2021 International Workshop on Privacy Engineering (IWPE‘21)

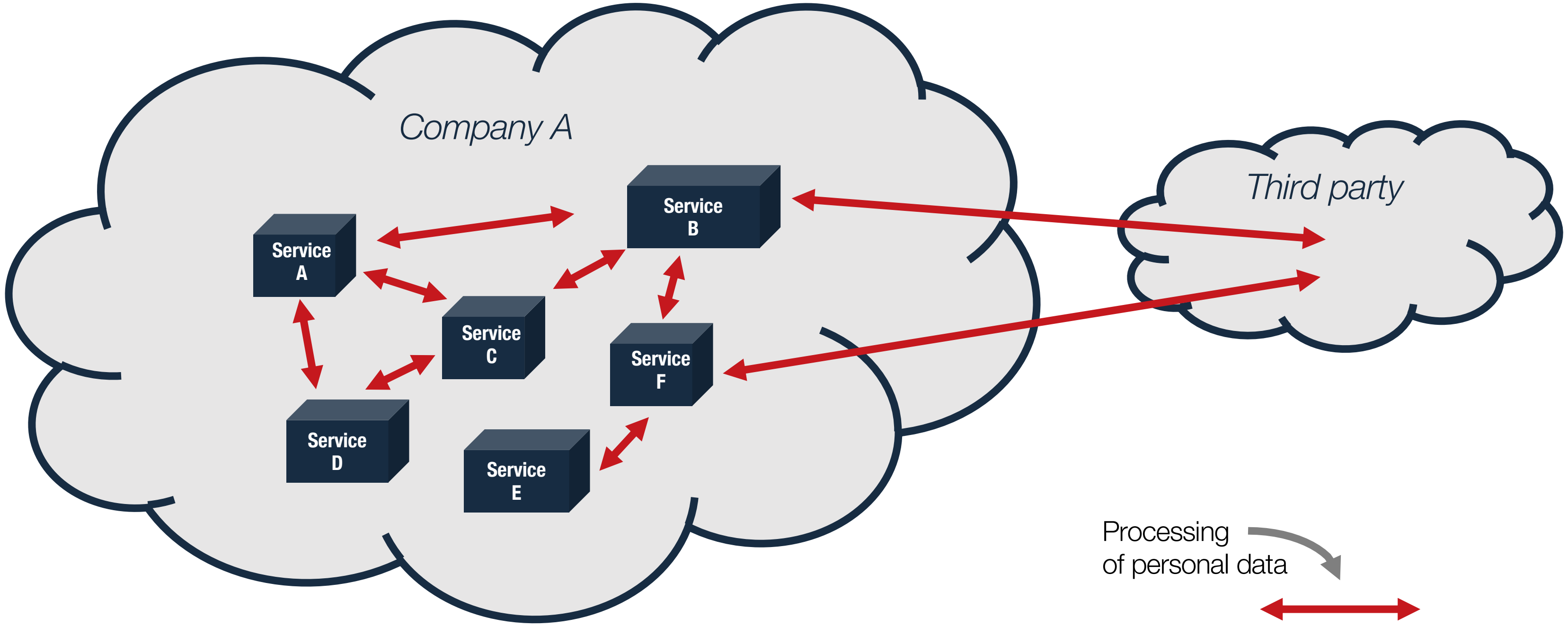
Co-located with 6th IEEE European Symposium on Security and Privacy

# TIRA: An OpenAPI Extension and Toolbox for GDPR Transparency in RESTful Architectures

**Elias Grünewald**, Paul Wille, Frank Pallas, Maria C. Borges, Max-R. Ulbricht

Information Systems Engineering
TU Berlin

# In a nutshell



What personal data is collected for which purposes?
How long is it stored?
Which third parties is it transferred to?
…

# Agenda

1. Introduction

2. Background

    Privacy and Transparency

    APIs, DevOps & RESTful Architectures

3. Requirements & General Approach

4. Transparency-focused **OpenAPI Extension** (incl. vocabulary)

5. **Toolbox** for aggregating transparency information (incl. CI/CD integration)

6. Discussion & Conclusion

# Privacy and Transparency

Art. 5(1) GDPR

> Personal data shall be
> (a) processed lawfully, fairly and in a **transparent** manner
> in relation to the data subject (‘lawfulness, fairness and **transparency**’);

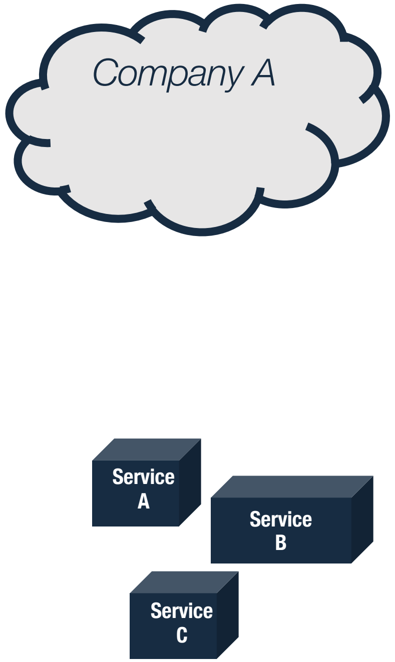
Art. 12(1) GDPR

> The controller shall take **appropriate measures** to provide any information [according to Art. 13, 14, 15-22, 34] relating to processing to the data subject in a concise, **transparent**, intelligible and easily accessible form

Art. 25 GDPR

**Data protection by design and by default**

# Privacy and Transparency (contd.)



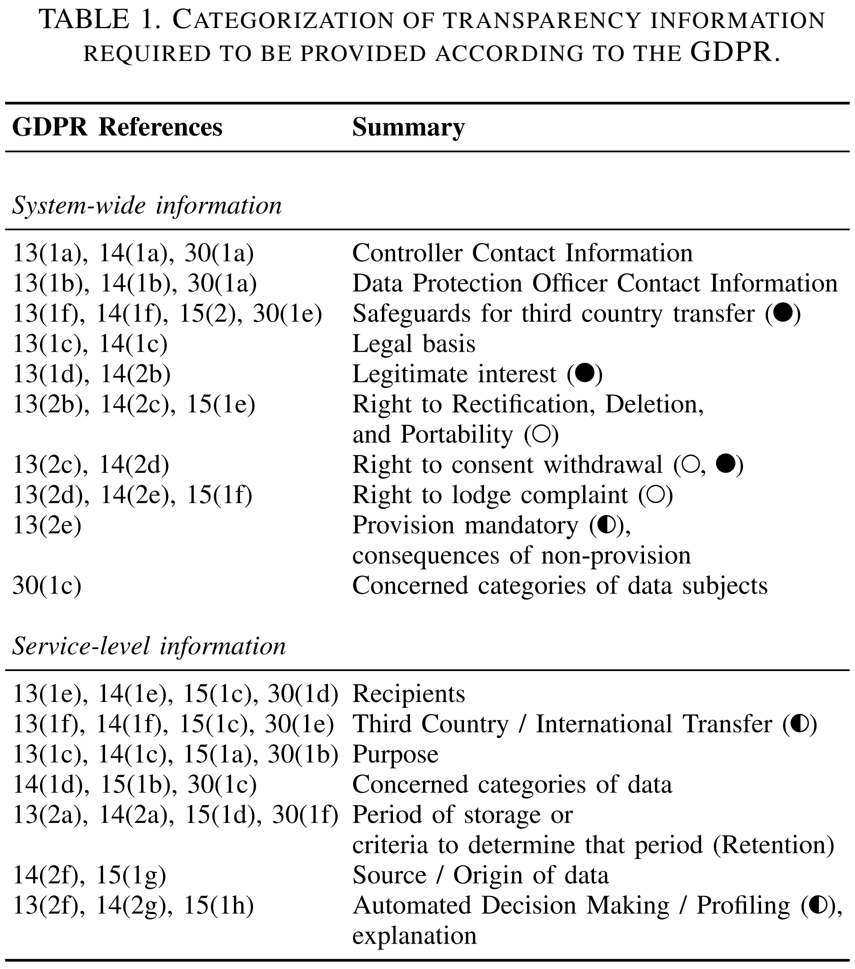
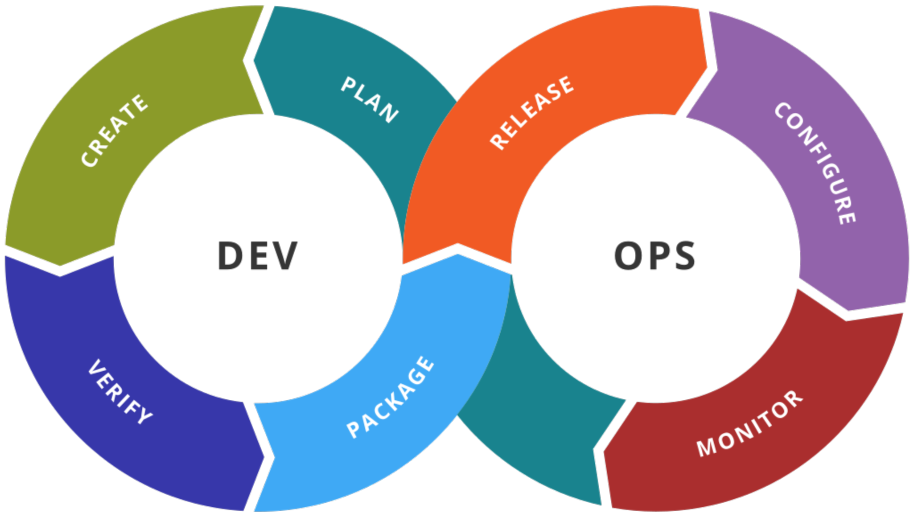


TABLE 1. CATEGORIZATION OF TRANSPARENCY INFORMATION REQUIRED TO BE PROVIDED ACCORDING TO THE GDPR.

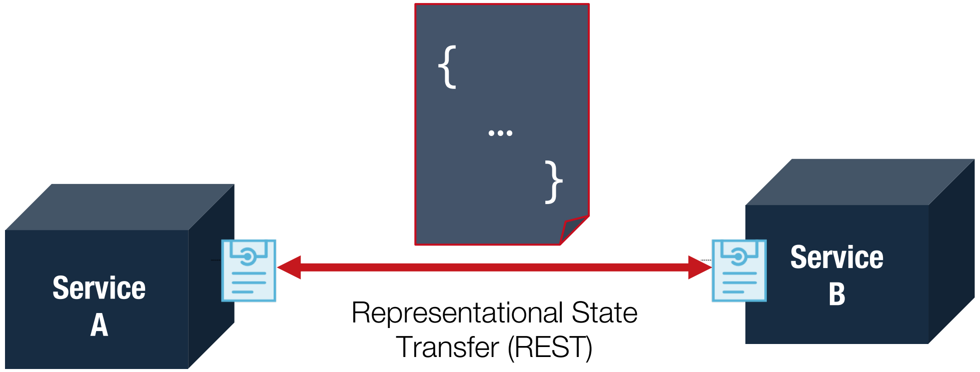
| GDPR References | Summary |
|---|---|
| *System-wide information* | |
| 13(1a), 14(1a), 30(1a) | Controller Contact Information |
| 13(1b), 14(1b), 30(1a) | Data Protection Officer Contact Information |
| 13(1f), 14(1f), 15(2), 30(1e) | Safeguards for third country transfer (●) |
| 13(1c), 14(1c) | Legal basis |
| 13(1d), 14(2b) | Legitimate interest (●) |
| 13(2b), 14(2c), 15(1e) | Right to Rectification, Deletion, and Portability (○) |
| 13(2c), 14(2d) | Right to consent withdrawal (○, ●) |
| 13(2d), 14(2e), 15(1f) | Right to lodge complaint (○) |
| 13(2e) | Provision mandatory (◐), consequences of non-provision |
| 30(1c) | Concerned categories of data subjects |
| *Service-level information* | |
| 13(1e), 14(1e), 15(1c), 30(1d) | Recipients |
| 13(1f), 14(1f), 15(1c), 30(1e) | Third Country / International Transfer (◐) |
| 13(1c), 14(1c), 15(1a), 30(1b) | Purpose |
| 14(1d), 15(1b), 30(1c) | Concerned categories of data |
| 13(2a), 14(2a), 15(1d), 30(1f) | Period of storage or criteria to determine that period (Retention) |
| 14(2f), 15(1g) | Source / Origin of data |
| 13(2f), 14(2g), 15(1h) | Automated Decision Making / Profiling (◐), explanation |

Legend: ○ indication only, ● where applicable, ◐ yes/no

# APIs, DevOps & RESTful Architectures



[i]



- ↗️ Agile development practices with short release cycles in diverse teams
- ↗️ Numerous microservices process personal data
- 😵‍💫 Traditional privacy policies can only provide static information ➡️ new TETs needed

# Requirements

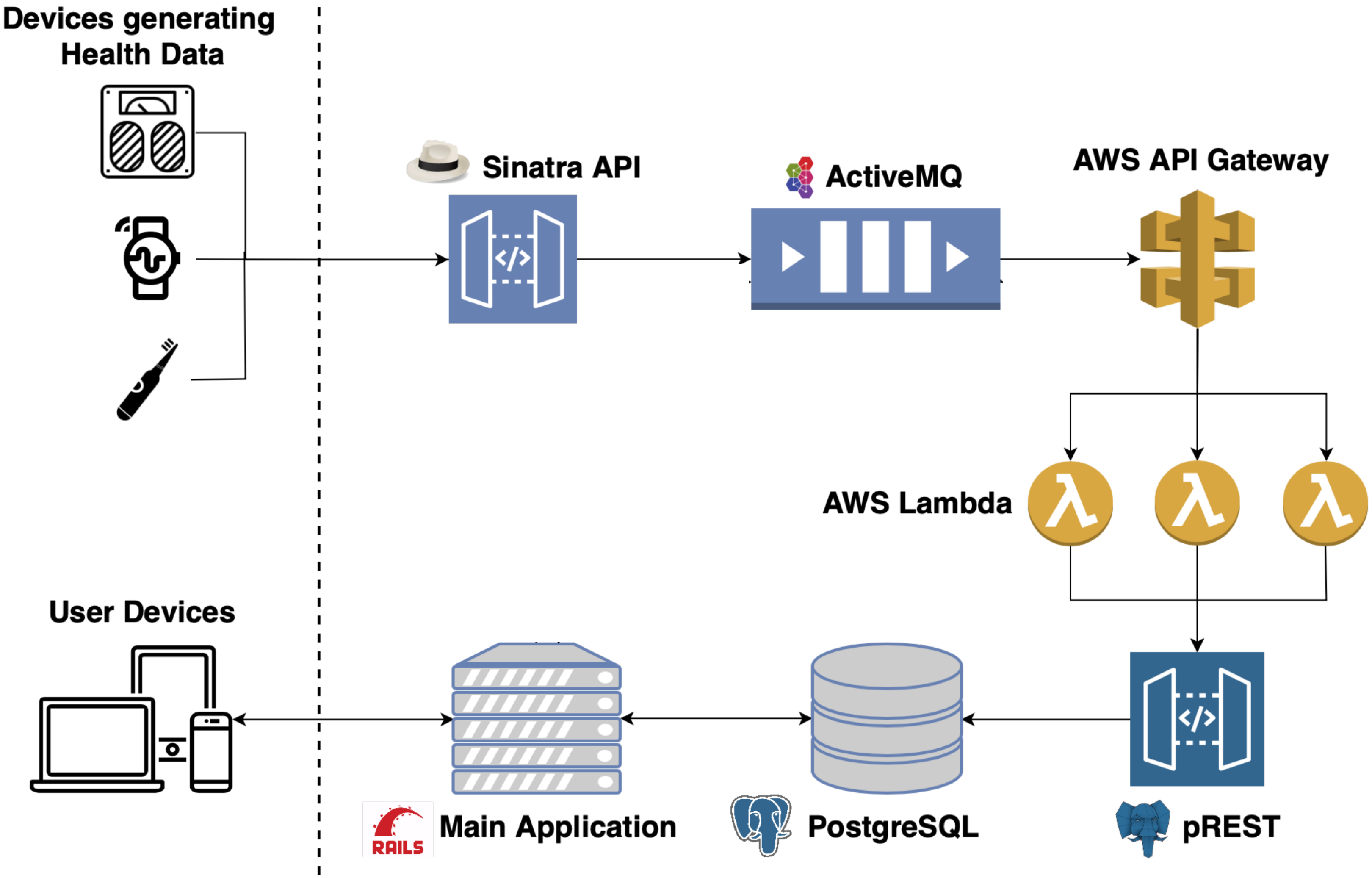
FR

1. Express all legal transparency obligations
2. Service-focused approach (bottom-up)
3. Automated, dynamic, and aggregated perspective (system and services)

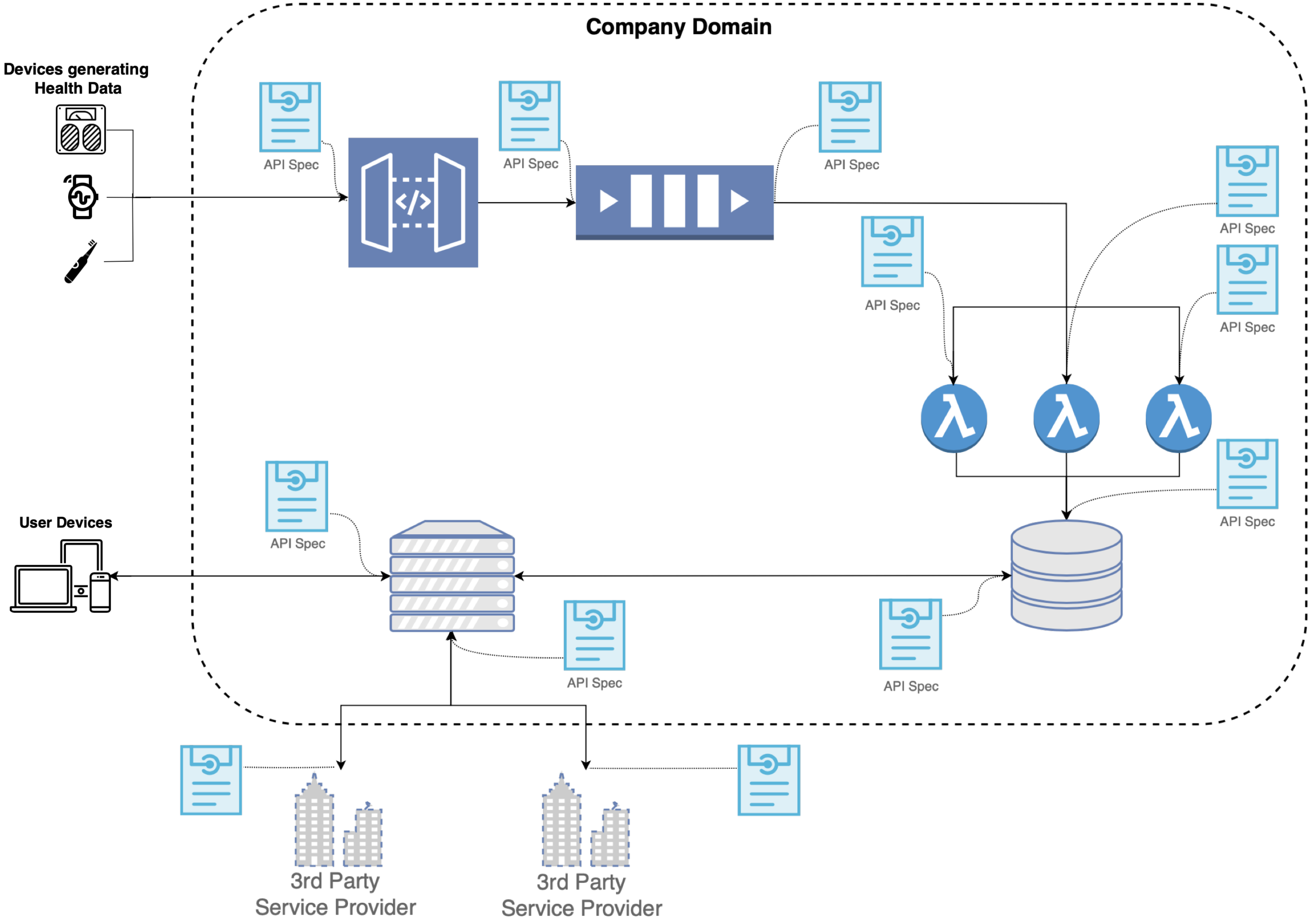
NFR

4. Integrate with **well-known development practices and toolchains**
5. Developer-friendliness, low implementation overhead
6. Re-usable artifact for consistent adoption

# General Approach



# Open API Specifications

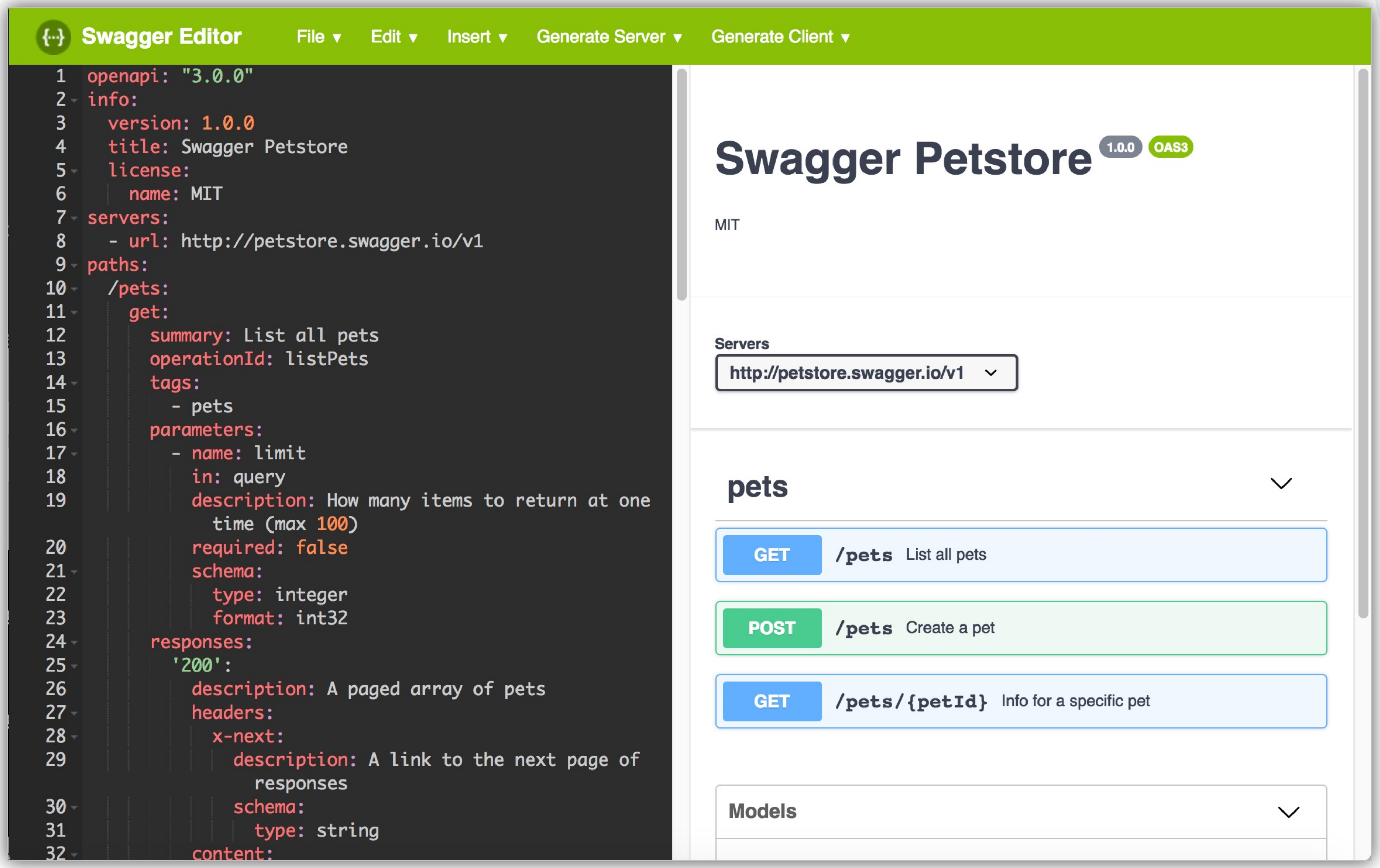


Open API Specification

Swagger

https://swagger.io/specification/

API Spec



```yaml
openapi: "3.0.0"
info:
  version: 1.0.0
  title: Swagger Petstore
  license:
    name: MIT
servers:
  - url: http://petstore.swagger.io/v1
paths:
  /pets:
    get:
      summary: List all pets
      operationId: listPets
      tags:
        - pets
      parameters:
        - name: limit
          in: query
          description: How many items to return at one time (max 100)
          required: false
          schema:
            type: integer
            format: int32
      responses:
        '200':
          description: A paged array of pets
          headers:
            x-next:
              description: A link to the next page of responses
              schema:
                type: string
          content:
```

## Swagger Petstore 1.0.0 OAS3

MIT

Servers

http://petstore.swagger.io/v1

pets

GET /pets List all pets

POST /pets Create a pet

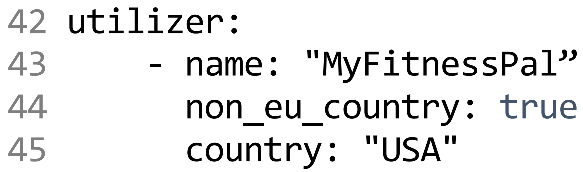
GET /pets/{petId} Info for a specific pet

Models

[ii]

# Open API Specifications

+



```
utilizer:
      - name: "MyFitnessPal”
        non_eu_country: true
        country: "USA"
```

Add personal data indicators *(PD indicators)*

Open API
Specification

Swagger

https://swagger.io/specification/

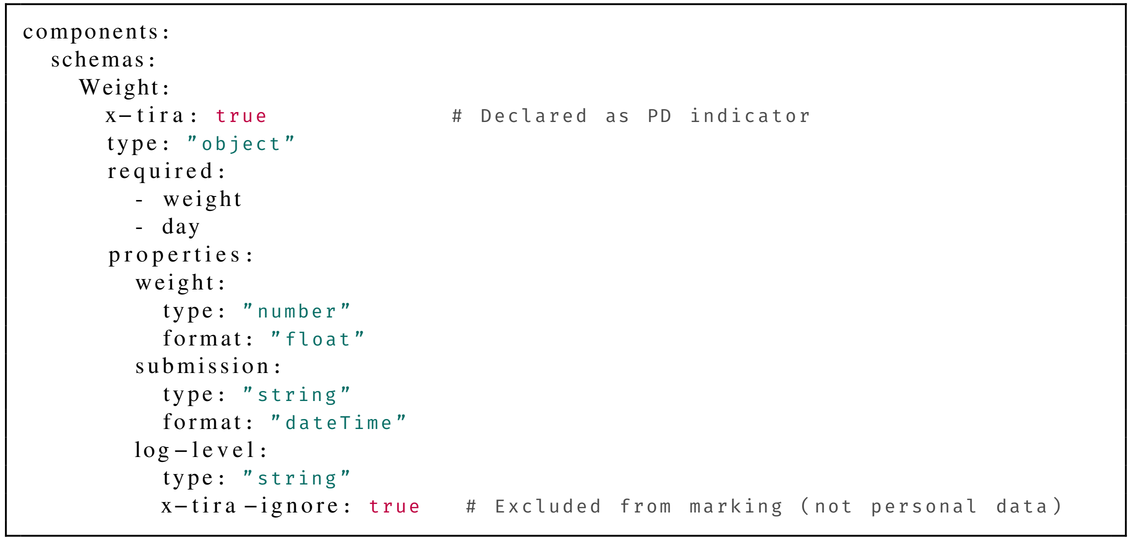
# Extending OpenAPI

1 Declare any data field as *Personal Data indicator*

2 Further annotate each *PD indicator*

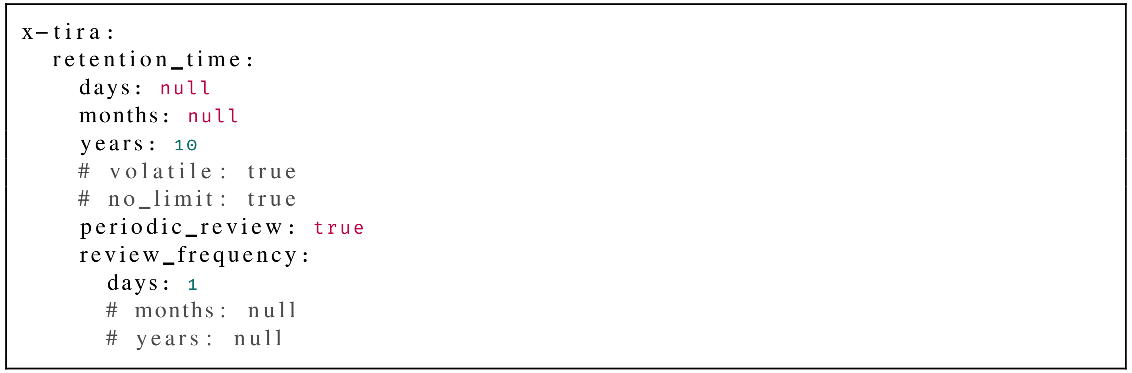
3 Specify transparency properties of a whole service *(not shown)*

1



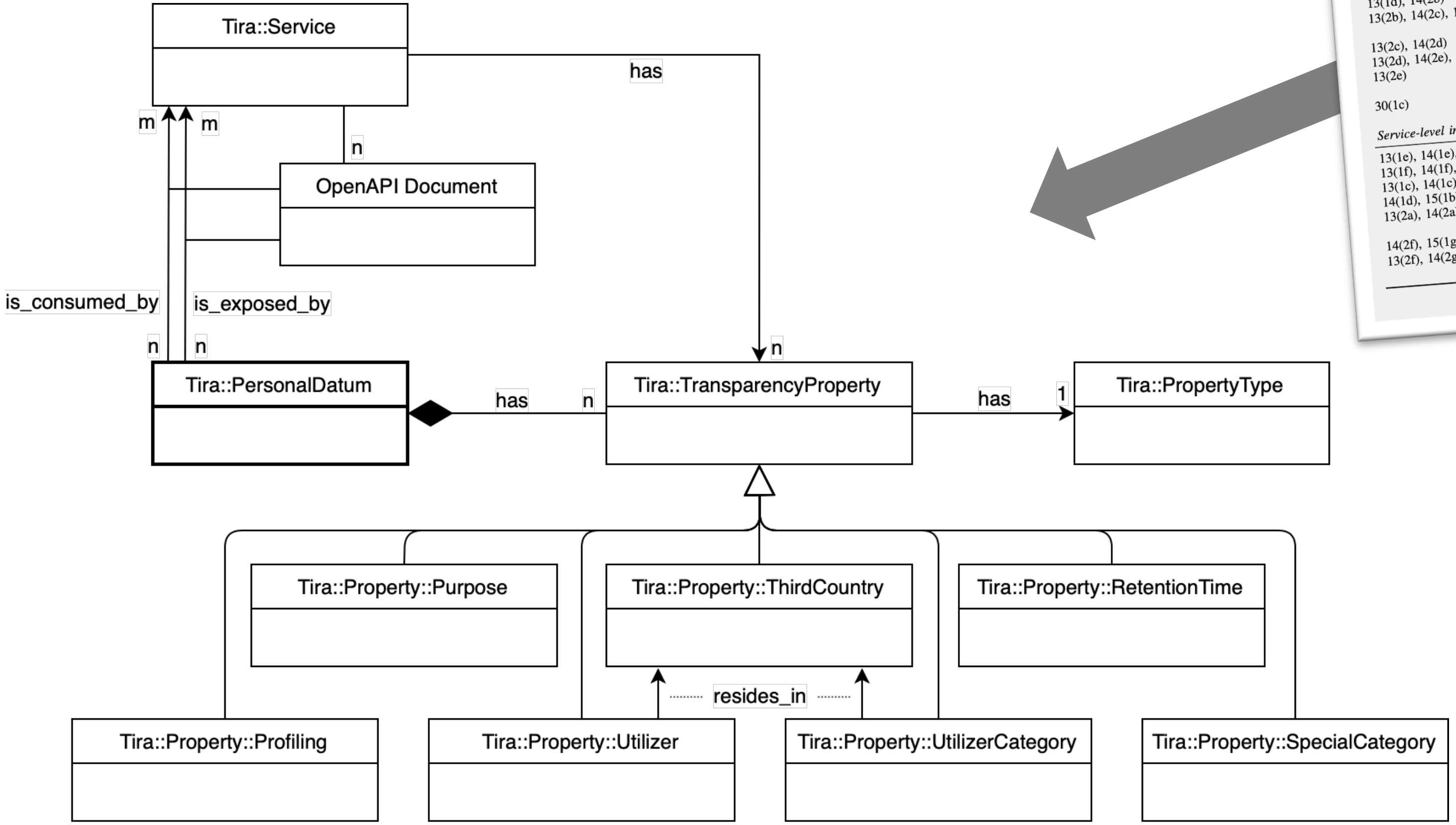
```
components:
  schemas:
    Weight:
      x-tira: true                # Declared as PD indicator
      type: "object"
      required:
        - weight
        - day
      properties:
        weight:
          type: "number"
          format: "float"
        submission:
          type: "string"
          format: "dateTime"
        log-level:
          type: "string"
          x-tira-ignore: true   # Excluded from marking (not personal data)
```

2



```
x-tira:
  retention_time:
    days: null
    months: null
    years: 10
    # volatile: true
    # no_limit: true
    periodic_review: true
    review_frequency:
      days: 1
      # months: null
      # years: null
```

# Vocabulary



Tira::Service

OpenAPI Document

Tira::PersonalDatum

Tira::TransparencyProperty

Tira::PropertyType

Tira::Property::Purpose

Tira::Property::ThirdCountry

Tira::Property::RetentionTime

Tira::Property::Profiling

Tira::Property::Utilizer

Tira::Property::UtilizerCategory

Tira::Property::SpecialCategory

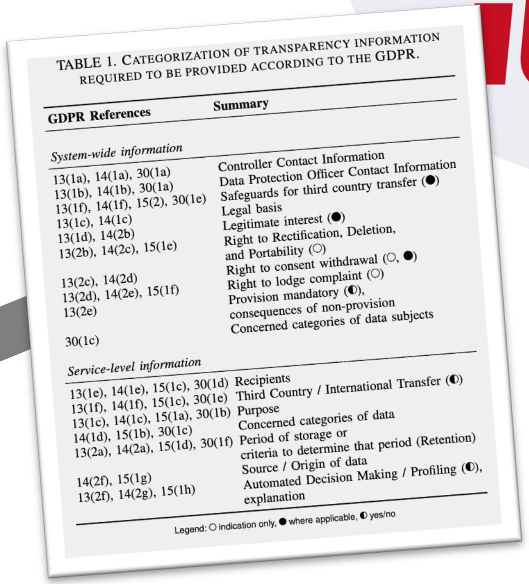


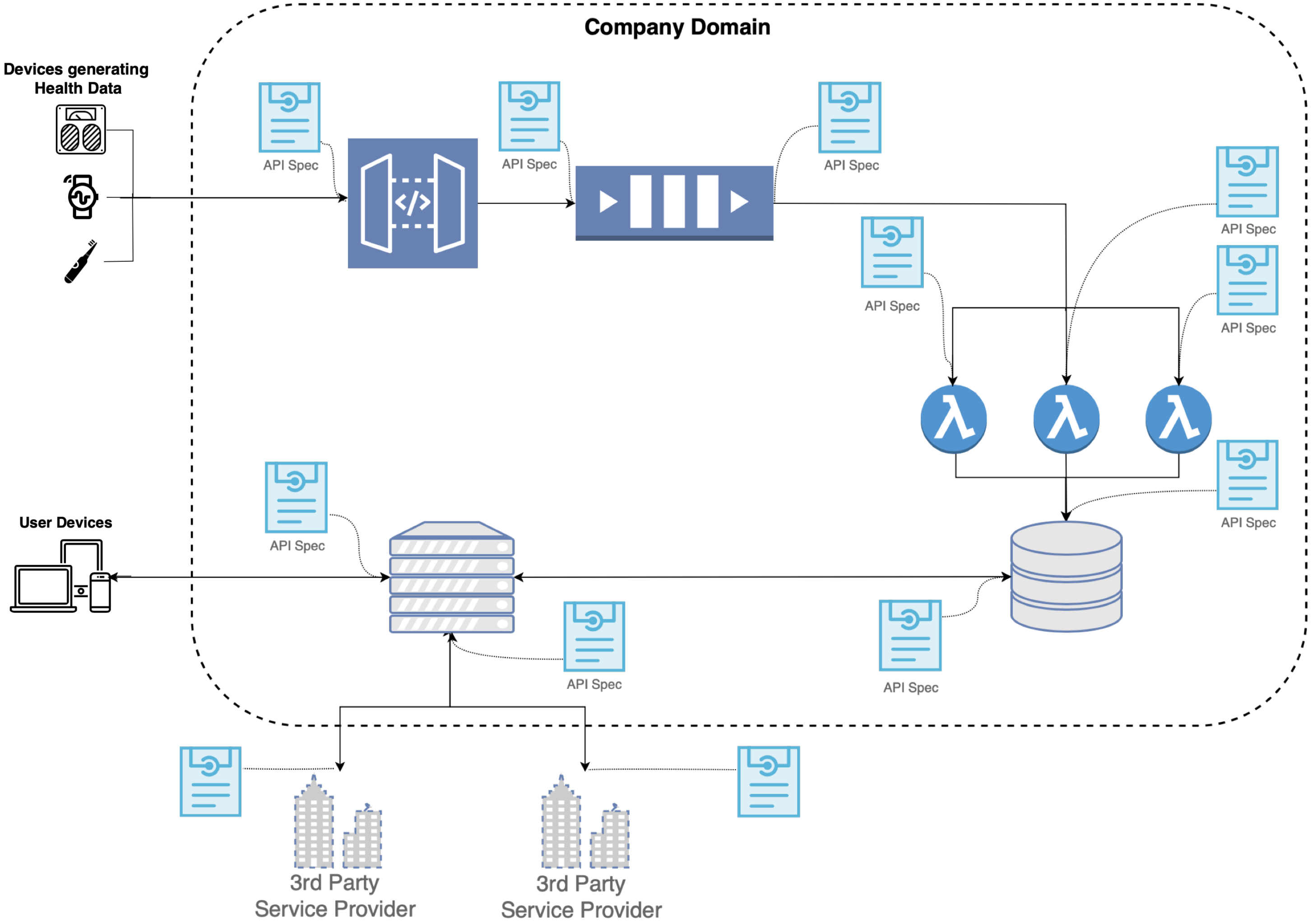
TABLE 1. CATEGORIZATION OF TRANSPARENCY INFORMATION REQUIRED TO BE PROVIDED ACCORDING TO THE GDPR.

| GDPR References | Summary |
|---|---|
| *System-wide information* | |
| 13(1a), 14(1a), 30(1a) | Controller Contact Information |
| 13(1b), 14(1b), 30(1a) | Data Protection Officer Contact Information |
| 13(1f), 14(1f), 15(2), 30(1e) | Safeguards for third country transfer (●) |
| 13(1c), 14(1c) | Legal basis |
| 13(1d), 14(2b) | Legitimate interest (●) |
| 13(2b), 14(2c), 15(1e) | Right to Rectification, Deletion, and Portability (○) |
| 13(2c), 14(2d) | Right to consent withdrawal (○, ●) |
| 13(2d), 14(2e), 15(1f) | Right to lodge complaint (○) |
| 13(2e) | Provision mandatory (◐), consequences of non-provision |
| 30(1c) | Concerned categories of data subjects |
| *Service-level information* | |
| 13(1e), 14(1e), 15(1c), 30(1d) | Recipients |
| 13(1f), 14(1f), 15(1c), 30(1e) | Third Country / International Transfer (◐) |
| 13(1c), 14(1c), 15(1a), 30(1b) | Purpose |
| 14(1d), 15(1b), 30(1c) | Concerned categories of data |
| 13(2a), 14(2a), 15(1d), 30(1f) | Period of storage or criteria to determine that period (Retention) |
| 14(2f), 15(1g) | Source / Origin of data |
| 13(2f), 14(2g), 15(1h) | Automated Decision Making / Profiling (◐), explanation |

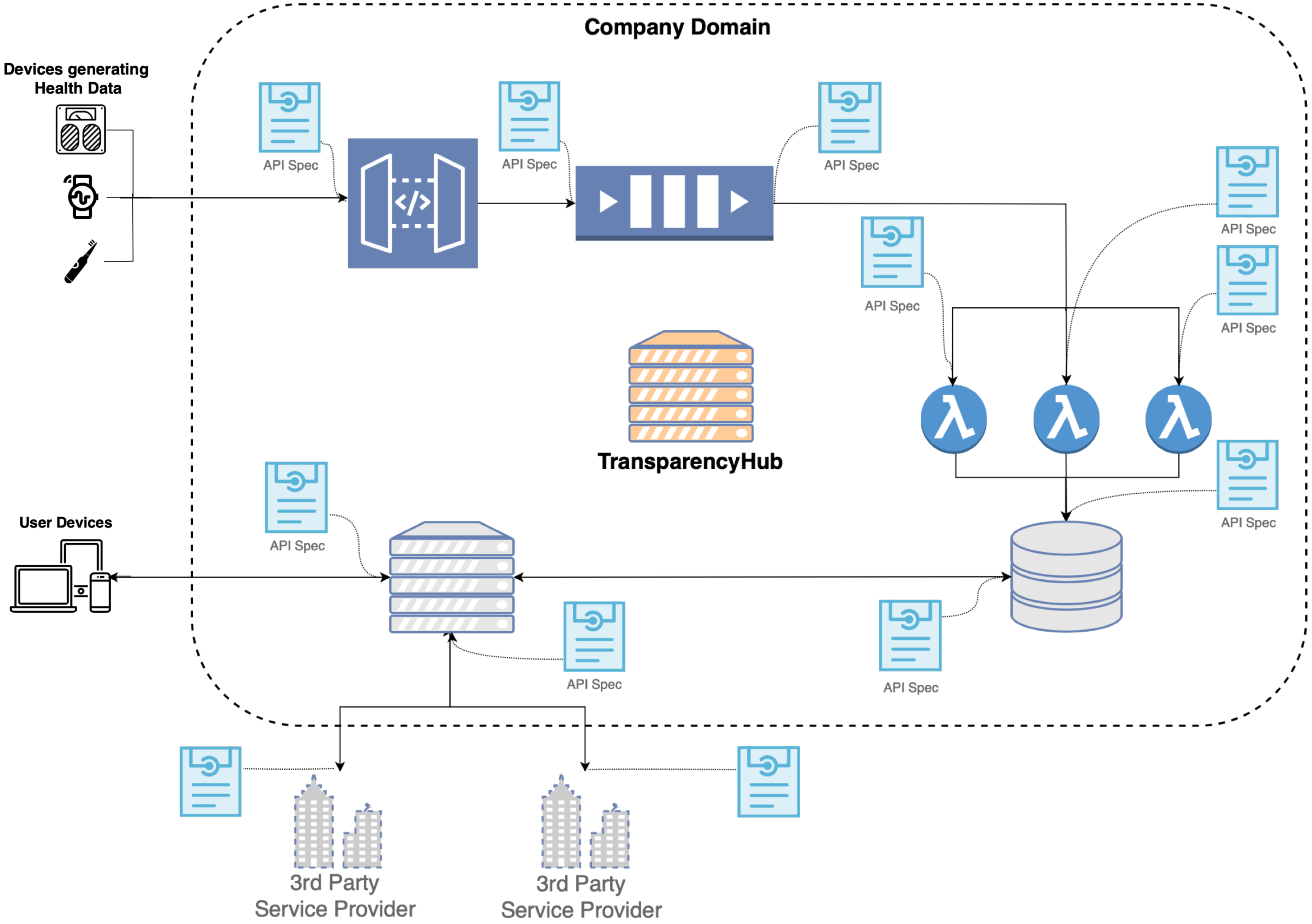
Legend: ○ indication only, ● where applicable, ◐ yes/no

https://github.com/PrivacyEngineering/tira/blob/main/docs/VOCABULARY.md

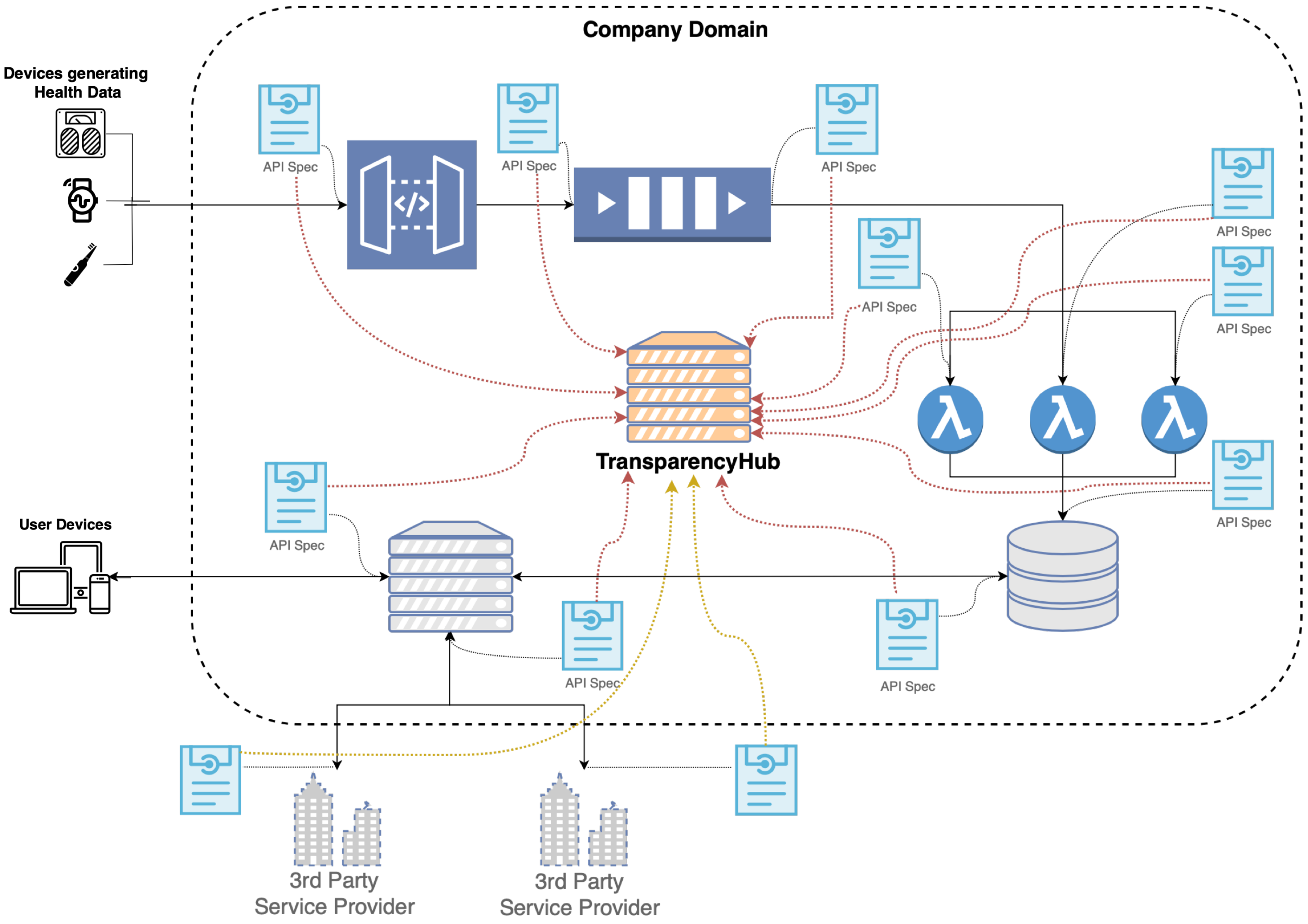
# Managing system-wide transparency



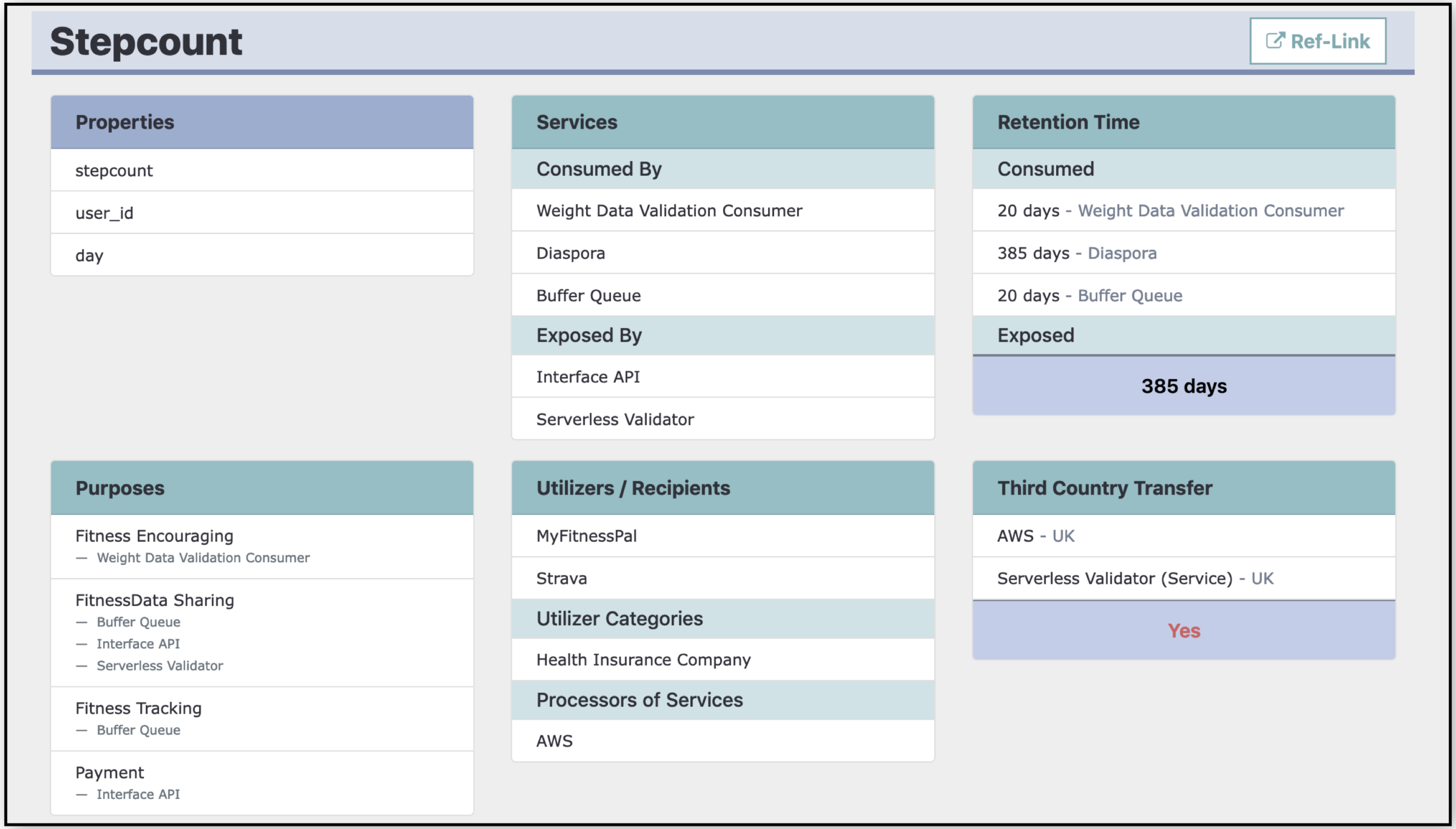
# Introducing *TransparencyHub*



# Introducing *TransparencyHub*



# Aggregating transparency information



## Stepcount

### Properties

- stepcount
- user_id
- day

### Services

**Consumed By**

- Weight Data Validation Consumer
- Diaspora
- Buffer Queue

**Exposed By**

- Interface API
- Serverless Validator

### Retention Time

**Consumed**

- 20 days - Weight Data Validation Consumer
- 385 days - Diaspora
- 20 days - Buffer Queue

**Exposed**

**385 days**

### Purposes

Fitness Encouraging
— Weight Data Validation Consumer

FitnessData Sharing
— Buffer Queue
— Interface API
— Serverless Validator

Fitness Tracking
— Buffer Queue

Payment
— Interface API

### Utilizers / Recipients

- MyFitnessPal
- Strava

**Utilizer Categories**

- Health Insurance Company

**Processors of Services**

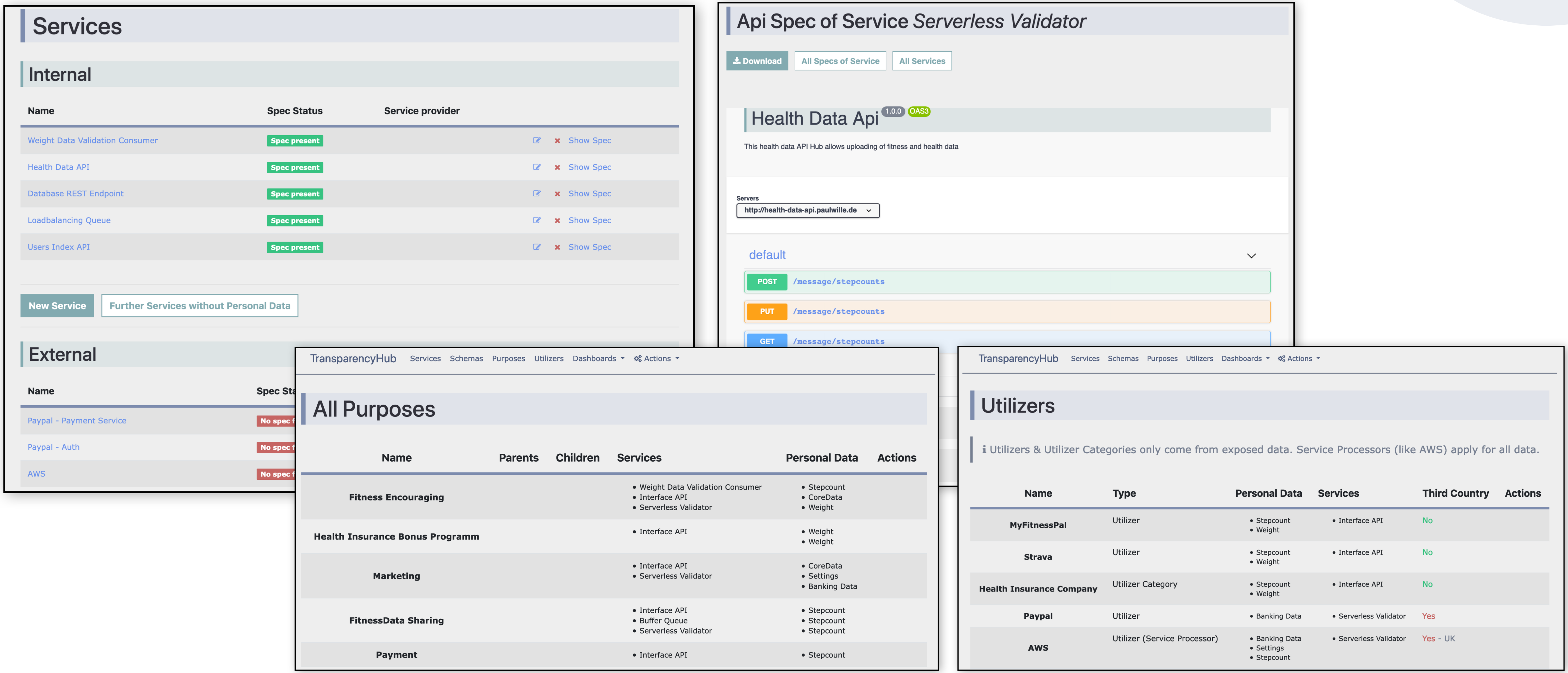
- AWS

### Third Country Transfer

- AWS - UK
- Serverless Validator (Service) - UK

Yes

# Further insights and management



## Services

### Internal

| Name | Spec Status | Service provider | |
|---|---|---|---|
| Weight Data Validation Consumer | Spec present | | ✖ Show Spec |
| Health Data API | Spec present | | ✖ Show Spec |
| Database REST Endpoint | Spec present | | ✖ Show Spec |
| Loadbalancing Queue | Spec present | | ✖ Show Spec |
| Users Index API | Spec present | | ✖ Show Spec |

### External

| Name | Spec Sta |
|---|---|
| Paypal - Payment Service | No spec f |
| Paypal - Auth | No spec f |
| AWS | No spec f |

## Api Spec of Service *Serverless Validator*

### Health Data Api 1.0.0 OAS3

This health data API Hub allows uploading of fitness and health data

Servers
http://health-data-api.paulwille.de

TransparencyHub Services Schemas Purposes Utilizers Dashboards Actions

## All Purposes

| Name | Parents | Children | Services | Personal Data | Actions |
|---|---|---|---|---|---|
| Fitness Encouraging | | | • Weight Data Validation Consumer<br>• Interface API<br>• Serverless Validator | • Stepcount<br>• CoreData<br>• Weight | |
| Health Insurance Bonus Programm | | | • Interface API | • Weight<br>• Weight | |
| Marketing | | | • Interface API<br>• Serverless Validator | • CoreData<br>• Settings<br>• Banking Data | |
| FitnessData Sharing | | | • Interface API<br>• Buffer Queue<br>• Serverless Validator | • Stepcount<br>• Stepcount<br>• Stepcount | |
| Payment | | | • Interface API | • Stepcount | |

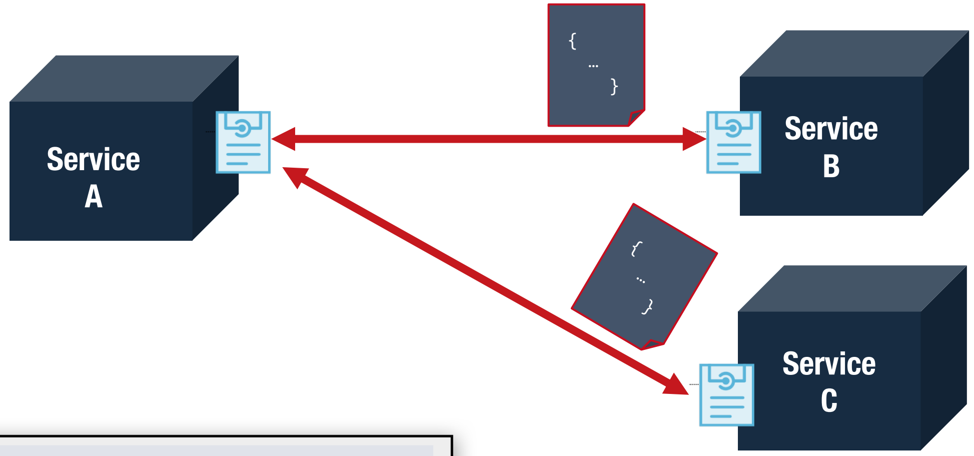
TransparencyHub Services Schemas Purposes Utilizers Dashboards Actions

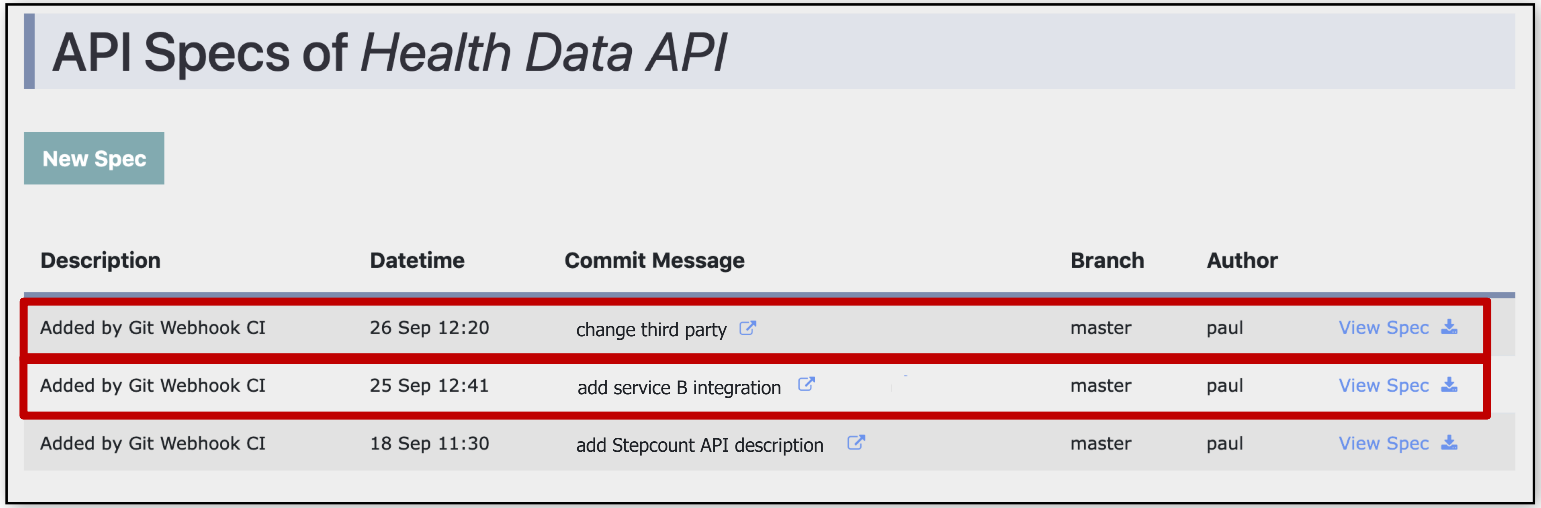
## Utilizers

ℹ Utilizers & Utilizer Categories only come from exposed data. Service Processors (like AWS) apply for all data.

| Name | Type | Personal Data | Services | Third Country | Actions |
|---|---|---|---|---|---|
| MyFitnessPal | Utilizer | • Stepcount<br>• Weight | • Interface API | No | |
| Strava | Utilizer | • Stepcount<br>• Weight | • Interface API | No | |
| Health Insurance Company | Utilizer Category | • Stepcount<br>• Weight | • Interface API | No | |
| Paypal | Utilizer | • Banking Data | • Serverless Validator | Yes | |
| AWS | Utilizer (Service Processor) | • Banking Data<br>• Settings<br>• Stepcount | • Serverless Validator | Yes - UK | |

# DevOps / Continuous Integration and Delivery





## API Specs of *Health Data API*

New Spec

| Description | Datetime | Commit Message | Branch | Author | |
|---|---|---|---|---|---|
| Added by Git Webhook CI | 26 Sep 12:20 | change third party | master | paul | View Spec |
| Added by Git Webhook CI | 25 Sep 12:41 | add service B integration | master | paul | View Spec |
| Added by Git Webhook CI | 18 Sep 11:30 | add Stepcount API description | master | paul | View Spec |

# Discussion & Conclusion

*First of its kind* **developer-focused and GDPR-aligned** OpenAPI extension

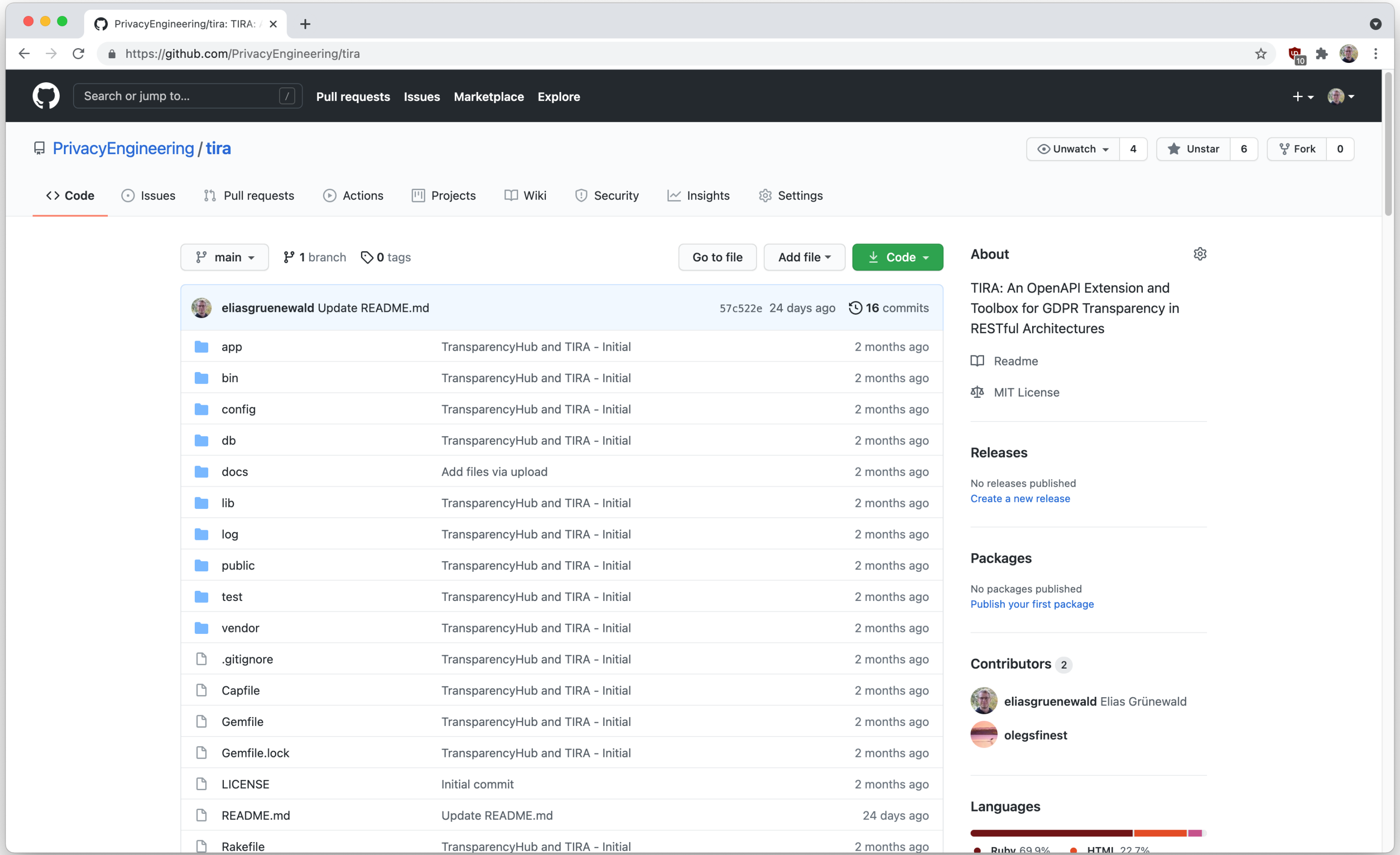
**DevOps-driven** approach for transparency

—

Future work includes
**other service description formats and service registries**, integration of **advanced vocabularies** (such as **TILT***), **presentation means for data subjects**…

* **Transparency Information Language and Toolkit** (Grünewald and Pallas 2021): https://dl.acm.org/doi/10.1145/3442188.3445925

# Open Source Software (MIT License) – Get involved!



PrivacyEngineering / tira

eliasgruenewald Update README.md 57c522e 24 days ago 16 commits

| | | |
|---|---|---|
| app | TransparencyHub and TIRA - Initial | 2 months ago |
| bin | TransparencyHub and TIRA - Initial | 2 months ago |
| config | TransparencyHub and TIRA - Initial | 2 months ago |
| db | TransparencyHub and TIRA - Initial | 2 months ago |
| docs | Add files via upload | 2 months ago |
| lib | TransparencyHub and TIRA - Initial | 2 months ago |
| log | TransparencyHub and TIRA - Initial | 2 months ago |
| public | TransparencyHub and TIRA - Initial | 2 months ago |
| test | TransparencyHub and TIRA - Initial | 2 months ago |
| vendor | TransparencyHub and TIRA - Initial | 2 months ago |
| .gitignore | TransparencyHub and TIRA - Initial | 2 months ago |
| Capfile | TransparencyHub and TIRA - Initial | 2 months ago |
| Gemfile | TransparencyHub and TIRA - Initial | 2 months ago |
| Gemfile.lock | TransparencyHub and TIRA - Initial | 2 months ago |
| LICENSE | Initial commit | 2 months ago |
| README.md | Update README.md | 24 days ago |
| Rakefile | TransparencyHub and TIRA - Initial | 2 months ago |

**About**

TIRA: An OpenAPI Extension and Toolbox for GDPR Transparency in RESTful Architectures

Readme

MIT License

**Releases**

No releases published
Create a new release

**Packages**

No packages published
Publish your first package

**Contributors** 2

eliasgruenewald Elias Grünewald
olegsfinest

**Languages**

● Ruby 69.9% ● HTML 22.7%

➡️ `https://github.com/PrivacyEngineering/tira`

2021 International Workshop on Privacy Engineering (IWPE‘21)

# TIRA: An OpenAPI Extension and Toolbox for GDPR Transparency in RESTful Architectures

**Elias Grünewald**
eg@ise.tu-berlin.de
@eliasgruenewald

**Paul Wille**
pw@ise.tu-berlin.de

**Dr.-Ing. Frank Pallas**
fp@ise.tu-berlin.de
@sallapf

**Maria C. Borges**
mb@ise.tu-berlin.de
@blablablorges

**Max-R. Ulbricht**
mu@ise.tu-berlin.de
@maroulb

The work behind this paper was partially conducted within the project DaSKITA (https://daskita.github.io), supported under grant no. 28V2307A19 by funds of the Federal Ministry of Justice and Consumer Protection (BMJV) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme.

ISEngineering

# Sources

**See paper for complete bibliography.**

[1] Joel R. Reidenberg, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T. Graves, Fei Liu, Aleecia McDonald, Thomas B. Norton, and Rohan Ramanath. 2015. Disagreeable Privacy Policies: Mismatches between Meaning and Users' Understanding. *Berkeley Technology Law Journal* 30, 39.

[2] Elias Grünewald and Frank Pallas. 2021. TILT: A GDPR-Aligned Transparency Information Language and Toolkit. In: *Proceedings of the 2021 Conference on Fairness Accountability and Transparency (FAccT'21),* ACM, *pp. 636-646*.

[3] Marit Hansen. Data protection by design and by default à la European General Data Protection Regulation. In: *IFIP Summer School on Privacy and Identity Management*. Springer, pp. 27-38.

[4] Seda Gürses and Joris van Hoboken. 2018. Privacy after the Agile Turn. Ser. Cambridge Law Handbooks. Cambridge University Press, pp. 579-601.

[i] Illustration showing stages in a DevOps toolchain. CC-BY-SA 4.0. Kharnagy. https://commons.wikimedia.org/wiki/File:Devops-toolchain.svg

[ii] OpenAPI/Swagger UI. https://idratherbewriting.com/learnapidoc/pubapis_openapi_tutorial_overview.html