Improving the Transparency of Privacy Terms Updates

Opinion paper

Alexander Railean and Delphine Reinhardt
Motivation

- Lengthy privacy terms
- Complicated language, vagueness
- Updated terms are lengthy too
- Fatigue
- The GDPR calls for transparency solutions

Schaub et al., “A Design Space for Effective Privacy Notices”
McDonald et al., “The Cost of Reading Privacy Policies”
Okoyomon et al., “On The Ridiculousness of Notice and Consent”
When paper authors compare versions

- The table in Fig.\ref{fig:consent-table} was originally conceived as a component of an Online Interface for IoT Transparency Enhancement (OnLITE), which summarizes privacy facts, and makes it easy to compare different IoT devices side by side, as shown in Fig.\ref{fig:onlite-diff} \cite{railean_onlite_2020}

+ The table in Fig.\ref{fig:consent-table} was originally conceived as a component of an Online Interface for IoT Transparency Enhancement (OnLITE), which summarizes data collection practices and privacy information, and makes it easy to compare different IoT devices side by side, as shown in Fig.\ref{fig:onlite-diff} \cite{railean_onlite_2020}
When programmers compare versions

@@ -36,6 +36,9 @@
 class MqttClient:
     def on_request(self, client, userdata, msg):
         """Invoked when a message is received on t_requests""
         if self.external_handler:
-           self.external_handler(client, userdata, msg)
+           try:
+               self.external_handler(client, userdata, msg)
+           except Exception as err:
+               log.exception('ERR %err %s %s', err, msg.topic, msg.payload)
         else:
             log.debug("MQTT IN %s %s", msg.topic, msg.payload)
When users compare versions...
Issues with the status quo

- Not even a side by side comparison
- Can the “summary of changes” be trusted?
- Normalization of deviance
Objectives

- Make before/after comparisons easier
- Reduce the number of update prompts
- Reduce volume of displayed information
- Standardize representation
Roots: printed label for IoT devices

- Tested with 31 participants
- Easy to interpret
- Positive feedback
- Published at MobileHCI 2018 [1]

Roots: online interface

- Cross-disciplinary design (legal, usability, security, privacy)
- Iterative approach
- Heuristic evaluation with experts
- Tested with 15 participants
- Task analysis, interview
- Published at NordSec 2020 [1]

<table>
<thead>
<tr>
<th>Collected data</th>
<th>Data flows</th>
<th>Sent</th>
</tr>
</thead>
<tbody>
<tr>
<td>customer nr.</td>
<td></td>
<td>hourly</td>
</tr>
<tr>
<td>temperature</td>
<td></td>
<td>to Tesami GmbH</td>
</tr>
<tr>
<td>humidity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>device internet address</td>
<td></td>
<td></td>
</tr>
<tr>
<td>customer nr.</td>
<td></td>
<td>daily</td>
</tr>
<tr>
<td>temperature</td>
<td></td>
<td>to Aster SRL</td>
</tr>
<tr>
<td>humidity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UV radiation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wind speed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wind speed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparing updated privacy terms

- Compare a device against itself
- Generalize: IoT devices → any privacy terms

**Collected data**
- customer nr.
- temperature
- humidity
- device Internet address

**Sent**
- hourly to Tesami GmbH

**Stored for**
- 3 years in France
Formal notation of privacy terms

Data type $\Delta$  
- type of collected data

Purpose $\Pi$  
- purpose of collection

Time $T$  
- the retention period

Company $C$  
- a company that gets the data

Location $\Lambda$  
- location of said company

Frequency $\Phi$  
- how often the data are transmitted

Term $\Theta$  
- a tuple of the form $(\Delta, \Pi, C, \Lambda, T, \Phi)$

Consent $K$  
- a set of terms accepted by the user, e.g., $K = \{\Theta_1, \Theta_2, \Theta_3, \ldots, \Theta_i\}$.
## Tuples, tables, and set algebra

<table>
<thead>
<tr>
<th>Data type</th>
<th>↑↓</th>
<th>Purpose</th>
<th>↑↓</th>
<th>Company</th>
<th>Country</th>
<th>Duration</th>
<th>↑↓</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>temperature</td>
<td>↑↓</td>
<td>research</td>
<td></td>
<td>Minerva LTD</td>
<td>🇨🇦Canada</td>
<td>1y</td>
<td></td>
<td>daily</td>
</tr>
<tr>
<td>humidity</td>
<td></td>
<td>marketing</td>
<td></td>
<td>ThirstFirst LTD</td>
<td>🇺🇸 USA</td>
<td>1y</td>
<td></td>
<td>hourly</td>
</tr>
</tbody>
</table>

![Venn diagram showing set operations]
Omitting update prompts

- Formal rules (e.g. new terms are more strict)
Omitting update prompts

- Formal rules (e.g. new terms are more strict)
- When is the best time to show a prompt?

A restart has been scheduled

If you want, you can restart now. Or, you can reschedule the restart to a more convenient time. Be sure your device is plugged in at the scheduled time. The install may take 10-20 minutes.

- We’ll schedule a restart during a time you usually don’t use your device (right now 3:30 AM tomorrow looks good).

Select a restart time

Time:

| 3 | 30 | AM |

Day:

Tomorrow
Omitting update prompts

- Formal rules (e.g. new terms are more strict)
- When is the best time to show a prompt?
- Privacy protection gradient
Omitting update prompts

- Formal rules (e.g. new terms are more strict)
- When is the best time to show a prompt?
- Privacy protection gradient
- Decouple feature, security and privacy updates

Broomix v3.4 update

- The update wants to change the way data are handled
- You can keep the current version, if you want to
- You can decide later

Accept the new processing terms?

No, keep current  View new terms  Ask me later
Reduce information volume

**Inline prompts**

---

### Broomix v3.4 update

- The update wants to change the way data are handled
- You can keep the current version, if you want to
- You can decide later

**What has changed:**

<table>
<thead>
<tr>
<th>Data type</th>
<th>Purpose</th>
<th>Company</th>
<th>Country</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>temperature</td>
<td>research</td>
<td>Minerva LTD</td>
<td>Can</td>
<td>1y</td>
</tr>
<tr>
<td>humidity</td>
<td>marketing</td>
<td>ThirstFirst LTD</td>
<td>USA</td>
<td>1y - 3y</td>
</tr>
</tbody>
</table>

Accept the new processing terms?

- No, keep current
- Accept new terms
- Ask me later

---

A. Railean and D. Reinhardt, Improving the Transparency of Privacy Terms Updates
Reduce information volume

- Inline prompts
- Tabular representation and information efficiency

“We care about your privacy, therefore our smart indoor temperature and humidity meter only collects and shares your data with 2 companies. Temperature data are shared on a daily basis with Minerva LTD, located in Canada. The data are retained for a period of 1 year and are used for research purposes. Humidity is shared on an hourly basis with ThirstFirst LTD, and retained by them for 1 year, in the USA. Humidity data are used for marketing purposes.”

<table>
<thead>
<tr>
<th>Data type</th>
<th>↑↓</th>
<th>Purpose</th>
<th>↑↓</th>
<th>Company</th>
<th>Country</th>
<th>Duration</th>
<th>↑↓</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>🌡 temperature</td>
<td></td>
<td>research</td>
<td></td>
<td>Minerva LTD</td>
<td>🇨🇦 Canada</td>
<td>1y</td>
<td></td>
<td>daily</td>
</tr>
<tr>
<td>🌧 humidity</td>
<td></td>
<td>marketing</td>
<td></td>
<td>ThirstFirst LTD</td>
<td>🇺🇸 USA</td>
<td>1y</td>
<td></td>
<td>hourly</td>
</tr>
</tbody>
</table>
Other benefits and potential applications

- Granular permissions
- Search engine – e.g., “devices that store temperature in the EU for research“
- Archive of policies
- Privacy assistant
Summary of our contribution

- Formal notation for privacy terms
- Standardized representation of terms
- Information efficiency as a benchmark
- Concrete steps to reduce
  - Effort to compare
  - Volume of displayed information
Call to action

- **Try our UI:** privacy-facts.eu
- **Provide feedback**
  - weaknesses
  - improvements
  - new use cases
- **Talk to your friends** about it
- **Mention it to policy-makers**

- **Check out our related work**
  - Let there be LITE: design and evaluation of a label for IoT transparency enhancement
  - OnLITE: On-line Label for IoT Transparency Enhancement
  - More at s.gwdg.de/ISn2Eg

arailea@cs.uni-goettingen.de

This research has received funding from the H2020 Marie Skłodowska-Curie EU project "Privacy&Us" under the grant agreement No 675730.
Bonus slides

"You’ve unlocked a secret area!"
### Collected data

<table>
<thead>
<tr>
<th>Hausio T1000</th>
<th>VS</th>
<th>Casami FX</th>
<th>Domowoj</th>
</tr>
</thead>
<tbody>
<tr>
<td>customer nr.</td>
<td>温度</td>
<td>customer nr.</td>
<td>温度</td>
</tr>
<tr>
<td>湿度</td>
<td>Humidity</td>
<td>湿度</td>
<td>Humidity</td>
</tr>
<tr>
<td>设备互联网地址</td>
<td></td>
<td>紫外线</td>
<td>风速</td>
</tr>
</tbody>
</table>

### Sent

- **hourly** to Tesami GmbH
- **daily** to Aster SRL
- **daily** to Domotics s.r.o.
customer number = 481-AHR-1831
temperature = 22 C
humidity = 34%
device Internet address = 93.184.216.34

URL to OnLITE

Privacy facts
Collected data
- customer nr.
- temperature
- humidity
- device Internet address

Sample

Sent hourly to Tesami GmbH

Stored for 3 years in France

All data accessed by
- You
- Tesami GmbH
- 9 partners

Purpose of collection
- your personal use
- scientific research
- targeted advertisements
- product improvement

Received data
Software updates

www.privacy-facts.eu/43dy-kf75
Quantifying useful information

<table>
<thead>
<tr>
<th>Data type</th>
<th>DPV items</th>
<th>bits</th>
<th>P3P items</th>
<th>bits</th>
<th>Apple items</th>
<th>bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>161</td>
<td>8</td>
<td>17</td>
<td>5</td>
<td>32</td>
<td>5</td>
</tr>
<tr>
<td>Duration</td>
<td>31</td>
<td>5</td>
<td>16</td>
<td>4</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Useful information in a privacy term

Δ temperature
Π research
C Minerva LTD
Λ Canada
T 1 year
Φ daily

161 data types
31 purposes
20 symbols per company
193 countries
8 units + n
8 frequencies

39-symbol alphabet {a..z, 0..9, \t, \n, \space}
6 bit/symbol

8 bits
5 bits
20x6=120 bits
8 bits
3+8=11 bits
3 bits

Σ= 155 bits per privacy term
Encodings, their efficiency and interpretability

- Canada
- CA
- CAN
- The world’s second-largest country

🍁 🇨🇦
Privacy rules, like email filters
That’s all, fellow scholars!