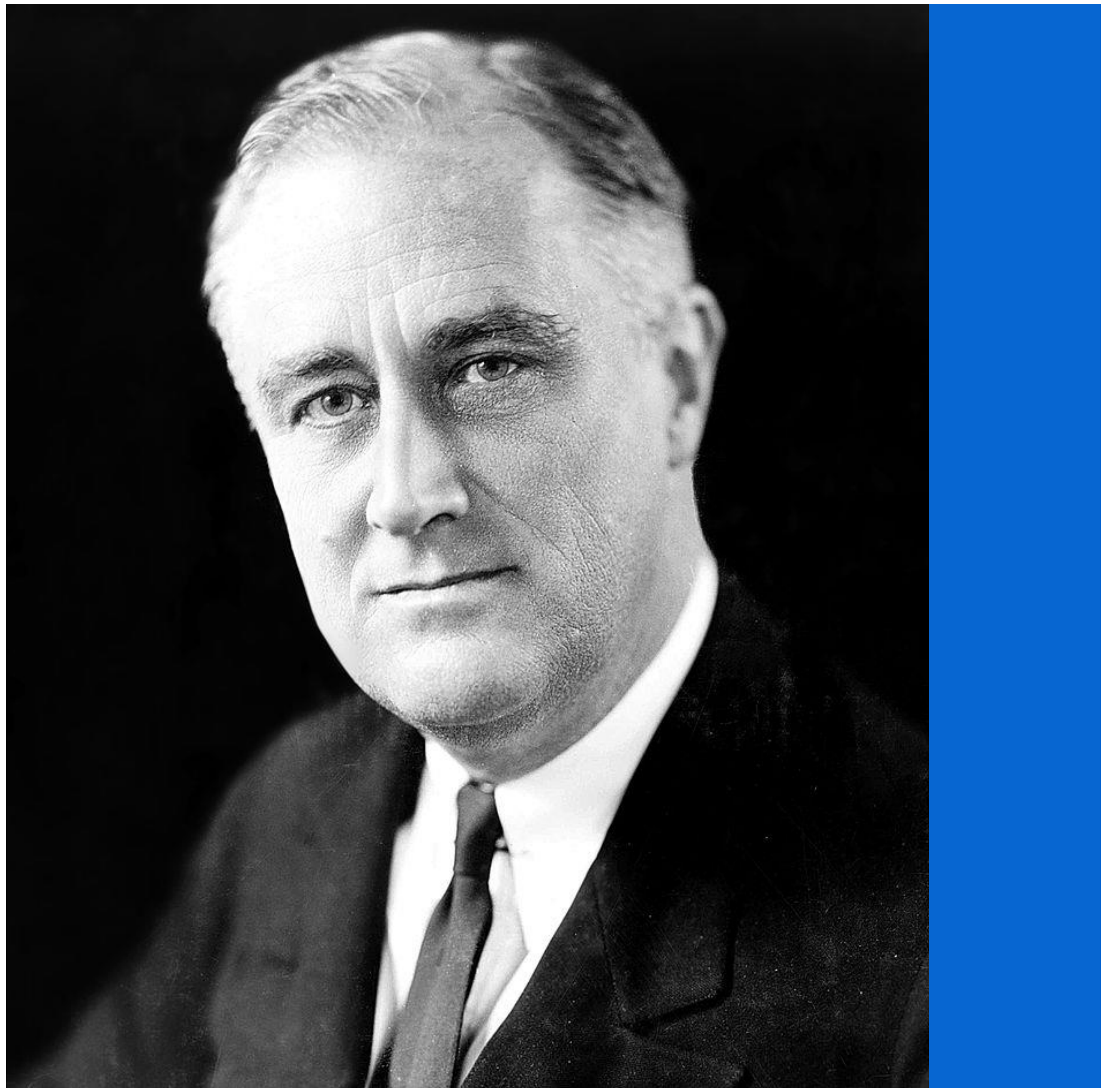
How to Predict Elections and Combat Automated Liars

(for Fun and Profit)

Simon Edwards
Head of AI Platform Engineering and Architecture
www.linkedin.com/in/simonjamesedwards

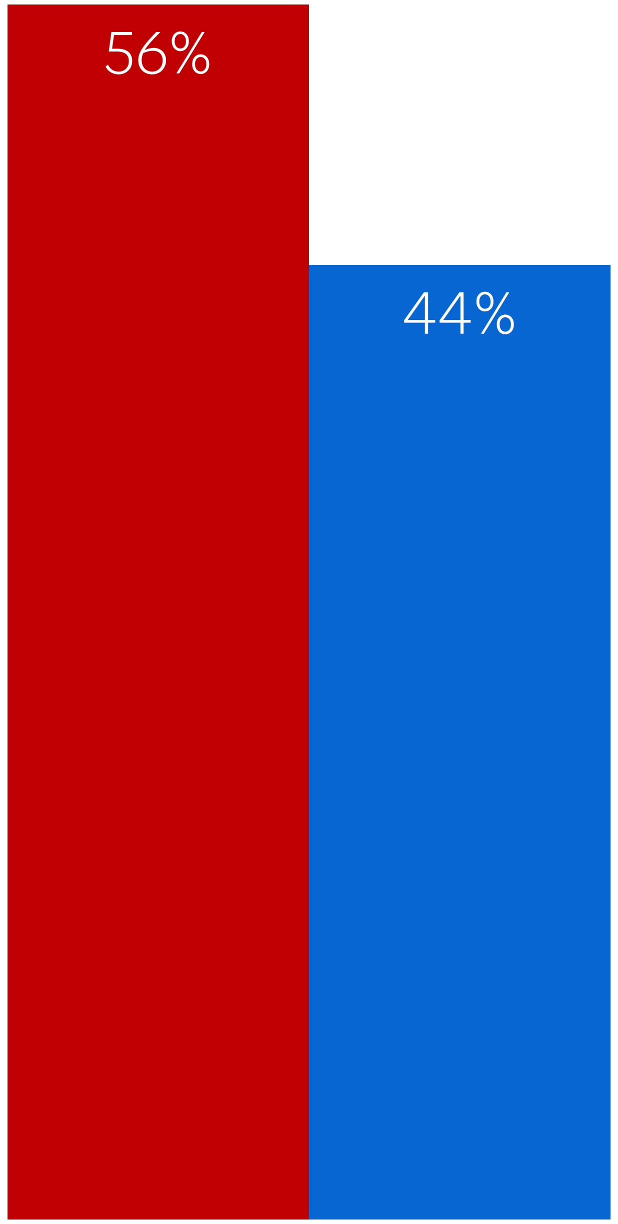
How to Predict an Election

1936 US presidential election



How to Predict an Election

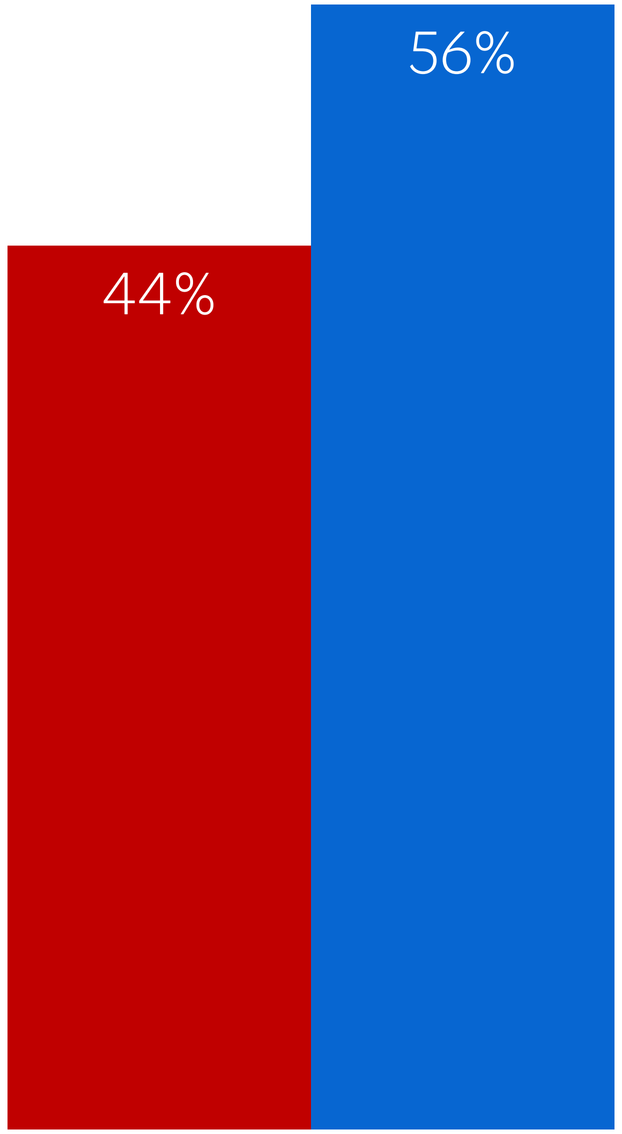
Prediction 1

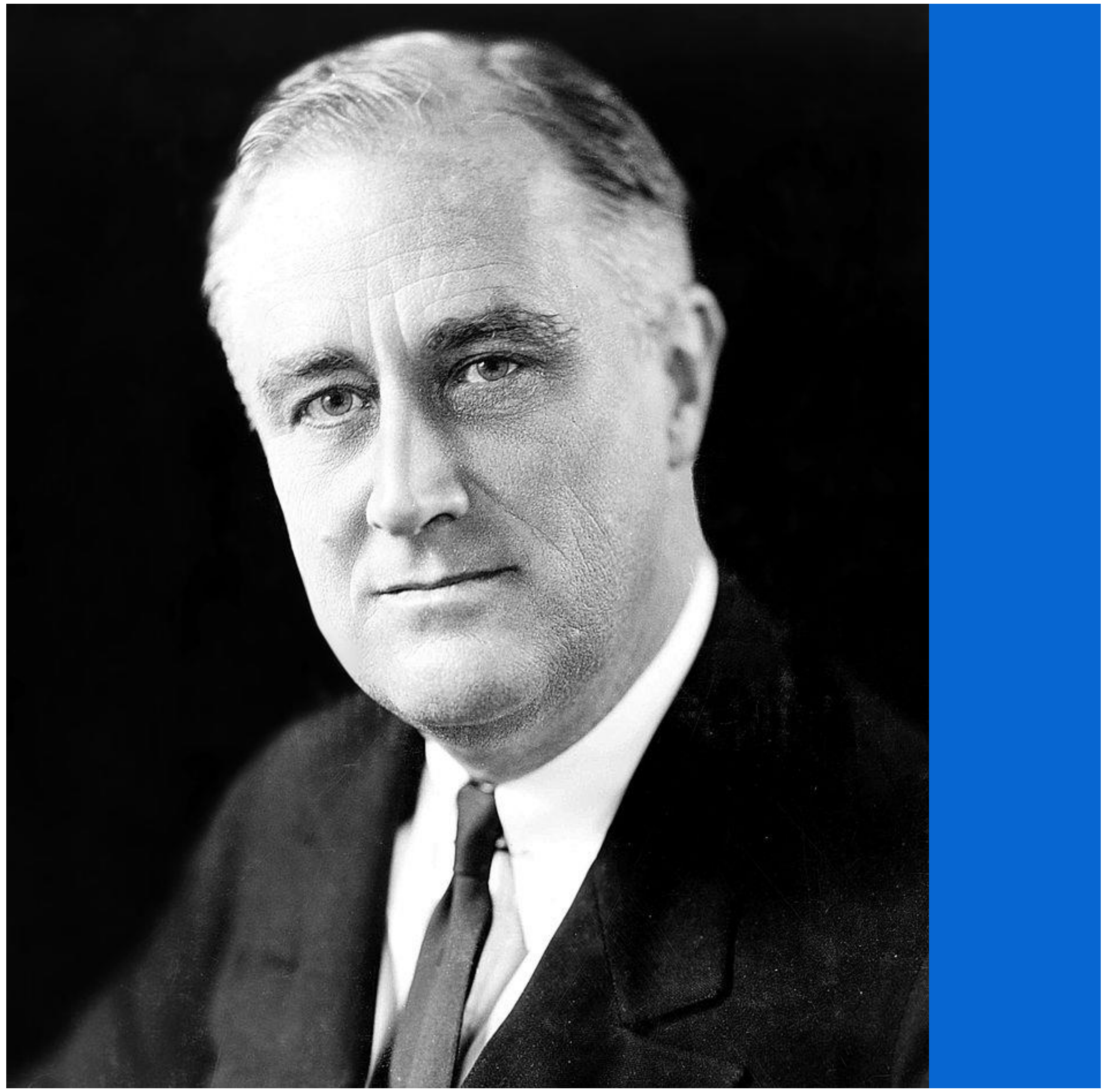


2,380,000
respondents

How to Predict an Election

Prediction 2



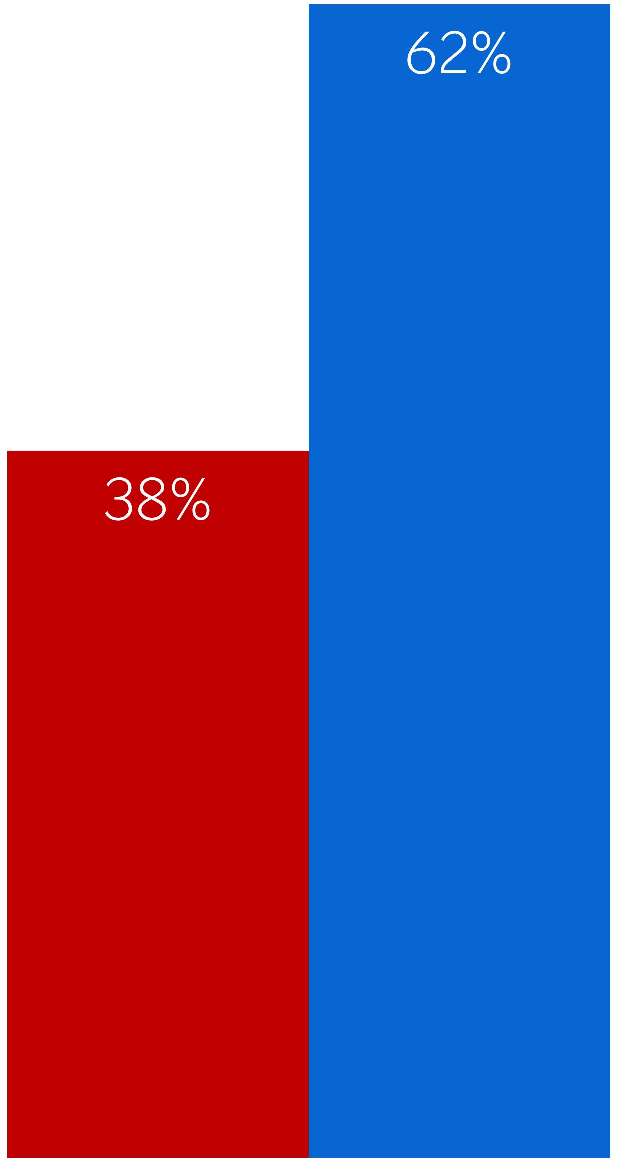


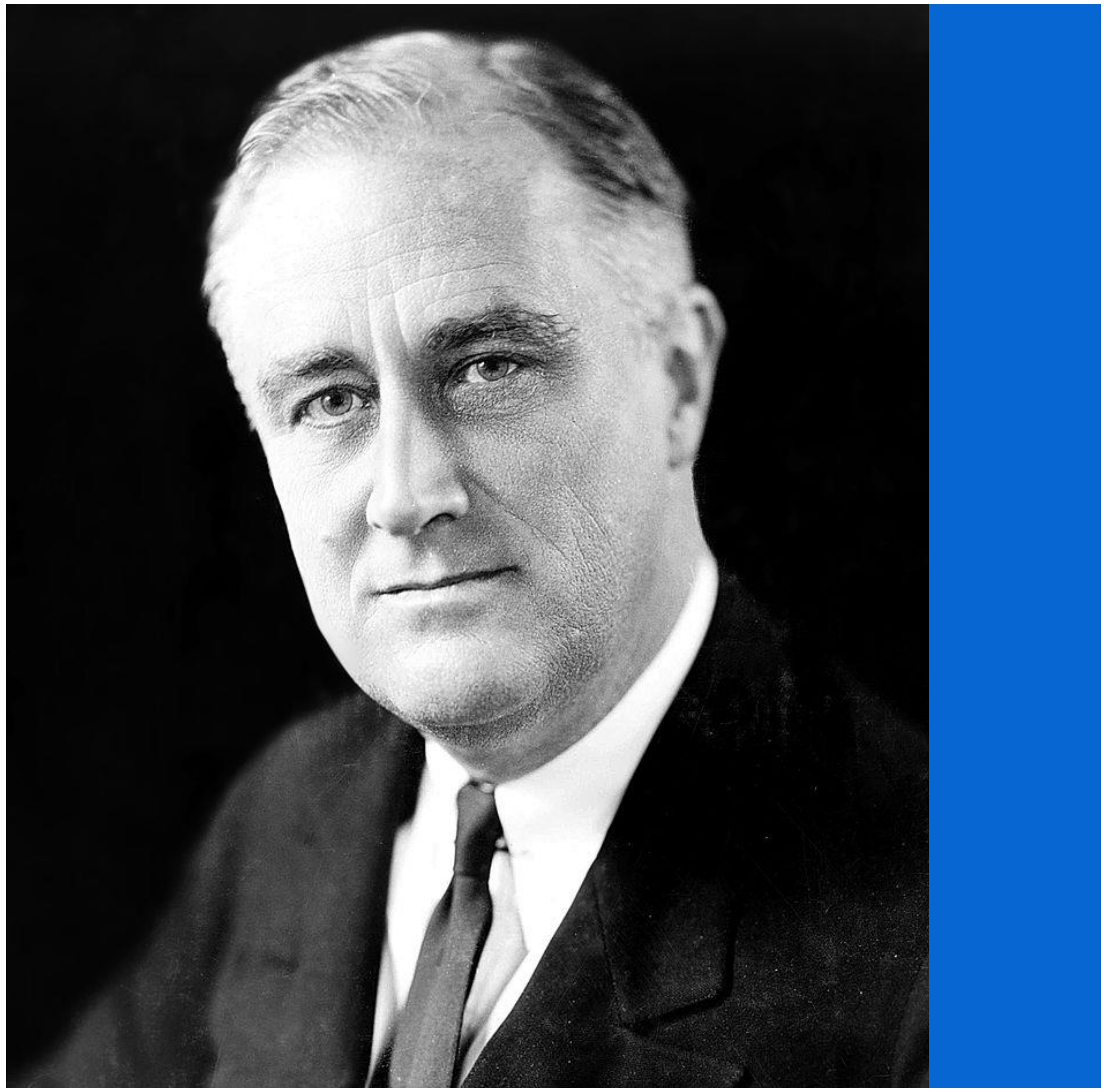
50,000
respondents

How to Predict an Election

Alf Landon

Actual Result

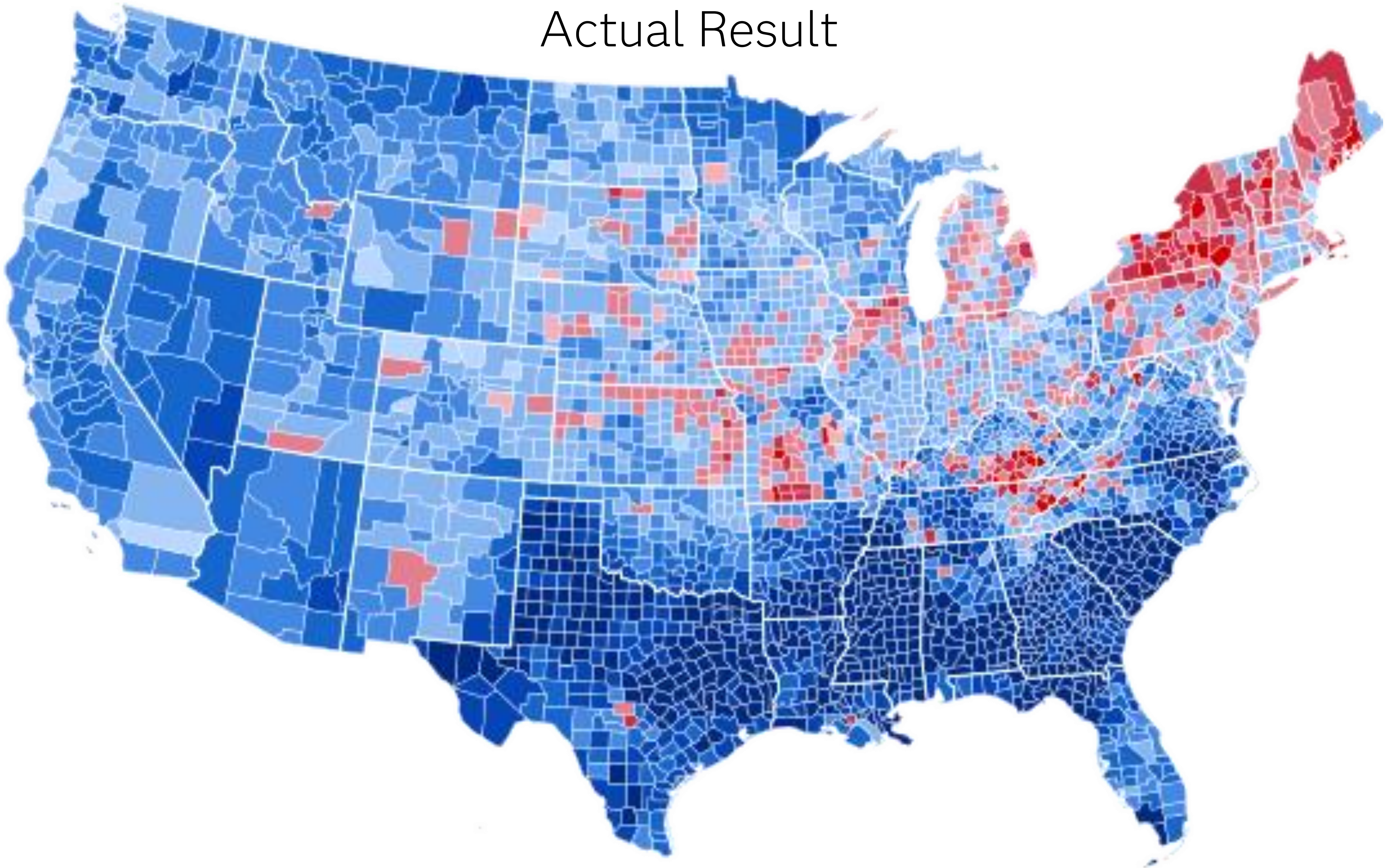




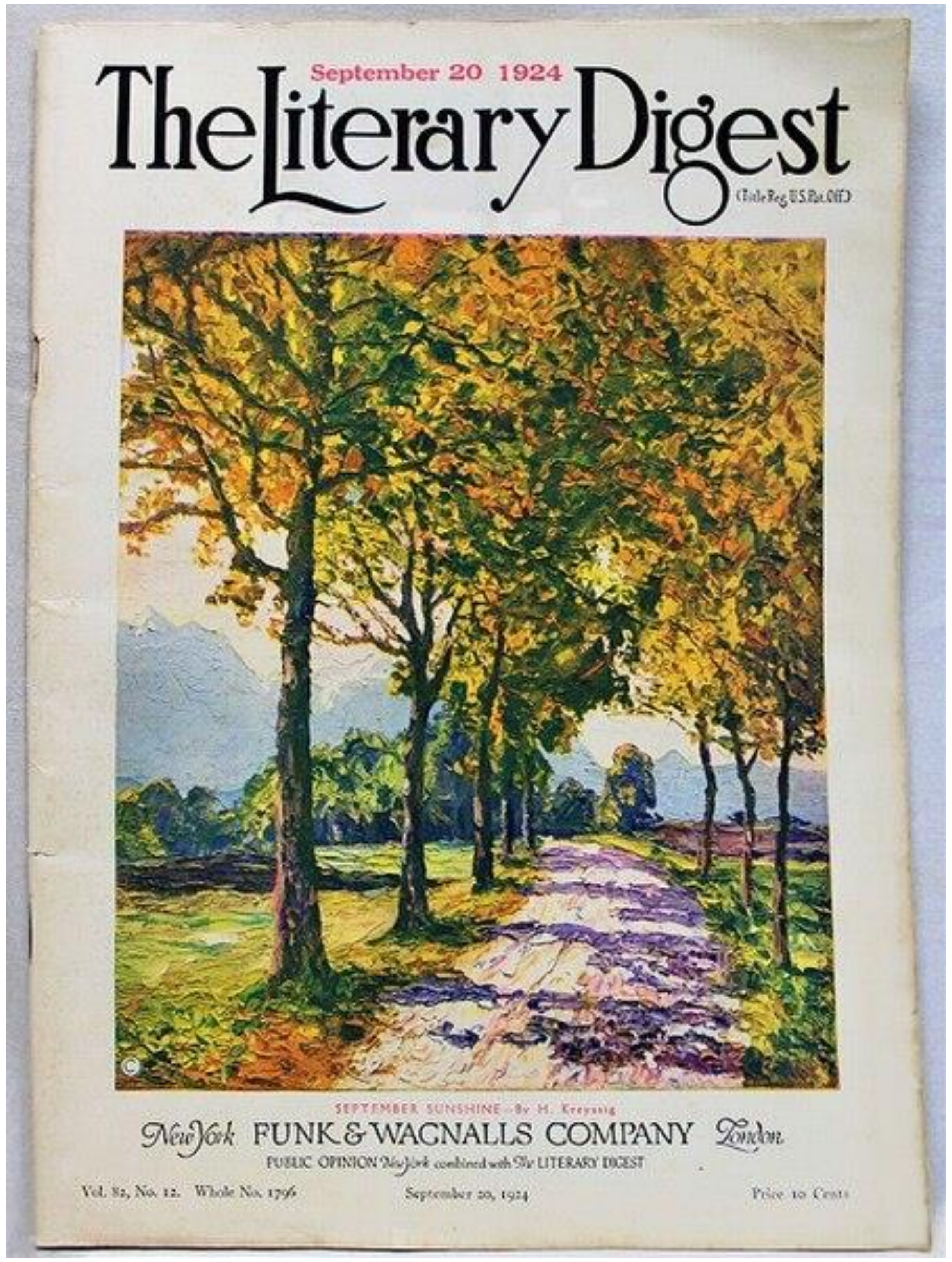
Franklin D. Roosevelt

How to Predict an Election

Actual Result



How (Not) to Predict an Election



Failed prediction #1 was from a poll run by the Literary Digest.

Sample size isn’t everything.

Recommended reading:

- https://www.britannica.com/topic/public-opinion/Allowance-for-chance-and-error#ref913886
- https://en.wikipedia.org/wiki/The_Literary_Digest

How to Predict an Election

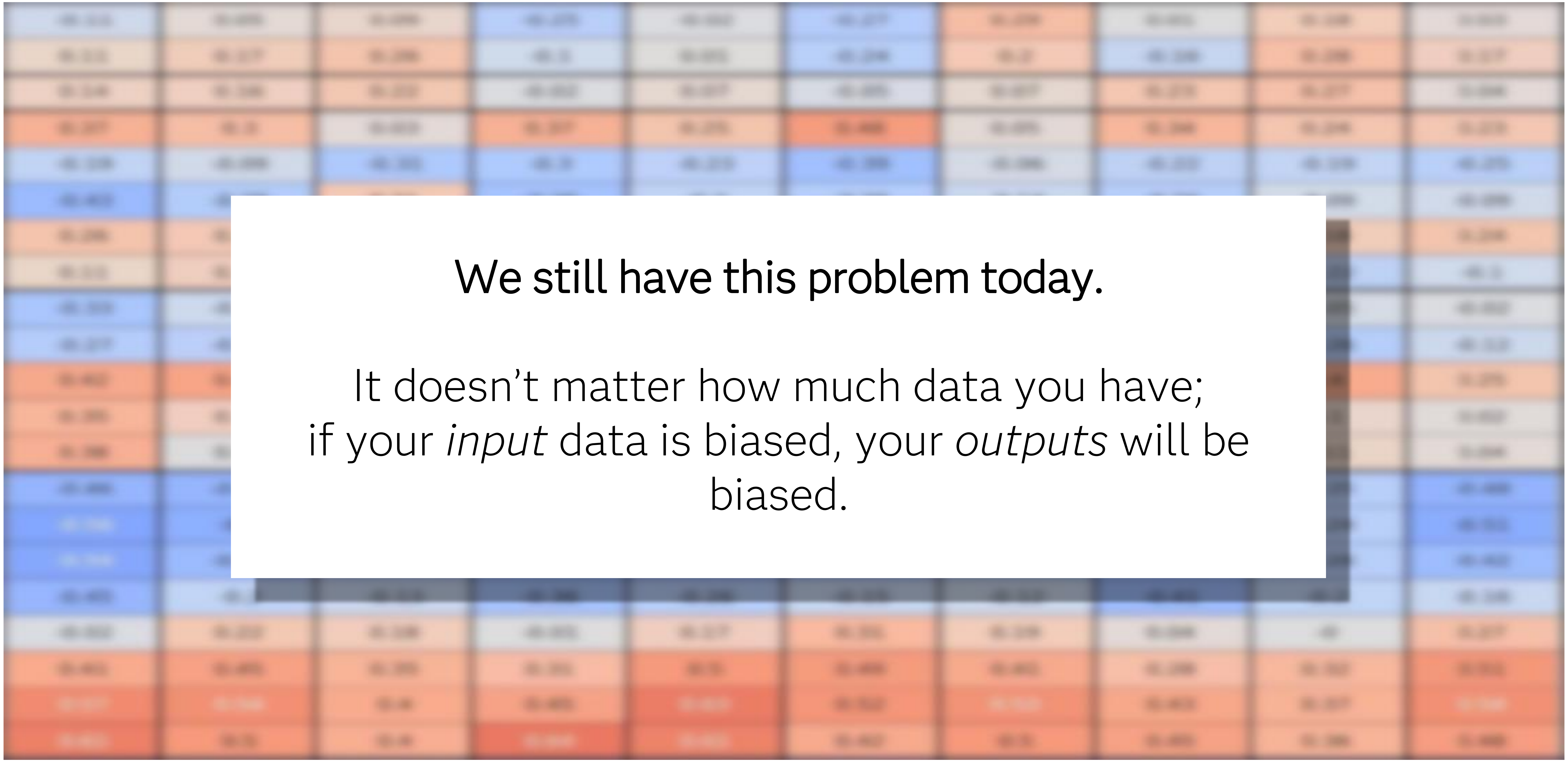
Prediction #2 was from a poll run by **George Gallup**.

George Gallup was a talented statistician who:

- Understood **basic statistical bias**
- Understood **basic sampling techniques**
- **Correctly predicted** the actual result to **within 1.4%**

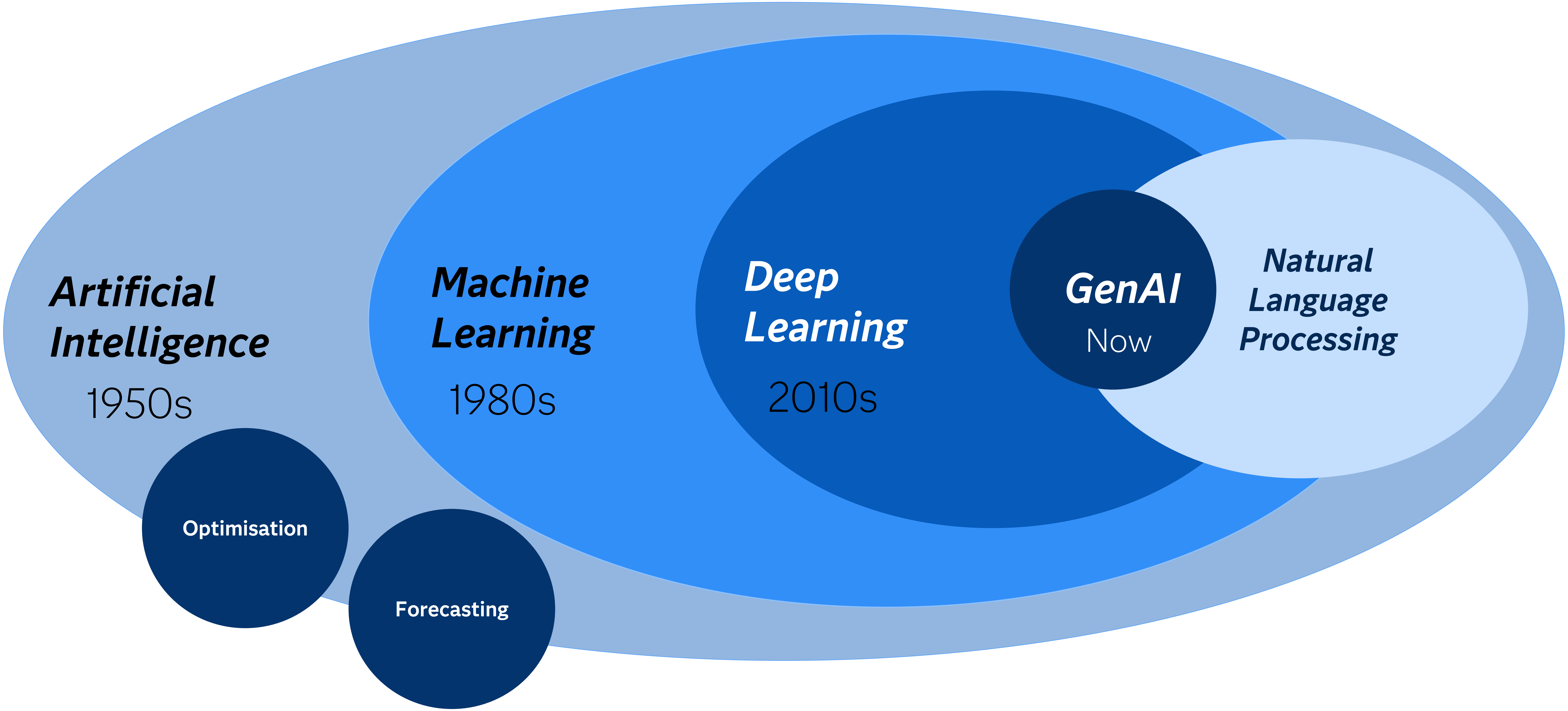
This was **career making** for George Gallup.

How to Predict an Election



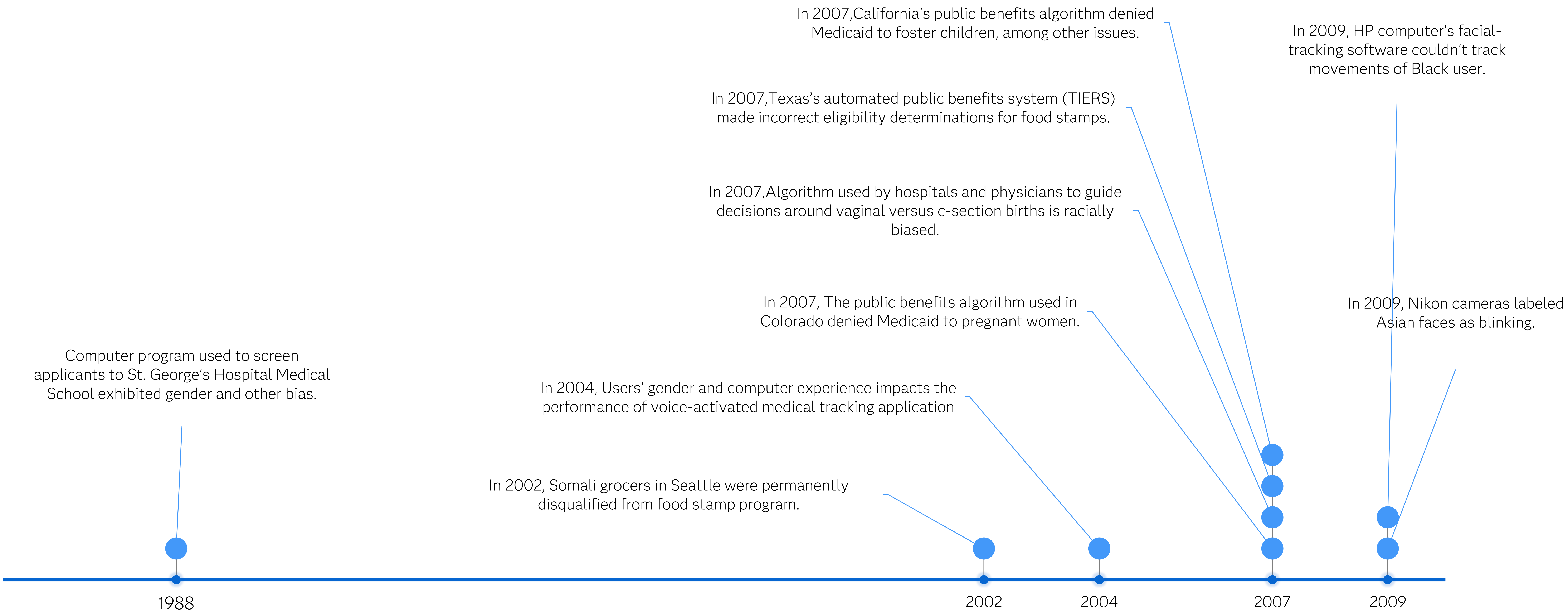
We still have this problem today.

It doesn’t matter how much data you have;
if your *input* data is biased, your *outputs* will be biased.



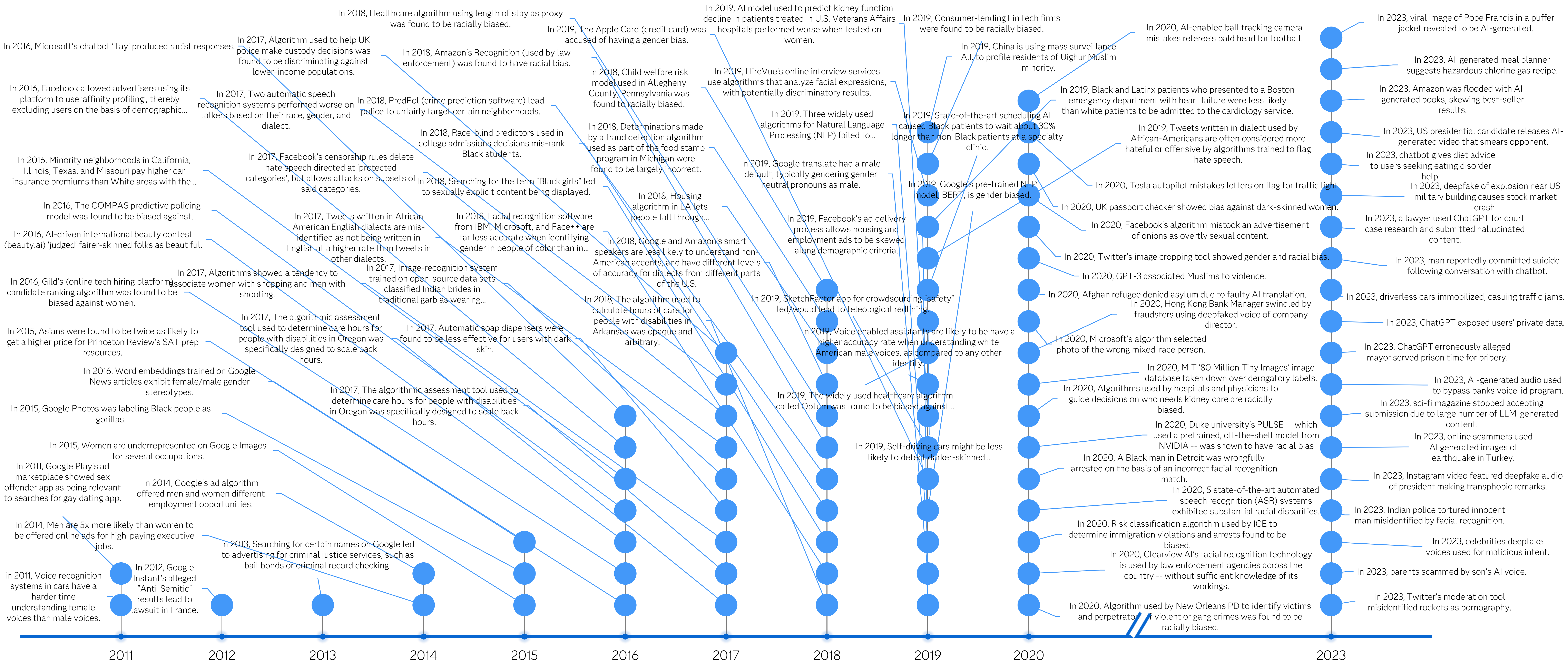
Benefits are huge, but AI and models are only as good as the data they are built on.
If data is biased, models will be biased, and the results will be biased.

Whoops...



Source: BERKELEY HAAS CENTER FOR EQUITY, GENDER & LEADERSHIP Examples of Bias in Artificial Intelligence

...WHOOPS



Source: BERKELEY HAAS CENTER FOR EQUITY, GENDER & LEADERSHIP Examples of Bias in Artificial Intelligence

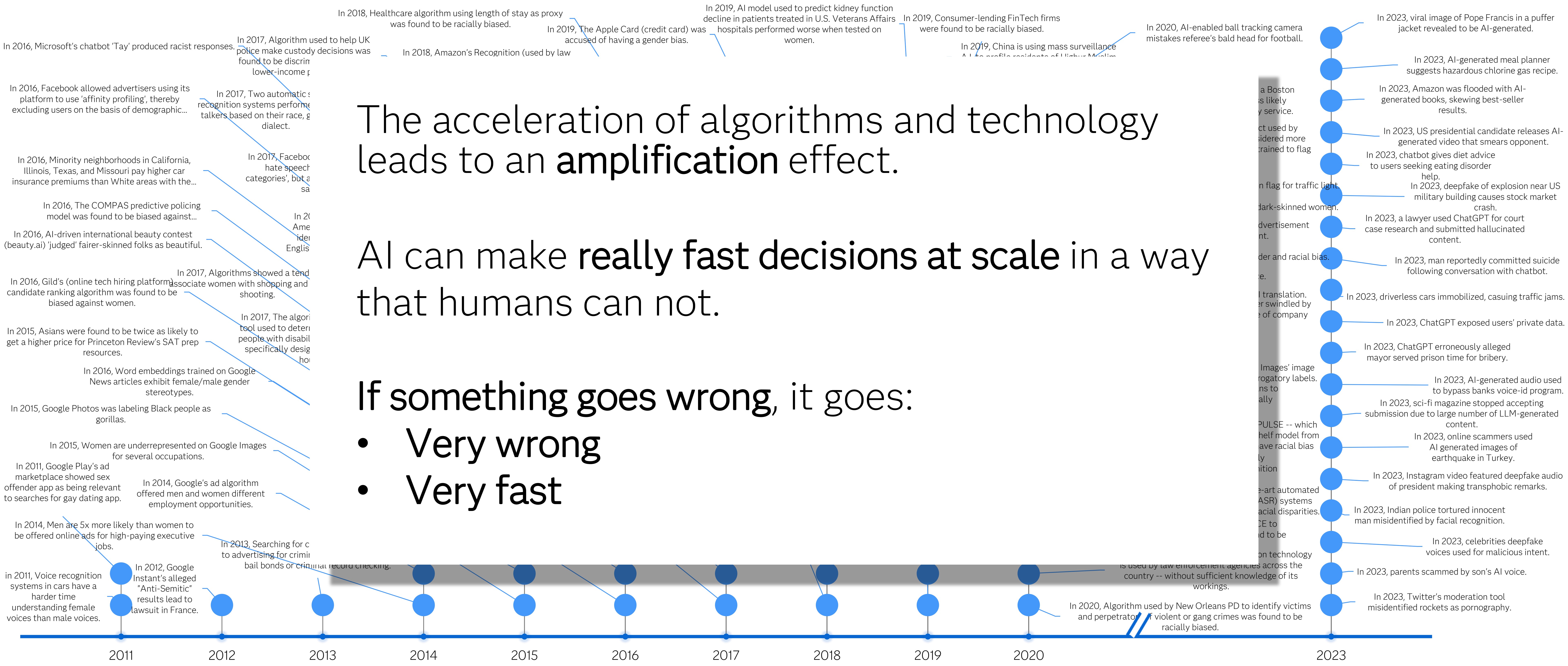
...WHOOPS

The acceleration of algorithms and technology leads to an **amplification** effect.

AI can make **really fast decisions at scale** in a way that humans can not.

If something goes wrong, it goes:

- **Very wrong**
- **Very fast**



Source: BERKELEY HAAS CENTER FOR EQUITY, GENDER & LEADERSHIP Examples of Bias in Artificial Intelligence

Nobody wants to make headlines

ABC NEWS

This US lawyer used ChatGPT to research a legal brief with embarrassing results. We could all learn from his error

ABC RN / By Damien Carrick and Sophie Kesteven for the Law Report, with additional reporting from Reuters.

Posted Sat 24 Jun 2023 at 11:15am

IBM Abandons Facial Recognition Products, Condemns Racially Biased Surveillance

June 9, 2020 · 8:04 PM ET

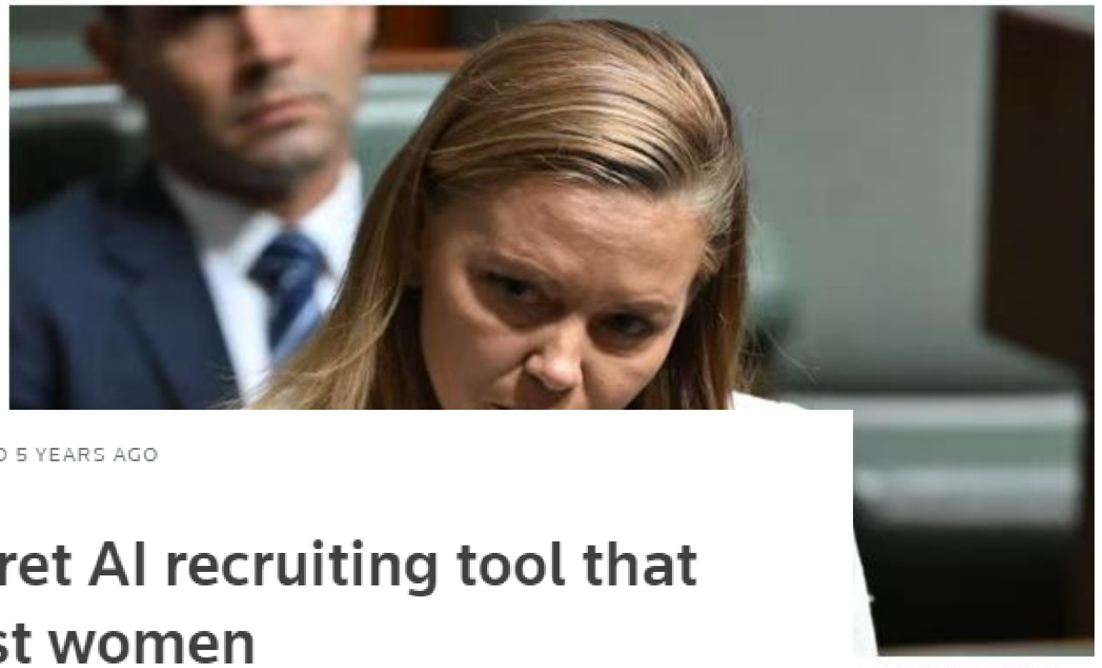
'There is no standard': investigation finds AI algorithms objectify women's bodies

Guardian exclusive: AI tools rate photos of women as more sexually suggestive than those of men, especially if nipples, pregnant bellies or exercise is involved

Liberal backbencher Zoe McKenzie says robodebt scheme caused 'avoidable human suffering'

Victorian MP joins Bridget Archer and Keith Wolahan in criticising Coalition's Centrelink debt recovery scheme

- Follow our Australia news live blog for latest updates
- Get our morning and afternoon news emails, free app or daily news podcast



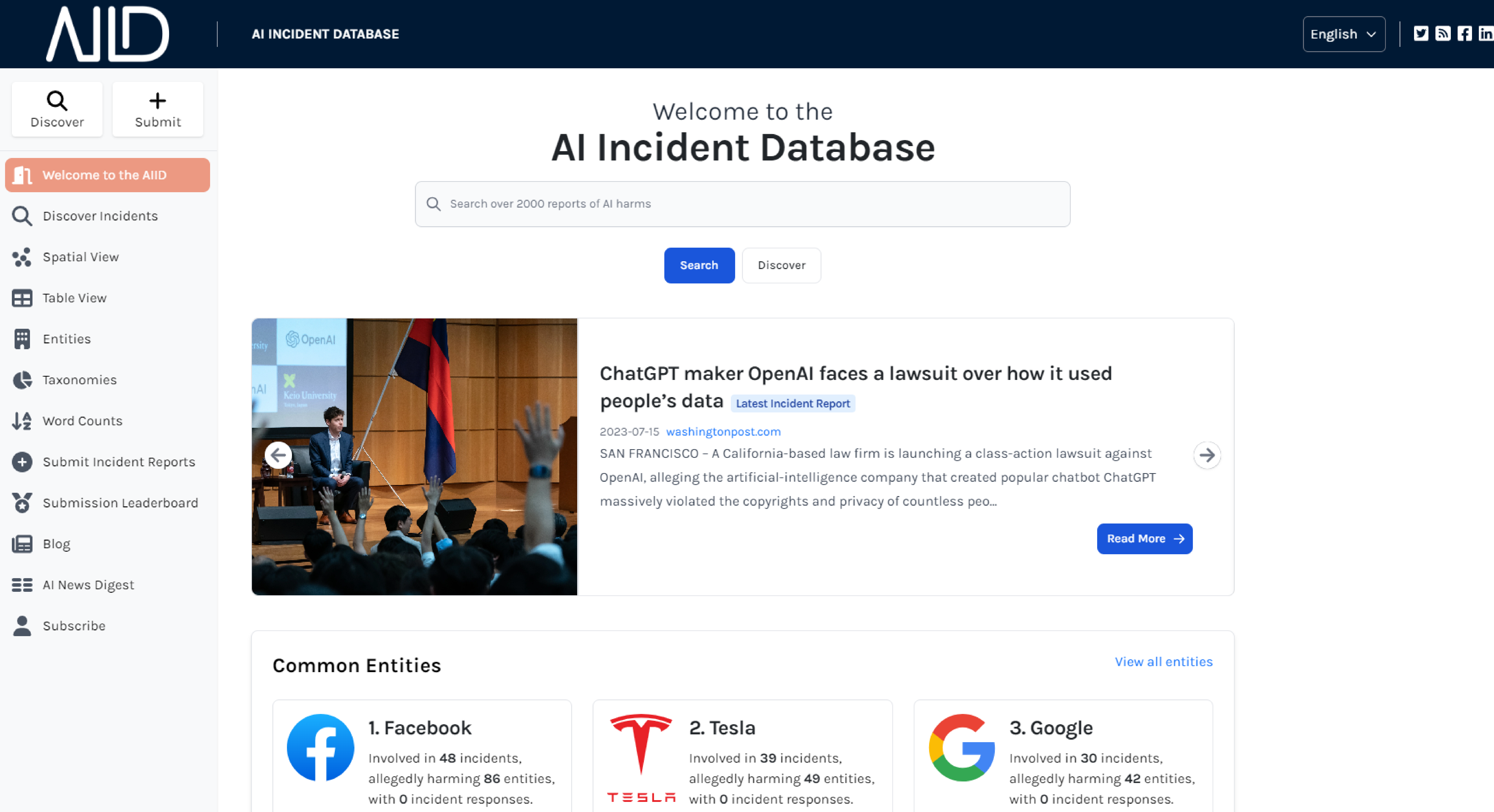
n ... paints a picture

RETAIL OCTOBER 11, 2018 / 10:04 AM / UPDATED 5 YEARS AGO

Amazon scraps secret AI recruiting tool that showed bias against women

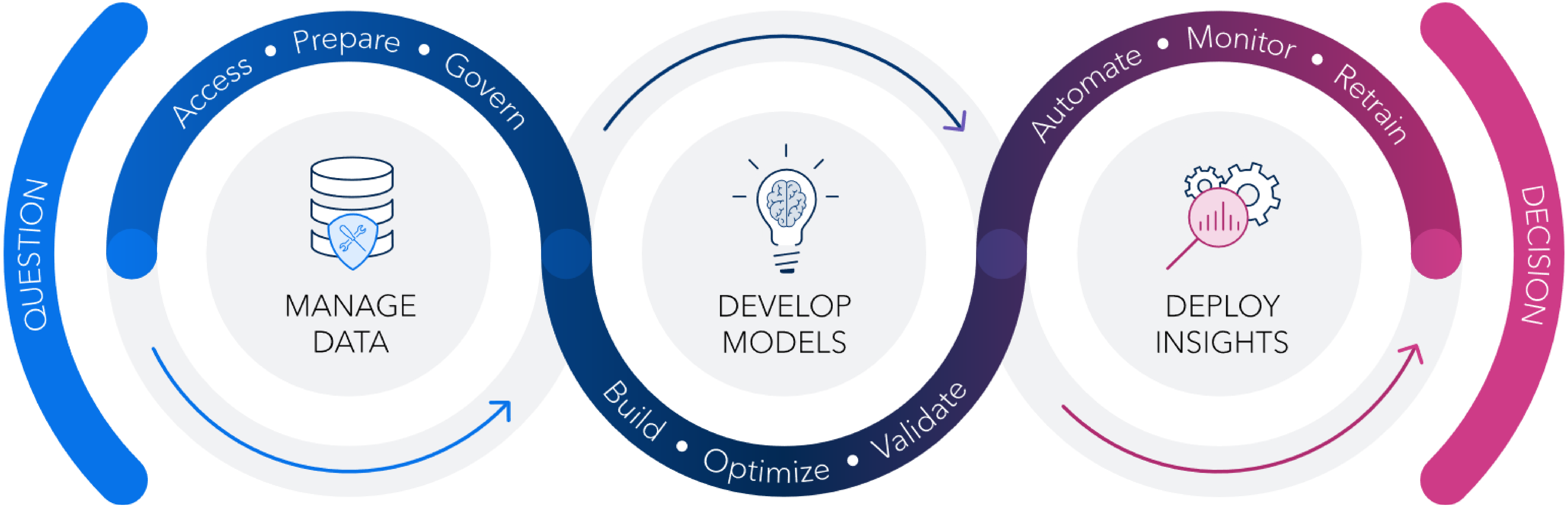
By Jeffrey Dastin 8 MIN READ

SAN FRANCISCO (Reuters) - Amazon.com Inc's AMZN.O machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.



Source: https://incidentdatabase.ai

Data & AI Lifecycle: Bias Everywhere



Availability Bias
Recall Bias

Exclusion Bias
Pre-processing Bias
Measurement Bias
Time-interval bias
Historical Bias
Sample Selection:
Selection Bias
Attrition Bias
Proxy Bias

Confirmation Bias
Cause/ Effect Bias
Confounding Bias
Collider Bias
Prediction Bias
Performance Bias
Hindsight Bias
Chronological Bias
Funding Bias

Automation bias
Deployment Bias
Drift Bias
Reporting Bias

Accepting the Truth and Finding Solutions

ALL real data is biased.

This is a problem that has no <u>mathematical</u> solution.

You can't get rid of bias, only mitigate against it.

Use statistical techniques!

How to Avoid Creating Automated Liars

1. Question and understand your data

2. Use the right data for the right purposes

3. Monitor and Modify your usage as you go

4. Get the whole company on board

1. Question and understand your data

Quality and qualities

- How clean is your data? Does it contain errors?
- How old is your data?
- Where and how was it recorded?
- Are there similar sources of data available for comparison?

Quantities

- Do you have enough data for what you want to use it for?
- Should you consider supplementing it?
- Do you need to collect more before you can start?

Trusted source? Lineage?

- Do you trust the source of your data?
- Where and how is the data going to be used?

Diversity

- Is the data diverse enough?
- What groups of people might be missing from the data?
- What happens if minority groups an under-represented?
- What happens if some groups an over-represented? (more vocal)
- Maybe you are missing people who dropped off before purchasing?
- Maybe missing customer data for declined loans because they never became customers?
- Homeless people missing from census data?

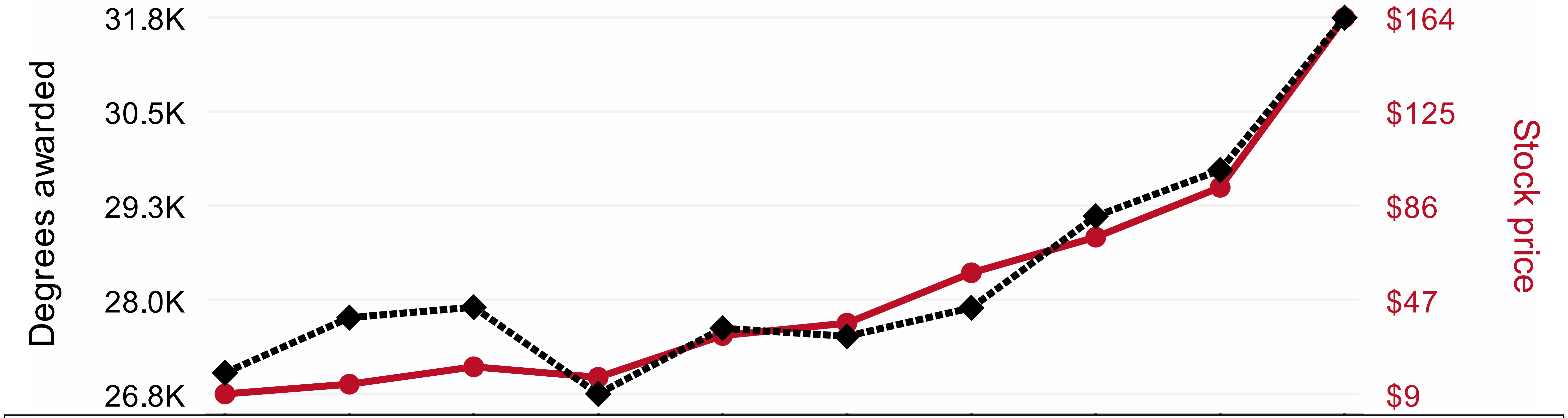
Semantics and labelling

- How have you labelled your data? (very important for GenAI)
- Is the meaning of it clear?
- Can you sense test it with a subject matter export?

Master's degrees awarded in Psychology

correlates with

Amazon.com's stock price (AMZN)



"If you torture the data long enough, it will confess to anything."

Ronald Coase, Economist (Nobel Prize in Economics 1991)

Source: https://tylervigen.com/spurious/correlation/2539_masters-degrees-awarded-in-psychology_correlates-with_amazoncoms-stock-price

2. Use the right data for the right purposes

(and vice versa)

Work with what you have:

- Don't let your data scare you off - but make sure you've questioned and understood it first.
- If you understand your data, you can **correct** for bias as you go.

Right data for the model

- AI and analytics use cases and techniques dictate what data you need for them
- E.g. Generative AI requires different data inputs to a predictive model.
- Horses for courses!

Make it the right data

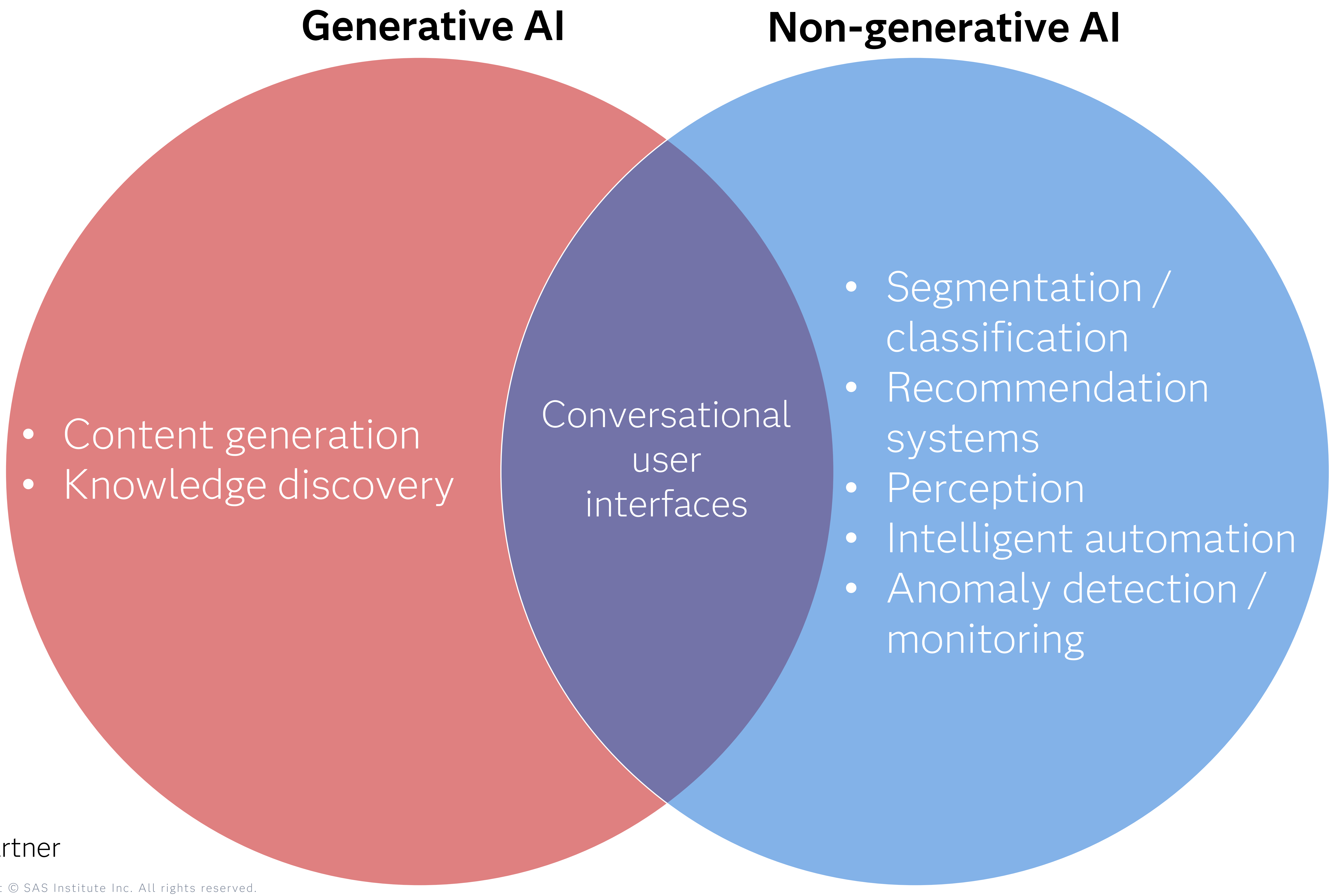
- If you haven't got the right data, make it the right data.
- Consider how you are going to deal with outliers and missing values
- Consider using techniques like boosting samples, stratifying test data, ensemble modelling, and synthetic data if needed.

"Fairness" varies in definition

- Consider fairness, based on different relevant definitions of fairness.
- What's fair in one context might not be what's fair in another. Sometimes you need to give minority groups a boost in your data to make sure they don't get lost. Sometimes you need to counter-correct for a vocal minority. It depends on the context.

Use the right models for the right purposes

- Understand where an AI technique or model technique's strengths and weakness lie
- Horses for courses
- Don't use LLMs and GenAI for things like numerical analysis, predicting trends, predicting elections!



Source: Gartner

3. Monitor and Modify your usage as you go

Implement versioning

- Making sure use of your data is versioned and revertible to an older version if possible
- Important because models can drift or break
- You might need different versions of data or models for different scenarios

Audit usage

- Audit usage of data
- Audit usage of models
- Security and Privacy concerns need to be covered

Watch for inference and derivation

- I.e, Where output of one model can be used as the input to another
- Valid but it can amplify issues – see GenAI
- Track how your models are combined and whether that combination is acceptable
- Governance important

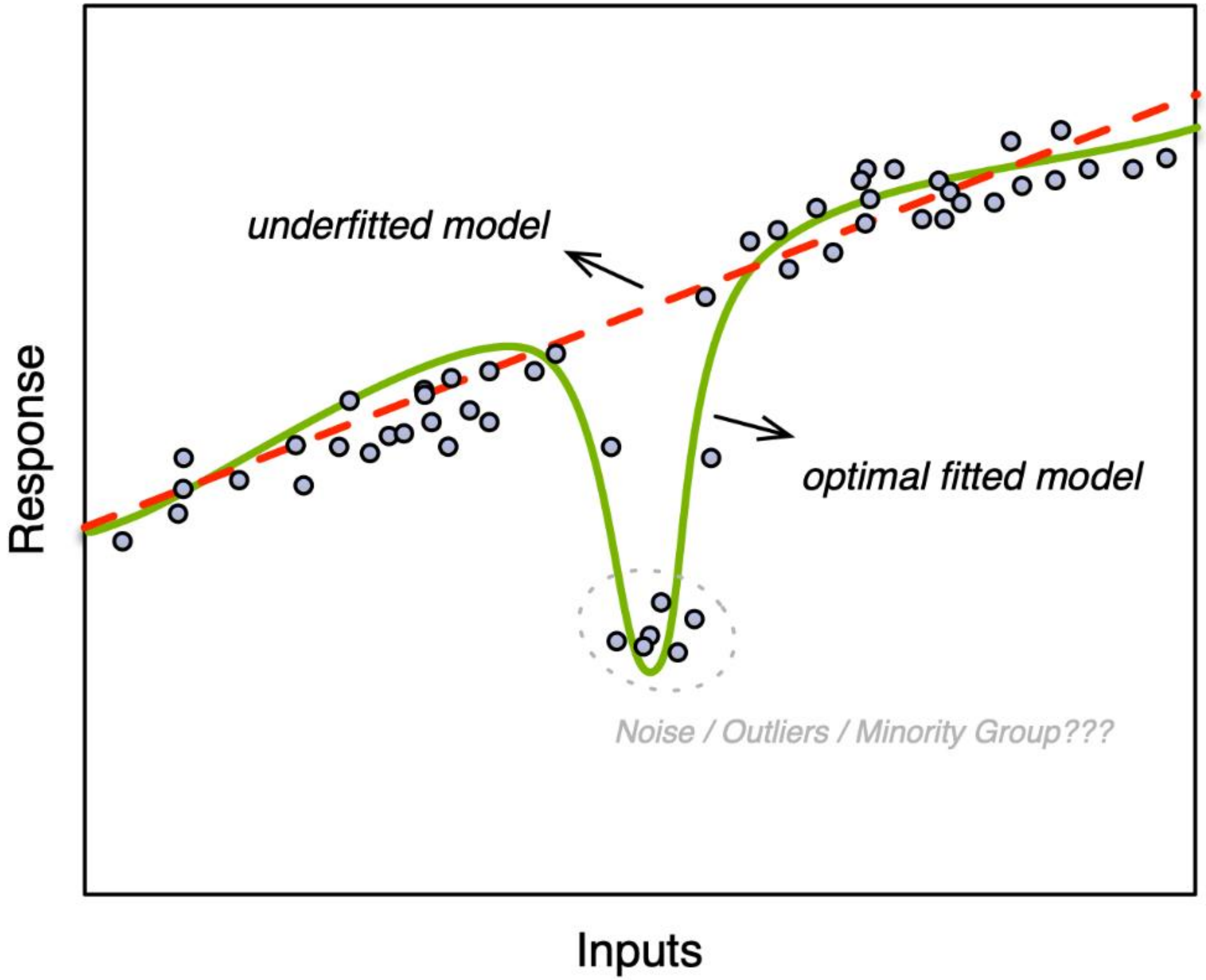
Observe and monitor model performance

- Performance in terms of accuracy
- Performance in terms of bias
- Robustness of model based on inputs
- Performance in terms of general health
- You can't fix an issue you can't see

Continuously test AI for retraining

- Lots of ways successful models can go off the rails
- E.g. Drift of data as compared to training data can be an indication that a model needs retraining.

3. Monitor and Modify your usage as you go



4. Get the whole company on board

"Push left"

- Refers to the left side of the data and AI lifecycle
- The more you can do on the **left** to get the data right to start with, the more trouble you avoid on the **right** later on
- Establish **data stewardship** to ensure data used correctly across the full life cycle.

Establish ethical oversight

- Establish an **oversight** committee to provide guidance on AI ethical dilemmas
- Should cover everything from sales to procurement

Ensure compliance

- Know what your obligations
- More regulations are coming
- Manage **risk**

Drive cultural fluency

- Look at initiatives that drive **cultural fluency**
- **Data and AI literacy**
- **Training** and **coaching** employees in **trustworthy AI** principles, methods, and tools

Operations alignment

- Standard operating procedures
- Strong technology platforms
- Strong practices

March 4, 2020

Trustworthy AI in Aotearoa – The AI Principles

Artificial intelligence (**AI**) can drive significant economic and social benefits for New Zealand. But it also introduces a range of risks and challenges to New Zealand society that cannot be overlooked.

To help maintain public trust in the development and use of AI in New Zealand, the Law, Society and Ethics Working Group of the AI Forum has published a set of guiding principles for "Trustworthy AI in Aotearoa New Zealand" (**the AI Principles**). Those principles are designed to provide high-level guidance for anyone involved in designing, developing and using artificial intelligence in New Zealand (**AI stakeholders**), with the goal of ensuring New Zealanders have access to trustworthy AI.

A key focus of the group has been to make sure the AI principles are simple, succinct and user friendly.

VIEW NOW

The SAS Trustworthy AI Landscape

ACTIVATION

Oversight
AI Governance, Strategy, and Enforcement

Operations
SOPs w/ Supporting Infrastructure

Compliance
Performance and Risk Management

Culture
Ethically Systemic Norms & Practices

TECHNOLOGY

Data Management

- Data Quality
- Variable Metadata
- Data Preparation
- Data Asset Catalog

Explanation

- Natural Language Explanation
- Explainable ML
- Counterfactual Explanation
- Surrogate Model Interpretation
- Causal Inference

Detection

- Bias Detection
- Fairness Assessment

Privacy and Security

- Privacy Preservation
- Model Security
- Autonomy Preservation
- Consent & Control

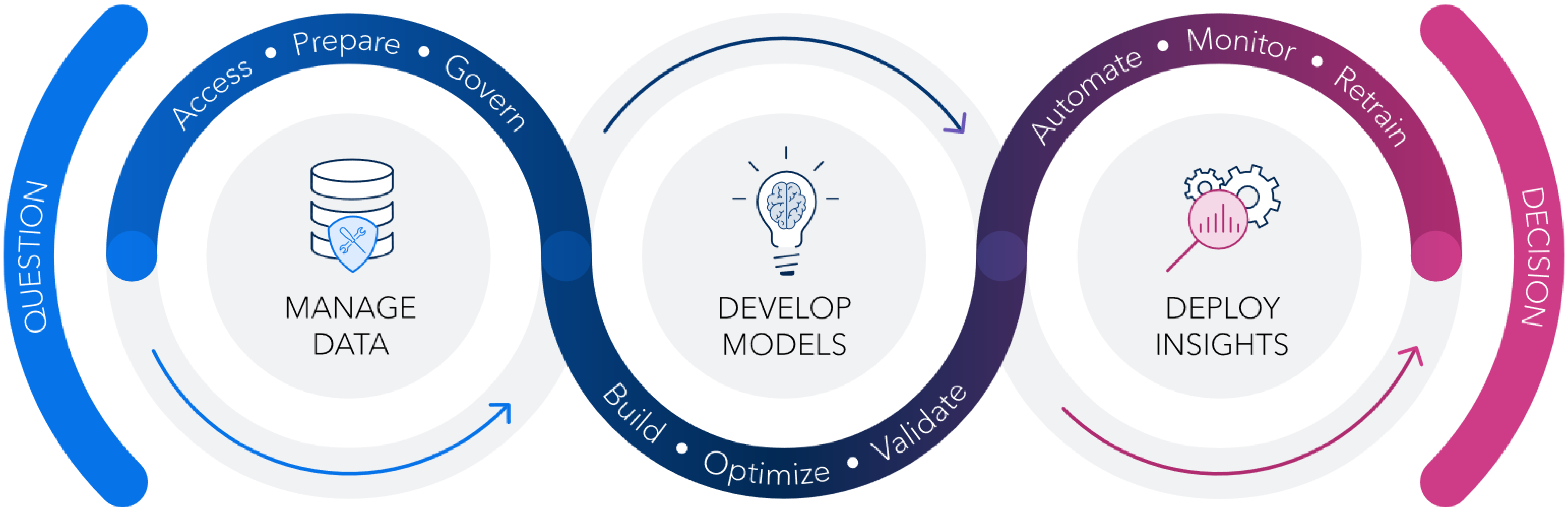
Mitigation

- Bias Mitigation
- Bias Prevention
- Synthetic Data Generation

Model Ops

- Model Cards
- Decisioning
- Lifecycle Management
- Metric Monitoring
- Model Robustness
- Model Oversight

Trustworthy AI Workflow



Data Quality
Data Exploration
Information Privacy
Data Masking
Data Suppression
Data Lineage
Synthetic Data Generation
Semantic Type Remediation

Natural Language Insights
Model Interpretability
Fairness Assessment
Bias Mitigation

Model Governance
Model Monitoring
Decision Accountability
Model Cards

Recap

1. Question and understand your data

(Understand your bias)

2. Use the right data for the right purposes

(Correct for bias, and use horses for courses)

3. Monitor and Modify your usage as you go

(You can't fix an issue you don't see)

4. Get the whole company on board

("Push left")

Takeaways

1. All real data is biased. You can't get rid of bias, only mitigate against it.
2. Understand the bias you have in your data, and compensate based on the purpose you're using it for
3. Use the right AI and models with the right data (and vice versa)
4. Come talk to us!

Thank you!

Simon Edwards
Head of AI Platform Engineering and Architecture
in www.linkedin.com/in/simonjamesedwards