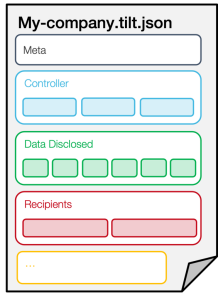
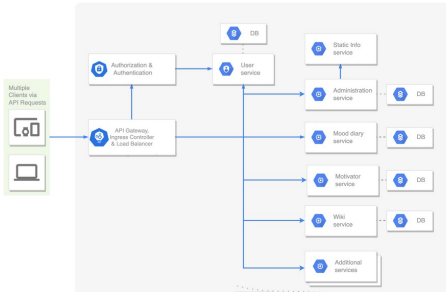
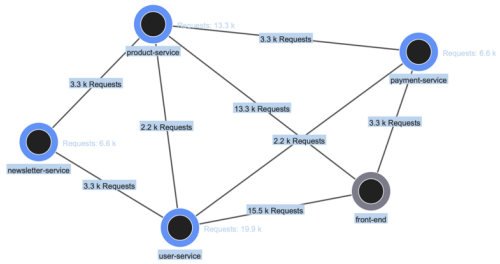
IEEE CLOUD 2023

Technische Universität Berlin







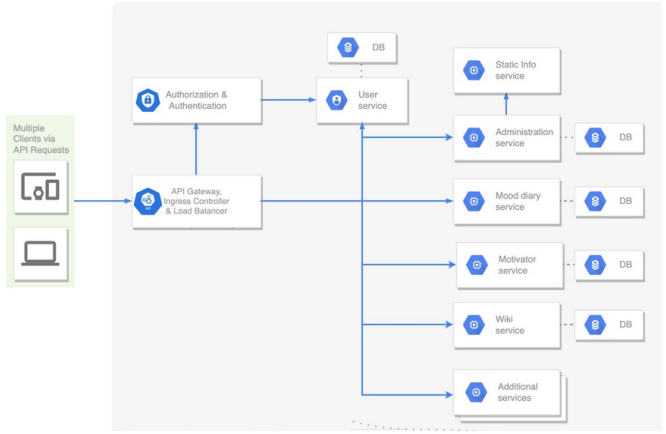
Hawk: DevOps-driven Transparency and Accountability in Cloud Native Systems

Elias Grünewald, Jannis Kiesel, Siar-Remzi Akbayin, and Frank Pallas

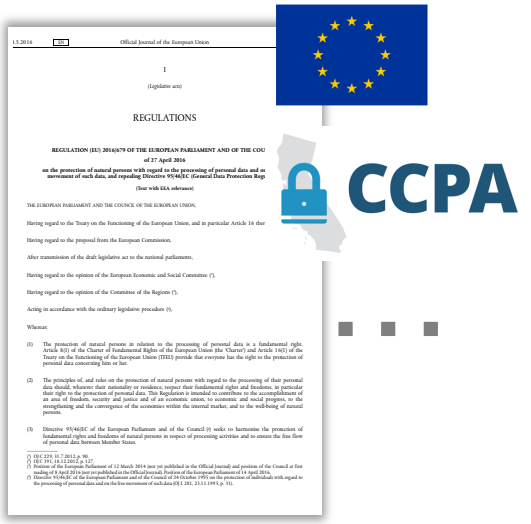
Information Systems Engineering
TU Berlin

ISEngineering

Cloud Native Privacy Engineering



CLOUD NATIVE COMPUTING FOUNDATION



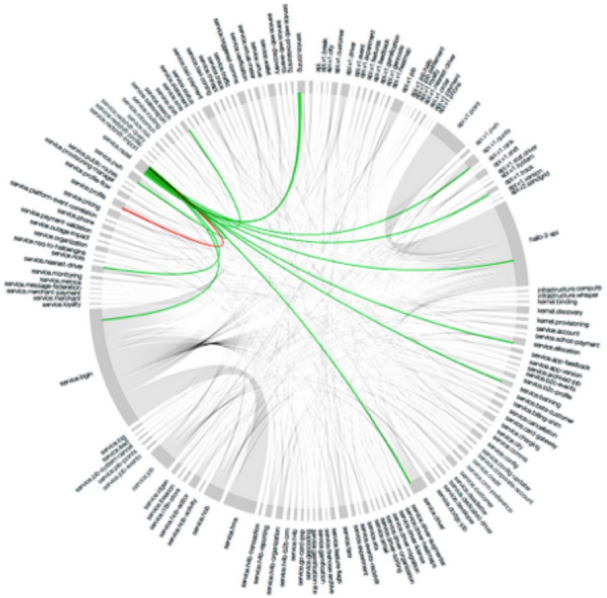
CCPA

. . .

ISEngineering

How modern software architectures look like

450+ microservices

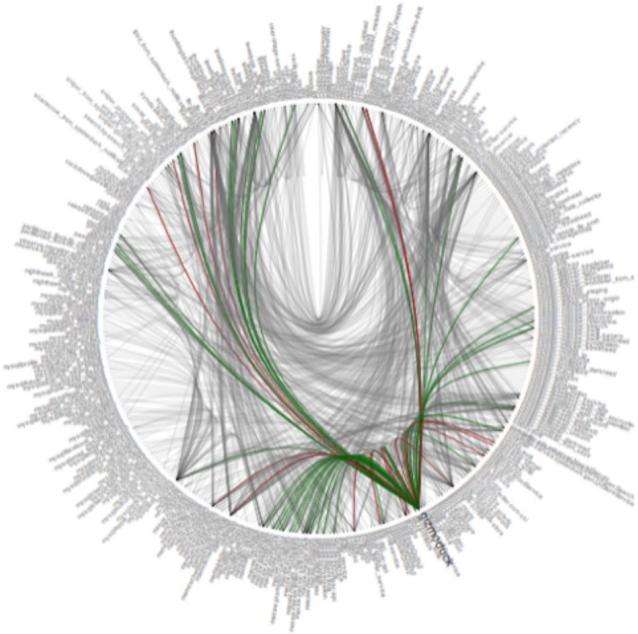


HAIL

500+ microservices

NETFLIX

500+ microservices



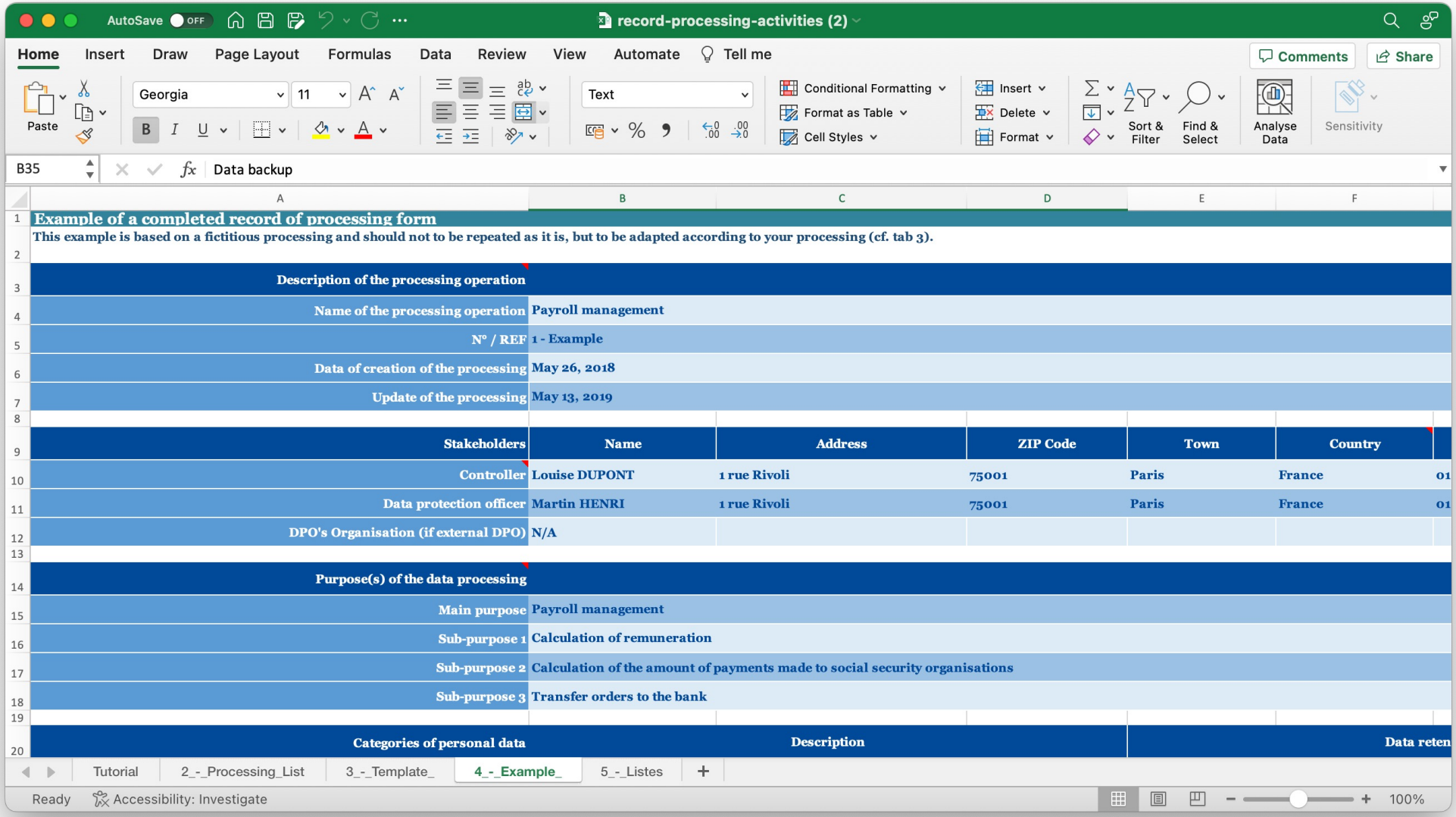
https://sp-ao.shortpixel.ai/client/to_auto,q_glossy,ret_img,w_530,h_262/https://www.peerislands.io/wp-content/uploads/2020/07/microservices-death-star.png

How current transparency measures look like

https://www.fastcompany.com/90171107/printing-out-the-privacy-policies-of-facebook-snap-and-others

https://www.fastcompany.com/90171107/printing-out-the-privacy-policies-of-facebook-snap-and-others

How current accountability measures look like



AutoSave OFF — record-processing-activities (2)

Home Insert Draw Page Layout Formulas Data Review View Automate Tell me — Comments Share

Paste | Georgia | 11 | Text | Conditional Formatting | Format as Table | Cell Styles | Insert | Delete | Format | Sort & Filter | Find & Select | Analyse Data | Sensitivity

B35 — Data backup

Example of a completed record of processing form

This example is based on a fictitious processing and should not to be repeated as it is, but to be adapted according to your processing (cf. tab 3).

Description of the processing operation						
Name of the processing operation	Payroll management					
N° / REF	1 - Example					
Data of creation of the processing	May 26, 2018					
Update of the processing	May 13, 2019					
Stakeholders	**Name**	**Address**	**ZIP Code**	**Town**	**Country**	
Controller	Louise DUPONT	1 rue Rivoli	75001	Paris	France	01
Data protection officer	Martin HENRI	1 rue Rivoli	75001	Paris	France	01
DPO's Organisation (if external DPO)	N/A					
Purpose(s) of the data processing						
Main purpose	Payroll management					
Sub-purpose 1	Calculation of remuneration					
Sub-purpose 2	Calculation of the amount of payments made to social security organisations					
Sub-purpose 3	Transfer orders to the bank					
Categories of personal data			**Description**			**Data reten**

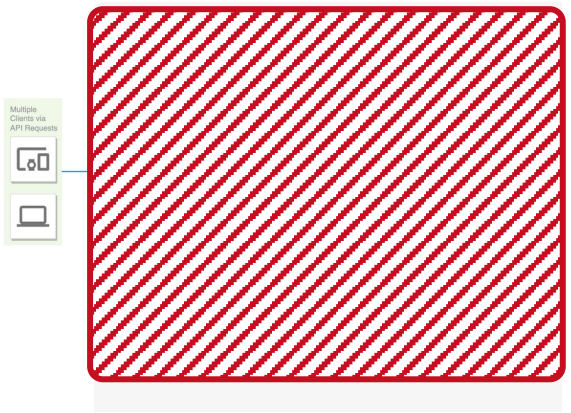
Tutorial | 2_-_Processing_List | 3_-_Template_ | 4_-_Example_ | 5_-_Listes

Ready — Accessibility: Investigate — 100%

https://www.cnil.fr/en/record-processing-activities

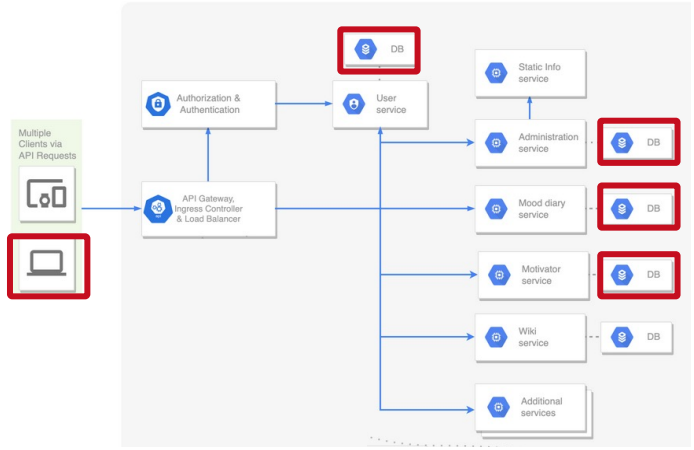
Architect's / Privacy Engineer's perspectives

1 "Black box" / Overview



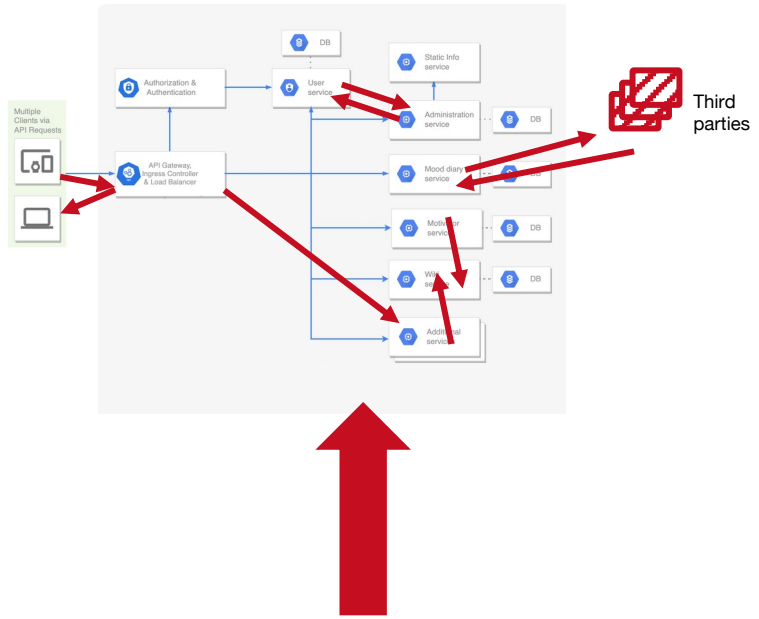
Elias Grünewald and Frank Pallas. 2021. TILT: A GDPR-Aligned Transparency Information Language and Toolkit for Practical Privacy Engineering. In *Proceedings of the 2021 ACM FAccT, ACM, New York*

2 Personal data *at rest*



Elias Grünewald and Leonard Schurbert. 2022. Scalable Discovery and Continuous Inventory of Personal Data at Rest in Cloud Native Systems. In: *Proceedings of the International Conference on Service-Oriented Computing (ICSOC)*, Springer

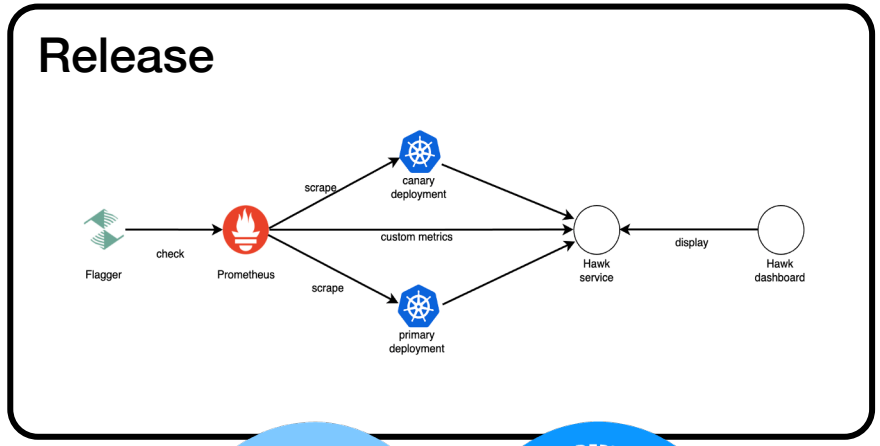
3 Personal data *in transit*

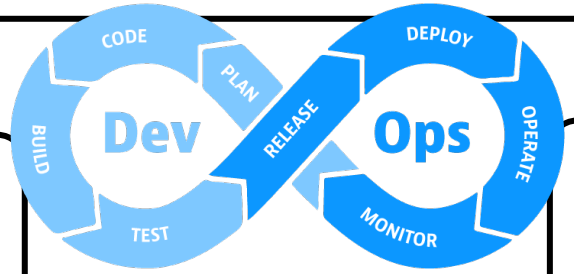


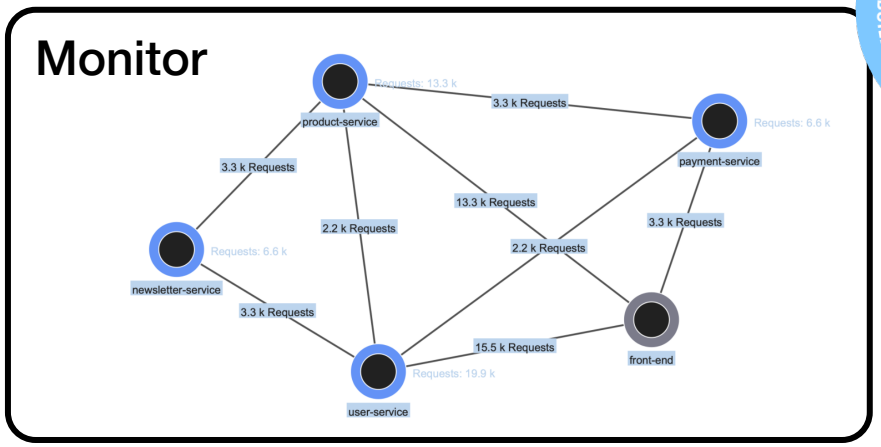
See also for API transparency:

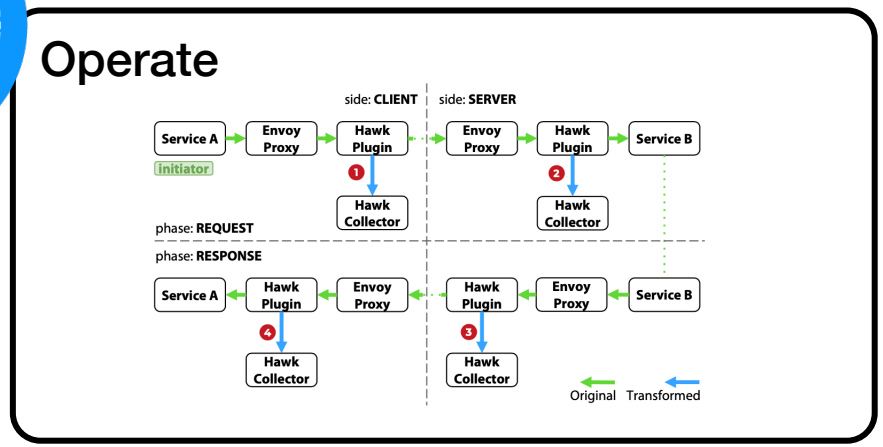
Elias Grünewald, Paul Wille, Frank Pallas, Maria C. Borges and Max-R. Ulbricht. 2021. TIRA: An OpenAPI Extension and Toolbox for GDPR Transparency in RESTful Architectures. In: *Proceedings of the 2021 EuroS&PW (IWPE)*, pp. 312-31

Hawk: DevOps-driven Transparency and Accountability in Cloud Native Systems



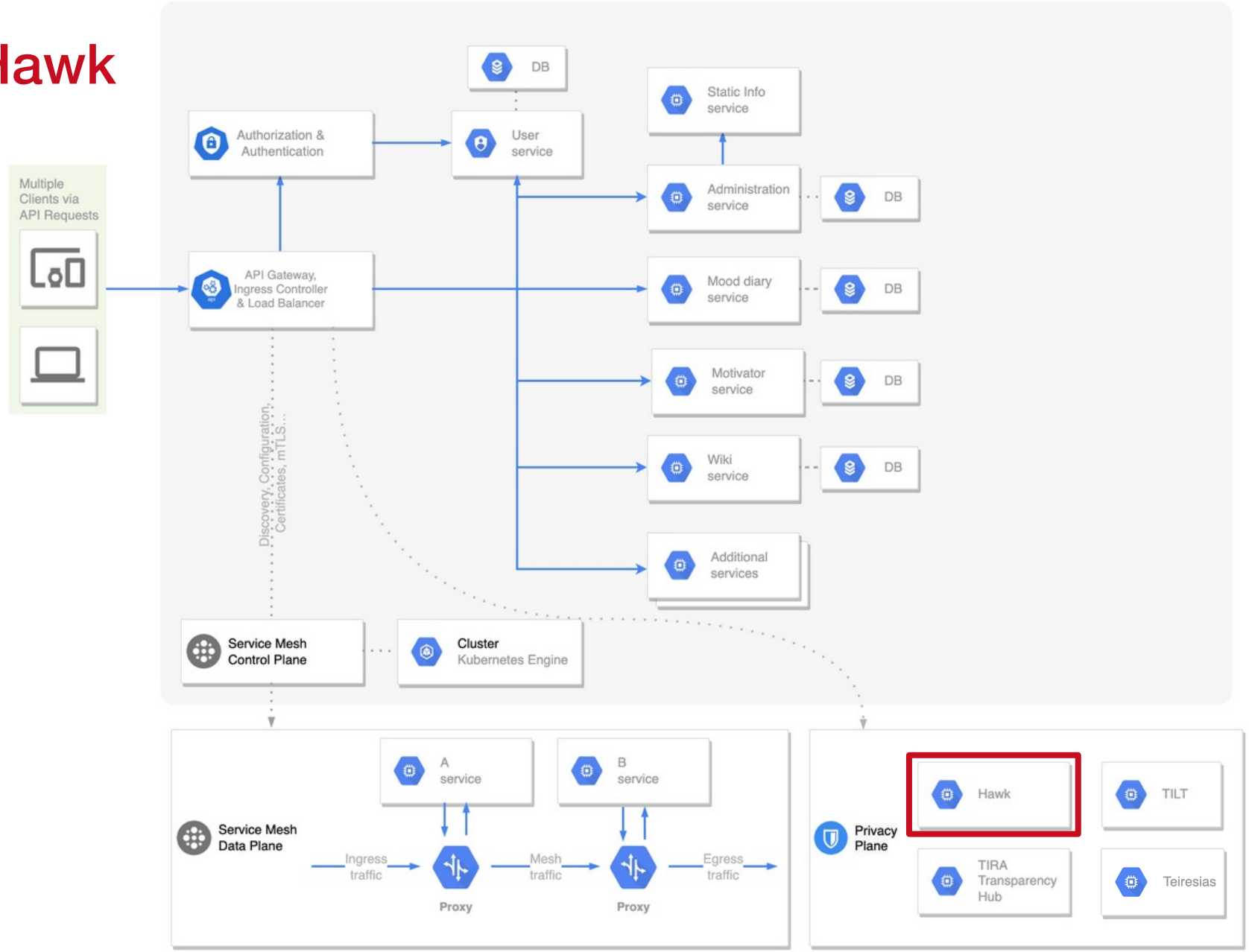






Elias Grünewald, Jannis Kiesel, Siar-Remzi Akbayin, and Frank Pallas. 2023 (under review at IEEE CLOUD). "Hawk: DevOps-driven Transparency and Accountability in Cloud Native Systems".

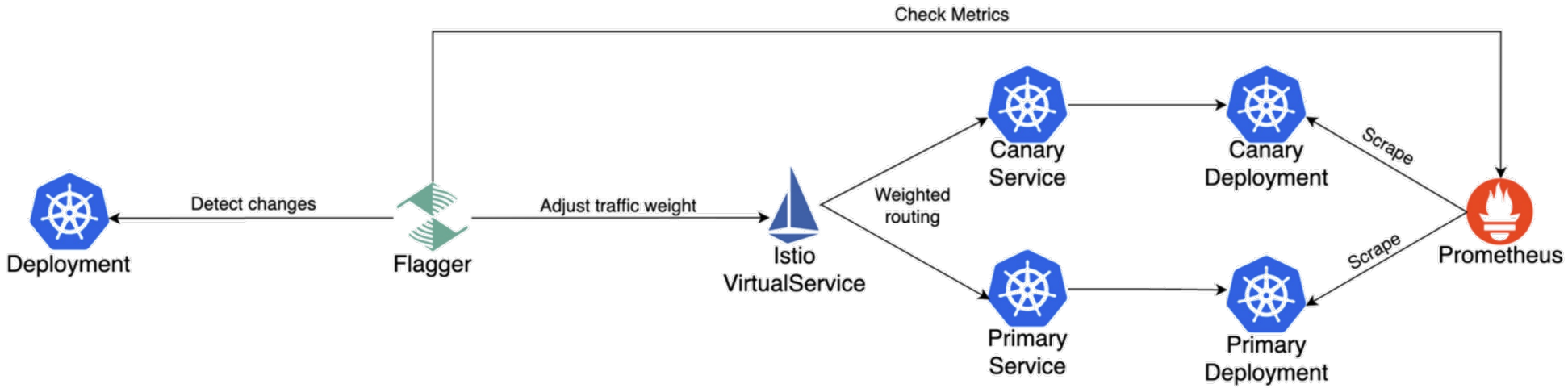
Hawk



Exemplary microservice infrastructure

Privacy Plane

ISEngineering



Hawk Release

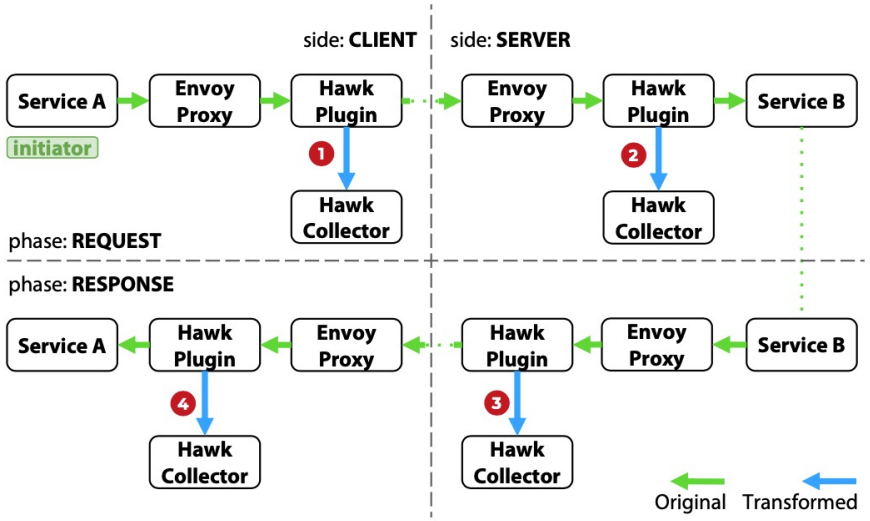
Challenges: Fast deployments, potentially new personal data processing activities

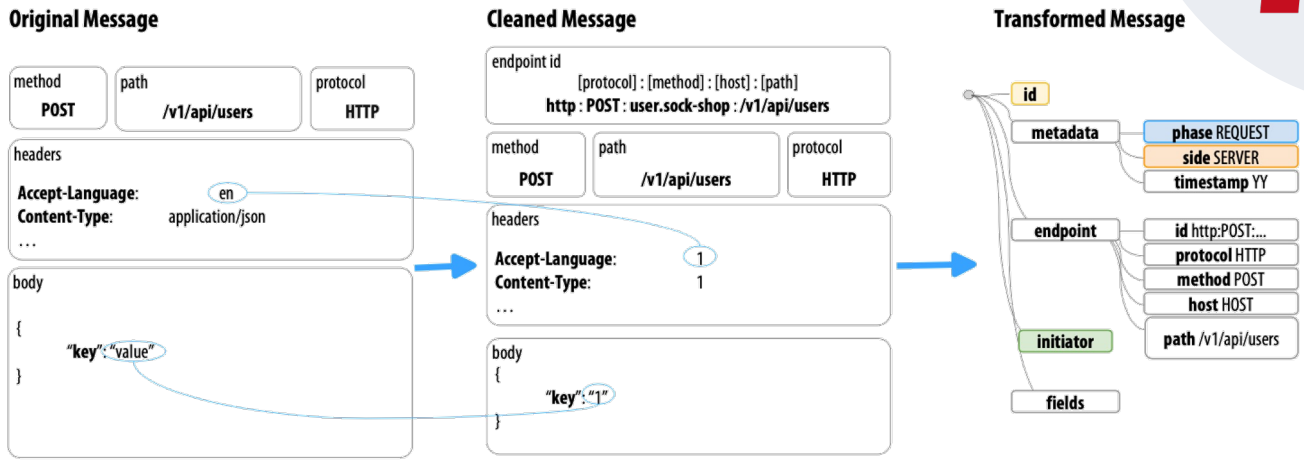
Approach: Canary releases (incl. Deployment, Load shifting, Observation, Clean-up[1])

Implementation: GitOps (Flux, Flagger, Kustomize, Helm), custom metric templates

Limitations: Other deployment strategies, additional safeguards, reporting, legal evaluations

[1] Ernst, Becker, Tai. 2019. (ICSA-C)





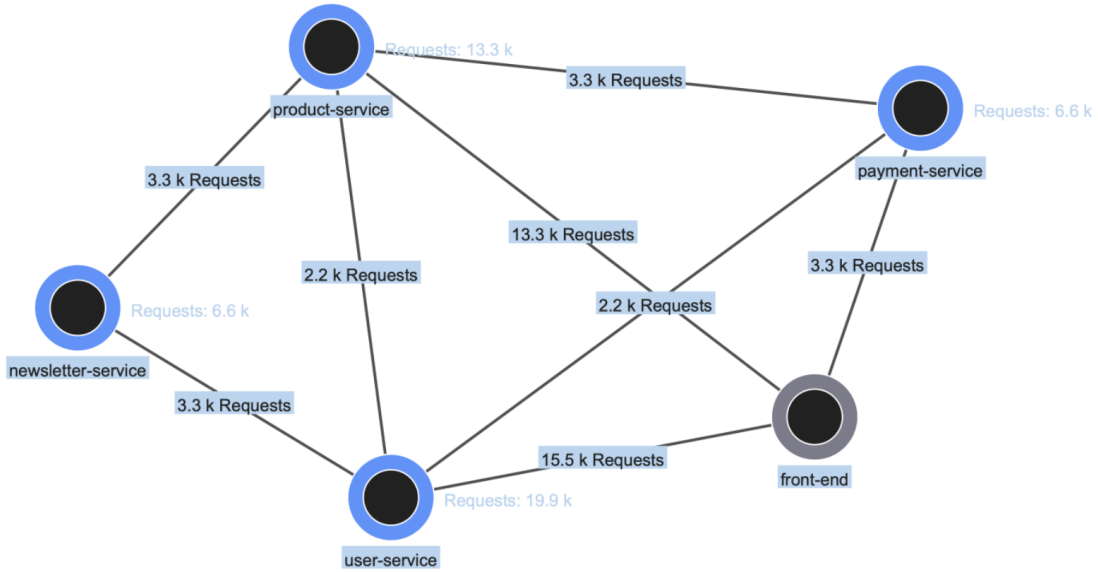
Hawk Operate

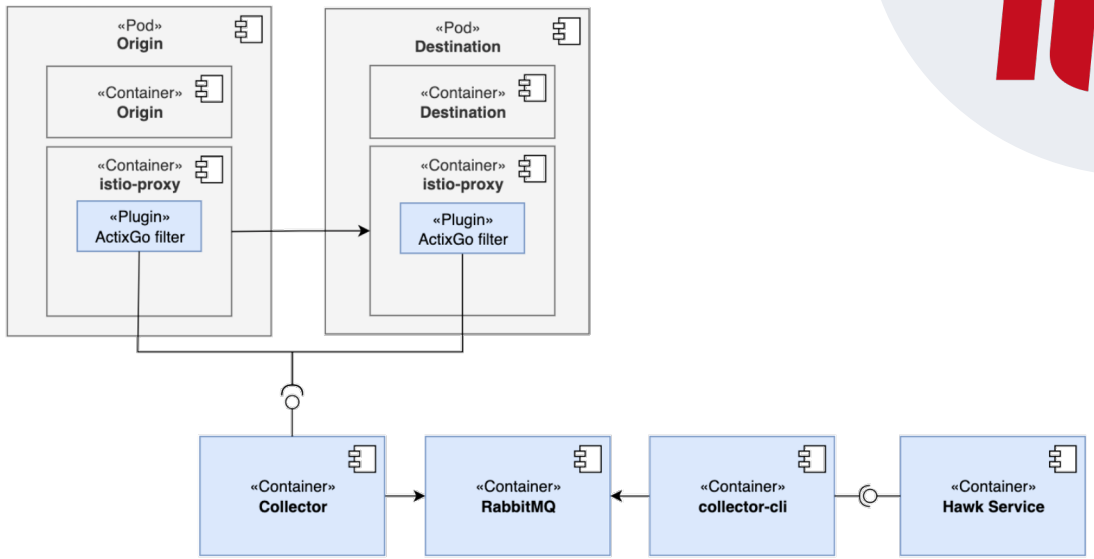
Challenges: Polyglot services, runtime information, incoming/outgoing requests, etc.

Approach: Service Mesh Extension, Collector, Message transformation

Implementation: Istio (WASM Extension), Collector (Java, RabbitMQ, Yugabyte), Java lib

Limitations: JSON over HTTP, Performance overhead, (One-time) manual labelling





Hawk Monitor

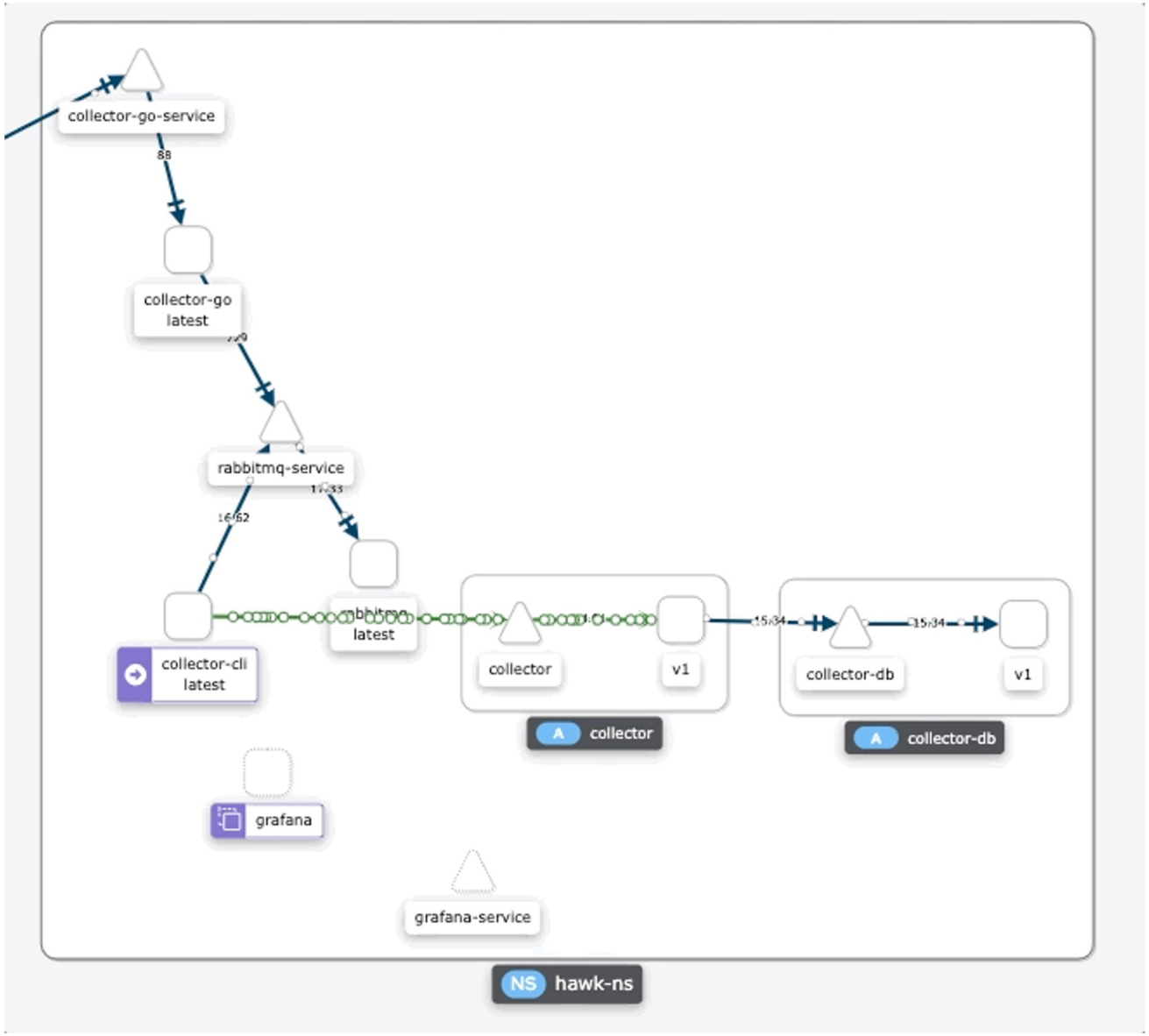
Challenges: Comprehensive transparency, records of processing activities

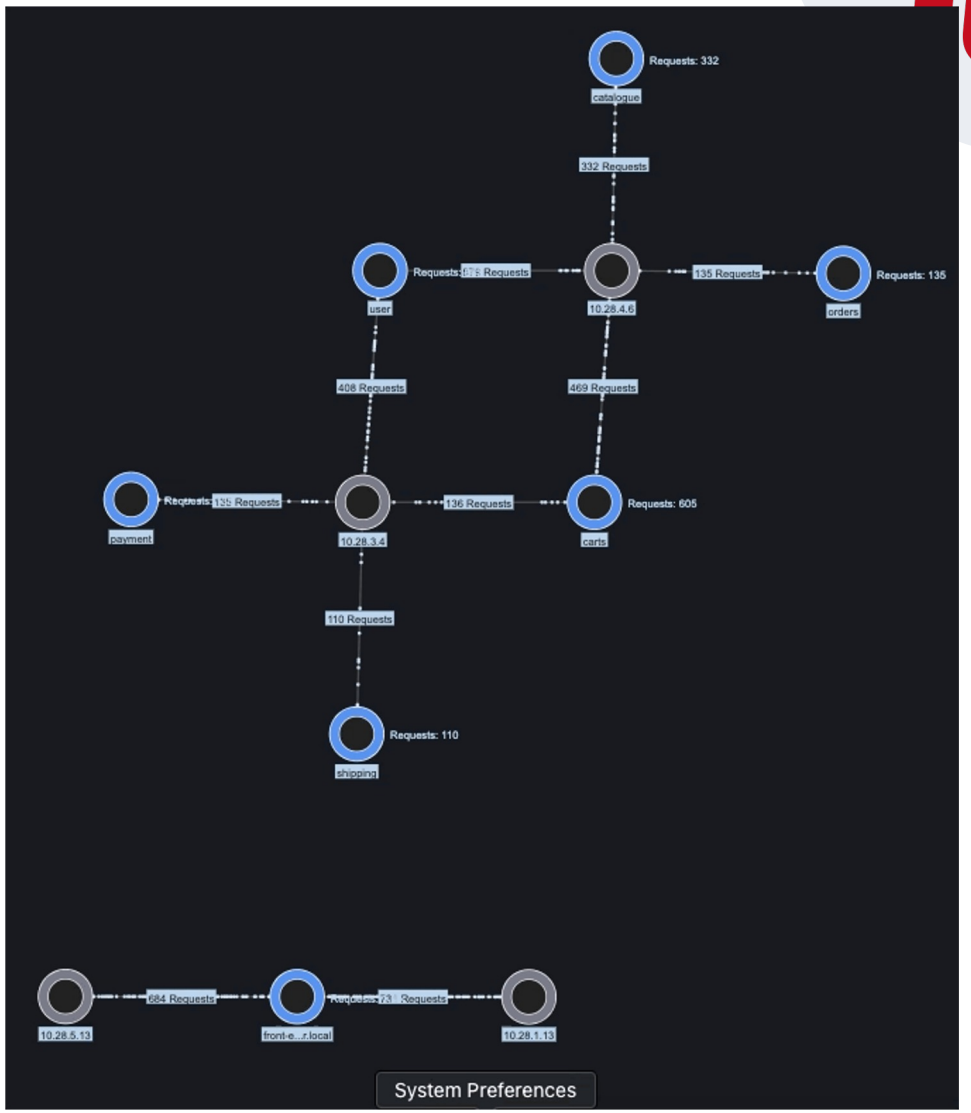
Approach: Collector + Metrics, Labelling templates, Config + reporting dashboards

Implementation: Prometheus metrics, Grafana dashboards, PromQL queries,…

Limitations: Mapping/Labelling overhead, Simple service graphs

[1] Ernst, Becker, Tai. 2019. (ICSA-C)


collector-go-service
88
collector-go
latest
rabbitmq-service
16/62
collector-cli
latest
collector
v1
collector-db
v1
A collector
A collector-db
grafana
grafana-service
NS hawk-ns


Requests: 332
catalogue
332 Requests
user
135 Requests
10.28.4.6
135 Requests
orders
Requests: 135
408 Requests
469 Requests
payment
135 Requests
10.28.3.4
136 Requests
carts
Requests: 605
110 Requests
Requests: 110
shipping
10.28.5.13
684 Requests
10.28.1.13
System Preferences

Performance overhead

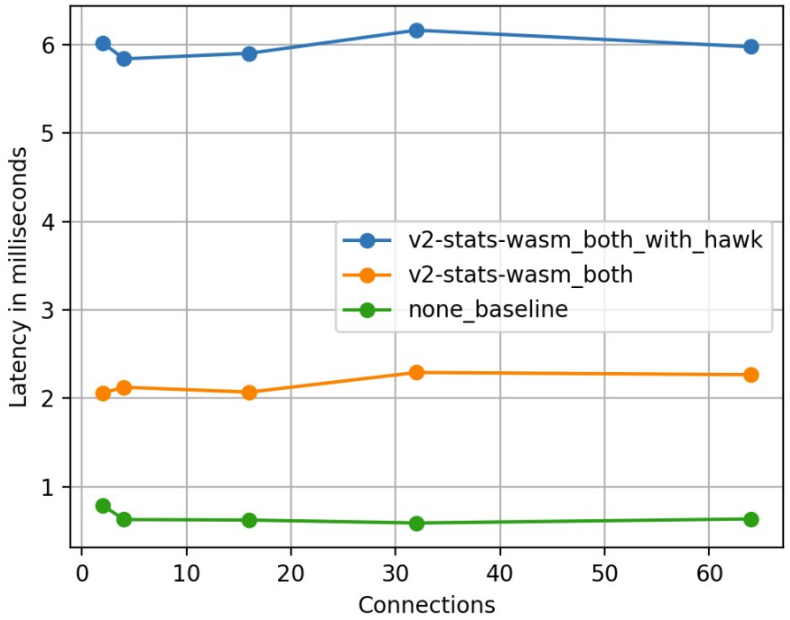


Fig. 4: Latency plots against the number of client connections. Comparison between no sidecar applied on both services (*none_baseline*), v2 telemetry stats WASM filter on both services without the *Hawk plugin* (*v2-stats-wasm_both*), and v2 telemetry stats WASM filter on both services with the *Hawk plugin* (*v2-stats-wasm_both_with_hawk*).

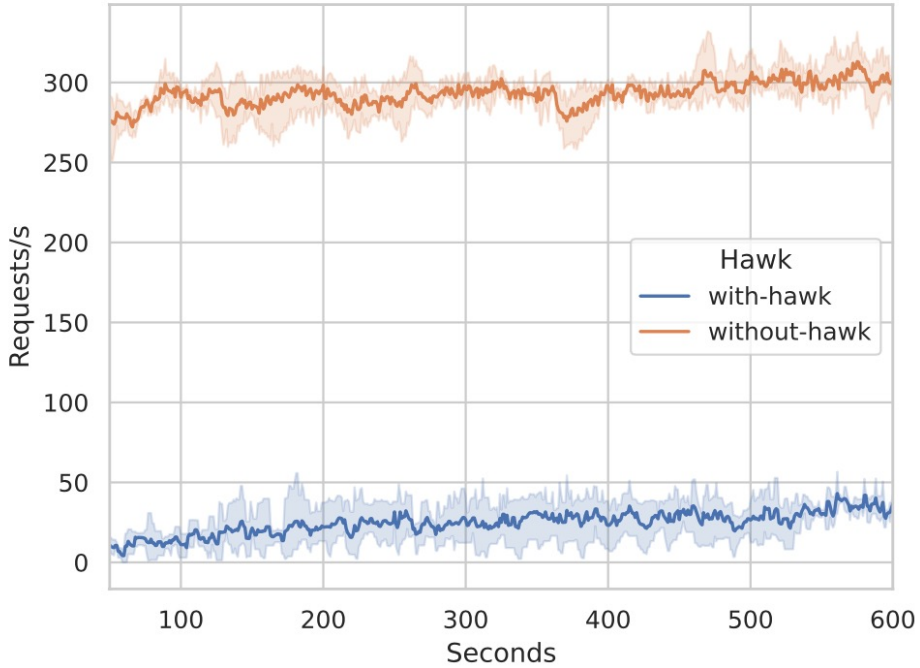
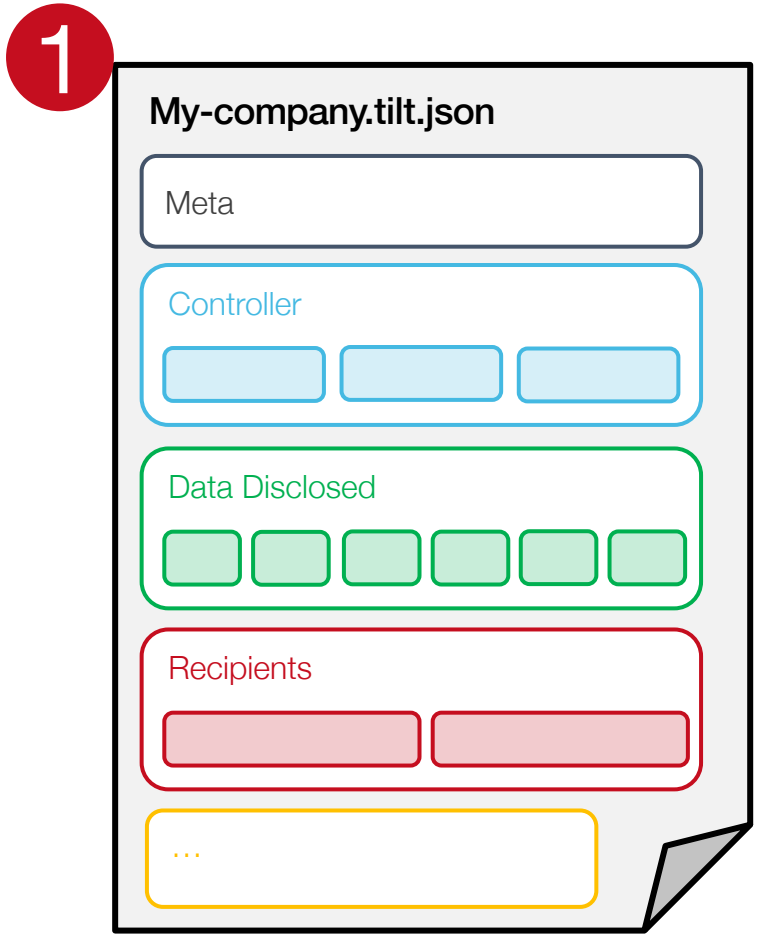


Fig. 6: Throughput for the */customers* endpoint with 100 concurrent users.

Machine-readable transparency information (e.g., TILT)

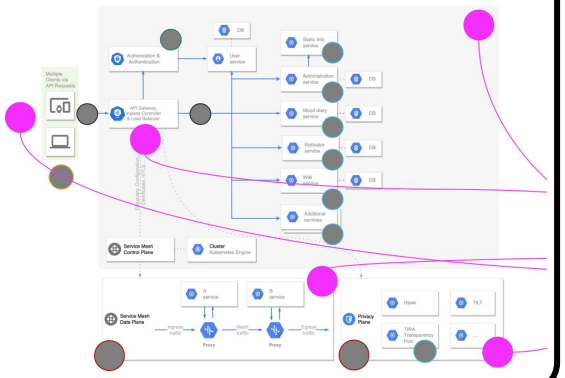


- One **structured, machine-readable format for all transparency information** according to Art. 12, 13, 14, 15, 30 GDPR
- To be integrated into **state-of-the-art developer tools**
- **Effective collaboration, communication, auditing, reporting etc.**

Elias Grünewald and Frank Pallas. 2021. TILT: A GDPR-Aligned Transparency Information Language and Toolkit for Practical Privacy Engineering. In Proceedings of the 2021 ACM Conference on Fairness, Accountabilityn, and Transparency (ACM FAccT), ACM, New York

Outlook & future work

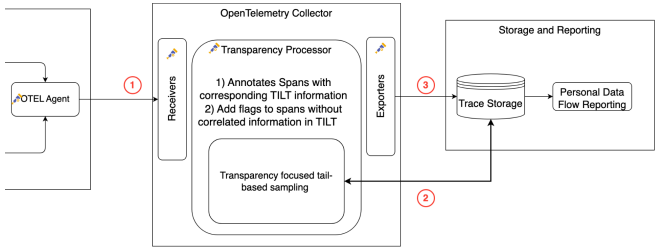
Hawk++



Privacy-aware deployments

Infrastructure transparency

Tail-based Tracing



Hawk: DevOps-driven Transparency and Accountability in Cloud Native Systems

Elias Grünewald, Jannis Kiesel, Siar-Remzi Akbayin, and Frank Pallas

eg@ise.tu-berlin.de
@eliasgruenewald
in/eliasgruenewald
tu.berlin/ise/eg

TOUCAN: **T**ransparency in Cl**ou**d-Native Ar**c**hitecture **an**d Engineering

⏩ https://tu.berlin/ise/toucan

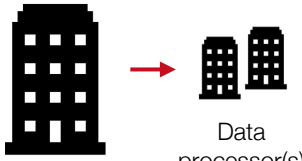
Federal Ministry of Education and Research

DLR Projektträger

Download these slides.

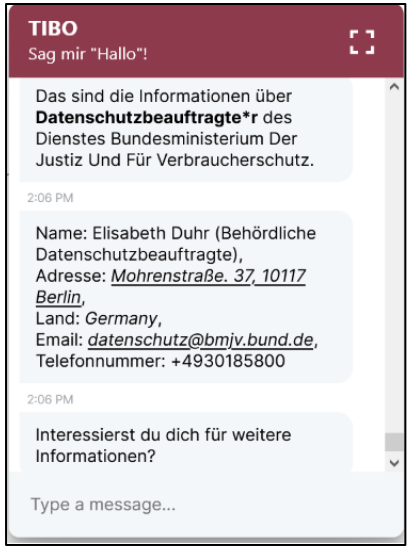
Supervisory authorities

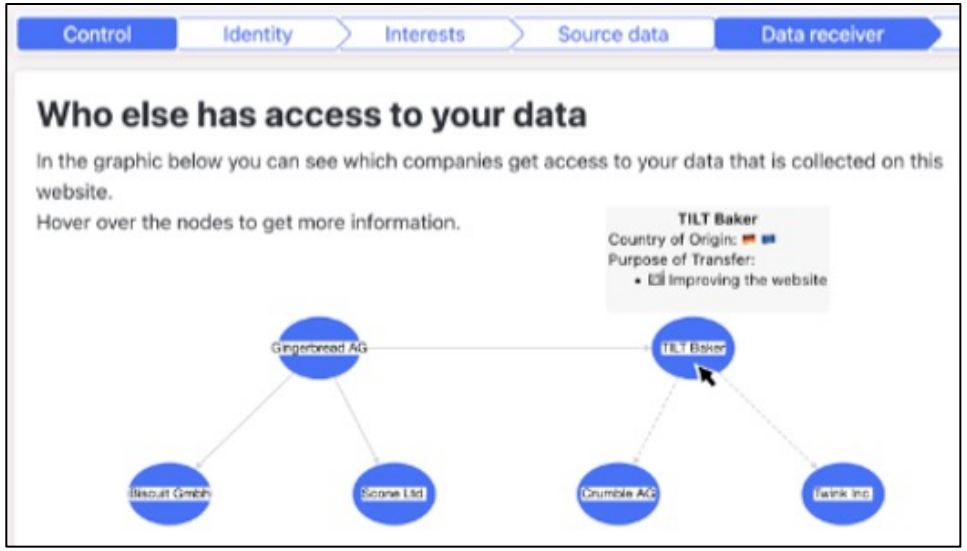
Data controllers



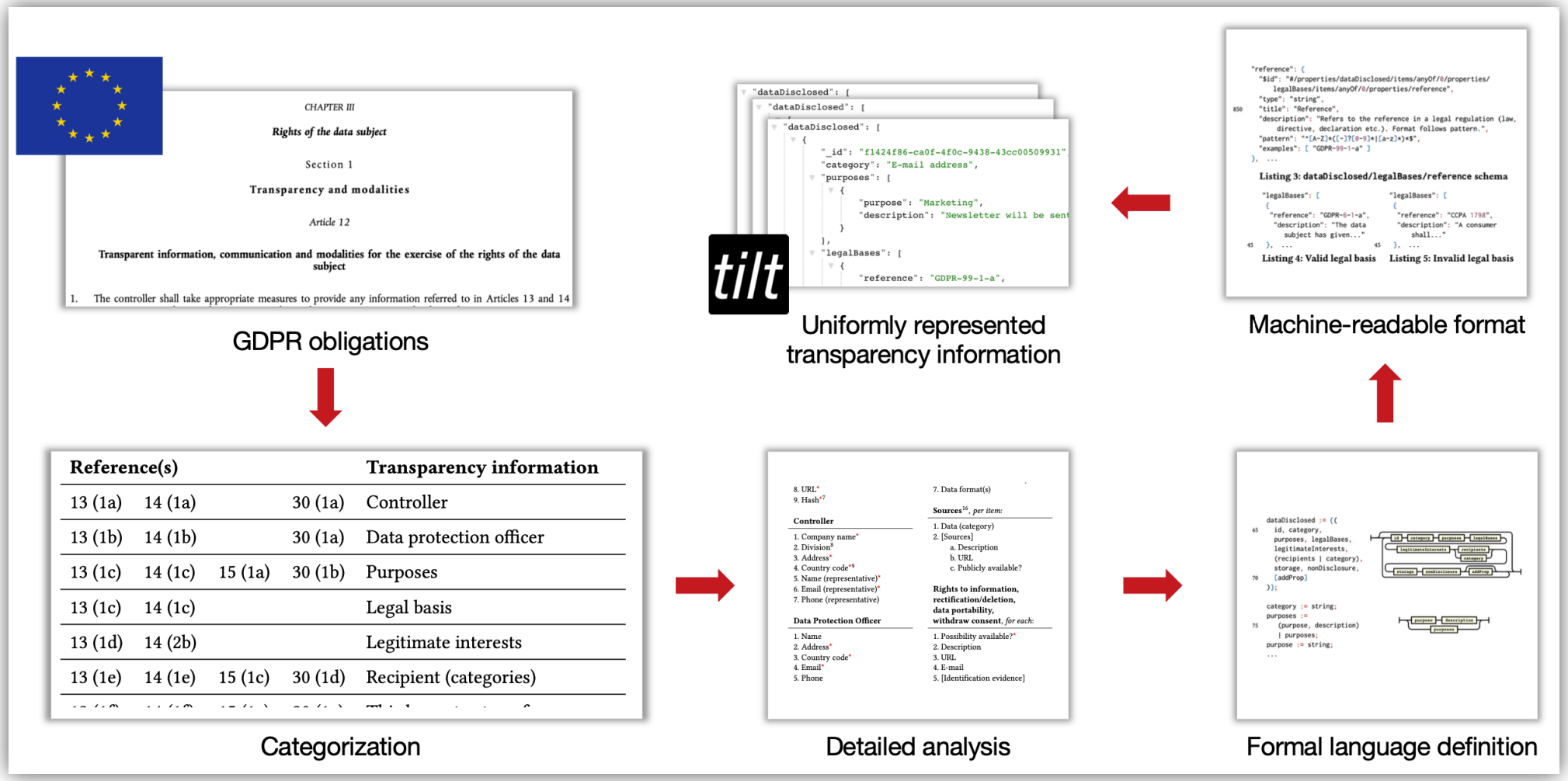
Data processor(s)

Data subjects





TILT | Transparency Information Language and Toolkit



Elias Grünewald and Frank Pallas. 2021. TILT: A GDPR-Aligned Transparency Information Language and Toolkit for Practical Privacy Engineering. In Proceedings of the 2021 ACM FAccT, ACM, New York

for

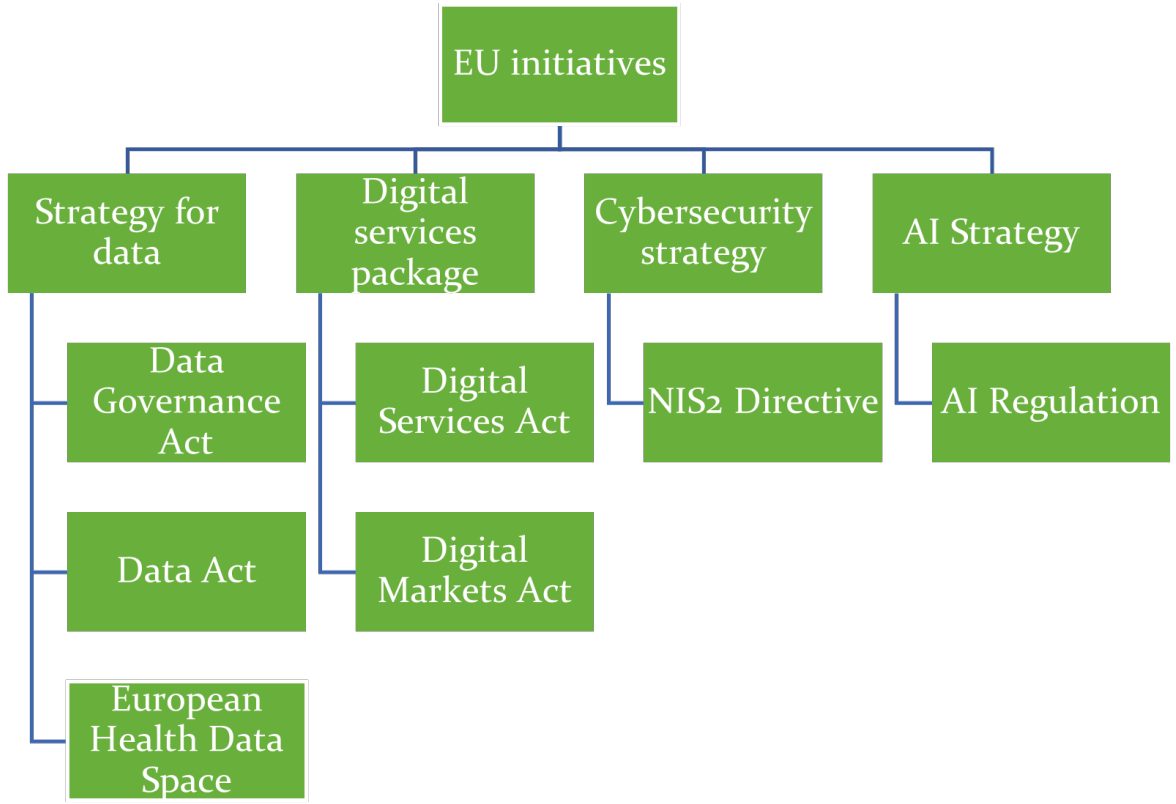
Transparency and Accountability

information and communication relating to the processing of personal data (concise, intelligible, and easily accessible) for the data subject
(e.g., Art. 5, 12 GDPR)

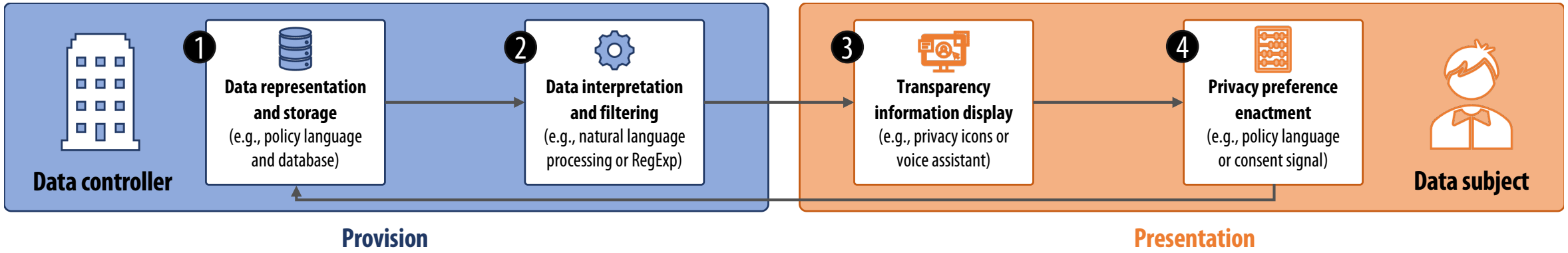
Responsibility and ability to demonstrate compliance by the controller
(e.g., Art. 5 GDPR)

Transparency information overview

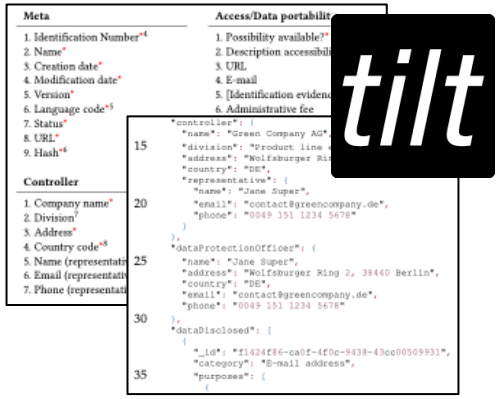
Reference(s)				Transparency information
13 (1a)	14 (1a)		30 (1a)	Controller
13 (1b)	14 (1b)		30 (1a)	Data protection officer
13 (1c)	14 (1c)	15 (1a)	30 (1b)	Purposes
13 (1c)	14 (1c)			Legal basis
13 (1d)	14 (2b)			Legitimate interests
13 (1e)	14 (1e)	15 (1c)	30 (1d)	Recipient (categories)
13 (1f)	14 (1f)	15 (1c)	30 (1e)	Third country transfer
13 (1f)	14 (1f)	15 (2)	30 (1e)	Adequacy (third country)
13 (1f)	14 (1f)	15 (2)	30 (1e)	Access and Data portability
13 (2a)	14 (2a)	15 (1d)	30 (1f)	Retention or storage criteria
13 (2b)	14 (2c)	15 (1e)		Right to request access
13 (2b)	14 (2c)	15 (1e)		Right to correction or deletion
13 (2b)	14 (2c)	15 (1e)		Right to data portability
13 (2c)	14 (2d)			Right to withdraw consent
13 (2d)	14 (2e)	15 (1f)		Right to complaint
13 (2e)				Necessity/Non-disclosure
13 (2f)	14 (2g)	15 (1h)		Automated decision making
	14 (2f)			Sources
13 (3)				Notification on purpose change
			30 (1c)	Data subjects/Data disclosed

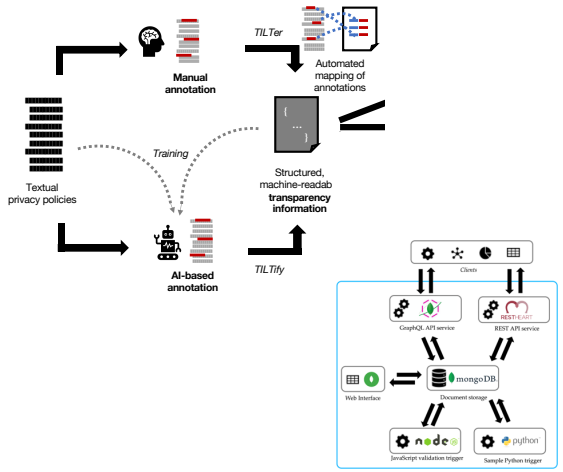


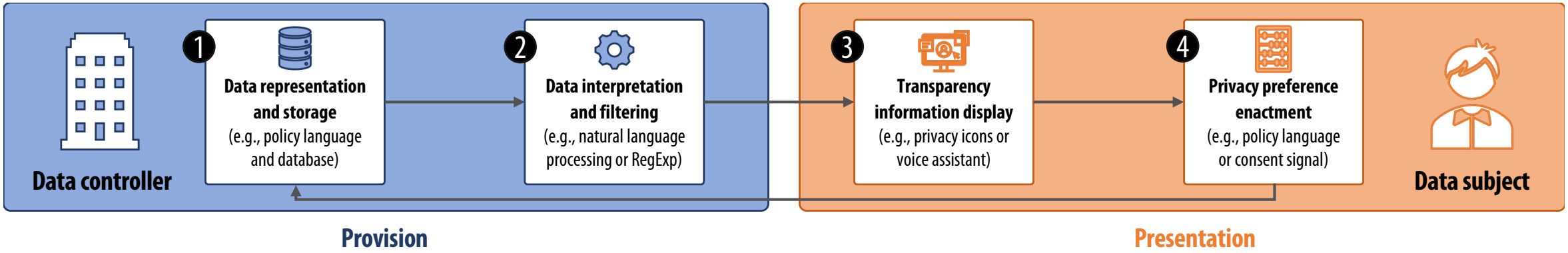
https://privacyoutloud.ro/articles/eu-legislative-initiatives-under-the-strategies-on-data-ai-cybersecurity/


Data controller
1
Data representation and storage
(e.g., policy language and database)
2
Data interpretation and filtering
(e.g., natural language processing or RegExp)
3
Transparency information display
(e.g., privacy icons or voice assistant)
4
Privacy preference enactment
(e.g., policy language or consent signal)
Data subject
Provision
Presentation

Enabling Versatile Privacy Interfaces

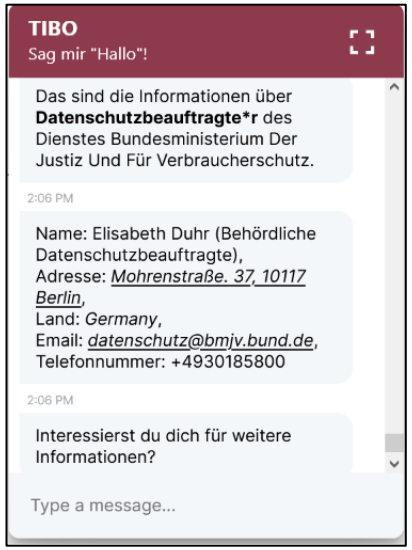


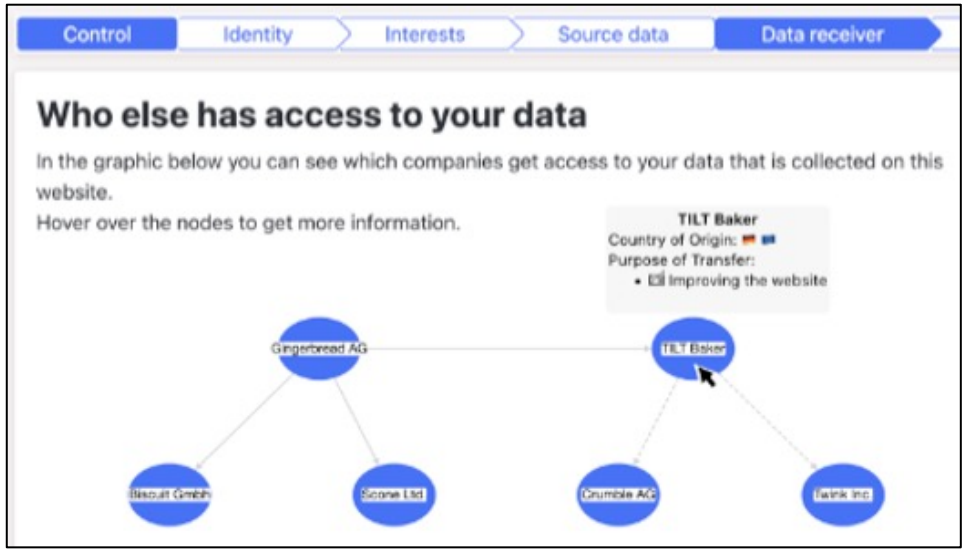


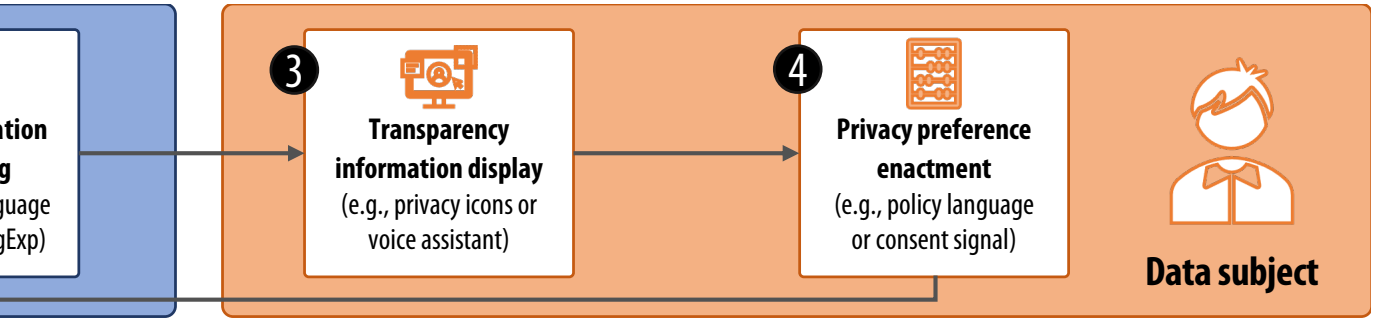


E. Grünewald, J. Halkenhäußer, N. Leschke, J. Washington, C. Paupini, and F. Pallas. “Enabling Versatile Privacy Interfaces Using Machine-Readable Transparency Information”. 2023. *To appear in*: Proceedings of the Privacy Symposium. Springer.

Enabling Versatile Privacy Interfaces







E. Grünewald, J. Halkenhäußer, N. Leschke, J. Washington, C. Paupini, and F. Pallas. “Enabling Versatile Privacy Interfaces Using Machine-Readable Transparency Information”. 2023. *To appear in*: Proceedings of the Privacy Symposium. Springer.

Operational perspectives

Supervisory authorities

Data controllers

Data processor(s)

Data subjects

E. Grünewald, J. Halkenhäußer, N. Leschke, and F. Pallas. “TAP: A Platform for Cross-Provider Analysis of Transparency Information”.

Elias Grünewald. 2022. Cloud Native Privacy Engineering through DevPrivOps. In: *Privacy and Identity Management. Between Data Protection and Security. Privacy and Identity 2021*. IFIP Advances in Information and Communication Technology, vol 644. Springer, Cham.

Elias Grünewald and Frank Pallas. 2021. TILT: A GDPR-Aligned Transparency Information Language and Toolkit for Practical Privacy Engineering. In *Proceedings of the 2021 ACM FAccT, ACM, New York*

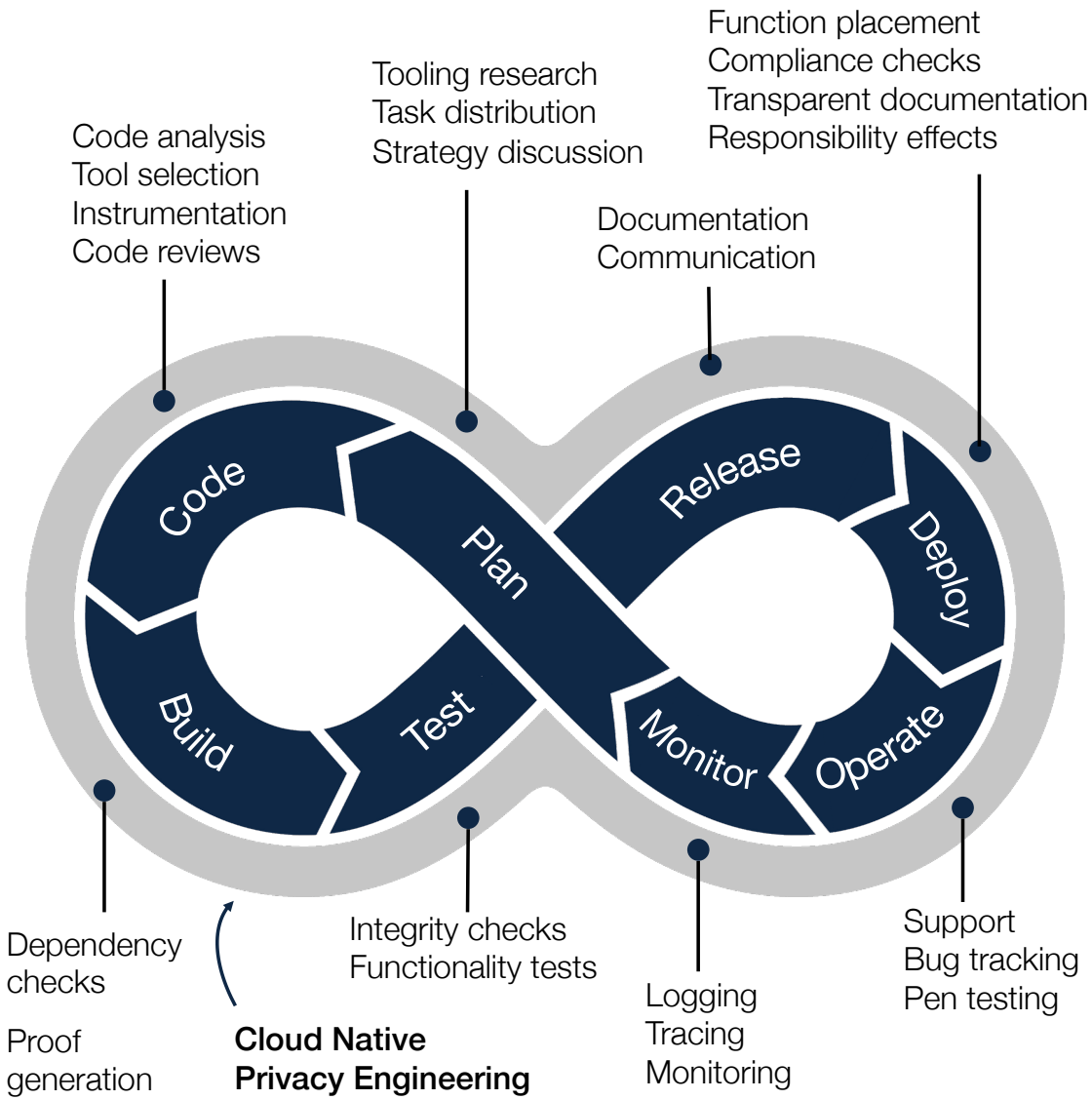
Elias Grünewald and Leonard Schurbert. 2022. Scalable Discovery and Continuous Inventory of Personal Data at Rest in Cloud Native Systems. *To appear in:* Proceedings of the International Conference on Service-Oriented Computing., Springer

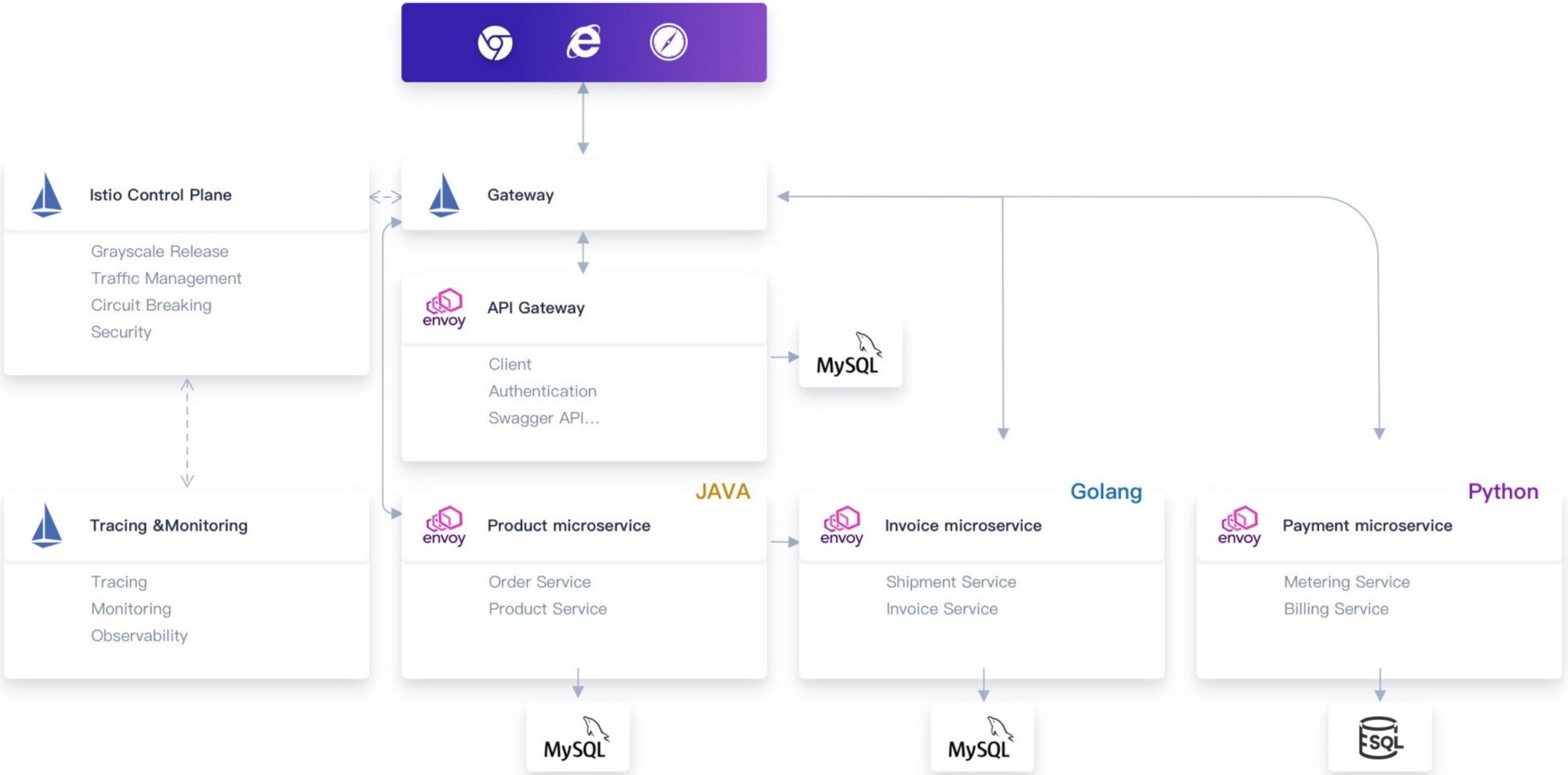
Elias Grünewald, Paul Wille, Frank Pallas, Maria C. Borges and Max-R. Ulbricht. 2021. TIRA: An OpenAPI Extension and Toolbox for GDPR Transparency in RESTful Architectures. 2021 EuroS&PW, pp. 312-31

Elias Grünewald, Jannis Kiesel, Siar-Remzi Akbayin, and Frank Pallas. 2023 (IEEE CLOUD). “Hawk: DevOps-driven Transparency and Accountability in Cloud Native Systems”.

E. Grünewald, J. Halkenhäußer, N. Leschke, J. Washington, C. Paupini, and F. Pallas. “Enabling Versatile Privacy Interfaces Using Machine-Readable Transparency Information”. 2023. To appear in: Proceedings of the Privacy Symposium.

DevPrivOps


Code analysis
Tool selection
Instrumentation
Code reviews
Tooling research
Task distribution
Strategy discussion
Function placement
Compliance checks
Transparent documentation
Responsibility effects
Documentation
Communication
Code
Plan
Release
Deploy
Build
Test
Monitor
Operate
Dependency checks
Proof generation
Integrity checks
Functionality tests
Logging
Tracing
Monitoring
Support
Bug tracking
Pen testing
Cloud Native
Privacy Engineering



https://kubesphere.io/service-mesh

ISEngineering