AKTU · 3RD YEAR · SEMESTER 5 · BCDS051

Business Intelligence and Analytics

Unit 5 — Implementation of BI

- Business Activity Monitoring
- Complex Event Processing
- Business Process Management
- Metadata
- Root Cause Analysis



DECODED BY

Chirag Jain

Founder, Ycotes · Gold Medalist in Artificial Intelligence, IIT Madras · taught 10,000+ students

Unit 5: Implementation of BI - Overview

Key operational paradigms and architectural components for implementing Business Intelligence.

- Business Activity Monitoring (BAM) for real-time tracking of operational metrics.
- Complex Event Processing (CEP) for detecting patterns in high-velocity event streams.
- Business Process Management (BPM) for systematic workflow optimization.
- Metadata Management for structured data governance and context tracking.
- Root Cause Analysis (RCA) for diagnostic analytics and operational problem-solving.

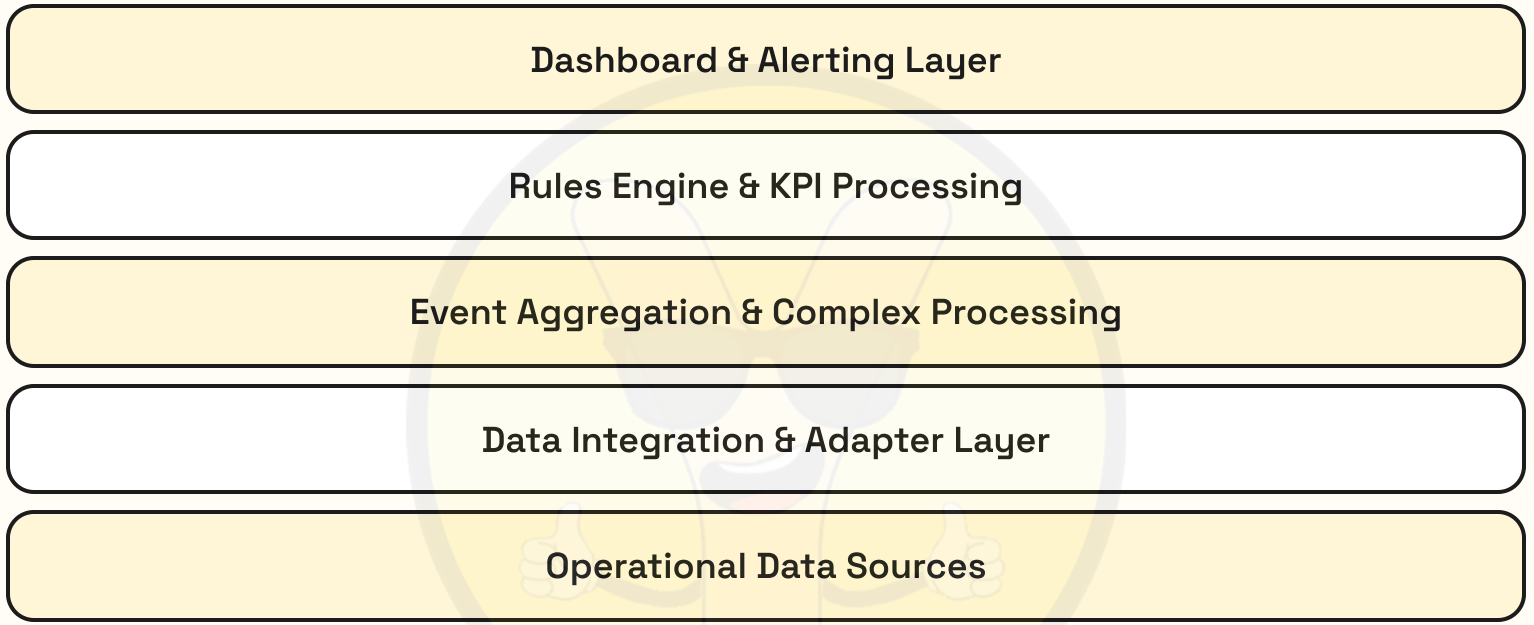
Business Activity Monitoring (BAM)

BUSINESS ACTIVITY MONITORING (BAM)

BAM is a software paradigm that enables real-time tracking, analysis, and management of key operational business activities and metrics across an enterprise.

- Continuously aggregates operational data directly from active business applications.
- Computes real-time Key Performance Indicators (KPIs) to identify process bottlenecks.
- Generates proactive alerts when metric values breach predefined thresholds.
- Reduces time latency between operational event occurrence and decision execution.

Architectural Layers of BAM



Layered architecture of Business Activity Monitoring system.

BAM Key Components and Advantages

Core features enabling operational intelligence in real time.

- Real-time Dashboards: Displays live operational metrics and graphical status indicators.
- Event Correlator: Combines multiple operational events to calculate business metrics.
- Alerting Engine: Triggers notifications via email, SMS, or API calls upon KPI threshold breach.
- Latency Reduction: Minimizes data latency, analysis latency, and decision latency.
- Automated Escalation: Escalates unresolved operational bottlenecks to management.

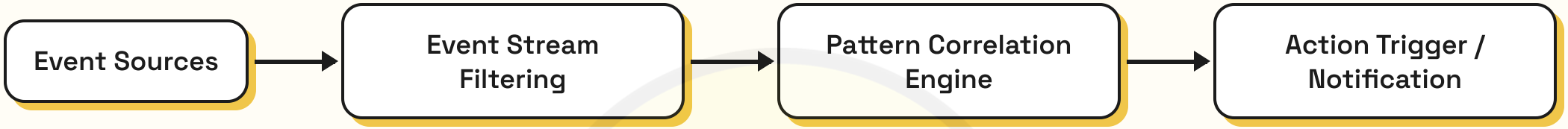
Complex Event Processing (CEP)

COMPLEX EVENT PROCESSING (CEP)

CEP is an event-driven technology that processes multiple continuous streams of low-level events to detect complex patterns, relationships, and actionable insights in real time.

- Treats data as continuous, high-velocity streams rather than static tables.
- Evaluates temporal, spatial, and causal relationships among independent events.
- Employs Event Processing Languages (EPL) to define detection rules.
- Enables immediate automated responses to critical event patterns.

CEP Event Processing Pipeline



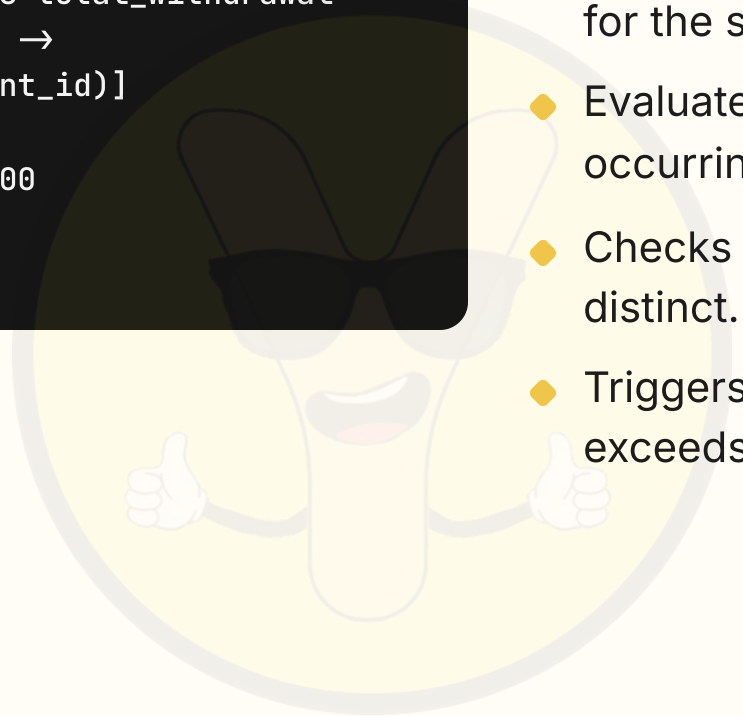
Flow of events through a Complex Event Processing system.

Comparison: CEP vs Traditional DBMS

Feature	Traditional Database (DBMS)	Complex Event Processing (CEP)
Data Model	Static persisted data in tables	Dynamic continuous event streams
Query Model	On-demand queries over static data	Continuous standing queries over streaming data
Latency	High latency (store then query)	Ultra-low latency (sub-millisecond processing)
Data Storage	Stores data first, processes later	Processes data first, selectively stores results
Focus	Historical analysis and reporting	Real-time pattern detection and immediate action
Execution	Explicit user execution	Event-driven automatic execution

Event Pattern Matching Query (EPL)

```
SELECT e1.account_id, SUM(e1.amount) AS total_withdrawal
FROM Pattern [every e1=WithdrawalEvent →
e2=WithdrawalEvent(account_id=e1.account_id)]
WHERE e1.location ≠ e2.location
  AND e2.timestamp - e1.timestamp ⩽ 300
GROUP BY e1.account_id
HAVING SUM(e1.amount) > 10000;
```



- Monitors two consecutive ATM withdrawals for the same account.
- Evaluates temporal condition: events occurring within $t \leq 300$ seconds.
- Checks spatial anomaly: physical locations are distinct.
- Triggers alert if cumulative monetary value exceeds defined limit.

CEP Fraud Detection Calculation

An ATM transaction stream receives two events for Account A: Event 1 at $t = 0$ min for $\$6,000$ in New York, and Event 2 at $t = 3$ min for $\$5,000$ in London. Calculate time window Δt, evaluate velocity anomaly v, and determine if a fraud pattern alert triggers given threshold velocity $v_{\max} = 1000\ \text{km/h}$ and distance $d = 5500\ \text{km}$.

1. Step 1: Calculate time interval $\Delta t = 3\ \text{minutes} = 0.05\ \text{hours}$.
2. Step 2: Compute implied travel velocity: $v = \frac{d}{\Delta t} = \frac{5500\ \text{km}}{0.05\ \text{h}} = 110,000\ \text{km/h}$.
3. Step 3: Compare velocity with threshold: $v = 110,000\ \text{km/h} > v_{\max} = 1000\ \text{km/h}$.
4. Step 4: Evaluate aggregate monetary threshold: $\text{Total} = \$6,000 + \$5,000 = \$11,000 > \$10,000$.

Both physical impossibility (velocity breach) and financial threshold breaches occur within $\Delta t \leq 5$ min. High-priority Fraud Alert triggered immediately.

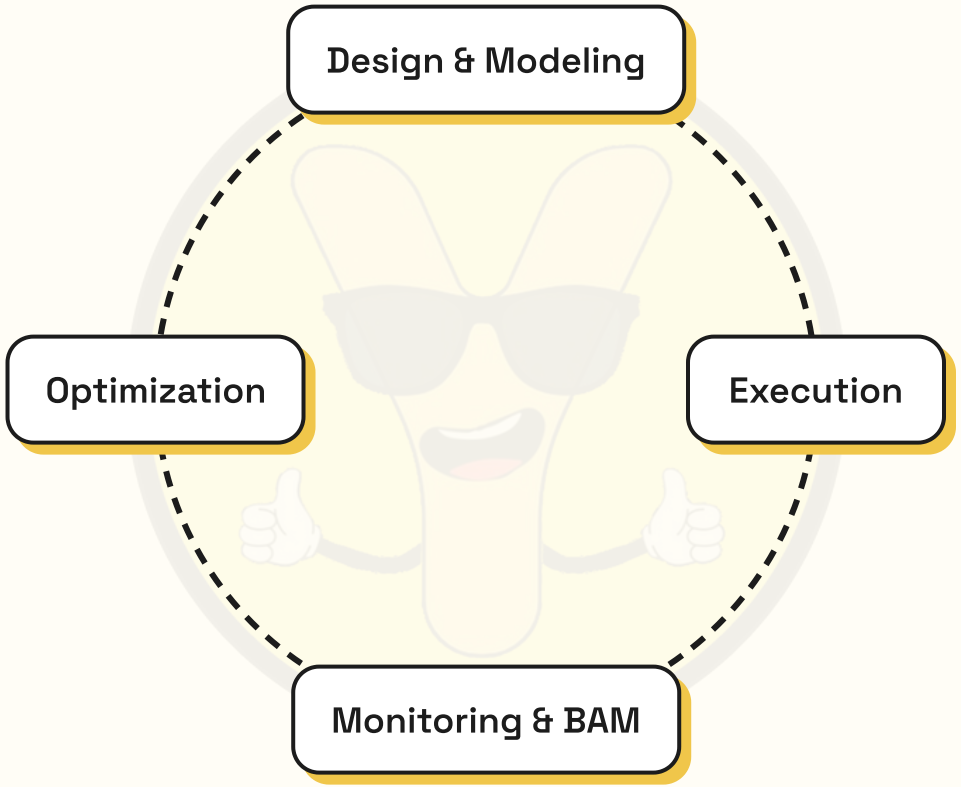
Business Process Management (BPM)

BUSINESS PROCESS MANAGEMENT (BPM)

BPM is a structured discipline that uses methods, techniques, and software tools to discover, model, analyze, execute, measure, and optimize end-to-end business processes.

- Aligns operational workflows with explicit organizational strategic goals.
- Integrates human tasks, automated operations, and legacy software systems.
- Employs standardized notations such as Business Process Model and Notation (BPMN).
- Provides quantitative feedback to drive continuous process improvement.

Business Process Management Lifecycle



Continuous iterative lifecycle of Business Process Management.

Integration of BPM and Business Intelligence

How BPM and BI converge to deliver operational and strategic value.

- Process-Aware BI: Embedded analytical reports directly inside operational workflow interfaces.
- Closed-Loop BI: Analytical insight triggers automatic execution of optimized workflow steps.
- Performance Tracking: Uses BI dashboards to track BPM Key Performance Indicators.
- Bottleneck Identification: Uses process mining to reveal delays and operational friction.
- Data-Driven Decision Making: Replaces intuition-based process changes with empirical data.

Comparison: BAM vs CEP vs BPM

Dimension	BAM	CEP	BPM
Primary Focus	Monitoring operational KPIs	Pattern matching on streams	Process workflow management
Data Type	Aggregated metric events	High-velocity atomic events	Structured process instances
Abstraction Level	Business metric level	Technical / Event level	Workflow / Operational level
Action Mechanism	Threshold alerts & dashboards	Complex event generation	Workflow task execution
Latency Target	Seconds to minutes	Milliseconds to seconds	Minutes to days
Key Tool / Method	KPI Dashboards & Alerting	EPL engines (Esper, Flink)	BPMN 2.0 workflow engines

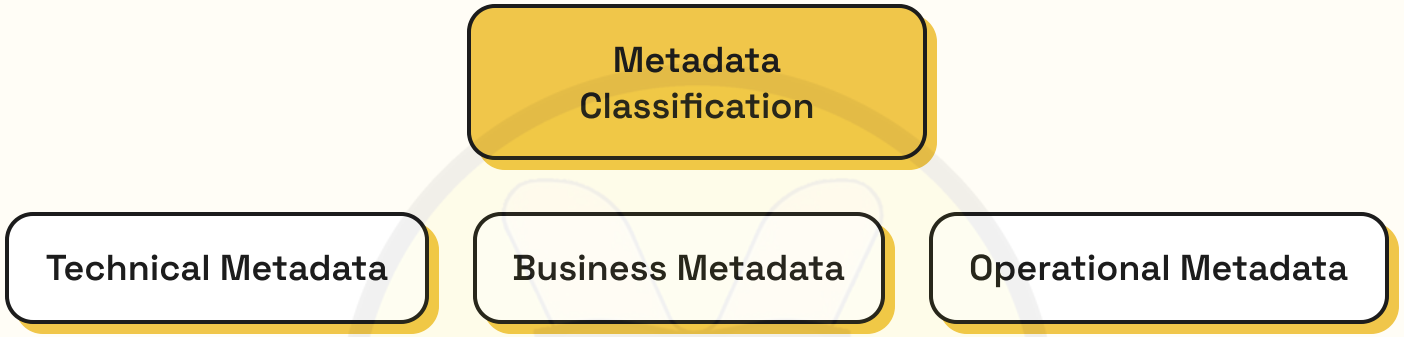
Metadata in BI Systems

METADATA

Metadata is structured informational data about data that describes the structure, context, origin, usage, transformational history, and governance rules of enterprise data assets.

- Acts as the central blueprint and dictionary for data warehousing and BI architectures.
- Enables data lineage tracking from source extract to analytical reporting.
- Enhances data quality management, trust, and enterprise compliance.
- Facilitates efficient semantic mapping between business terminology and technical schema.

Taxonomy of BI Metadata



Hierarchical breakdown of metadata types in a Business Intelligence system.

Metadata Repository & Governance

Architecting a centralized metadata management framework.

- Centralized Metadata Repository: Single source of truth containing schemas and mappings.
- Data Lineage: Tracks data origin $\text{Source} \rightarrow \text{ETL} \rightarrow \text{DW} \rightarrow \text{BI Dashboard}$.
- Impact Analysis: Evaluates downstream consequences before modifying source database schema.
- Business Glossary: Standardizes definitions of metrics (e.g., precise definition of Net Revenue).
- Data Cataloging: Automates discovery and indexing of multi-source enterprise data.

Root Cause Analysis (RCA)

ROOT CAUSE ANALYSIS (RCA)

RCA is a structured problem-solving methodology used in diagnostic analytics to identify the fundamental, underlying causes of operational failures or performance anomalies.

- Prevents symptom-only treatment by isolating deep underlying systemic flaws.
- Shifts analytics focus from descriptive (What happened?) to diagnostic (Why did it happen?).
- Combines qualitative frameworks with quantitative data mining and drill-down techniques.
- Reduces recurring operational failures and optimizes long-term process quality.

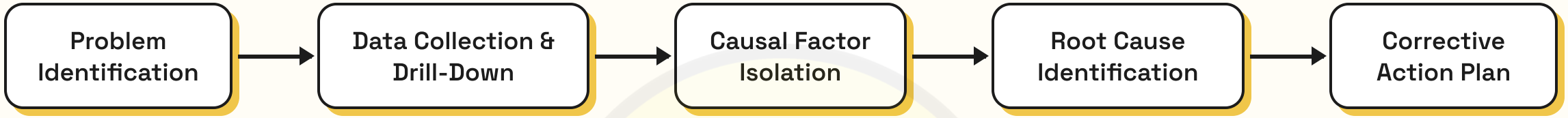
Worked Example: 5 Whys Methodology

A BI implementation experienced a complete dashboard sync failure during morning business hours. Conduct a 5 Whys Analysis to isolate the root cause.

1. Why 1: Why did the dashboard fail to update? Answer: The ETL pipeline failed to complete by 06:00 AM.
2. Why 2: Why did the ETL pipeline fail? Answer: The ETL job timed out waiting for the source database response.
3. Why 3: Why was the source database unresponsive? Answer: Database indexes were corrupted during nightly maintenance batch.
4. Why 4: Why were indexes corrupted? Answer: The automated maintenance script crashed due to running out of disk space.
5. Why 5: Why did disk space run out? Answer: Old log files were never auto-purged due to a misconfigured log rotation policy.

Root Cause identified as a misconfigured log rotation policy on the source database server. Action: Fix log rotation script and add automated disk alert.

Root Cause Analysis Diagnostic Process



Sequential steps in performing Diagnostic Analytics and Root Cause Analysis.

Exam Corner: Likely AKTU Questions

- **Q1** Explain Business Activity Monitoring (BAM). Discuss its architectural layers, core benefits, and role in real-time BI.
- **Q2** Define Complex Event Processing (CEP). How does CEP differ from traditional database processing?
- **Q3** Differentiate between BAM, CEP, and BPM with suitable operational criteria and technical examples.
- **Q4** Describe the Business Process Management (BPM) lifecycle. Explain how BI integrates with BPM for process optimization.
- **Q5** What is Metadata in BI? Classify Technical, Business, and Operational metadata with clear examples.
- **Q6** Explain the significance of Data Lineage and Impact Analysis in Metadata Management repositories.
- **Q7** What is Root Cause Analysis (RCA)? Illustrate the 5 Whys method and Fishbone diagram with a practical IT failure scenario.

Unit 5 done. You are exam-ready.

Revise the Exam corner once more, then open Unit 1 on theycotes.com. Every AKTU subject, unit by unit, free. Notes by Chirag Jain, Ycotes.

