He's making a list
He's checking it twice
He's gonna find out who's naughty or nice
Santa Claus is in contravention of article 4 of the General Data Protection Regulation (EU) 2016/679

9:39 AM - 20 May 2018

19,820 Retweets 45,057 Likes
The Seven GDPR Sins of Personal-Data Processing Systems

Supreeth Shastri, Melissa Wasserman, Vijay Chidambaram

TEXAS
The University of Texas at Austin
General Data Protection Regulation (GDPR)

**May 25, 2018**
Adopted after 2 years of public debate.
All but 2 EU countries have legislated.

**Fundamental right**
Grants all European people a right to protection and privacy of personal data

**Personal data**
Any information relating to a natural person;
Broad in scope unlike FERPA, HIPAA

**Covers entire lifecycle**
Collection, processing, protection, transfer and deletion; Regulated via 99 articles

**Hefty penalty**
Max penalty of 4% of global revenue or €20 million, whichever is greater
GDPR Entities

- **Controller** (e.g., Spotify)
  - share personal data
  - send personal data for external processing
  - store and process personal data internally

- **Processor** (e.g., Google cloud)
  - notification of data breaches
  - audit and investigate

- **Supervisory Authority**
  - report GDPR violations
  - audit and investigate

- **Data Subject** (e.g., Spotify user)
  - provide personal data
  - exercise GDPR rights

- **Other Controllers** (e.g., SoundCloud)
  - allow data sharing
  - share personal data

- **GDPR queries**
  - personal data
  - GDPR rights
Terminated

Adapted

Advertised compliance

Assumed compliance

<50% estimated compliance
By the end of 2018 [Gartner 2018]

94,622 complaints from people
In the first 9 months of GDPR rollout
Internet-era systems have *primarily* focused on reliability, scalability, and affordability.

**KEY OBSERVATION**

Relegating security and privacy as afterthoughts has given rise to **principles** and **practices** that are *at odds* with GDPR.

**The Seven GDPR Sins**
1. Storing Data Forever

§17: Right To Be Forgotten

(1) The data subject shall have the right to obtain from the controller the erasure of personal data without undue delay [...] 

§ 5(1)(E): Storage Limitation

“(…) kept for no longer than is necessary for the purposes for which the personal data are processed [...]”

Before

ars TECHNICA
YOUR REPUTATION PRECEDES YOU —
Google extends right-to-be-forgotten rules to all search sites
KELLY FIVEASH - 3/7/2016, 7:47 AM

After

180
days

Time that Google cloud requires to guarantee that a requested personal data item is fully deleted
2. Reusing Data Indiscriminately

Facebook is using your 2FA phone number to target ads at you

Reported by GIZMODO on 9/26/2018

On Jan 21st 2019, the French DPA levied the largest GDPR fine yet on Google for purpose bundling

§ 5(1)(B): Purpose Limitation
“Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes [..]”

§ 21: Right To Object
“(1) The data subject shall have the right to object at any time to processing of personal data concerning him or her [..]”
3. Creating Black Markets and Walled Gardens

**Before**

3000+

Unique data points per US consumer

**After**

Many programmatic ad exchanges shut down

§14: Information To Be Provided Where Personal Data Have Not Been Obtained From The Data Subject

“(1) (c) the purposes of the processing [...], (e) the recipients [...]. (2) (a) the period for which the personal data will be stored [...]. (f) from which source the personal data originate [...].”

§ 20: Right to Data Portability

“(1) The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller. (2) […] the right to have the personal data transmitted directly from one controller to another.”

Source: 2014 FTC report on 9 largest personal data brokers
4. Risk Agnostic Data Processing (a.k.a Move fast and break Things)

§ 35: Data Protection Impact Assessment
"Where processing, in particular using new technologies, is likely to result in a high risk to the rights of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing."

§ 36: Prior Consultation
"The controller shall consult the supervisory authority prior to processing where [...] it would result in a high risk in absence of measures taken by the controller to mitigate the risk."

User accounts hacked in 2018, after Facebook’s View-As feature was exploited.
5. Hiding Data Breaches

§ 33: Notification of A Personal Data Breach
(1) the controller shall without undue delay and not later than 72 hours after having become aware of it, notify the supervisory authority. [...] 
(3) The notification shall at least describe the nature of the personal breach. [...] likely consequences, and [...] measures taken to mitigate its adverse effects. 

Breaches in the real world

Reported data breaches 6 months before and after GDPR

<table>
<thead>
<tr>
<th>Before GDPR (worldwide)</th>
<th>After GDPR (only Europe)</th>
</tr>
</thead>
<tbody>
<tr>
<td>945</td>
<td>41,502</td>
</tr>
</tbody>
</table>
6. Making Unexplainable Decisions

A Popular Algorithm Is No Better at Predicting Crimes Than Random People

The COMPAS tool is widely used to assess a defendant’s risk of committing more crimes, but a new study puts its usefulness into perspective.

§ 15: Right of Access

“(1) The data subject shall have the right to obtain from the controller […] meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing.”

§ 22: Automated Individual Decision-Making

“(1) The data subject shall have the right not to be subject to a decision based solely on automated processing […]”

Workshop on Human Interpretability in ML

Workshop on Explainable AI
7. Security as a Secondary Goal

§ 25: Data Protection By Design and By Default

*(1) [...] design to implement data protection principles in an effective manner [...]*

§ 24: Responsibility of the Data Controller

"the controller shall [...] be able to demonstrate that processing is performed in accordance with this Regulation."

Security in the real world

Amazon Macie

ML-driven reactive security
Concluding Remarks

FUTURE DIRECTIONS

**GDPR-compliant Redis**
Exploring system-level tradeoff in achieving compliance

**Cloud consolidation**
Could compliance be better tackled at cloud provider level?

**Beyond GDPR**
California’s CCPA is going into effect 1/1/2020

We want to hear from you!

https://utsaslab.github.io/research/gdpr/