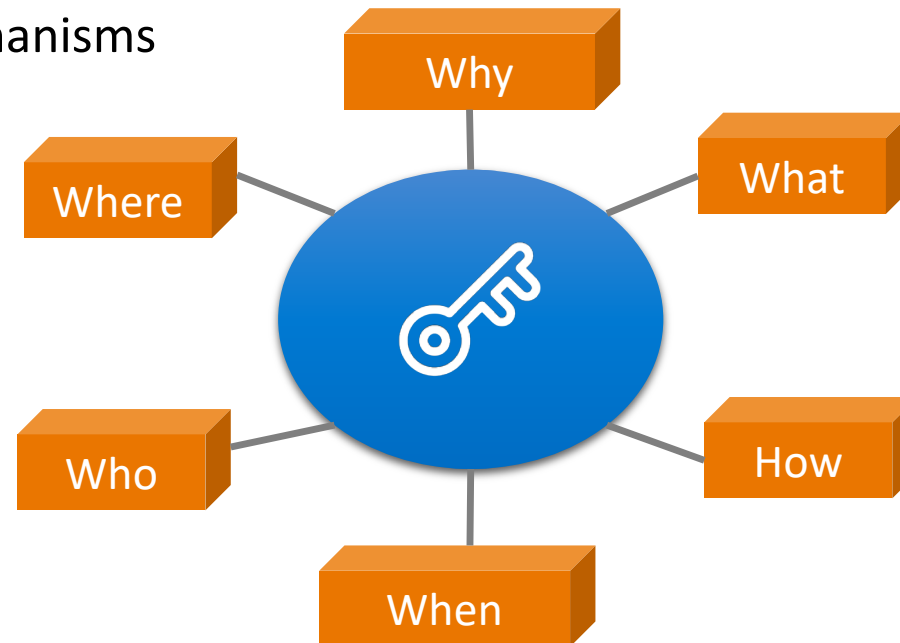


SURF DATA EXCHANGE: MANAGING THE WHO, WHAT, HOW, WHY, WHEN AND WHERE OF DATA ACCESS



The challenge: Data-sharing without sharing data

- There is a huge volume of valuable data which is not accessible for research purposes (because of e.g. privacy of commercial interests)
- Parties managing such data are often amenable – or even have an interest – to making such data accessible for research
- However, “just sharing” the data is not an option; appropriate conditions and control mechanisms are required
- Existing solutions are expensive, not scalable and/or carry risk



Some examples



Airliners with flight data: safety ↔ commercial interests



Patient organizations with wearables data: academic ↔ commercial interests



Health insurance with data re: dentist visits: public health ↔ privacy

Methods to Support Data Sharing

Agreements

- Stipulation of what can/cannot be done
- Signing of contract or NDA
- Dispute resolution process

Registration

- Authentication
- Verification of credential
- Reputation score
- Policy framework
- Audit trails

Pseudonymization

- Filtering
- Pruning
- Aggregation
- Coarse-grained buckets
- Alteration of data
- One-way hashing
- One-time identifiers

Data Vault

- Data source retains control
- Delegate permissions
- No central data lake
- Data marketplace

Secure Containers

- Bring algorithm to data
- Trusted third party
- Share output instead of data

Secure Computing

- Secure multi-party computation
- Homomorphic encryption
- Garbled Circuits
- Zero-knowledge proof

SURF Data Exchange (SURF Open Innovation Lab)



- **Goal: Exploring the dynamics and mechanics around sharing sensitive data for research**
 - What are the main barriers and concerns for data providers to share their data?
 - Can those concerns be addressed through technology and/or policy?
 - Can that be done in such a way that ‘it works’ for the researcher as data consumer?
- **How?**
 - Community engagement through e.g. webinars
 - Developing a functional prototype of a Data Exchange platform which **starts with the researcher** and places **data sovereignty** first.

Webinars

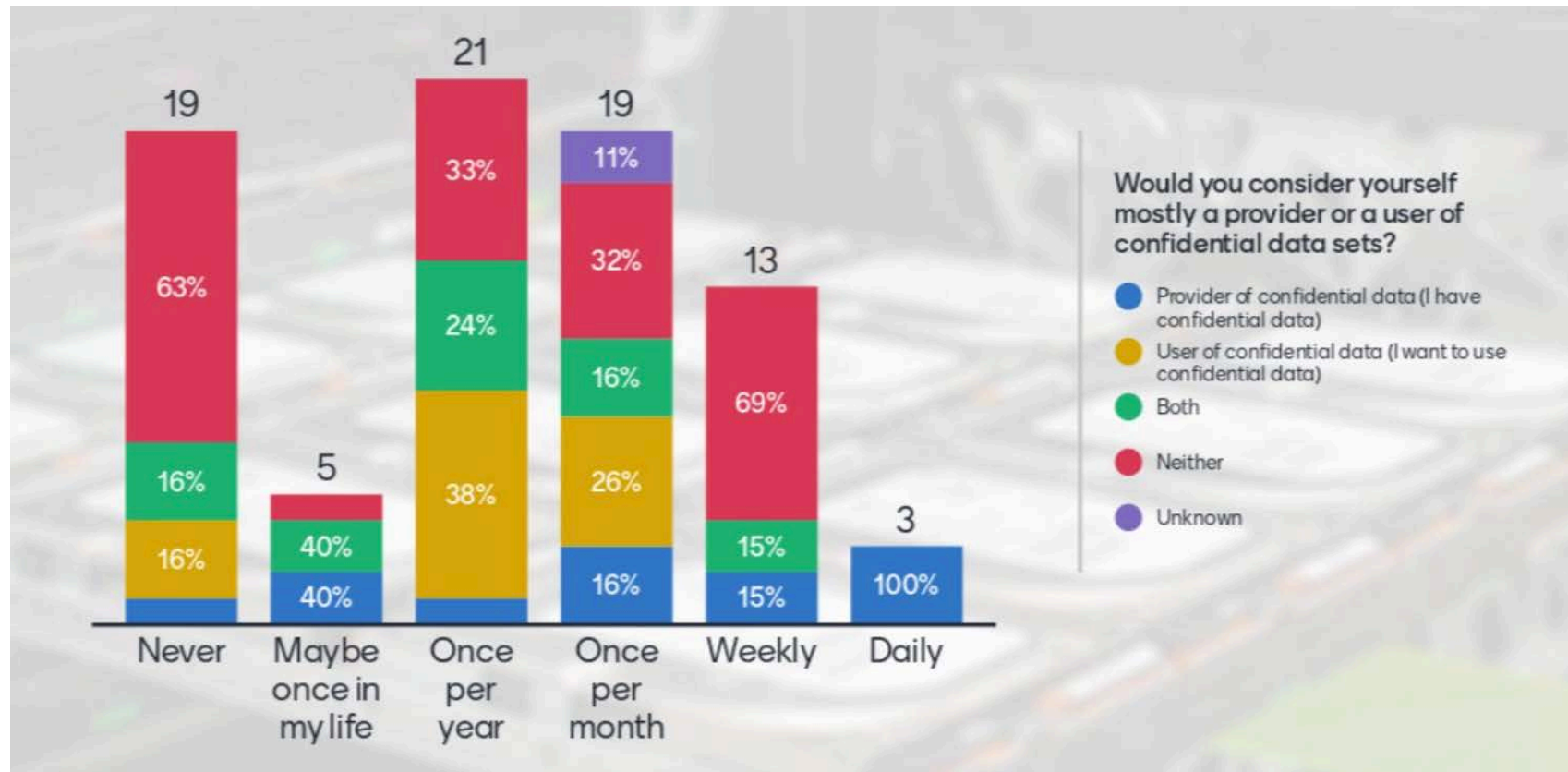
- *23 June 2020*: Sharing confidential data for research
- *22 Sept 2020*: Data sharing with privacy protection for research



<https://www.surf.nl/en/data-exchange-trusted-data-sharing>

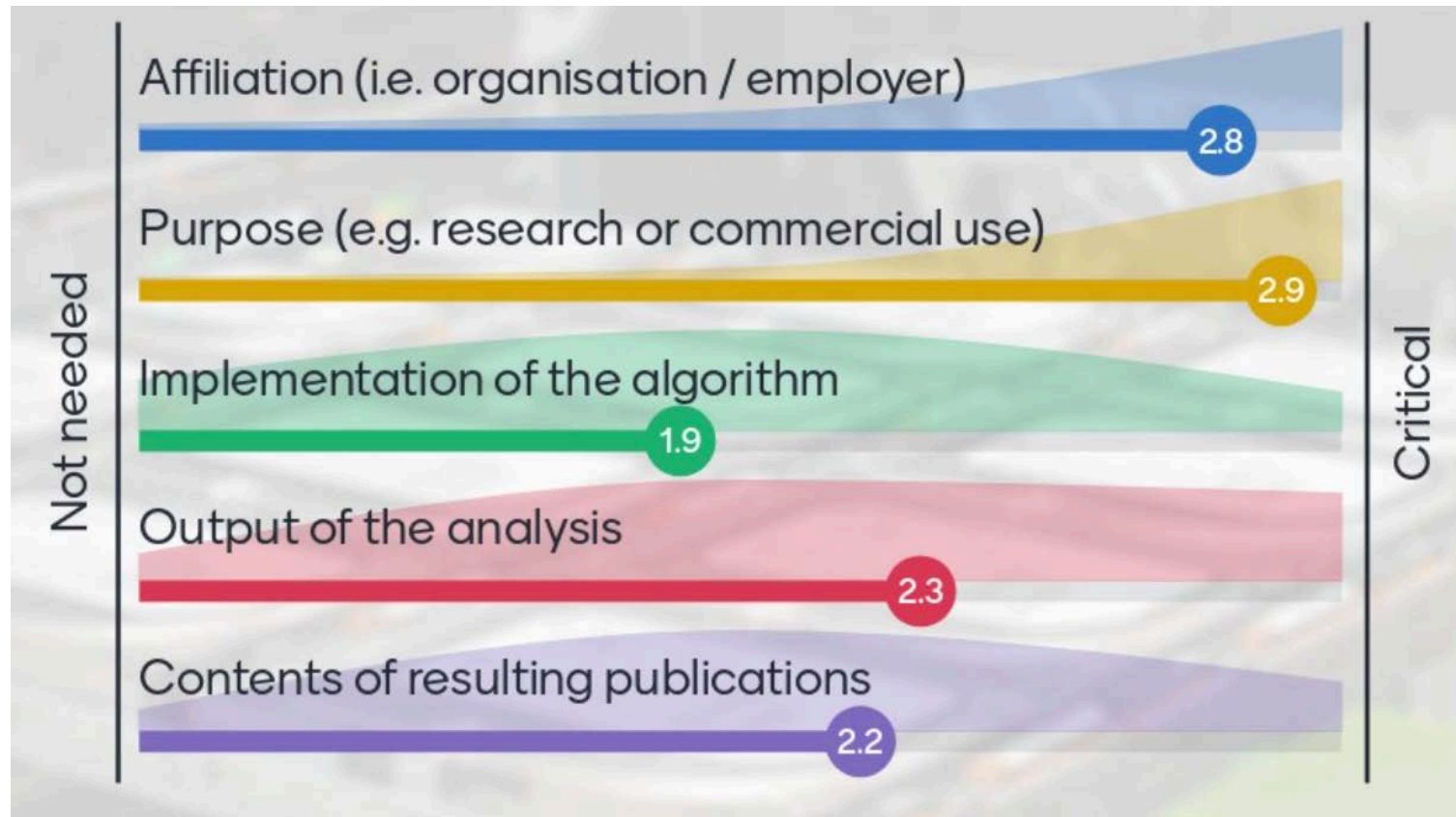
Webinars: Some data to share

Question: How often did you encounter a specific need or request to share confidential data sets at your work?



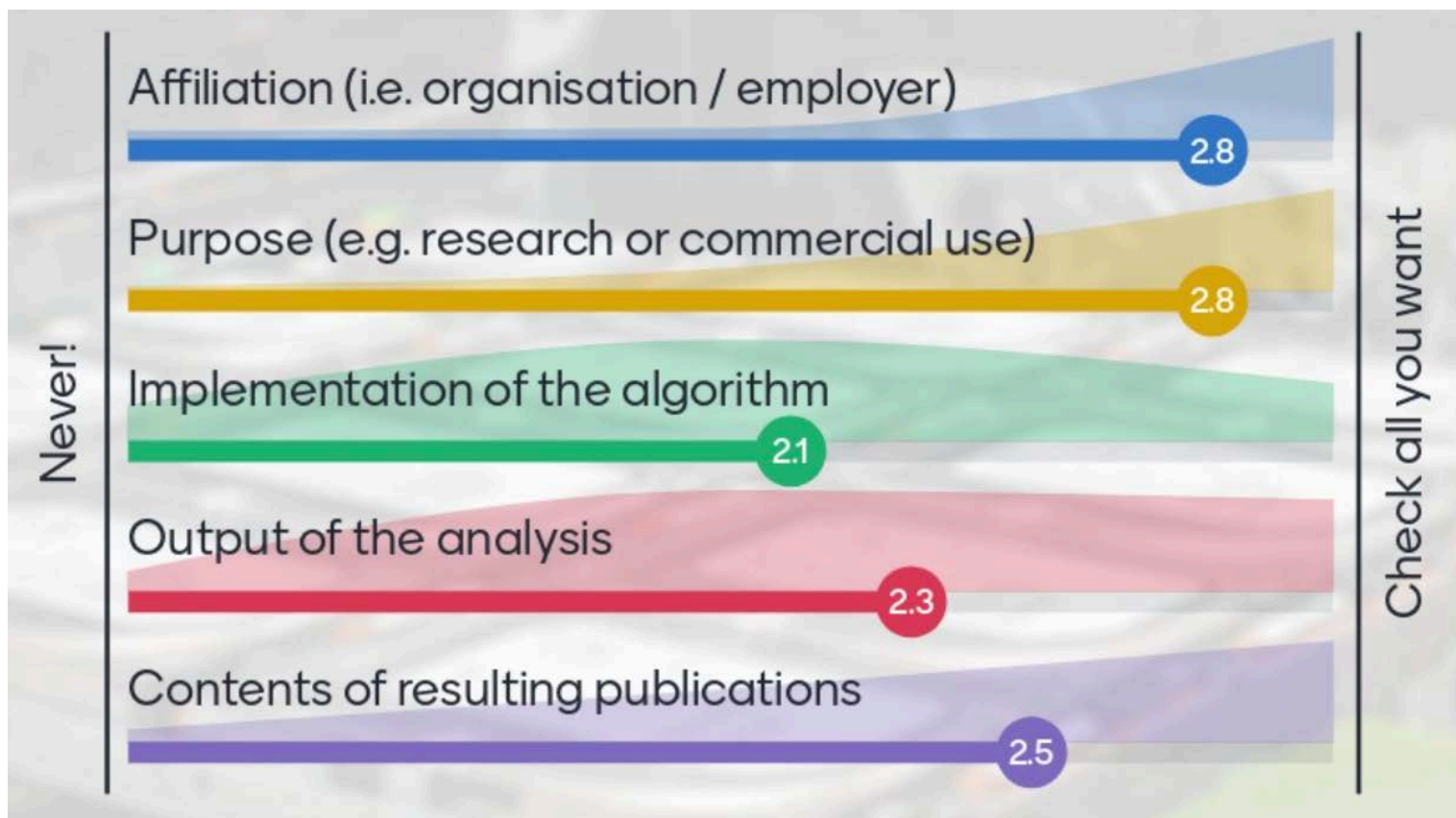
Webinars: Some data to share

Question: As a **provider** of confidential data, to what extent would you want to verify these aspects before sharing data?

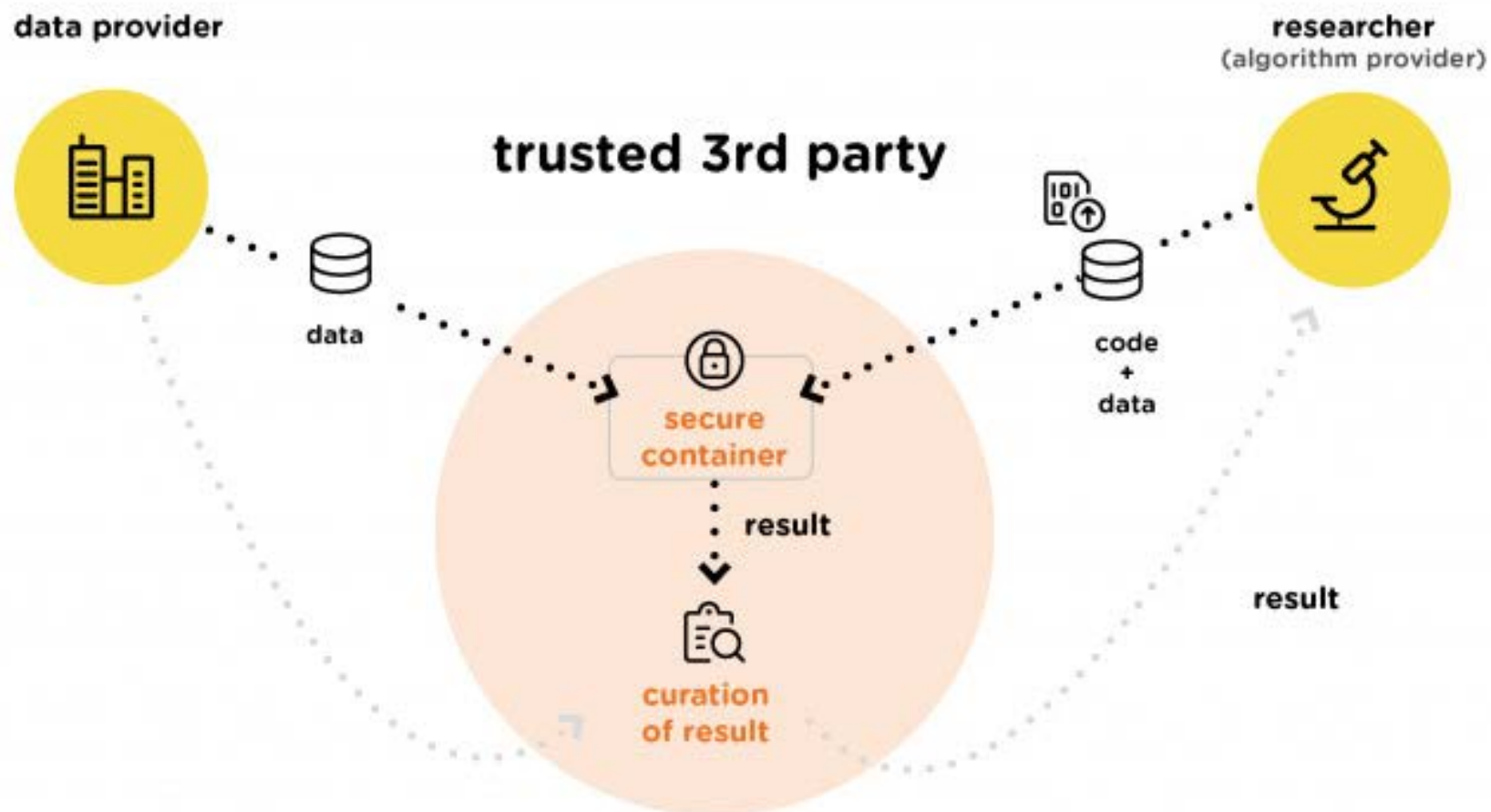


Webinars: Some data to share

*Question: As a **user** of confidential data, to what extent would you accept verification of your usage?*



Prototype: SURF Digital Data Exchange



Interested? We'd like to hear from you!

- Information and webinars available at: <https://www.surf.nl/en/data-exchange-trusted-data-sharing>
- Give it a try at <https://dataexchange.surfsara.nl/>
- Or email us at freek.dijkstra@surf.nl or hylke.koers@surf.nl