SoK: Three Facets of Privacy Policies

Victor Morel (contact@victor-morel.net), Raúl Pardo (raup@itu.dk)

SOUPS 2021

Motivation

- Privacy policies are the main way to obtain information related to personal data collection
- They are however **rarely read** nor understood
- Different means of expressing privacy policies arose from different communities, adressing different requirements, *i.e.*, **legal validity**, **understandability**, and **enforceability**
- So far, no work systematically studied the different means of expressing privacy policies

Methodology

- We use the term **facet** to denote each main way to express privacy policies, *i.e.*, in **natural language**, with **graphical representations**, and using **machine-readable** means.
- We categorize the content of each facet with a **taxonomy** inspired by Wilson *et al.* [3]:

1st Party Type of data collected, purpose and collection mode
3rd Party Type of data, purpose and collection mode for third parties
Legal Basis Ground which determines the lawfulness of processing
DS Rights Rights of the data subjects
Data Retention Duration of data storage
Data security Modalities of protection of data
Policy Change Modalities of notification for policy changes
Other Identity of DC, information related to DNT, to children . . .

- For each facet we also study specific **tools**, **benefits**, and **limitations**:

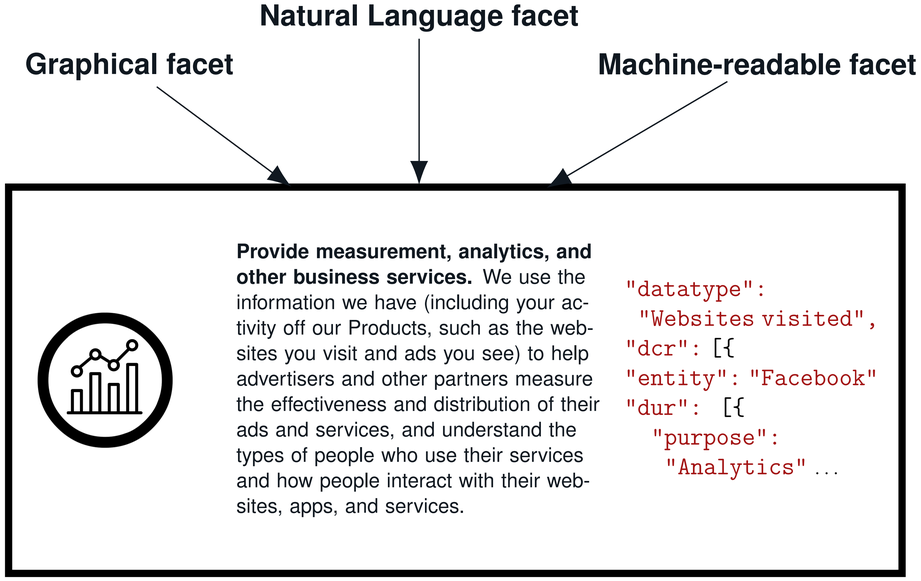
	Natural Language	Graphical	Machine-readable
Tools	Templates Generators Retrievers Analysis tools	Notifications Visual comparison	Enforcement engines Formal semantics Policy comparison Analysis tools
Benefits	Legal value	Understandability	Enforcement Auditability Correctness Automation
Limitations	Ambiguity Understandability Enforceability	Ambiguity Incompleteness	Understandability Lack of adoption

Insights

Existing means of expressing privacy policies *cannot be* **legally valid** and **understandable** and **automatically enforceable**

Recommendations

Privacy policies should be presented as **multi-faceted**, combining *natural language*, *graphical representations*, and *machine-readable* facets.



Example of a multi-faceted privacy policy

\+ Multifaceted policies overcomes the respective limitations of each facet by bringing together their benefits

\- In multifaceted policies, ensuring the consistency of facets may be challenging

Natural Language Facet

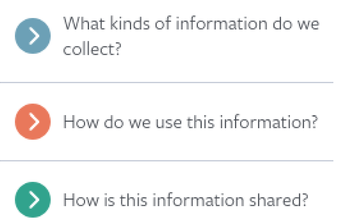


Figure 1: Excerpt of Facebook's privacy policy menu

Aimed at: **law practitioners**

Graphical Facet

Figure 2: Excerpt of the Privacy Tech icons [2]

Aimed at: **the general public**

Machine-readable Facet

Listing 1: Excerpt of a Pilot [1] privacy policy implemented in JSON

```
{
    "pilotRule": [{
    "datatype": "Wi-Fi MAC Address",
    "dcr": [{
    "entity": "Google",
    "dur": [{
      "purpose": "Marketing",
      "retentionTime": 30
    }]
  }]
 }]
}
```

Aimed at: **machines**

References

[1] Raúl Pardo and Daniel Le Métayer. "Analysis of privacy policies to enhance informed consent". In: *IFIP Annual Conference on Data and Applications Security and Privacy*. Springer. 2019, pp. 177–198.

[2] Privacy Tech. *Privacy Icons*. 2018. URL: https://www.privacytech.fr/privacy-icons/.

[3] Shomir Wilson et al. "The creation and analysis of a website privacy policy corpus". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2016, pp. 1330–1340.