing statistical virtual identities that do not necessarily match people's actual identities, taking control over one's identity out of one's hands and labelling people without their consent. This entire process is done without the users' consent and is beyond their control.
- IoT devices used in "smart cities" and processors built into more and more everyday appliances such as refrigerators, air conditioners and cars usually lack a human-machine interface through which one could express consent or non-consent to the collection of personal information. As a result, people's information is collected without their expressed consent and without any ability to control which information will be collected, or how and by whom it will be used.
- The various databases in which personal information is stored are interconnected. People do not provide consent to that—whether expressed or implied, nor can they control how and where their personal data will be stored, or how it will be processed and disseminated. The purposes for which our personal information is used are beyond our control, too, and we are not asked to provide consent for any such use.
- Devices that directly interface with our body a priori bypass the conscious

mechanisms of decision-making, so there can be neither control over the personal data nor consent to provide it.

- The scope of information available in the age of ICTs is too vast for people to make informed choices about which of their data they would like to control, and which they don't mind being collected. According to McDonald and Cranor's study, people do not read privacy policies, and if they did, the annual cost to the US market would be \$781 billion. [14].
- To sum up, **privacy regulation that relies on consent and control has proven impractical and unenforceable.** [15] The reason for its failure is over-abundance of information. People do not know all the information that is being collected about them—nor do they want to know. They do not understand all the consequences of information collection—nor do they want to. They do not have a real choice—or the will to choose—regarding the information that will be collected. As a result, people are making wrong decisions regarding the privacy of their personal information and selling it cheap—or giving it out for free—to Internet companies. People have no actual control over their personal information. The mechanisms of consent and control are being protected, but the actual information privacy is not.<sup>9</sup>

The conclusion from the above is that the current personal privacy paradigm has failed, and there is a need for a new paradigm of personal privacy. It seems to be impossible to protect **general personal privacy**, because this type of privacy has become a commercial product for the big Internet companies, and in capitalist regimes, economic interests always prevail over regulations and social norms. The governments themselves also have a strong incentive to use people's personal privacy information to enhance their control over their citizens. However, **deep personal privacy** is a fundamental human need that must be protected. This proves there is a crucial need for a new paradigm for protecting **deep personal privacy**.

## 5. Conclusion

The information revolution is creating a new civilization, in which both **general personal privacy** and **deep personal privacy** are constantly violated. In this article, I have demonstrated information and communication technologies' ability to expose personal privacy information on every level, particularly their unprecedented ability to expose deep personal privacy (e.g. our feelings, thoughts and desires). The existing legal regulations and social norms fail to protect personal privacy on all its levels. They fail to deal with the market forces that have strong interests in exploiting people's personal privacy, including Internet companies that use personal privacy information to maximize their profits, and government organizations that use privacy information to enhance their control and supervision over their citizens. This proves there is an urgent need for a new paradigm that will help protect deep personal privacy, for without personal privacy, society cannot exist.

<sup>9</sup>Government and regulatory agencies impose severe fines on Internet companies, but it does not help protect information privacy.

## Conflicts of Interest

The author declares no conflicts of interest.

## References

- [1] Toffler, A. (1980) *The Third Wave*. Morrow, 17.
- [2] Solove, D.J. (2009) *Understanding Privacy*. Harvard University Press, 1-2.
- [3] Nissenbaum, H. (2009) *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, 16. <https://doi.org/10.1515/9780804772891>
- [4] Schoeman, F.D. (1984) *Privacy: Philosophical Dimensions*. In: Schoeman, F.D., Ed., *Philosophical Dimensions of Privacy: An Anthology*, Cambridge University Press, 3.
- [5] Laurie, G. (2002) *Genetic Privacy: A Challenge to Medico-Legal Norms*. Cambridge University Press, 6. <https://doi.org/10.1017/cbo9780511495342>
- [6] Hongladarom, S. (2016) *A Buddhist Theory of Privacy*. Springer, 16.
- [7] Gavison, R. (1980) *Privacy and the Limits of Law*. *The Yale Law Journal*, **89**, 421-471. <https://doi.org/10.2307/795891>
- [8] Popper, K. (1978) *Three Worlds—The Tanner Lecture on Human Values*. University of Michigan. <https://tannerlectures.utah.edu/resources/documents/a-to-z/p/popper80.pdf>
- [9] Calderone, M. (2013) *Washington Post Began PRISM Story Three Weeks Ago, Heard Guardian's 'Footsteps'*. HUFFPOST. [https://www.huffpost.com/entry/washington-post-prism-guardian\\_n\\_3402883](https://www.huffpost.com/entry/washington-post-prism-guardian_n_3402883)
- [10] Hildebrandt, M. (2008) *Defining Profiling: A New Type of Knowledge?* In: Hildebrandt, M. and Gutwirth, S., Eds., *Profiling the European Citizen*, Springer, 17-45. [https://doi.org/10.1007/978-1-4020-6914-7\\_2](https://doi.org/10.1007/978-1-4020-6914-7_2)
- [11] Rokach, L. and Maimon, O. (2013) *Data Mining with Decision Trees*. 2nd Edition, World Scientific. <https://doi.org/10.1142/9097>
- [12] Quinlan, J.R. (1986) *Induction of Decision Trees*. *Machine Learning*, **1**, 81-106. <https://doi.org/10.1007/bf00116251>
- [13] European Parliament and of the Council (2016) *Regulation (EU) 2016/679 of the European Parliament and of the Council*.
- [14] McDonald, A. and Cranor, L.F. (2008) *The Cost of Reading Privacy Policies*. *A Journal of Law and Policy for the Information Society*, **4**, 543-568.
- [15] Kagal, L. and Abelson, H. (2010) *Access Control Is an Inadequate Framework for Privacy Protection*. MIT Computer Science and Artificial Intelligence Lab.