Why Has Usability Become Privacy’s Biggest Challenge and What Can We Do About It?

Norman Sadeh
Professor of Computer Science
Carnegie Mellon University

https://normsadeh.org
Privacy Threats Are Everywhere

Data-Hungry Economy
• AI/ML
• IoT sensors everywhere
• Myriads of APIs and dataflows
  and also:
• unscrupulous/ignorant data collectors/processors

Source: CSO online
New Regulations

• Increasingly more specific data practice disclosure requirements
• Increasingly specific data subject rights
• Emerging, yet loosely specified, usability expectations
• New, significantly steeper financial penalties
“All is for the Best in the Best of Possible Worlds” (Voltaire)

### Privacy Policy Before GDPR

<table>
<thead>
<tr>
<th>COMPANY</th>
<th>WORD COUNT</th>
<th>READING TIME*</th>
<th>READING LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>2722</td>
<td>12:54</td>
<td>14</td>
</tr>
<tr>
<td>Facebook</td>
<td>5420</td>
<td>26:24</td>
<td>11</td>
</tr>
<tr>
<td>Reddit</td>
<td>5524</td>
<td>26:42</td>
<td>12</td>
</tr>
<tr>
<td>Amazon</td>
<td>2627</td>
<td>12:48</td>
<td>12</td>
</tr>
<tr>
<td>Wikipedia</td>
<td>2881</td>
<td>13:49</td>
<td>15</td>
</tr>
<tr>
<td>Yahoo</td>
<td>1611</td>
<td>74:04</td>
<td>14</td>
</tr>
<tr>
<td>Twitter</td>
<td>3764</td>
<td>18:37</td>
<td>13</td>
</tr>
<tr>
<td>eBay</td>
<td>5244</td>
<td>25:27</td>
<td>18</td>
</tr>
<tr>
<td>Instagram</td>
<td>2931</td>
<td>14:26</td>
<td>12</td>
</tr>
<tr>
<td>Netflix</td>
<td>3046</td>
<td>15:10</td>
<td>15</td>
</tr>
</tbody>
</table>

### Privacy Policy After GDPR

<table>
<thead>
<tr>
<th>COMPANY</th>
<th>WORD COUNT</th>
<th>READING TIME*</th>
<th>READING LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>4056</td>
<td>19:11</td>
<td>14</td>
</tr>
<tr>
<td>Facebook</td>
<td>4233</td>
<td>20:41</td>
<td>13</td>
</tr>
<tr>
<td>Reddit</td>
<td>3444</td>
<td>16:39</td>
<td>12</td>
</tr>
<tr>
<td>Amazon</td>
<td>3837</td>
<td>18:24</td>
<td>13</td>
</tr>
<tr>
<td>Wikipedia</td>
<td>5617</td>
<td>27:06</td>
<td>14</td>
</tr>
<tr>
<td>Yahoo</td>
<td>2225</td>
<td>11:12</td>
<td>13</td>
</tr>
<tr>
<td>Twitter</td>
<td>4880</td>
<td>22:25</td>
<td>13</td>
</tr>
<tr>
<td>eBay</td>
<td>5666</td>
<td>27:32</td>
<td>20</td>
</tr>
<tr>
<td>Instagram</td>
<td>4221</td>
<td>20:38</td>
<td>13</td>
</tr>
<tr>
<td>Netflix</td>
<td>3477</td>
<td>16:39</td>
<td>16</td>
</tr>
</tbody>
</table>

* Minutes:Seconds

Source: https://www.varonis.com/blog/gdpr-privacy-policy
…And Who Has the Time to Review & Manage All these Privacy Settings?
The Human Bottleneck

Lack of:

- Expertise
- Time
- Attention
- Motivation
- etc.

What If Computers Understood the Text of Privacy Policies?
Annotation Tool

Select a category
Select an attribute
Select a value
Highlight text span for an attribute, value pair

Disclosure of Your Information  Sci-News.com does not sell, trade or rent your personal information to third parties. If we choose to do so in the future, you will be notified by email of our intentions, and have the right to be removed prior to the disclosure.

This policy segment discusses:

• Third Party Sharing/Collection
Automatic Identification of Data Practice Disclosures
Press Coverage – Notice the Irony

FastCompany informing their readers about their new policy

FastCompany’s article about our research

You’re Never Going To Read That Privacy Policy. Could AI Help?

This AI trained on legalese acts like a personal translator of confusing, opaque privacy statements.
Privacy Question Answering

• One-size-fits-all summaries of privacy policies only go so far
• Different people have different questions at different points in time
• Could we develop privacy question answering functionality?
• A number of challenges
  – Can people accurately articulate their questions. If not, how can we help them?
  – How do we provide useful answers – vague policies, inaccurate classifiers
  – etc.


User Choice Instance Extraction

- User choices often buried deep in the text of long policies
- Is it possible to **automatically extract information** about such “choice instances” from privacy policies?
- Use Natural Language Toolkit tokenizer to subdivide segments into sentences & build classifiers

**Results: Recall & Accuracy > 90%**

Opt-Out Easy Browser Extension

- Automatically identify and categorize opt-out choices in the text of privacy policies
- And present them in an easy-to-use interface to users as they browse the web
- Available in Google Chrome store and Firefox store - Watch our video
Can We Automatically Check for Potential Compliance Issues?

- Training **machine learning classifiers** to extract relevant policy statements
- Compare these statements against:
  - **Regulatory requirements**
  - What the software actually does
    - **Static and dynamic code analysis**

Reports View with Search and Filter

- **Search and Filter according to**
  - **Practices** (e.g., GPS location collection)
  - **Type** (Is GPS location not mentioned in the policy or does the policy explicitly state it is not occurring?)
  - **Specificity** (Perhaps “location” suffices in some jurisdictions not specifically requiring to mention “GPS.”)
  - **Many more ...**
# Flashlight App: Policy Results

Relevant parts of the policies are extracted and displayed alongside the analysis results

<table>
<thead>
<tr>
<th>Metadata</th>
<th>Potential Compliance Issues</th>
<th>Policy Analyses</th>
<th>Static App Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Complete Policy Analyses</td>
<td>Static App Analysis</td>
<td>Supporting Policy Segments</td>
</tr>
<tr>
<td></td>
<td>Completed Analysis (2 Sources)</td>
<td>Identifier 1st Party</td>
<td>2. COLLECTION &amp; TRANSMISSION OF PERSISTENT IDENTIFIERS. Goldenshores Technologies and its subsidiaries and agents may collect your persistent identifiers, such as a customer number held in a &quot;cookie,&quot; a static Internet Protocol (&quot;IP&quot;) address, a mobile device ID, or processor serial number. Goldenshores Technologies and its subsidiaries and agents may transmit that information to third-party services providers for the purpose of providing you better services in.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Policy Analysis Practice</th>
<th>Performed</th>
<th>Not Performed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifier 1st Party</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Identifier 3rd Party</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Identifier Cookie or similar Tech 1st Party</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Identifier Cookie or similar Tech 3rd Party</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Identifier Device ID 1st Party</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

© Copyright 2022, Norman Sadeh, All rights reserved.
Flashlight App: Static Analysis Results

Relevant parts of the code are extracted and displayed alongside the analysis results.

<table>
<thead>
<tr>
<th>Static Analysis Practice</th>
<th>Network Call</th>
<th>Third Parties</th>
<th>App Permissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifier SSID BSSID 3rd Party</td>
<td>3 Segments</td>
<td>com.millennialmedia, com.inmobi</td>
<td>INTERNET, ACCESS_WIFI_STATE</td>
</tr>
<tr>
<td>Identifier IMEI 3rd Party</td>
<td>4 Segments</td>
<td>com.flurry, com.amobee, com.tapit</td>
<td>READ_PHONE_STATE</td>
</tr>
<tr>
<td>Identifier Cookie or similar Tech 3rd Party</td>
<td>2 Segments</td>
<td>com.google</td>
<td>INTERNET</td>
</tr>
<tr>
<td>Location WiFi 3rd Party</td>
<td>11 Segments</td>
<td>com.flurry, com.millennialmedia, com.tapit, com.inmobi</td>
<td>ACCESS_FINE_LOCATION</td>
</tr>
<tr>
<td>Location GPS 3rd Party</td>
<td>11 Segments</td>
<td>com.flurry, com.millennialmedia, com.tapit, com.inmobi</td>
<td>ACCESS_FINE_LOCATION</td>
</tr>
<tr>
<td>Location Cell Tower 3rd Party</td>
<td>16 Segments</td>
<td>com.flurry, com.millennialmedia, com.tapit, com.inmobi</td>
<td>ACCESS_COARSE_LOCATION, ACCESS_FINE_LOCATION</td>
</tr>
</tbody>
</table>

© Copyright 2022, Norman Sadeh, All rights reserved.
Analysis of over 1 million Android Apps in Google Play Store

• Average number of potential compliance issues per app is 3.47 and the median is 3
• Requires manual vetting – both policy and app behavior to confirm potential compliance issue
Developers Struggle with 3rd Party APIs

- Lighter colors indicate greater transparency of practices. Darker colors indicate that practices are being performed but not disclosed.
- Cells with fewer than 25 apps performing the practice are annotated with the respective number of apps.
Other Collaborations

• Collaboration with California Attorney General’s office
• COPPA report compiled for Federal Trade Commission
  • Focusing on location, apps with a large number of downloads, and companies based in the US
• CDT report on mobile apps for connected cars
• Work with large European electronics manufacturer – checking for GDPR compliance of mobile apps
Tools for Developers

1. Template Provisioning
   E.g., the GDPR requires policies to notify users of their rights to request data access, rectification, erasure, restriction of processing, objection of processing, and portability (Art. 13(2)(b))

2. Code Analysis
  .plist Permissions and Swift API Usage

3. Wizard Fine Tuning
   Don’t worry! I got you covered. Here is the privacy analysis of your app. You can adjust the generated privacy policy via the checkboxes.

Compliance traceability: Privacy policies as software development artifacts, S Zimmeck, P Story, R Goldstein, D Baraka, S Li, Y Feng, N Sadeh, Privacy, Usability and Transparency Workshop (PUT 2019) at PoPETs 2019 conference
Could Computers Also Help Motivate People to Take Advantage of Privacy Settings?
Nudging Users: Surprise People with Something That Will Motivate Them to Pay Attention

Did you know?
Your location has been shared 5398 times with Facebook, Groupon, GO Launcher EX, and 7 other apps for the past 14 days.

Number of times your location has been shared with each app for the past 14 days.

<table>
<thead>
<tr>
<th>App</th>
<th>Times Shared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Play services</td>
<td>1603</td>
</tr>
<tr>
<td>Android System</td>
<td>1602</td>
</tr>
<tr>
<td>Groupon</td>
<td>1602</td>
</tr>
<tr>
<td>Weather &amp; Clock Widget</td>
<td>296</td>
</tr>
<tr>
<td>GO Launcher EX</td>
<td>255</td>
</tr>
<tr>
<td>Maps</td>
<td>18</td>
</tr>
<tr>
<td>Viber</td>
<td>11</td>
</tr>
<tr>
<td>Facebook</td>
<td>5</td>
</tr>
<tr>
<td>Google Search</td>
<td>3</td>
</tr>
<tr>
<td>MyFoodCoach Study</td>
<td>3</td>
</tr>
</tbody>
</table>

Let me change my settings
keep sharing my location


© Copyright 2022, Norman Sadeh, All rights reserved.
Nudging Pilot – 3-week study

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Week 2: Permission Manager Only**
- Permission Manager Helps (week 2)
- Nudges can make a big difference (week 3)
Could Computers Help People Manage their Privacy Controls?

Many privacy decisions are repetitive or similar... and tedious
Even simple solution with small number of clusters achieves high levels of accuracy

Privacy Assistants II

Generating recommendations rather than automating privacy decisions

Vast majority of recommendations accepted by users and kept despite nudges to reconsider. Successfully deployed in Google Play store for several users – rooted Android phones only…
Why Recommendations?

Agency is a major part of privacy: users should remain in charge of their decisions…but **AI can help** them make these decisions and can help overcome fundamental usability limitations

- **Major requirement:** the recommendations have to be **understandable** and **auditable**
Similar Results with Other Privacy Decisions

- Similar results with browser and IoT privacy decisions
- The challenge is that access to these privacy settings is generally not open
- **Would need regulation to make this possible**…
  - …just think about the number of times you answer the same cookie questions…

How about the Internet of Things?
IoT: Additional Usability Challenges

- How likely are you to notice this sign?
- Does this include facial recognition?
- What about facial expression or scene recognition?
- How long is the data retained?
- Do I get to opt in/opt out?
- Is this GDPR compliant?
Zhang, Y; Feng, L; Bauer, LF; Cranor, A; Das, N; Sadeh, N. "Did you know this camera tracks your mood?": Understanding Privacy Expectations and Preferences in the Age of Video Analytics", Proceedings on Privacy Enhancing Technologies, 2, 1, Apr 2021 [pdf]

© Copyright 2022, Norman Sadeh, All rights reserved.
Good Luck with this…

QR code for CCPA compliance at knee height…
An IoT Privacy Infrastructure

- Portal to help people publicize the presence and data practices associated with IoT systems, incl. privacy controls (e.g. opt-in/out, deletion, etc.)
- IoT Assistant app (Android and iOS) for users to discover nearby IoT systems and their data practices, incl. accessing any available privacy controls

https://www.iotprivacy.io

US Patents 10,956,586 with additional USPTO and EPO patents pending
IoT Privacy Infrastructure

- The IoT Assistant App video
- IoT Privacy Infrastructure Overview video
- Hosting nearly 150,000 IoT system entries today

Concluding Remarks - I

• Privacy is becoming increasingly complex
  – Everyone is collecting our data, increasingly complex data flows
  – Smartphones, IoT, AI/ML
• New regulations have been introduced that are in great part motivated by these developments (e.g., GDPR, CCPA/CPRA)
• These regulations are an important step in the right direction
• Yet, in the absence of better technologies, they make usability even more challenging
Concluding Remarks - II

• My collaborators and I have been working on the development of technologies that aim to **mitigate these usability challenges**
• Our vision: **Privacy Assistants** that leverage
  • Techniques designed to empower people to take **advantage of more detailed privacy policy disclosures**
  • Techniques designed to motivate people and assist them with the **management of** an increasingly unmanageable number of privacy decisions
• Using techniques such as AI/ML or privacy nudges gives rise to complex **ethical issues** and requires **careful evaluation**
• Some of our efforts to help people manage their privacy settings call for **regulations that mandate privacy APIs** - **without these APIs privacy will remain unmanageable**.
The **Usable Privacy Policy Project** and the **Personalized Privacy Assistant Project** involve collaborations with a number of individuals.

More details at:

https://usableprivacy.org
https://privacyassistant.org
https://explore.usableprivacy.org
https://www.iotprivacy.io