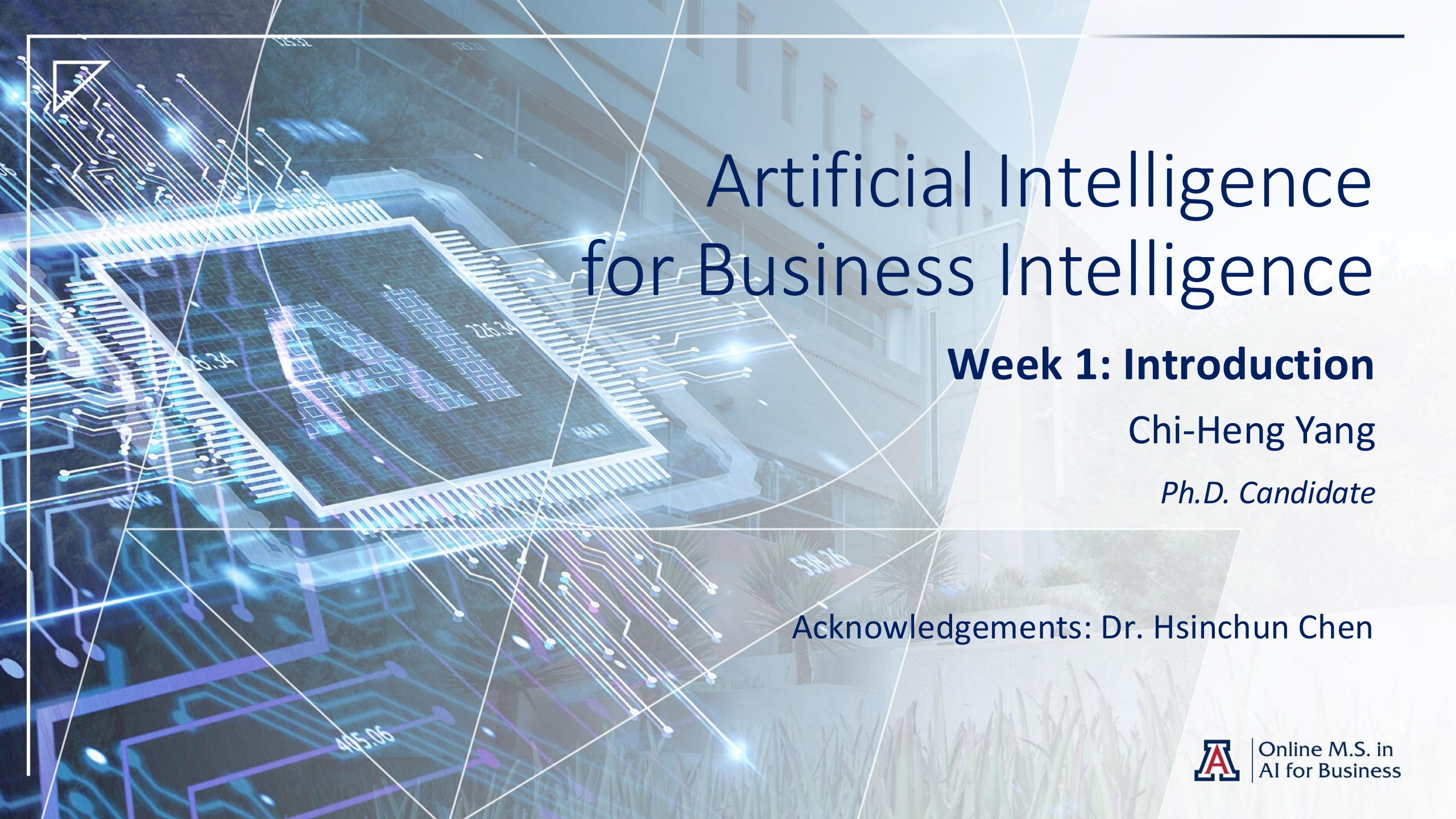




Artificial Intelligence for Business Intelligence

Week 1: Introduction

Chi-Heng Yang

Ph.D. Candidate

Acknowledgements: Dr. Hsinchun Chen



Online M.S. in
AI for Business

Introduction – Course Scope

Overview (Week 1)



Introduction to Business Intelligence



AI for Business Intelligence (AI4BI)
Development Process

Artificial Intelligence (Week 2 to 4)



Machine
Learning



Deep
Learning



Large Language
Model



Graph
Neural
Network



Time Series
Analysis



Reinforcement
Learning

Business Intelligence (Week 5 to 7)



Customer analytic



Supplier analytic



Operational analytic



Grading Policy

- ▶ Exams

- ▶ Exam 1 75 pts (15%)
- ▶ Exam 2 75 pts (15%)

- ▶ Course Project

- ▶ Milestone 1 100 pts (20%)
- ▶ Milestone 2 100 pts (20%)
- ▶ Milestone 3 100 pts (20%)
- ▶ Milestone 4 50 pts (10%)



Exam

- ▶ Two Exams
 - ▶ Exam 1 covers week 1 to week 4 (Introduction + Artificial Intelligence)
 - ▶ Exam 2 covers week 5 to week 7 (Business Intelligence)
- ▶ Each exam will contain 6-8 essay-format questions
 - ▶ Open book and open note
 - ▶ One week to complete the exam from when it opens

Course Project

- ▶ **Members:** 1-4 students per group
- ▶ **Task:** Develop AI analytics to provide BI insights by going through an entire AI4BI process.

Milestone	Task
Proposal	• Project team selection • Industry selection
Milestone 1	• Problem identification (stakeholder, scenario, current approaches, and limitaitons)
Milestone 2	• Data collection and extraction • Summary of collection and extraction
Milestone 3	• Analytics development and preliminary visualization
Milestone 4	• Visualization and business value delivery



Today's Learning Objectives

- ▶ **Understand the role of Business Intelligence (BI)** in transforming raw data into actionable insights.
- ▶ **Explain how Artificial Intelligence (AI)** enhances BI through prediction, automation, and personalization.
- ▶ **Understand the AI4BI development process**, including problem identification, data collection, analytics, and visualization.
- ▶ **Recognize key characteristics of good BI**, such as relevance, timeliness, and actionability.

Motivation

- ▶ A retail company has large volumes of data from multiple sources.
 - ▶ Sales transactions
 - ▶ product categories
 - ▶ Customer demographics
 - ▶ Store performance
 - ▶ Inventory levels
 - ▶ Marketing activities
- ▶ But these are raw numbers and records:
 - ▶ Hard to get a unified view of the business?
 - ▶ Difficult to spot issues early (e.g., low stock, declining sales)!
- ▶ How to know what has happened and what is likely to happen next?



Motivation

- ▶ We need to analyze the **data** to provide **actionable insights**.

Scenario	Data	Insights and Actions
Understanding which products are performing well	• Sales by product • Customer purchase counts	• Promote high-performing products • Improve or discontinue low performers • Adjust pricing or bundling
Forecasting future product demand	• Historical sales data • Seasonality patterns	• Prepare inventory before spikes • Reduce excess inventory costs • Optimize warehouse staffing
Identifying customers likely to churn	• Customer purchase frequency • Past churn patterns	• Launch retention campaigns • Offer personalized incentives • Improve service for at-risk groups

Business Intelligence

- ▶ This is where Business Intelligence comes in
 - ▶ Integrates data from multiple sources
 - ▶ Provides clear, interactive insights for decisions
 - ▶ Helps track performance, detect issues, and forecast outcomes
- ▶ **Business Intelligence** is the process of transforming data into actionable insights to support strategic, tactical, and operational decisions.



Business Intelligence

- ▶ BI is traditionally produced by
 - ▶ Analysts gather data from multiple systems
 - ▶ Teams build dashboards and reports to explain what happened (e.g., Tableau, PowerBI)
- ▶ Insights depend on manual exploration and interpretation.
- ▶ However,
 - ▶ This approach relies heavily on humans to spot patterns or anomalies.
 - ▶ Data volumes are getting larger and larger nowadays.
- ▶ Managers get information, but not **answers**, **predictions**, or **recommendations**.





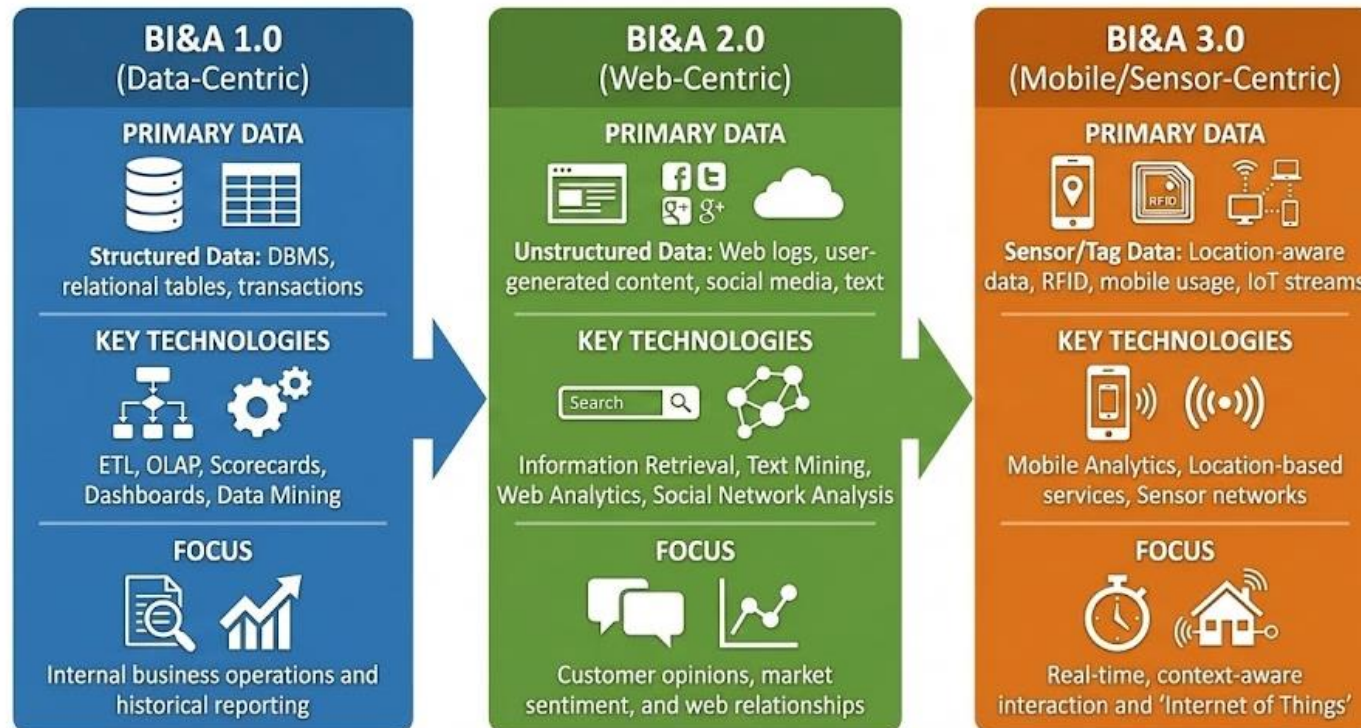
Artificial Intelligence for Business Intelligence

- ▶ **Artificial Intelligence (AI)** refers to computational techniques that allow machines to **learn patterns**, **make predictions**, and **automate decisions** from data.
- ▶ AI enhances BI by adding intelligence and insights on top of data and dashboards:
 - ▶ **Forecast** future trends such as product demand, sales, or churn
 - ▶ **Detect anomalies** early (e.g., sudden drops in sales, abnormal inventory levels)
 - ▶ **Segment customers** automatically based on behavior and value
 - ▶ **Recommend actions** such as optimal pricing, promotions, or inventory adjustments
 - ▶ **Personalize decisions** for specific products, stores, or customer groups

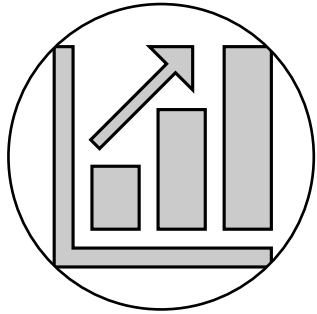
Evolution of BI&A

Business Intelligence & Analytic (BI&A) is the techniques, technologies, systems, practices, methodologies, and applications that analyze critical business data to help an enterprise (Chen et al. 2012):

- Better understand its business and market
- Make timely business decisions.

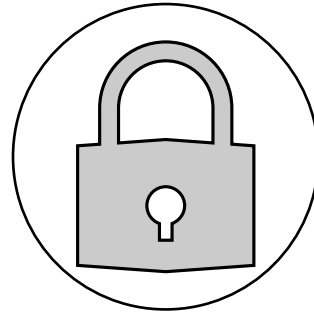


BI&A Applications



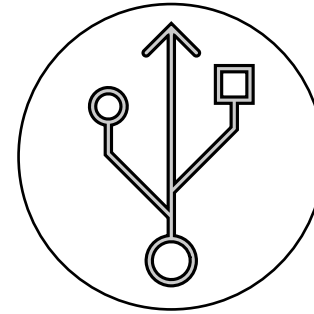
E-Commerce & Market Intelligence

- Analyzing clickstreams, social media, sentiment for marketing & recommendations.
- **Examples:** Personalization engines (Netflix, Amazon), advertising optimization (Google).



E-Government & Security

- Leveraging text mining & social network analysis for national security, crime.
- **Focus:** "Terrorism Informatics," public safety, intelligence.



Science & Technology

- Analyzing massive datasets for accelerated discovery & research.
- **Examples:** Genomics, astronomical data, materials science.



Health & Smart Systems

- Applying analytics to Electronic Health Records (EHRs), smart grids, logistics.
- **Focus:** Predictive diagnostics, operational efficiency.

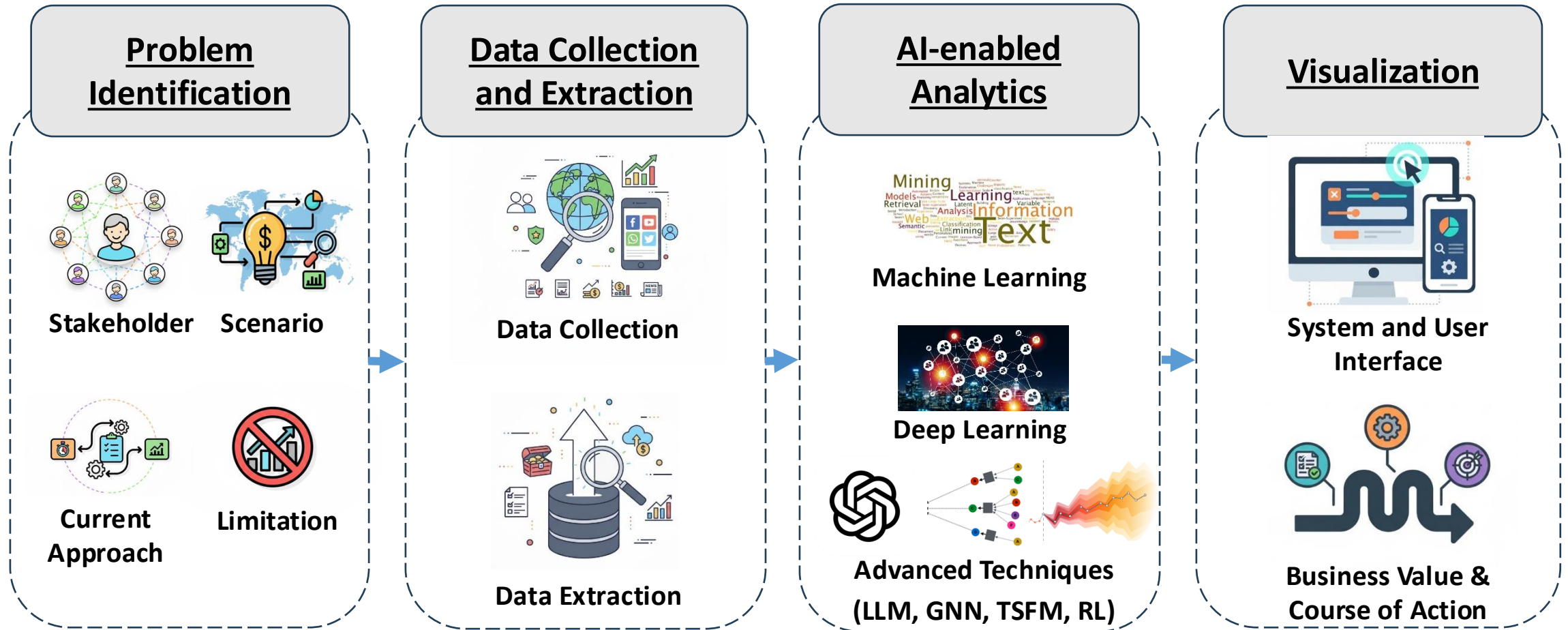




Characteristics of Good Business Intelligence

- ▶ In this course, we will be focusing on the BI in **E-commerce & Market Intelligence**.
- ▶ **BI must be:**
 - ▶ Formatted in a consumable format
 - ▶ Relevant to the business questions being asked
 - ▶ Actionable
 - ▶ Timely
 - ▶ Able to help an organization forecast, plan, and make data-driven decisions
- ▶ BI loses **significant** value over time if data becomes outdated.
- ▶ As such, it must be produced in a **systematic, repeatable, and efficient** manner.

AI4BI Development Process



Problem Identification



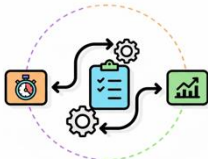
Stakeholder

- **Primary user:** The executives, managers, and analysts who will actively use the solution
- **Responsibility:** What are the responsibilities of the stakeholders?



Scenario

- **Task:** What key decision does the user need to make?
- **Process:** What part of the workflow are they trying to improve or optimize?



Current approach

- **Data source:** What type of data do they currently use?
- **Current method:** Do they use a specific rule-based method, an analytic method, or do it manually?



Limitation of this approach

- **Data-wise:** Are there large amount of data to process? Are there complex patterns in the data?
- **Process-wise:** Is the process highly manual and time-consuming?

Problem Identification (Example)



Stakeholder

- **Primary user:** Digital Marketing Analyst
- **Responsibility:** Oversees campaign performance and optimization of ad spend.



Scenario

- **Task:** Identify which campaigns to scale, pause, or adjust to improve ROI.
- **Process:** Continuous monitoring of multi-channel marketing performance (Google Ads, Meta Ads, Email).



Current approach

- **Data source:** Ad platform dashboards (Google, Meta) + email tool exports data daily.
- **Current method:** Analyst manually copies metrics into Google Sheets and builds weekly performance reports.



Limitation of this approach

- **Data-wise:** Multiple fragmented metrics across channels prevent unified insights
- **Process-wise:** Data extraction and manual tracking take several hours and delays insights.

Problem Identification (Example)



Stakeholder

- **Primary user:** Inventory planning manager
- **Responsibility:** maintaining optimal inventory levels across retail stores



Scenario

- **Task:** Determine which products need to be reordered and in what quantity
- **Process:** Weekly inventory review and replenishment planning across all store locations



Current approach

- **Data source:** POS system exports daily sales and stock data into CSV files.
- **Current method:** Planner manually merges CSVs in Excel and uses pivot tables to estimate reorder needs.



Limitation of this approach

- **Data-wise:** Complex demand patterns, seasonality, or anomalies across stores
- **Process-wise:** Reorder decisions rely heavily on manual judgment and basic pivot tables, which can be subjective



Data Collection and Extraction

- ▶ Our next step is to collect the data that is used in current approaches.
- ▶ Before collection, we explore and prioritize the data source to identify their:
 - ▶ Metadata and their data type
 - ▶ Data volume and frequency
 - ▶ Collection method
 - ▶ Extraction method
- ▶ This helps understand the key data sources for our AI-enabled analytics.

Data Collection and Extraction (Example)

Data Source/Platform	Metadata & Data Types	Data Volume & Frequency	Collection Method	Extraction
News Articles (e.g., BBC, Reuters)	- Title (text) - Author (text) - Publish date (datetime) - Article body (text)	100+ a day daily or hourly	Public news APIs	- Keyword extraction - Named entity extraction (people, places, orgs)
Social Media Posts (e.g., Reddit, Twitter/X)	- Username (text) - Post text (text) - Timestamp (datetime) - Upvotes/likes (numerical)	1M+ a day updated in real time	Public APIs (Reddit API, Twitter API)	- Hashtag/keyword extraction - Entity extraction - Sentiment extraction
Open Government Data (e.g., NOAA weather data)	- Station ID (text) - Latitude/Longitude (numerical) - Observation date/time (datetime) - Temperature (numerical) - Measurement quality flags (text)	100+ a day hourly	- Manual downloads (CSV, JSON) - Public APIs	- Unit normalization (°C/°F) - Missing-value handling - Aggregation (daily → monthly)

Data Collection and Extraction

- ▶ Each data source contains many types of metadata
 - ▶ News articles have a title, headline, author, date, and article body.
 - ▶ Sales transaction data has timestamps, product ID, customer ID, and payment method.
- ▶ Each metadata field has an associated data type

Data Type	Description	Examples
Numerical	Data representing quantities that can be measured or counted.	Sales revenue, customer age, product quantity, website visits.
Textual	Human language data, composed of words and sentences.	Customer reviews, news title, product descriptions.
Categorical	Data representing groups or categories.	Product category, payment method, satisfaction rating.
Datetime	Specific points in time, typically combining a calendar date with a time of day	Transaction timestamps, login times

Data Collection and Extraction

- ▶ There are four common methods for data collections.

Method	Description	Common Tools and Packages	Characteristics	Examples
Database Querying	Direct access to structured data from databases.	SQL clients, Python	Highly, structured, efficient for large volumes.	Internal customer data, transaction data
Web Scraping / Crawling	Programmatically extracts data from public websites by parsing HTML.	Python (bs4, Scrapy)	Accesses vast external web data, flexible.	Webpages
API Integration	Programmatic access to services via defined Application Programming Interfaces (API).	Python (requests), API-specific libraries	Structured & reliable, efficient, but sometime with usage limits	News articles, social media
Manual	Direct collection from individuals via forms, or downloading from websites.	Custom web forms, spreadsheets.	Targeted info, captures qualitative data	Customer feedback, market survey, government reports



Data Collection and Extraction

- ▶ The raw collected data can often contain too much information, in unstructured or semi-structured forms.
- ▶ Data extraction isolates the useful pieces needed for analysis. For example:
 - ▶ Use keywords to identify relevant news articles and sentences for sentiment analysis
 - ▶ Identify entities mentioned in social media to understand discussion trends
 - ▶ Extract customer information from PDF or images.
- ▶ **Common tools:** Python (pandas, spacy, nltk, Gensim), OCR tools, ETL platforms

Data Collection and Extraction

- ▶ **Coding example:** web data extraction
 - ▶ Refer to the supplemental material.
- ▶ Given a raw HTML webpage of a news article
 - ▶ Includes title, author, date, body text, headings, etc.
- ▶ We use **Beautiful Soup (bs4)** in Python to extract them systematically.

```
<html>
  <head>
    <title>News Article</title>
  </head>
  <body>
    <div class="article">
      <h1>Tech Giant Announces New AI Chip</h1>
      <p class="author">By Jane Doe</p>
      <p class="date">October 27, 2023</p>
      <div class="content">
        <p>Tech Giant announced its new AI chip today...</p>
        <p>The chip promises...</p>
      </div>
    </div>
  </body>
</html>
```


AI-enabled Analytic



Descriptive analytic:

- **Purpose:** Summarize and interpret historical data to understand past events.
- **Question to answer:** What happened?
- Often supported by unsupervised learning



Predictive analytic:

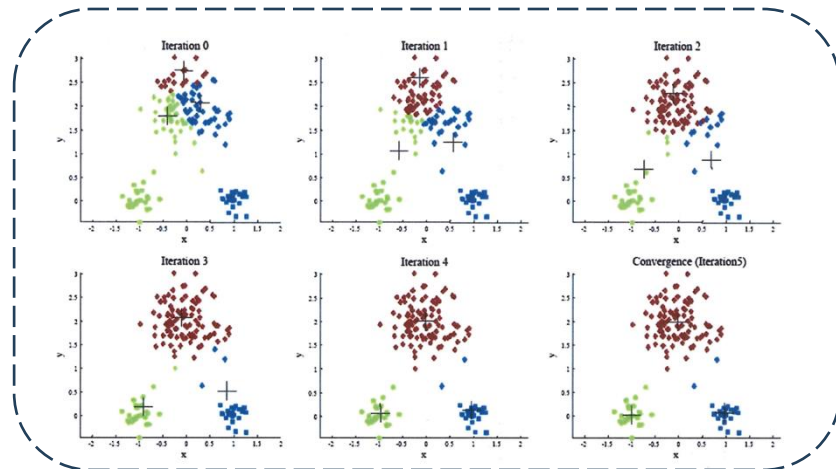
- **Purpose:** Forecast future outcomes or probabilities based on historical data
- **Question to answer:** What will happen?
- Often supported by supervised learning



Prescriptive analytic:

- **Purpose:** Recommend specific actions to take to achieve desired outcomes or optimize performance.
- **Question to answer:** What should I do?
- Often supported by reinforcement learning

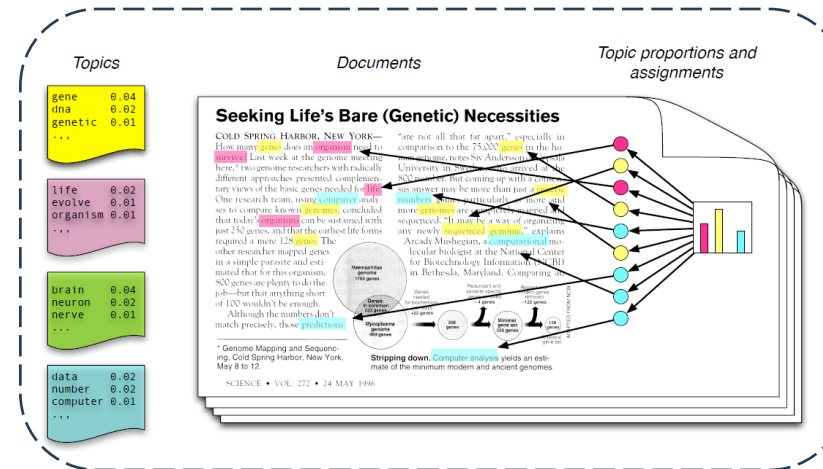
AI-enabled Analytic – Descriptive



Clustering

Examples

- **Customer Segmentation:** Grouping existing customers by purchasing habits or demographics.



Topic Modeling

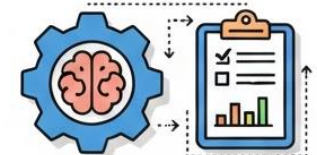
Examples

- **Customer Feedback Analysis:** Identifying themes and specific pain points from online reviews to guide product development.

1. Data



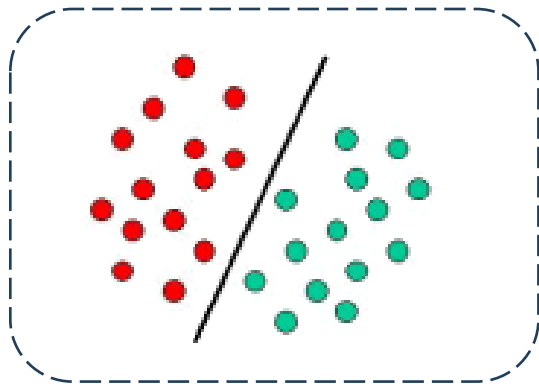
2. Training



3. Insights



AI-enabled Analytic – Predictive

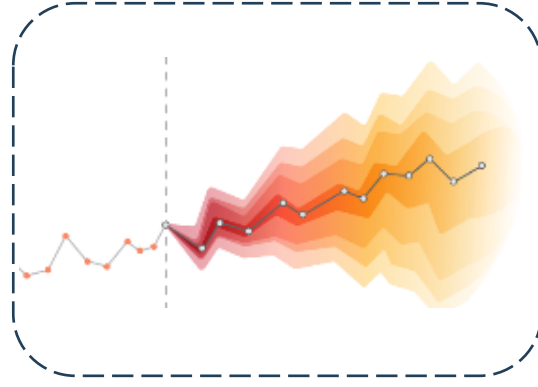


Classification



Examples

- **Customer Churn Prediction:** Identifying which customers are most likely to leave.
- **Fraud Detection:** Flagging transactions that have a high probability of being fraudulent.

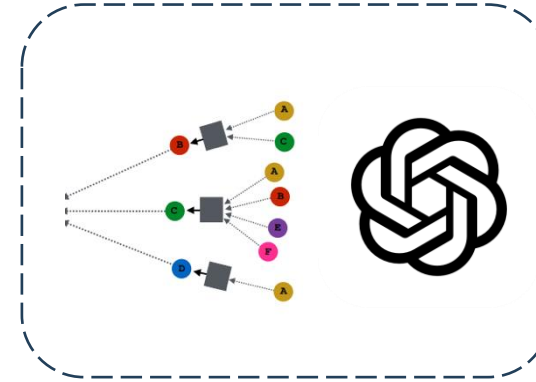


Time series forecasting



Examples

- **Sales Forecasting:** Predicting next quarter's revenue or demand for specific products
- **Energy Capacity Forecasting:** Predicting electricity demand for upcoming hours or days to optimize grid operations.



Emerging Analytics (GNN, LLM)



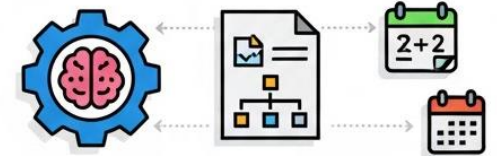
Examples

- **Product Recommendation:** Suggesting items a customer is likely to buy based on past behavior and similar users.
- **Report Summarization:** Generating concise summaries from long documents, reports, or articles.

1. Data



2. Training & Testing



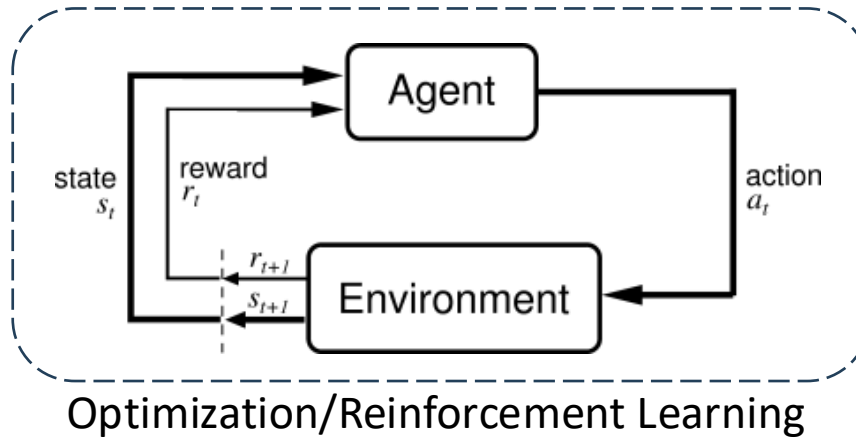
3. Prediction & Forecasting



Altmasfive Outcome

Online M.S. in AI

AI-enabled Analytic – Prescriptive



Examples

- **Production Optimization:** Determine the schedule of production machine to minimize idle time or maximize throughputs

1. Simulation



2. Optimization



3. Recommended Action



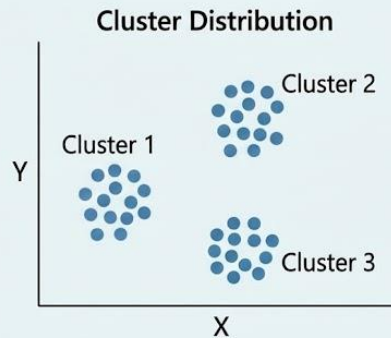


Visualization

- ▶ Visualization and value delivery are the final steps after analytics.
 - ▶ It is the bridge between analytic output and human decision
- ▶ **BI visualization principles:**
 - ▶ **Overview first:** Provide a high-level view before showing details
 - ▶ **Use the right visual form:** Match representation to the data
 - ▶ **Reduce cognitive load:** Use pre-attentive visual cues; avoid clutter
 - ▶ **Preserve context:** Show details while keeping the big picture visible
 - ▶ **Highlight patterns and relationships:** Use visualization + information analysis
 - ▶ **Enable actionable decisions:** Ensure visual insights translate into understanding

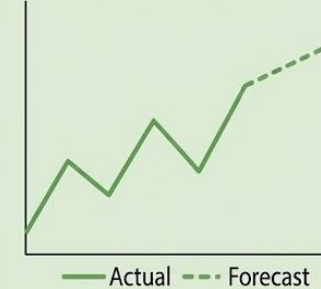
Visualization – Basic Diagrams

Descriptive Analytics → What happened?

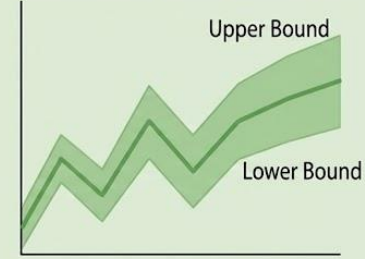


Predictive Analytics → What will happen?

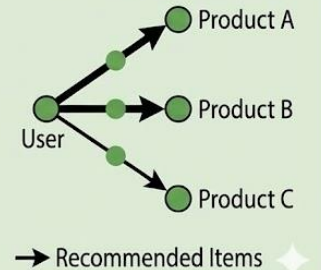
Forecast vs Actual



Confidence Interval Ribbons

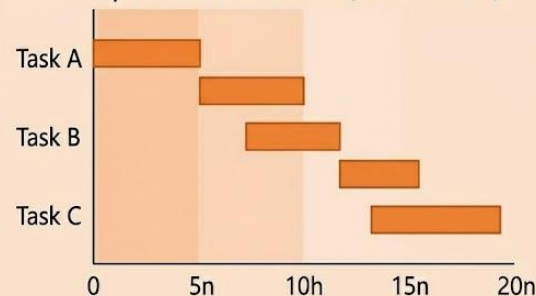


Product Recommendation Engine

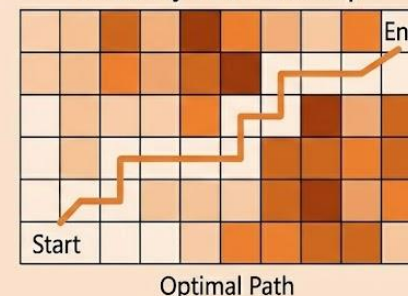


Prescriptive Analytics → What should we do?

Optimization Results (Gantt Chart)

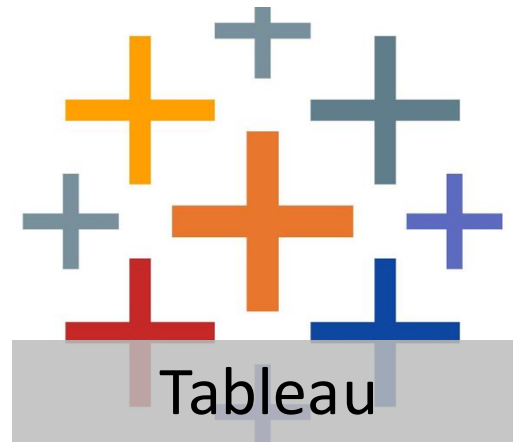


RL Policy Path Heatmap

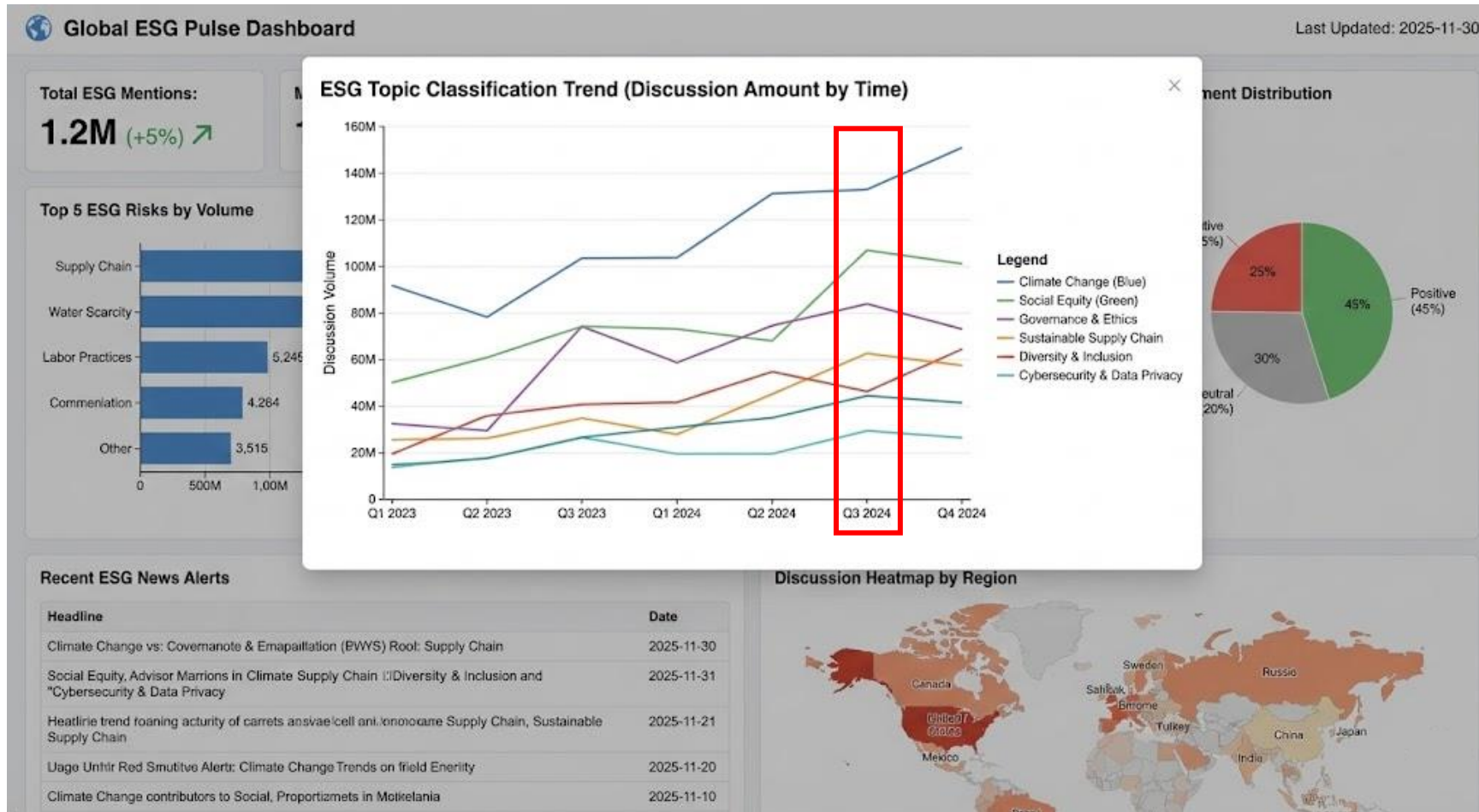


Visualization – User Interface

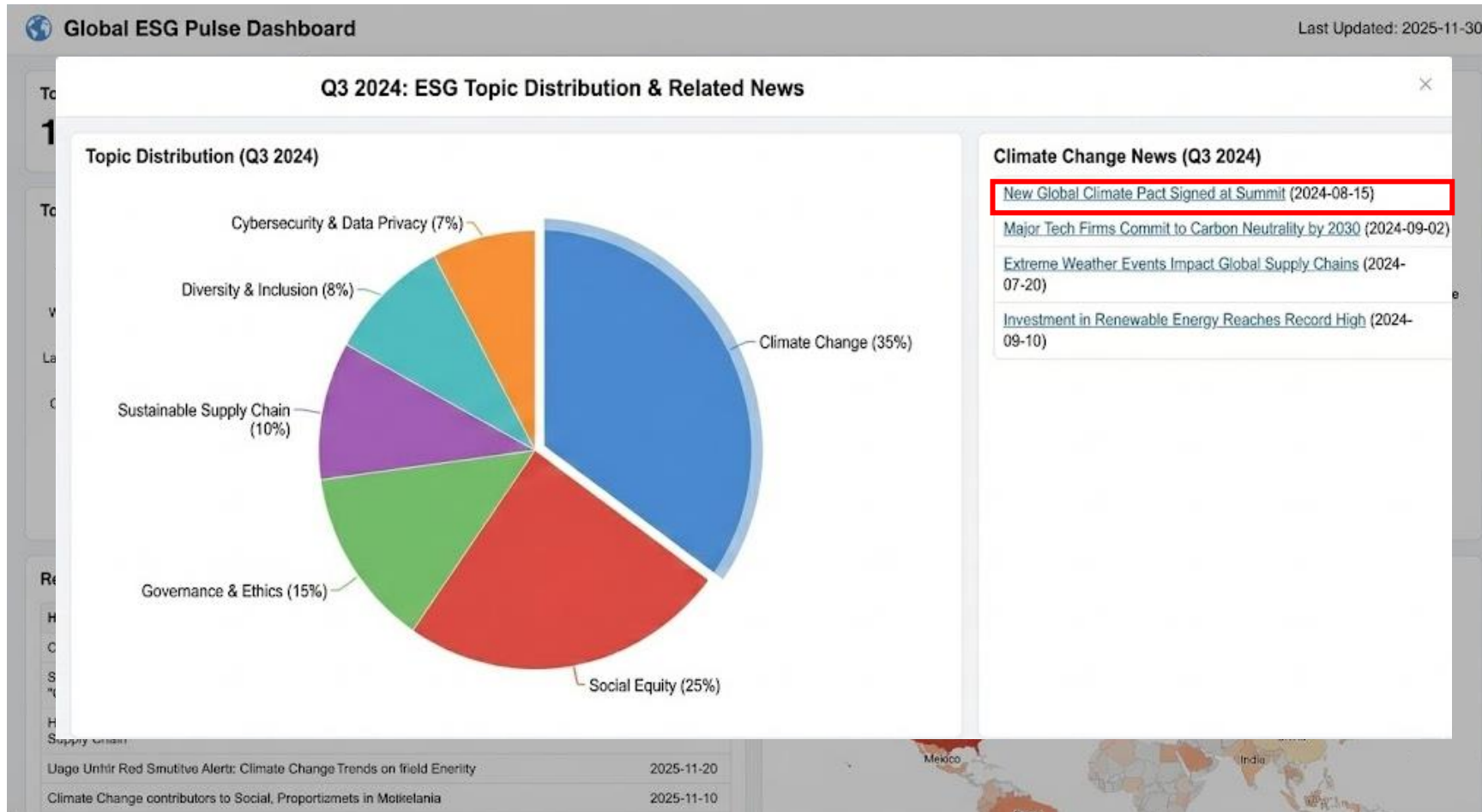
- ▶ Visualization can be presented via a **system and user Interface (UI)**
- ▶ A good visualization allows users to:
 - ▶ **Filter** by product, time, customer type, or risk category
 - ▶ **Drill down** from high-level patterns to specific data records
 - ▶ **Hover or click** to see explanations, raw data, or related analytics



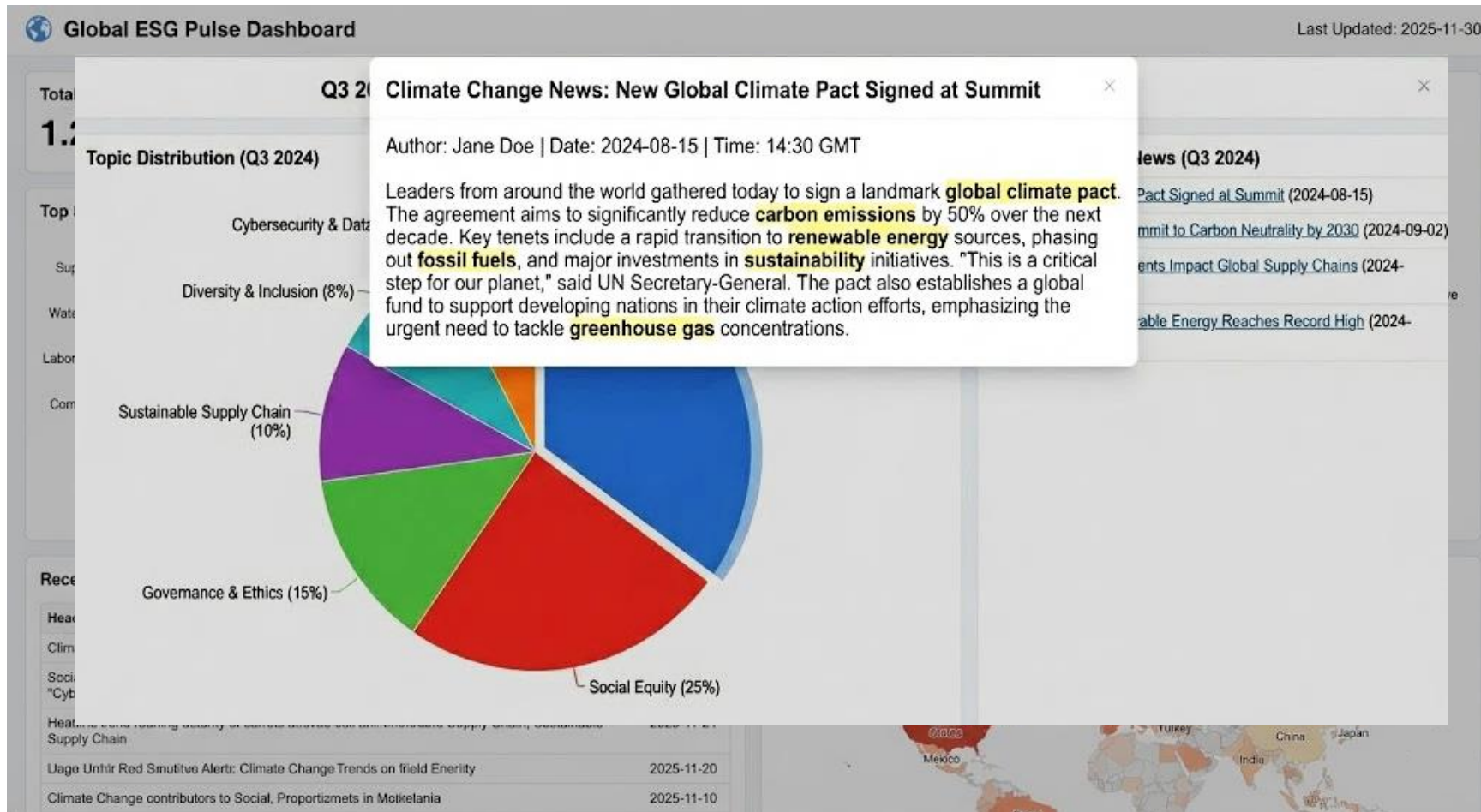
Visualization – User Interface (Example)



Visualization – User Interface (Example)



Visualization – User Interface (Example)





Review Questions

- ▶ In the AI4BI development process, what are the key components of the problem identification step?
- ▶ Give one example each of descriptive, predictive, and prescriptive analytics.
- ▶ Why is visualization critical for supporting decision-making in business environments?

Next Week

