Privacy at Speed: Privacy by Design at Uber

Engin Bozdag
Senior Privacy Architect, Uber
Outline

• Introduction
• Part 1: GDPR and Privacy by Design
• Part 2: Challenges and Strategies
• Part 3: Q&A
Part 1

Privacy By Design
Privacy by Design (Since 1995)
GDPR Art 25: Which factors?

- **Appropriate Measures and Principles**
- **State of the art**
- **Nature, scope, context and purpose**
- **Effective (Demonstrate with metrics)**
- **Costs**
- **Risk for the individual**
When to implement the controls?

- Decision making on the tech stack, vendor
- Abstract, engineering design document, prototypes

- Implementation is effective
- Periodic vendor reviews
- Data breaches
- Data deletion
Part 2

Challenges and Strategies
Existing Guidelines

EDPB Plenary meeting, 12-13 November 2019

Guidelines 4/2019 on Article 25
Data Protection by Design and by Default

Adopted on 13 November 2019

<table>
<thead>
<tr>
<th>Transparency</th>
<th>Accuracy</th>
<th>Data Minimization</th>
<th>Privacy Rights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lawfulness</td>
<td>Purpose Limitation</td>
<td>Storage Limitation</td>
<td>Privacy by Default</td>
</tr>
</tbody>
</table>
Microservices

iOS App  Android App  Web Service
Write Code

CI Testing

Build a Release

Land Code

Deploy Release

Test Release

another service

Uber iOS App

another service

Uber Android App

another service

uber-web service
Agile Development

Waterfall vs. Agile

- Requirements
- Design
- Implementation
- Testing
- Maintenance

Agile
- Backlog
- Design
- Test
- Analyze
- Deploy

Waterfall
- Requirements (fixed)
- Time (fixed)

Agile
- Requirements (variable)
- Time (variable)
- Resources (variable)
Challenge 1: System Characterization

- Distributed data
- Mix of structured + unstructured data
- Off-the-shelf tools don’t scale
- No (stable) architectural documentation
### Data Classification

- **Tier 1: Highly Restricted**
- **Tier 2: Restricted**
- **Tier 3: Confidential**
- **Tier 4: Public**

### Example Category

- **Government Identifier & location**
  - Driver’s License
- **Vehicle Data**
  - License Plate Number
  - Proof of Insurance
- **Non-Identifying Vehicle Data**
  - Make and Model Color
- **Public Information**
  - Product Brochures
Uber’s Approach: Data Classification
Uber’s Approach: Data Inventory

- Unified data category tags
- Automatic tagging and verification
- Maturity Levels: Tagging at DB level, tagging at column level, identify ALL data of an individual
- Use Data Inventory results to improve processes
Challenge 2: Threats and Mitigation

• Privacy threat of the service vs the whole chain
• Where to place the control?
• Privacy can be slow vs Agile fast
• Legacy systems and privacy debt
• Resource and costs
Modular Reviews: Technical Privacy Consulting

- Which personal data?
- Purpose and Access
- Retention
- Source And Recipients
- Storage Location
- Onboarding to PrivEng Services
Modular Reviews: Technical Privacy Consulting

- **Outcome:**
  - Technical privacy requirements for this specific project
  - Mitigation prioritization
  - Input to Privacy Legal

- **After the review:**
  - Further analysis of platforms based on discovered knowledge
  - Embed privacy into platforms
  - Update Data Classification and Handling Standard
Uber Approach: Education and Privacy Champions

Privacy Expert

Privacy Network

Champion

Domain 1

Champion

Domain 2

Champion

Domain 3

Champion

Domain 4

Champion

Domain 5
Challenge 3: Doing Privacy at Scale: Deletion

- Multiple use cases: user initiated account deletion, inactive account deletion, time-based deletion
- Variety of data stores
- Scalable, reliable, adaptable, demonstrable
Uber’s Approach to Data Deletion

- Support scale of data, data stores, and microservices
- Privacy Impact Assessment and Technical Privacy Reviews
- Vetting process combines legal and technical privacy
- Automate onboarding process for new services
Conclusion

• Privacy controls need to be chosen by the organization
• Existing PbD guidelines do not address the challenges of agile development in complex environments
• Solution:
  • Fix what you can in the project
  • Discover the bigger challenges
  • Monitor and iterate
• Provide privacy tools/services that are easy to adapt by engineers