# Privacy Rating: visualising privacy attributes of online services

Susanne Barth, Dan Ionita and Pieter Hartel

University of Twente, The Netherlands
Email: s.barth@utwente.nl, d.ionita@utwente.nl, pieter.hartel@utwente.nl

***Abstract*—Privacy concerns regarding online services are on the rise. Recent leaks and breaches have increased the awareness of both users and regulators. Many countries now mandate transparency and consent when personal data is handled. Nevertheless, it is increasingly difficult for consumers of online services to assess the privacy risks of using an online service or compare services based on their privacy level. In this paper we propose a rating system for privacy which provides a yardstick for measuring and comparing the privacy of online services. We also describe how the rating can be used to produce a concise visual representation which conveys the most relevant aspects of privacy.**

## I. Introduction

Online services range from social media (e.g. Facebook and Instagram) and entertainment (e.g. Netflix and Spotify) to shopping (e.g. Amazon and eBay) and banking (e.g. PayPal and Bunq). All of these services handle personal data. The pervasiveness of digital technologies in day-to-day life has resulted in a "Semantic Web" built almost entirely upon personal information.

Because of the complexity and interconnectedness of these services, users often have difficulties in assessing their privacy. Privacy policies describe how online services handle user data but most users never read them and those that do have difficulties understanding them.

Researchers, regulators, and industry are calling for a more visual representation of how e-services handle personal data [1], [2], [3], [4], [5] to communicate privacy risks to users in a clear and concise manner. The European GDPR in fact mandates "standardized icons" as an overview of the intended data processing [6]. In parallel, Privacy by Design (PbD) is also gaining a foothold in the software development industry. PbD aims to incorporate privacy into the fabric of digital services instead of bolting it on. A large variety of PbD standards and guidelines have emerged, such as those provided by ISO and The Privacy Company.

In [7] we reviewed eight different approaches to visualising privacy and eight established PbD guidelines to distill Unified List of Privacy Attributes which provide a basic ontology for discussing privacy. We also found that users and experts tend to rank all of the attributes as important or very important. In this whitepaper, we build upon those findings by operationalizing the privacy attributes so that they can be used as metrics for assessing the privacy of an online service. We then design a way to visually communicate the values of these metrics to users. More concretely, we present:

1) A *privacy rating* system which determines the level (Good/Bad/Neutral) of 12 privacy attributes and assigns a privacy class (A-to-G);
2) A *privacy label* with provides an interactive visual summary the privacy rating;
3) A *Proof-of-Concept* web-application which generates an embed-able privacy label based on an interactive questionnaire.

## II. Background

As part of the SERIOUS project (NWO grant number 628.001.011) we first looked at existing approaches to visualising privacy, namely: Mehldau's data-privacy declarations[1], KnowPrivacy's Policy Coding Methodology[8], CyLab's privacy nutrition label [9], Mozilla's privacy icons [2], The PrimeLife project [4], TrustArc's Privacy Short Notice [10], GDPR's draft privacy icons, [6], and CLEVER FRANKE's privacy label [3]. Except for CleverFRANKE's privacy label which is currently under development, all of the other privacy visualization projects have been abandoned. Furthermore, we notices significant differences between the approaches in terms of which aspects of data handling they represent, as well as in terms of their level of detail.

In order to shed light on which privacy attributes are important, we looked at Privacy by Design (PbD) guidelines. PbD is an approach to software development aimed at considering privacy risks during development in order to design applications which have privacy considerations built-in. We examined: Langheinrich's - Principles of Privacy-Aware Ubiquitous Systems [11], The Global Privacy Standard /citecavoukian2006creation, Cavoukian's 7 Foundational Principles [12], The Generally Accepted Privacy Principles [13], the ISO29100 Privacy Framework [14], OECD's privacy principles [15], Hoepman's Privacy Design Strategies"[16], and the Privacy Company's Privacy by Design Framework [17]. Our research revealed a significant gap between privacy attributes represented in proposals for privacy visualisations and those mandated by PbD principles.

[1]The full list of icons is available under a CC-BY license from: https://netzpolitik.org/wp-upload/data-privacy-icons-v01.pdf.

[2]https://wiki.mozilla.org/Privacy_Icons

[3]https://medium.com/sensor-lab/crafting-a-universal-privacy-system-681c04d45689

So we distilled a complete list of privacy attributes from all of the documents listed above. We discussed this list with experts, and performed several online surveys to rank these attributes in terms of importance. We ended up with a total of 15 privacy attributes:

**Accountability** = can the service provider be held accountable for violations?

**Anonymization** = are all identifiable markers completely removed so that data can never be traced back to a single person? *

**Collection** = which data is collected?

**Control** = is the data subject able to choose or decide which data to share and for which purpose, and how difficult is it to do so?

**Correctness** = are there mechanisms for preventing and fixing incorrect data?

**Disclosure** = what is the attitude of the service provider towards requests from law enforcement?

**Functionality** = is the user forced to choose between functionality and privacy?

**Purpose** = what is the collected data used for?

**Pseudonymization** = are personally identifiable markers replaced by artificial identifiers, or pseudonyms, such that data could can only be traced back to individual users with the help of additional information?*

**Retention** = how long is the collected data stored?

**Right to be forgotten** = can data subjects request that all personal data be removed?

**Sale** = is any of the data sold to third parties?

**Security** = what technical measures are taken to ensure that data is protected from unauthorized or malicious access?

**Sharing** = does any of the collected data leave the ownership of the service provider?

**Transparency** = is the user able to obtain information regarding how their personal data is handled?

## III. Privacy rating

In our online surveys, we learned that several experts experts and users found ”anonymisation” and ”pseudonymisation” hard to distinguish. Since they serve the same purpose, namely to de-identifying data, we decided to collapse the two for the purpose of our privacy rating. We view pseudonymisation as an inferior approach to de-identifying data, so for our rating, we view ”pseudonymisation as incomplete anonymisation. We also decided not to include the ”functionality” attribute because several user and experts in our previous online survey [7] found it ambiguous and because it is partly covered by Control. Finally, we do not rate the Transparency attribute because we want to design a label which in itself is an indicator of great transparency.

The system rates the remaining 12 privacy attributes from good to bad as described in Table I. We opted for only three levels per attribute in order to keep the rating simple and generic. We can then average the 12 attributes in order to obtain a privacy class. To do so, we give 0 points for every Bad attribute, 1 point for every Neutral attribute and 2 points for every Good attribute. This gives a total of between 0 and 24 points which we map to an A-to-G class, as show in Fig. 1.

*Rephrased in the online survey for simplicity

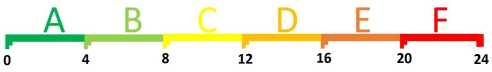



Fig. 1. Mapping the privacy rating to a privacy class

## IV. Privacy label

We went through several iterations while designing our label. We tried to base our approach on previous visualisations as well as the EU Enery Label. From our research, we know that privacy is subjective and context-dependent and that users vary in terms of their privacy concerns and technical knowledge. Therefore, we decided to specify all 12 attributes on the label, but group them into categories so as to allow *three* levels of details. After intensive research, we were able to delineate the following four categories: collection, sharing, control, security. The resulting label consists of 3 collapsible layers, allowing the user to obtain the right level of information for their needs and context. The 1st level consists of an A-to-G-(green-to-red) privacy class. The 2nd layer shows the four groups, rated using the same color scale. Each category can be expanded to show the 3rd layer: a concise and easy-to-read summary of the 12 individual metrics which go into the rating. Figure 2 shows some example labels.

## V. Proof-of-Concept

In order to show how we envision the rating and label to work in practice, as well as to be able to test them, we developed a Proof-of-Concept web-application. This application consists of a questionnaire which walks the user through a serious of questions. Each question corresponds to one of the levels of each attribute. The questionnaire is interactive: Once a question confirms the level of an attribute, the remaining questions corresponding to that attribute are skipped. The flow of the questionnaire is described in Figs. 3 and 4.

Once all attributes have been evaluated, the application produces an HTML snippet of the label as well as a PNG image of a small version of the label, which can be embedded into web pages or apps. For example, the small version of the label can be added to the footer of the page or to the cookie notice and the full version can be included in the privacy policy.

## VI. Conclusion and future work

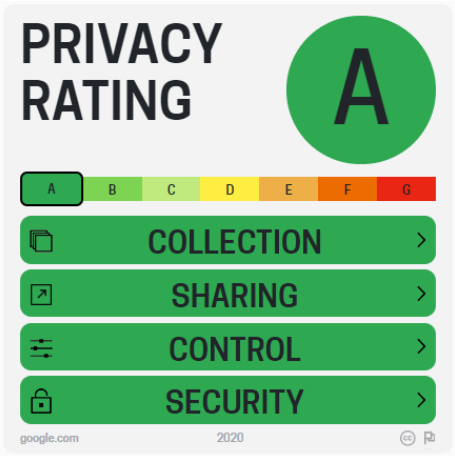
What are our plans?

## Acknowledgment

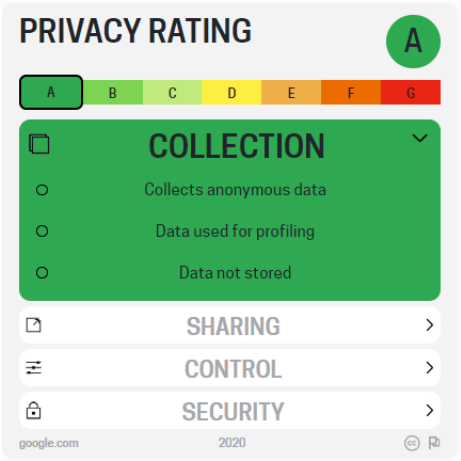
This research was supported by The Netherlands Organisation for Scientific Research (NWO) in the context of cybersecurity research (grant number 628.001.011).

| Attribute | Level | | |
|---|---|---|---|
| | Good | Neutral | Bad |
| Accountability | Provider is legally accountable for privacy breaches | Provider is legally binding privacy policy | Provider is legally accountable for privacy breaches |
| Anonymization | Provider anonymizes user data | Provider pseudonymizes user data | Provider store user data as-is |
| Collection | Provider collects anonymous user data | Provider collects personal user data | Provider collects sensitive user data |
| Control | Users must opt-in for data collection | Users can opt-out of data collection | Users cannot opt-out of data collections |
| Correctness | Users can amend all data | Users can amend some data | Users cannot amend data |
| Disclosure | Provider only discloses user data to government agencies when legally required | Provider cooperates with government agencies within the user's jurisdiction | Provider subject to disclosure requests from government agencies outside the jurisdiction of the end-user |
| Purpose | Service only uses user data to provide basic functionality | Service provides personalized content based on user data | Service performs profiling based on user data |
| Retention | All user data is removed after each session | User data is stored for a limited amount of time | User data is stored indefinitely |
| Right to be forgotten | All user data is deleted upon request | All user data is removed upon request | User data cannot be removed by request |
| Sale | Provider does not sell user data | Provider sells anonymous user data | Provider sells user data |
| Security | Provider is NIST 800-53 or ISO 27001 certified | Service is OWASP compliant (or equivalent) | No established security standards or guidelines |
| Sharing | No user data will leave the ownership of the service provider | Anonymous user data may leave the ownership of the service provider | Personal user data may leave the ownership of the service provider |

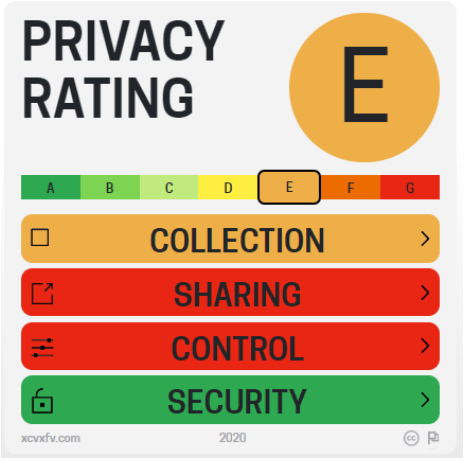
TABLE I
PRIVACY ATTRIBUTES CONTAINED IN THE RATING THE THEIR LEVELS



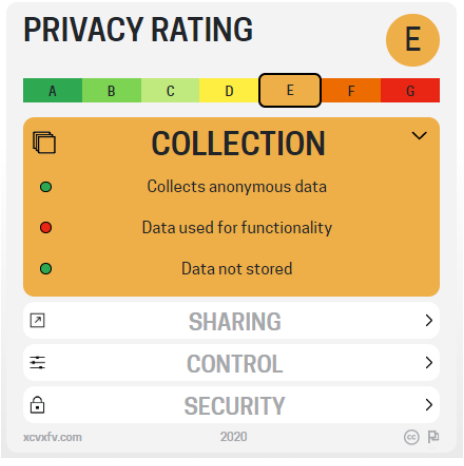

(a) Class A




(b) Class A expanded




(c) Class E




(d) Class E expanded

Fig. 2. Some sample privacy labels

## REFERENCES


[1] A. I. Anton, E. Bertino, N. Li, and T. Yu, "A roadmap for comprehensive online privacy policies," Purdue University, Tech. Rep., 2004.

[2] A. I. Anton, J. B. Earp, Q. He, W. Stufflebeam, D. Bolchini, and C. Jensen, "Financial privacy policies and the need for standardization," *IEEE Security & privacy*, vol. 2, no. 2, pp. 36–45, 2004.

[3] L. Edwards and W. Abel, "The use of privacy icons and standard contract terms for generating consumer trust and confidence in digital services," CREATe working paper series. 10.5281/zenodo. 12506, Tech. Rep., 2014.

[4] L.-E. Holtz, K. Nocun, and M. Hansen, "Towards displaying privacy information with icons," in *IFIP PrimeLife International Summer School on Privacy and Identity Management for Life*. Springer, 2010, pp. 338–348.

[5] A. Rossi and M. Palmirani, "A visualization approach for adaptive consent in the european data protection framework," in *Proceedings of the 7th International Conference for E-Democracy and Open Government, CeDEM 2017*, 2017, pp. 159–170, cited By 7.

[6] Council of European Union, "regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation)."

[7]

[8] J. Gomez, T. Pinnick, and A. Soltani. (2009) Privacy coding methodology. KnowPrivacy. [Online]. Available: http://knowprivacy.org/policies_methodology.html

[9] P. G. Kelley, J. Bresee, L. F. Cranor, and R. W. Reeder, "A nutrition label for privacy," in *Proceedings of the 5th Symposium on Usable Privacy and Security*. ACM, 2009, p. 4.

[10] T. Pinnick. (2011) Privacy short notice design. TrustArc Inc. [Online]. Available: https://www.trustarc.com/blog/?p=1253

[11] M. Langheinrich, "Privacy by design — principles of privacy-aware ubiquitous systems," in *Ubicomp 2001: Ubiquitous Computing*, G. D. Abowd, B. Brumitt, and S. Shafer, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 273–291.

[12] A. Cavoukian *et al.*, "Privacy by design: The 7 foundational principles," *Information and Privacy Commissioner of Ontario, Canada*, vol. 5, 2009.

[13] AICPA and CICA, "Generally accepted privacy principles: Cpa and ca practitioner version," The American Institute of Certified Public Accountants (AICPA) and The Canadian Institute of Chartered Accountants (CICA), Tech. Rep., 2009.

[14] ISO, "Information technology — Security techniques — Privacy framework," International Organization for Standardization, Geneva, CH, Standard, Nov. 2011.

[15] "OECD privacy principles," http://oecdprivacy.org/, accessed: 2019-09-17.

[16] J.-H. Hoepman, "Privacy design strategies," in *IFIP International Information Security Conference*. Springer, 2014, pp. 446–459.

[17] "Data protection by design framework," https://www.privacycompany.eu/files/DPbD_Framework.pdf, The Privacy Company B.V.

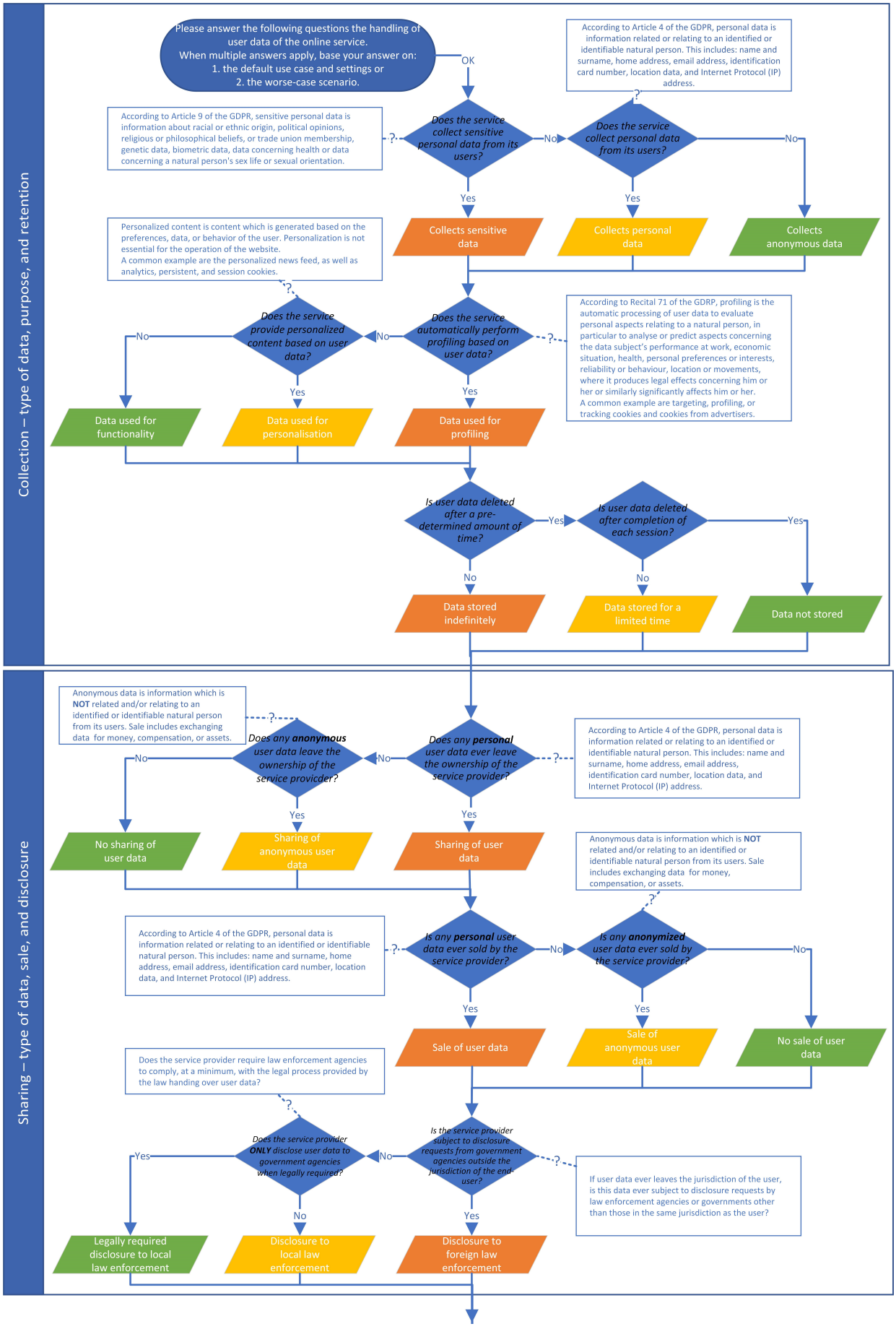



Fig. 3. Process of generating a label - Part 1 of 2

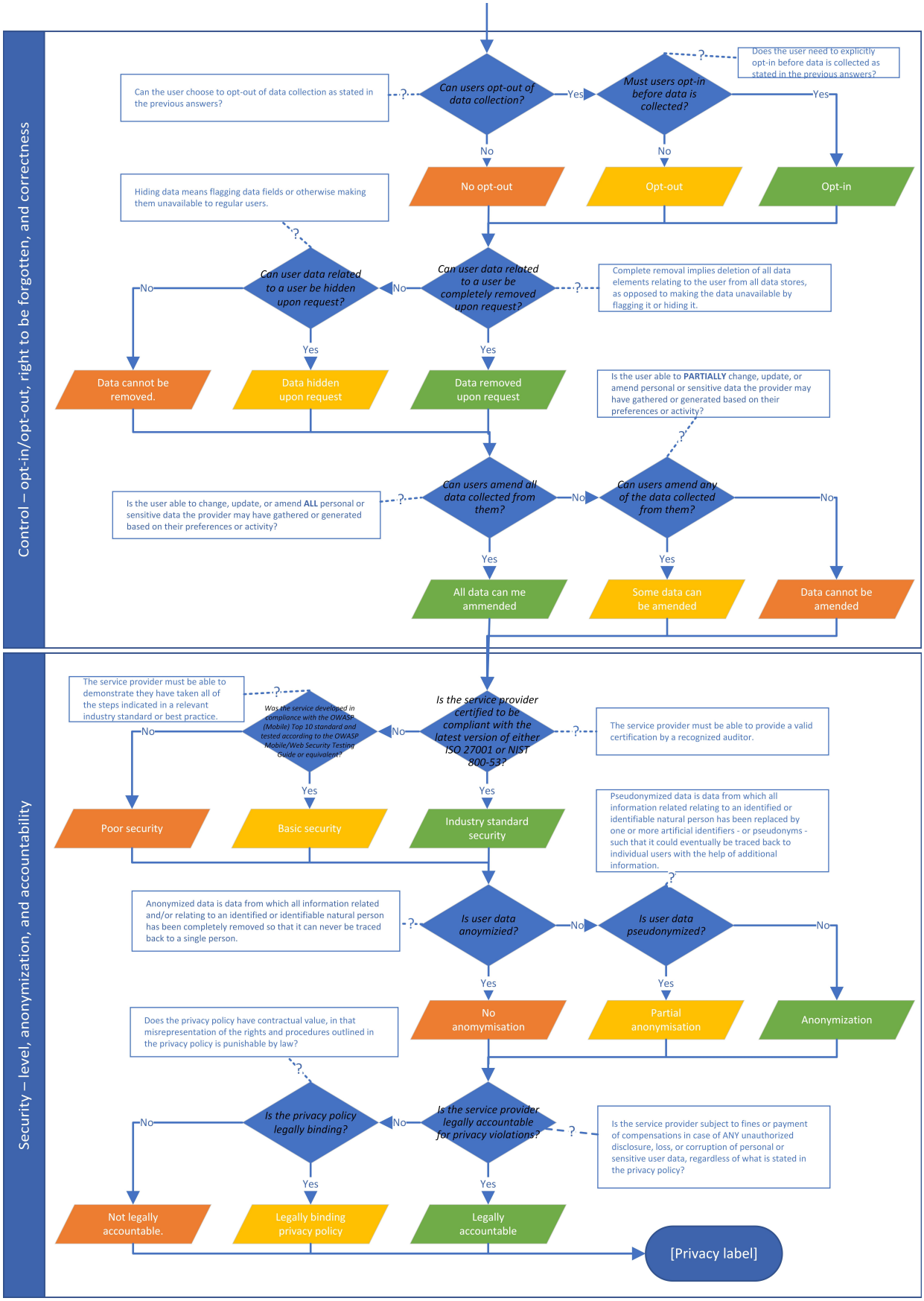



Fig. 4. Process of generating a label - Part 2 of 2