

# gesis

Leibniz Institute  
for the Social Sciences



## Re-examining Data Governance

Deb Wiltshire, Secure Data Center, Cologne

Linking Survey & Digital Data, NatCen 2024

# Roadmap

1. The role of Trusted Research Environments
2. Re-examining data governance
3. Where does that leave us?



Sample Footer Text

**DATA**

PRO

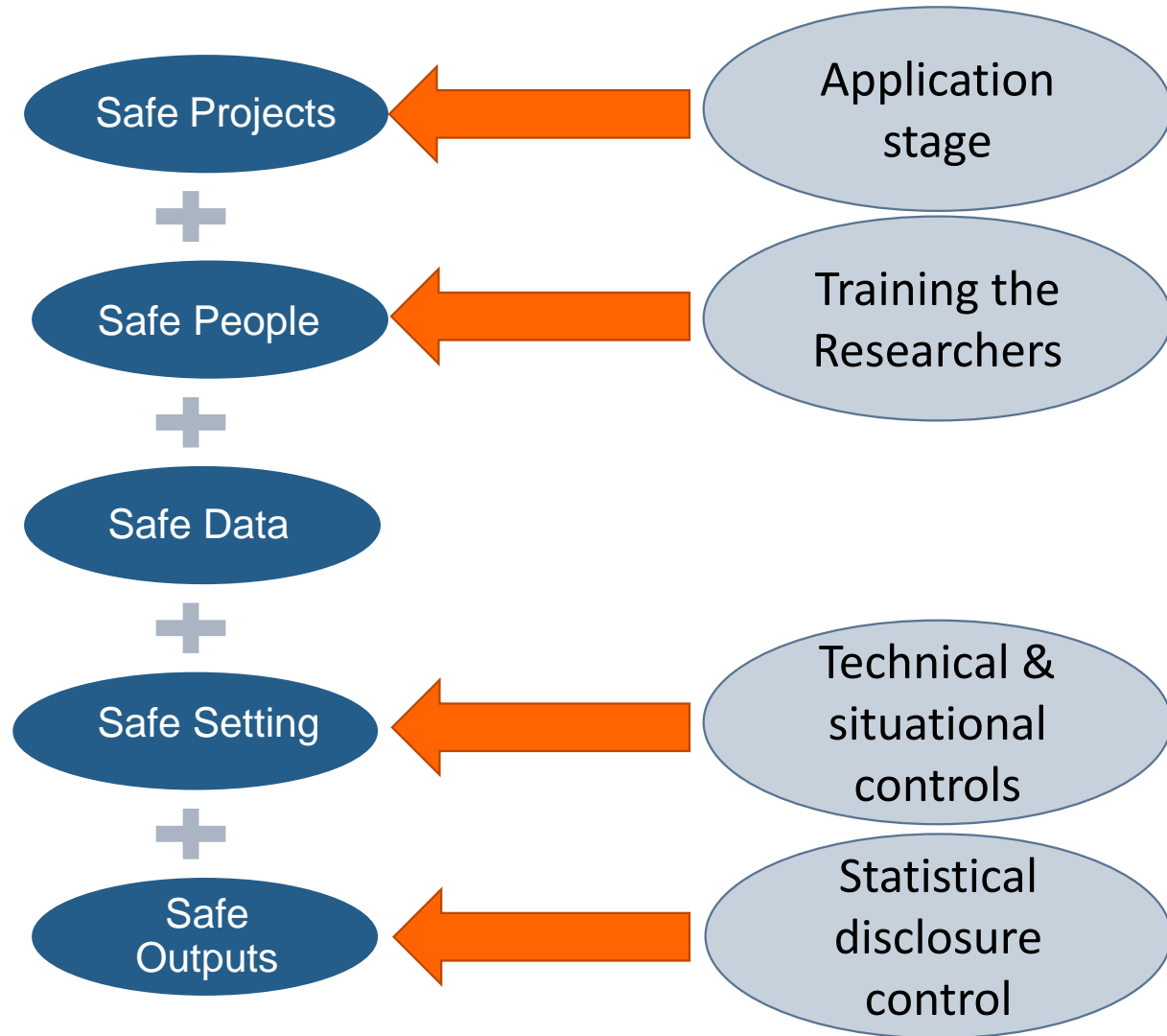
This Photo by Unknown Author is licensed under CC BY

# The role of Trusted Research Environments



- Not all data can be anonymised & remain useful
- Need to find ways to balance the risks and benefits
- Data governance & Trusted Research Environments

# The 5 Safes Framework



# The landscape is constantly changing...

- Country differences
- Pandemic = new rules
- Emerging new data forms – digital behavioural data



[This Photo](#) by Unknown Author is licensed under [CC BY-NC](#)

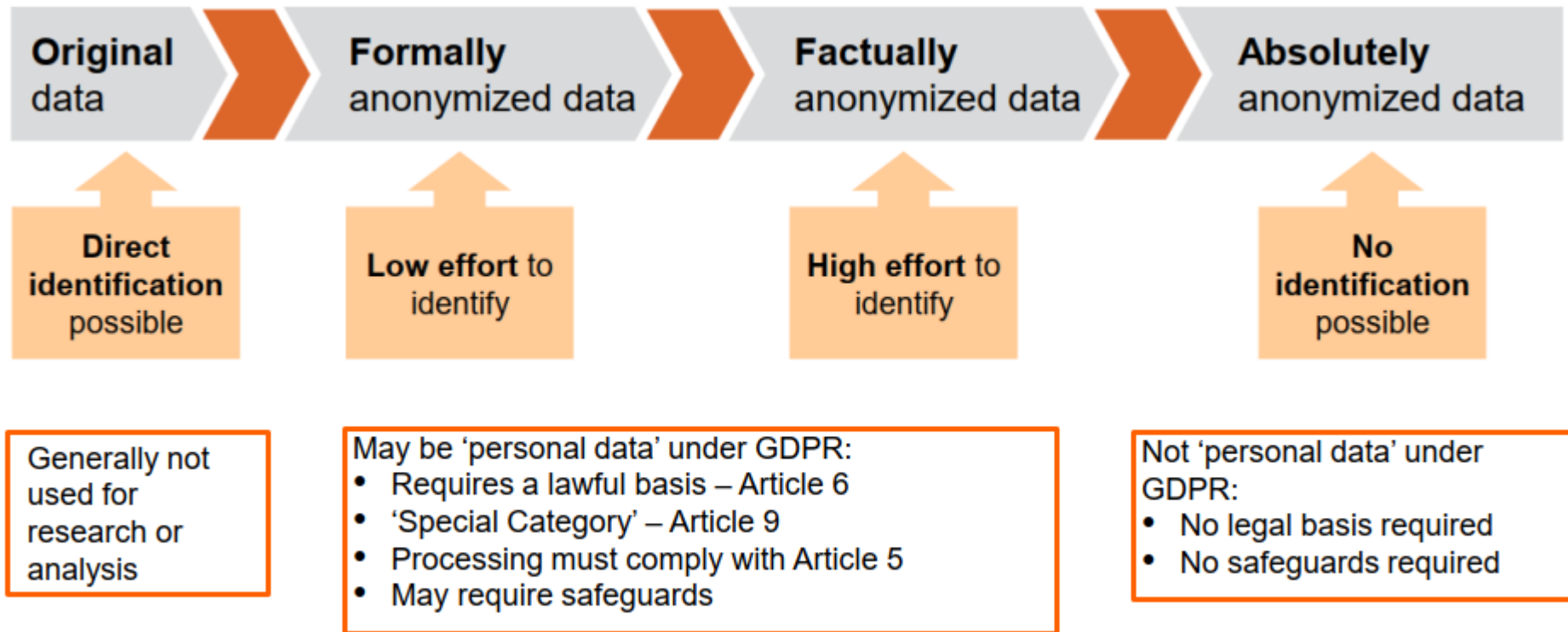
Do we know what is  
personal data?

# What is personal data?

‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); **an identifiable natural person is one who can be identified, directly or indirectly,** in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person; (GDPR Article 4, Definitions 13).

Recital 39 adds -“Personal data should be processed in a manner that ensures **appropriate security and confidentiality** of the personal data...”

# What influences data access





The legal report for Breyer v Bundesrepublik Deutschland adds an important interpretation:

“its wording suggests that, for information to be treated as ‘personal data’ within the meaning of Article 2(a) of that directive, **it is not required that all the information enabling the identification of the data subject must be in the hands of one person.**”



[This Photo](#) by Unknown Author is licensed under [CC BY](#)

Do we know  
where the  
risk comes  
from & why  
does that  
matter?



[This Photo](#) by Unknown Author is licensed under [CC BY](#)

Are our  
systems  
based on  
control or  
trust?

## The Intruder Model of Data Security

---

Well documented model

---

Based on premise of a deliberate  
intention to breach data confidentiality

---

Leads to very cautious risk  
management

---

‘Worse case scenario’

---

Leads to greater restrictions

# The Human Model of Data Security

---

Based on the premise that people are human

---

Design system to address human factors

---

Well designed system AND well trained researchers, can allow greater access

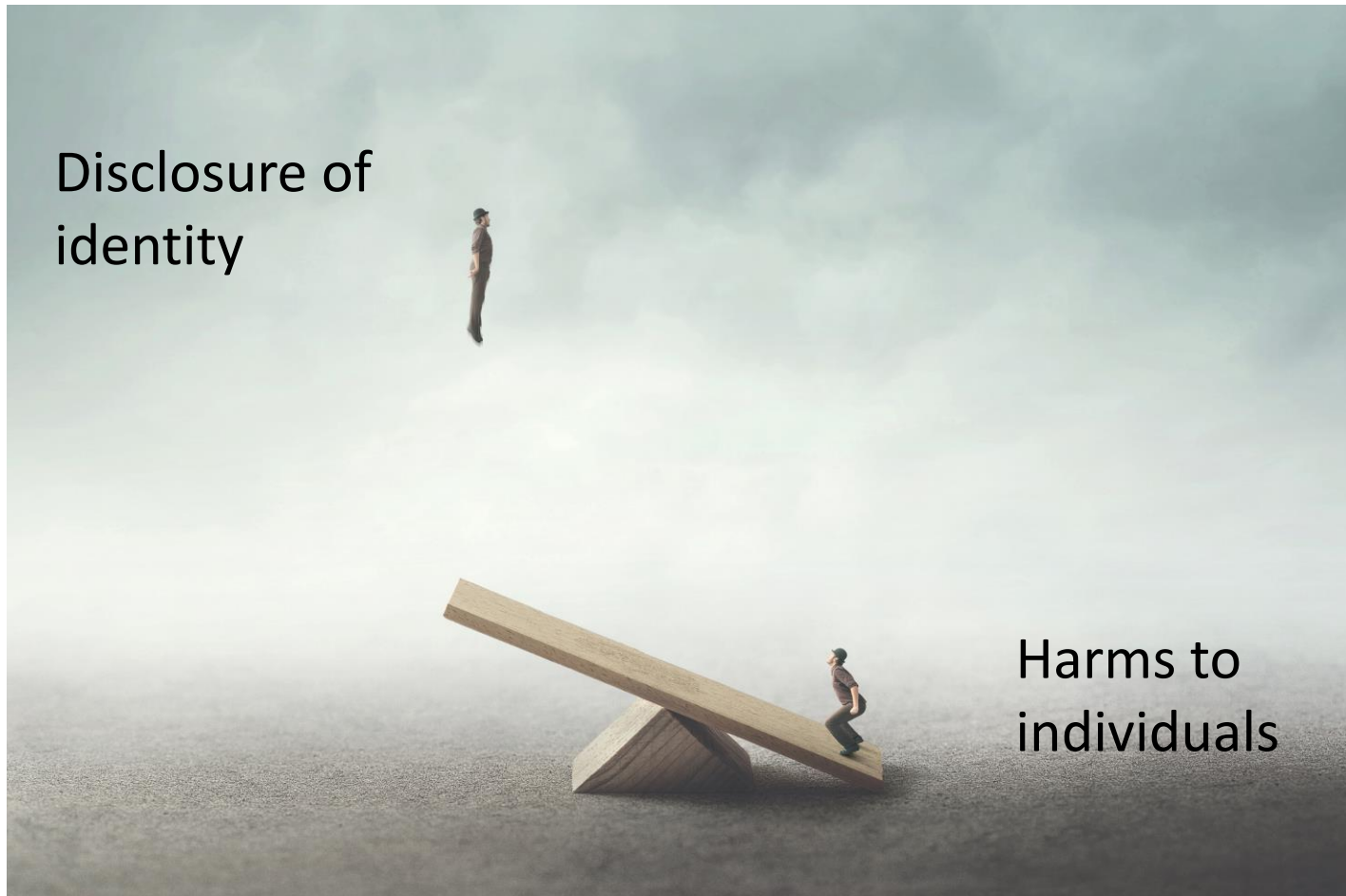
---

Allows researchers to demonstrate that they can use data safely

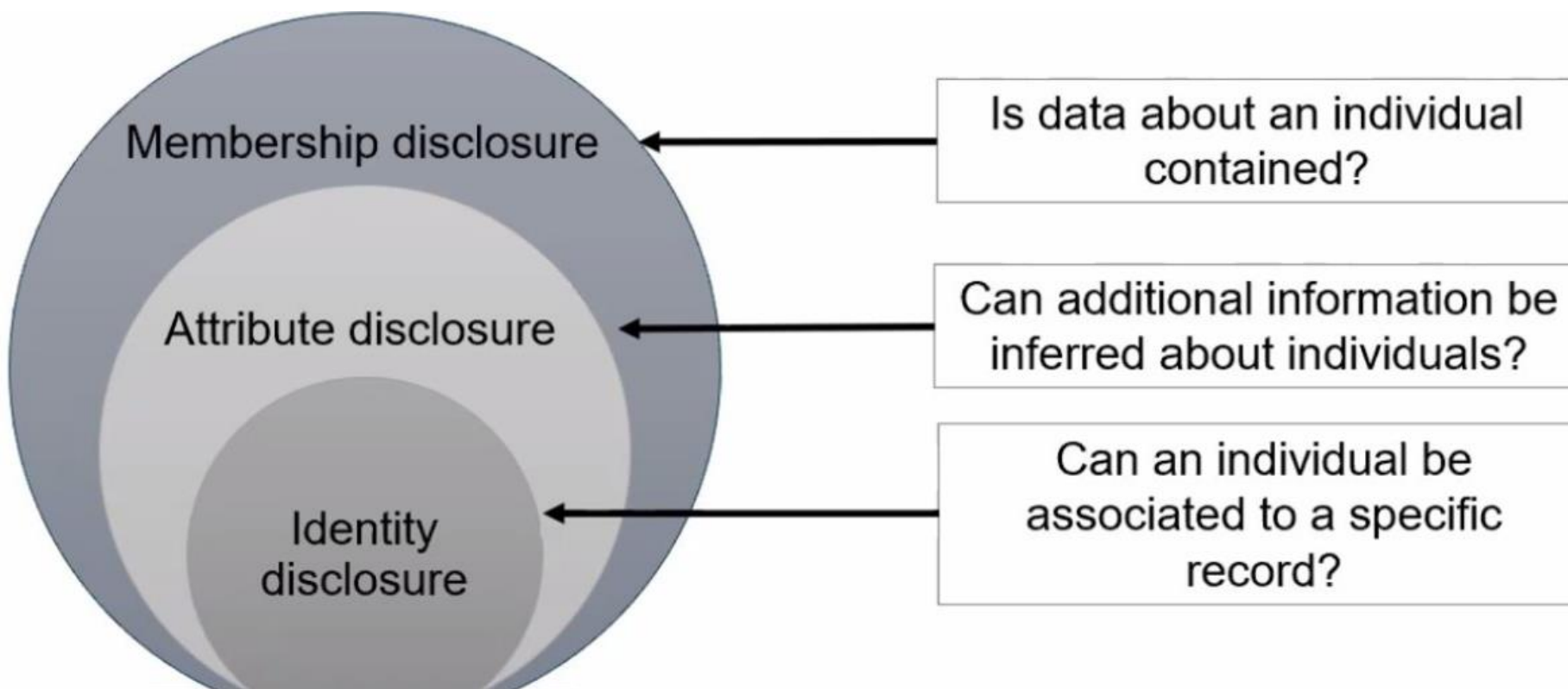
---

‘Accidental intruder’

# Do we know what we're trying to achieve?



# Thinking about disclosure and harm



# Where does that leave us?

1. Things are constantly changing
2. There are some things we cannot know



Sample Footer Text

PRO This Photo by Unknown Author is licensed under CC BY



## Where does that leave us?

“A fundamental aspect of this is that one does not know...how data will be used in the future, or what other data they will be linked with. This means we cannot usefully characterize data sets as public...or by potential use...and that the intrinsic nature of the data cannot be used as an argument that they are not risky. It is not the data per se that raise ethical issues, but the use to which they are put and the analysis to which they are subjected.”

gesis

Leibniz Institute  
for the Social Sciences



[deborah.wiltshire@gesis.org](mailto:deborah.wiltshire@gesis.org)