OSS in the quest for GDPR compliance

Pass the Salt 2019
Errata

- this talk was proposed by Cristina DeLisle
- I'm filling in due to a scheduling conflict
- I Am Not A Lawyer
Agenda

1. XWiki & CryptPad, who we are, what we do
2. what we talk about when we talk about privacy
3. about CryptPad
4. GDPR: our experience, implications for open-source
$ whoami

- Aaron MacSween 🍁
- Privacy engineer & researcher, applied cryptographer
- CryptPad project lead
- XWiki SAS 🇫🇷 (Paris, France)
What is XWiki?

- ~40 person organization
  - France, Romania, Spain*, Germany*, Belgium*
- enterprise knowledge management software
  - the open-source XWiki platform
- in business for 15 years
- ...but how does this fit into *Pass the Salt*?
Privacy and security are often "added at the end"

- ...and it doesn't work
- and it has terrible consequences
- and we'd like to change that
- but...
There's no single fix

- privacy and security are complicated
- they're context dependent
XWiki knows a lot about knowledge management

- it's one small piece of the puzzle (privacy)
- we research how to advance the state of the art
Privacy & Security

- from whom? the NSA? your little brother?
- for how long? until you're out of the country?
- what are you protecting or hiding?
- what's your threat model?

In short, the two don't always go together.
Security with less privacy

- anti-fraud policies
  - protection via surveillance
- 2FA
  - something you know, something you have
Privacy with less security

- "zero knowledge" web services
  - pastebins, file upload, *X but with encryption*
- no 2FA, but no third parties
CryptPad: c'est quoi?

- real-time like Etherpad or Google docs, but with encryption
- e2ee collaboration suite
- fully open-source (AGPL), 250+ instances in the wild
Our architecture

- browser-based "thick client"
- p2p conflict resolution with pluggable encryption
- multiple editors with compatible APIs and UIs
- mostly dumb websocket store-and-forward server
  - like IRC channels but with history
- append-only logs on the server filesystem
- cryptographic keys and document ids shared as URLs
Extensions

- "CryptDrive" (just another document)
- cryptographic login (via Scrypt)
- read/write/delete capabilities
- public-key authenticated RPCs
- encrypted files embedded in documents
- shared folders
- "Friends" and write-only "Mailboxes"
- private messaging and embedded group chat
Our users

- The pirate party of Germany (self-hosted)
- C3W (CCC Vienna, self-hosted)
- various other activist groups, hackerspaces
- 12K registered on our instance
- about 10K unique IPs each week
Funded by...

- French R&D grants (merci BPI France)
- NLnet Foundation (NGI PET)
- donations: opencollective.com/cryptpad
- subscriptions on CryptPad.fr
But that makes us responsible for other people's data...
Handling data

- General Data Protection Regulation (GDPR)
- in effect since May 2018
- unified set of data protection laws
- formal recognition of encryption as best practices
Our strategy

- **Privacy by Design**
  - read the docs: "Seven foundational principles"
- data minimization
  - "who needs to know?"
- challenge conventional wisdom, find alternatives to PII
  - (Personally Identifying Information)
Roles and definitions

- Data Protection Officers
- Data controllers
- Data processors
- Lawful processing
DPOs

- **Data Protection Officer**
- one of Cristina's roles at XWiki
- can be adversarial in nature
- audits policies, keeps inventories of PII
- formalize access control strategies
- 30 days to respond to queries
Data controllers

- the organization which employs the DPO and holds the data
- set privacy policies and strategies for the data's lifecycle
- proactively demonstrate compliance
- process PII lawfully, with informed consent
Data processors

- third parties involved in handling your data
- defined in a Data Processor Agreement
- For us:
  - OVH (hosting)
  - Stripe (payments)
  - Quaderno (invoicing and regional tax rates)
Lawful processing

- compliance with the law
- contractual reasons
- involving consent of the data subject
- legitimate* interest
Fines for violations

- coerced or forced "consent"
- not reporting confidentiality or availability breaches
- up to 4% of annual global turnover or €20 million
  - whichever is greater.
GDPR and OSS

• forces cloud infrastructure to be more accountable
• protects and empowers data subjects
• raises awareness of privacy and the risks of proprietary platforms
Uncertainty

- at what point does a self-hoster become a controller?
- what schemes are best?
- what's the right way to handle data?
- how do we challenge "legitimate interest"?
- what can be considered a reasonable effort?
Conclusions

Privacy advocates still need lots of help:

- from dedicated security experts
- from domain expert
- POC implementations for different problems
Questions?

Come say hi after:

- if you want **stickers** or...
- if you're interested and eligible for **EU R&D** projects
• Aaron | ansuz
  ○ https://social.privacytools.io/@ansuz
  ○ https://twitter.com/fc00ansuz
• Cristina
  ○ cristina.rosu@xwiki.com
  ○ https://mastodon.social/@redchrision
• CryptPad
  ○ https://twitter.com/cryptpad
  ○ https://social.weho.st/@cryptpad