École Pour l'Informatique et les Techniques Avancées – EPITA

Masters program – 29 April 2022

Course: Data Privacy by Design

Course instructor: M Salman Nadeem

mohammad-salman.nadeem@epita.fr

Data Privacy by Design (PbD)

Course schedule (tentative)

Date & Time	No.	Topics	Duration (in hours)
04/03/2022 14:30-17:30	1	Data & its types, Information & knowledge, Introduction to Data Privacy by Design (PbD)	3 hours
18/03/2022 14:30-17:30	2	DPbd Case studies, Data privacy risks & solutions	3 hours
02/04/2022 10:00-13:00	3	Privacy Enhancing Technologies (PET's)	3 hours
22/04/2022 14:30-17:30	4	General Data Protection Regulation (GDPR), PbD and GDPR	3 hours
29/04/2022 14:30-17:30	**5**	**Open session, Putting it all together, Quiz, Final project presentation**	**3 hours**
		Total Lecture (hours)	***15***

Evaluation: 10% Class attendance + 10% Class participation + 30% Class/home exercises + 50% Final Evaluation

Lecture 5 Outline

- **Review**
- Open session
  - Protecting yourself
  - Spreading awareness
  - Q/A
- Closing
  - Evaluation
  - In conclusion

Review

Recalling previous lectures

Privacy by design (Objective, Strategies, Activities)

+

Always remember the Crypto package!

+

Use of appropriate anonymization/pseudonymisation techniques and PETs

+

Ensure threat detection and security controls

+

GDPR (Key definitions, Lawful basis for processing, Data Privacy by design & by default, Individual (subject) rights, Accountability and governance, Cross-border data transfers, Security, Data breaches, Sanctions & Fines, Other aspects)

Lecture 5 Outline

- Review
- **Open session**
  - **Protecting yourself**
  - **Spreading awareness**
  - **Q/A**
- The end
  - Evaluation
  - In conclusion

Present and future!

- "33 bits of entropy are sufficient to identify an individual uniquely among the world's population"
- Attacks only get better with time
  - Privacy should rest on provable guarantees rather than the absence of known attacks
    - Burden of proof be on the data controller to affirmatively show that anonymized data cannot be linked to individuals, rather than on privacy advocates to show that linkage is possible
- Paul Ohm warned of the "database of ruin", a single, massive database containing secrets about every individual, formed by linking different companies' data stores
  - Today there is a booming market for these linkages between different companies' data stores
  - Some companies also display privacy theater

Open session (1/2)

- Protecting yourself: **Some** recommendations….
  - Choosing an app/service: making informed decisions.
    1. Make clear separation between your work and private apps/services/tools.
       - In case you prefer or follow different online/digital identities, then make sure to isolate them properly.
    2. Consider different factors: Opensource? Company? License? Based in? Security & Privacy? What data is required to be shared?...
    3. How many different data points a given app/service will have on you once you start using the app/service?
    4. Keep a backup plan (what to do in case the service/app gets breached/or goes rogue, which GDPR rights to exercise, …).
       - Stay vigilant: Follow news – act swiftly, …
    5. Upon no longer using a given app/service, get rid of your digital traces.
  - Following good security/privacy practices:
    - Guides/How-to's: ssd.eff.org, securityinabox.org, datadetoxkit.org, myshadow.org, ftxreboot.wiki.apc.org, communitydocs.accessnow.org, digitalfirstaid.org, securityplanner.consumerreports.org…
    - Tools: privacytools.io, …
    - Be consistent!

Open session (2/2)

- Spreading awareness:
  - **Open-exercise**: Following existing group setting:
    - Reflect on the knowledge you have obtained so far in this course (regarding data privacy risks and their mitigation) to:
      1. Propose **one** approach (per group) for spreading the word/awareness, that you think would be effective
      2. Write it on the board
- Play your role:
  - Educate and empower others

Full package!

Data is a commodity

+

Privacy by design (Objective, Strategies, Activities)

+

Remember the Crypto package

+

Use of appropriate Anonymization/pseudonymisation techniques and PETs

+

Ensure general threat detection and security controls

+

GDPR (Key definitions, Lawful basis for processing, Data Privacy by design & by default, Individual (subject) rights, Accountability and governance, Cross-border data transfers, Security, Data breaches, Sanctions & Fines, Other aspects)

+

Protecting yourself and spreading awareness!

Lecture 5 Outline

- Review
- Open session
  - Protecting yourself
  - Spreading awareness
  - Q/A
- **The end**
  - **Evaluation**
  - **In conclusion**

Evaluation (1/2) (individual assignment)

- Check your assignment topic in 'Class notebook' (Microsoft Teams)

1. Prepare a 3-page assignment document (.doc*)
   - **First page**: Introduction, Details of the breach (when/how) using public information
   - **Second page**: What factors lead to that breach (use your course knowledge, OWASP Top-10 privacy risks project), etc.
   - **Third page**: What would you suggest to avoid such breach in the future (use your course knowledge, OWASP Top-10 privacy risks project, etc.

   Note: Any out-of-scope security assessments/recommendations will NOT be accepted. Keep your focus towards data privacy risks and their mitigation

   1. Export your document as firstname_lastname.doc*
   2. Submit it in MS Teams: Assignment section

Inspire from case studies done in class, use your crypto package and PETs

Deadline: See 'Teams' Assignment section

Evaluation (2/2) (group presentation)

- Case study (Create a **group of 3**) and choose a timeslot (in class notebook)
  1. Pick one of the following applications:
     1. Cryptpad (https://cryptpad.fr -> collaboration, productivity)
     2. Signal (https://signal.org/ -> instant messaging)
     3. Mastodon (https://mastodon.social -> social media)
     4. Mailfence (https://mailfence.com -> email-suite)
  2. Study the application (privacy policy/terms of service, working, processes/procedures, features, …)
  3. Propose a basic reference model, and identify **data privacy-based risks**
  4. Propose a strategy and/or techniques to mitigate those data privacy risks (System design, Tools/Techniques, Procedures, Roadmap, …)
- Presentation slides order: **(presentation time: 10 mins)**
  - Slide 1: Introduction (study the app, include relevant info.)
  - Slide 2: Draw **basic** reference model (based on your study of the app)
  - Slide 3: Perform 1st, 2nd and 3rd PbD activities:
  - Slide 4: Perform 4th PbD activity: Final proposed solution
    - Don't forget to apply your crypto package, and PETs (that we have discussed in this course)
    - Any out-of-scope security assessments/recommendations will NOT be accepted!
    - Keep your focus towards possible data privacy risks and their mitigation!

CryptPad

MASTODON.SOCIAL

MailFence

Inspire from case studies done in class, use your crypto package and PETs

Deadline: See 'Teams' Assignment section

In conclusion

- The consequences of getting it wrong are severe
  - Equally, however, are the positive consequences of getting it right
- Do more with your data, without the risk of having to stop
  - Strong internal data protection and security controls
- Be future-proof internationally
  - Countries adopting GDPR-style rules (e.g., Sep 2018: Colorado Data Privacy Act, Feb 2020: Brazil LGPD, …)
- Gain individual's trust
  - Privacy increasingly important for consumers
  - Foster trust with customers and partners alike

Lecture 5 ends here

- Course Slides: Go to MS Teams:
  'Data Privacy by Design Spring (S1) Spring 2022' -> Files section
- Send your questions by email:
  mohammad-salman.nadeem@epita.fr
  OR via direct message using MS Teams
- Thank You!

Course references

- Privacy by Design in Law, Policy and Practice – A White Paper for Regulators, Decision-makers and Policy makers [http://www.ontla.on.ca/library/repository/mon/25008/312239.pdf]
- Engineering Privacy by Design Reloaded – KU Leuven [https://www.esat.kuleuven.be/cosic/publications/article-2589.pdf]
- Systematic Privacy by Design engineering, Systematic design of privacy-preserving systems: Privacy by Design Reloaded – Carmela Troncoso [http://carmelatroncoso.com/]
- GDPR [http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679]
- European commission: Rights for citizens [https://ec.europa.eu/info/law/law-topic/data-protection/reform/rights-citizens_en]
- Wikipedia [https://www.wikipedia.org/]
- OWASP [https://www.owasp.org]
- Access now [https://www.accessnow.org]
- School of data [https://schoolofdata.org]
- Tactical Tech [https://tacticaltech.org/]
- Cloudfare Blog [https://blog.cloudflare.com/validating-leaked-passwords-with-k-anonymity/]
- Robust de-anonymization of large sparse datasets: a decade later [randomwalker.info/publications/de-anonymization-retrospective.pdf]