

DEEL 10

DATA



Data

- The Bigger Picture
- Questions to ask
- Exercise

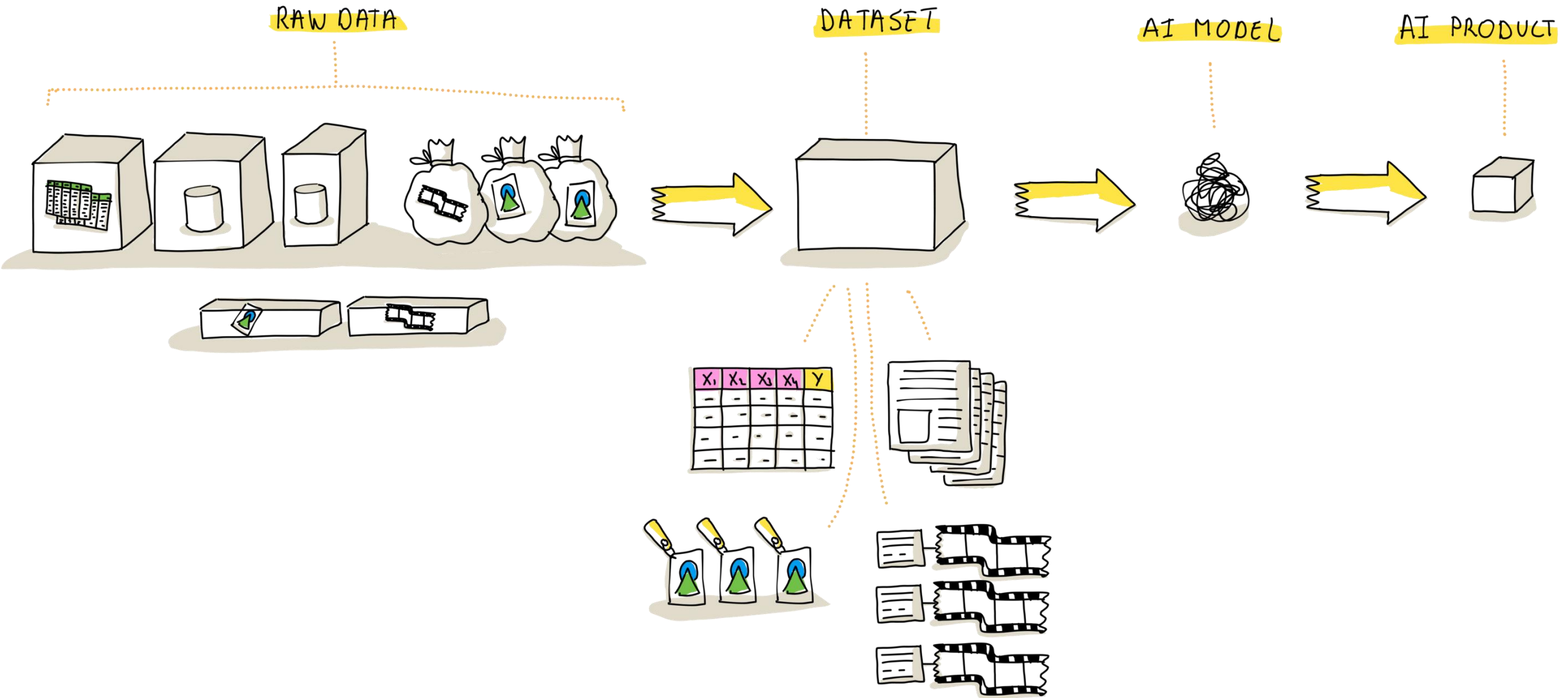


Data

- **The Bigger Picture**
- Questions to ask
- Exercise



The Bigger Picture





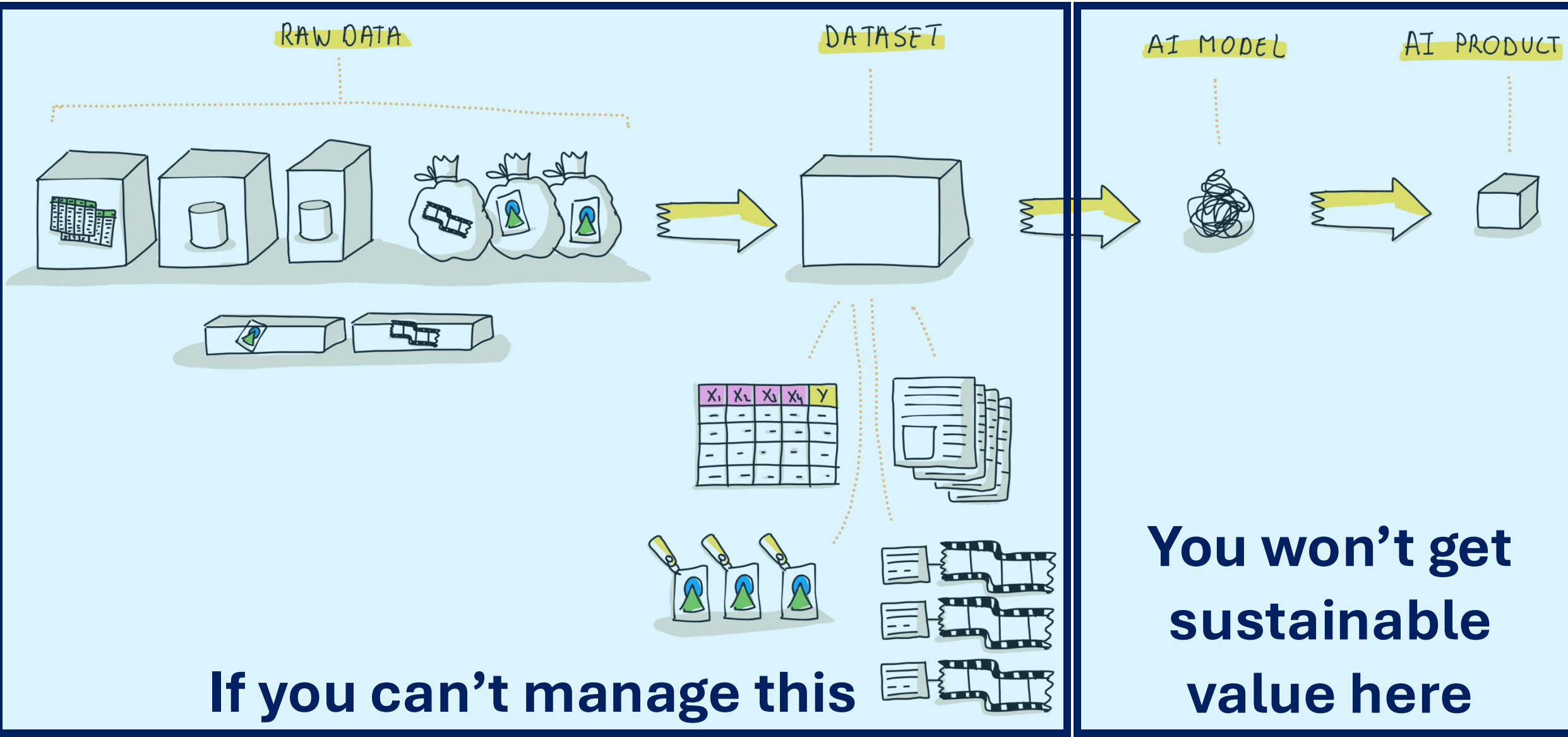
The Bigger Picture

RAW DATA

DATASET

AI MODEL

AI PRODUCT



If you can't manage this

You won't get sustainable value here



CIO JOURNAL

Rush to Use Generative AI Pushes Companies to Get Data in Order

Data management is under the spotlight again as companies seek to out-innovate competitors with large language models

By [Belle Lin](#) [Follow](#)

June 8, 2023 7:00 am ET



Share



Resize



Listen (2 min)



Corporate technology chiefs are under pressure to ensure companies' data is stored, filtered, and protected for use with AI. PHOTO: I-HWA CHENG/BLOOMBERG NEWS



Barriers to Preventing Organizations From Delivering More Value From Data, Analytics, and AI

Q11: Please rank all barriers preventing you from delivering more value from data, analytics, and AI, where 1 is the biggest/most present barrier (Note: Other (n=41) not shown):

% Ranking Each 'Rank 1 - Biggest/Most Present Barrier,' 'Rank 2,' or 'Rank 3'

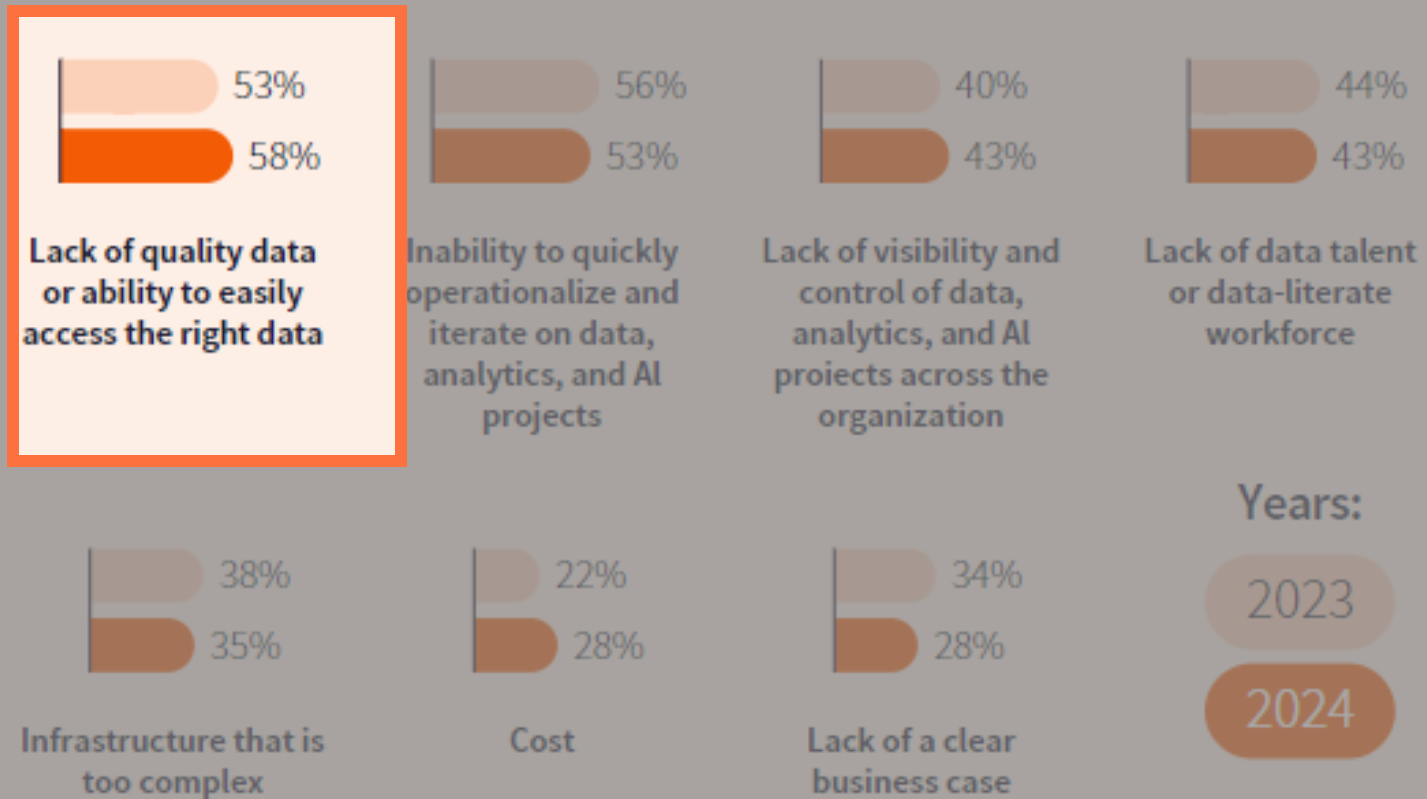




Barriers to Preventing Organizations From Delivering More Value From Data, Analytics, and AI

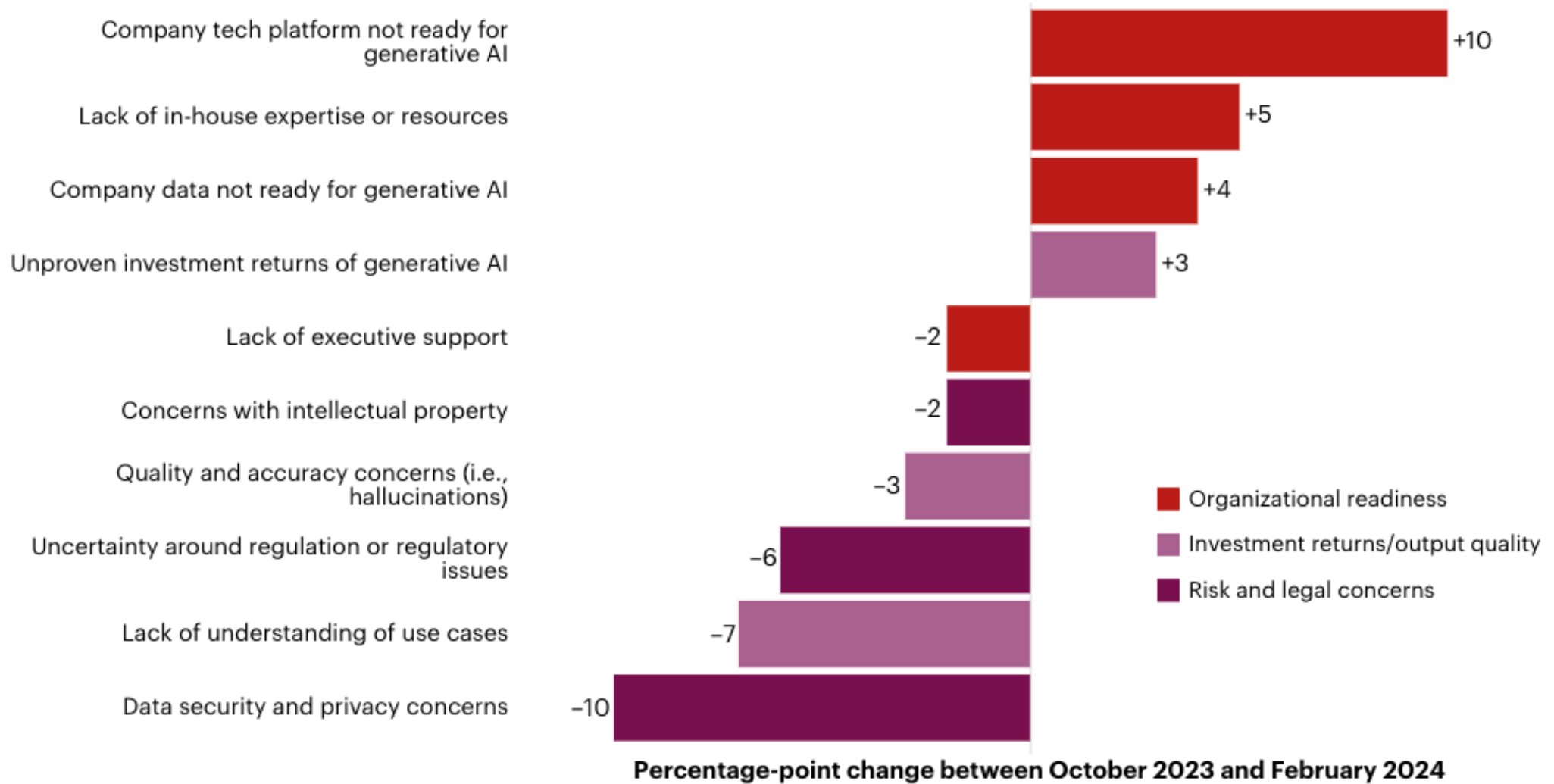
Q11: Please rank all barriers preventing you from delivering more value from data, analytics, and AI, where 1 is the biggest/most present barrier (Note: Other (n=41) not shown):

% Ranking Each 'Rank 1 - Biggest/Most Present Barrier,' 'Rank 2,' or 'Rank 3'





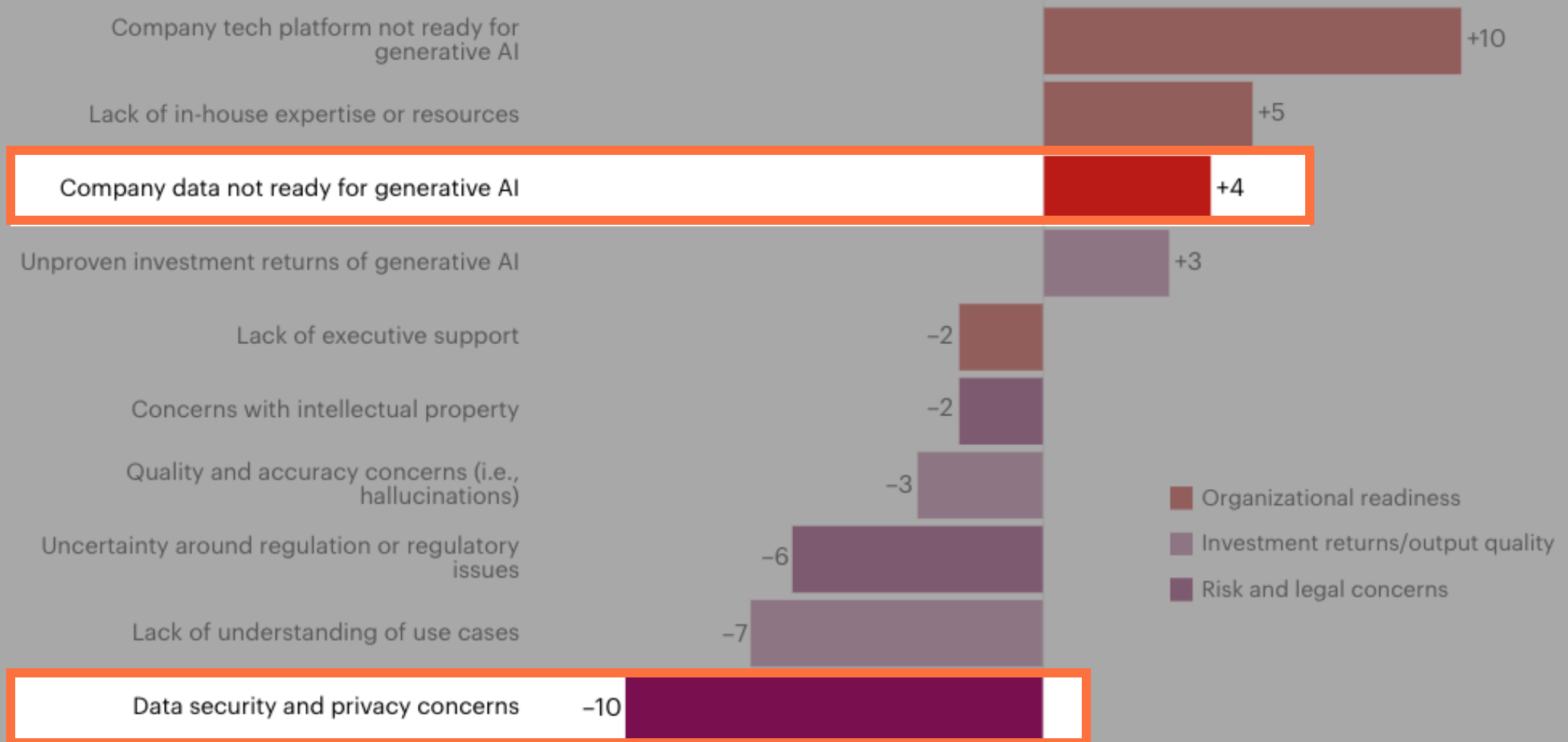
What are the top reasons preventing your company from moving faster with generative AI?



Sources: Bain Generative AI Surveys, October 2023 (N=198) and February 2024 (N=200)



What are the top reasons preventing your company from moving faster with generative AI?



Percentage-point change between October 2023 and February 2024

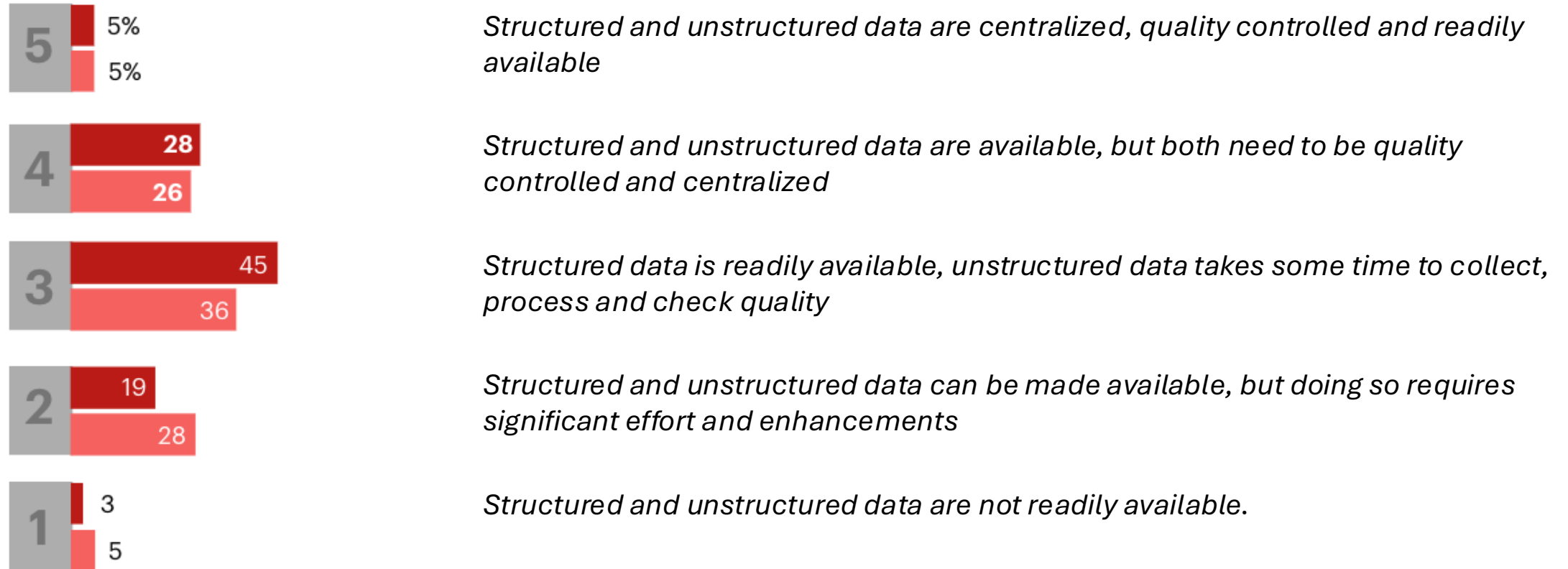
Sources: Bain Generative AI Surveys, October 2023 (N=198) and February 2024 (N=200)



Data Readiness for Generative AI

Tech Nontech

Data readiness for generative AI



Source: Bain Generative Artificial Intelligence Survey, February 2024 (N=200)



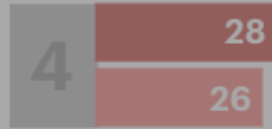
Data Readiness for Generative AI

Tech Nontech

Data readiness for generative AI



Structured and unstructured data are centralized, quality controlled and readily available



Structured and unstructured data are available, but both need to be quality controlled and centralized



Structured data is readily available, unstructured data takes some time to collect, process and check quality



Structured and unstructured data can be made available, but doing so requires significant effort and enhancements



Structured and unstructured data are not readily available.



Data

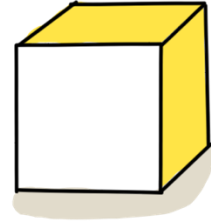
- The Bigger Picture
- **Questions to ask**
- Exercise



DATA



**AI USE
CASE**



2 Do we have the needed data?

3 Which type of data?

4 Can we trust the data?

5 Is the data sensitive?

1 Which data do we need to make AI work?

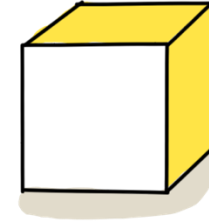
6 Can we access the data?



DATA



**AI USE
CASE**



2 Do we have the needed data?

3 Which type of data?

4 Can we trust the data?

5 Is the data sensitive?

1 Which data do we need to make AI work?

6 Can we access the data?



3 Which Type of Data?

STRUCTURED

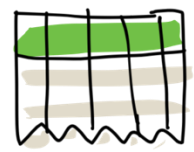


DATABASE

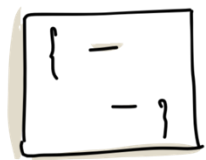


APPLICATION

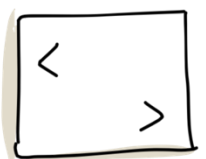
SEMI-STRUCTURED



SPREADSHEET
CSV, TSV

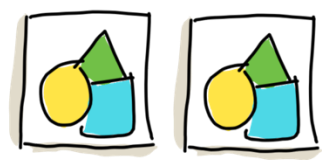


JSON

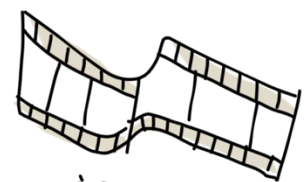


XML

UNSTRUCTURED



IMAGES



VIDEO



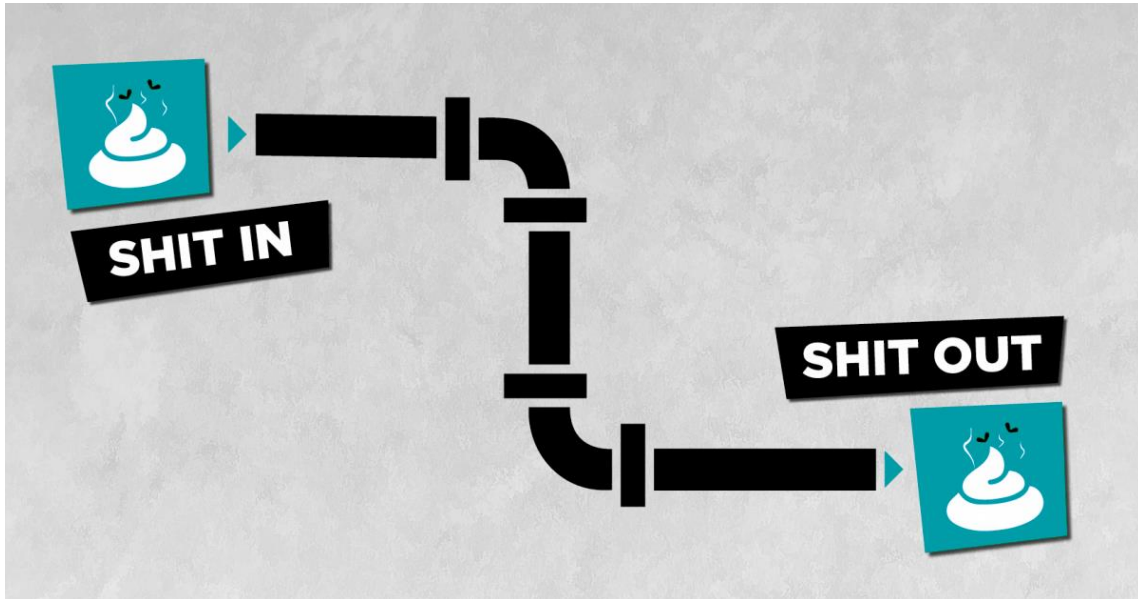
DOCUMENTS





SOUND

4

Can we Trust the Data?



 + Machine Learning = 

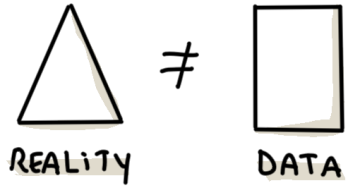
 + Artificial Intellifence = 

 + Generative AI = 

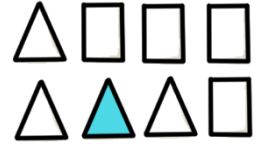
4

Can we Trust the Data?

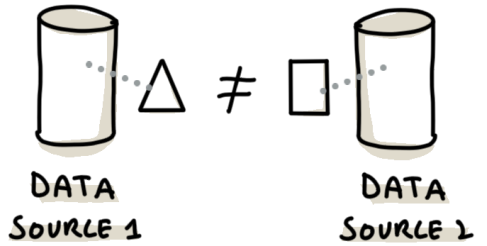
ACCURACY



VALIDITY

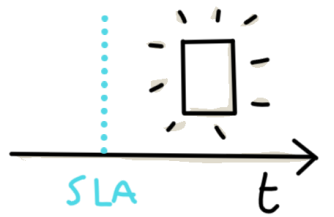


CONSISTENCY

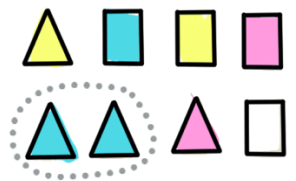


5 DIMENSIONS
OF DATA QUALITY

TIMELINESS



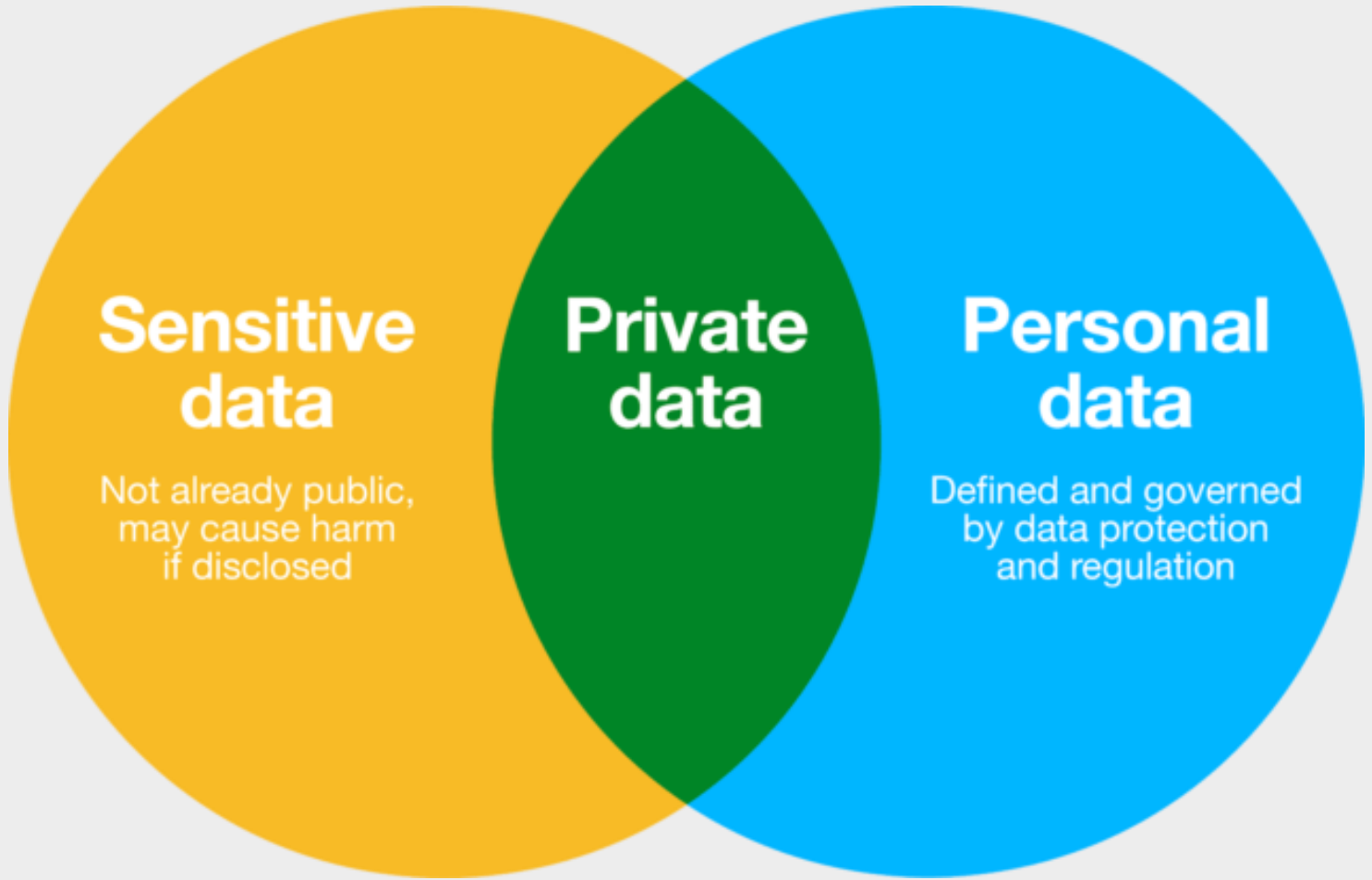
UNIQUENESS





5

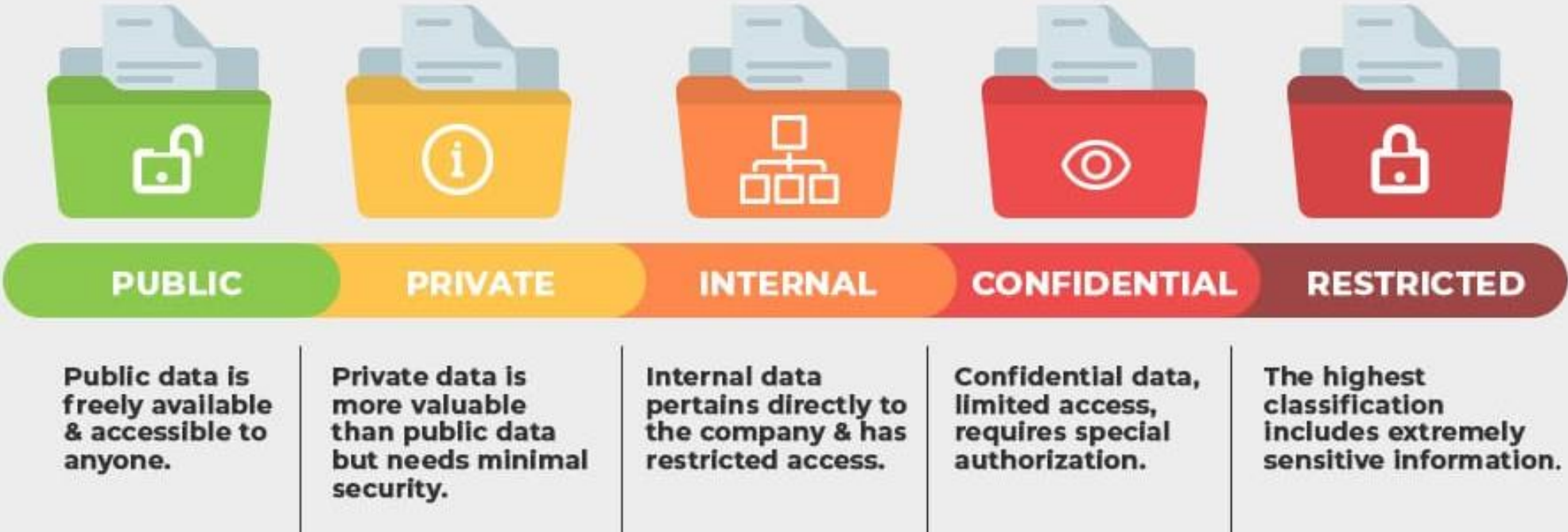
Is the Data Sensitive?






5

Is the Data Sensitive?





5 Is the Data Sensitive?

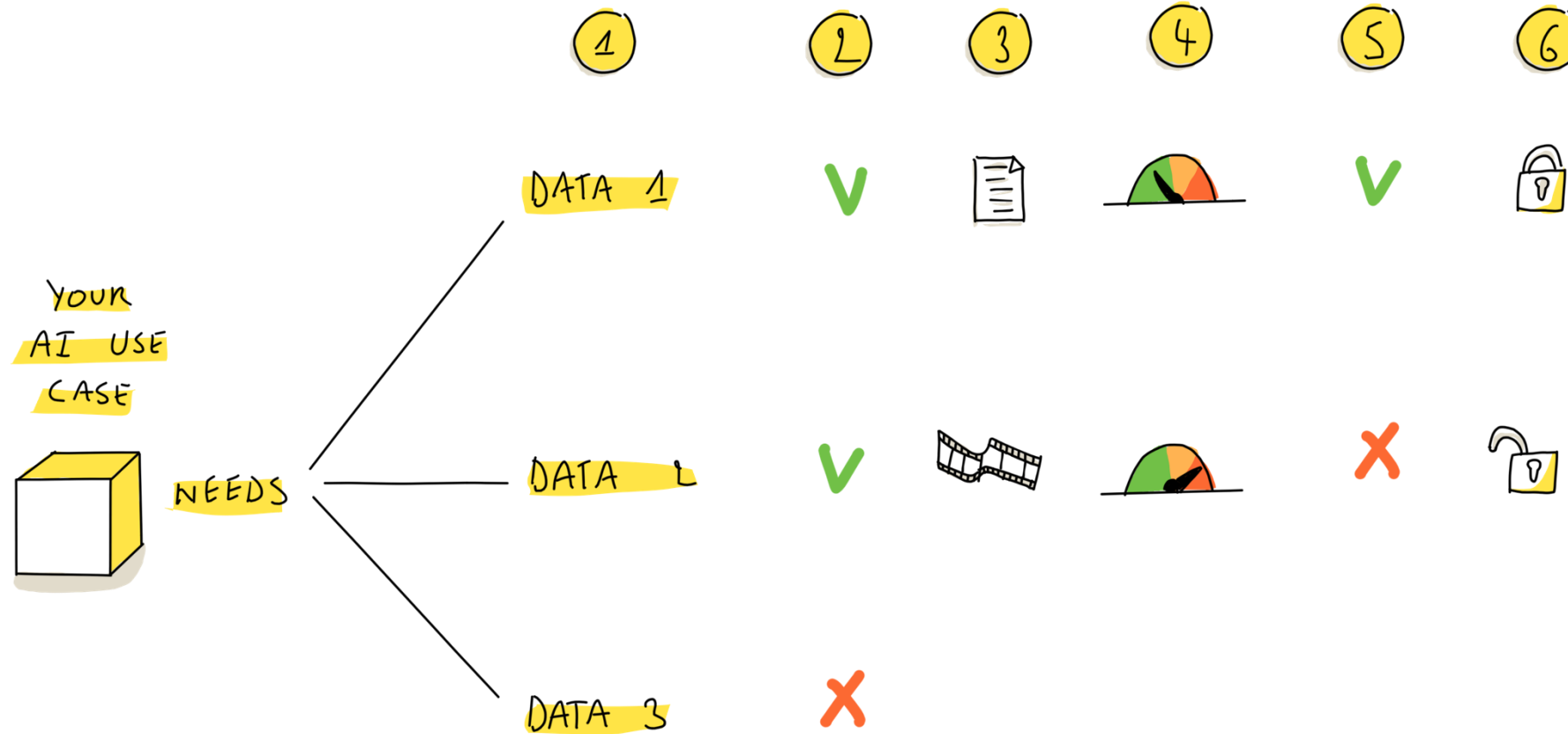
If you can't put your data on 

Don't put it on  as well



Data

- The Bigger Picture
- Questions to ask
- **Exercise**



1 Which data do we need to make AI work?

2 Do we have the needed data?

3 Which type of data?

6 Can we access the data?

4 Can we trust the data?

5 Is the data sensitive?

Exercise: Apply this model to Your Case

