Privacy Tools and Techniques for Developers

-Amber Welch
Amber Welch

MA, CISSP, CISA, CIPP/E, CIPM, FIP, CCSK, and ISO 27001 Lead Auditor

linkedin.com/in/amberwelch1

github.com/msamberwelch

@MsAmberWelch
Menu

- Privacy Engineering Intro
- Privacy by Design
- Privacy Enhancing Technologies

bit.ly/2WXJTeR
First, an apology.
Legal teams have often kept tech out of privacy.
Developers don’t know privacy concepts. Privacy teams haven’t taught them.
Privacy Impact Assessment
A Privacy Impact Assessment (PIA) is a method to:

- Identify privacy risk
- Map personal data flows
- Document privacy risk mitigations
- Fulfill regulatory requirements
Privacy Impact Assessment throughout an initiative

IDEA

DEFINITION

DESIGN

ACTION (FOR EXAMPLE SOFTWARE DEVELOPMENT, PURCHASE, CONTRACT, POLICY DRAFTING)

TEST

USE

Decide whether to do a PIA

PIA and ongoing updates with change

Review
Use Cases

- New applications
- Adding functions and features
- Collecting new sensitive personal data
- Annual reviews or audits
Tasting Notes

Benefits

- Legal compliance
- Identify and reduce privacy risks
- Catch privacy errors
Tasting Notes

Benefits

● Legal compliance
● Identify and reduce privacy risks
● Catch privacy errors

Limitations

● High time investment
● Ineffective if not completed well
● Not a security risk assessment

bit.ly/2qbrnu5
Data Minimization and Retention
Data minimization is:

- Collecting only necessary data
- Maintaining and updating data
- Deleting old data that isn’t needed
Use Cases

- New applications
- API integrations
- Adding functions and features
- Collecting new personal data
- Customer termination
Tasting Notes

Benefits

- Legal compliance
- Minimize volume of data to be breached
- Improve data quality
Tasting Notes

Benefits
● Legal compliance
● Minimize volume of data to be breached
● Improve data quality

Limitations
● Users may be frustrated
● Companies like to keep all the data

bit.ly/2qbrnu5
Default Settings
Default settings for privacy should:

- Minimize personal data collected
- Prevent default data sharing
- Require enabling of intrusive settings
- Avoid making data public by default
Less than 5% of general users change any default settings, while programmers change 40% of settings.
Manage your preferences for Enhanced Relevant Advertising

Enhanced Relevant Advertising uses information generated by all users of the AT&T products and services (Internet, video, and mobile) on your account to deliver a more personalized experience. The ads you see will be tailored to your likes and interests. You won't see more ads. This information includes: TV viewing, Web browsing, app usage, location, call detail records, and other Customer Proprietary Network Information (what is CPNI, including my rights and AT&T's duties?). We may share this information with third parties. If we do, we won't directly identify you.

By choosing Yes below, you as the account holder agree to the terms and conditions of the Enhanced Relevant Advertising program. Your choice applies to all users of your account. Your choice doesn't affect anyone on your account's ability to use our products and services. You may revoke your consent at anytime. It may take up to 7 days to complete your request.

Please note, your choice for Relevant Advertising is separate.
Tasting Notes

Benefits

- Reputation for privacy
- Reduce user frustration
- Protect less educated users

bit.ly/2qbrnu5
Tasting Notes

Benefits
- Legal compliance
- Reputation for privacy
- Reduce user frustration
- Protect less educated users

Limitations
- Companies may want to monetize intrusive apps
- Requires privacy awareness at design
Encryption
Encrypt these:

- TLS
- Email and messaging
- Databases
- Cloud storage
- Backups
- Password management
- Endpoint devices
Don’t:

- Make your own crypto
- Use deprecated crypto (i.e., SHA1)
- Hard code keys
- Store keys on the same server as the data
- Use one key for everything
- Skip password hash and salt
- Forget to restore certificates after testing
- Use old crypto libraries
Differential Privacy
Differential privacy:

- Adds statistical noise to a data set
- Prevents identification of one individual’s record
- Provides the same results as the raw data would, with or without one record
<table>
<thead>
<tr>
<th>Privacy Model</th>
<th>Record Linkage</th>
<th>Attribute Linkage</th>
<th>Table Linkage</th>
<th>Probabilistic Attack</th>
</tr>
</thead>
<tbody>
<tr>
<td>k-Anonymity</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MultiR k-Anonymity</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ℓ-Diversity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Confidence Bounding</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>(α, k)-Anonymity</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(X, Y)-Privacy</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>(k, ε)-Anonymity</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>(ε, m)-Anonymity</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Personalized Privacy</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>t-Closeness</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>δ-Presence</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>(c, t)-Isolation</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>ε-Differential Privacy</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(d, γ)-Privacy</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Distributional Privacy</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Tasting Notes

Benefits

- Limit insider threats
- Increase data usability
- Allows for collaboration without exposing data

bit.ly/2qbrnu5
Tasting Notes

Benefits

- Legal compliance
- Limit exposure from security incidents
- Limit insider threats

Limitations

- Works best on large databases
- Must be tuned well

bit.ly/2qbrnu5
Privacy Preserving Ad Click Attribution
Privacy preserving ad click attribution:

- Allows ad attribution monetization
- Prevents user ad click tracking
- Uses the browser to mediate ad clicks
<a adDestination="shop.example" adCampaignID="55">search.example</a>
Delayed, ephemeral HTTP POST to search.example/.well-known/ad-click-attribution/20/55 with referrer set to shop.example
Available now as an experimental feature

- Accessibility Object Model
- Ad Click Attribution Debug Mode
- Ad Click Attribution
Tasting Notes

Benefits

- Allows websites to still monetize content
- Could become a W3C web standard
Tasting Notes

Benefits

- Allows websites to still monetize content
- Could become a W3C web standard

Limitations

- Needs widespread adoption to be effective
- Users may not believe any ads respect privacy

bit.ly/2qbrnu5
Federated Learning
Description

Federated learning:
- Trains a central model on decentralized data
- Never transmits device data
- Sends iterative model updates to devices which return new results
- Uses secure aggregation to decrypt only the aggregate and no user data

bit.ly/2KmuLyl
Your phone personalizes the model locally, based on your usage (A). Many users' updates are aggregated (B) to form a consensus change (C) to the shared model, after which the procedure is repeated.
Use Cases

- Android’s Gboard prediction model
- Health diagnostics
- Behavioral preference learning
- Driver behavior
Tasting Notes

Benefits

- Speeds up modeling and testing
- Minimally intrusive
- Individual data is not accessible to the central model

bit.ly/2qbrnu5
Tasting Notes

Benefits

- Speeds up modeling and testing
- Minimally intrusive
- Individual data is not accessible to the central model

Limitations

- Errors could cause private data leakage
- Requires a large user base

bit.ly/2qbrnu5
Homomorphic Encryption
Description

Homomorphic encryption:
- Allows computation on ciphertext
- Enables collaboration without disclosing confidential data
- Only the calculation results can be decrypted
Use Cases

- Computations on data shared across organizations
- Research using highly sensitive records
- Processing by employees with a lower clearance
- Google’s open source Private Join and Compute
Tasting Notes

Benefits

- Reduces insider threat
- Increases collaboration
- Increases data usability

bit.ly/2qbrnu5
Tasting Notes

Benefits
- Reduces insider threat
- Increases collaboration
- Increases data usability

Limitations
- Resource-intensive
- Limited functions
- No fully homomorphic encryption available yet
Becoming a Privacy Champion
Amber Welch

MA, CISSP, CISA, CIPP/E, CIPM, FIP, CCSK, and ISO 27001 Lead Auditor

linkedin.com/in/amberwelch1

github.com/msamberwelch

@MsAmberWelch