

Today's Data for Tomorrow's Pork Production System

Caleb J. Grohmann, Ph.D.

Data Scientist

Carthage Innovative Swine Solutions, LLC

Carthage Conference 2026

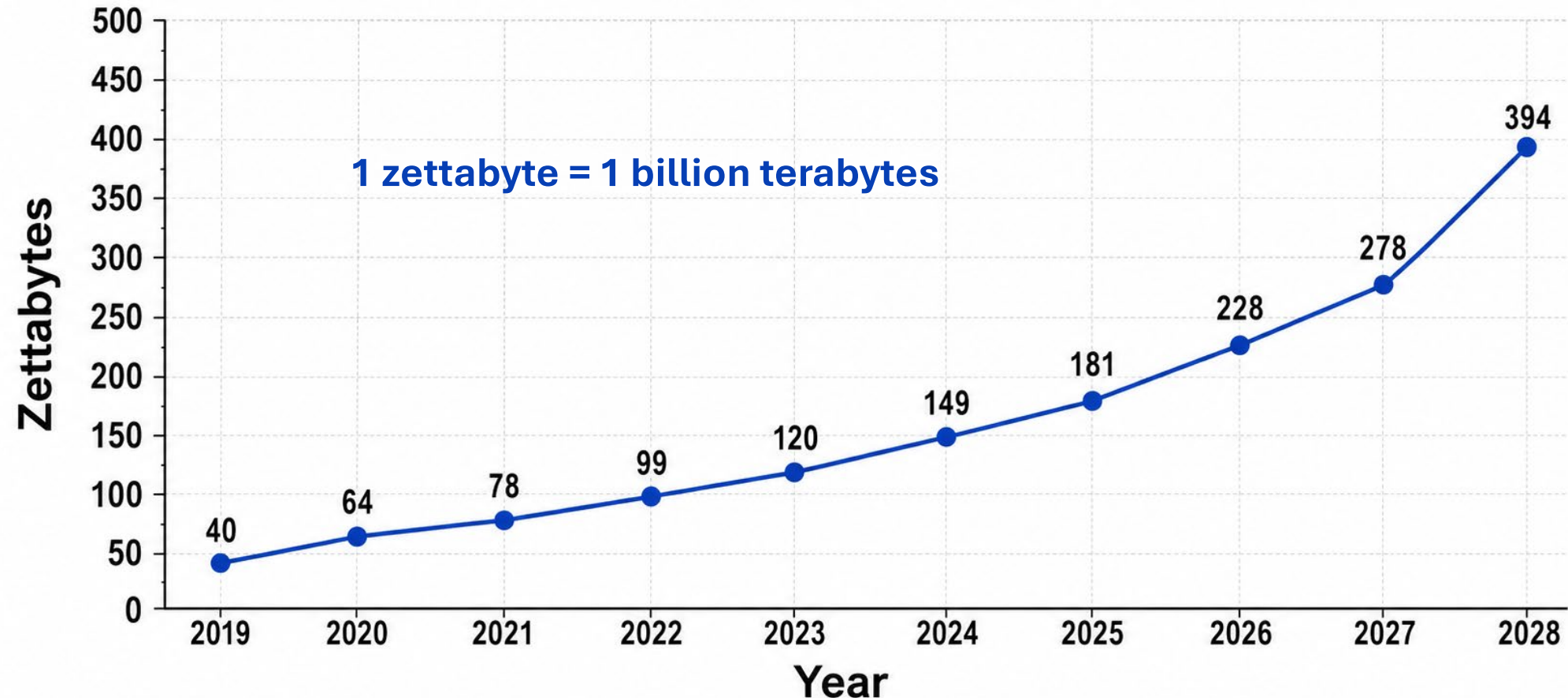
September 1st, 2026



An estimated 90% of the world's data has been generated within the past two years

Growth in Global Data Volumes, 2019–2028

Annual worldwide data created, captured, copied, and consumed



Source: IDC, reproduced in ITU/UNESCO Broadband Commission, State of Broadband 2025.

Advances in technology have enabled a rapid increase in the amount of data available to commercial pig producers

Group Level Production Data

- Inventory
- Movements
- Entry and exit weights
- Feed deliveries
- Market sales
- Mortality
- Treatments

Sensor Data

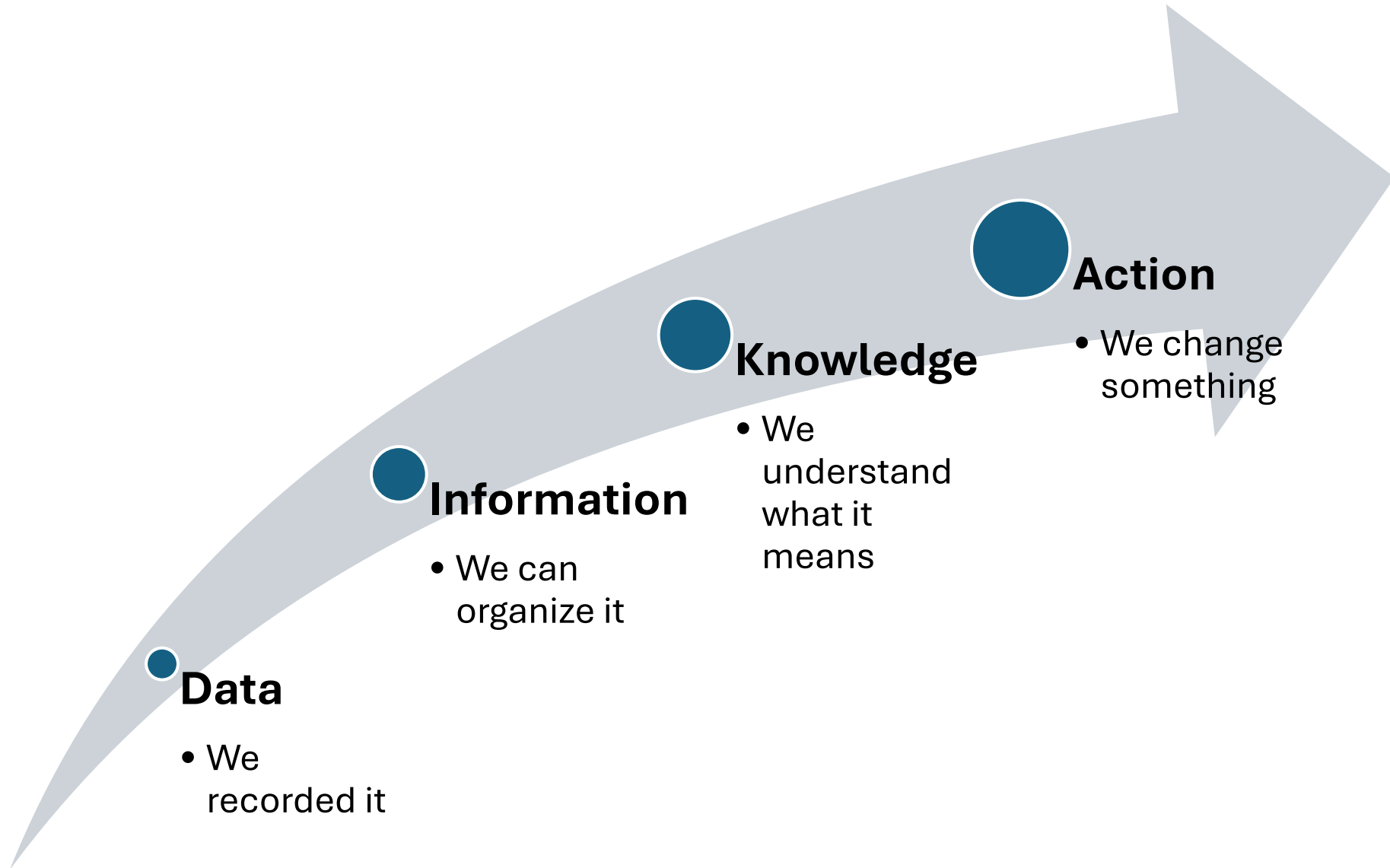
- Temperature
- Water disappearance
- Cough monitors
- Air quality measurements
- Fan runtimes
- Feed bin levels
- Feed line runtimes

Camera Data

- Feeding activity
- Drinking activity
- Injury detection
- Aggression detection
- Weight estimation
- Body temperature monitoring
- Social interaction

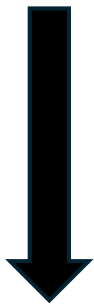


More data does not automatically mean better decisions



The “data ecosystem” in pork production **today** is **disjointed**

**Electronic
Records
Platforms**



**Feed
Management
Systems**



**Barn
Controllers
and Sensors**



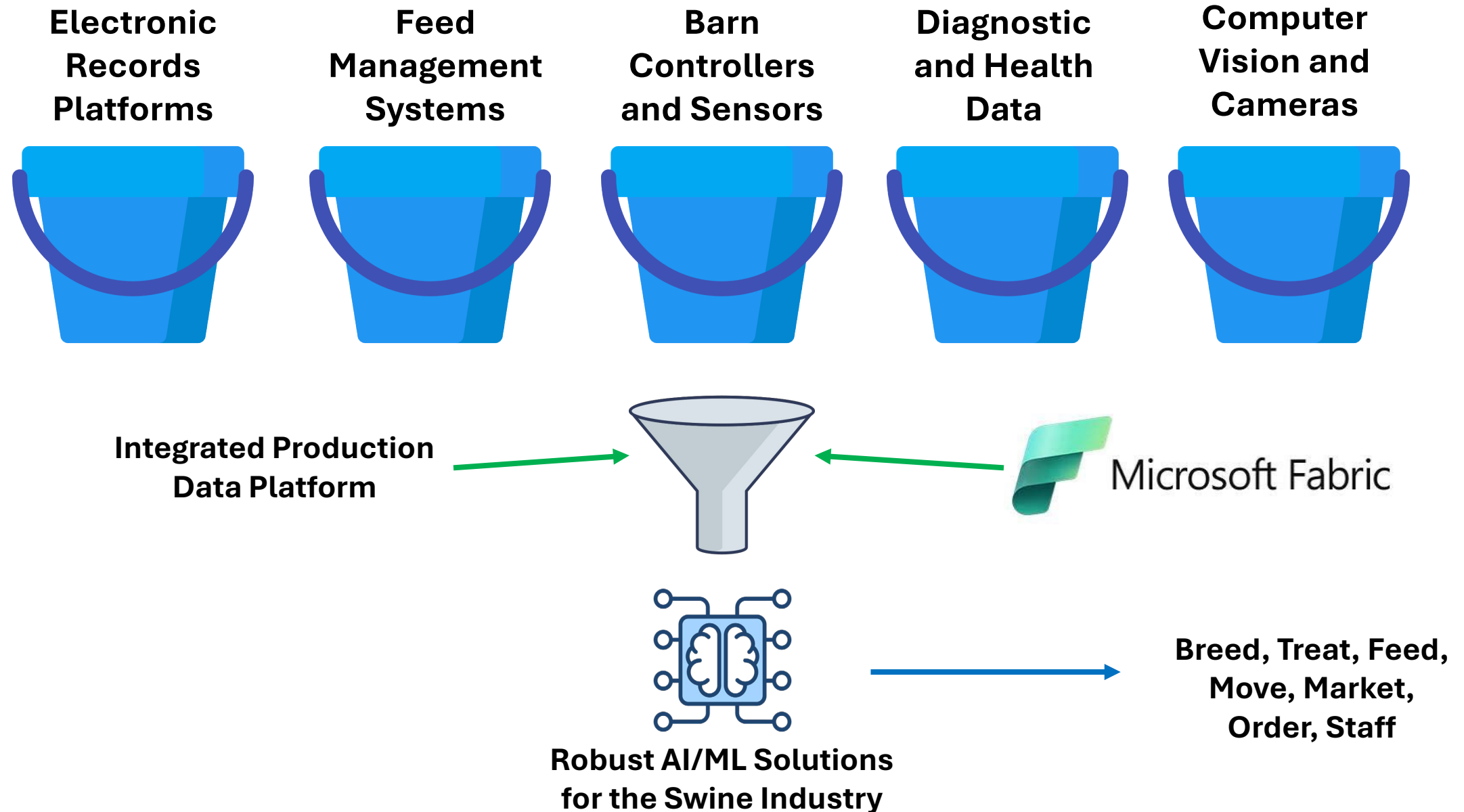
**Diagnostic
and Health
Data**



**Computer
Vision and
Cameras**



The “data ecosystem” in pork production **tomorrow** is **integrated**



Example questions tomorrow's data systems should answer:

- **Finishing**

- What date and at what weight should each load be marketed to maximize profitability?

- **Breed-to-wean**

- How many and which sows should I breed to optimize farrowing rate and reduce non-productive sow days?

- **Health**

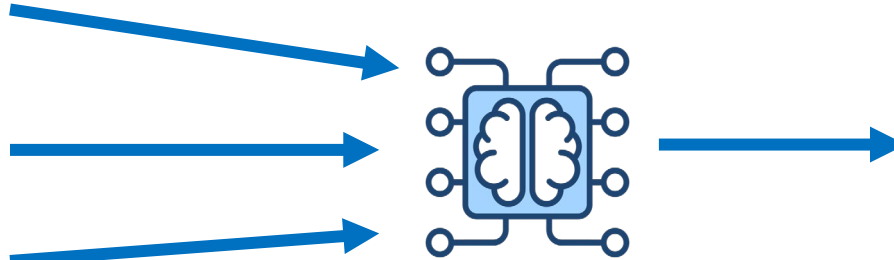
- Is a health event emerging before humans realize it? If, so, how can its impact be mitigated?

Prediction only matters when it changes a decision

Case Study: Forecasting Marketing Outcomes

What we know today

- Starting inventory
- Starting weight
- Feed delivered
- Mortality
- Pig movements
- Weather
- Historical growth patterns
- Current market load weights



Outcomes

- Load by load:
 - Expected weight
 - Expected market date
 - Expected profitability
- As soon as a group is placed on feed

What can production do differently?

- Schedule trucks earlier
- Coordinate packer space
- Optimize barn turns
- Identify groups falling behind
- Reduce overweight/underweight pigs

Garbage In, Garbage Out at Production Scale

Model thinks:

Start weight = 100
pounds

Reality:

Start weight = 55
pounds

Result:

Pigs marketed
too soon

No algorithm can fix inaccurate source data

2000 pigs
recorded, 3000
actually present

Feed delivery
recorded a
month late

Missing latitude
and longitude
values

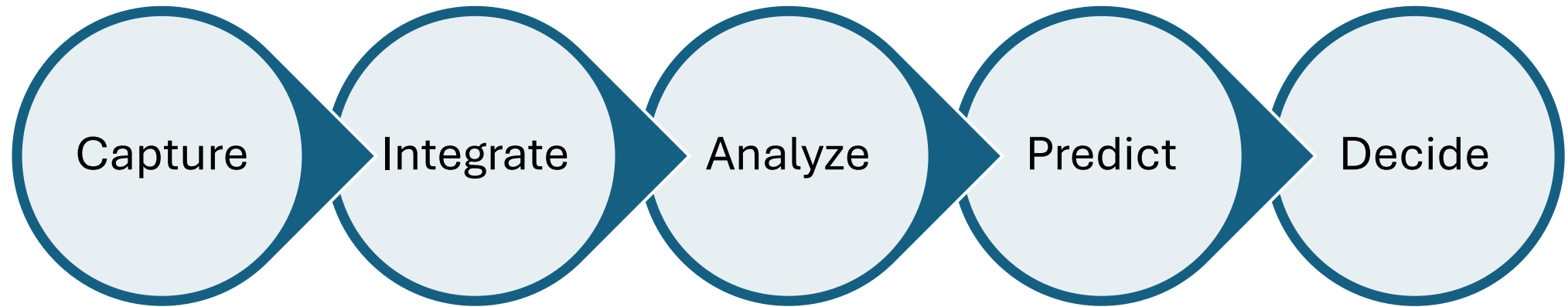
Mortality
recorded
infrequently

Movements
assigned to
wrong group

A stylized illustration of a pink piggy bank. The piggy bank is light pink with darker pink legs and a small, curled tail. It has a simple black dot for an eye and a small ear.

The diagram illustrates a decision tree for the 'Salary' variable. The root node is labeled 'True' and splits based on 'Age' (0.681). The left branch (True) leads to a node with 'EstimatedSalary < 0.006' (gini = 0.253, samples = 215, value = [175, 46], class = 0). The right branch (False) leads to a node with 'Age < 0.825' (gini = 0.275, samples = 85, value = [54, 71], class = 1). The tree continues to split based on 'Age' and 'EstimatedSalary' until it reaches leaf nodes. The nodes are color-coded: orange for 'True' and blue for 'False'. Each node contains the following information: the split variable and threshold, the gini index, the number of samples, the value distribution, and the predicted class.

Tomorrow's production system starts with today's data



Questions?

Contact: cgrohmann@hogvet.com