A First Look at the General Data Protection Regulation (GDPR) in Open-Source Software

Franke, Liang, Brantly, Davis & Brown

PRESENTER
James C. Davis
Asst. Prof., ECE@Purdue

BACKGROUND

Enacted 2016, enforced from 2018
BILLINGS IN FINES

OSS + GDPR?

- Majority of OSS is for commercial use
- Synopsys: In 1700 codebases across 17 industries, ~75% of code was OSS
- Software supply chain – users want compliance

OSS software engineers find GDPR compliance challenging.

GDPR slows them down, it is hard to validate, and they want tools.

Example question	Example codes
What impact has GDPR had on data security and privacy?	Rights to users, data privacy, fines
What GDPR concepts do you find the most difficult to implement?	Data minimization, Privacy by design, User experience
Have you sought legal counsel on GDPR-related issues?	Positive and negative views Effects (cost, time, data storage, data processing)

METHODS

1. Pilot study and data familiarization (3 interviews)
2. Survey design
3. Participant recruitment (GitHub repos with GDPR PRs && Twitter/Reddit) – N=47
4. Data analysis (open coding, 2 analysts)

QUOTES

IMPACT ON PRODUCTS AND ORGS.
- "The risk of...hefty fines has made companies take privacy and security more proactively"
- "The cost of protection should not go over the cost of consequence of data breach"
- "Weakens small- and medium-sized enterprises"
- "Jobs for lawyers...[not] people trying to solve real problems"
- "Open-source projects can't afford even to sustain maintainers, not even speaking about legal team"
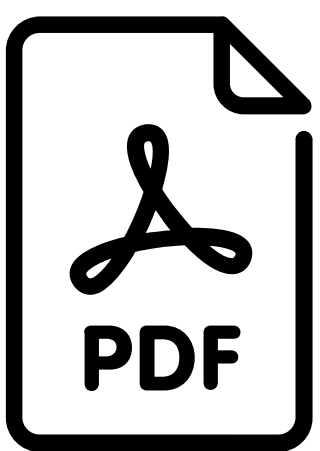
IMPACT ON DEVELOPMENT PROCESS
- "It slows things down as code has to be reviewed and objectives rvised"
- "Consider whether you really need all the data you collect"
- "[GDPR work] showed things we had not considered before...logging...access restrictions"

TECHNOLOGY GAPS
- "Really no good way to evaluate [compliance]"
- "Essentially incompatible with big data and AI"

WHAT'S NEXT?

- Second iteration of survey
- Mine OSS: Patterns, Defects, Validation

Lab website
(Want to join?)