

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](https://www.linkedin.com/in/amberwelch1)

github.com/msamberwelch


@MsAmberWelch



Menu



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

A tall, slender glass filled with golden beer, showing numerous small bubbles rising to the surface. The glass is set against a blurred background of warm, bokeh lights. The text "First, an apology." is overlaid on the left side of the image.

First, an apology.

Legal teams have often kept tech out of privacy.



Developers don't know privacy concepts. Privacy teams haven't taught them.



Figure 1: Participants' Formal Knowledge on Privacy Concepts

Privacy Impact Assessment



Description

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



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

Data Minimization and Retention



Description

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

Default Settings



Description

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.

Music, Video, & Books

Lists

Reminders & Alarms

Contacts

Alexa Devices

Routines

Smart Home

Skills

Settings 1

Things to Try

Help & Feedback

bit.ly/2Hic0qm

Settings

Notifications >

Routines >

Voice Purchasing >

Recognized Voices >

Household Profile >

History 2

General

About Alexa app >

SIGN OUT

Settings

History

History shows your voice interactions with Alexa. Tap a line to see details, hear recordings, provide feedback, or delete recordings.

Learn more.

Filter by Date
None

buy a ten pound bucket of sea salt 3

Today at 6:06 PM on Todd's Echo

alexa >

Today at 6:06 PM on Todd's Echo

stop >

Today at 6:06 PM on Todd's Echo

alexa >

Today at 6:06 PM on Todd's Echo

play britney spears >

Today at 6:06 PM on Todd's Echo

alexa >

Today at 6:06 PM on Todd's Echo

stop

Settings

History

Today at 6:06 PM

▶ buy a ten pound bucket of sea salt

The top search result for 10 lbs bucket of sea salt is Redmond Real Salt, Nature's First Sea Salt, Fine Salt, 10 Pound Bucket

It's \$44.21 before tax

Would you like to add it to your cart?

▶ alexa stop

4

DELETE VOICE RECORDINGS

Sometimes Alexa does not understand you perfectly, and the

Manage your privacy and automated messaging choices

[Read our privacy policy](#)

How AT&T Communicates with Customers

Automated Messages

How AT&T Uses Customer Data

External Marketing & Analytics Reports

DNS Error Assist

Relevant Advertising

Enhanced Relevant Advertising

➤ [Manage your preferences](#)

Third Party Services

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.

Service

Wireless number:

Wireless number:

Allow use of information?

 No No

Tasting Notes

Benefits

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

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



Description

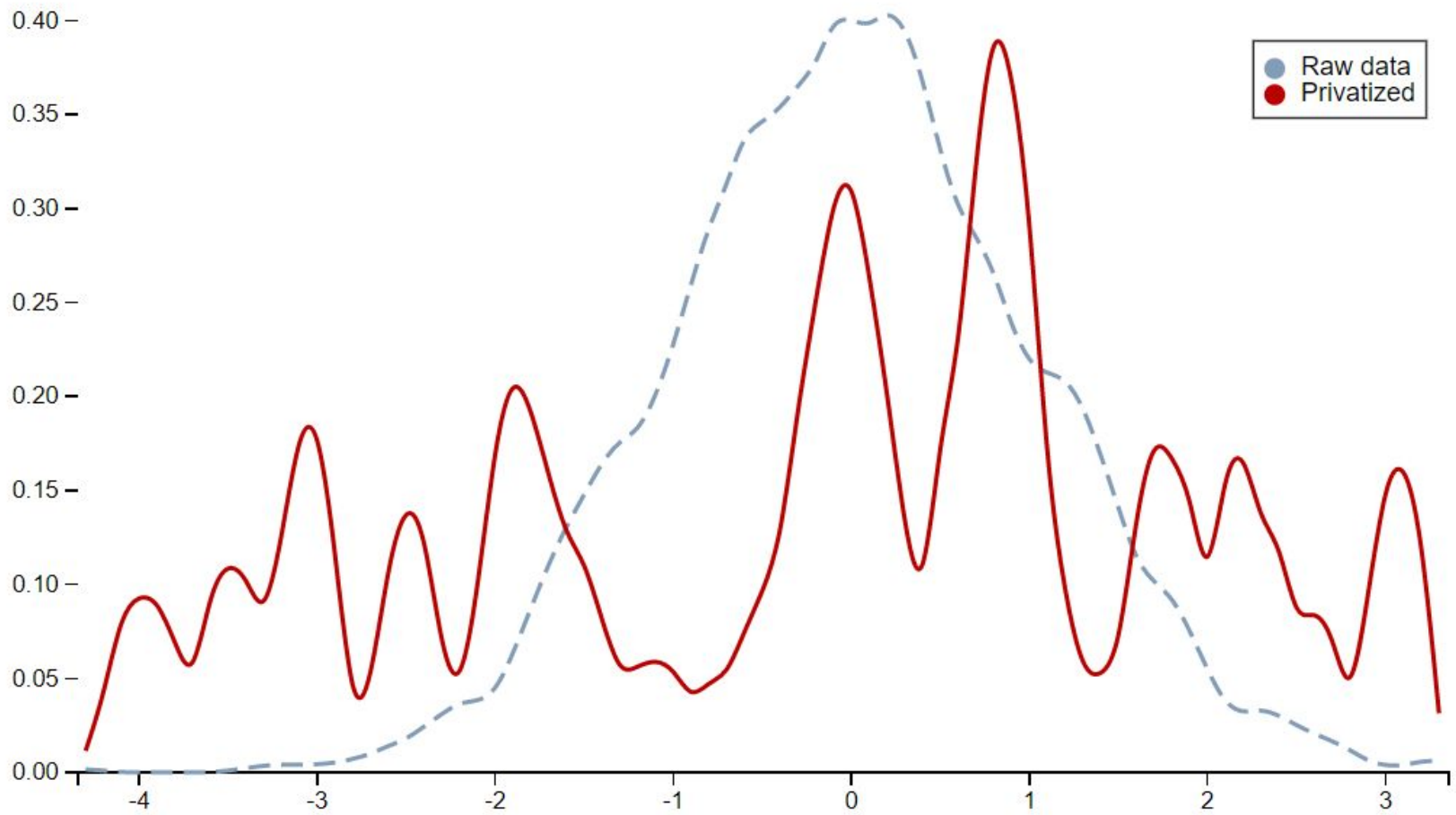
Differential privacy:

- 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



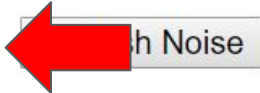
Table I. Privacy Models

Privacy Model	Attack Model			
	Record Linkage	Attribute Linkage	Table Linkage	Probabilistic Attack
k -Anonymity	✓			
MultiR k -Anonymity	✓			
ℓ -Diversity	✓	✓		
Confidence Bounding		✓		
(α, k) -Anonymity	✓	✓		
(X, Y) -Privacy	✓	✓		
(k, e) -Anonymity		✓		
(ϵ, m) -Anonymity		✓		
Personalized Privacy		✓		
t -Closeness		✓		✓
δ -Presence			✓	
(c, t) -Isolation	✓			✓
ϵ -Differential Privacy			✓	✓
(d, γ) -Privacy			✓	✓
Distributional Privacy			✓	✓

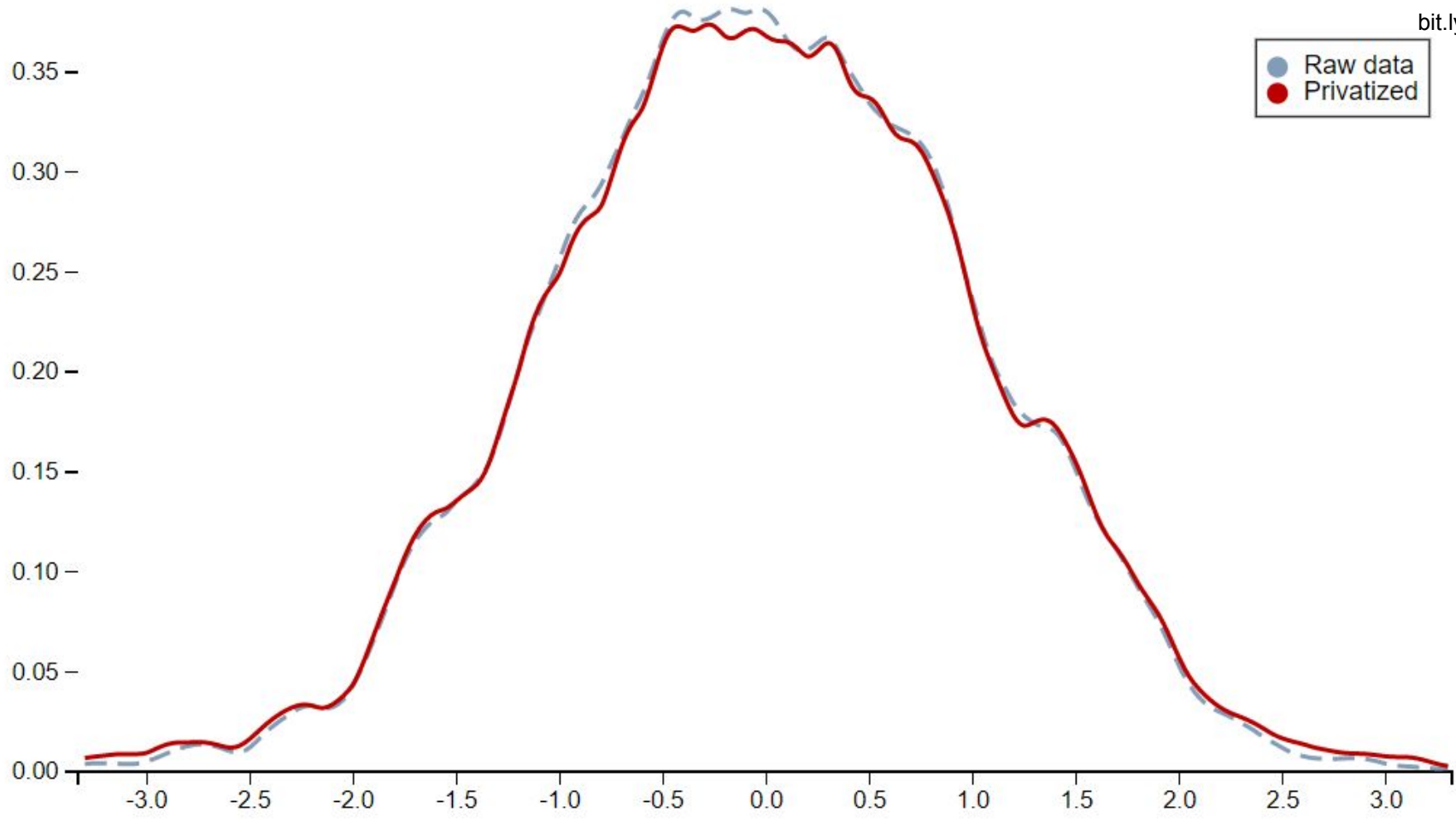


Random Variables --> Normal(0,1) ▼

ϵ 0.02



h Noise Refresh Distribution



Random Variables --> Normal(0,1) ϵ 2

Tasting Notes

Benefits

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

Tasting Notes

Benefits

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

Limitations

- Works best on large databases
- Must be tuned well



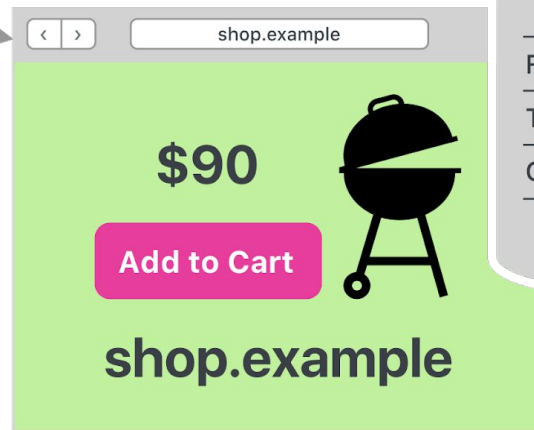
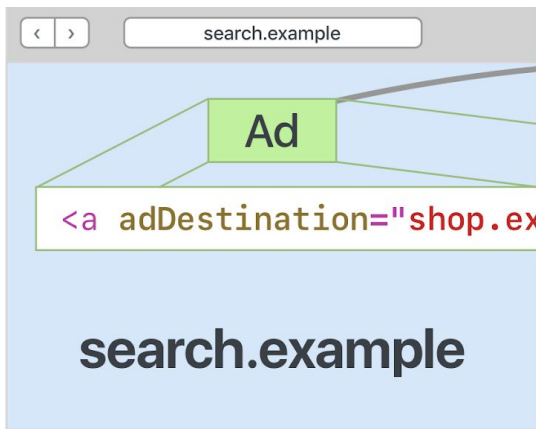
Privacy Preserving Ad Click Attribution

Description

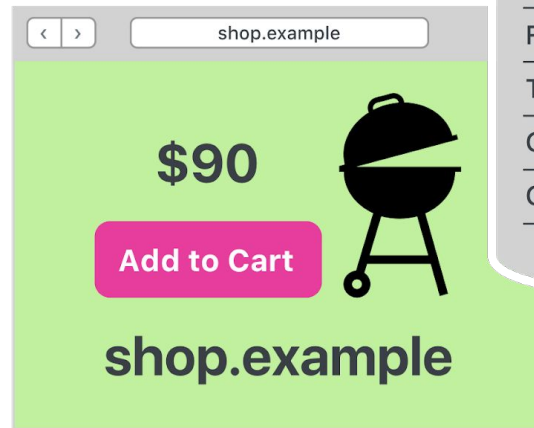
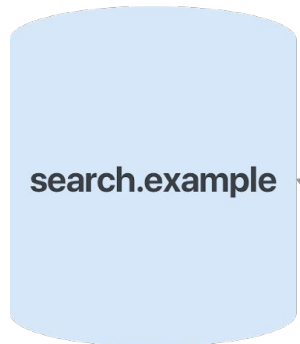
Privacy preserving ad click attribution:

- Allows ad attribution monetization
- Prevents user ad click tracking
- Uses the browser to mediate ad clicks





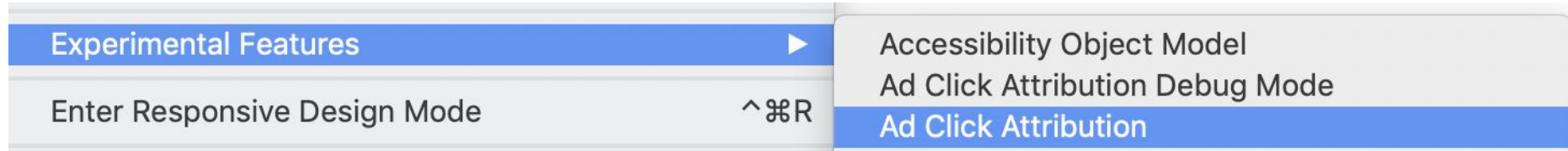
From:	Search
To:	Shop
Campaign:	55



From:	Search
To:	Shop
Campaign:	55
Convert:	20

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



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

Federated Learning

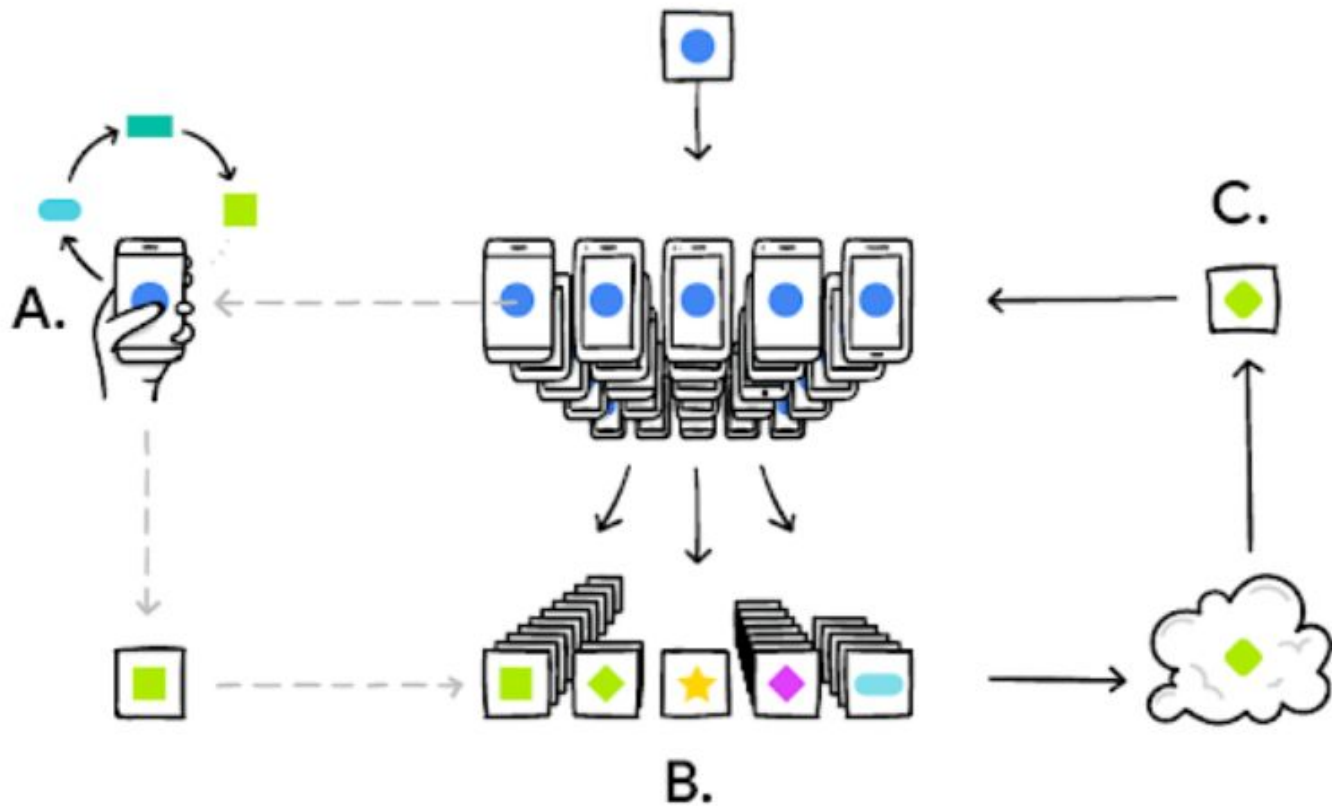


Description

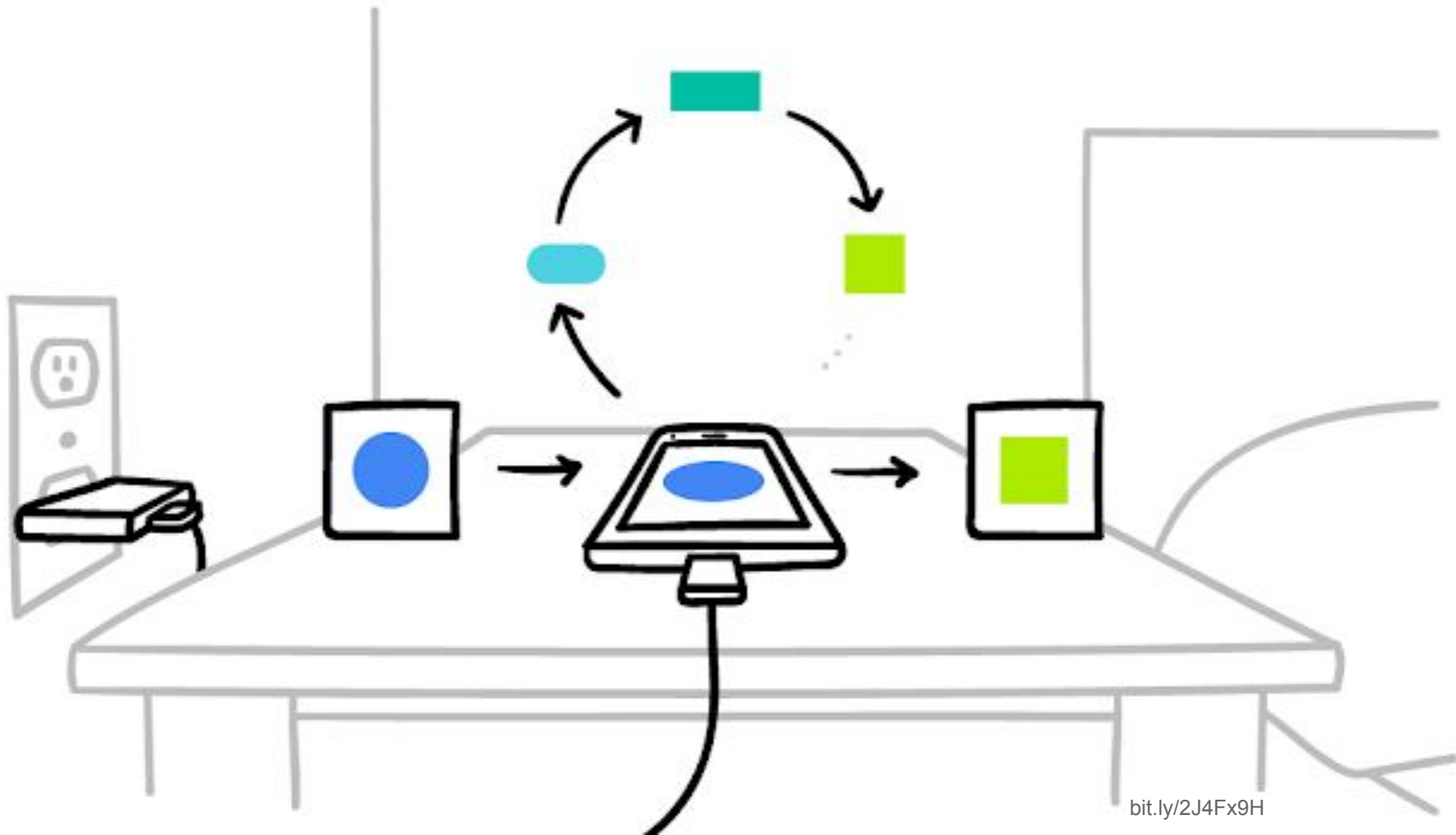
Federated learning:

- 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





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

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

Homomorphic Encryption



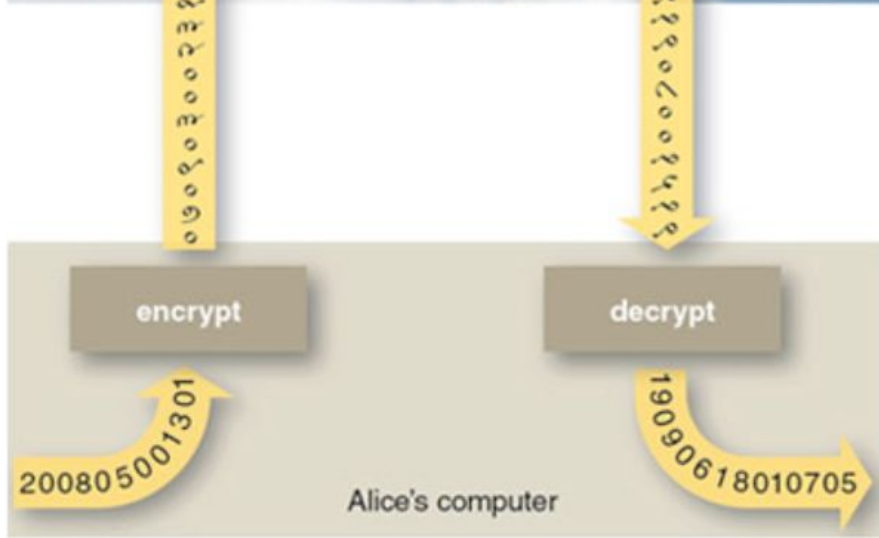
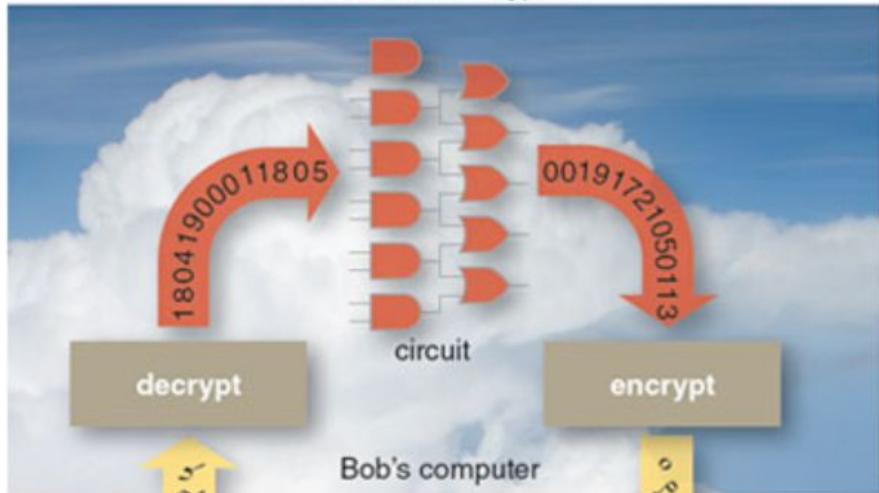
Description

Homomorphic encryption:

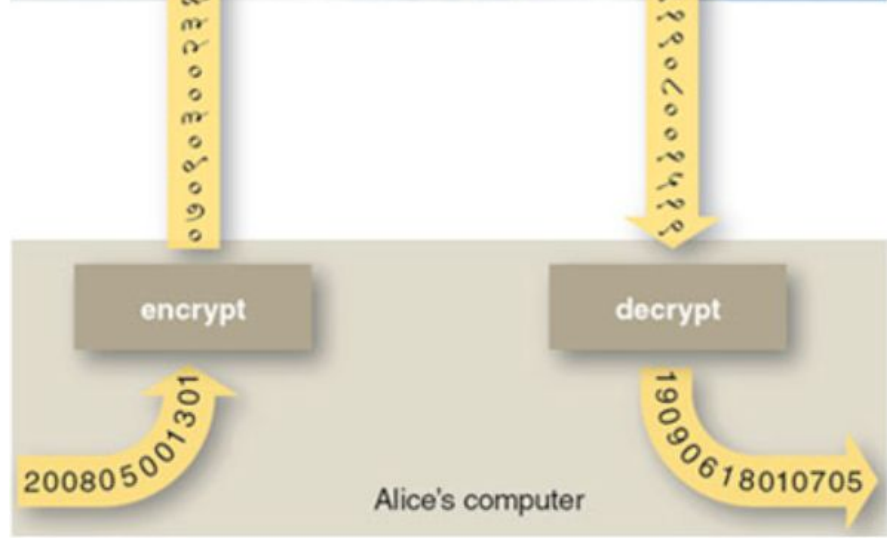
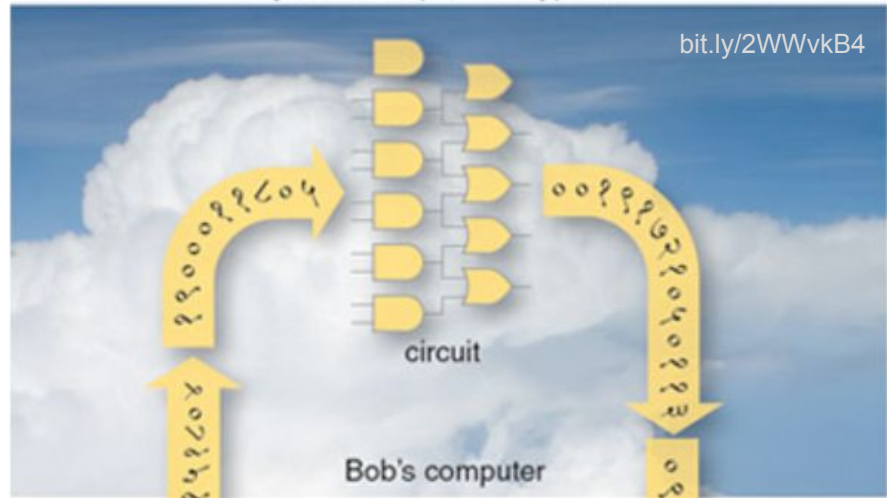
- Allows computation on ciphertext
- Enables collaboration without disclosing confidential data
- Only the calculation results can be decrypted



conventional encryption



fully homomorphic encryption



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


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](https://www.linkedin.com/in/amberwelch1)

github.com/msamberwelch

@MsAmberWelch

