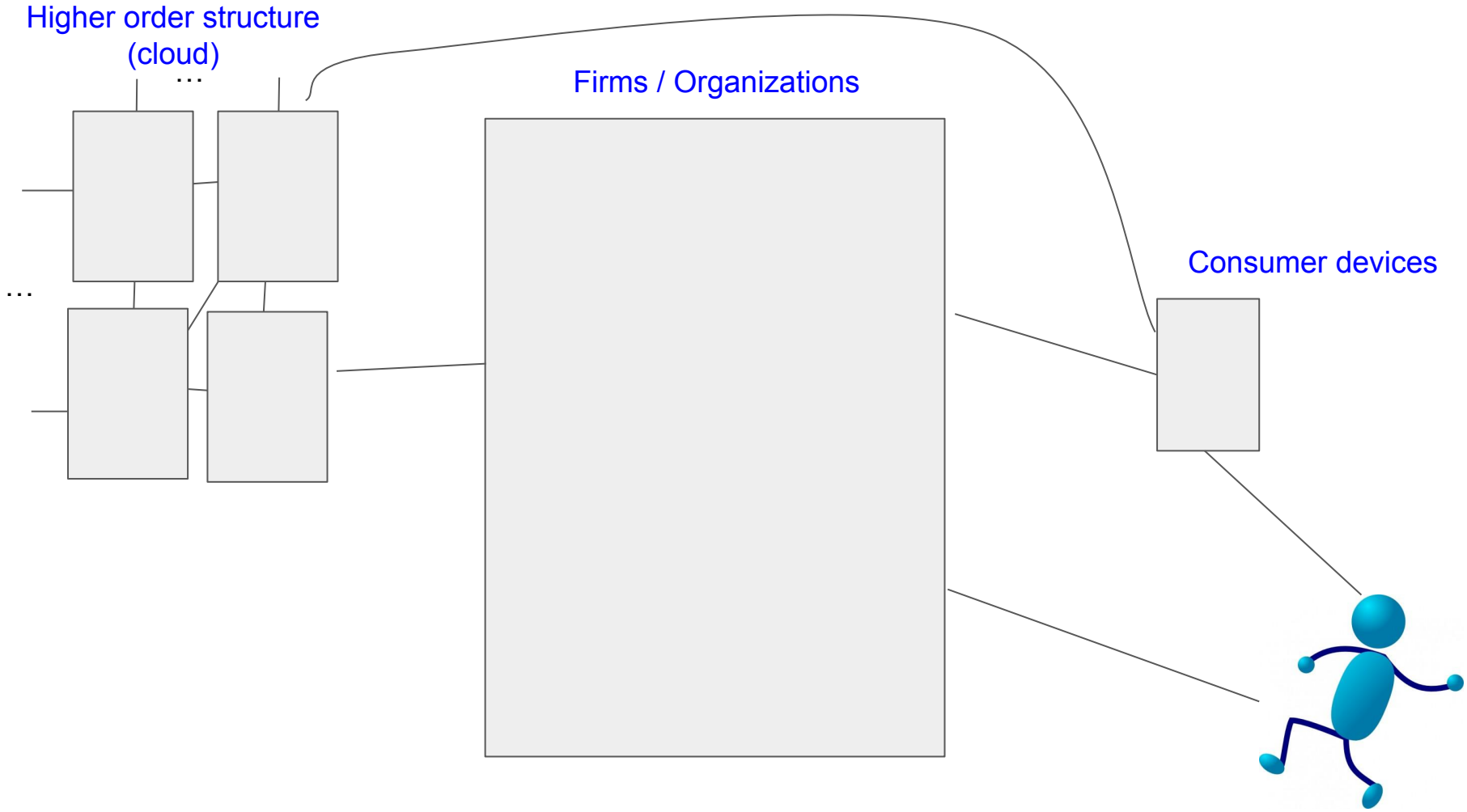
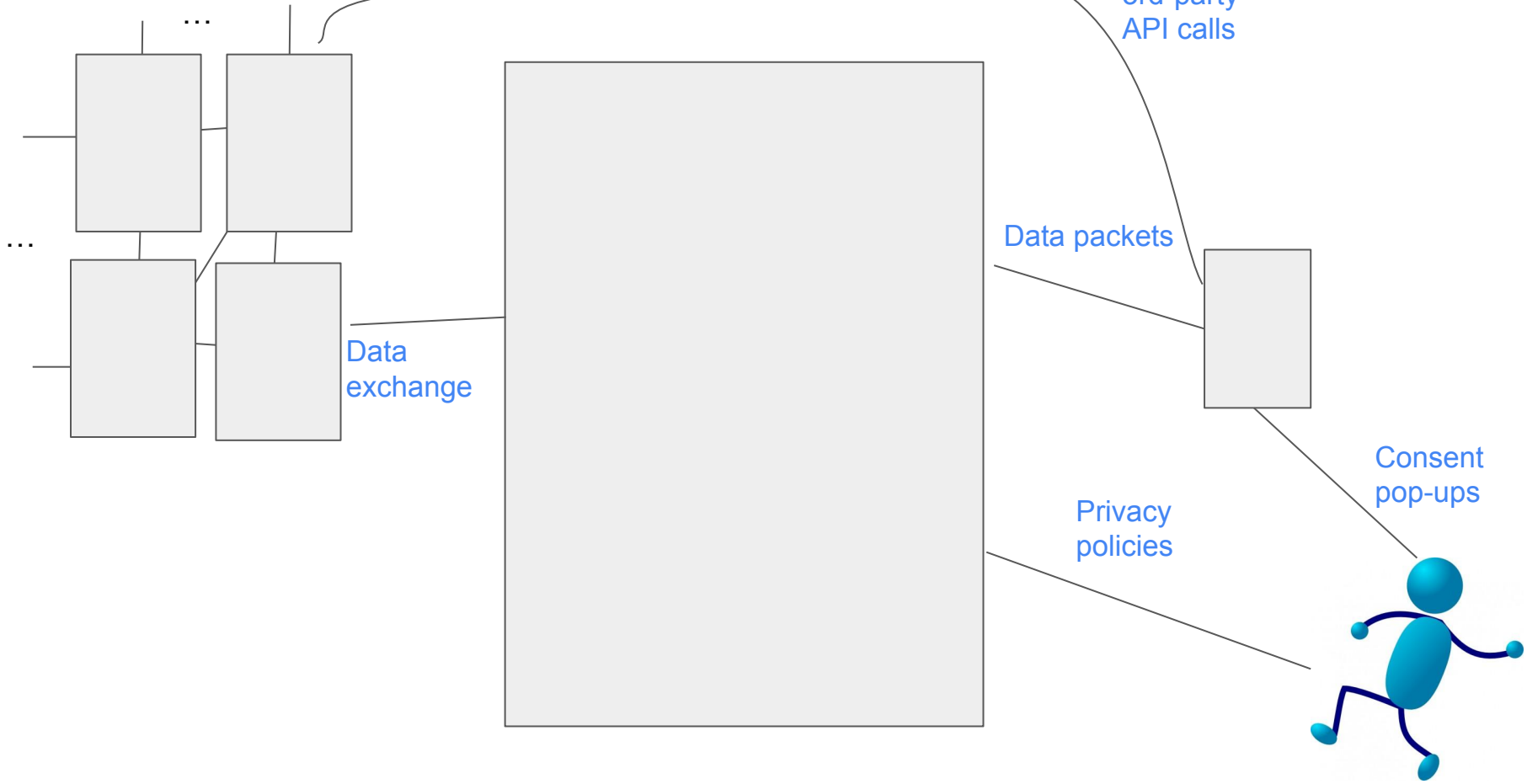
Compliance Effort

Responses to Data Protection Statutes in Open Source Development

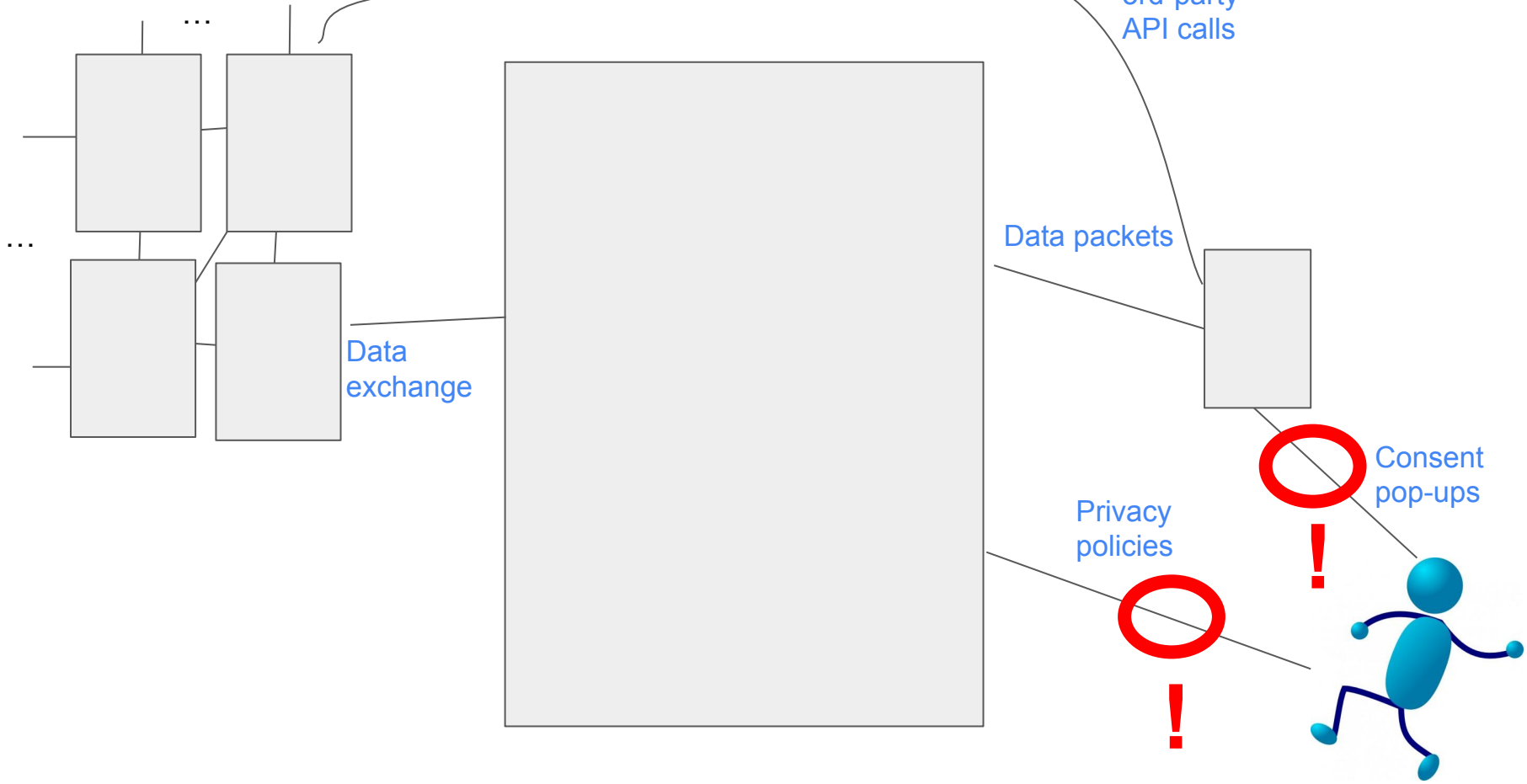
Aileen Nielsen, Karel Kubicek, Elias Datler
ETH Zurich


Higher order structure
(cloud)
Firms / Organizations
Consumer devices

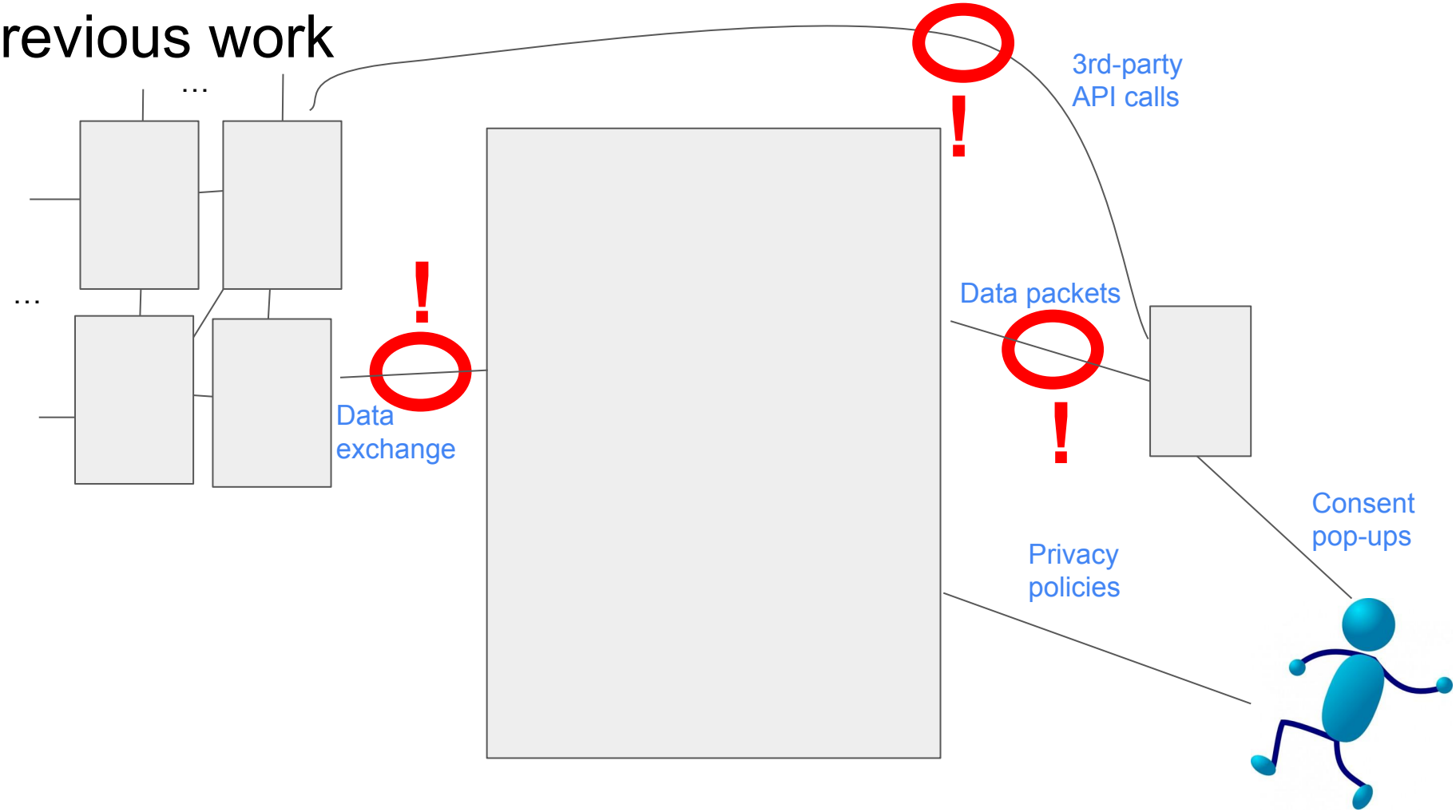
Previous work



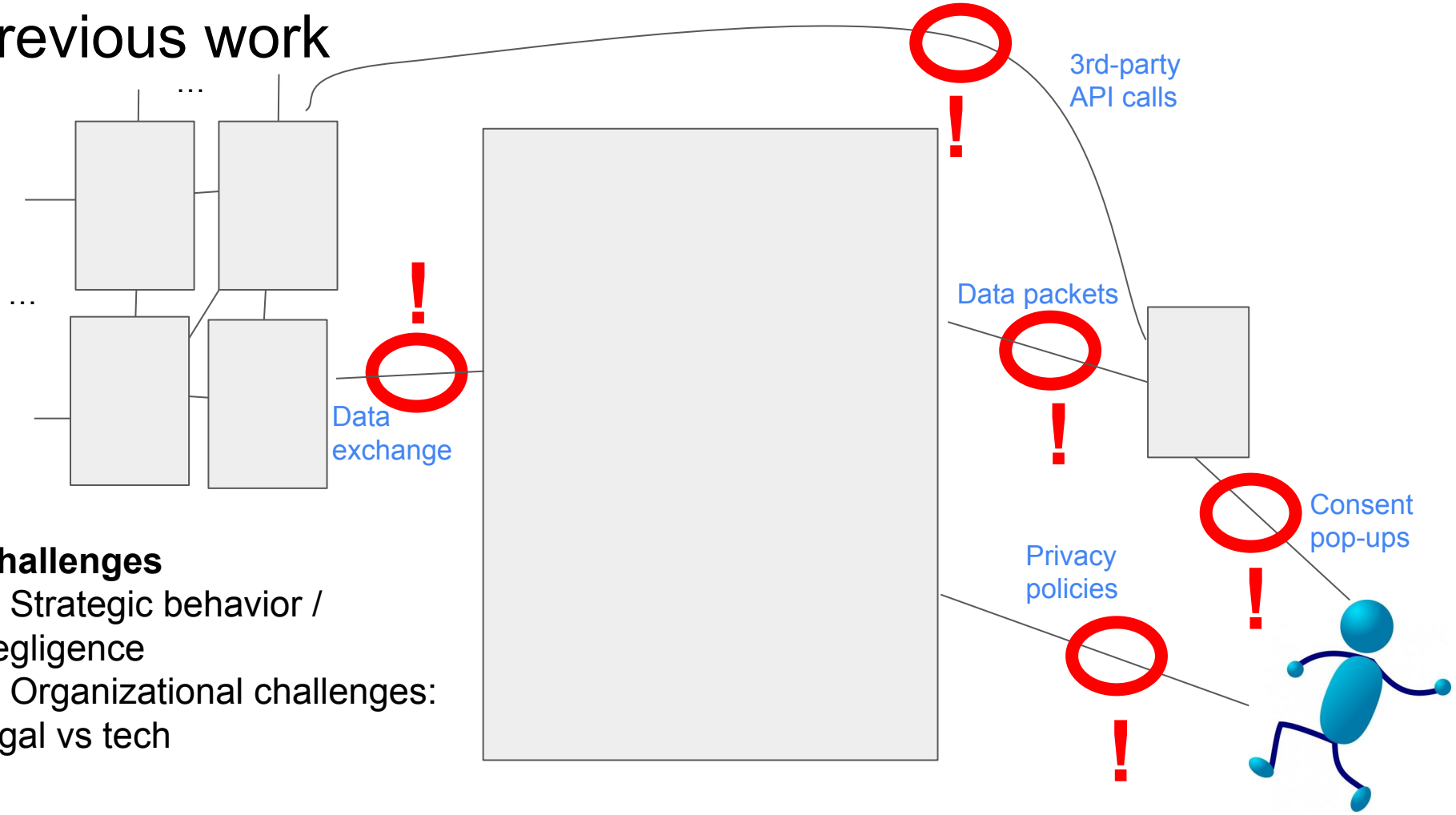
Previous work



Previous work



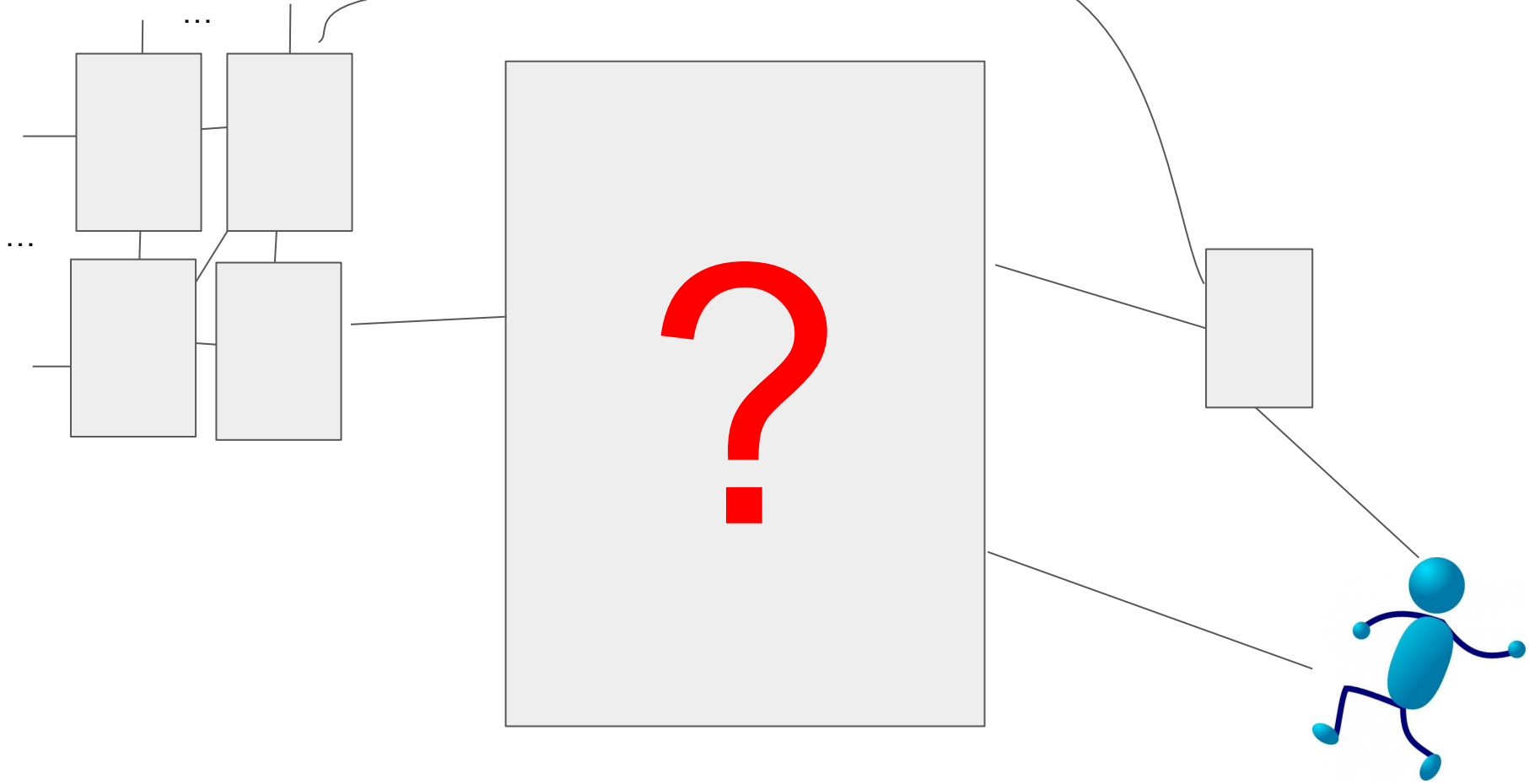
Previous work



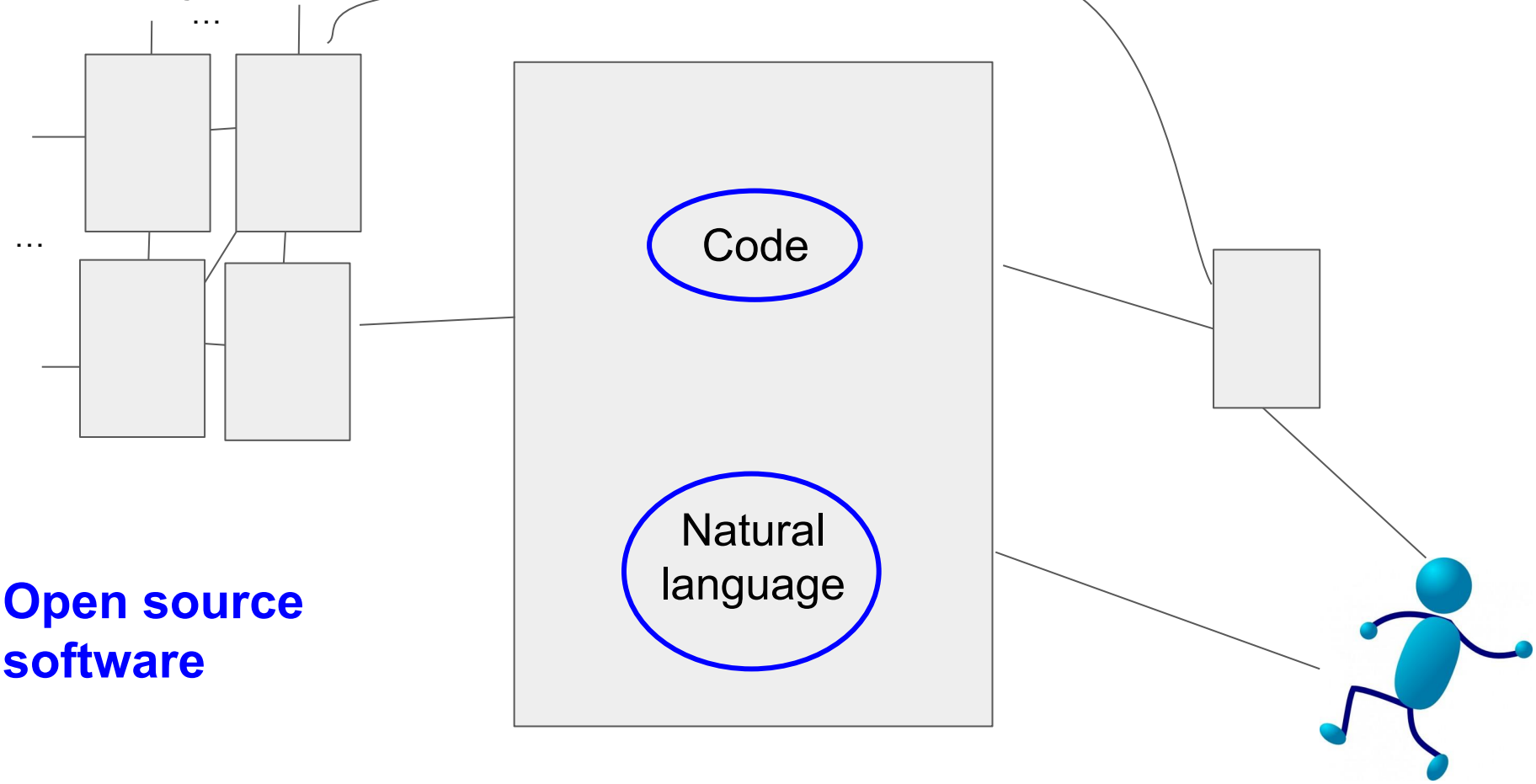
Challenges

1. Strategic behavior / negligence
2. Organizational challenges: legal vs tech

This project

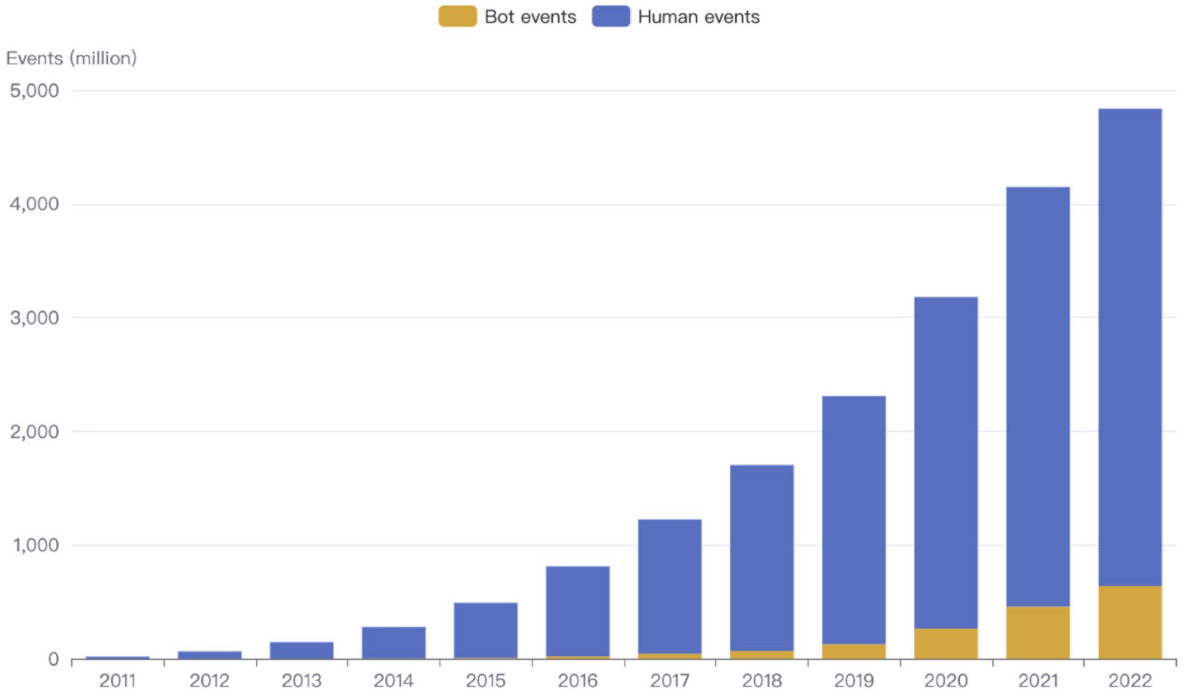


This project



Open source software

Open source software



https://ossinsight.io/blog/github-data-is-booming/

Repository

Open city profile

Summary

Open city profile is used to store common information (name, contact information, ...) about the citizens of the city of Helsinki.

When a citizen is using a service which is connected to the profile, the service can query for the citizen's information from the profile so that the citizen doesn't have to enter all of their data every time it is needed. The services may also provide a better user experience using the profile's data, for example by returning more relevant search results based on the citizen's interests.

The same data may also be queried by the employees of the city of Helsinki while performing their daily duties, for example using the administrative functions of services.

Open city profile is implemented using Django and it provides a GraphQL API.

Open city profile

Continuous integration passing codecov 91%

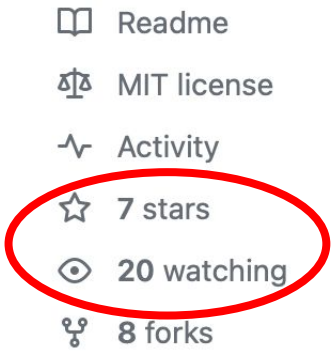
Summary

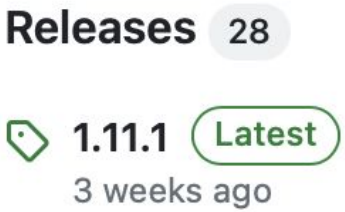
Open city profile is used to store common information (name, contact information, ...) about the citizens of the city of Helsinki.

When a citizen is using a service which is connected to the profile, the service can query for the citizen's information from the profile so that the citizen doesn't have to enter all of their data every time it is needed. The services may also provide a better user experience using the profile's data, for example by returning more relevant search results based on the citizen's interests.

The same data may also be queried by the employees of the city of Helsinki while performing their daily duties, for example using the administrative functions of services.

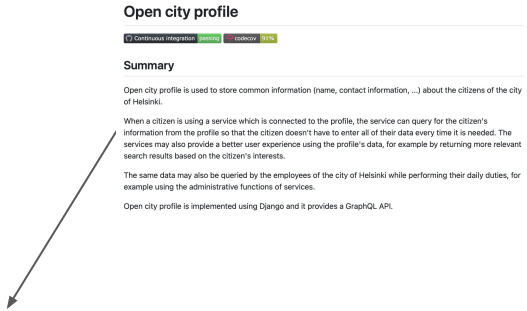
Open city profile is implemented using Django and it provides a GraphQL API.





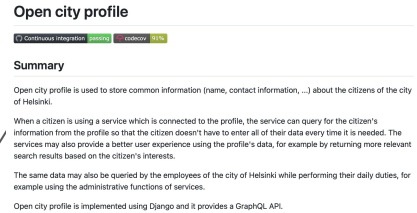
Issues/commits

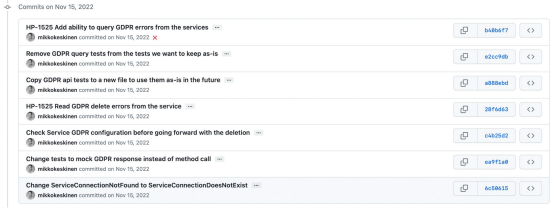
Code +
Natural Language



Commits on Nov 15, 2022

HP-1525 Add ability to query GDPR errors from the services ··· **mikkokeskinen** committed on Nov 15, 2022 ✕	b40b6f7
Remove GDPR query tests from the tests we want to keep as-is ··· **mikkokeskinen** committed on Nov 15, 2022	e2cc9db
Copy GDPR api tests to a new file to use them as-is in the future ··· **mikkokeskinen** committed on Nov 15, 2022	a888ebd
HP-1525 Read GDPR delete errors from the service ··· **mikkokeskinen** committed on Nov 15, 2022	28f6d63
Check Service GDPR configuration before going forward with the deletion ··· **mikkokeskinen** committed on Nov 15, 2022	c4b25d2
Change tests to mock GDPR response instead of method call ··· **mikkokeskinen** committed on Nov 15, 2022	ea9f1a0
Change ServiceConnectionNotFound to ServiceConnectionDoesNotExist ··· **mikkokeskinen** committed on Nov 15, 2022	6c50615





Example implementation of **right to data deletion**

```
1420          -       if not dry_run:
      1485    +       errors = [error for result in results for error in result.errors]
      1486    +       if not dry_run and not errors:
1421  1487                    delete_profile_from_keycloak(profile)
1422  1488                    profile.delete()
1423  1489                    info.context.user.delete()
1424  1490
```

Examples

Self-regulation

- A national government agency identified nonconsensual Google Analytics tracking in an open source repo in 2020

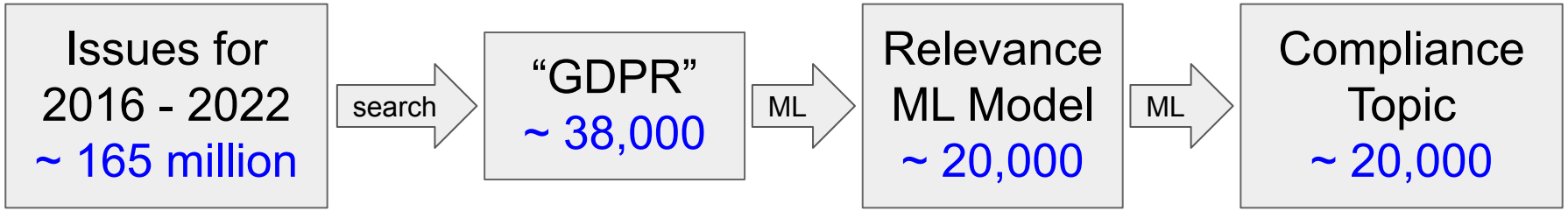
Statutory interpretation and divergent implementations

- Data deletion implemented as removal of data
- Data deletion implemented as anonymization of data

Loss of functionality

- External 3rd-party service was discontinued due to volume of GDPR takedown requests

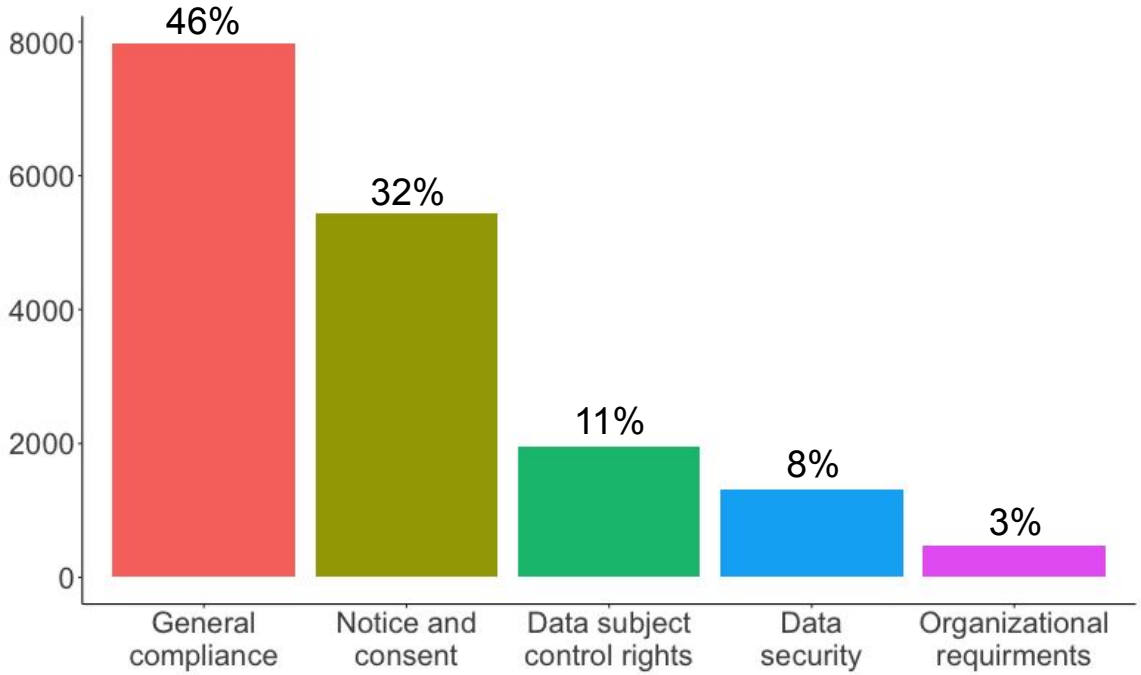
Machine learning for automated labeling



Focus of developers

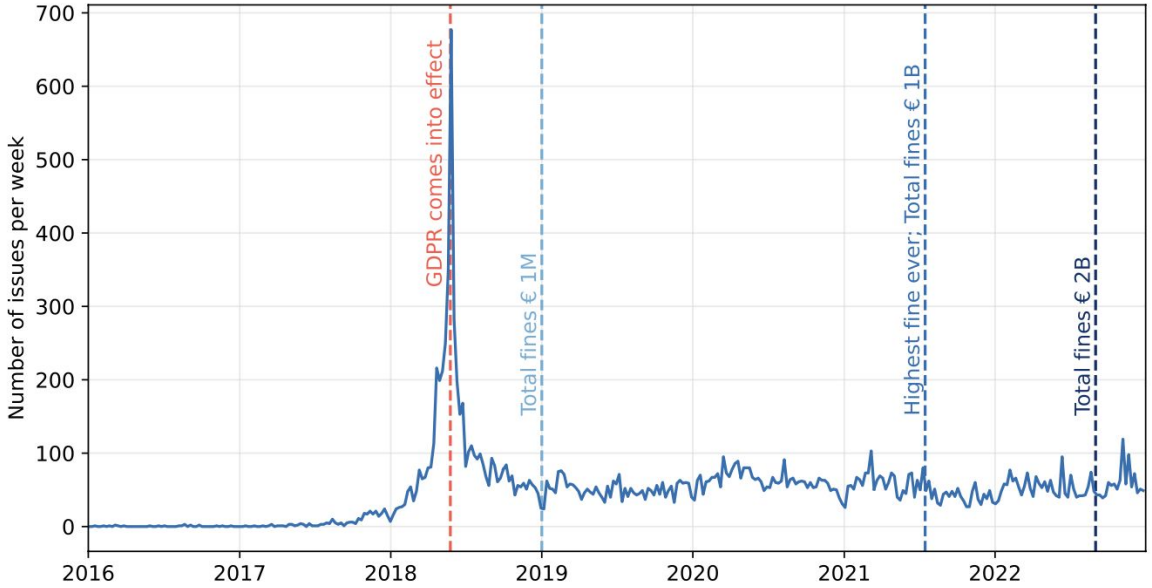
- Focus of many issues is unclear even to a human coder
  General compliance

- Largest topic of interest is one most visible
  Notice and consent

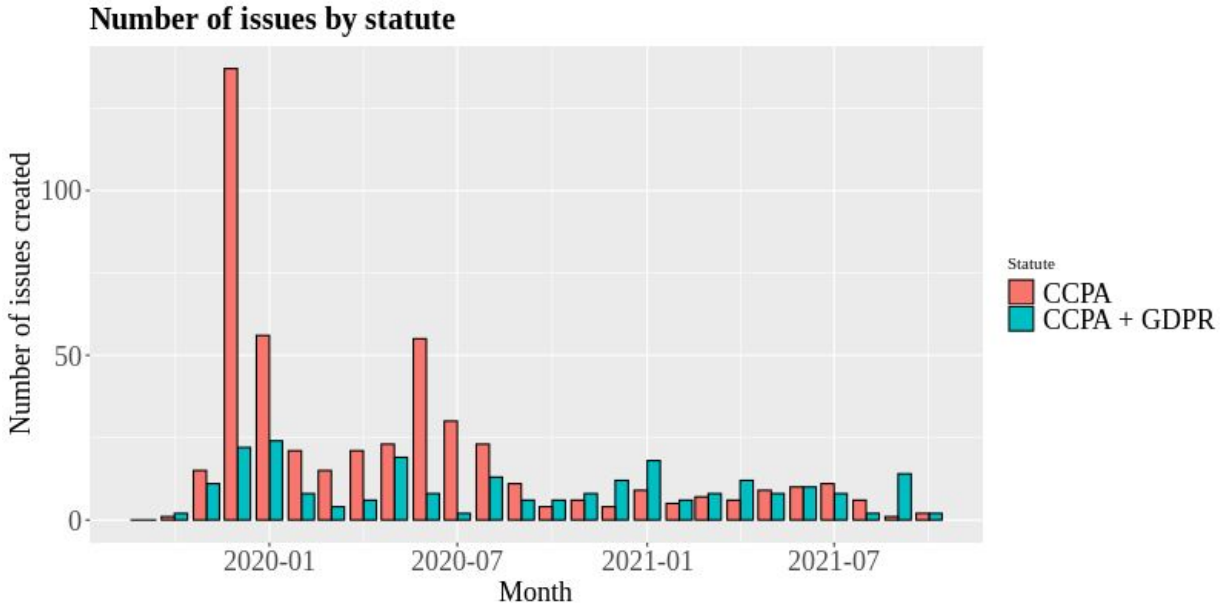


Impact of statutory dates and enforcement actions

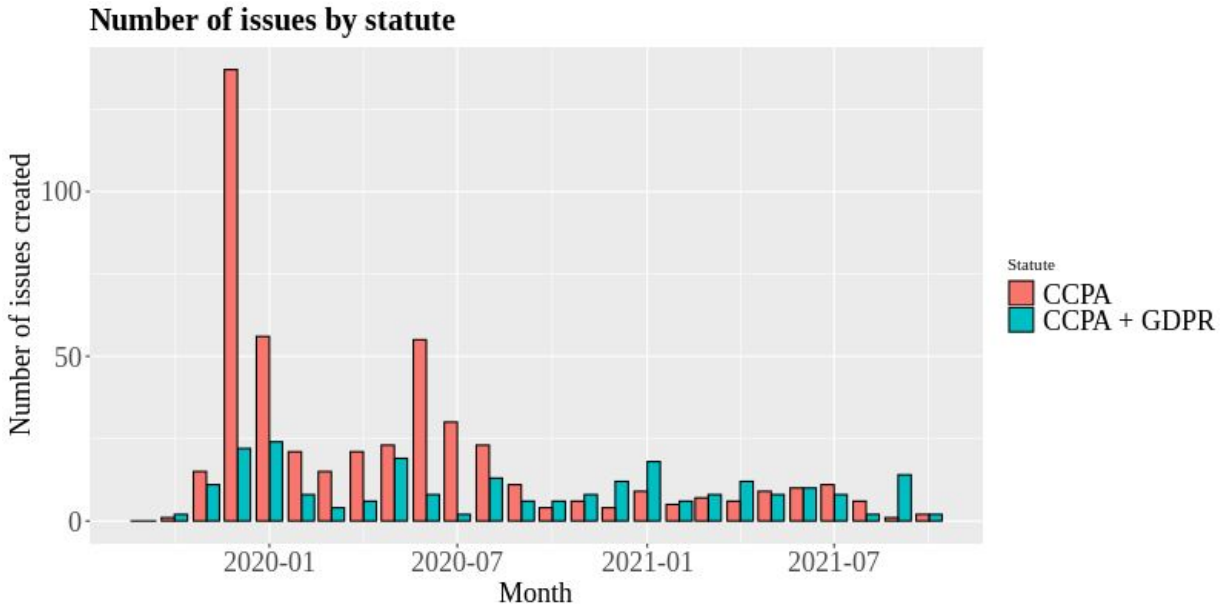
- Largest burst of activity came around May 2018
- No effect of significant enforcement actions



Looking to California's 2020 experience with CCPA



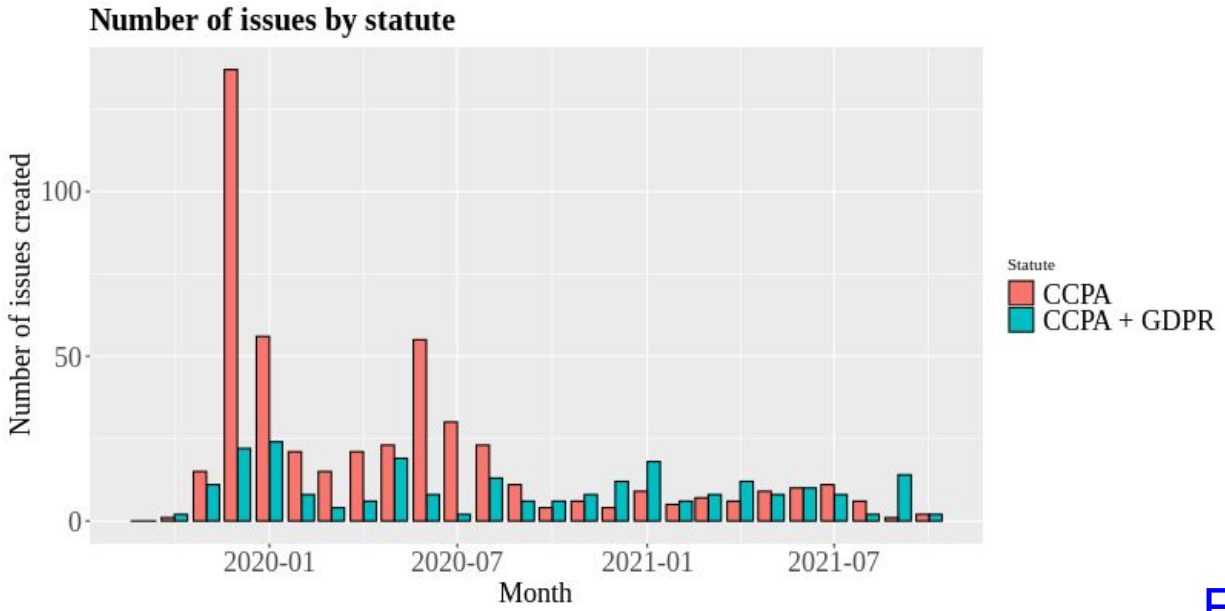
A nuanced Brussels effect



Merging of CCPA with GDPR compliance

```
- self.gdprConsent
+ self.privacyConsent
```

A nuanced Brussels effect



Merging of CCPA with GDPR compliance

```
- self.gdprConsent
+ self.privacyConsent
```

Future work: ***How*** is the Brussels effect rolling out?

Policy insights

- Interpretation and divergent implementations
- Emphasis on visible compliance
- Mixed effects of statutory dates and enforcement actions
- Nuanced Brussels effect

Authors: Aileen Nielsen, Karel Kubicek, Elias Datler