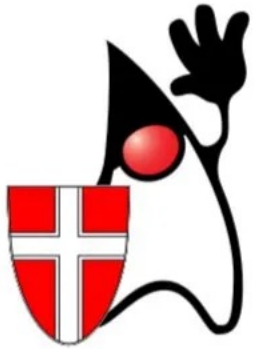


```
public static final Meetup JAVA_VIENNA =
new Meetup("Java", "Vienna", "Beer", "Fun");
  //TODO visit https://www.java.wien/
```

June 2024 Meetup #1

dynatrace

POSEDIO

openFORCE

dynatrace

Privacy Leaks Beyond The Class-Level

Protecting Personal Data

PRESENTER

Cat Easdon
Senior Privacy Engineer

whoami

- Senior Privacy Engineer and TechEvangelist at Dynatrace
- ⛰️-obsessed Brit based in Innsbruck
- Outside of work: research and tech policy, ~~probably~~ an excessive amount of trail running, hiking, skiing…
- Previously: hacking CPUs at TU Graz

ECCRI EUROPEAN CYBER CONFLICT RESEARCH INITIATIVE

Internet Society

Privacy leaks at the class-level

Example of privacy leaking “getter”

```
public class Leaking {
   private Counter leakedCounter;
...
   /* Getter generated by Eclipse: */
   public Counter getLeakedCounter() {
      return leakedCounter;
   }
```

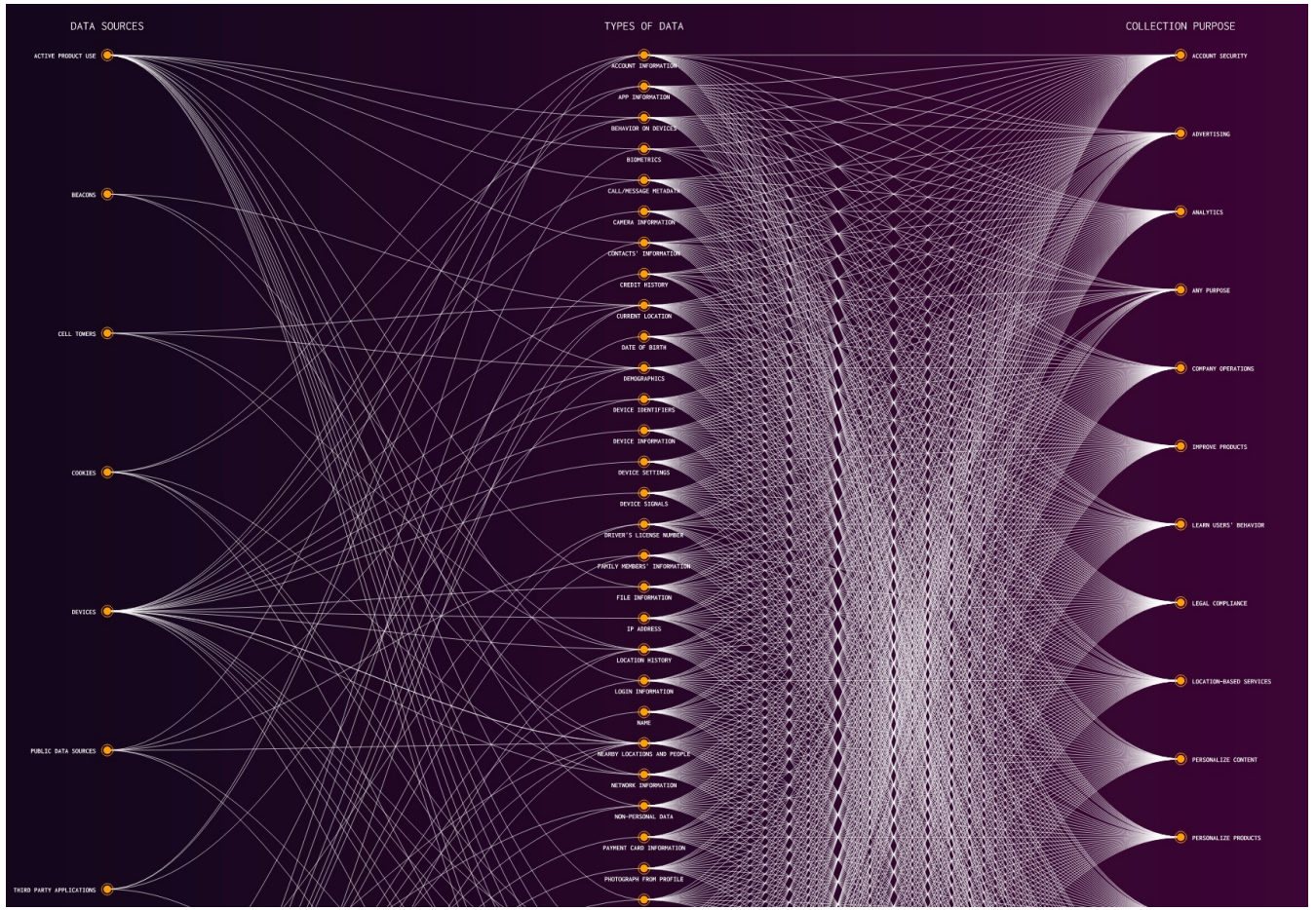
Notice how the “getter” generated by Eclipse simply returns a copy of the **reference** to the private attribute; giving the client access to modify the internal state of an object in the class directly.

Full source code: `http://repositories.jacob-sparre.dk/privacy-leaks-examples/`

Jacob Sparre Andersen | Privacy Leaks in Java Classes

Privacy Leaks in Java Classes, Ada-Europe 2014, by Jacob Sparre Andersen: https://jacob-sparre.dk

Privacy leaks *beyond* the class level



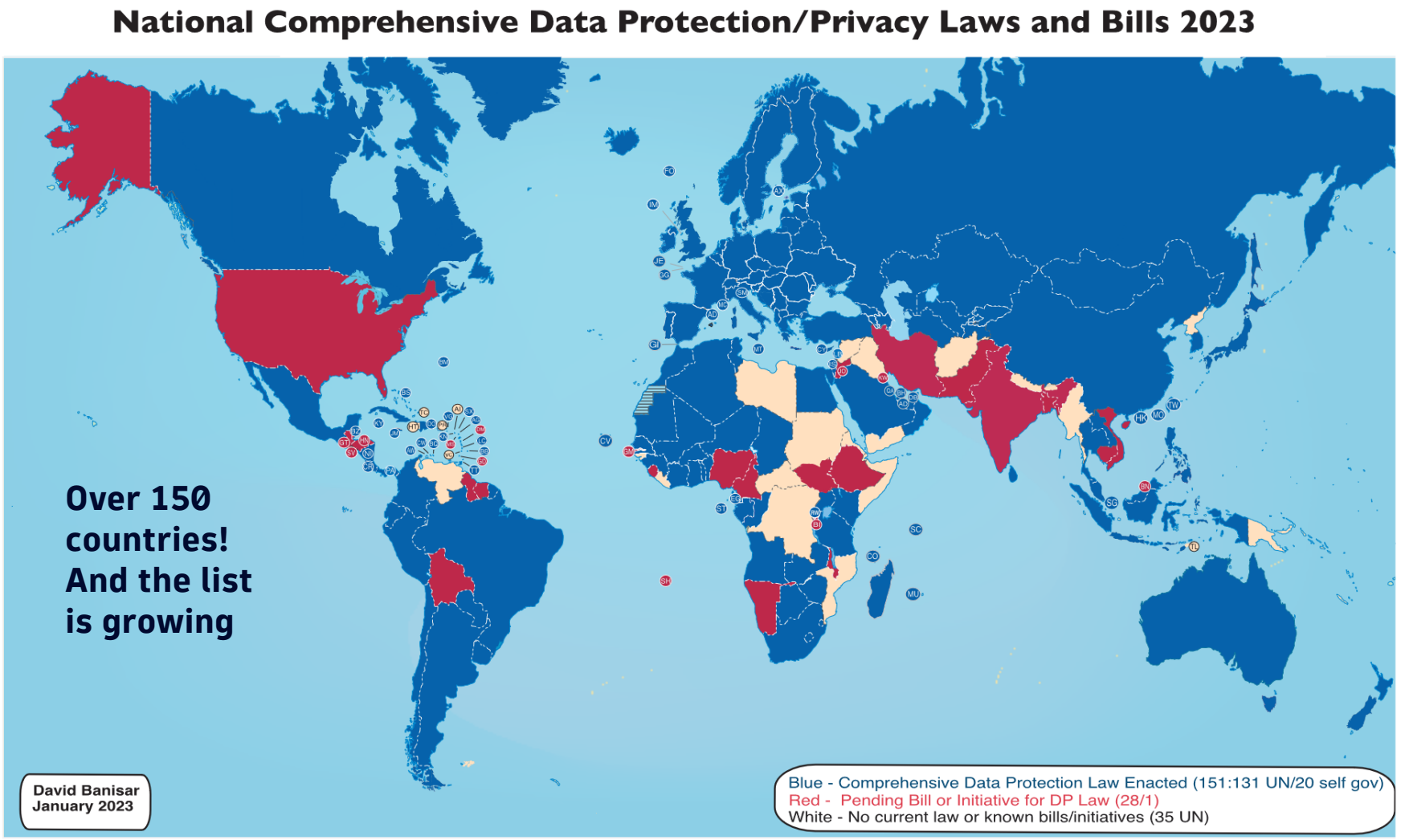
Mapping Data Flows, SIPA – Columbia University: https://mappingdataflows.com

Outline

- Motivation
  - What does privacy mean?
  - Why does it matter?
- Privacy at the code level
- Privacy at the infrastructure level
- Privacy by design and how to build it into the SDLC

What does privacy mean to you?

What does privacy mean?



Source: David Banisar

What does privacy mean?

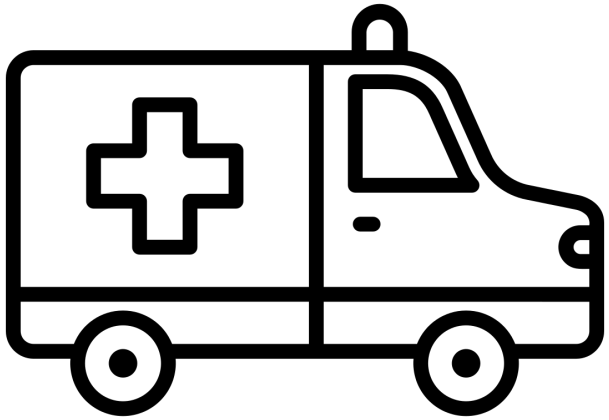
- Right to know, access, export
- Right to update (rectify)
- Right to delete (“be forgotten”)
- Right to data portability
- Plus more…
  - Right to restriction of processing
  - Right not to be subject to automated decision-making, incl. profiling (GDPR)
  - Right to opt out of sensitive data processing (CPRA)

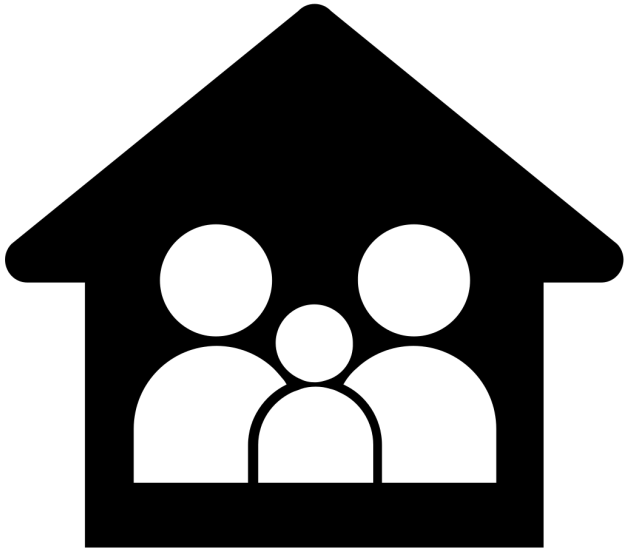
What does privacy mean?

Image: Gatha Varma

What does privacy mean?

- Privacy isn't about hiding everything!
- Society runs on information flows, but these flows should be *context-appropriate*
- Data type, data subject, sender, recipient, legal basis, and purpose all matter





Why does privacy matter?

Privacy threats

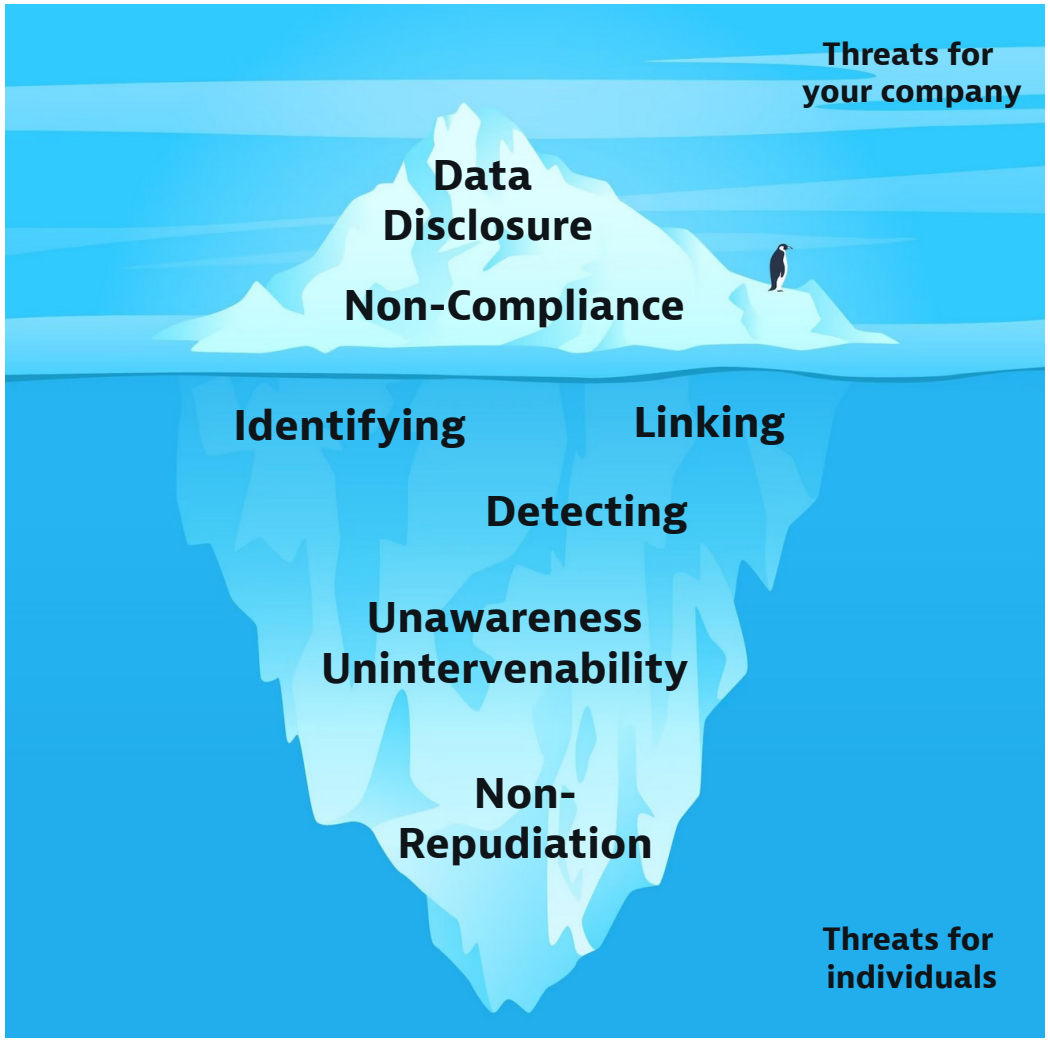


Image: Freepik, threats: LINDDUN

Privacy harms

Privacy Threat
Linkability
Identifiability
Non-Repudiation
Detectability
Data Disclosure
Unawareness, Unintervenability
Non-Compliance

Privacy Harms
Physical Harms
Economic Harms
Reputational Harms
Autonomy Harms
Psychological Harms
Discrimination Harms
Relationship Harms
Harm to the Business

Autonomy Harms
Coercion
Manipulation
Failure to Inform
Thwarted Expectations
Lack of Control
Chilling Effects

Privacy Harms, Citron & Solove, Boston University Law Review (2022)

Privacy harms: manipulation

Neuroticism - Trapped | Neuroticism - Stress Reactors | Neuroticism - Self Lovers
Neuroticism - Easily Deflated | Neuroticism - Internal Escapists
General Attitudes - I generally get a raw deal out of life | Dealing with Stress - Hot and Cold
Dealing with Stress - Emotional | Dealing with Stress - Bottled Up

Clickagy > Health > **Addict**ions > Alcohol

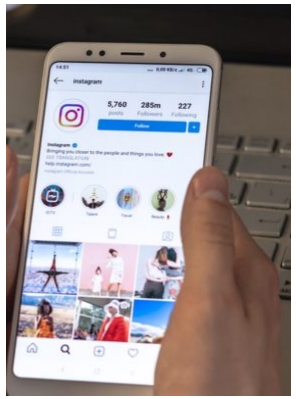
Skydeo > ConditionGraph > Disease Propensity by Type > **Depression** Diagnosis

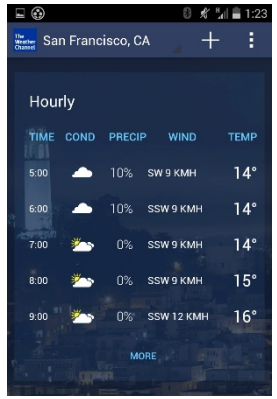
VisualDNA Lifestyle - Lifestyle - Health - Trying to cut down on **Alcohol**
Data provider: Nielsen Marketing Cloud

TransUnion - Demographics - Marital Status - Likely Recently **Divorced**

AlikeAudience: United States > Interest > Entertainment > Party And **Night Club** Enthusiasts

Eyeota - US Acxiom - CPG - **Alcoholic** Drinks - Vodka Brand - Grey Goose for age 21+ - Likely





Source: The Markup

Privacy at the code level

Privacy leaks at the class-level: an alternative perspective

```
final var response = startQueryForUser(user, tenantInfo).block();
final LocalDateTime retrievedAt = LocalDateTime.now(ZoneOffset.UTC);
if (response == null) {
    log.error("Request to tenant {} for user {} failed!",
            tenantInfo.url(), user.email());
    return new ResultAtTime<>(retrievedAt, result: null);
} else if (response.requestToken() != null) {
    // ...
```


```
{
  "object": "user",
  "id": "uuid-redacted",
  "email": "a.sample.email@email.com",
  "name": "Cat Easdon",
  "picture": "https://lh3.googleusercontent.com/a/url-redacted",
  "created": 1702548524,
  "phone_number": "+431234567890",
  "groups": {"object": "list"...}
}
```

Privacy in API design

Beyond the DTOs, also ensure that personal data is not included in endpoint paths:

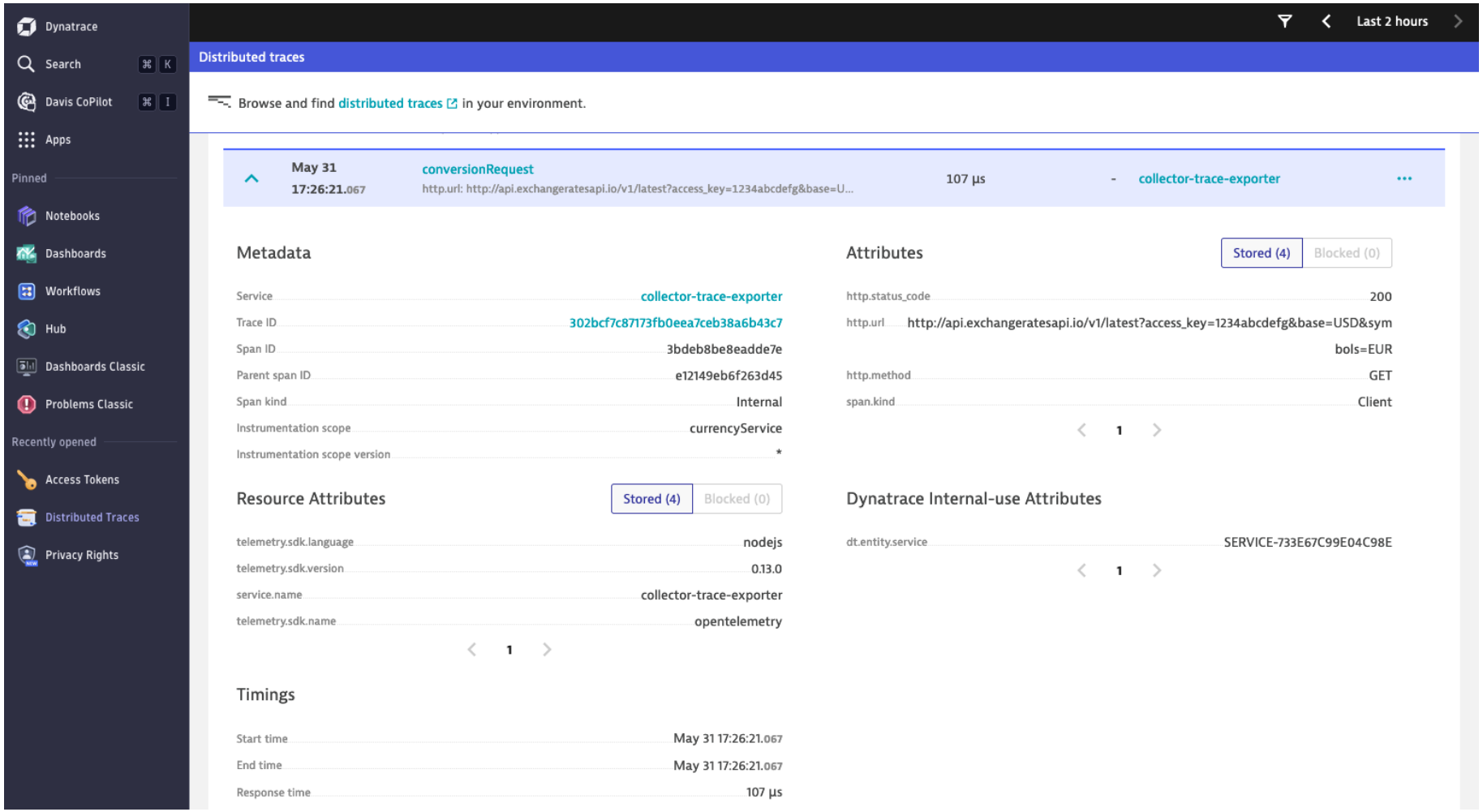
```
privacy-rocks.com/settings/maria@email.com
```

```
privacy-rocks.com/confirmation?user=andreas@email.com&token=ab456d
```

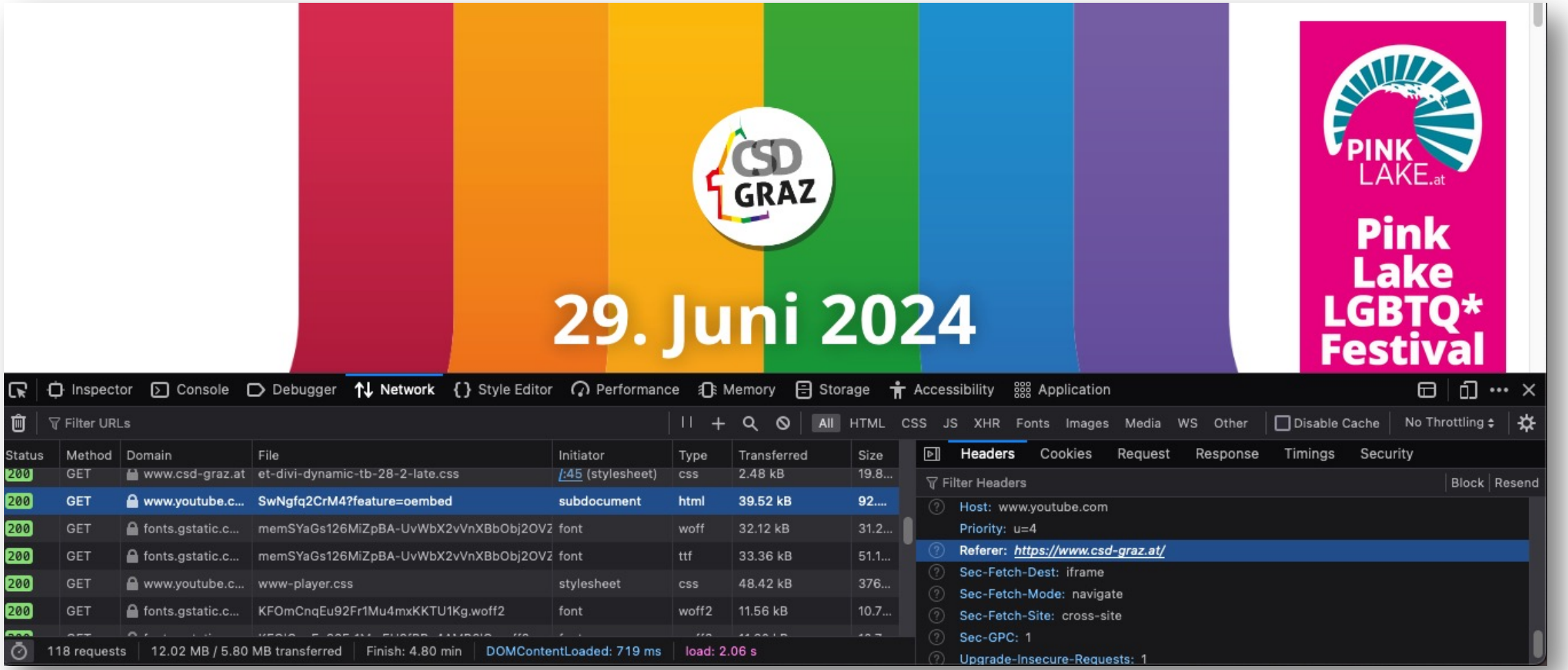
Otherwise, this data may be stored in the browser history and server logs, and gets sent to:

- Proxies and CDNs
- Google (or other ad vendor) and analytics providers
- Observability tooling
- …

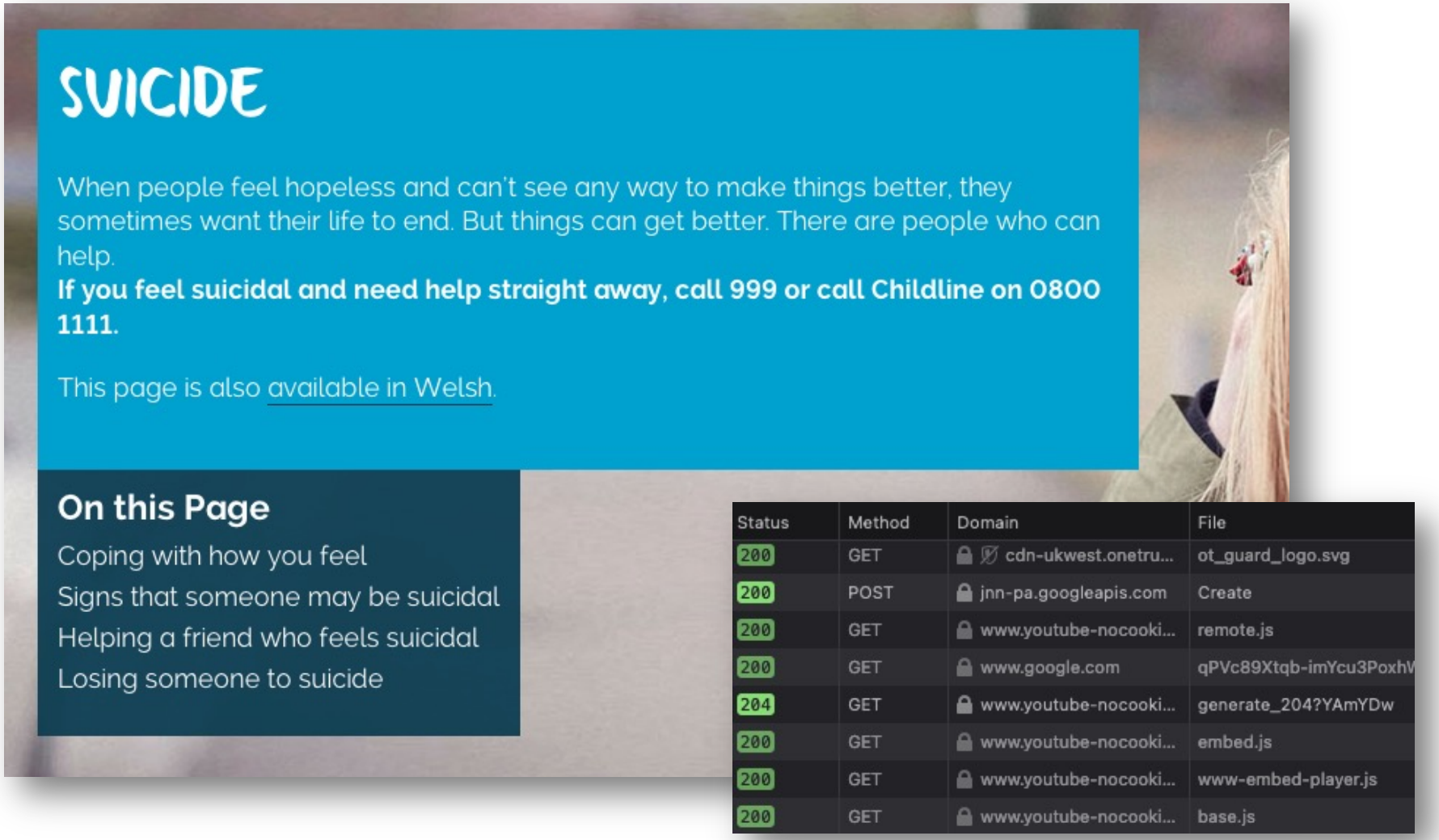
Privacy in observability tooling



Privacy in sensitive contexts



Privacy in sensitive contexts

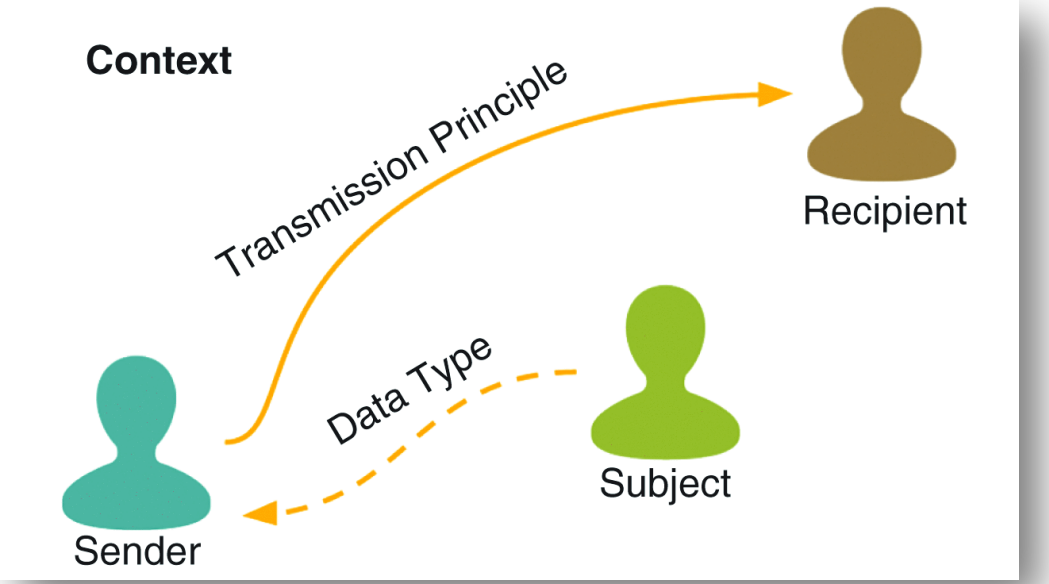


Privacy in sensitive contexts

In some contexts, every third-party request your web page or app makes may be an inappropriate flow of personal data!

Starting points:

- `<meta name="referrer" content="no-referrer" />`
- Don't include trackers (Google Analytics, Facebook Pixel / embedded social media or sharing links, videos…)
- Beware hidden tracking (e.g. Stripe)
- These are even easier to overlook when developing a mobile app!



Source: IEEE Security & Privacy

Privacy at the infrastructure level

Privacy operations: rights at scale

How to keep track of personal data?

- Internal policies
- Code-level annotations
- Data and software catalogues
- Personal data detection and data-flow mapping

privado

fides lang

Semgrep

```
public @interface MicrophoneAnnotation {
    String ID();
    Visibility[] visibility();
    MicrophoneDataType[] dataType();
    MicrophonePurpose[] purpose();
    String [] purposeDescription();
}
```

```
system:
  - fides_key: demo_analytics_system
    name: Demo Analytics System
    description: A system used for analyzing customer behaviour.
    system_type: Service
    administrating_department: Engineering
    egress:
      - fides_key: another_demo_system
        type: system
        data_categories:
          - user.contact
    ingress:
      - fides_key: yet_another_demo_system
        type: system
        data_categories:
          - user.device.cookie_id
    privacy_declarations:
      - name: Analyze customer behaviour for improvements.
        data_categories:
          - user.contact
          - user.device.cookie_id
        data_use: improve.system
        data_subjects:
          - customer
        egress:
          - another_demo_system
        ingress:
          - yet_another_demo_system

```

All open source! Check them out here: Privado, Fides Lang, Semgrep, Coconut IDE

Privacy operations: rights at scale

- Once you've found the data, you then need to deal with incomplete exports or deletions due to:
  - Eventual consistency, caching, synchronization
  - Outages, timeouts
  - Immutable or "immutable in practice" data stores
  - Data warehousing and long-running data pipelines
  - Backups
- Two complementary approaches
  - Pseudonymize personal data -> centralize storage
  - Event-based pipelines
    - ...perhaps using Change Data Capture! ☺

Source: Imgflip

Privacy operations: rights at scale

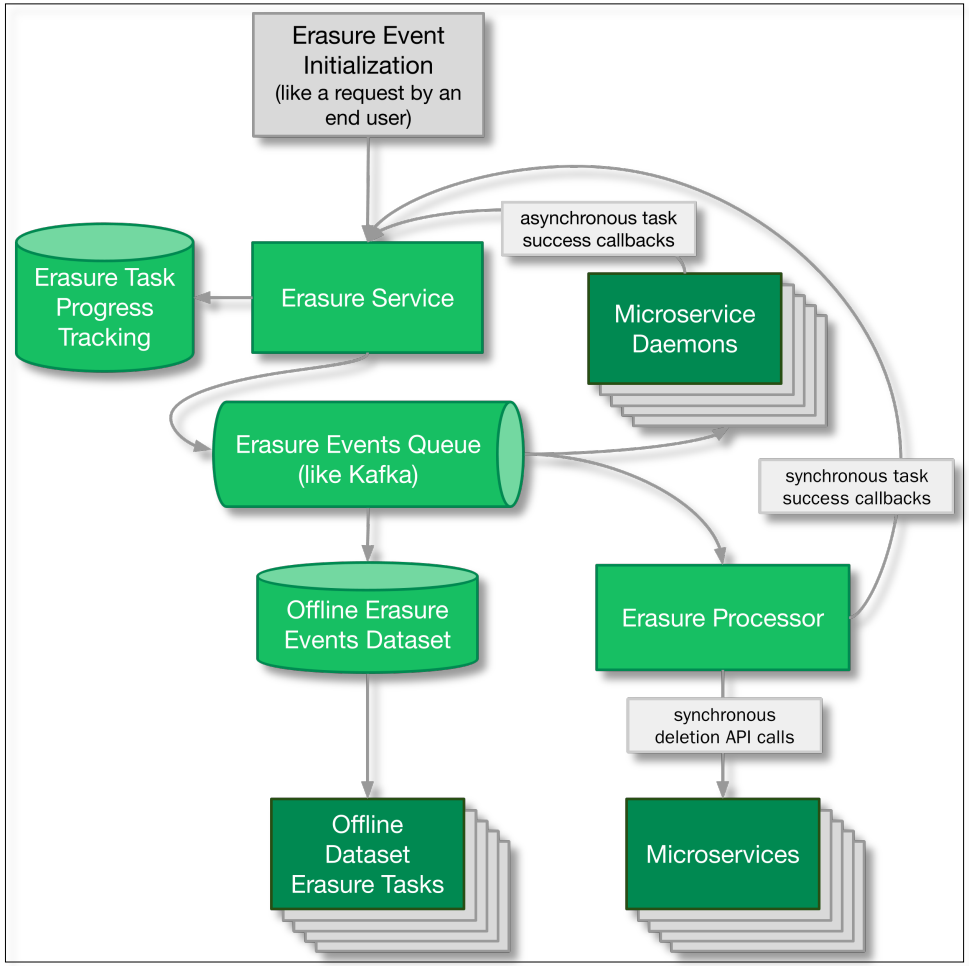
Deletion methods:

- API calls to each service (synchronous or async)
- Event-driven
- Offline / manual

Deletion needs to be *replayable* to handle:

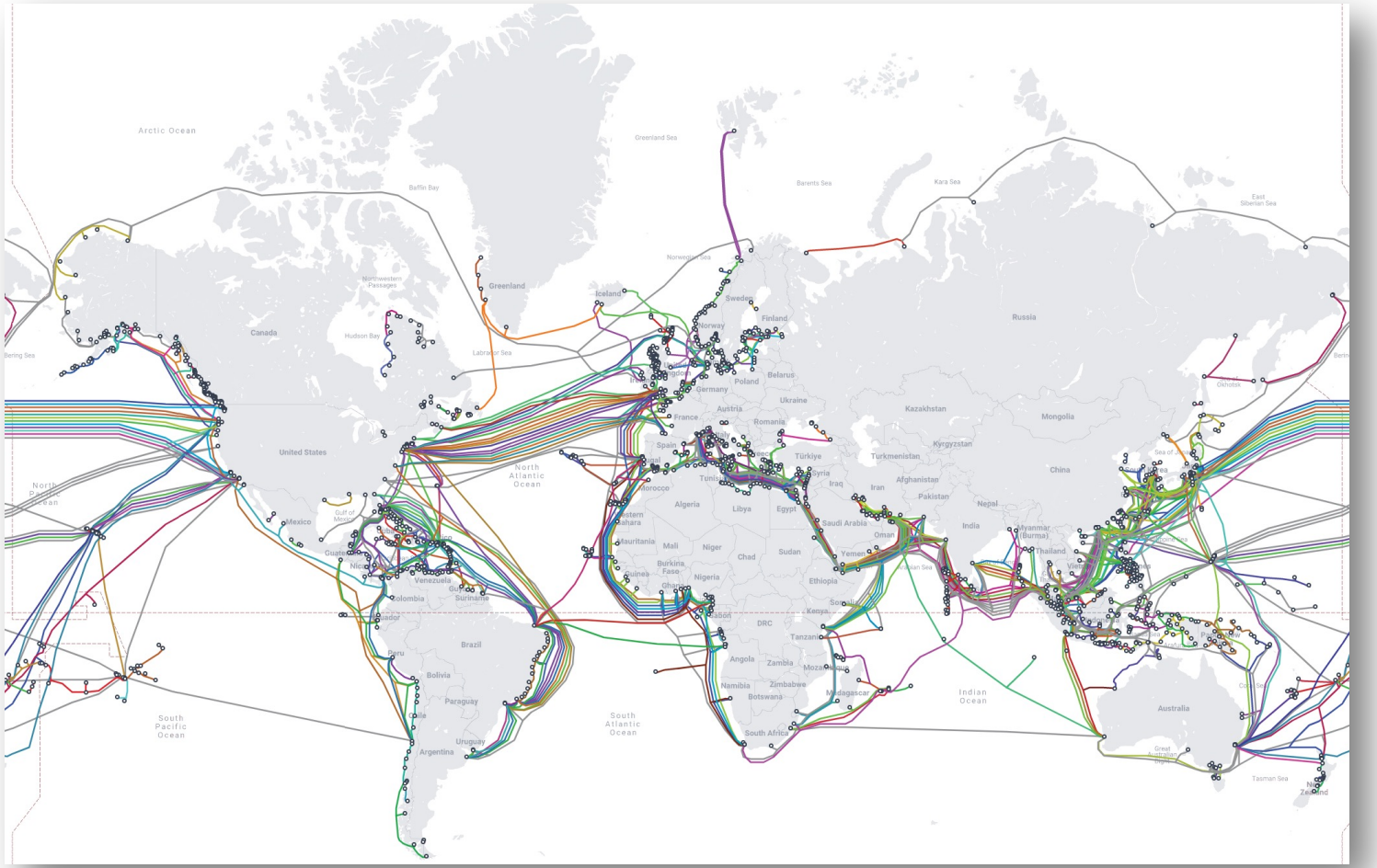
- Outages
- Disaster recovery

Best if the team that owns each system also owns their role in export/deletion



Source: [Twitter Engineering]

Cross-border data transfers in the cloud

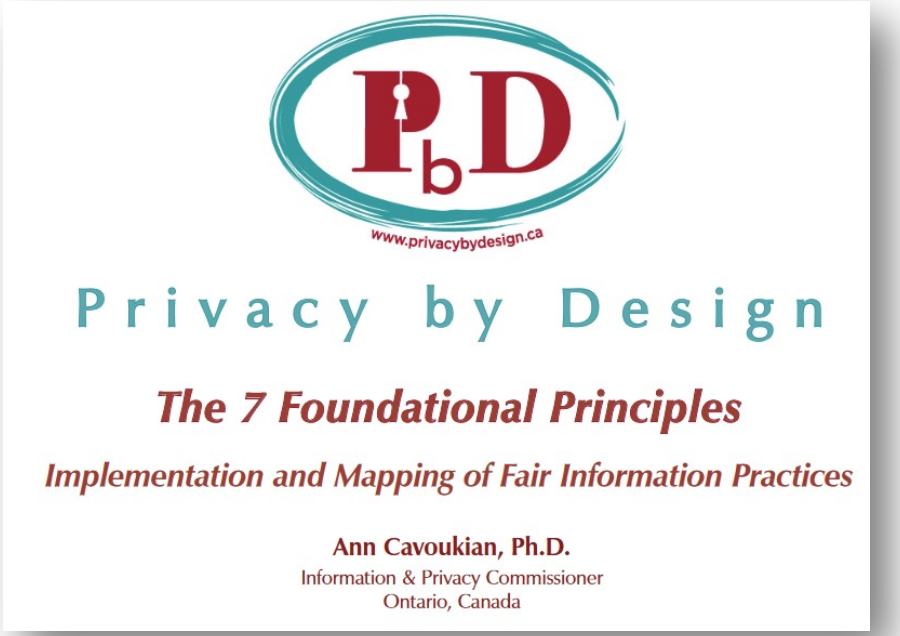


Source: [Submarine Cable Map by TeleGeography]

Privacy by design in the SDLC

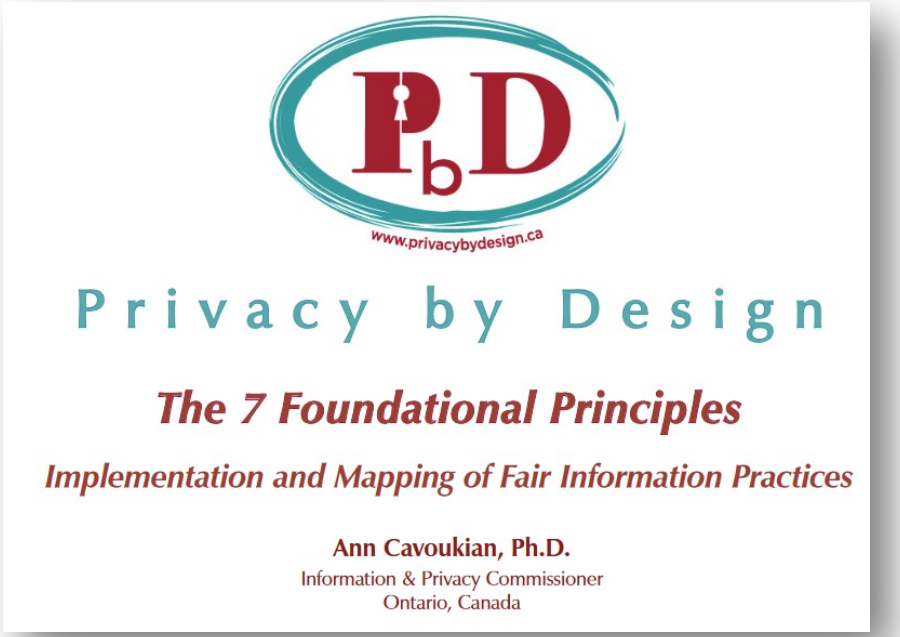
Privacy by design

1. Proactive not reactive; preventative not remedial
2. Privacy as the default setting
3. Privacy embedded into design
4. Full functionality – positive-sum, not zero-sum

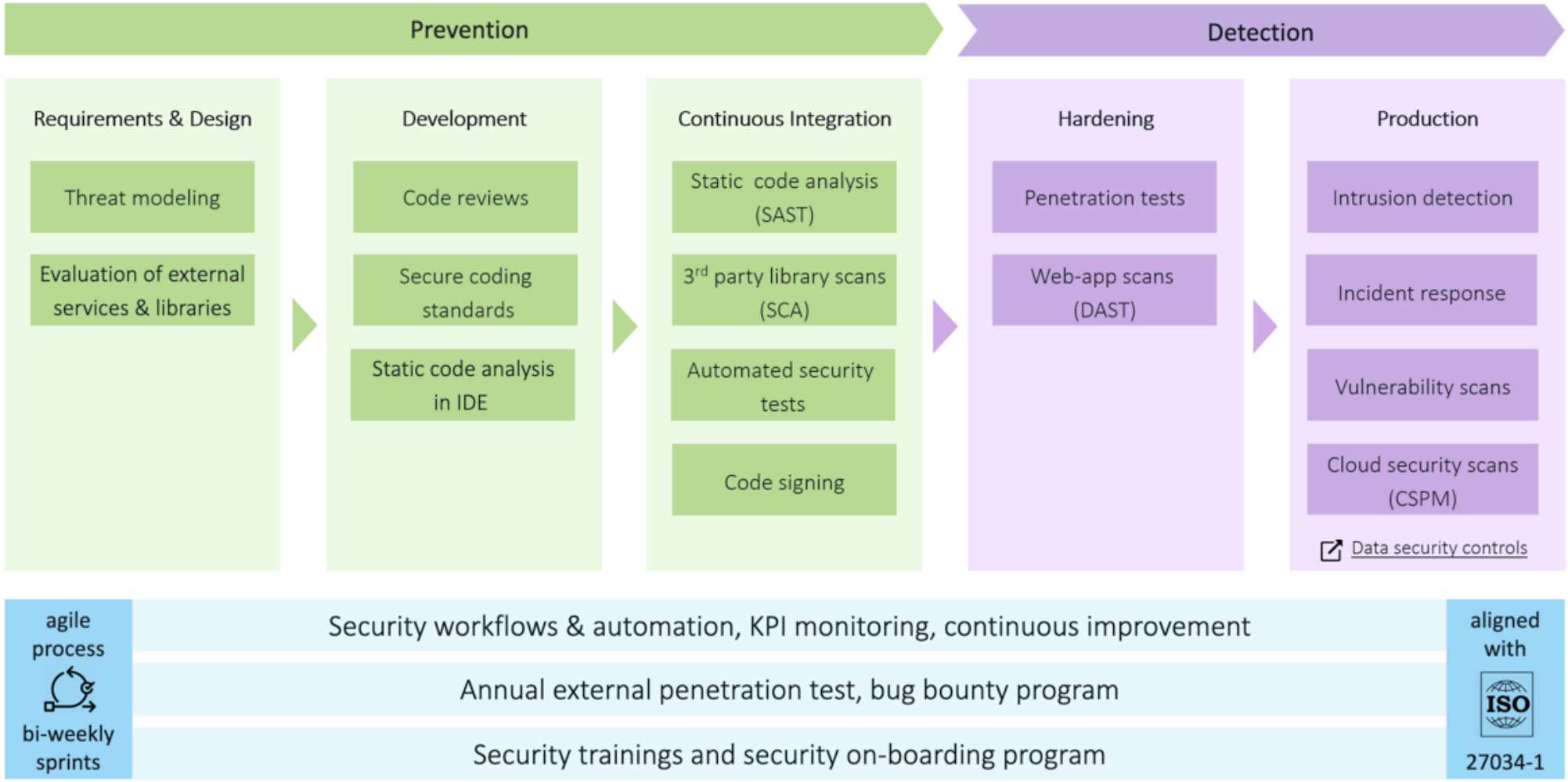


Privacy by design

5. End-to-end security – full lifecycle protection
6. Visibility and transparency – keep it open
7. Respect for user privacy – keep it user-centric



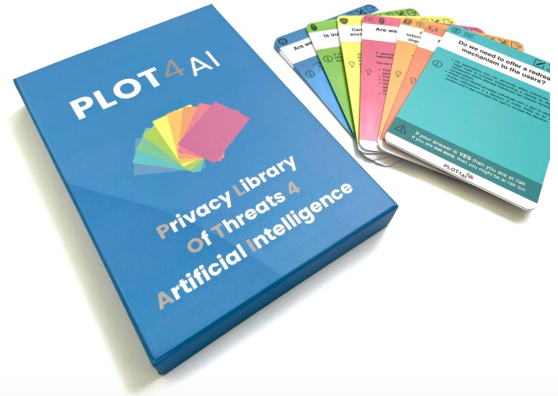
Security controls in the SDLC

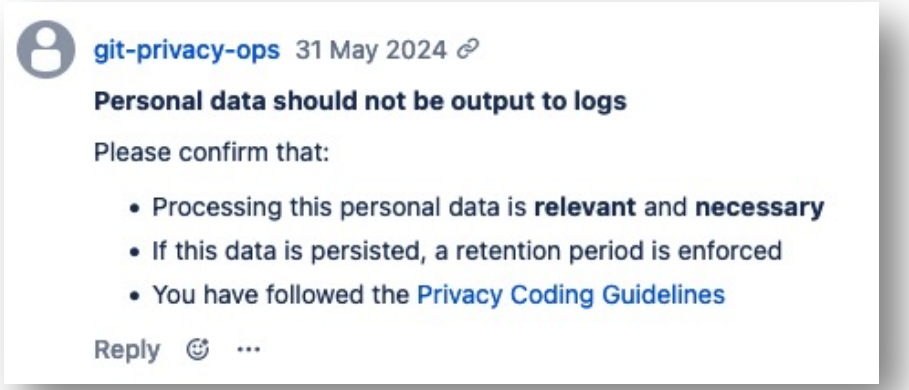


Privacy by design in the SDLC with privacy controls

- Threat modeling from initial design phase onwards
- Training and internal enablement
- Static code analysis
- Data and software cataloguing
- Privacy vulnerability handling
- Runtime scanning
  - PII detection
  - Data Loss Prevention (DLP)
- Privacy penetration testing and red-teaming

LINDDUN





Thanks for listening!

Introduction to Privacy Engineering

Open University OpenLearn Create

Cat Easdon

Privacy Engineering at Dynatrace | Bridging the gap between tech and policy