Outline

- What is data?
  - Types of datasets
  - Examples
- Why share data?
  - Why share data on GBIF
  - Barriers to sharing data
- How to identify data for publishing
- Mentimeter Quiz

What is Data?

- Understanding the GBIF mission - to facilitate free and open access to biodiversity data worldwide to underpin sustainable development can help us identify what sort of data we are referring to.
  - Eg: invasive species data for biosecurity and IS management
  - Biodiversity data on threatened bird species for protection and conservation
- Data can be complex and have different origins
- In this context, when discussing data, it is primarily *distribution* data on plants, animals, fungi, and microbes for the world, and scientific names data.

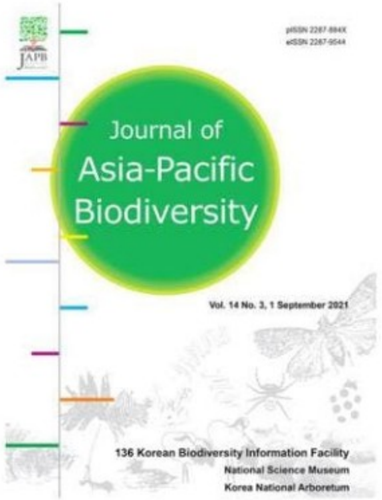
Types of Datasets

Different types of data that can be called primary biodiversity data and shared within GBIF. These include:

- Species Checklists - *a list of taxa. In Darwin Core terms, these data only have taxonID (UUID) and scientificName.*
- Occurrence-only datasets – *records of the existence of an organism at a specified place and time.*
- Sampling-event datasets – *provide greater detail about a species occurring at a given location and date, including the methods, events, and relative abundance of species recorded in a sample.*

What are examples of Data?

- Fieldwork data – survey reports, Environmental Impact Assessment studies, field notes.
- Literature documents – scientific articles, Masters or Ph.D. thesis, reports
- Specimen and collection materials.



Why Share Data

- Sharing data allows us to make ourselves known to the biodiversity research and education community. Sharing makes the statement that we are an active member of the community, that we want to help, and want to be engaged within the field of biodiversity.
- By sharing, you gain recognition for your data collection and sharing efforts through data citations in research, demonstrating the value you are adding to global knowledge about our biodiversity. Data about tropical island species can be challenging to obtain and store safely, but this data is valuable to Pacific communities, environmental managers, and the global research community.
- Storing data in a shared repository keeps the data safe from server crashes, computer loss, and file corruption. The metadata and explanatory information required for data sharing also help keep the data valid after staff turnover.
- For a small island nation, publishing our data collections from our home environment celebrates development in areas of science and research on a global stage.
- Informed decision-making depends on quality information. Sharing data helps support research and sustainable development, especially in areas of:
    - Conservation
    - Food Security
    - Climate Change
    - Human Health

Why share data on GBIF?

- GBIF is widely recognized as the major international repository for biodiversity information. It meets high standards for data access, preservation, resource stability, and suitability for use by all researchers.
- By publishing data through GBIF , you make it easy for the international research community to discover and access it. This facilitates integrative global analyses of biodiversity and ecosystems.

Barriers to sharing data

- **Psychological and behavioural barriers** such as:
  - unwillingness to share for commercial reasons
  - barriers resulting from concerns over how data, information and knowledge might be used,
- **Barriers relating to describing information and data**, which range from:
  - lack of widely agreed classification systems for some types of biodiversity,
  - to insufficient use of contextual and explanatory information linked to datasets.

- **Practical barriers**, which range from knowing:
  - how to make data, information and knowledge you hold available in meaningful ways
  - to locating the information that you need from amongst the plethora of data, information and knowledge available.
- **Inadequate strategies and resources** that result in data, information and knowledge often being made available in an opportunistic manner rather than being focused on need, often without sufficient resources being made available.

How to identify data for publishing

- While new data can and should be published, the focus of this project is for national partners to identify existing data for publishing.
- Some valuable data remains in old files waiting to be dug up again and are worth sharing.
- Data in spreadsheets and data records are ideal
- However, data in any format can be extracted and put into columns and rows. For example, you may have a report that verbally describes encounters with certain species in a particular area or areas are all important.
- Data on endemic species or new areas as well as species found in common areas are all important.
- You can decide whether your data is occurrence data, checklist data, or sample event data. Each of these types requires different amounts and details of data.


QUIZ
time

Recap & Quiz

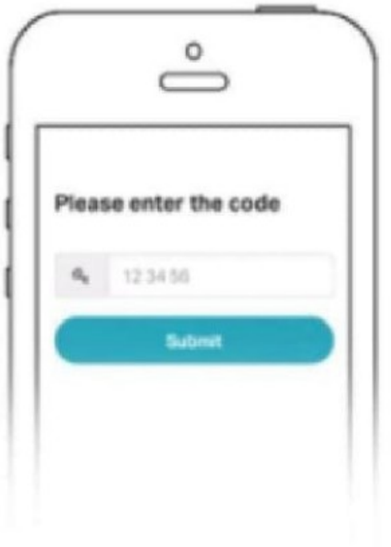
1

Grab your phone

www.menti.com

2

Go to www.menti.com



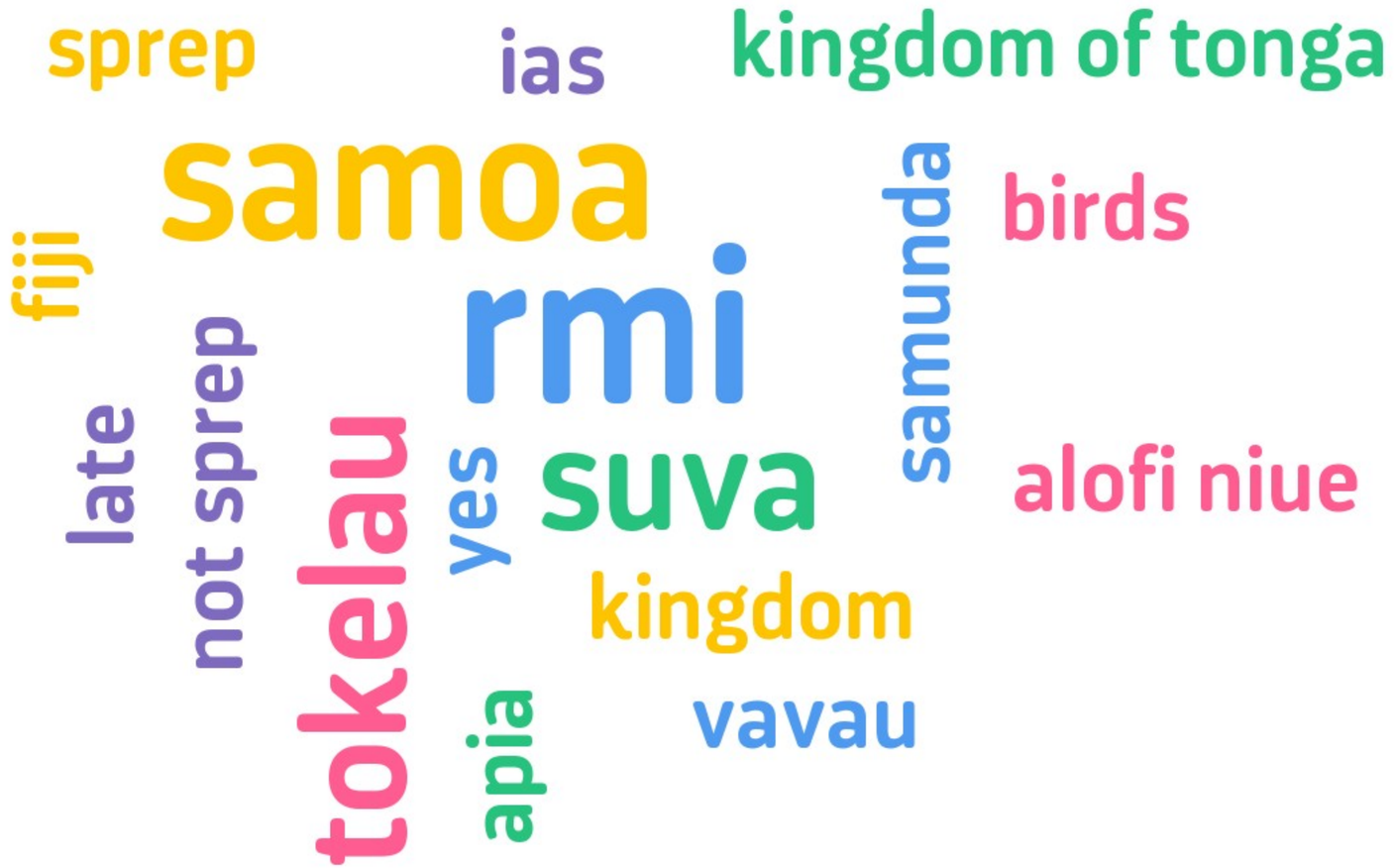
3

Enter the code 90501646

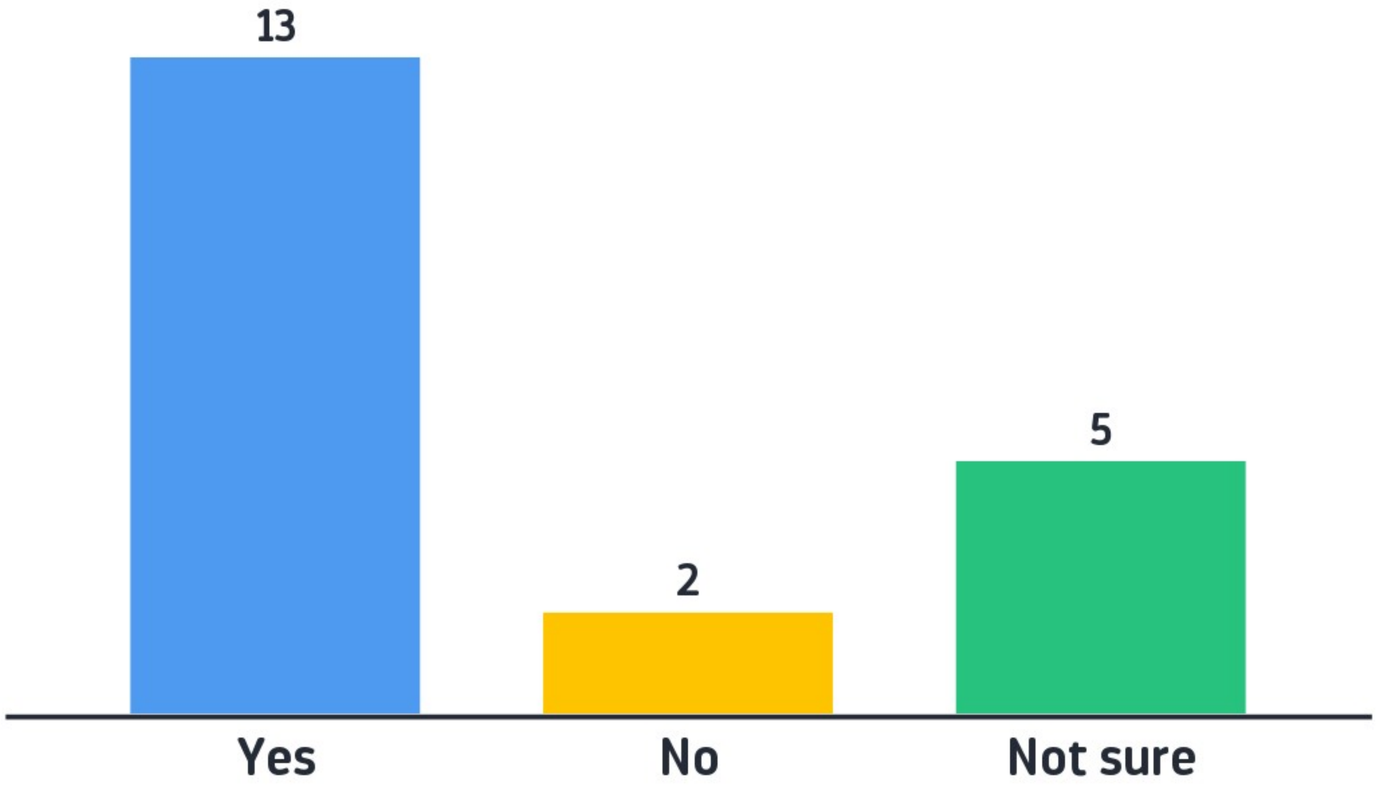
20

Send us a greeting in your language

Where are you joining from?



Does your organisation have biodiversity datasets?



If you answered yes, can you name one?

Toloa Rainforest

Reef fish

Biocontrol agents

Species-specific datasets

marine species records from Fiji

Whale strandings

turtles

Mt talau...

Coral

If you answered yes, can you name one?

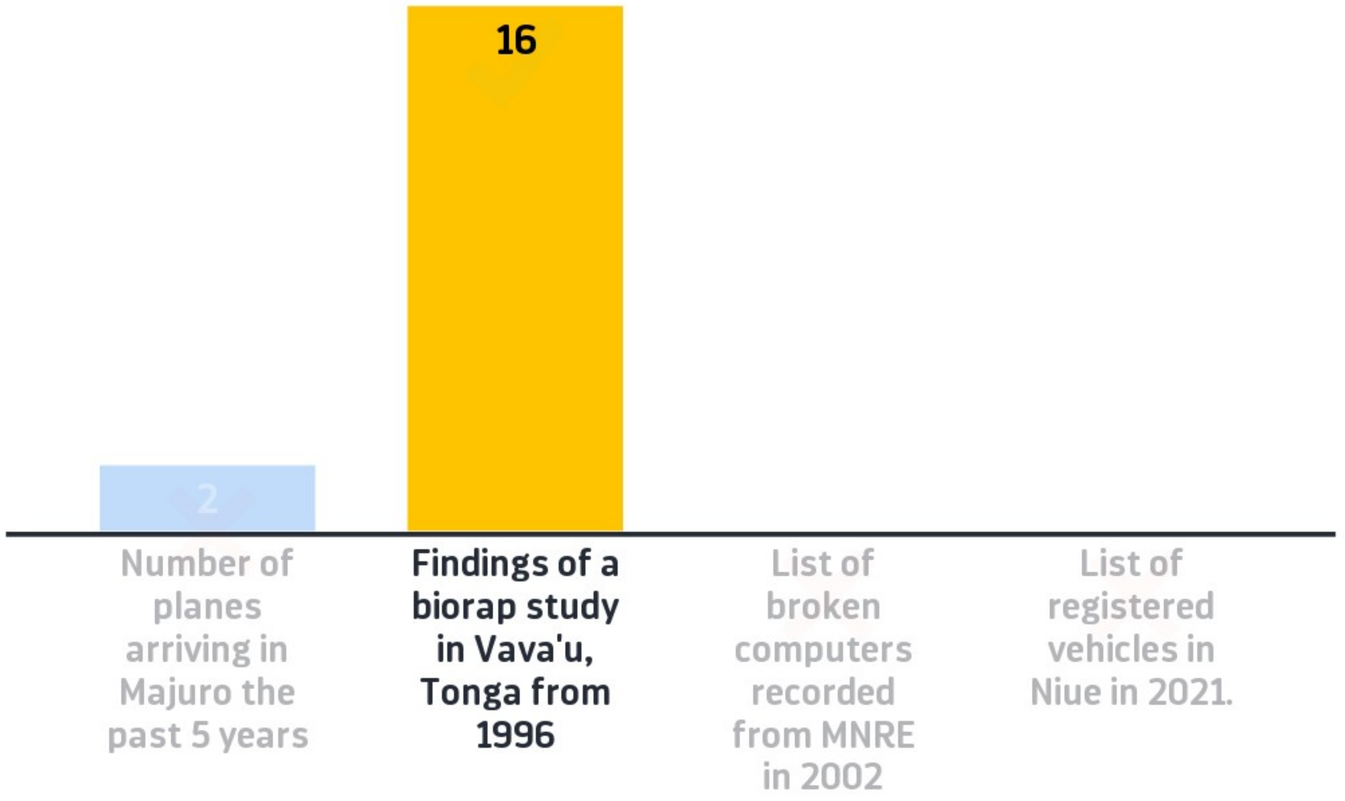
Algae

Important Bird & Biodiversity Areas

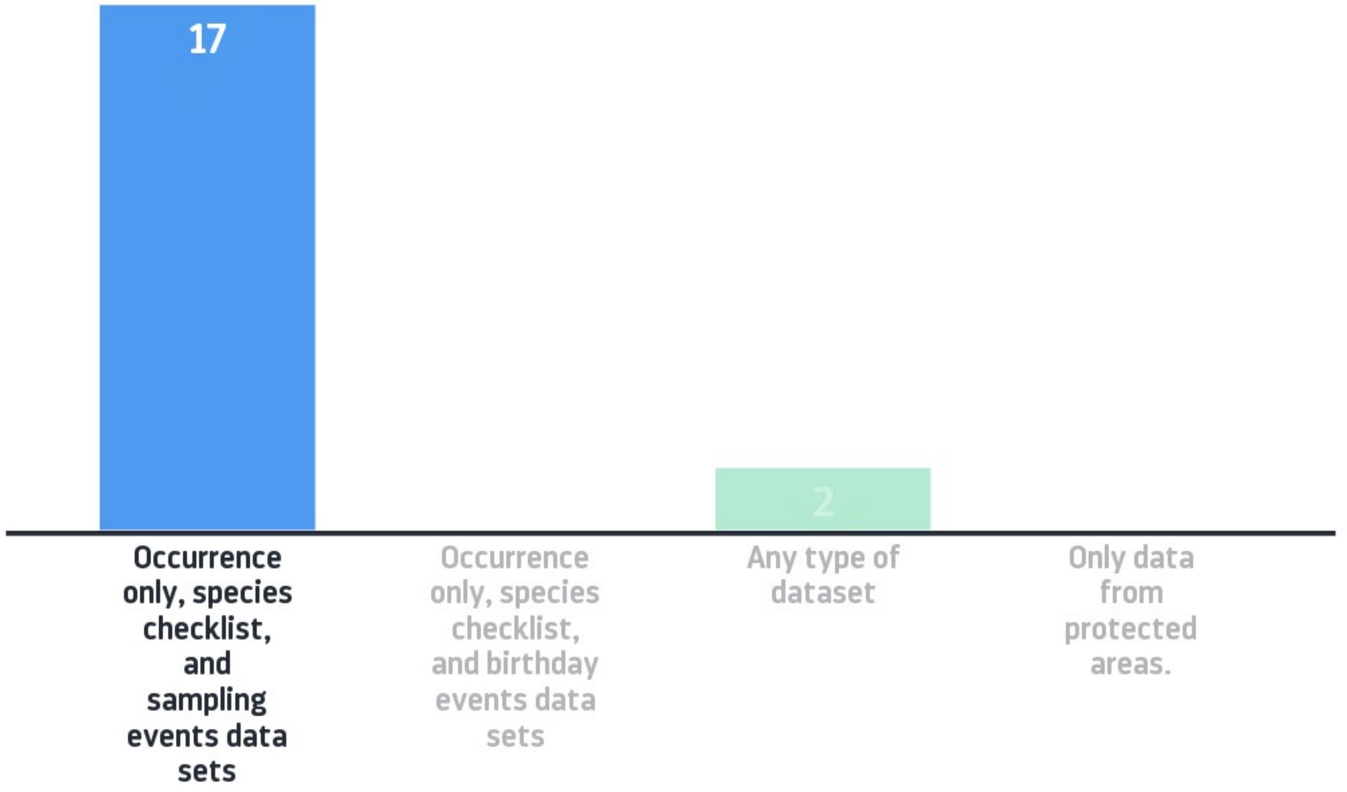
priority terrestial invasive species

Tonga SMA

Identify a suitable dataset you can publish on GBIF

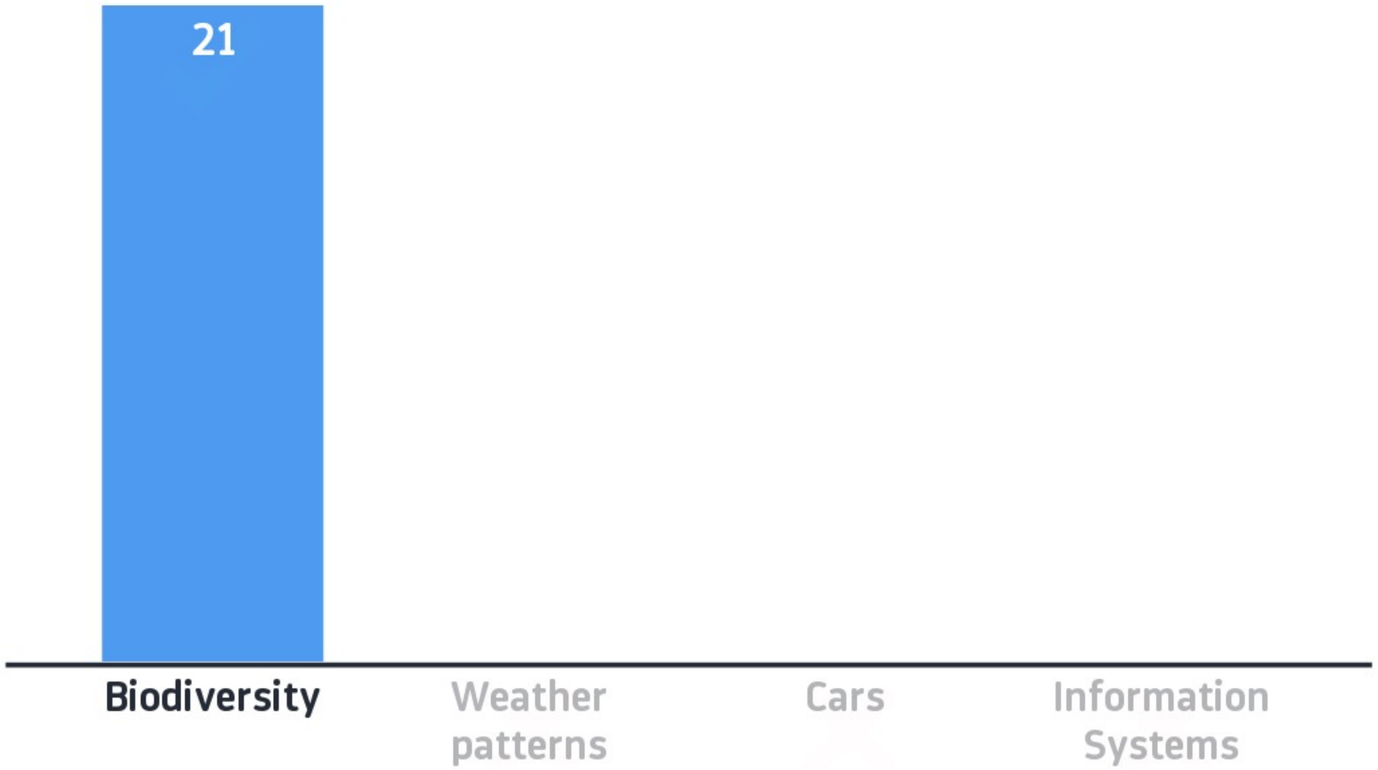


Identify the correct types of datasets supported by GBIF



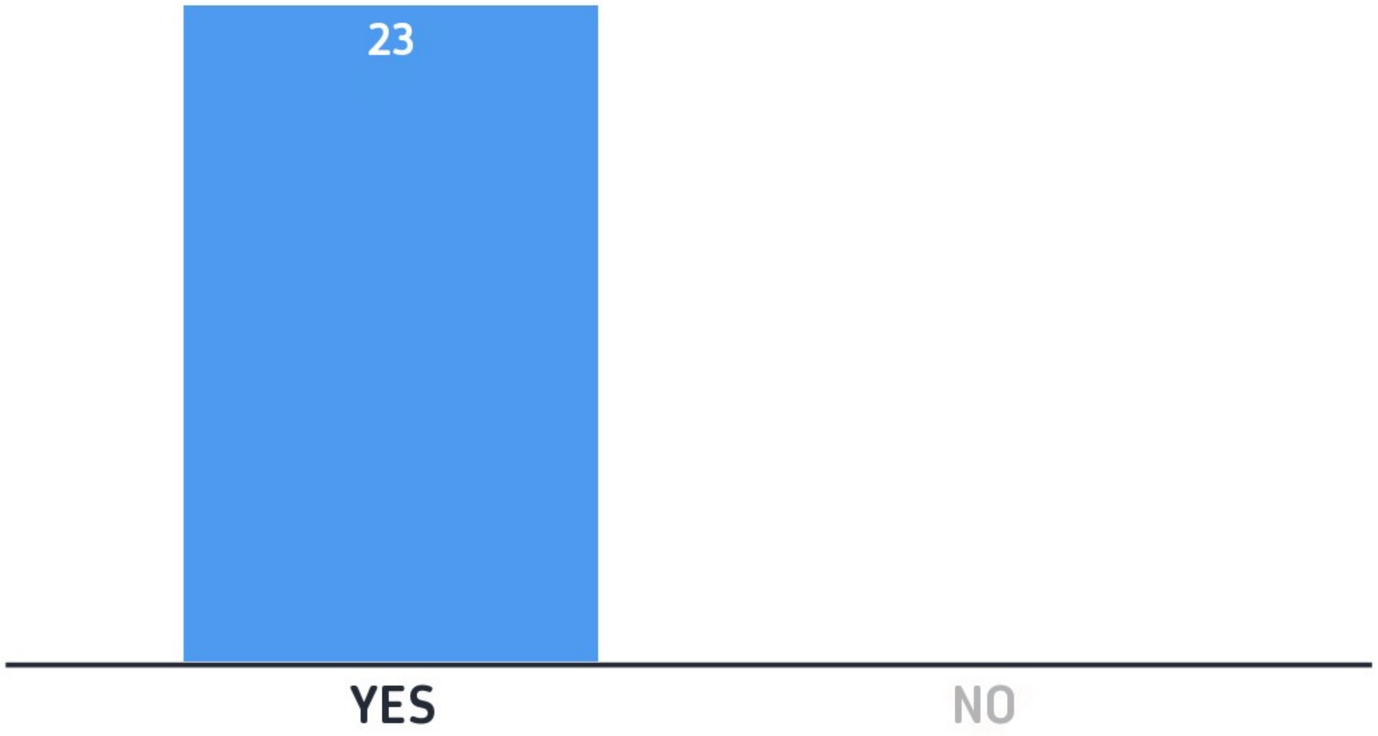
What does data mobilisation mean to you?

When identifying data for mobilization, what should the data be about?

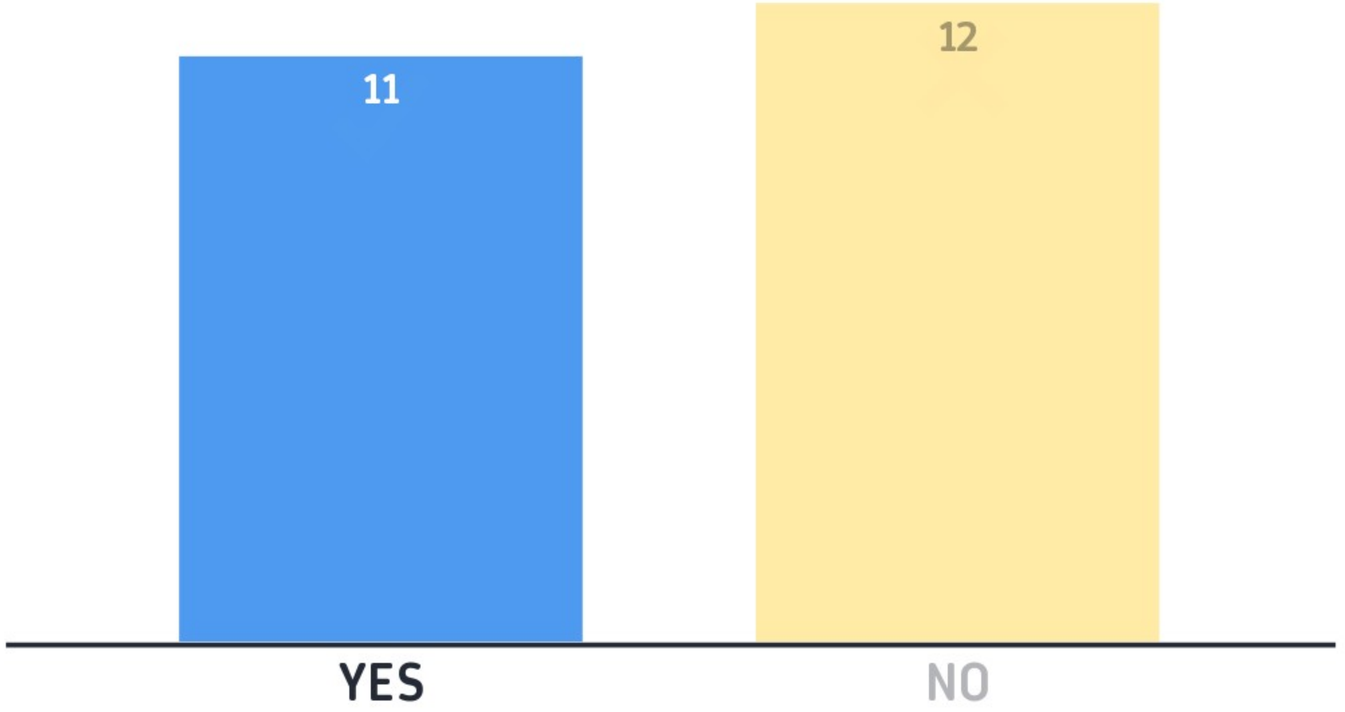


21

When identifying data for mobilization, is “old data” relevant?

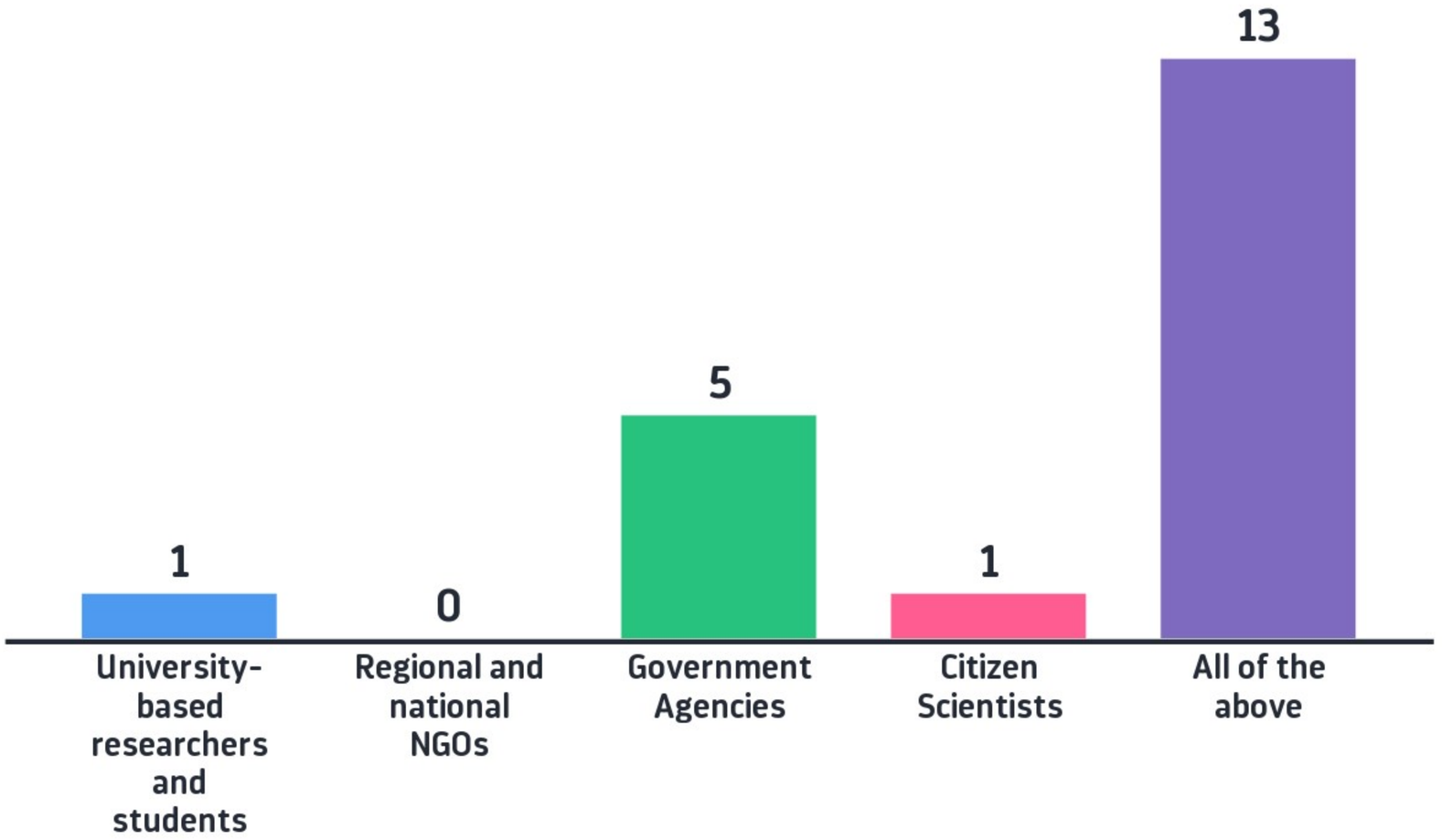


Can I publish data that does not belong to me?



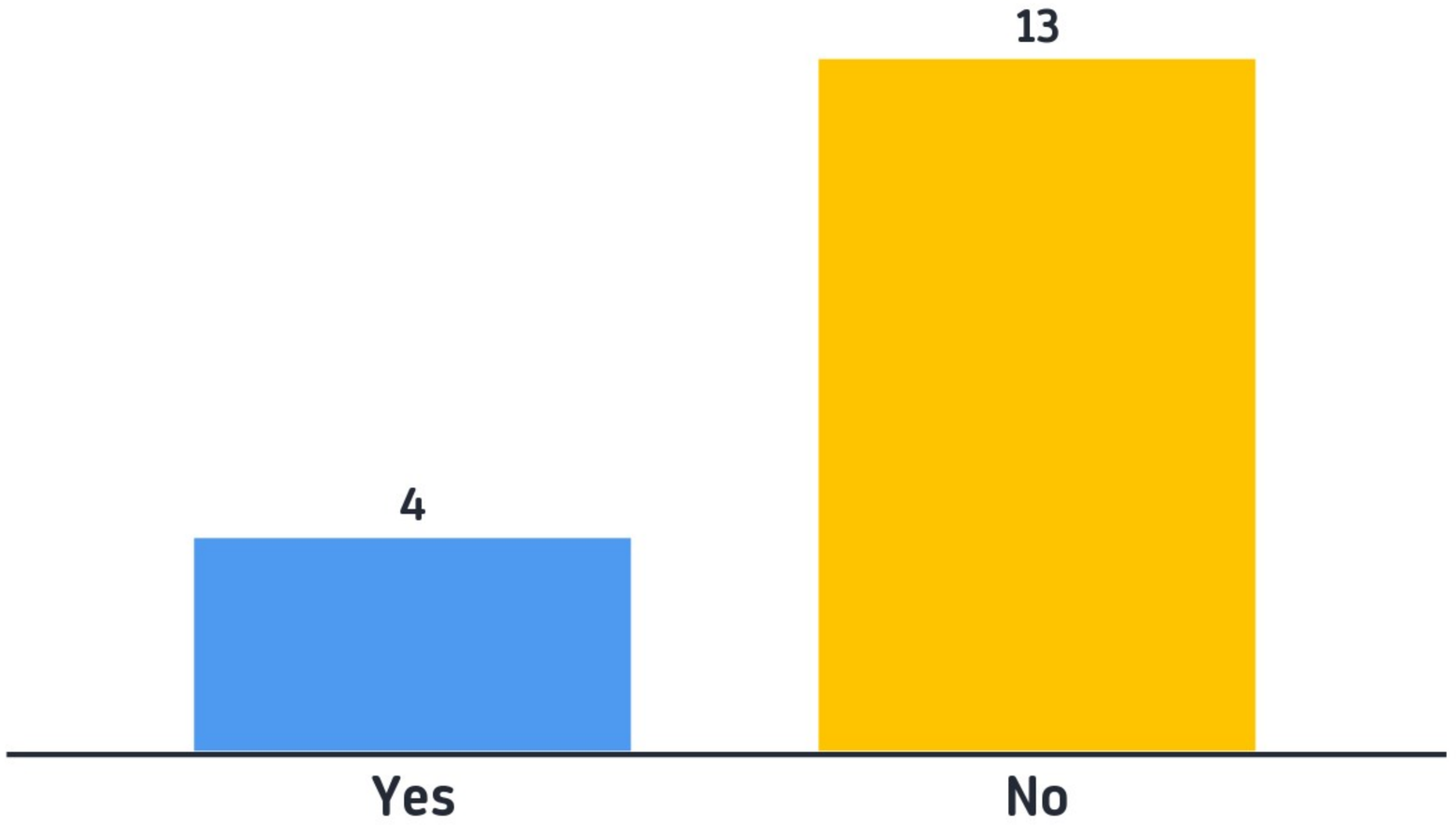
Please provide one benefit of sharing data

Identify a group of data holders at national level

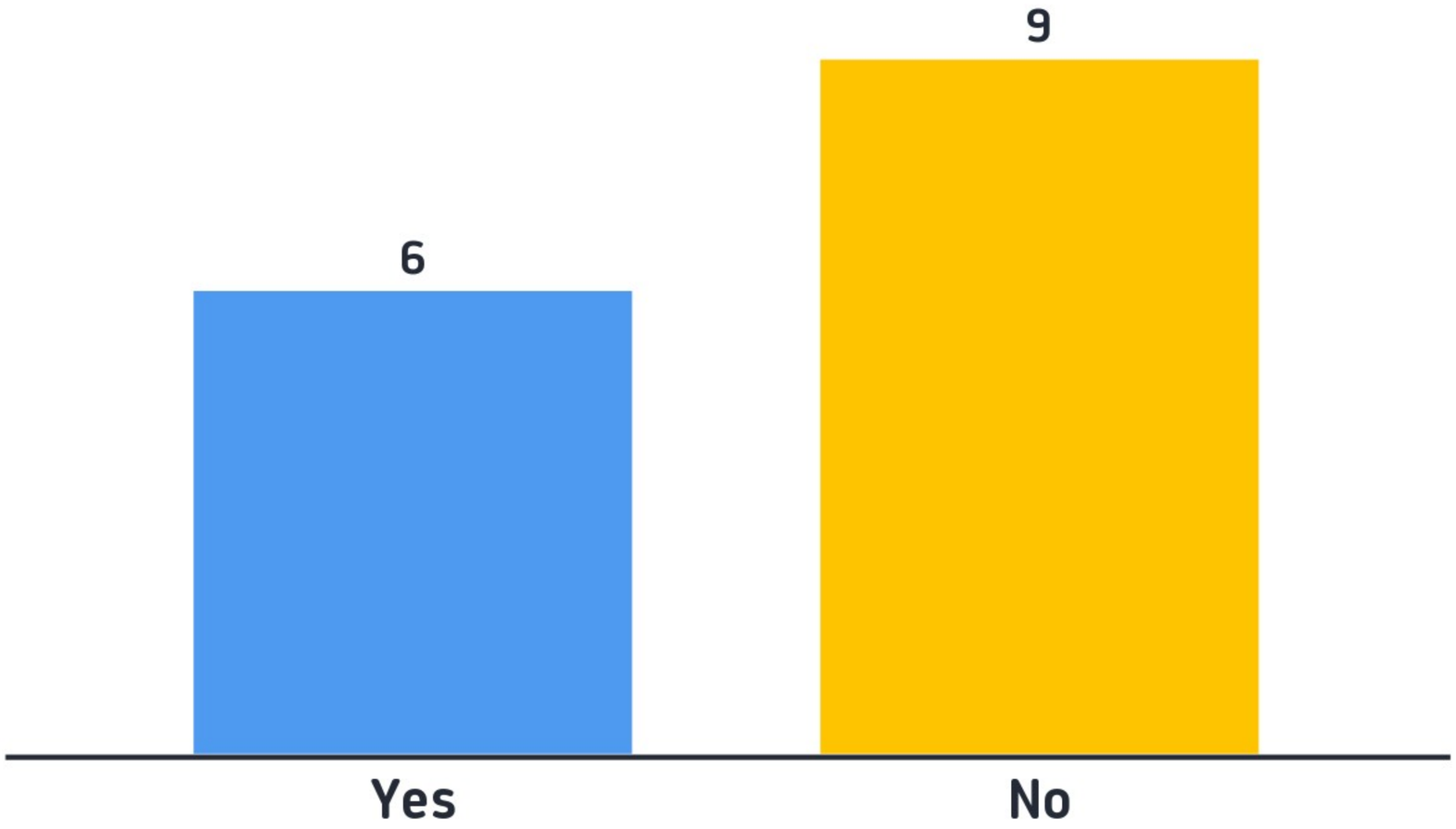


When publishing data, are any of the following a concern to you?

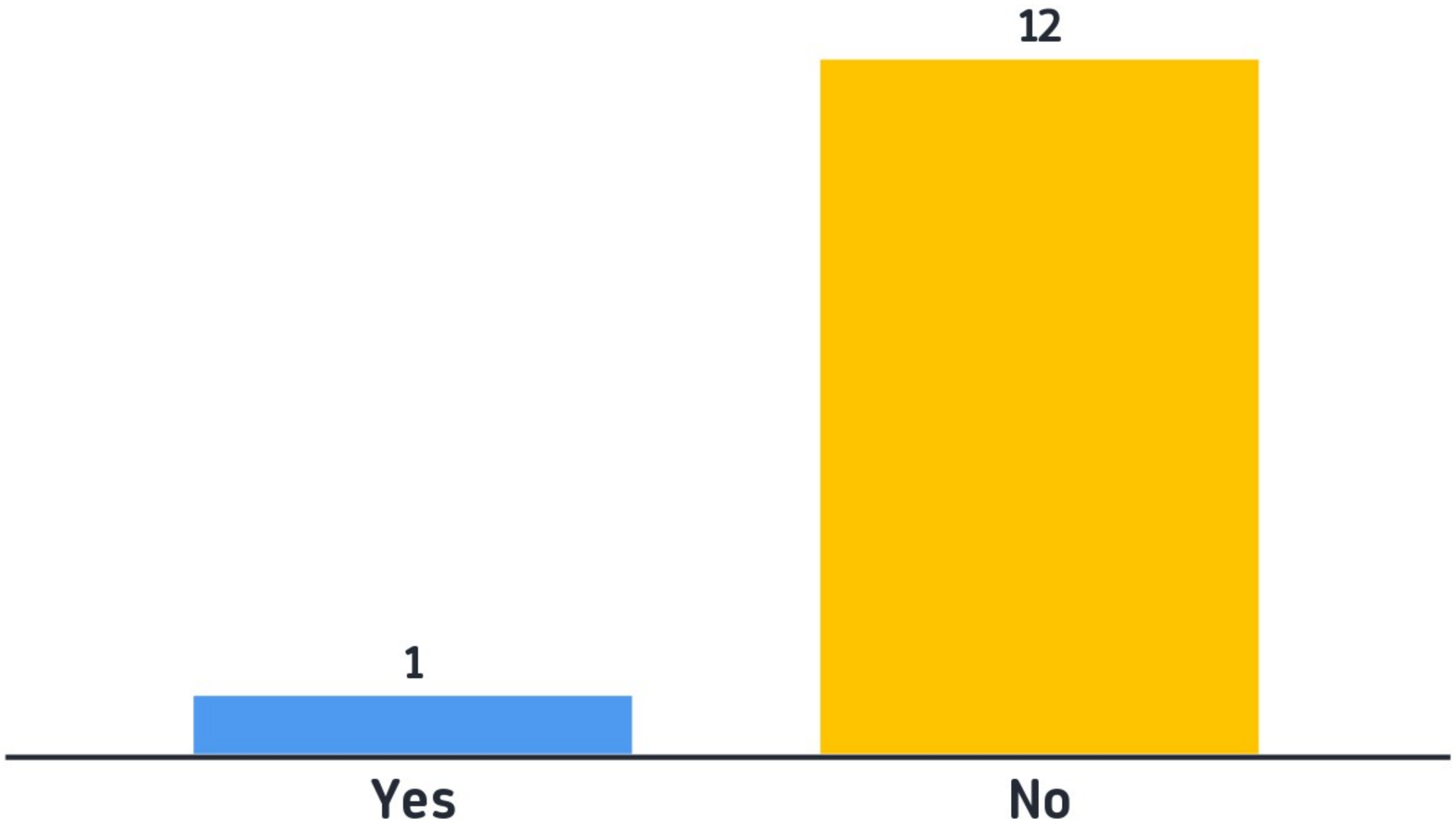
Someone will steal my data



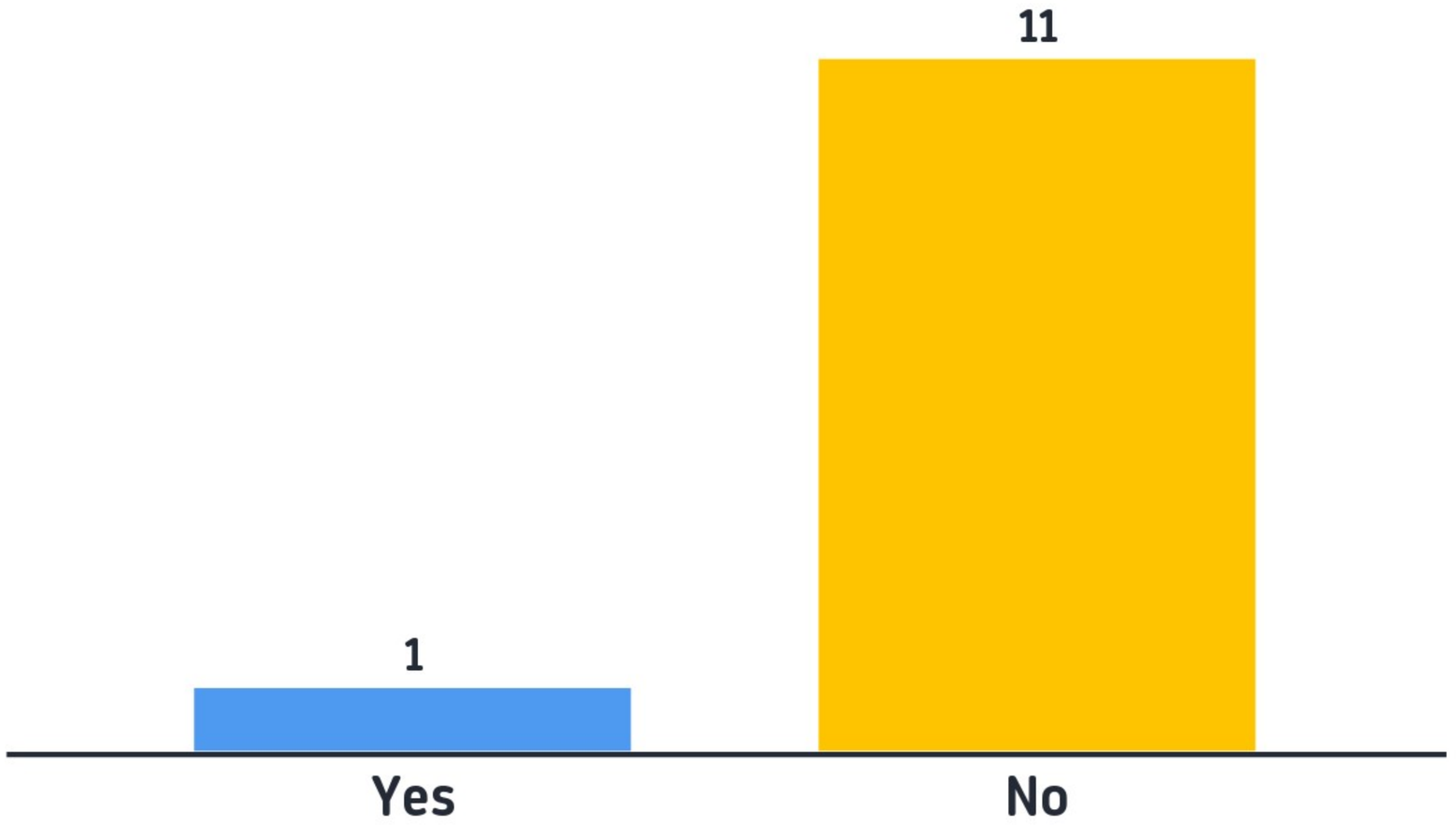
I want to share data, but it does not belong to me.



My data will not belong to me after publishing .



I don’t know how to publish data




THANK YOU!