How to Avoid Breeding Bad AI Agents

Data Integrity ensures Accurate AI for Managed Risks

Curtis O’Dell

Global Director Data Integrity SME

A bit about me

- Passionate leader in Tech for data's business value realization, with proven successes. IT Business leader with over 120m in ARR built over career. IPO and M&A veteran.
- I also have a proven track record of creating and delivering value from data analytics with AI, such as identifying fraud and opportunities, generating revenue, and reducing costs. I built my own version of probability for the data work I created and own the IP.

Curtis O'Dell

Global Business Manager @ Tricentis | Driving Global Busi...

IT Executives response to AI

The hype train has left the station

67%

Integrate Gen AI

of enterprises will integrate gen AI into their AI strategy

79%

ERP cloud migration

of executives are using generative AI in app modernization projects

65%

Build custom apps

will integrate Gen AI in building, managing and delivering new apps

$6.6T

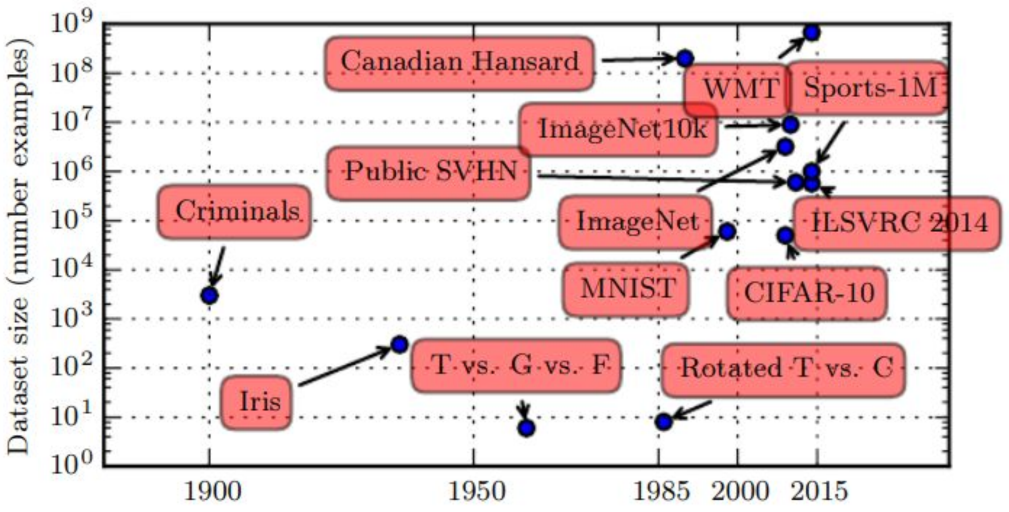
Improve productivity

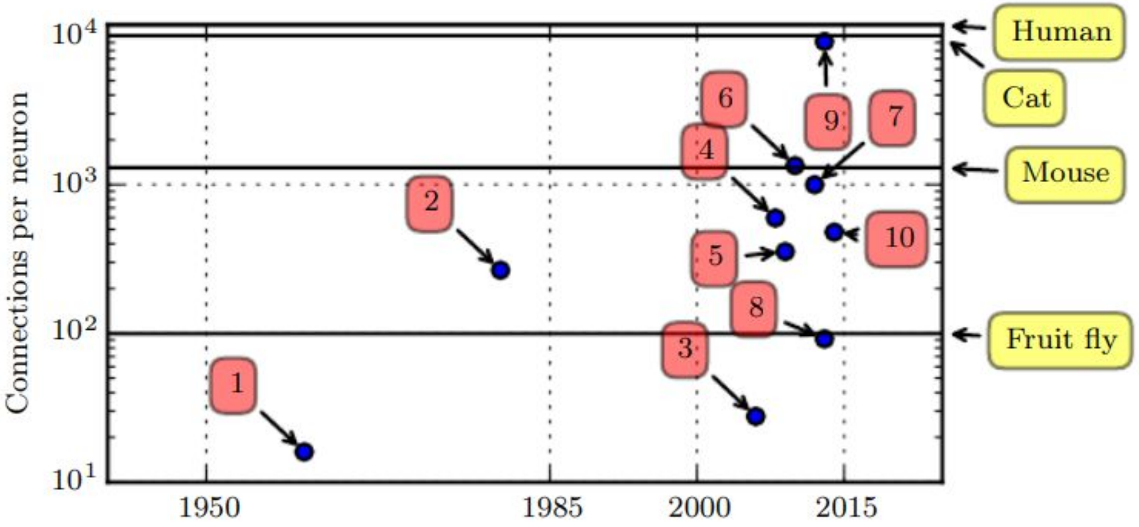
USD contributed by 2030 from increased productivity, 80% of total benefit*

* PwC's Global Artificial Intelligence Study | PwC

AI / ML Why Now?

- Datasets move up in size to astronomical levels Data Explosion for training data catalyst #1
- Neuron connections has exponential levels

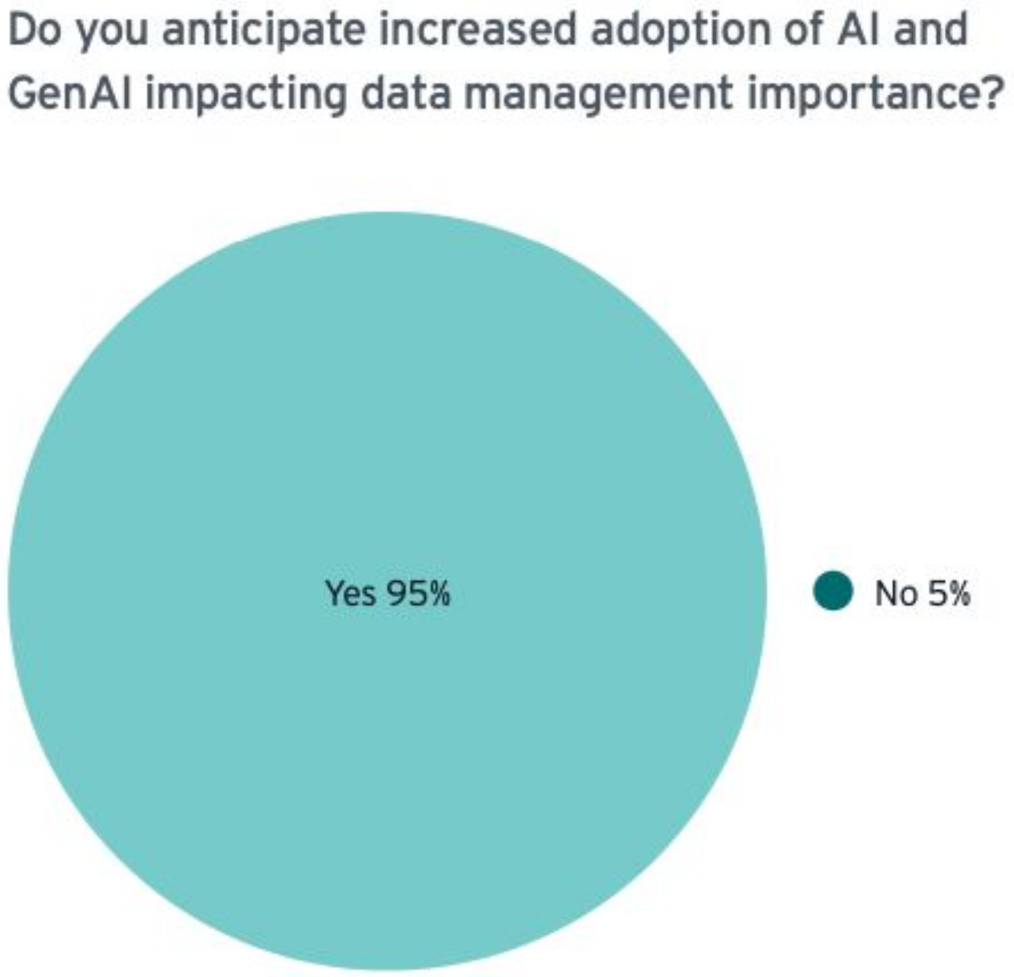




Deep Learning
by Aaron Courville, Yoshua Bengio, Ian Goodfellow

AI Readiness and Data Quality

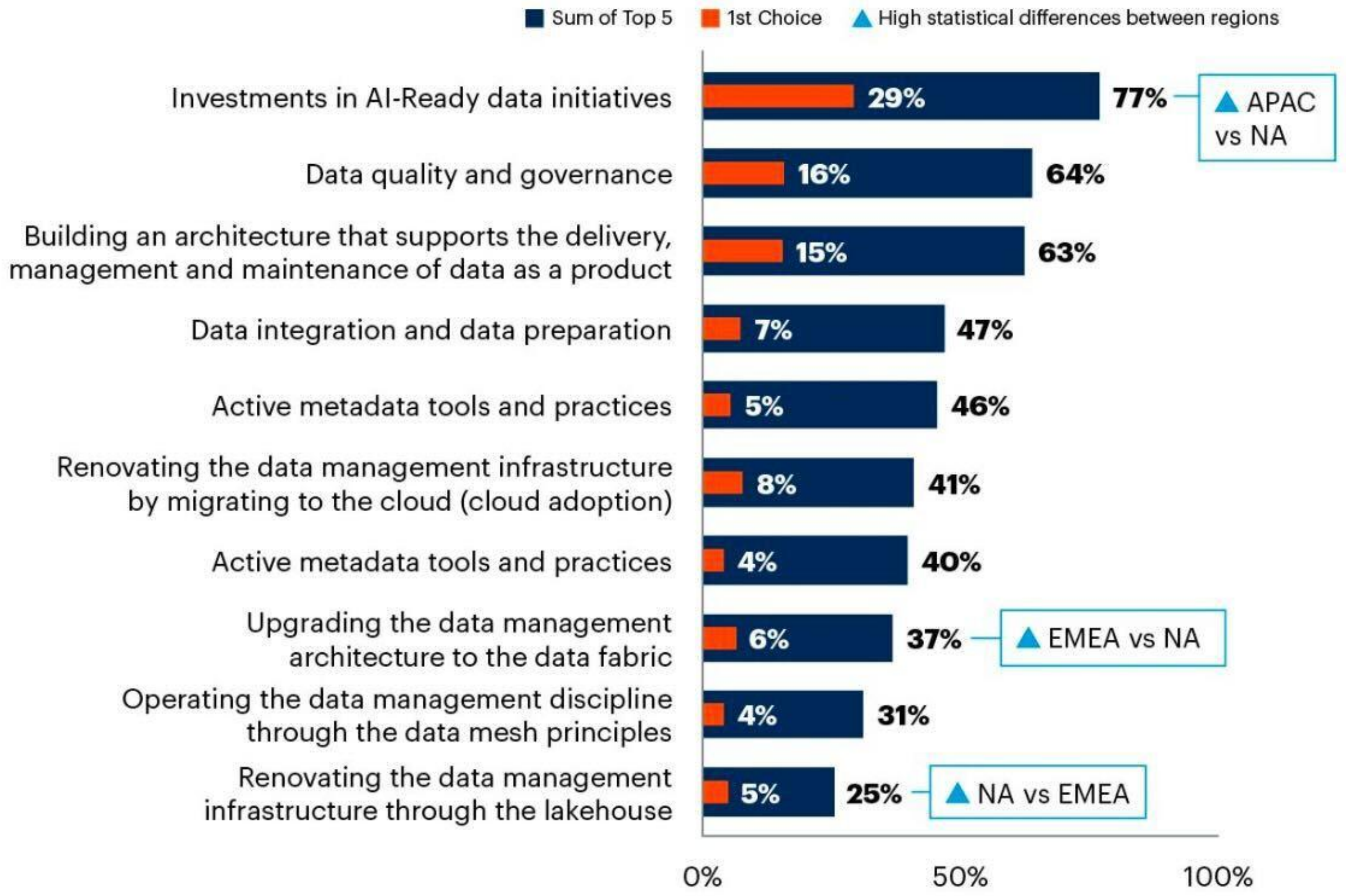
- 7 Key Points to Consider when thinking about AI Ready Data
  - AI-ready data strategy
  - Knowledge management
  - Data governance
  - Master data management
  - Data risk and compliance
  - Data quality
  - AI-ready data architecture


Do you anticipate increased adoption of AI and GenAI impacting data management importance?

NASCIO report: Your AI Blueprint: 12 Key Considerations as States Develop Their Artificial Intelligence Roadmaps

#2 AI Data Management Investment by CIOs

Top 5 Investment Trends in the Next 2-3 Years



n = 247; All Respondents excluding Not Sure

Q10: What do you think are the top 5 investment trends for data management leaders in the next 2-3 years?

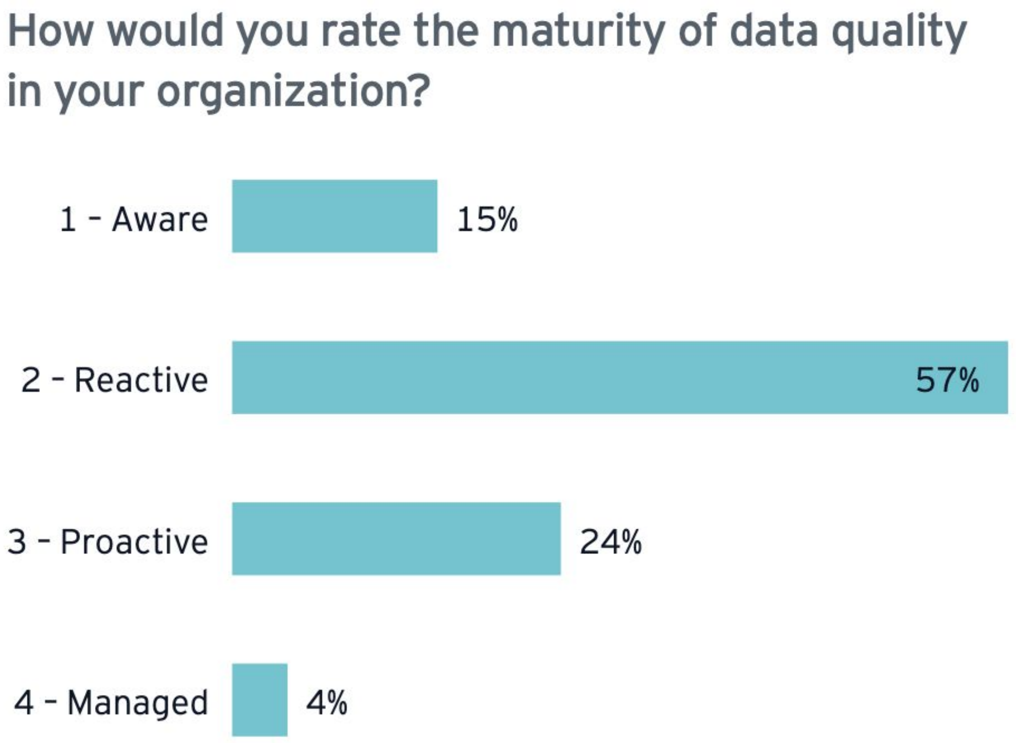
Source: 2024 The Evolution of Data Management Survey

821077_C

Gartner

Data Quality Needs Data Governance

- Data errors can negatively impact mission delivery, constituent trust, and increase costs
- Without a data governance program in place, your organization runs the risk of utilizing poor quality data across the organization leading to fines, poor constituent experience, and failed initiatives.



How would you rate the maturity of data quality in your organization?

Data quality – vital to optimizing GenAI: A Survey of state chief information officers and chief data officers

Massive Data, Massive Models, Massive potential

Model Sizes

Classical Intelligence:

80bn Neurons

2020 AI:

1.5bn Neurons

2024 AI:

1.35T Neurons

Dataset Sizes

Standard 2010's:

15k data

2020's Fraud:

2 million records

2024 AI:

15 Trillion tokens

185m ChatGPT users

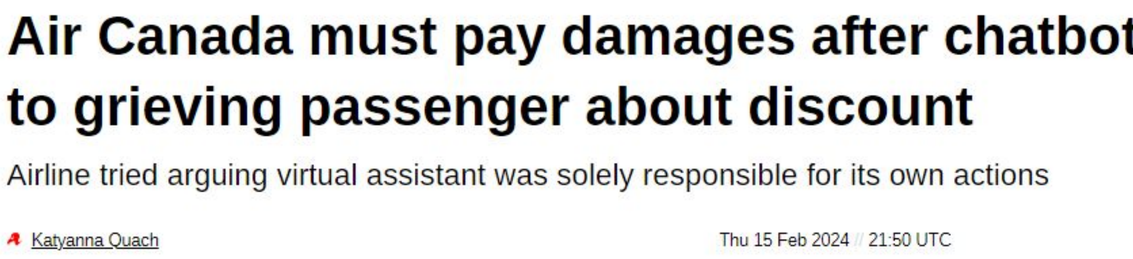
Data Management Practices – is your data ready?

- Good Data Management looks like….
  - Addressing potential risk early and often
  - Effectively preparing for AI/ML journeys
  - Establishing Robust data practices

- But are you there yet?
  - Can your data keep up?
  - Can you trust accuracy of your data and thus your AI?
  - Is your data compliant?

DATA ENGINEERS WHEN ASKED..
'OUR DATA QUALITY IS READY FOR AI?'

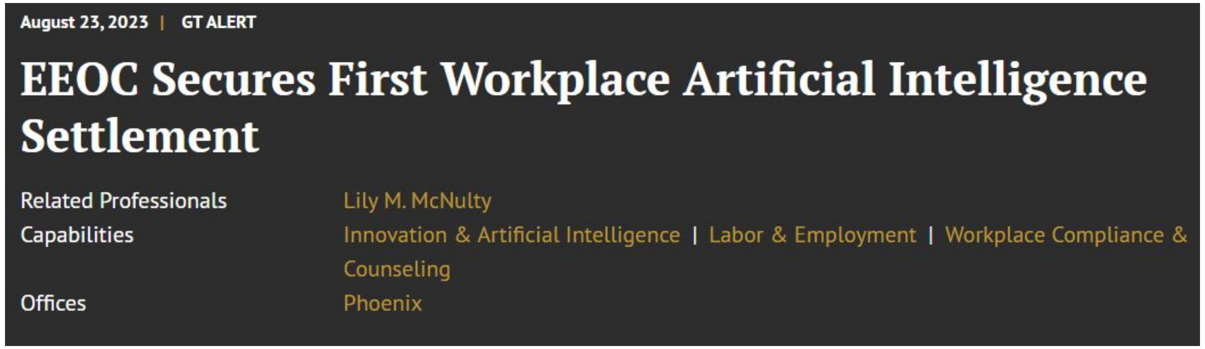
But it can also go pretty badly wrong…



Air Canada must pay damages after chatbot lies to grieving passenger about discount

Airline tried arguing virtual assistant was solely responsible for its own actions

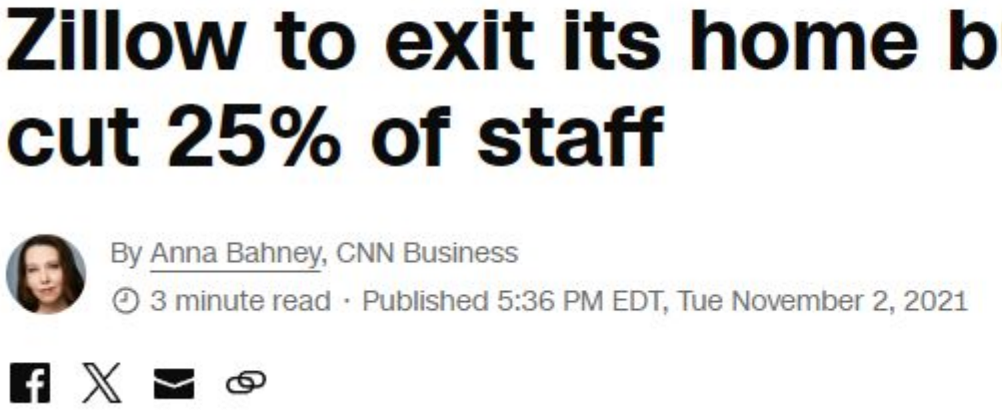
Katyanna Quach Thu 15 Feb 2024 // 21:50 UTC



August 23, 2023 | GT ALERT

EEOC Secures First Workplace Artificial Intelligence Settlement

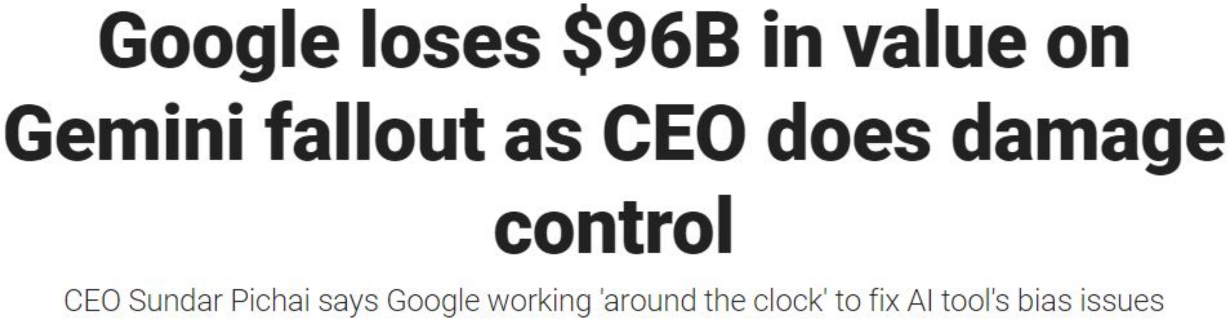
Related Professionals	Lily M. McNulty
Capabilities	Innovation & Artificial Intelligence | Labor & Employment | Workplace Compliance & Counseling
Offices	Phoenix



Zillow to exit its home buying business, cut 25% of staff

By Anna Bahney, CNN Business

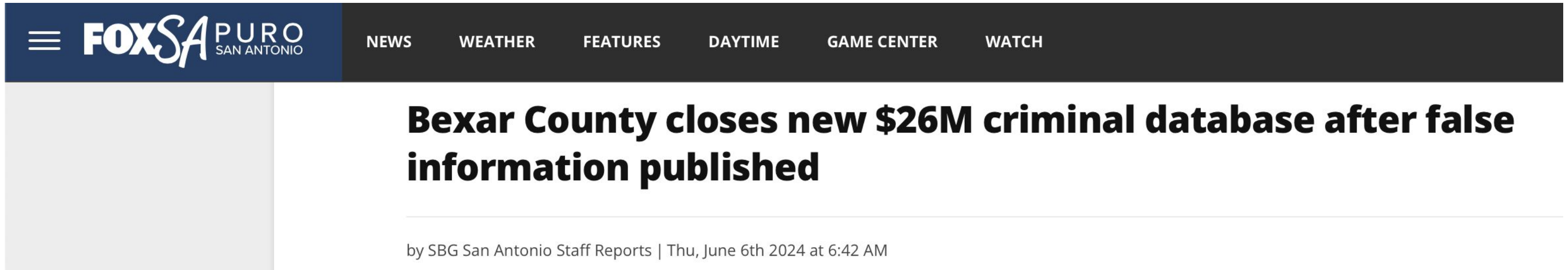
3 minute read · Published 5:36 PM EDT, Tue November 2, 2021



Google loses $96B in value on Gemini fallout as CEO does damage control

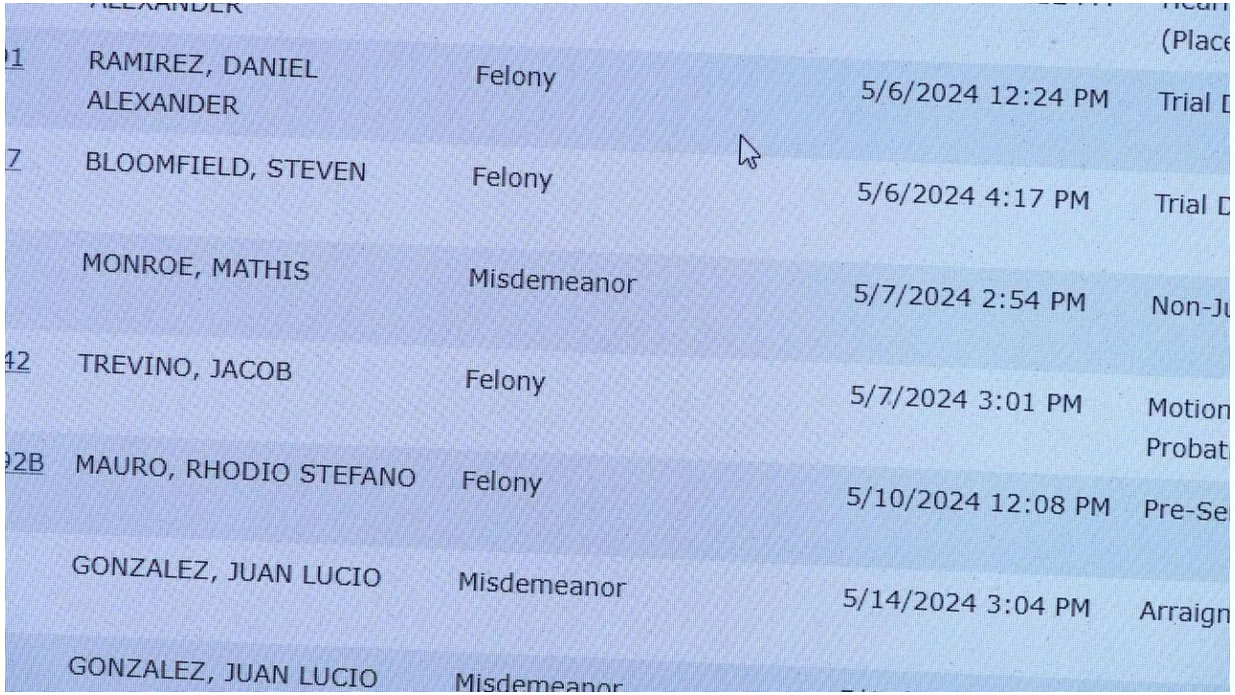
CEO Sundar Pichai says Google working 'around the clock' to fix AI tool's bias issues

Bad data not only affects your internal initiatives, but also your reputation



Bexar County closes new $26M criminal database after false information published

by SBG San Antonio Staff Reports | Thu, June 6th 2024 at 6:42 AM



	ALEXANDER			(Place
1	RAMIREZ, DANIEL ALEXANDER	Felony	5/6/2024 12:24 PM	Trial D
7	BLOOMFIELD, STEVEN	Felony	5/6/2024 4:17 PM	Trial D
	MONROE, MATHIS	Misdemeanor	5/7/2024 2:54 PM	Non-Ju
42	TREVINO, JACOB	Felony	5/7/2024 3:01 PM	Motion Probat
2B	MAURO, RHODIO STEFANO	Felony	5/10/2024 12:08 PM	Pre-Se
	GONZALEZ, JUAN LUCIO	Misdemeanor	5/14/2024 3:04 PM	Arraign
	GONZALEZ, JUAN LUCIO	Misdemeanor		

Some frustrations are — dismissed cases are showing active, cases where people have been acquitted, or showing they're convicted.

How can you identify a quality Data Governance tool?

A quality data governance tool is one that is integrated into your current testing process that is continuous, automated, and end-to-end

Data Testing Report Card

Accuracy

Do data objects correctly represent the values?

Completeness

How much data is missing?

Conformity

Does the data match a specified format?

Consistency

Are there data conflicts related to an object?

Integrity

Are data relationships maintained?

Timeliness

Is data up-to-date for the task at hand?

Uniqueness

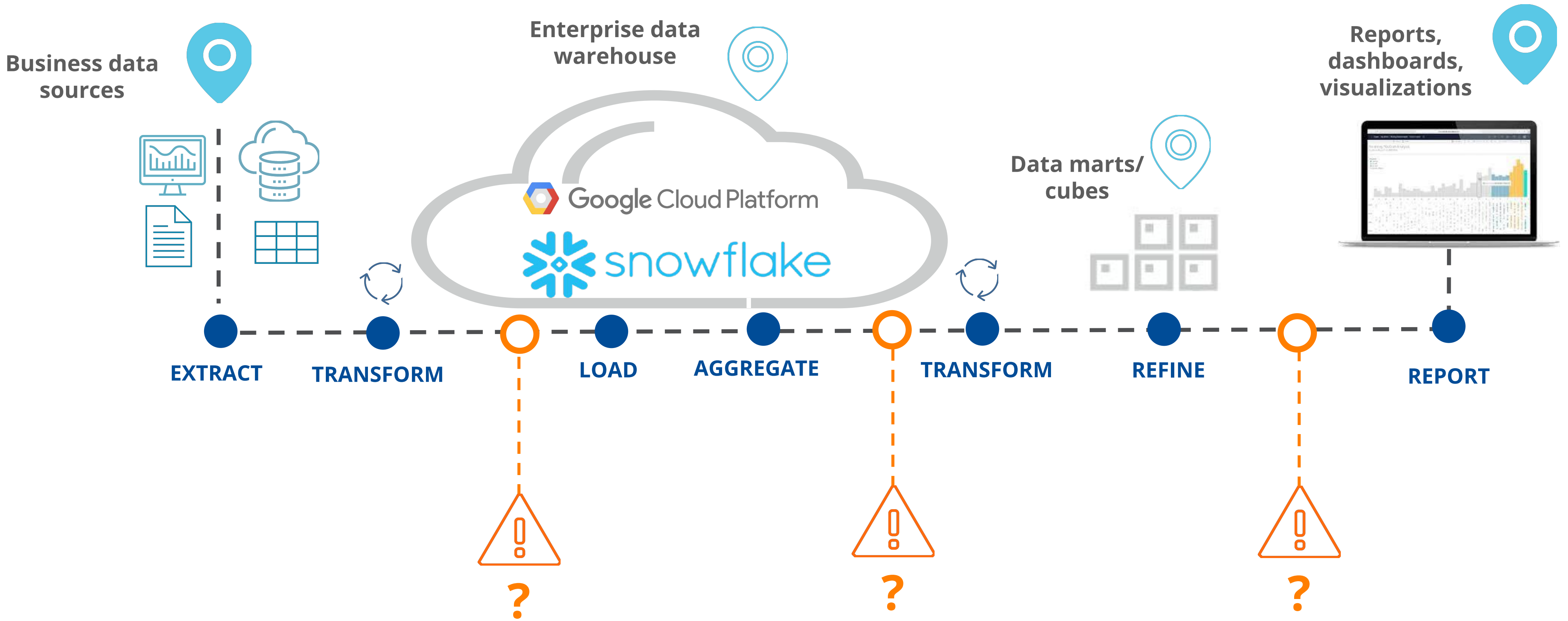
Does data repeat where it shouldn't?

AI is not **Magic**

It is an **high quality parrot**

It needs **high quality training data**

The race is on to find the data errors



Business data sources

Enterprise data warehouse

Google Cloud Platform

snowflake

Data marts/ cubes

Reports, dashboards, visualizations

EXTRACT

TRANSFORM

LOAD

AGGREGATE

TRANSFORM

REFINE

REPORT

Did the problem originate in the source data?

Was there an issue with a data load?

Did a transformation job go wrong?

Did a job fail to run or run too many times?

Were there issues with the transformation logic?

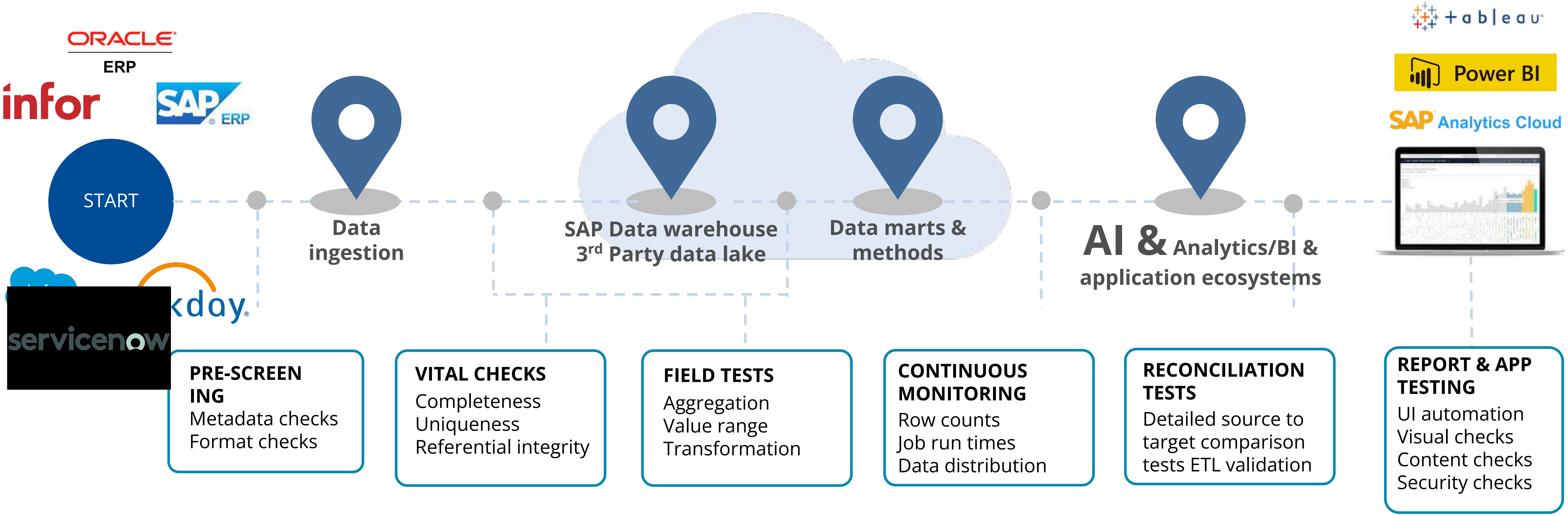
Is the report pulling from the right data mart?

Is there a problem with the report logic?

Is the report rendering incorrectly?

Deliver trustworthy data through a Complex Process

Example scenario: Data Pipeline for analytics and dashboards



ORACLE ERP

infor

SAP ERP

servicenow

kday

START

Data ingestion

SAP Data warehouse
3rd Party data lake

Data marts & methods

AI & Analytics/BI & application ecosystems

tableau

Power BI

SAP Analytics Cloud

PRE-SCREEN ING
Metadata checks
Format checks

VITAL CHECKS
Completeness
Uniqueness
Referential integrity

FIELD TESTS
Aggregation
Value range
Transformation

CONTINUOUS MONITORING
Row counts
Job run times
Data distribution

RECONCILIATION TESTS
Detailed source to target comparison tests ETL validation

REPORT & APP TESTING
UI automation
Visual checks
Content checks
Security checks

SAP Solution Extension ORACLE Microsoft Azure aws servicenow salesforce

Improving your data quality improves your dev time

Before:

- 2 Months dev effort
- Limited progress

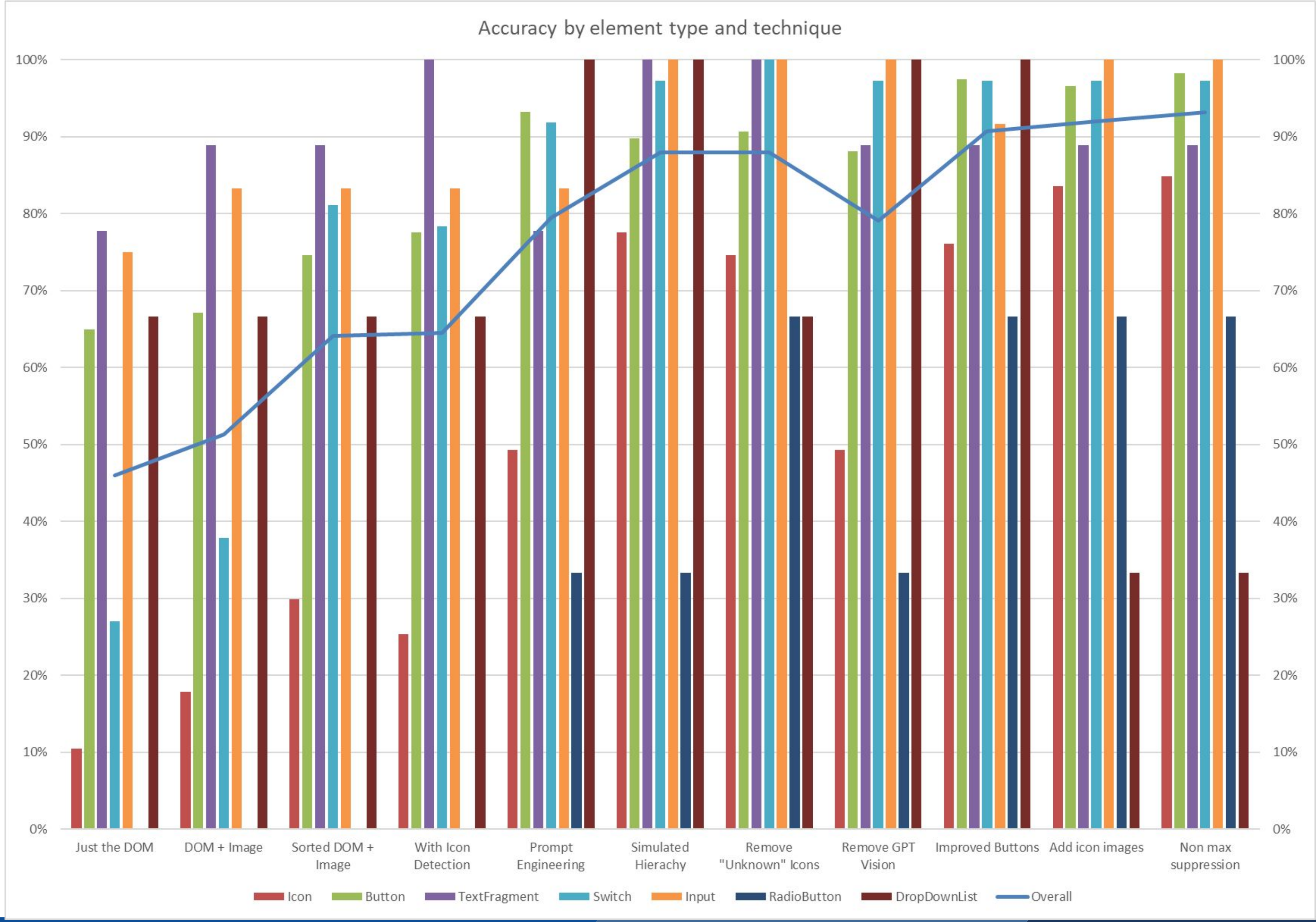
Actions:

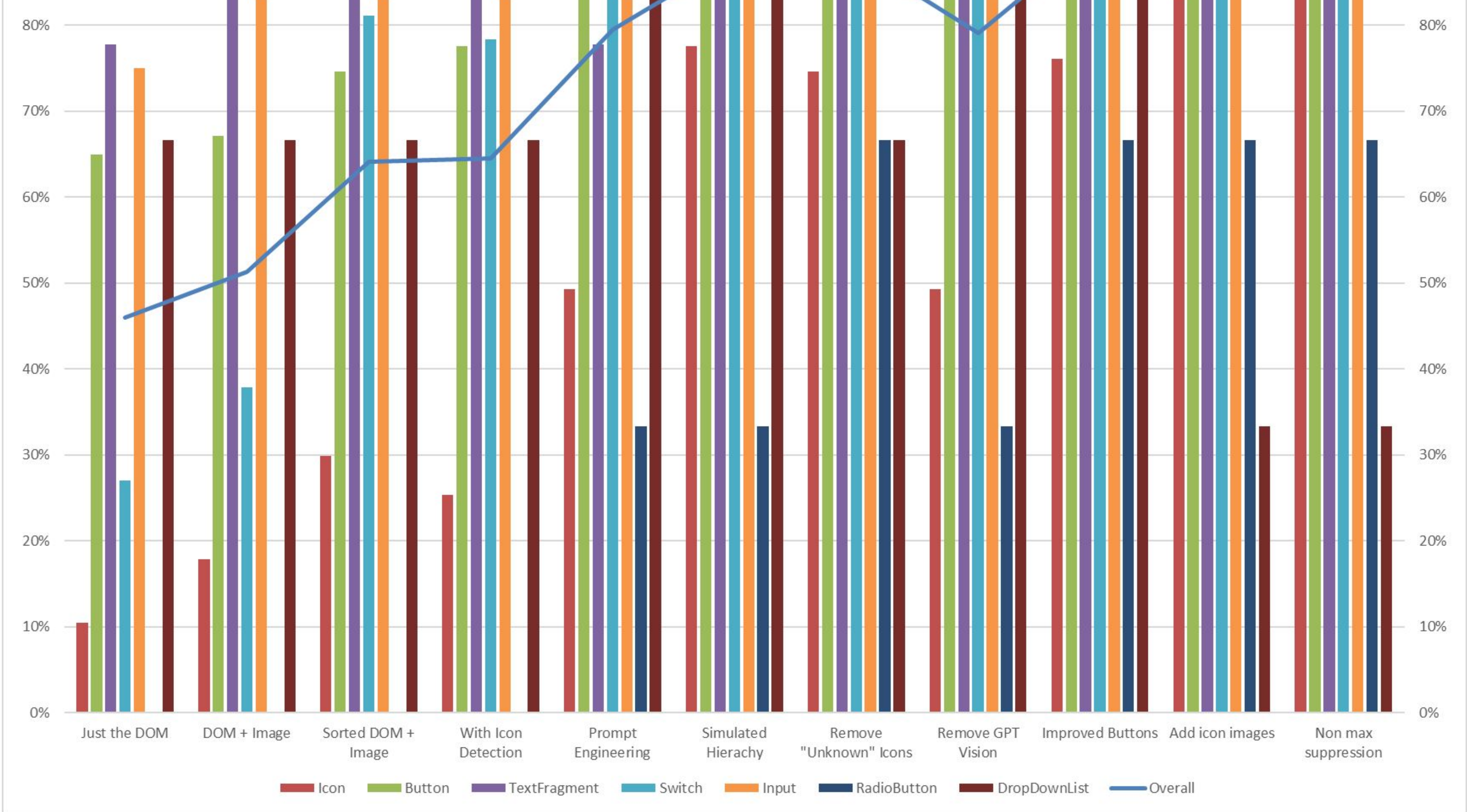
□ Cleanse input data

□ Re-categorise and validate

After:

✔ 1 week from 46% to 93%




80%
70%
60%
50%
40%
30%
20%
10%
0%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Just the DOM
DOM + Image
Sorted DOM + Image
With Icon Detection
Prompt Engineering
Simulated Hierachy
Remove "Unknown" Icons
Remove GPT Vision
Improved Buttons
Add icon images
Non max suppression
Icon
Button
TextFragment
Switch
Input
RadioButton
DropDownList
Overall

Testing AI Challenges

UNPREDICTABILITY:

CHALLENGE OF TESTING AI-GENERATED CONTENT, WHICH CAN BE UNPREDICTABLE AND VARIABLE.

LACK OF GROUND TRUTH:

THE DIFFICULTY OF ESTABLISHING GROUND TRUTH (YOUR BENCHMARKS) FOR AI-GENERATED CONTENT, MAKING IT HARD TO EVALUATE ACCURACY.

CONTEXTUAL UNDERSTANDING:

THERE IS A NEED FOR TESTERS TO UNDERSTAND THE CONTEXT IN WHICH AI-GENERATED CONTENT IS BEING USED.

BIAS AND FAIRNESS:

THE IMPORTANCE OF TESTING FOR BIAS AND FAIRNESS IN AI-GENERATED CONTENT.

Approaches for Testing AI

1) **Black Box Testing**: Use black box testing to evaluate AI-generated content without knowledge of the underlying model.
2) **White Box Testing**: Use white box testing to evaluate AI-generated content with knowledge of the underlying model.
3) **Hybrid Approaches**: Hybrid approaches that combine black box and white box testing.
4) **Evaluation Metrics**: Use evaluation metrics such as accuracy, precision, recall, and F1 score to assess AI-generated content.
5) **Human Evaluation**: *Automation* of the human evaluation in assessing the quality and relevance of AI-generated content.

Tools and Techniques for Testing AI

Manual: Human-in-the-Loop Testing: Use human-in-the-loop testing to evaluate AI-generated content

Better: Testing Frameworks: Testing frameworks such as PyTorch, TensorFlow, Keras

Best: Automated Testing Tools: Automated testing tools with low code automation

Deeper AI Testing Use Cases

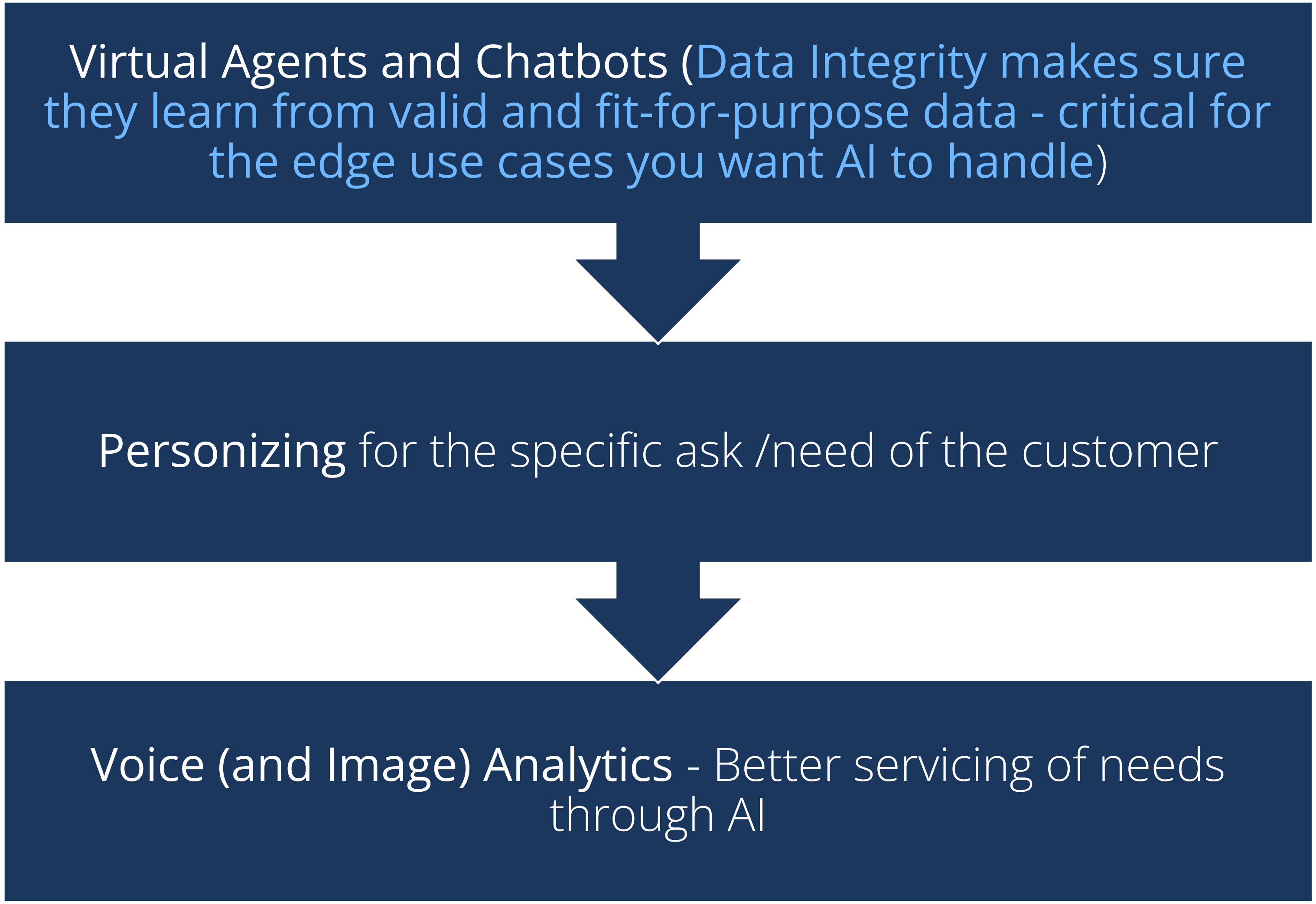
AI Benefits from Data Integrity

Create a solid data foundation for your analytics and AI projects by feeding them trustworthy data in a simple, efficient and cost-effective way. Use clean, fit-for-purpose data

- 2 major areas for AI impacts on a business
  - Gen AI:
    - Improving the customer experience
    - Employee productivity boosted with AI
  - ML/Predictive AI: Optimizing Business Operations Outcomes
    - Data Migrations
    - Innovation
    - Compliance

Improving the customer experience: (Customer Contact) And Employee productivity boosted with AI (Generative AI)

Create and Validate Content data integrity with automated, end to end and continuous data validations for your employee's AI processes



Virtual Agents and Chatbots (Data Integrity makes sure they learn from valid and fit-for-purpose data - critical for the edge use cases you want AI to handle)

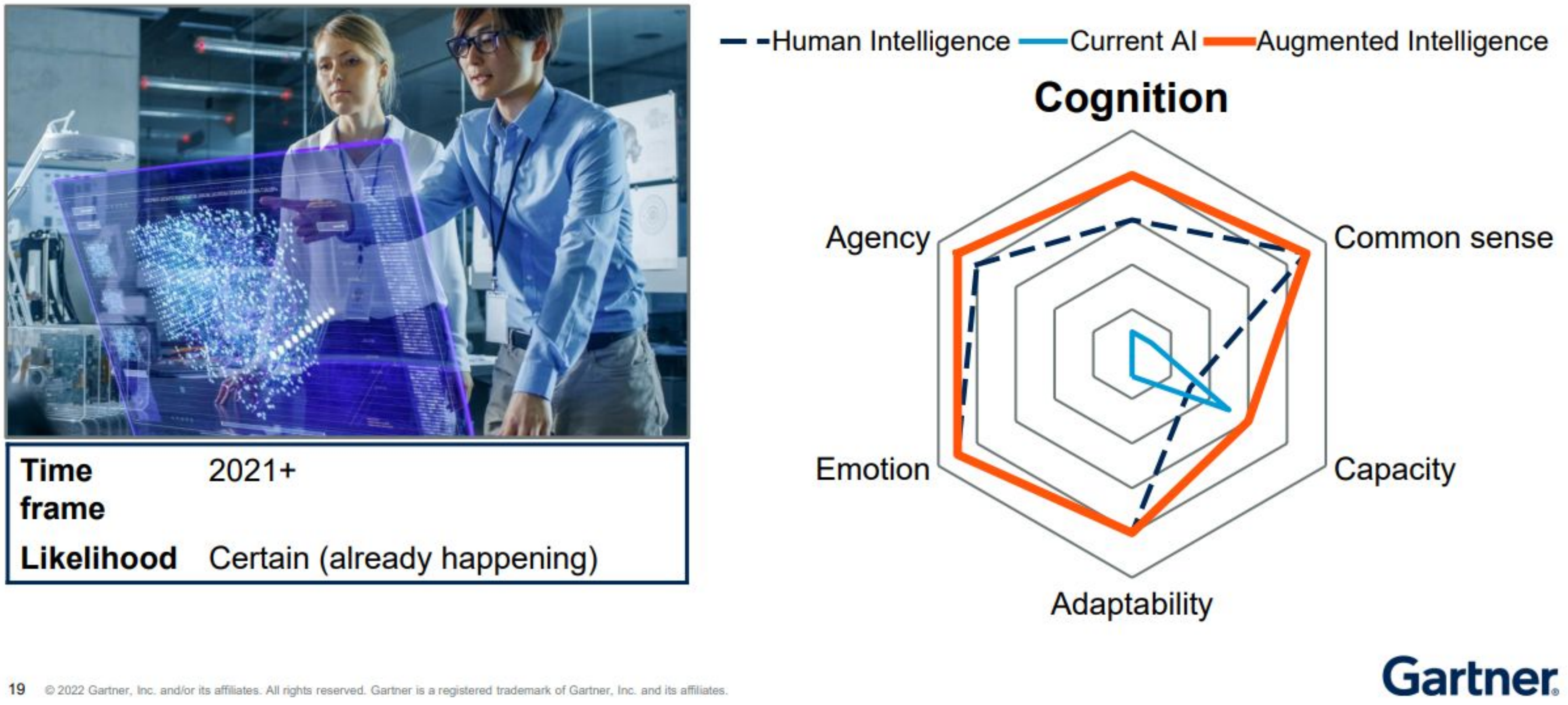
Personizing for the specific ask /need of the customer

Voice (and Image) Analytics - Better servicing of needs through AI

Optimizing Agency Operation's Outcomes

- AI/ML
  - Risk
    - Fraud Detection
    - Compliance requirements
  - Predictive Innovation
    - Business Process Optimization
  - Gen Innovation
    - Intelligent Document Processing

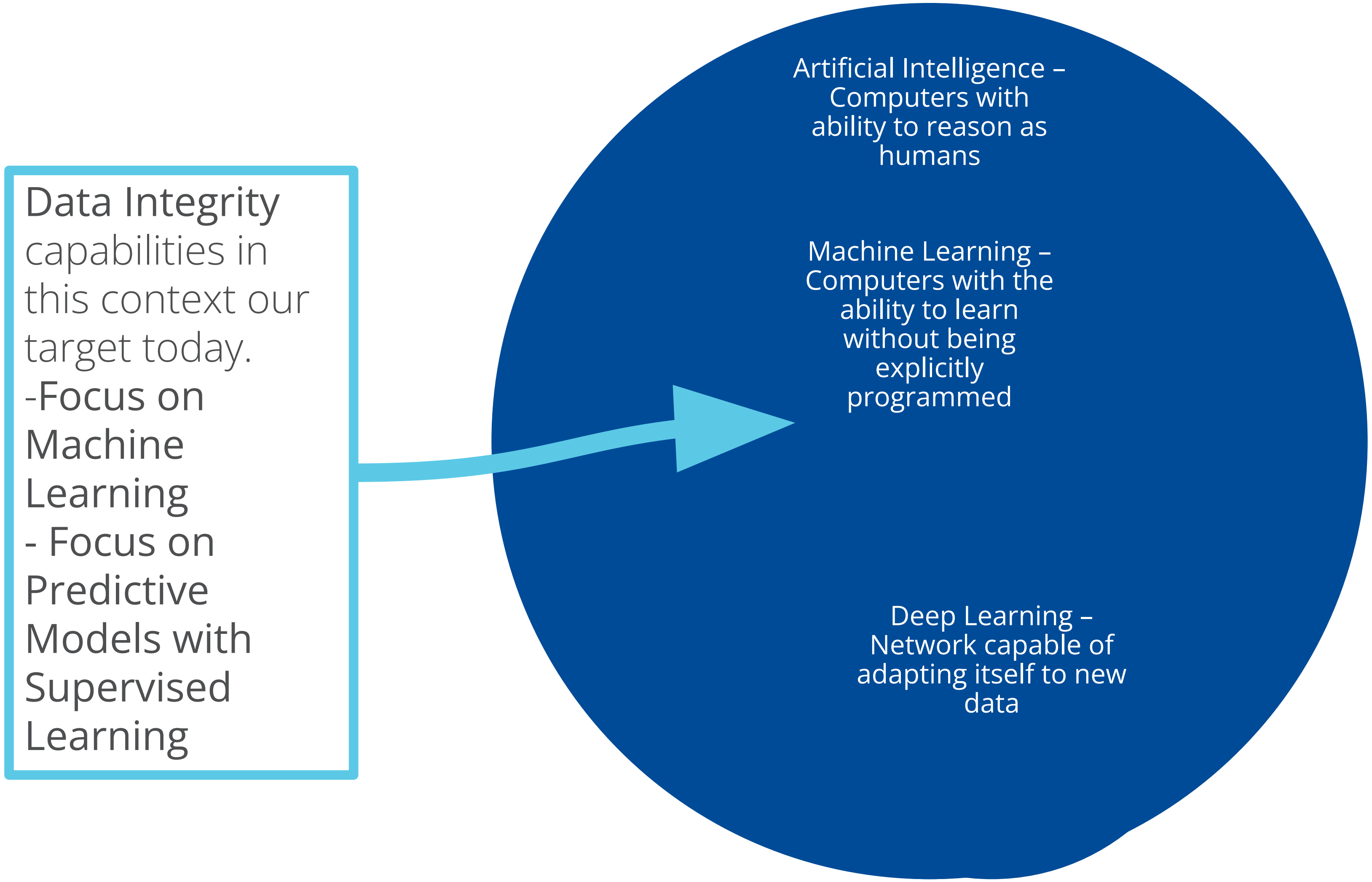
Augmented Intelligence: Complementing Human Strengths Using Automated Data Integrity as the Humans



Time frame	2021+
Likelihood	Certain (already happening)

19 © 2022 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates.

Where Data Integrity fits in AI/ML



Data Integrity capabilities in this context our target today.

-Focus on Machine Learning

- Focus on Predictive Models with Supervised Learning

Artificial Intelligence – Computers with ability to reason as humans

Machine Learning – Computers with the ability to learn without being explicitly programmed

Deep Learning – Network capable of adapting itself to new data

(Predictive) Discriminate Models are:

- Regression
- Classification
- Logistic Regression
- Support Vectors Methods
- Convolutional Neural Networks
- Reinforcement Learning
- Federated Learning
- Ensemble Learning
- xgBoost

Point □ Labeled data most important here - Supervised

Generative Models are:

- Gan – Gen Adversarial Models
- LLM – Large Language Models

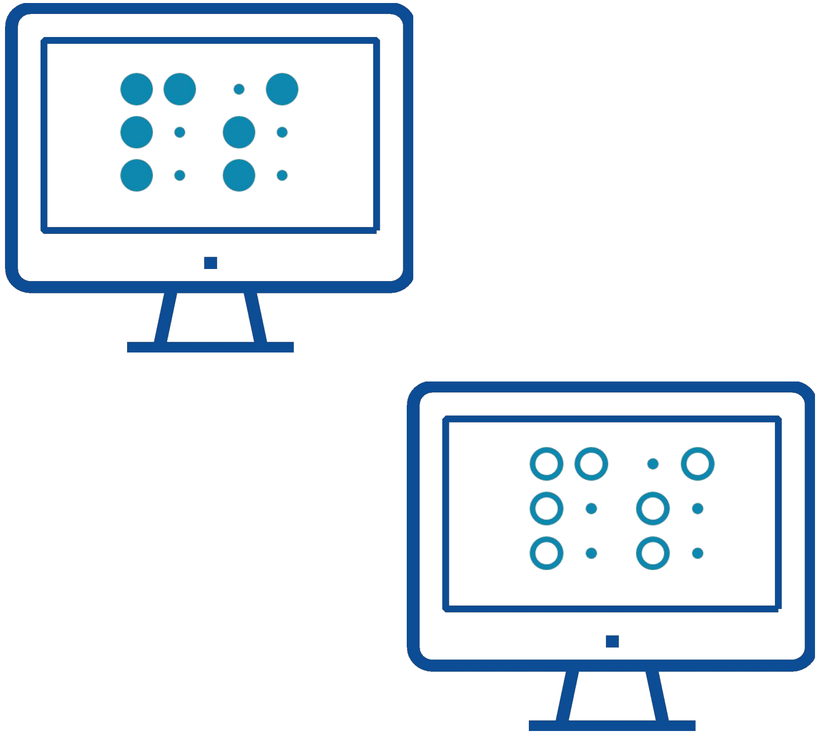
Point □ Unlabeled data drives these - Unsupervised

Manual “stare and compare” is slow and doesn’t scale.

And is not a great use of your team’s brainpower.

They are Data Scientists and Engineers not Janitors

10^9th power is 1bil records! Years to manually check!



Required to ensure data integrity Trust □ A data TESTING solution that's…

Includes data, UI, and API testing for any data type — across your entire landscape.

End-to-end

Automation Augments AI with Automated Humans with Training data coverage > 90%

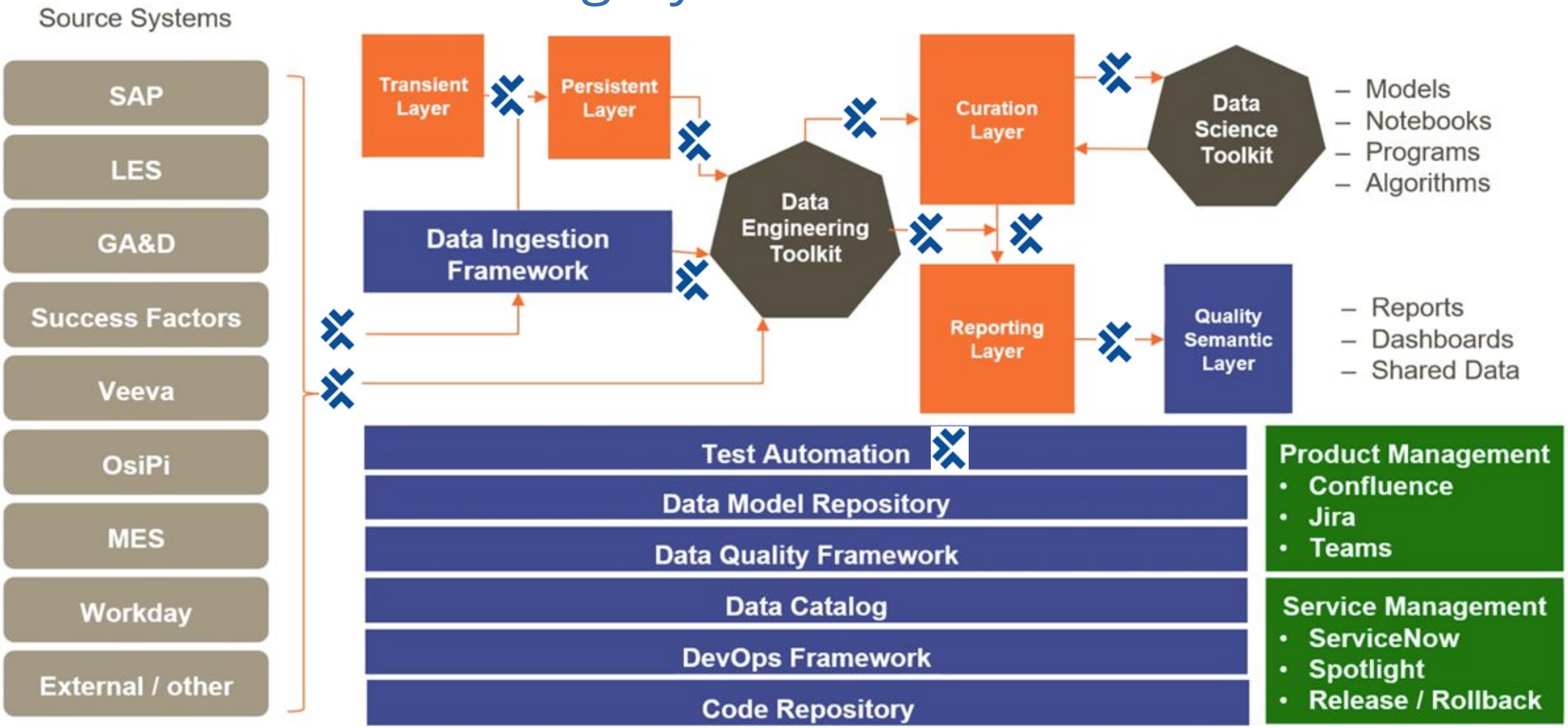
Automated

Utilize DevOps □ DataOps □ MLOps □ capabilitiess for Continuous TRAINING and DRIFT FIX

Continuous

Example Reference Architecture for ML with Data Integrity

- ▨ Source Data
- 🟧 Cloud Data Stores
- 🟦 D&A Governed Capabilities
- ⬣ D&A Governed Toolkits
- 🟩 Plan, Monitor, Manage & Control



Source Systems

SAP

LES

GA&D

Success Factors

Veeva

OsiPi

MES

Workday

External / other

Transient Layer

Persistent Layer

Data Ingestion Framework

Data Engineering Toolkit

Curation Layer

Data Science Toolkit

- Models
- Notebooks
- Programs
- Algorithms

Reporting Layer

Quality Semantic Layer

- Reports
- Dashboards
- Shared Data

Test Automation

Data Model Repository

Data Quality Framework

Data Catalog

DevOps Framework

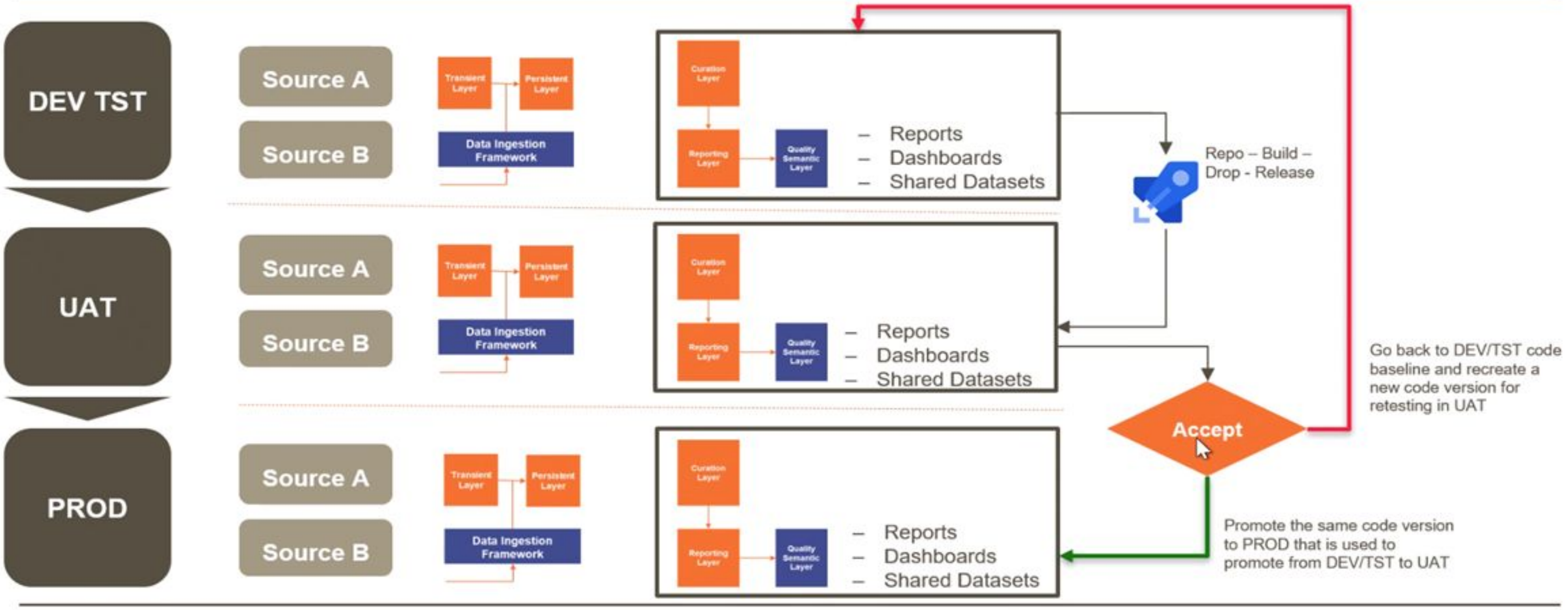
Code Repository

Product Management
- Confluence
- Jira
- Teams

Service Management
- ServiceNow
- Spotlight
- Release / Rollback

Example strategy – Deployments/DataOps architecture for ML with data integrity

Pipelines automatically builds and tests code projects to make them available to others



- Reports
- Dashboards
- Shared Datasets

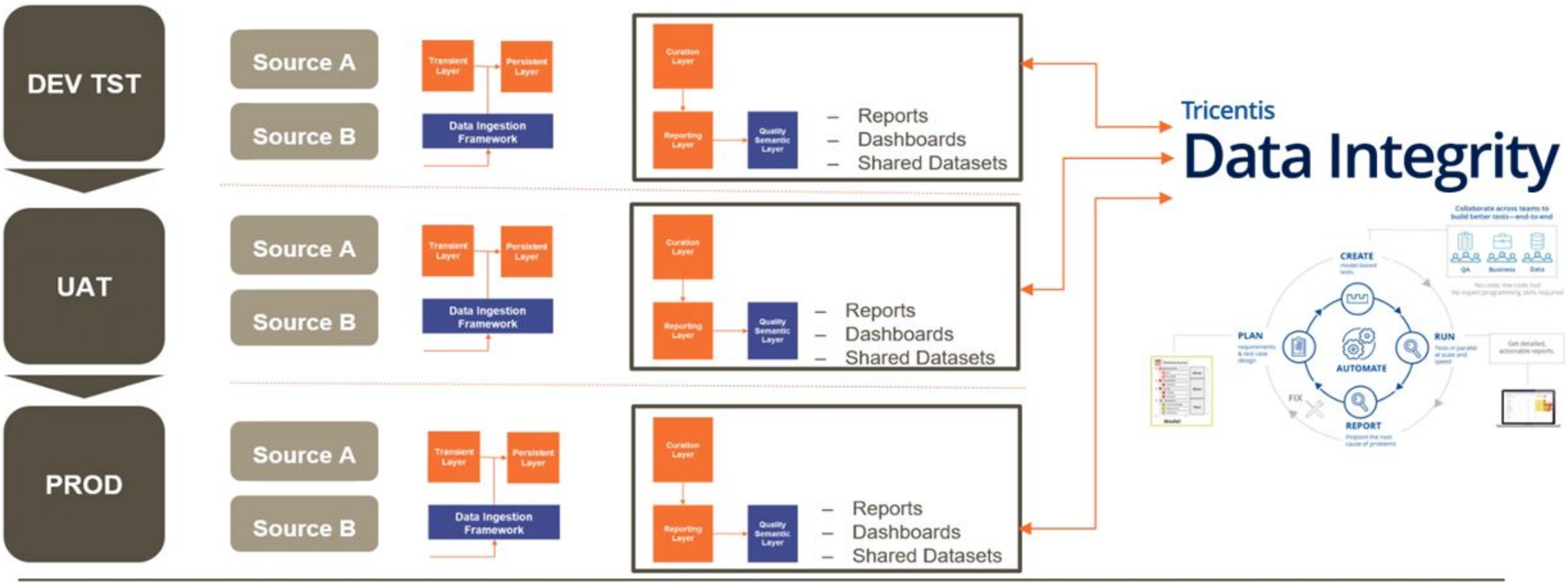
Repo – Build – Drop - Release

Go back to DEV/TST code baseline and recreate a new code version for retesting in UAT

Promote the same code version to PROD that is used to promote from DEV/TST to UAT

Example strategy – Test automation architecture for ML with data integrity

Keep the test automations easy, faster and at scale with 'Model-based' tests



Optimizing Agency Operation's Outcomes – Risk / Compliance

- Example Banking Compliance
  - AML & AML + KYC compliance
    - Anti – Money Laundering and Adversarial Machine Learning
    - Compliance reporting with AI.ML produced data
  - Backdoor attacks can happen purpose or by accident

Integrity of the Data in the Models

Includes:

1. Schema and Metadata Checks
2. Parsing Checks
3. Clean-Label and Backdoor Poisoning

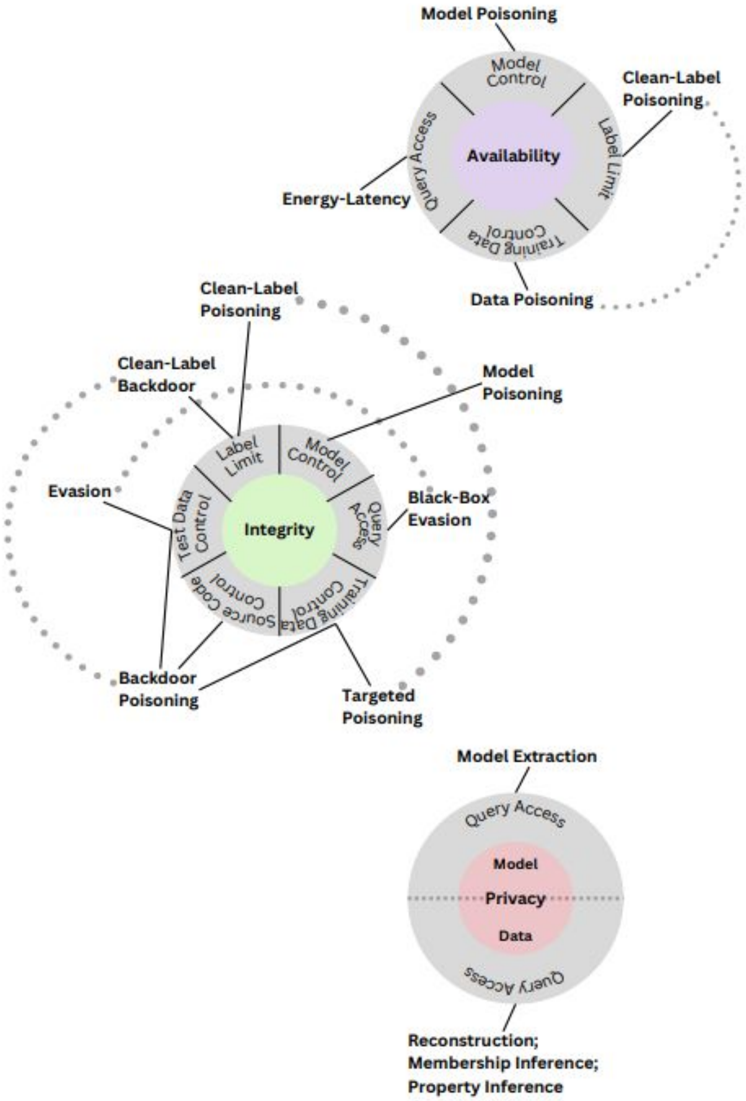
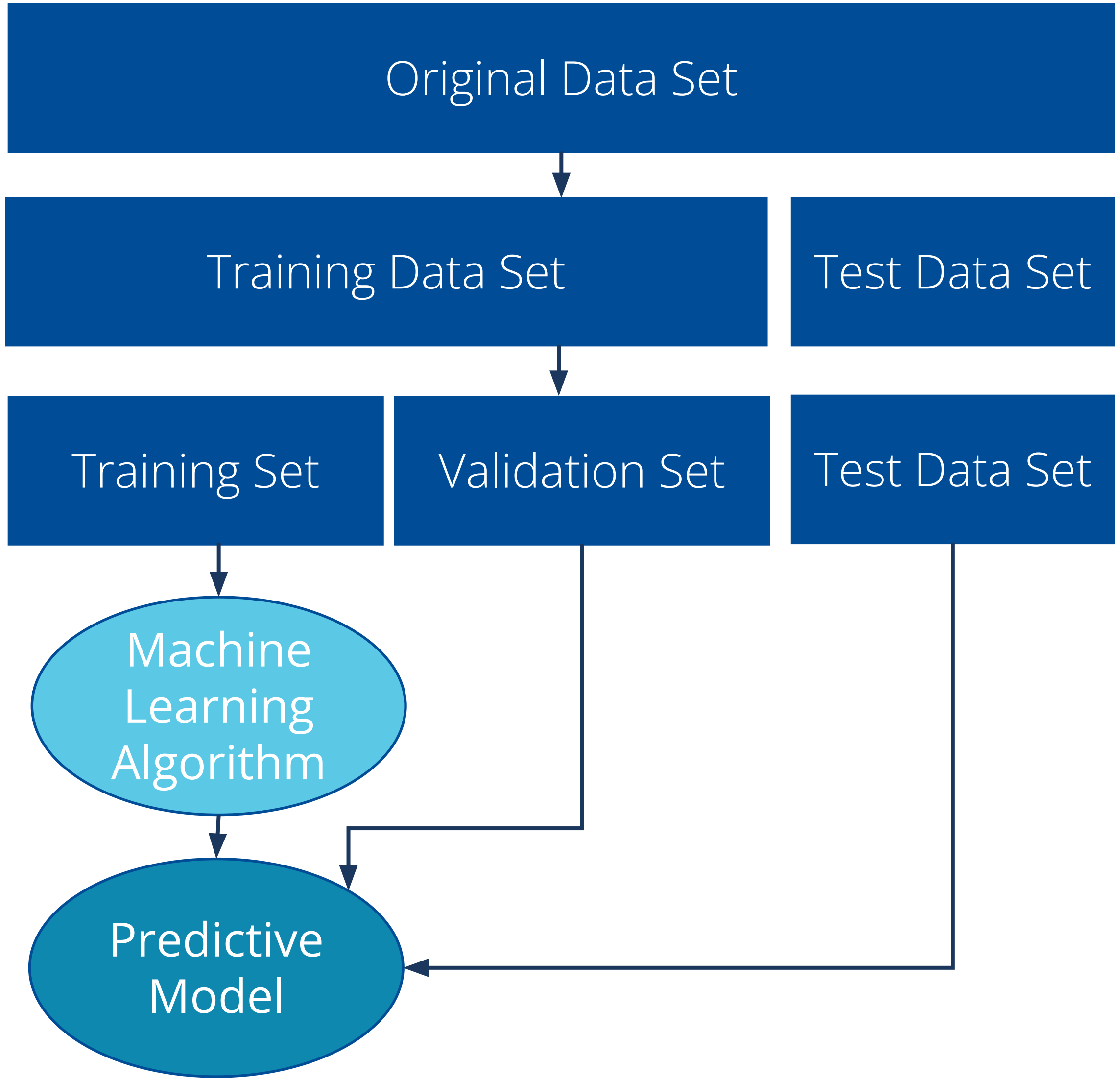


Figure 1. Taxonomy of attacks on Predictive AI systems.

NIST Trustworthy and Responsible AI NIST AI 100-2e2023 Diagram

How we can assist with Training Data as you move to AI

- Testing Data Can be a great primer for Training data and Test Data
- Feature engineering enhanced by great data and business analysts understanding
- If you get public datasets they are often laden with data problems
  - Ex. Our CMS data set



Data Integrity Benefits for AI - Recap

1. Feature Engineering augmented with Application Test Parameters
    - Business Value associated with all AI/ML comes from the quality of Features and Hyperparameters
2. Training Data augmented with Test Data from Data and Application's Tests
    - Efficient Creation of Test Data for testing, utilized for training data, can decrease TTM for AI/ML solutions
3. Data Migrations from older Unified Data Model 2.0 to ML Data Model 2.0
    - Move data for AI/ML at 2x the speed
4. Validation of integrated data models into Data Science
    - A solid automated MLOps process will ensure results and TTM
5. Validation of pipeline for delivery of Data Science results
    - See #4

Thank You!

C.ODell@tricentis.com

214-616-0853