Legal Knowledge and Information Systems
S. Villata et al. (Eds.)

doi:10.3233/FAIA200846

A Taxonomy for the Representation of Privacy and Data Control Signals

Kartik CHAWLA [a,1], Joris HULSTIJN [a]
[a] *Department of Information Management, TiSEM, Tilburg University*

Abstract. In interacting with digital apps and services, users create digital identities and generate massive amounts of associated personal data. The relationship between the user and the service provider in such cases is, *inter alia*, a principal-agent relationship governed by a 'contract'. This contract is provided mostly in natural language text, however, and remains opaque to users. The need of the hour is multi-faceted documentation represented in machine-readable, natural language and graphical formats, to enable tools such as smart contracts and privacy assistants which could assist users in negotiating and monitoring agreements.

In this paper, we develop a Taxonomy for the Representation of Privacy and Data Control Signals. We focus on 'signals' because they play a crucial role in communicating how a service provider distinguishes itself in a market. We follow the methodology for developing taxonomies proposed by Nickerson et al. We start with a grounded analysis of the documentation of four smartphone-based fitness activity trackers, and compare these to insights from literature. We present the results of the first two iterations of the design cycle. Validation shows that the Taxonomy answers (10/14) relevant questions from Perera et al.'s requirements for the knowledge-modelling of privacy policies fully, (2/14) partially, and fails to answer (2/14). It also covers signals not identified by the checklist. We also validate the Taxonomy by applying it to extracts from documentation, and argue that it shows potential for the annotation and evaluation of privacy policies as well.

1. Introduction

In interacting with digital apps and services in what Hildebrandt [1] terms the modern '*on*life', users create digital identities and generate massive amounts of associated personal data. The interaction between the users of these devices and services, and their service providers is characterised by a variety of roles and relationships (e.g., user-service provider, consumer-trader, data subject-data controller).

Crucially, one of these relationships is that of a principal (the user) and an agent (the service provider) [2], as the user must rely on the service provider performing its task of protecting and enabling her 'privacy' with care and effort. The linchpin of any such relationship is a contract between the parties; in this case, this is quintessentially represented by the service provider's 'documentation', which hereinafter refers to the terms and conditions, privacy policy, and linked legal or technical documents [3]. Ideally,

[1] PhD Candidate, Department of Information Management, TiSEM, Tilburg Universty, Warandelaan 2, 5037 AB Tilburg, The Netherlands; E-mail: k.chawla@uvt.nl. My thanks to Bert-Jan Butijn, Mariana Evram, Rajesh Chawla, and Ruud van Cruchten for their comments and feedback.

these digital contracts should be negotiated and their implementation monitored to the benefit of both parties. In practice, however, these contracts are often ignored, and even if they are not, are difficult to comprehend, note or manage [4,5]. Consequently, one of the biggest issues in the contemporary privacy debate is enabling users to negotiate the default conditions and monitor the actions of all of the apps, services and devices that collect their data. This is a problem for users, but also for service providers, data protection authorities, and the market as a whole [6].

In research and in practice, we find a variety of initiatives to address this issue. One stream of research focuses on negotiation protocols such as the P3P [7], or the creation of privacy assistants [8], analogous to the idea of a 'butler' [9]. Others investigate the automatic annotation or evaluation of privacy policies [10,11,12,13]. A third stream focuses on the development of 'Personal Data Stores' (PDSs),[2], which are systems that provide an architecture allowing users retain and manage their own data. There are almost certainly other initiatives as well.

Each of these proposed solutions needs to work with a representation of the 'Documentation', or at least the privacy policy, a role currently fulfilled largely by natural language text. Morel and Pardo [14] survey the means of representation of privacy policies and find three main dimensions: natural language, graphical and machine-readable, each fulfilling some particular needs of the communities they originate from. However, none can single-handedly fulfill the requirements of all communities (e.g., legal compliance, understandability and enforceability). Morel and Pardo argue that what is needed instead is a *multi-faceted privacy policy*, one that covers all three dimensions simultaneously [14]. We agree. Multi-faceted documentation would allow users to process and manage their interactions with digital services according to their privacy preferences better than natural language documentation alone. A secondary benefit of machine-readable taxonomies is easier enforcement [14]. We would add that machine-readable policies also allow for the creation of privacy management tools for the management of all the policies a user 'agrees' to, similar to current password managers like LastPass[3], and for the customisation of 'notice'.

For the creation of machine-readable and graphical documentation, a pre-requisite is a categorisation and coherent representation of a service provider's data practices. This is not an easy task [11]. The objective of this research is to develop a *Taxonomy for the Representation of Privacy and Data Control Signals*. The term 'signal' comes from contract theory (law and economics), and refers to credible information conveyed by the agent to the principal, in a market with asymmetric information [15]. A signal is meant to reveal certain information about the agent's behaviour (here: data handling and control practices) to the principal, so that they can react accordingly. The natural language documentation from the service provider is supposed to convey some of these signals to the user, and in its final version it represents a 'meeting of the minds' as to what happens to a user's data.

In this paper, we report on our efforts to identify a taxonomy to represent such 'privacy and data control signals', as communicated in the Documentation. We employ design science for this task [16], specifically Nickerson et al.'s [17] methodology for the development of taxonomies. This research answers the knowledge question 'What infor-

[2]SOLID (`https://inrupt.com/solid`), Hub of All Things, (`https://www.hubofallthings.com/`)
[3]`https://www.lastpass.com`

mation should a multi-faceted documentation be able to represent?'. The design question 'How should this information be represented?' will be the focus of further research.

We present the results from the first two iterations of the method. We identify a complex and multi-layered Taxonomy, based on four empirical samples. We evaluate the Taxonomy at this stage using Perera et al's [18] checklist of questions that a knowledge-based modelling of privacy policies should be able to answer. In the long run, however, as Nickerson et al. [17] note, a taxonomy is only useful if it is used. We aim to bring this Taxonomy to a level where it can actually be used in practice, *inter alia*, for multi-faceted privacy documentation, and annotation schemes, and subsequently for the creation of smart contracts and privacy assistants. The next steps in this project consist of two interlinked stages: running further iterations of the Taxonomy's design cycle, and using the Taxonomy for the design and implementation of Multi-faceted Privacy Policies.

2. Theory

We live today in what Mirelle Hildebrant [1] calls a 'new animism': a transformative '*on*life' situated *"beyond the increasingly artificial distinction between online and offline."* In interacting with this '*on*life', we create digital identities and generate massive amounts of associated data from the increasing number of devices and services that we use in the course of our daily lives, from the banal to the exceptional. This creates significant challenges for privacy-conscious users.

2.1. Notice and Choice

Each digital service comes with its own documentation, its own 'contract'. The user's consent to this agreement rests, precariously, on the infamous principle of 'notice and choice' [19]. This is currently implemented largely by a service provider's natural language documentation, though various parts thereof can be scattered throughout a user's experience with a service.

This mechanism simply cannot keep up with the evolution of technology and the cornucopia of information that needs to be conveyed. A significant amount of literature that the failure of the 'notice and choice' mechanism in general [20,19,5], and the failure of privacy policies in specific [21]. The issue is not limited to privacy – 'online' or 'digital' contracts are generally associated with significant information and negotiation power asymmetries [4].

As Calo [19] notes, however, the problem doesn't lie in the idea of 'notice and choice' but in its implementation. There will necessarily always be a component of 'notice' and 'choice' in such interactions. Even the GDPR requires transparency about processing operations and their purposes (Recital 60), and requires consent to be *"freely given, specific, informed and unambiguous"* (Art. 4(11)) where it is necessary. At the normative level, this is necessary. At the practical level, such transparency and active choice is difficult to implement for service providers, and difficult to comprehend and use for end-users. It results in documentation being written for the purpose of legal compliance [22], rather than for the communication of privacy signals to users [14]. It also creates a potential opportunity for exploitation by rent-seekers. For effective notice and choice, the documentation, and the privacy and data control signals it contains, needs to evolve into a more manageable format.

2.2. Privacy and Privacy Signals

In the market for digital services, a privacy-conscious user, (seen here as principal [2]) must identify the service providers that match her privacy preferences and communicate such preferences to them. As the agent, the service provider's must communicate 'signals' about the quality of their service and about how they accommodate and respect their users' privacy preferences. Neither of these tasks are easy.

Before we can identify them, we must define what we mean by 'privacy and data control signals'. Vila et al. [6] analysed the market for privacy in websites and described it as a market with asymmetric information – exactly the type of market where 'signals' become relevant. They define 'signals' as *"a means by which privacy-respecting sites can differentiate themselves from their defecting competitors"* [6].

Vila et al. [6] mention a 'strong' privacy policy as an example of such a signal. But what is a 'strong' privacy policy? The fact of the matter is that a strong privacy policy is often the one that matches the user's preferences. A policy that one user may consider 'weak' might be entirely acceptable by the standards of another user. The GDPR provides a set of 'default rules', a minimum standard that every policy must comply with. Beyond that, however, there is still a lot of room left for negotiation, or rather selection. A practical example of this is the collection of user data on websites via cookies, and the options given to users for their consent to different types of data collection.

A signal can therefore be defined as information that allows users to identify the whether the data management practices of a service provider matches their expectations or preferences or not, allowing them to adjust their behaviour accordingly. So any type of information that reveals how the service provider handles its user data, and which 'choices' it allows users, will count as such a signal. This is a rather broad definition, but we will limit our scope by focusing only on signals in the service provider's documentation. We also focus exclusively on the user-service provider relationship, even though third-party integration of services and devices creates important consequences for a user's privacy.

We focus specifically on the signals that correlate with Westin's [23] definition of privacy as an individual's right *"to control, edit, manage, and delete information about them*[selves] *and decide when, how, and to what extent information is communicated to others."* 'Signals' here include legal and technical information. An illustration of such signals is Naeini et al. [24]'s work on standardised 'labels' for privacy in IoT devices.

2.3. Open Texture

Legal documents, often suffer from the 'open texture' of language; i.e., they employ open-ended terms (sometimes intentionally) in order to account for as many potential eventualities as possible [25] and to comply with legal requirements [22]. These are not written, primarily, for the communication of these 'signals'. For instance, privacy policies often indicate that an action 'may' be conducted without specifying whether it is actually conducted or not, or use inclusive rather than exhaustive lists at many places (e.g. *"...and other information you might share with us"*).

This has two important consequences. First, there is necessarily some loss of information between 'open-textured' legal documentation and the specificity of a taxonomy-based representation. Second, this along with the dynamic nature of the context means that any taxonomy would need to be flexible, frequently updated, and carefully applied.

3. Related Work

Our focus is on privacy and data control signals, their representations, taxonomies, and evaluations. There is related work in legal (e.g. [26]) and technical [27] research. Directly relevant is research on the annotation or evaluation of privacy policies [11], and the development of knowledge-based languages for their representation [18]. A search of the keywords 'privacy policy' & 'annotation' on SCOPUS, Web of Science and IEEE identifies 21 papers that seem relevant based on their titles and abstracts. Of these, 10 contain or utilise annotation schemes or other categorisations of privacy policies content. Most use a novel scheme, though two reuse Wilson et al.'s corpus and scheme [12].

There are two main limitations of these categorisations. First, many are limited to *the* privacy policy, ignoring additional documentation and the settings and choices offered in the software itself. Second, many tend to focus on the readability, comprehensibility or compliance of the text of the privacy policies. Few provide an analysis of the content of privacy policies, with exceptions such as Antón & Earp [27] (requirements engineering perspective) for and Wilson et al. [12] (computational linguistics).

4. Methodology

We adopt Nickerson et al.'s [17] method for the development of taxonomies in a domain of interest, due to its suitability for the task at hand. The method consists of two cycles, Empirical-to-Conceptual (E2C) and Conceptual-to-Empirical (C2E), in which either dimensions, or characteristics are added, on the basis of empirical material and literature respectively. Table 1 summarises the the application of the methodology in this project.

Table 1. Nickerson et al.'s [17] Methodology, As Applied

Methodology Components	Methodology Application
1. Meta Characteristic	**Signals**: Information regarding a service provider's data handling practices and the control allowed to users, relevant for ascertaining compatibility with a user's privacy preferences. **Expected Use**: The design of a multi-faceted privacy policy, privacy-policy annotation. **Purpose**: The categorisation of the information conveyed via a natural language privacy policy and associated documentation
2. Ending Conditions	**Objective Condition 1**: No new dimensions or characteristics added in the last iteration **Objective Condition 2**: No dimensions or characteristics merged or split in the last iteration **Subjective conditions**: Conciseness, Robustness, Comprehensiveness, Extendibility, and Explanability
3. Empirical to Conceptual Approach	3.1 The sampling of four diverse examples of 'smartphone-based fitness activity tracking' applications and their documentation. 3.2 The coding and organisation of 'signals' from the documentation with Atlas.ti. 3.3 The categorisation of the signals, providing a first iteration of the Taxonomy.
4. Conceptual to Empirical Approach	4.1 The identification of 5 selected papers relevant to the taxonomy. 4.2 Identifying additional dimensions and characteristics from the selected literature. 4.3 The addition or reorganisation of the Taxonomy taking into account the insights from the literature.

4.1. Sample Selection

For the first iteration of this project we selected four 'fitness activity tracking' applications on the Google Play Store: Strava (socially-oriented), Runkeeper (was subject to a complaint for privacy violations by the Norwegian Consumer Council), Adidas Runtastic (Affiliated entity), and OpenTracks (privacy-oriented). The documentation of these services was obtained via the links provided on their Google Play Store pages, other linked pages as needed, and the relevant 'settings' page in the applications. Where the policies linked to policies of third-party service providers, the latter were not analysed.

In the first cycle, the documentation collected from the samples was analysed and coded with Atlas.ti using a grounded approach. The identified codes were then structured to identify the relevant dimensions and attributes, taking into account definitions and concepts from the GDPR. The second cycle took into account the categorisations used in literature on the topic, specifically: Wilson et al. [12,11], Morel and Pardo [14], Bhatia et al. [28] and Contissa et al. [10]. These papers were selected for their relevance. The Taxonomy resulting from the E2C cycle most closely resembles Wilson et al. [11]'s work though it is structured differently and adds a few dimensions, which can be taken as at least a partial validation of it. Similarly, categorising the 'types of data collected' is tricky, and Bhatia et al. [28] offer an alternative to the results of the E2C cycle. After these two cycles, further iterations will follow. In the next iteration, we aim to include a larger sample of applications and documentation and literature.

5. Results: A Taxonomy for Privacy and Data Control Signals

This exercise results in a complex and multi-layered Taxonomy for the Representation of Privacy and Data Control Signals. Due to space constraints, we can only present the first two levels of the dimensions of the resulting Taxonomy in (Fig 1) and explain some of the important dimensions below. The full Taxonomy is available in graphical and tabular representations on Github [4].

The taxonomy contains three levels of dimensions. The first level includes: 'Policy Meta-Data', 'Data and Control', and 'Processing and Usage'. Data and Control includes, *inter alia*, signals about what data is collected and what controls users are allowed; Processing and Usage relates to processing activities, entities and purposes; and Policy Meta-data covers some more abstract, contextual information, and information regarding the policy itself. To allow for documentation's open texture, we keep the 'Types of data collected' agnostic to the specific characteristics of the data or the sensors from which it is collected. These are covered under a separate dimension ('Data Characteristics'). Bhatia et al. [28] provide an alternative lexicon for this. We also identify signals relating to User Control, further divided into control 'Options', 'Channels' and 'Limitations'. This is similar to the factors in the design space for privacy notices noted by Schaub et al.[22]. We separate the anonymity of a user's profile on a platform ('Active Audience') from the risk that the service provider will share their data with a third party ('Data Shared with').

We identify three distinct but related types of signals in the Documentation: legal basis, purpose and functionality. Each conveys some information about why a user's data is collected, but at different levels of abstraction. The 'basis' is the most abstract, relating

[4] https://github.com/KartikChawla-droid/Taxonomy_Privacy_Data_Control_Signals

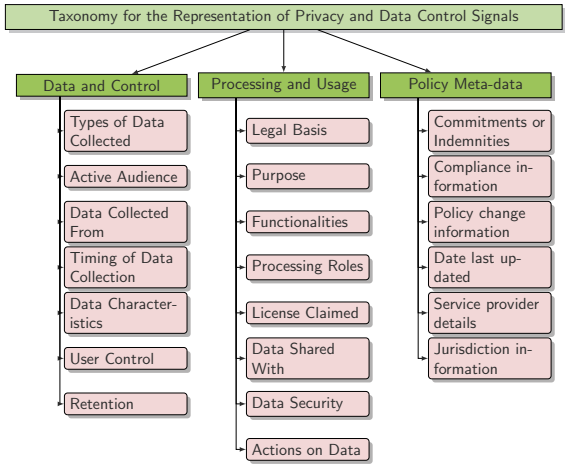

Figure 1. The First Two Layers Of The Dimensions Of The Taxonomy

to Art. 6(1) of the GDPR. The 'purpose' is slightly more specific but still vague, while 'functionality' is the most specific and relates closely to the technical aspects of the service. For instance, the purpose for the collection of a user's account data and location data, both, is 'provision of service'. The distinction lies in the functionality: 'account creation' and 'fitness activity tracking' respectively. Functionality, furthermore, allows for a comparison between services: if two unrelated services both offer a 'social network' functionality, even if they cannot be compared as a whole, the implementation of this functionality and the data it collects can be compared between the two services.

6. Evaluation and Discussion

The Taxonomy identifies a variety of factors not identified by previous research: the distinction between user control options, limitations and channels; the commitments and indemnities; whether the data is licenseable and whether such a license is claimed or not; and, crucially, the distinction between the 'legal basis' and 'purpose' of data collection and the 'functionality' it links to. The variety of dimensions identified by it verifies the depth and complexity of the information conveyed via the documentation.

The purpose of the Taxonomy is to enable representation of privacy signals in a multi-faceted format. The individual elements of the Taxonomy already identify some relevant signals, but combinations thereof identify even more, making explicit the links between different types of information. For instance, 'Types of data collected' is a signal in and of itself but its combination with 'functionalities' or 'data shared with' communicates a different, but still crucial, type of a signal altogether. Clustering 'types of data collected' with 'functionalities' tells the reader what data is funneled into which functionalities. Keeping in mind that the same data may go into multiple functionalities, this allows a user to evaluate the 'exchange'; i.e., it allows users to see which functionalities require which types of data, and evaluate whether they are willing to forgo with such data to receive these functionalities. An analogue to this in practice is the separation of 'cookies' by functionality (such as: 'necessary', 'marketing' or 'analytics') available in cookie consent managers, and the separation of microservices in cloud computing.

The results of the Taxonomy also justify the extension of its scope beyond the privacy policy, and beyond compliance with regulations such as the GDPR. The dimensions relating to the licenseability of the data would not have been identified from the privacy policy alone, and the 'functionalities' and 'applicable jurisdiction' are more evident in the Terms and Conditions in some cases. A variety of signals go beyond pure compliance with the GDPR. For instance, Table 2 illustrates a possible application of the Taxonomy, using the OpenTracks privacy policy as an example[5]. Note that this policy would fail a test for GDPR compliance, but then it doesn't need to comply because there is no third-party processing of data! However, even this two-sentence documentation contains important information that is captured by the Taxonomy. A more extended list of examples for evaluation is available in the Github repository.

Given the diversity and dynamic nature of the information conveyed by the sample space, we would argue that rather than specifying all the information that could potentially be conveyed, it would be more efficient to specify a flexible code 'library' or 'package', based on this taxonomy, that can enable the writers of the documentation to add new information on the fly.

Table 2. Coding From Text and Application Of Taxonomy To Opentracks Sample

Sample Text	**Taxonomy Coding**
OpenTracks does only store data on the local device	retention_location: local
that is relevant for tracking your sport exercise .	functionality: activity tracking
Stored data is not transmitted from the app itself to a third party.	type_of_data_collected: [none]

We further evaluate the Taxonomy against Perera et al.'s [18] checklist about the information a representation language for privacy policies should be able to convey. The checklist contains a total of 17 questions, 14 of which are relevant for the 'content' of the representation. The Taxonomy presented here can answer 12 questions completely, 2 partially, and fails to answer 2. For further details, please refer to Github. One unanswered question tells us that we need to add 'Methods of data collection', in the next cycle. The second missing question asks what information is data controllers expect to discover from the user's data, but this information is not present in the sample documentation.

7. Limitations and Further research

This research has certain limitations. First, an application's effect on a user's privacy must take into account its technical context. Applications are necessarily deployed on a hardware and software stack ('vertical stack') and may be integrated with third-party applications and services working in parallel ('horizontal integration'). Both affect the functioning of the application and the user's privacy. We have not taken these vertical and horizontal interfaces into account, but a useful taxonomy needs to be 'modular' to accommodate this layering. Second, the open texture of legal documents means that a certain loss of information or ambiguity in the the taxonomy is perhaps inevitable (e.g., 'open-ended' as an attribute for 'purpose'). Third, our analysis is limited to privacy signals contained in the documentation and technical implementation. There are further

[5] https://opentrack.run/about/privacy.html

market-oriented signals which have not been included here, such as reputation, size, business model, and of course the code (as much as is observable). These and further signals regarding the context [29] and consumer rights [30] should be included in further research. Particularly, information regarding the API calls made or enabled by an application, if available, should also be included in the Taxonomy. Fourth, the Taxonomy is limited to the *representation* of 'signals'. The natural follow-up question is whether the communicated signals are legitimate or not. This would require a more elaborate system for monitoring a service provider's behaviour and testing compliance with the agreements [6]. That makes a good topic for future research.

8. Conclusion

For online services, we look at the relationship between users and service providers from the perspective of principal-agent theory [2]. This relationship exists in a market with asymmetric information, which means that 'signals' about the digital service are crucial for users. From the empirical analysis it is evident that a service provider's documentation provides a lot of privacy and data control signals in a relatively unstructured form. However, currently, signals about privacy and data control tend to get lost in natural language documentation. The negotiation and monitoring costs the user must bear to ensure an optimal contract are too high without support tools. Even if a user retains technical control over her data with a PDS system, she would still need legal support tools for negotiating and monitoring access to her data. The depth and complexity of the information, even when viewed through the lens of the Taxonomy, makes the need for machine readable or annotated privacy policies self-evident even without taking into account the behavioural issues pointed out by Acquisti et al. [31].

This paper presents the results of the first two iterations of a design science project for the development of a 'Taxonomy for the Representation of Privacy and Control Signals' that allows for a machine-readable representation of these signals. We identify crucial dimensions not covered by previous taxonomies, based on an empirical analysis of four sample documentations. This answers the knowledge question 'What information should be represented in a multi-faceted documentation on privacy and data controls?'. The Taxonomy still requires further iterations, which are planned. At the same time, we will attempt to use this knowledge model for the annotation and evaluation of privacy policies and the development and design of smart contracts and privacy assistants. That is, we will attempt to use this to answer the design question 'How should this information be represented?' in further research. We will conduct a survey of relevant tools for the latter (e.g. with protégé, as XML, or as a library) as well.

References

[1] M. Hildebrandt. Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology. Cheltenham Edward Elgar Publishing; 2015.

[2] EA Posner. Agency Models in Law and Economics. John M. Olin Program in Law and Economics Working Paper No. 92. 2000.

[3] M. Pavis. Paris Tribunal Guts Twitter'S T&Cs... Including The Copyright Clause For User-Generated Content. [online] The IPKat. Available at: https://ipkitten.blogspot.com/2018/09/paris-tribunal-guts-twitters-t.html [Accessed 23 October 2020].

[4] R. Momberg. Standard Terms and Transparency in Online Contracts. In A. De Francheschi, editors, *Standard Terms and Transparency in Online Contracts*. Intersentia, 2016:189-206.
[5] K. Martin, Transaction costs, privacy, and trust: The laudable goals and ultimate failure of notice and choice to respect privacy online. FM [Internet]. 2013 Dec. 15 [cited 2020 Oct. 23]; 18(12). Available from: https://firstmonday.org/ojs/index.php/fm/article/view/4838.
[6] T. Vila, R. Greenstadt, and D. Molnar. Why We Can't Be Bothered to Read Privacy Policies - Models of Privacy Economics as a Lemons Market.*Proceedings of the 5th ICEC*. 2003.
[7] I. Reay, S. Dick, and J. Miller. An analysis of privacy signals on the World Wide Web: Past, present and future. Inf Sci. 2009 Mar 29; 179:1102–15.
[8] A. Das, M. Degeling, D. Smullen, N.M. and Sadeh. Personalized Privacy Assistants for the Internet of Things: Providing Users with Notice and Choice. IEEE Pervasive Computing. 2018; 17:35–46.
[9] Lawrence Lessig. Code and Other Laws of Cyberspace. Basic Books; 1999.
[10] G. Contissa, K. Docter, F. Lagioia, M. Lippi, H.W. Micklitz, P. Palka, G. Sartor, P. and Torroni. Claudette Meets GDPR : Automating the Evaluation of Privacy Policies Using Artificial Intelligence. SSRN Electronic Journal. 2018 Jan 1.
[11] S. Wilson, F. Schaub, F. Liu, K.M. Sathyendra, D. Smullen, S. Zimmeck, R. Ramanath, P. Story, F. Liu,N. Sadeh, and N.A. Smith. Analyzing Privacy Policies at Scale: From Crowdsourcing to Automated Annotations. ACM Trans Web. 2018 Dec 13.
[12] S. Wilson, F. Schaub, A.A. Dara, F. Liu, S. Cherivirala, P. Giovanni Leon, M. Schaarup Andersen, S. Zimmeck, K.M. Sathyendra, N.C. Russell, T.B. Norton, E. Hovy, J. Reidenberg, and N. Sadeh. The Creation and Analysis of a Website Privacy Policy Corpus.*Proceedings of the 54th Annual Meeting of the ACL (Volume 1: Long Papers)*. Berlin, Germany, 2016 Aug :1330–40.
[13] D. Audich, R. Dara, and B. Nonnecke. Privacy Policy Annotation for Semi-automated Analysis: A Cost-Effective Approach. 2018 Jun 30;29–44.
[14] V. Morel and R. Pardo. Three Dimensions of Privacy Policies. Research Report 9287. Inria, Project-Teams Privatics, 2019 Nov.
[15] M. Spence. Job Market Signaling In P. Diamond and M. Rothschild. Uncertainty in Economics. Academic Press, 1978 Jan 1:281–306.
[16] R.J. Wieringa. Design Science Methodology for Information Systems and Software Engineering. Springer, 2014.
[17] R.C. Nickerson, U. Varshney, and J. Muntermann. A method for taxonomy development and its application in information systems. European Journal of Information Systems. 2013;22:336–59.
[18] C. Perera, C. Liu, R. Ranjan, L. Wang, and A. Zomaya. Privacy Knowledge Modelling for Internet of Things: A Look Back. Computer. 2016; 49.
[19] MR Calo. Against Notice Skepticism in Privacy (and Elsewhere). Notre Dame Law Review. 2011;87:1027.
[20] FH Cate and V. Mayer-Schönberger. Notice and consent in a world of Big Data. International Data Privacy Law. 2013 May 1; 3:67–73.
[21] R. W. Proctor, M. Athar Ali, and L. Kim-Phoung L. Vu. Examining Usability of Web Privacy Policies. International Journal of Human-Computer Interaction. 2008; 24:307–28.
[22] F. Schaub, R. Balebako, L.F. and Cranor LF. Designing Effective Privacy Notices and Controls. IEEE Internet Computing. 2017 May; 21:70–7.
[23] A.F. Westin. Privacy and Freedom. Bodley Head, 1967.
[24] P.E. Naeini, Y. Agarwal, L. Cranor, and H. Hibshi. Ask the Experts: What Should Be on an IoT Privacy and Security Label? 2020 Feb 11.
[25] H.L.A. Hart. The Concept of Law. Third. Oxford University Press, 2012.
[26] D.J. Solove. A Taxonomy of Privacy. Univ Pa Law Rev. 2006 Jan; 154:477–564.
[27] AI Antón and JB Earp. A requirements taxonomy for reducing Web site privacy vulnerabilities. Requirements Engineering. 2004 Aug 1; 9:169–85.
[28] J. Bhatia and T.D. Breaux. Towards an information type lexicon for privacy policies. 2015 :19–24.
[29] H. Nissenbaum. A Contextual Approach to Privacy Online. Daedalus. 2011 Oct 1; 140:32–48.
[30] M. Loos and J. Luzak. Wanted: a Bigger Stick. On Unfair Terms in Consumer Contracts with Online Service Providers. Journal of Consumer Policy. 2016 Mar; 39:63–90.
[31] A. Acquisti, L. Brandimarte, and G. Loewenstein. Privacy and human behavior in the age of information. Science. 2015; 347:509–14.