

Making AI Practical

How to Get Started Today

Mark S. Struck, FCAS, FCIA, MAAA
Principal Consultant
Online Business Systems Global





AGENDA

- 01 A Bit About Me
- 02 AI's Trajectory
- 03 AI-Assisted Analysis
- 04 Use Case Identification
- 05 Risks and Limitations
- 06 Think Strategically!

A Bit About Me

Expertise

- Actuarial Science/Mathematics
- Enterprise Risk Management
- Data Management
- Machine Learning/Advanced Analytics/AI
- System Design, Development, and Deployment
- Leading Teams to accomplish all of the above

Roles that shape my thinking

- Principal Consultant, Data & AI @ Online Business Systems
- Chief Analytics Officer, Max Analytics

Weapon of choice

- ChatGPT

 Founded in 1986
Privately held

 Over 400 professionals
in Canada, USA and
EMEA

 Balance between
Clients, Onliners and
Online



“ We know that when great people, who share a set of common values, work together, they can accomplish great things. ”

– Chuck Loewen
President and Chief Executive Officer

DIGITAL TRANSFORMATION

CYBERSECURITY

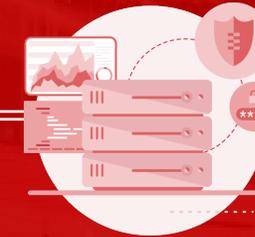
Digital Advisory
Services



Digital
Studio

AI & Data
Services

Advisory
Services



Assessment
Services

Customer
Engagement

Service
Management

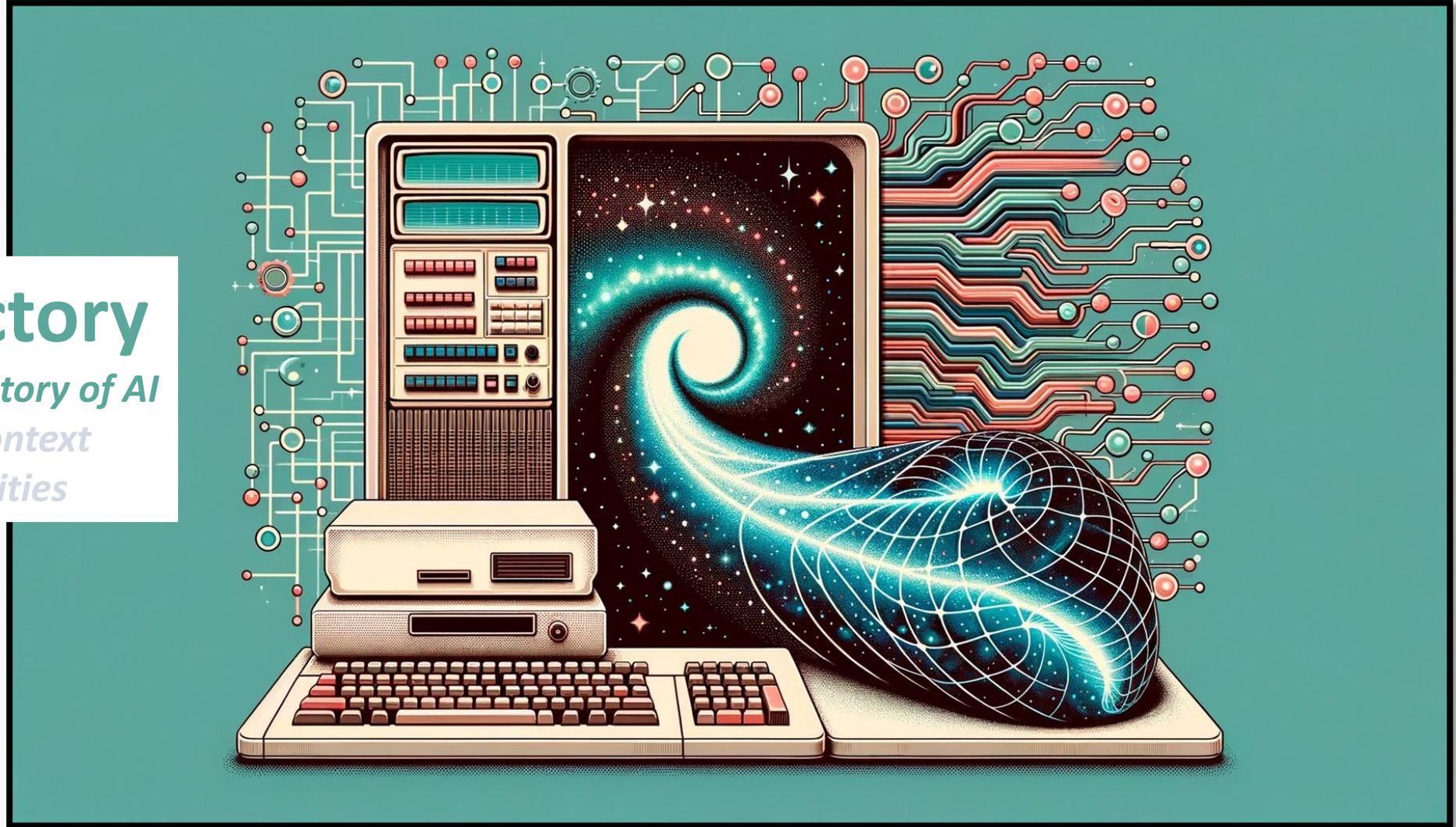
Security
Operations

Technical Security
Services

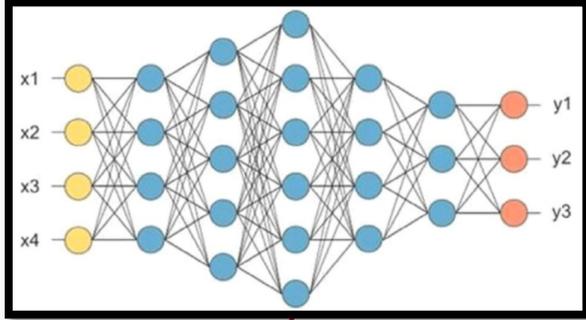
Organizational Change Management • Project Management • Quality Assurance • Cloud Deployment • Systems Integration • Application Development
Application Management Services • AI & Data Services

AI's Trajectory

-  *A Brief History of AI*
-  *Current Context*
-  *AI Capabilities*



Neural Network Conceived (1943)



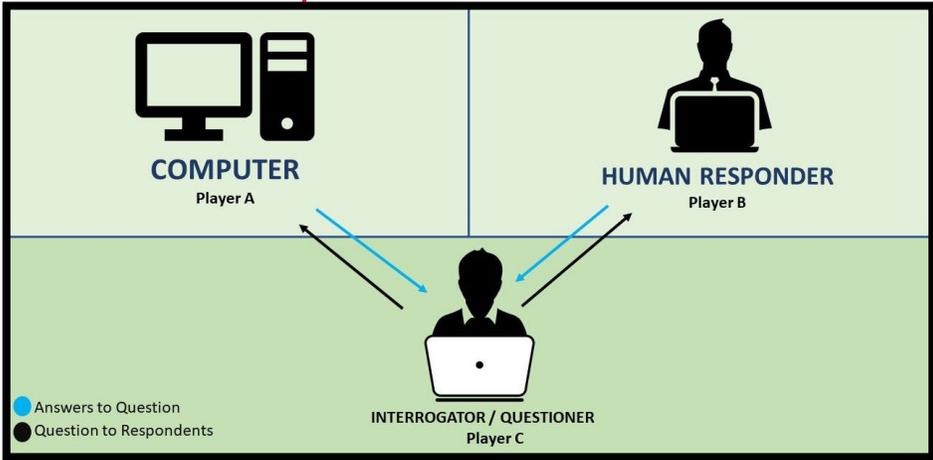
Deep Blue defeats G. Kasparov



Watson wins on Jeopardy!



Birth of AI as a field



Turing Test (1950)

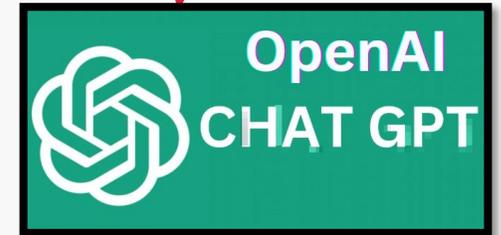
AI Winter #1

AI Winter #2

Backpropagation popularized



Roomba introduced

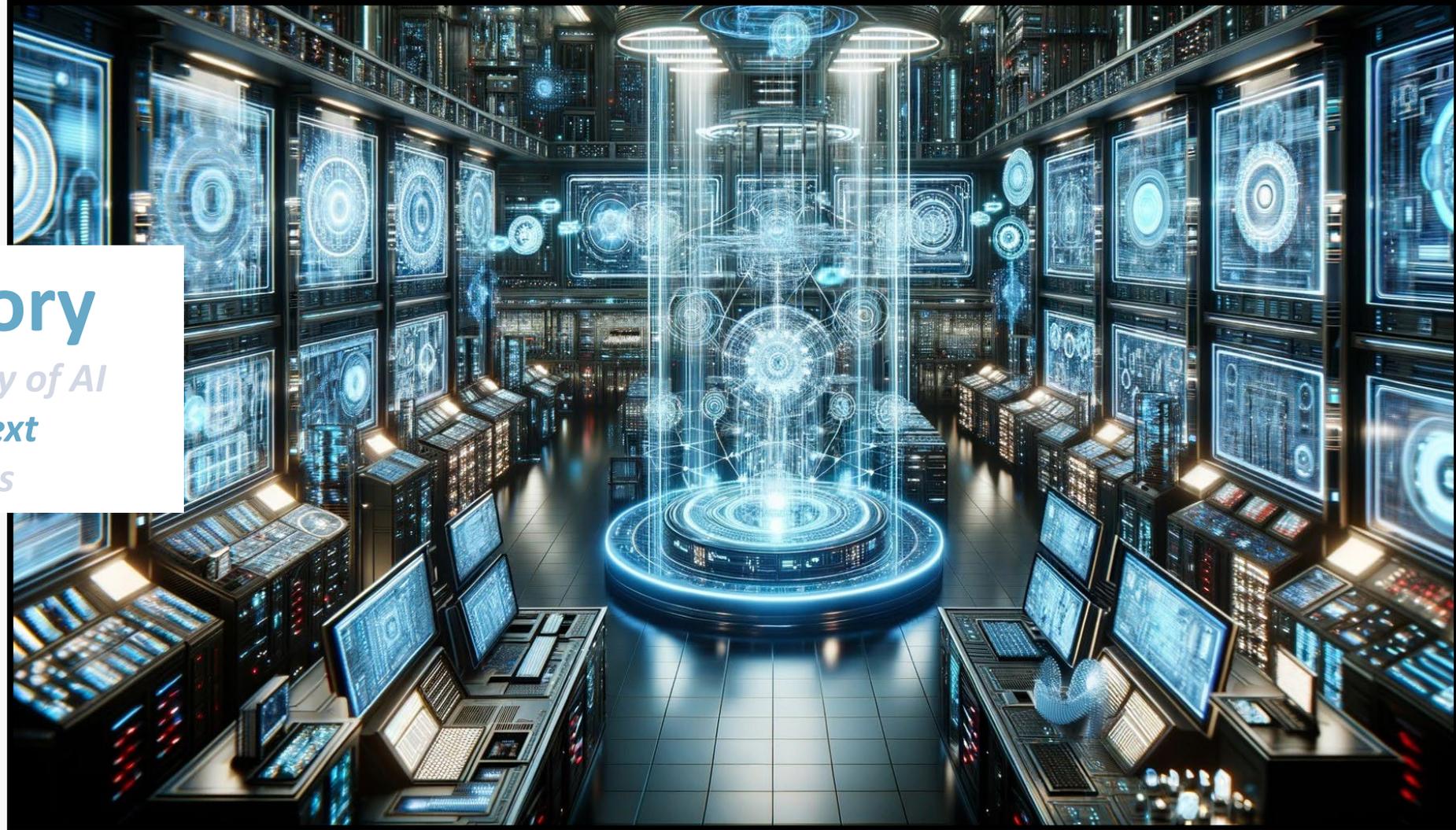


OpenAI introduces:

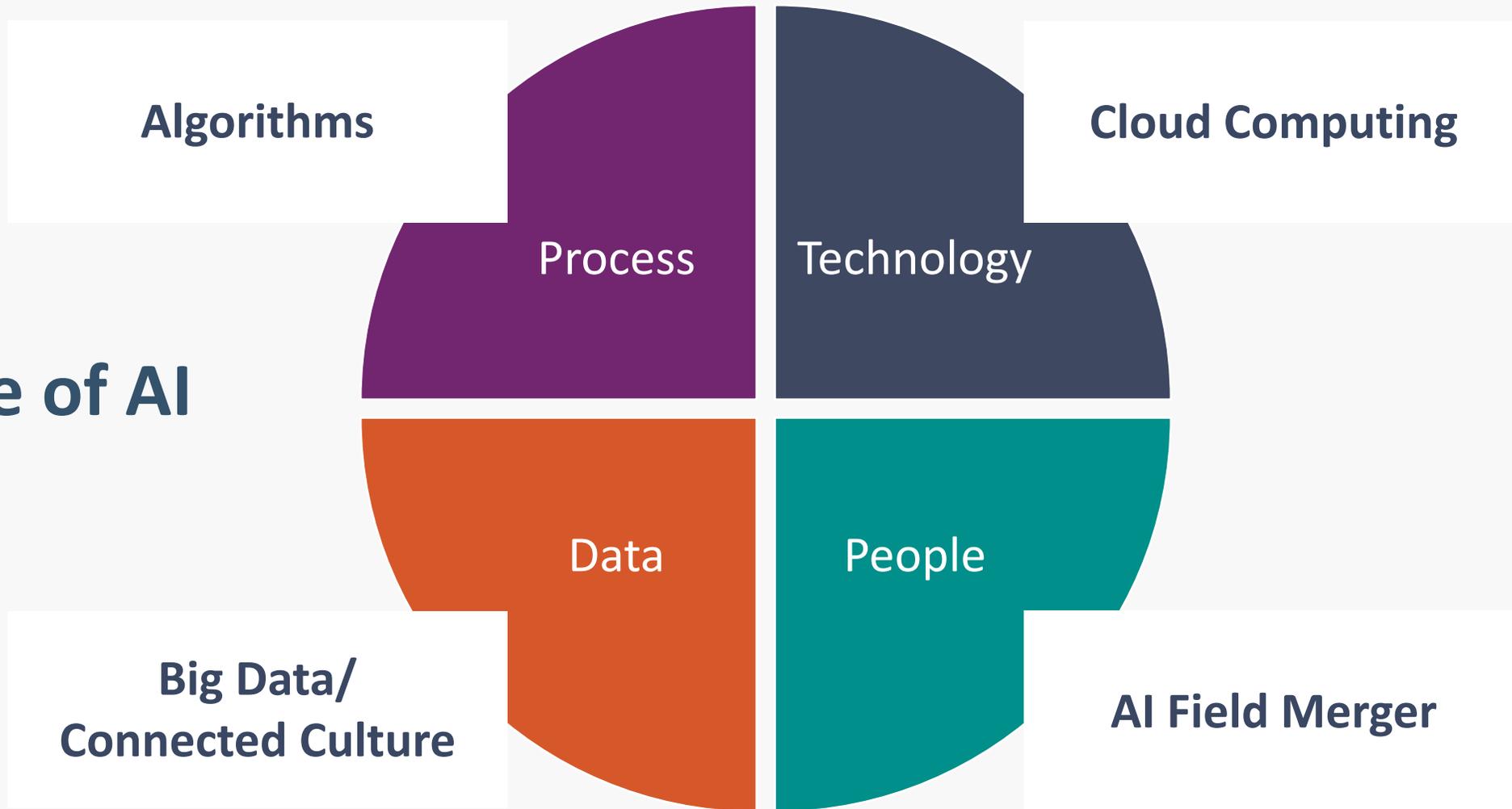
- 1) GPT-2 (2018)
- 2) GPT-3 (2020)
- 3) GPT-4 (2023)

AI's Trajectory

- 🌀 *A Brief History of AI*
- 🌀 **Current Context**
- 🌀 *AI Capabilities*



State of AI



ChatGPT Capabilities ^{1/2}

1

Answer general knowledge questions

Recommend strategies for efficient learning and information retention

2

Provide summaries of complex topics

Simplify intricate subjects into concise overviews or explanations

3

Translate text between several languages

Convert text from one language to another using my multi-lingual training

4

Assist with math problems

Help solve mathematical questions from basic arithmetic to advanced calculus.

5

Generate creative writing prompts

Offer unique ideas or starting points for storytelling and writing exercises

6

Help with grammar and spelling checks

Correct and improve text by highlighting grammatical or spelling errors

7

Offer recommendations (e.g., books, movies, places)

Suggest items or experiences based on user preferences or general popularity.

8

Describe scientific concepts

Explain principles and ideas from various scientific disciplines in understandable terms

9

Provide coding help with programming languages

Assist with code writing, debugging, and explaining programming concepts

10

Assist with homework questions

Offer guidance on academic queries across various subjects to aid learning

ChatGPT Capabilities 2/2

11

Offer tips for effective study habits

Provide answers on a vast range of topics based on my trained data

12

Generate jokes or fun facts

Provide light-hearted humor or intriguing tidbits on various topics

13

Help with resume writing and tips

: Guide users in crafting a compelling CV and job application materials

14

Provide an overview of historical events

Summarize significant moments or periods from the past

15

Describe and define philosophical concepts

Explain ideas and thinkers from the realm of philosophy

16

Offer tips for better mental health

Recommend practices and insights for psychological well-being

17

Assist with basic medical questions

Offer general information on medical topics, not to replace professional advice

18

Recommend recipes based on ingredients

Suggest culinary creations based on specified or available food items

19

Give overviews of latest tech trends*

Summarize recent technological advancements and patterns

* up until January 2022

20

Provide information on workouts and exercises

Guide users on physical routines for fitness and health.

ChatGPT Capabilities – Insurance Specific

- 1. Claims Data Analysis:** Analyzing trends and patterns from claims data to identify fraud, inefficiencies, or areas for improvement.
- 2. Risk Assessment Guidance:** Providing explanations and insights on risk assessment methodologies and models.
- 3. Policy Terms Clarification:** Helping explain the intricacies of policy terms, conditions, and coverages.
- 4. Automated Chatbot Assistance:** Offering guidance on setting up and optimizing chatbots for customer inquiries about claims or policies.
- 5. Statistical Concepts Explanation:** Assisting in understanding statistical measures used in risk assessment and claims analysis.
- 6. Document Digitization Advice:** Guidance on digitizing and storing policy documents, claims forms, and other related paperwork.
- 7. Database Management:** Assisting with queries and management tasks related to databases storing policyholder information and claims records.
- 8. Machine Learning for Fraud Detection:** Offering insights on how machine learning can be utilized for detecting fraudulent claims.
- 9. Regulatory Compliance:** Providing information about insurance industry regulations and ensuring IT systems are compliant.
- 10. Cloud Computing:** Guidance on cloud platforms and tools suitable for storing policy and claims data securely.
- 11. Insurance Tech Trends:** Offering overviews of the latest technological advancements impacting the insurance sector. (To January 2022)
- 12. Policy Recommendation Systems:** Understanding algorithms that recommend policies based on customer profiles.
- 13. Data Security and Privacy:** Recommendations and best practices for exploring how IoT devices, like connected cars or smart homes, impact claims and policy management.
- 14. Claims Automation Processes:** Offering insights into automating parts of the claims process for efficiency.
- 15. Blockchain in Insurance:** Understanding the potential applications of blockchain for policy verification and claims.
- 16. Business Continuity and Disaster Recovery:** Guidance on IT strategies to ensure business operations are uninterrupted or quickly restored after disruptions.
- 17. Customer Relationship Management (CRM) Systems:** Assisting with the implementation and use of CRM tools in the insurance domain.
- 18. Predictive Analytics for Underwriting:** Exploring how predictive analytics can enhance the underwriting process.
- 19. Digital Transformation Strategy:** Guidance on transitioning traditional insurance operations to leverage digital tools and platforms.

Example: AI-Assisted Analysis



Analysis

Idea

LLMs can be used for far more than language-oriented tasks

Process

- 1) Procure some data that you would like to analyze
- 2) Upload the data to the LLM
- 3) Ask/interact with the LLM to perform analyses

Analysis

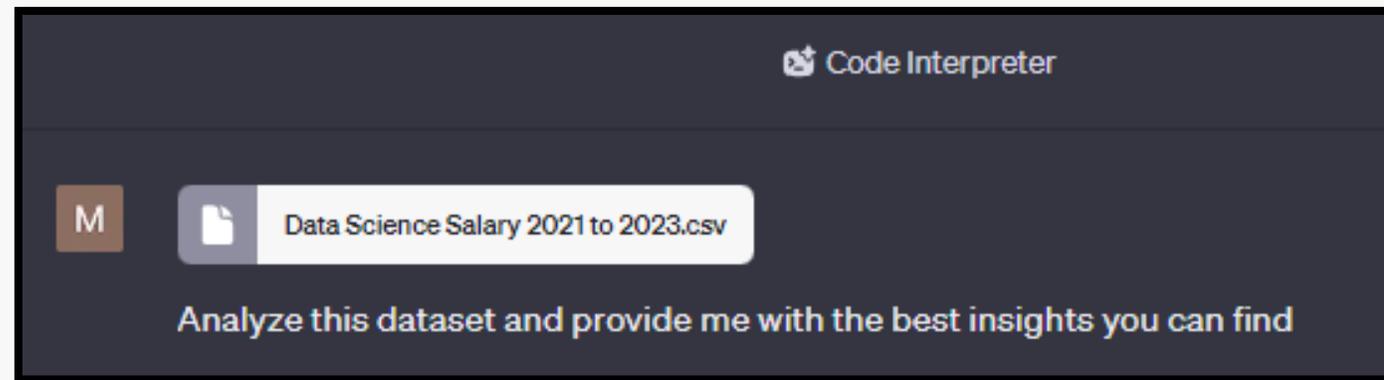
Focus

Data Science Salaries

Ref: Kaggle: <https://www.kaggle.com/datasets/harishkumardatalab/data-science-salary-2021-to-2023>

Process

- 1) Procure some data that you would like to analyze
- 2) Upload the data to the LLM
- 3) Ask/interact with the LLM to perform analyses



Analysis



First, I'll load the dataset and take a look at its initial structure to understand what kind of data we're dealing with.

Finished working Show work ▾

Based on the initial look at the dataset, we have the following columns:

1. ``work_year``: The year in which the salary was given.
2. ``experience_level``: Level of experience (the exact levels will need to be determined).
3. ``employment_type``: Type of employment (e.g., full-time, part-time, etc.).
4. ``job_title``: The specific job title.
5. ``salary``: Salary in the given currency.
6. ``salary_currency``: The currency of the salary.
7. ``salary_in_usd``: Salary converted to USD.
8. ``company_location``: Location of the company.
9. ``company_size``: Size of the company (exact size ranges will need to be determined).

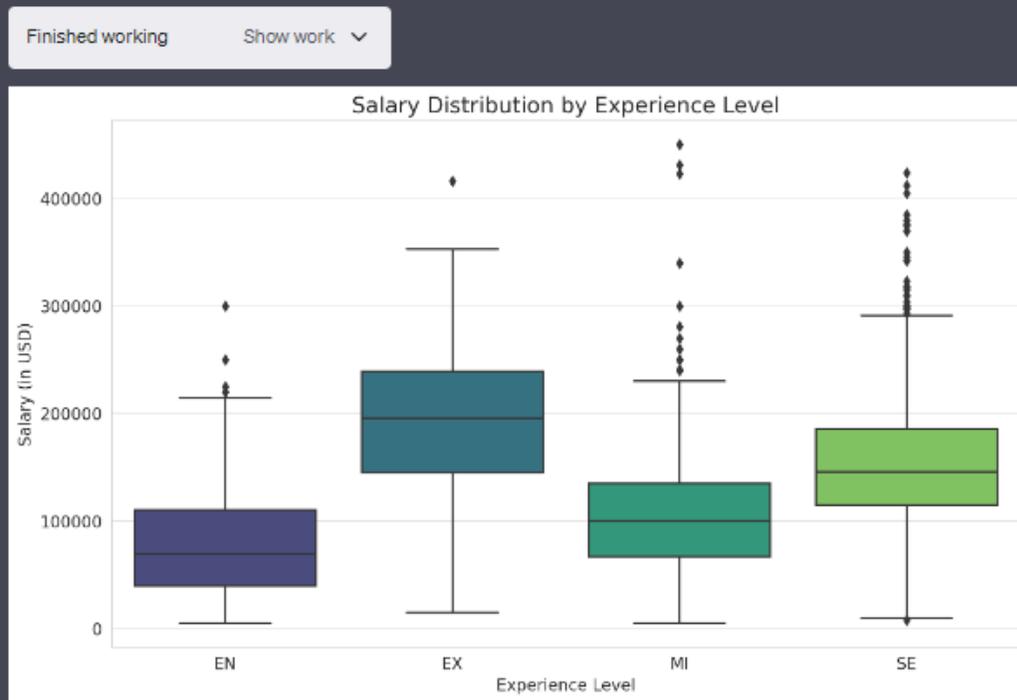
Ingestion → Metadata

Here are the key insights from the general descriptive statistics:

1. **Work Year:**
 - The dataset spans from 2020 to 2023.
 - Most of the data seems to be from 2022 based on the median value.
2. **Experience Level:**
 - There are four unique experience levels.
 - The most frequent experience level is "SE" (which we can infer as Senior Engineer or something similar, though we'd need to confirm).
3. **Employment Type:**
 - There are four unique employment types.
 - The most frequent employment type is "FT" (presumably "Full Time").
4. **Job Title:**
 - There are 93 unique job titles.
 - The most frequent job title is "Data Engineer".
5. **Salary:**
 - The average salary (across all currencies) is approximately 190,600, though this value is influenced by the salaries converted to USD.
 - Salaries range from 6,000 (minimum) to 30,400,000 (maximum). The max value seems exceptionally high and may be an outlier or error.
 - The median salary in USD is 135,000.
6. **Salary Currency:**
 - There are 20 unique currencies represented in the dataset.
 - The most frequent currency is USD, with 3,229 occurrences.
7. **Company Location:**
 - There are 72 unique company locations.
 - The most frequent location is the US.
8. **Company Size:**
 - There are three unique company sizes.
 - The most frequent company size is "M" (presumably "Medium").

Analysis

Next, I'll analyze the distribution of salaries across different categories. Let's begin by looking at the salary distribution based on the experience level.



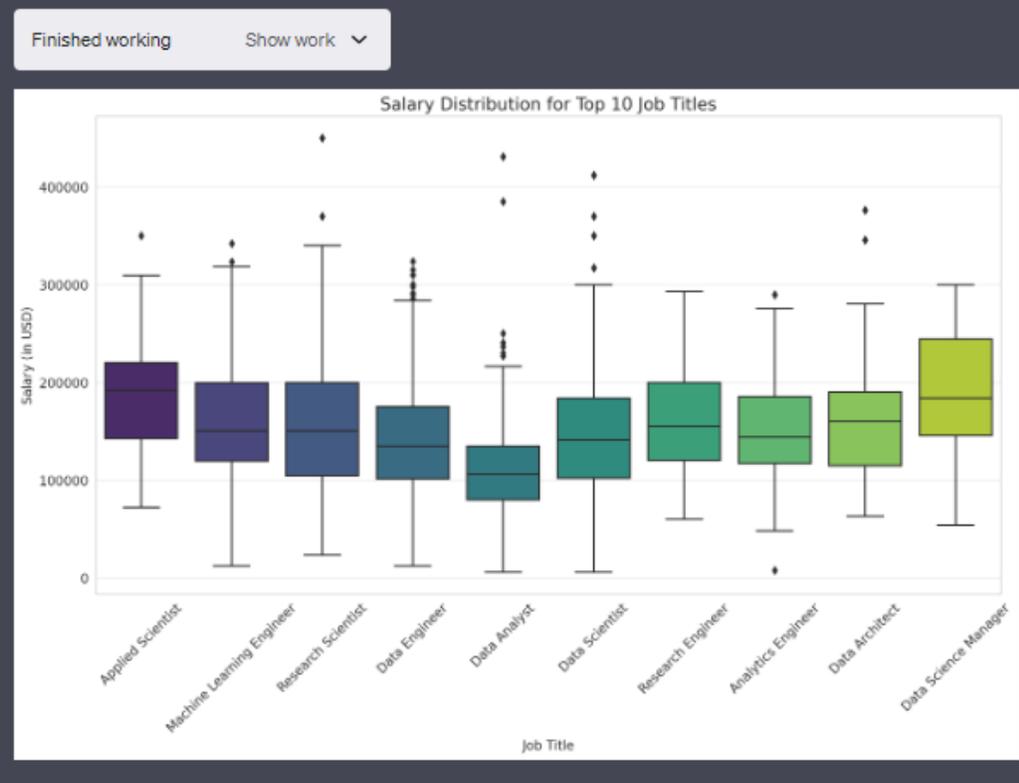
The boxplot provides insights into the salary distribution by experience level:

Tenure: $EX > SE > MI > EN$

- Outliers:** Each experience level has a few outliers, especially on the higher end of the salary range.
- Salary Range:**
 - The median salary seems to increase as we move from left to right, suggesting a trend where more experienced individuals earn higher salaries. However, we need to confirm the exact mapping of these experience level codes to understand the progression.
 - The interquartile range (IQR) appears wider for the experience levels on the left, indicating greater variability in salary for those levels.
- Highest Salaries:** The rightmost experience level (potentially the most senior level) has the highest salaries, as evidenced by its upper whisker.

Analysis

Next, let's analyze the salary distribution based on job titles. Given that there are 93 unique job titles, we'll focus on the top 10 most common job titles for clarity.



Job title specific insights

The boxplot provides insights into the salary distribution for the top 10 most common job titles:

- Variability in Salaries:** Different job titles have different salary ranges, with some roles, such as "Machine Learning Engineer" and "Data Scientist", generally having higher median salaries compared to others like "Data Quality Analyst".
- Outliers:** Several job titles, especially "Data Engineer" and "Data Analyst", have noticeable outliers on the higher end of the salary spectrum.
- Overlap in Salary Ranges:** While some roles have distinct salary distributions, there's overlap between others. For instance, the IQR (Interquartile Range) for "Machine Learning Engineer" and "Data Scientist" are somewhat similar.
- Width of the Boxes:** The width of the boxes (IQR) indicates the variability in salaries within that role. For example, "Data Scientist" has a wider IQR, suggesting a broader range of salaries within that role, whereas "Business Intelligence Developer" has a narrower IQR.

Analysis



Bigger may not be better!

The boxplot provides insights into the salary distribution by company size:

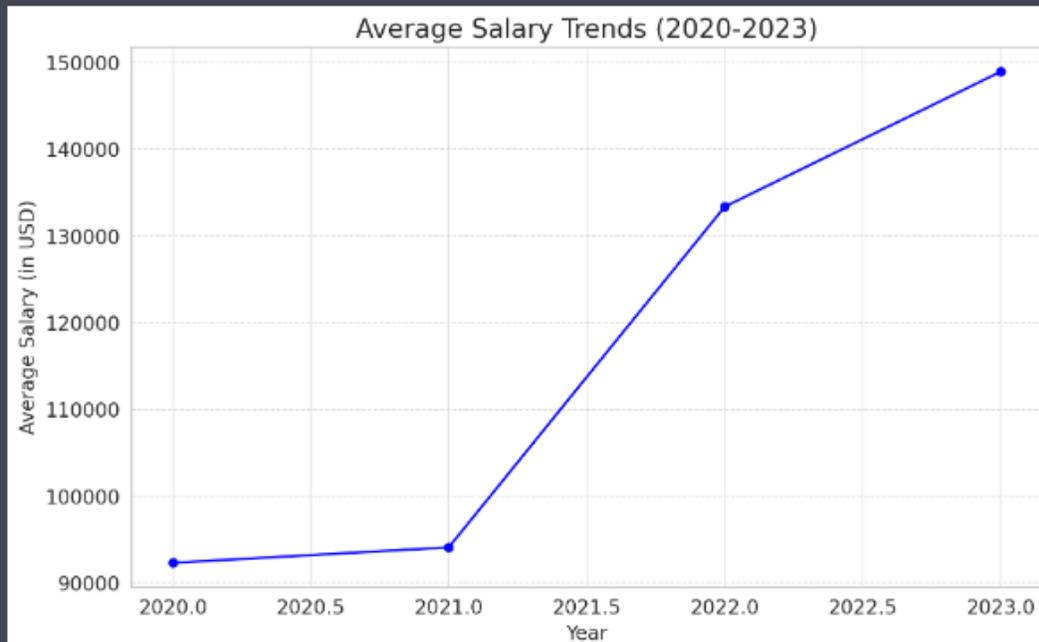
- Company Size and Median Salary:** The median salary seems to increase slightly with the size of the company. Larger companies (L) tend to offer higher median salaries compared to medium (M) and small-sized (S) companies.
- Variability:** The interquartile range (IQR) is broader for larger companies, indicating greater variability in salaries within that category. This might be due to the diverse range of roles and experience levels present in larger organizations.
- Outliers:** All company sizes have outliers, especially on the higher end of the salary spectrum. However, larger companies have a more extended range of outliers, suggesting that they might have specific roles with exceptionally high compensation.

Analysis

Next, let's explore how the average salary has changed over the years provided in the dataset.

Finished working

Show work ▾

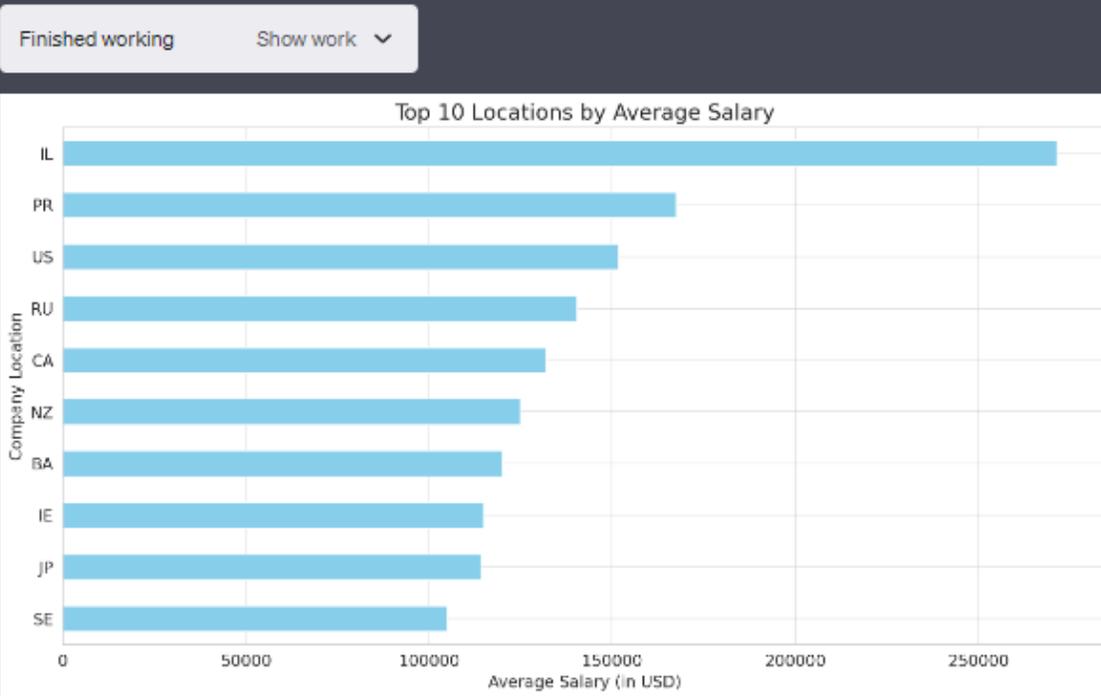


The line chart provides insights into the average salary trends from 2020 to 2023:

1. **Uptrend:** There's a clear upward trend in the average salary for data science roles from 2020 to 2023.
2. **Steady Increase:** The increase appears to be relatively steady each year, suggesting that the demand and compensation for data science roles have been consistently rising over this period.

Analysis

Lastly, let's look at the average salaries based on the company location. Given the potential diversity of locations, we'll focus on the top 10 locations by average salary for clarity.



The bar chart provides insights into the top 10 company locations by average salary:

- High Average Salaries:** The locations on this chart have notably high average salaries for data science roles. This might suggest that these regions have a high demand for data science talent or a higher cost of living that necessitates larger salaries.
- Diverse Locations:** The locations span different continents, indicating that high-paying opportunities for data science roles are not restricted to a single region or country.
- US Dominance:** The US, being a major tech hub, unsurprisingly has one of the highest average salaries for data science roles.

Analysis

Before concluding, it's essential to check for any missing data in our dataset to ensure our insights are comprehensive and unbiased. Let's do that.

Finished working Show work

Great news! There are no missing values in the dataset, which means our analyses are based on complete information.

To summarize the insights from the dataset:

1. **Experience and Salary:** Generally, individuals with higher experience levels tend to earn more.
2. **Job Titles and Salary:** Roles like "Machine Learning Engineer" and "Data Scientist" have higher median salaries compared to some other roles.
3. **Salary Trends:** The average salary for data science roles has been on the rise from 2020 to 2023.
4. **Company Size:** Larger companies tend to offer slightly higher median salaries than medium or small-sized companies.
5. **Top Locations:** Locations like the US, UK, and AU are among the top regions offering high average salaries for data science roles.

```
python
# Get the top 10 most common job titles
top_10_titles = data['job_title'].value_counts().head(10).index

# Filter the data for these top 10 job titles
filtered_data = data[data['job_title'].isin(top_10_titles)]

# Plot salary distribution for the top 10 job titles
plt.figure(figsize=(15, 8))
sns.boxplot(x="job_title", y="salary_in_usd", data=filtered_data, palette="v")
plt.title("Salary Distribution for Top 10 Job Titles")
plt.ylabel("Salary (in USD)")
plt.xlabel("Job Title")
plt.xticks(rotation=45)
plt.show()
```

Quick Rundown

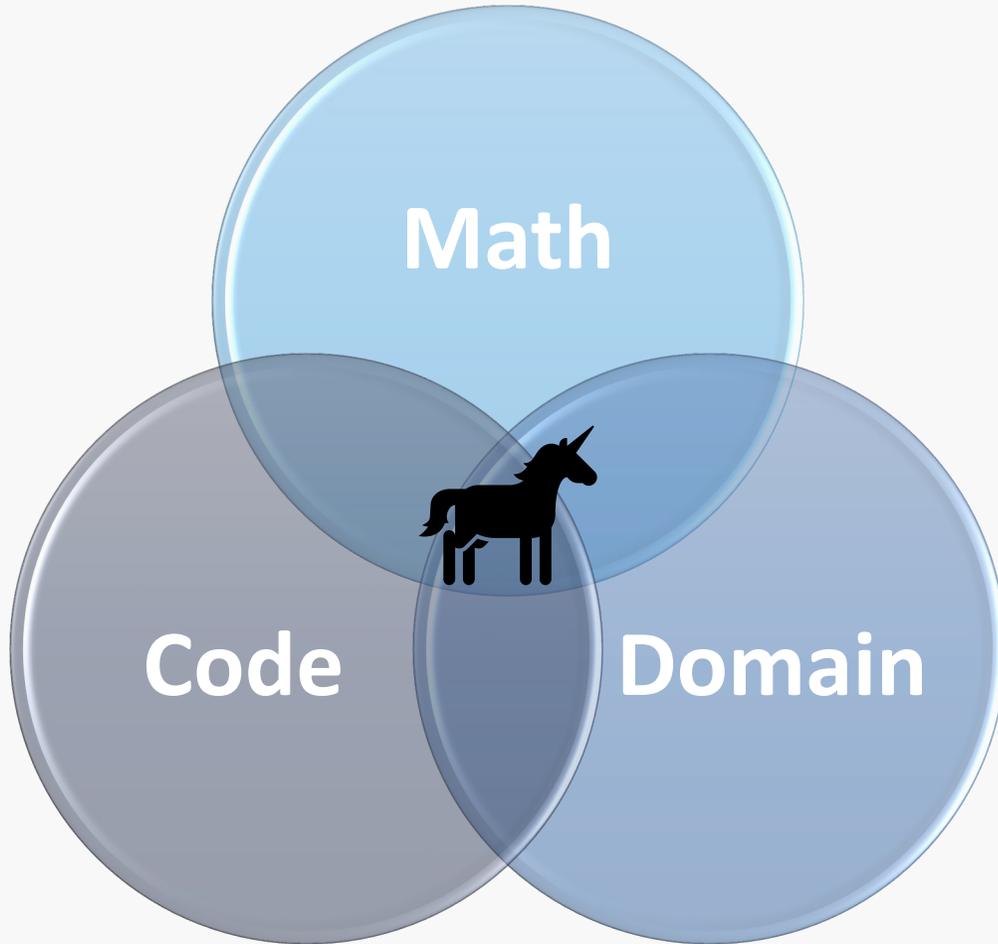
- Data ingestion
- Data defined (metadata)
- Experiential Influence Analysis
- Job Title Influence Analysis
- Company Size Influence Analysis
- Salary Trend
- Breakdown by Country
- Code generation

ALL IN LESS THAN 5 MINUTES!!

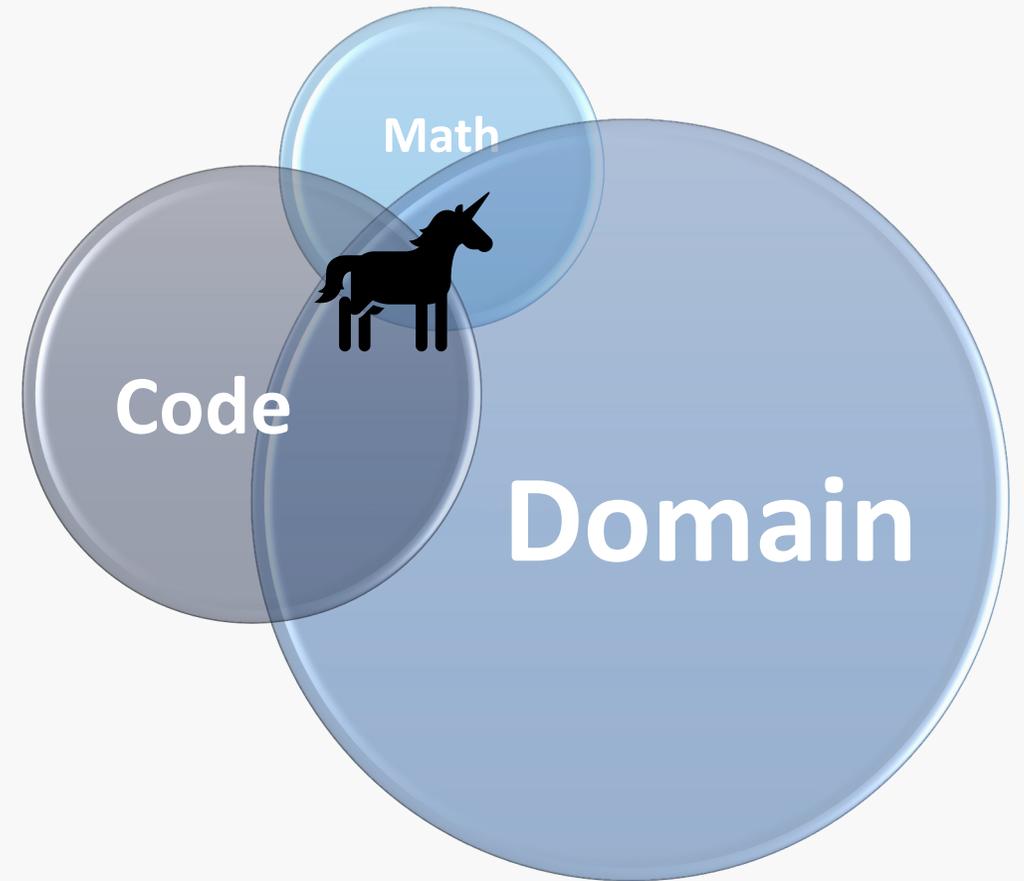
My work

- 1) Load a file
- 2) Ask the LLM to analyze the file

Current-ish



Future-ish?



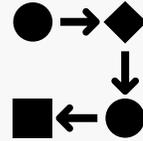
AI Use Case Identification





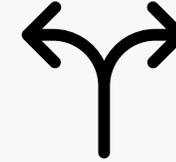
| Metrics

Identifying where D&AI can positively impact an enterprise starts by focusing on the metrics that matter the most. The workshop outcomes are largely dependent on finding the right targets.



| Process

Finding the Enterprise Processes that are most directly linked to the identified Metrics will point us in the direction of those teams that will serve as Subject Matter Experts – a requirement for success.



| Decisions

Impactful D&AI work typically aligns with the decisions made most frequently. It's also where the largest volumes of data can usually be found. Affecting the right decision can materially impact the target metric.

AI Use Case Identification

-  *Metrics*
-  *Processes*
-  *Decisions*



METRICS



Criticality

Selecting a metric that has a direct impact on the success of the enterprise is key – this impacts resource allocation discussions and buy-in.

Granularity

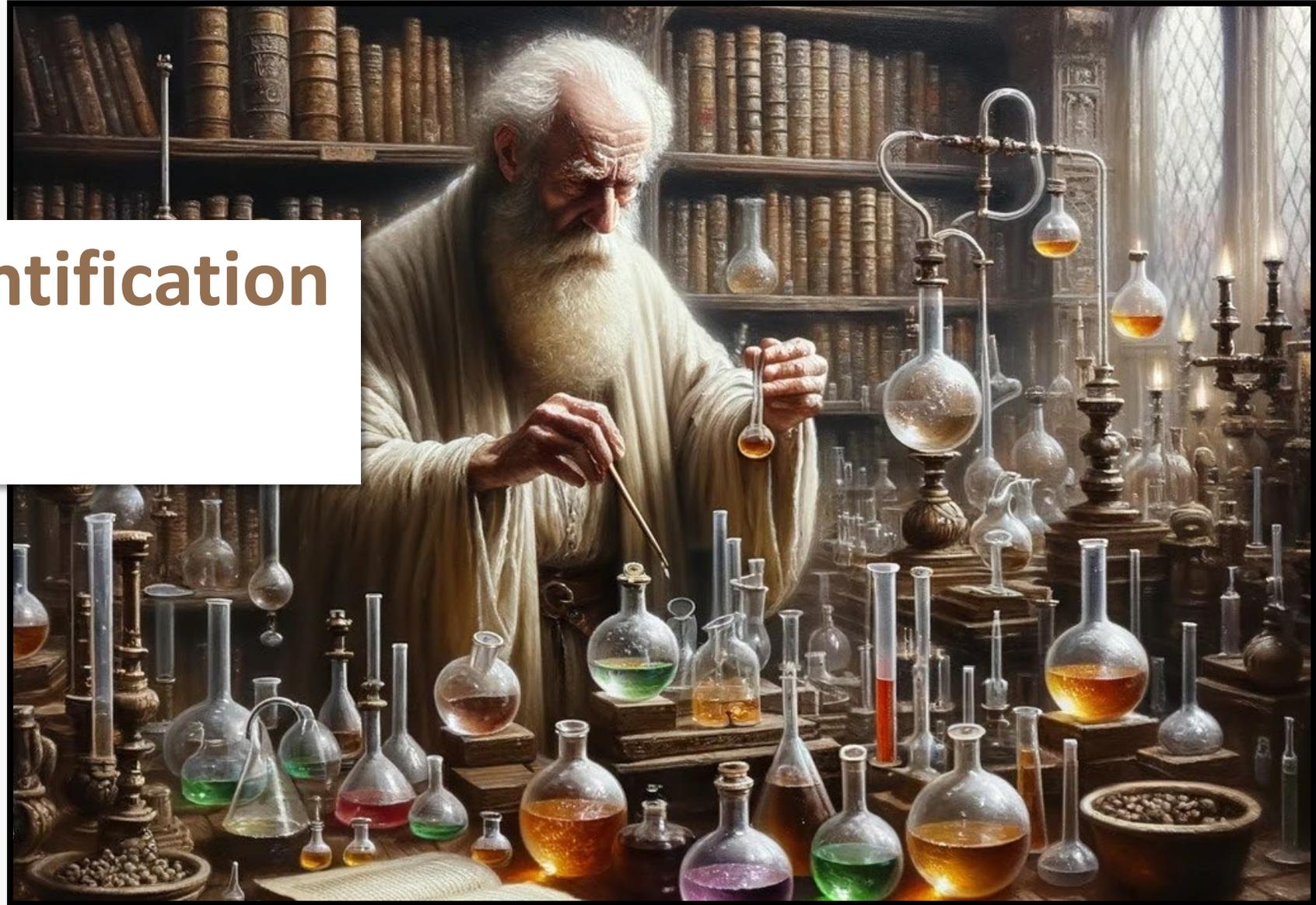
Striking a balance between the degree of granularity and meaningfulness to the enterprise ensures actionability and impact.

Enterprise Definition

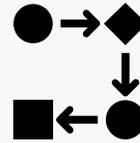
This ensures that everyone is working towards the same goals and interpreting the metrics in the same way.

AI Use Case Identification

-  *Metrics*
-  *Processes*
-  *Decisions*



PROCESS



Metric Impact

The bigger the better.

An enterprise process that has a significant impact on the selected metric should be prioritized over less impactful processes.

Data Availability

Available, accessible, and accurate data about the process can have a dramatic impact on identifying the best use cases.

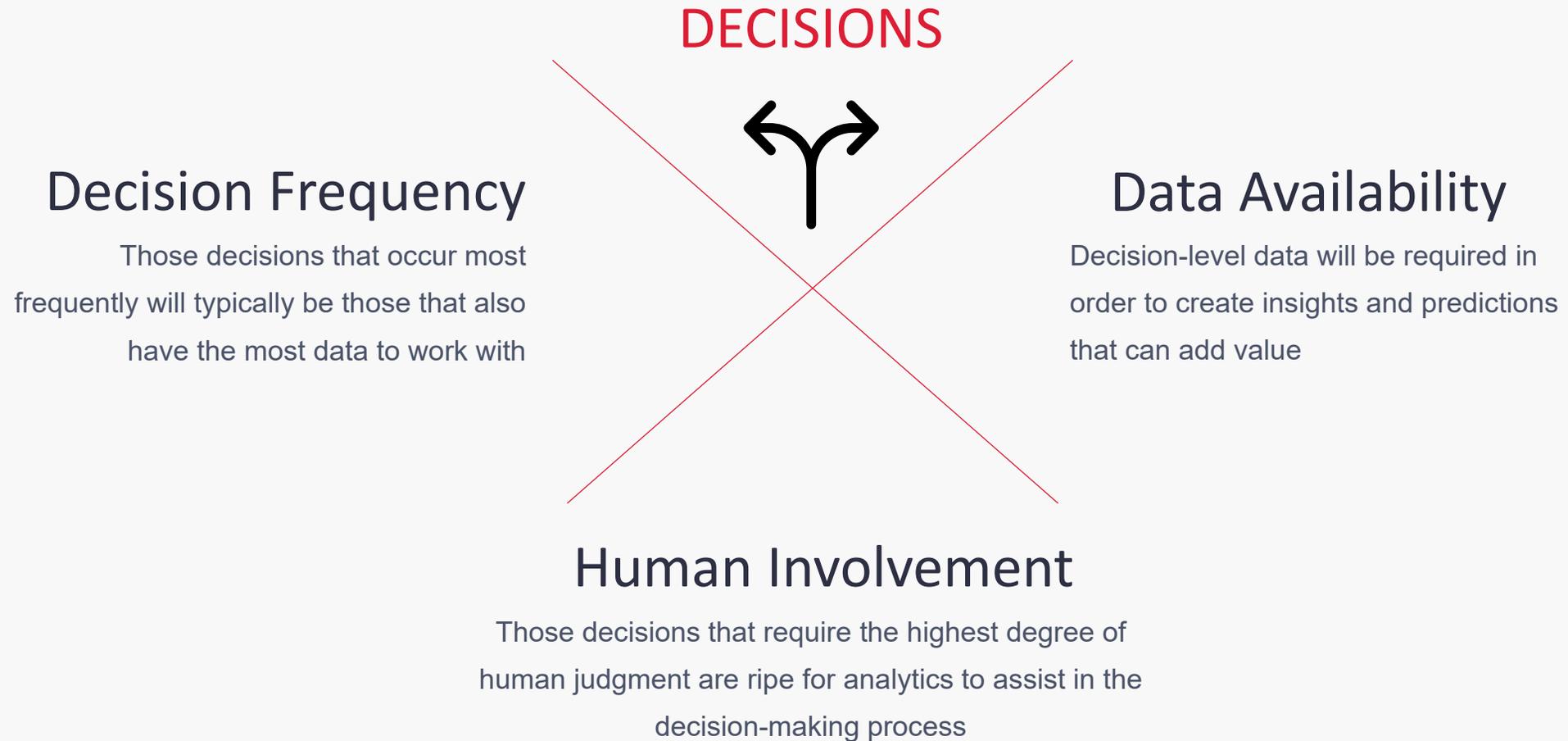
Complexity

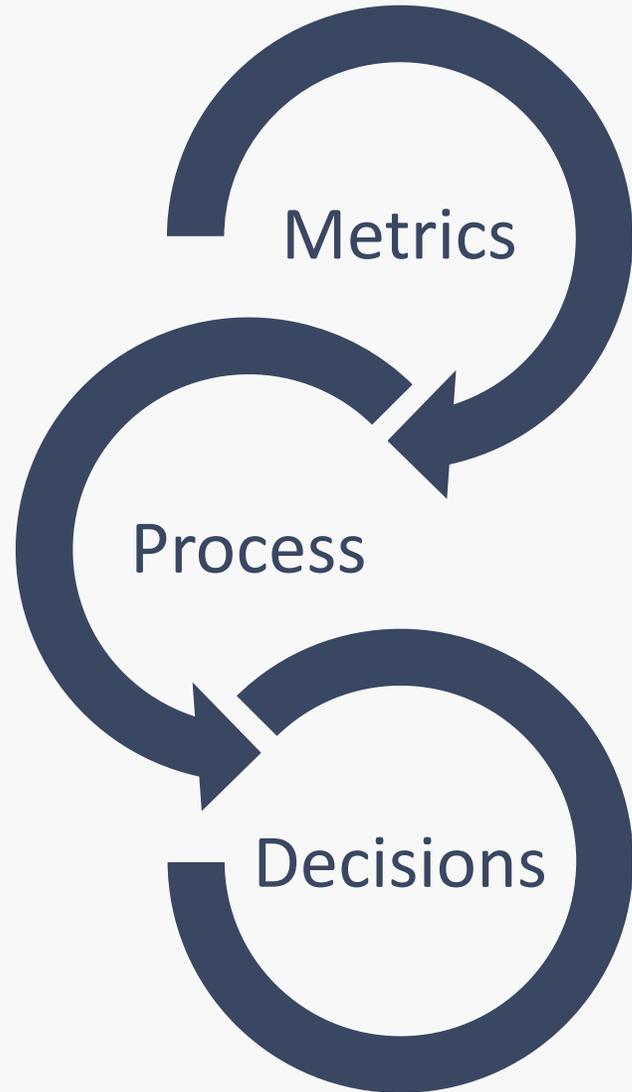
A highly complex process can make it difficult to identify the root causes, or predictors, of how the next step (key decisions) is directly linked to the selected metrics.

AI Use Case Identification

-  *Metrics*
-  *Processes*
-  *Decisions*







Metrics, Process, and Decisions are the key inputs to finding where the best use cases for Data & AI:

This will also:

- Identify what matters
- Identify who matters
- Identify where it matters

So that you can:

- Inform decisions that matter
- Affect processes that matter
- Impact metrics that matter

Risks and Limitations

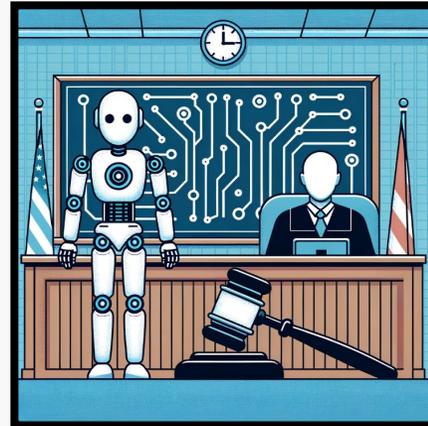
-  *Risks associated with AI*
-  *Limits associated with AI*



Ethical and Moral Risks

Economic Risks

Potential job displacement, economic inequality



Security Risks

Weaponization, adversarial attacks, cybersecurity

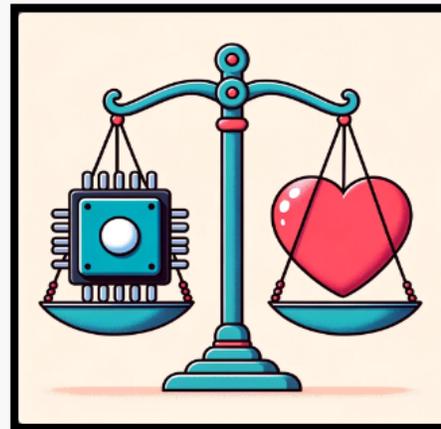


Technological Risks

Dependency

Regulatory and Legal Risks

Who's at fault when AI makes a mistake?
Can regulatory bodies and governments keep up?

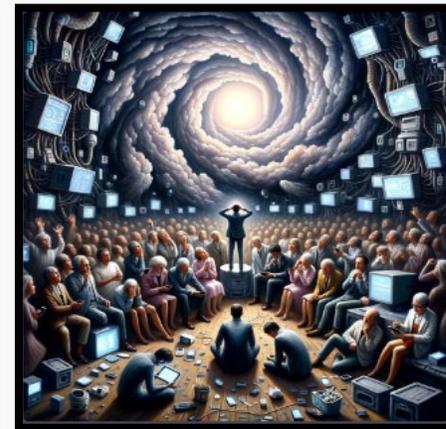


Existential Risks

In theory

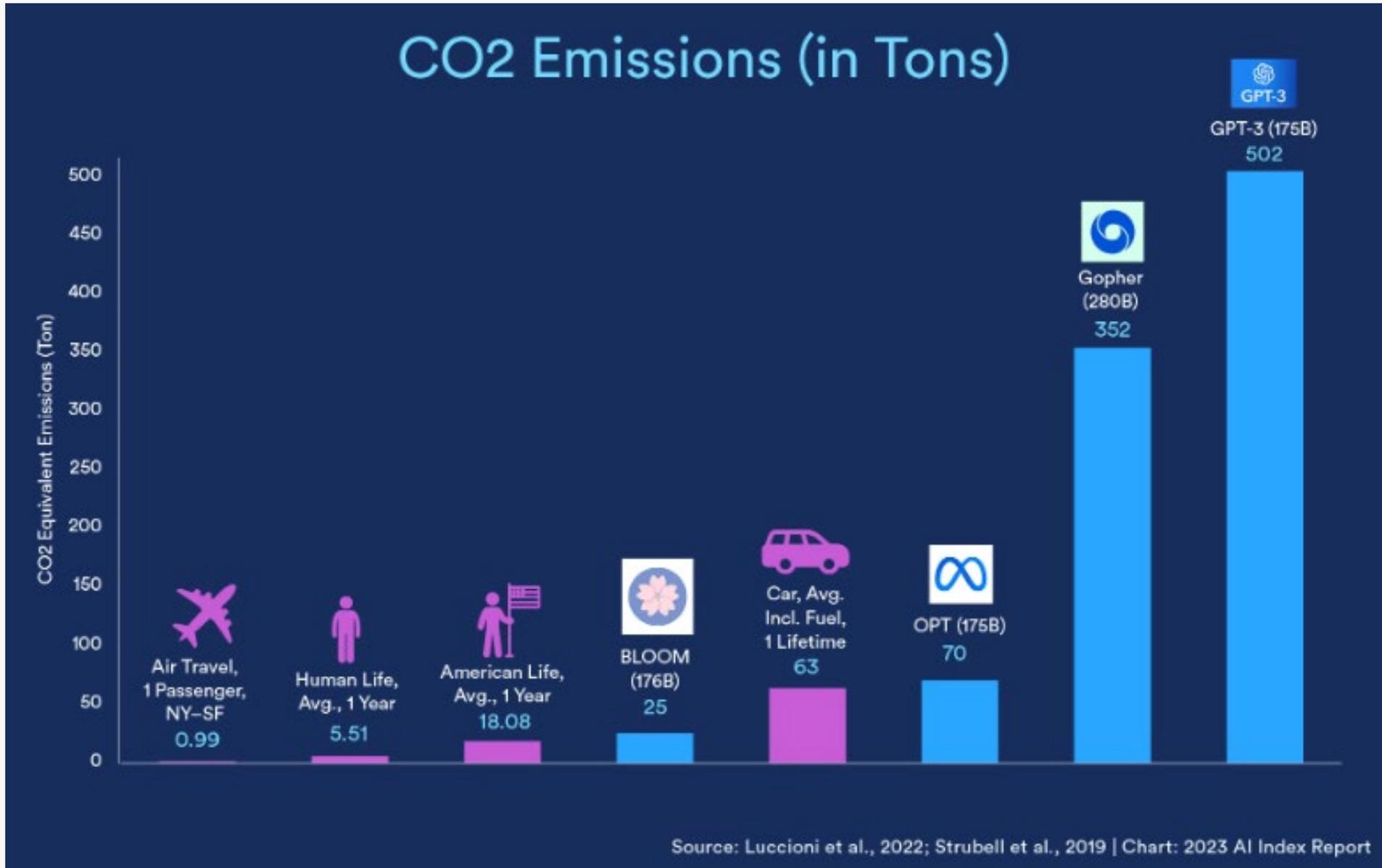
Environmental Risks

Resource consumption



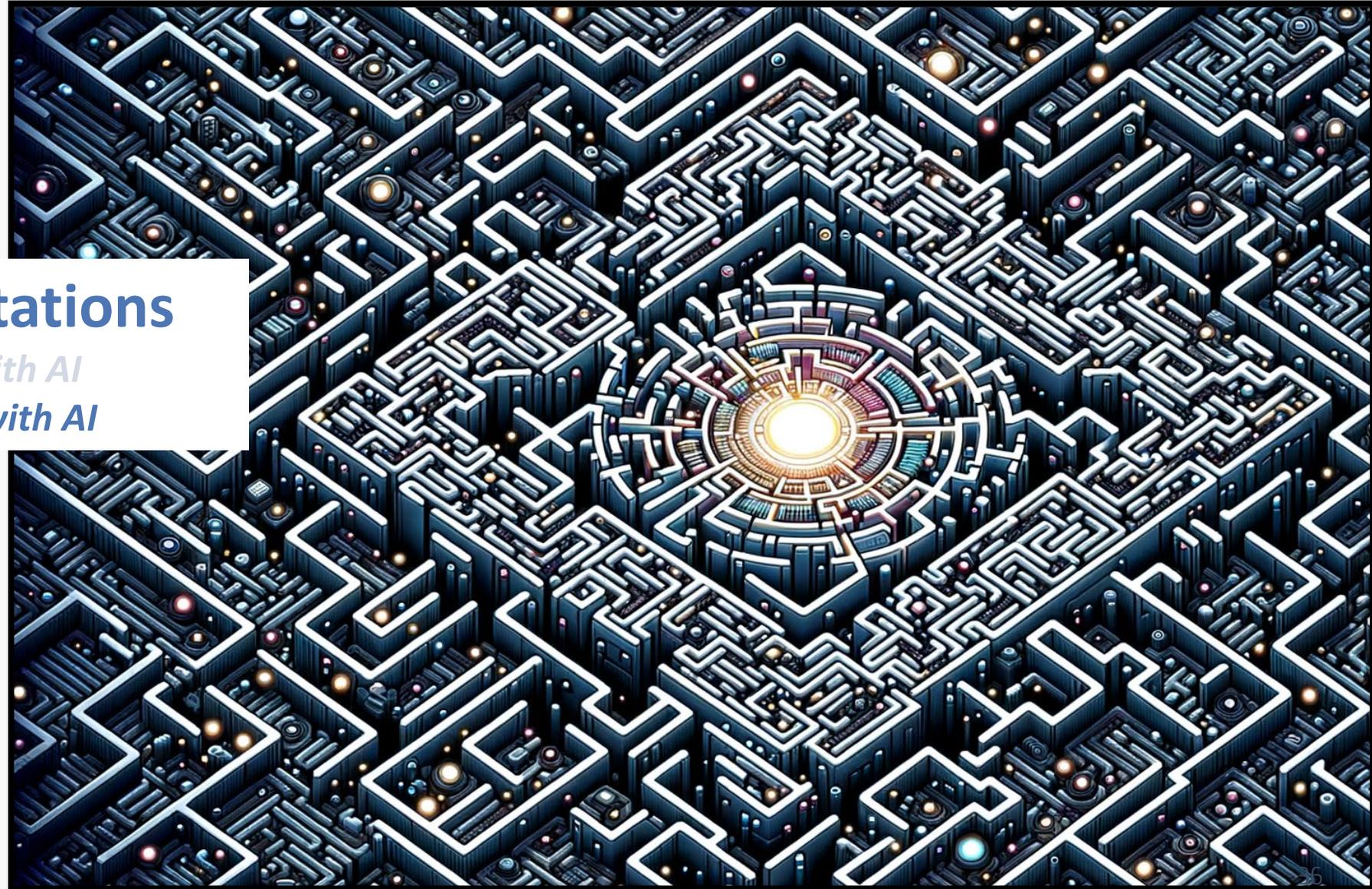
Societal Risks

Social manipulation, loss of human touch



Risks and Limitations

- ⌘ *Risks associated with AI*
- ⌘ *Limits associated with AI*



Common Limitations Associated with AI

LIMITED UNDERSTANDING OF CONTEXT

Patterns are not “understanding”

LACK OF CREATIVITY

Seemingly creative, but eerily
familiar

LACK OF COMMON SENSE

It can lead to irrational or incorrect
decisions (e.g., medical diagnosis “stickers”).
It’s like interacting with the smartest
dumb person

LACK OF ROBUSTNESS

(e.g., “jailbreaks”)

LACK OF EMOTION

Careful selection of where AI is used

DEPENDENCY

It can/will lead to biases. If past data
exhibits poor behaviour, so too will
the AI.

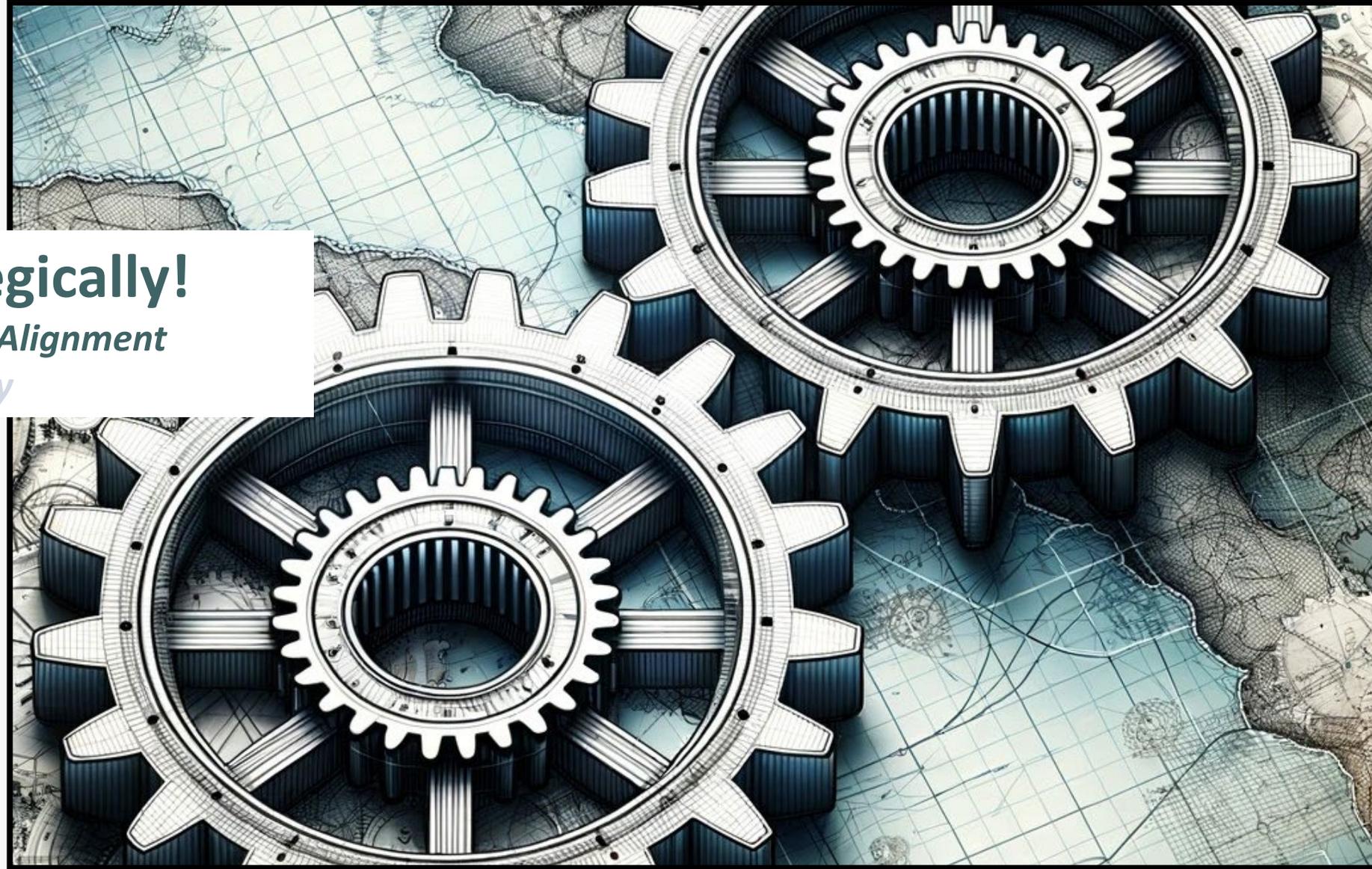
Think Strategically!



Strategic Alignment



AI Literacy



Data Strategy: The First Misstep

A Separate Data or AI Strategy tends to:

Inherently **separate your data** from our business objectives

Turn into a **Data Management initiative**, leading to a centralized bottleneck that will stifle data exploration and innovation

Create an **organization silo**

Lose sight of the **bigger picture** – that the data is intended to support the enterprise strategy and goals

An Enterprise Strategy supported by Data & AI tends to:

Align all **data-related** work with identified business objectives, and desired outcomes

Foster **enterprise level collaboration**

Identify key areas where **data can be reused** by multiple areas

Develop a **shared language** around your data assets, helping to **knock down silos**

Encourage a **broader audience** to become engaged, as they see how your data assets can support their work

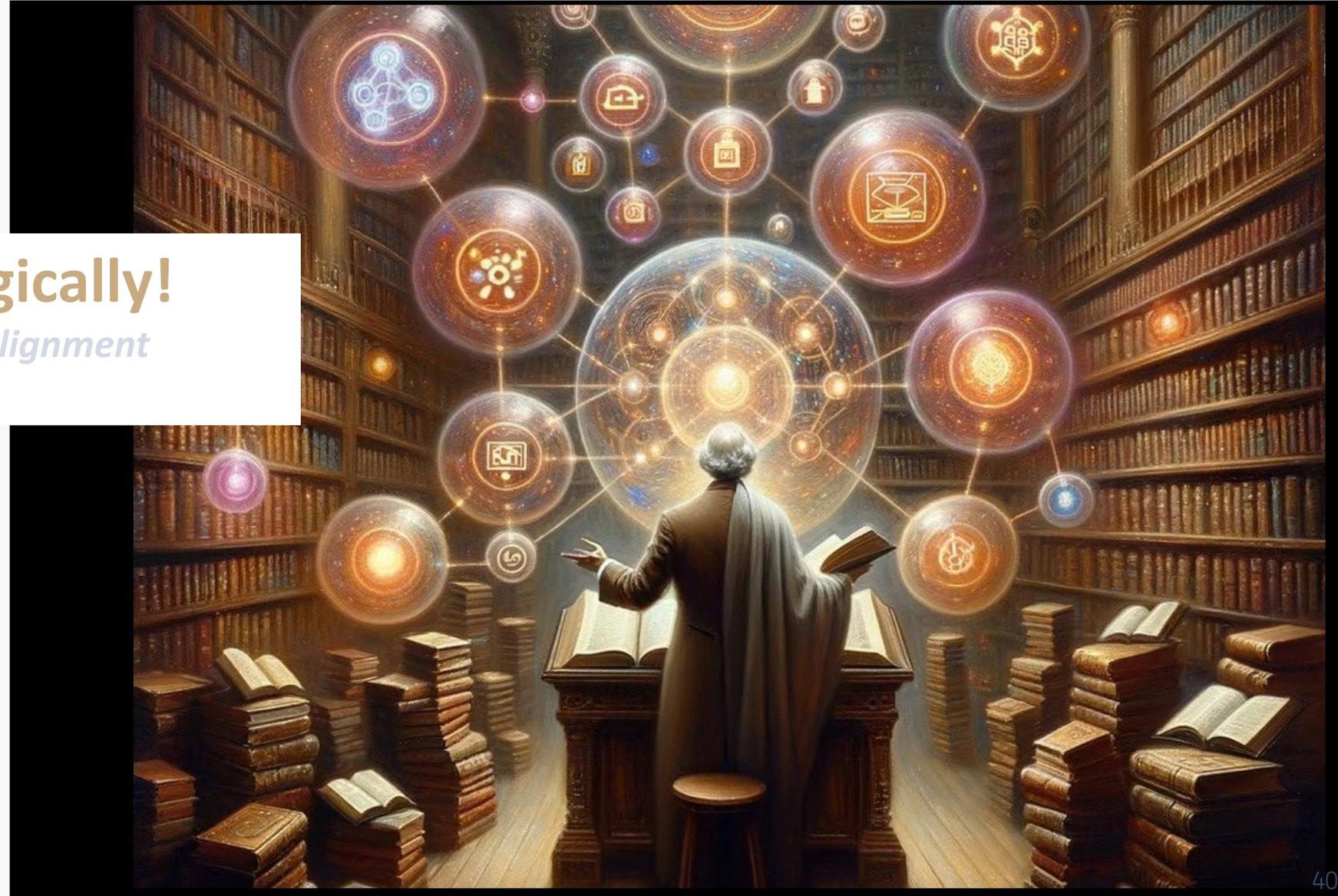
Think Strategically!



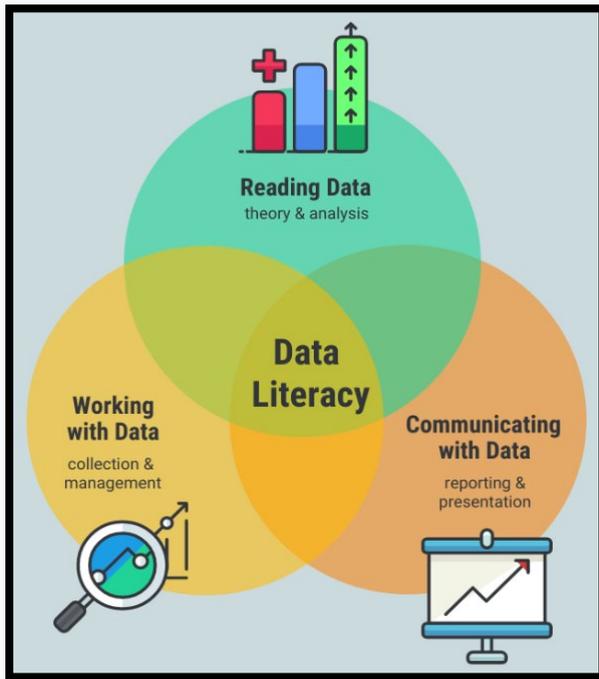
Strategic Alignment



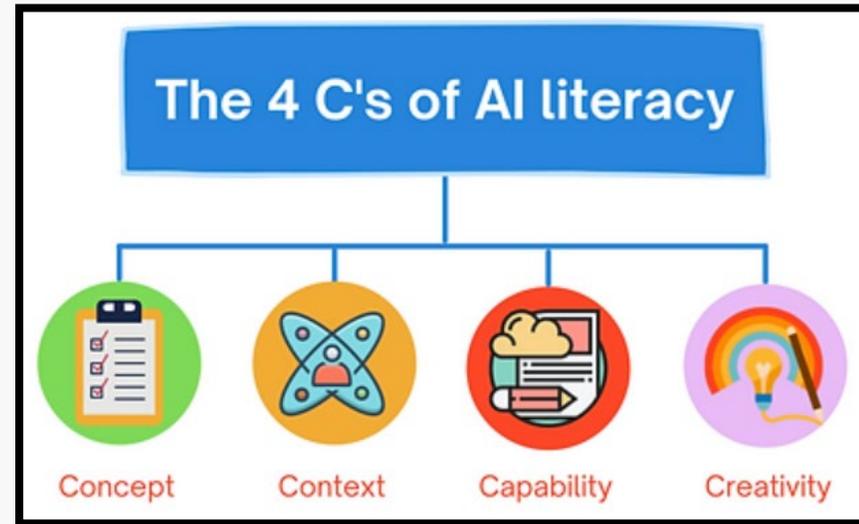
AI Literacy



Are You Ready for the Change AI will Bring?



— Data Literacy —



— AI Literacy —



— Domain Knowledge —



PARTING THOUGHTS



BUILD YOUR OWN FOUNDATION

In the “Age of AI”, what will matter most will be:

- Your range of experiences
- Your curiosity
- Your adaptability
- Your ability to ask better questions than your competition



KEYS TO SUCCESS

In the Data & AI space, success will come down to your ability to:

- 1 - Understand your Data
- 2 - Understand your Process
- 3 - Understand your Choices

THANK YOU

<https://info.obsglobal.com/csio-presentation-slides>

Mark Struck: mstruck@obsglobal.com