AWS First Cloud AI Journey – Project Plan
[Project team] – [University] – [Project Name]
[Date]

Table of Contents

1 BACKGROUND AND MOTIVATION

1.1 Executive Summary

Organization faces HR evaluation inefficiencies due to manual data handling, lack of transparency in evaluation processes and metrics tracking.

InsightHR delivers HR automation through flexible evaluation management, automated scoring, and AI insights. AWS provides serverless scalability, cost efficiency, security for sensitive data, AI Chatbot via Bedrock, and rapid deployment.

Custom KPI, automated performance scoring, multi-level dashboards, AI assistant for natural-language queries, automated notifications, role-based access (Admin/Manager/Employee), multi-tenant support.

End-to-end delivery including Well-Architected design, serverless backend (Lambda, DynamoDB, API Gateway), frontend (S3 + CloudFront), authentication/security (Cognito, IAM), KPI/formula builder, AI chatbot (Bedrock + Lambda query data from info tables), notifications (SNS, SES), CI/CD, monitoring, and knowledge transfer.

1.2 Project Success Criteria

Success is defined by demonstrating a functional MVP that proves the platform's capability to automate HR evaluations and deliver measurable business value.

1. **Functional Criteria:**
   - Authentication with role-based access (Admin/HR, Manager, Employee)
   - HR creates custom KPIs without technical support
   - CSV upload triggers automated Lambda scoring
   - Dashboard displays individual/team performance with charts
   - AI chatbot answers natural language queries from DynamoDB data
   - SES sends automated email notifications

2. **Technical Criteria:**

- 99.9%+ uptime
- <300ms API latency (95th percentile)
- 95%+ scoring accuracy vs manual calculations
- 90%+ AI response relevance
- Zero critical security vulnerabilities

3. **Performance & Cost:**
   - ~$33.14/month AWS cost
   - End-to-end workflow (upload → score → visualize) completes in <5 minutes

4. **Business Impact:**
   - Demonstrates 60%+ HR time reduction potential
   - Non-technical users operate KPI builder and chatbot independently

5. **Delivery:**
   - Week 8: MVP (authentication, KPI/formula management, scoring, basic dashboard)
   - Week 12: Full features (chatbot, notifications, advanced dashboard)

1.3 Assumptions

1. **Assumptions:**
   - The current AWS cost estimate of approximately $33.14/month is accurate for the projected initial load and usage.
   - The required data format and mapping logic for employee performance data can be clearly defined and provided by the HR team for the automated scoring engine.
   - The Large Language Model provided by Amazon Bedrock to support HR.
   - The automated scoring system is trained locally. The technical evaluation files of each team are assessed according to the companies' criteria and must follow the format provided by the customer.

2. Constraints:
- The project delivery must adhere to the 12-week timeline utilizing the Agile Scrum framework.
- The solution must be built entirely on serverless AWS services to meet the objectives of scalability, cost efficiency, and reduced operational overhead.
- The final production AWS cost must remain around the ~$33.14/month target.

3. Risks:
- Data Security/Compliance: Failure to fully understand or implement all of the customer's specific regulatory control validation requirements could impact the project's ability to meet security objectives.
- Feature Creep: Requests for features identified as "Out of Scope" could derail the 12-week MVP delivery timeline.

2 SOLUTION ARCHITECTURE / ARCHITECTURAL DIAGRAM

2.1 Technical Architecture Diagram

The InsightHR platform is built on a serverless architecture using AWS services, providing scalability, cost-effectiveness, and high availability, the architecture includes:

1. **Frontend & Content Delivery:**
   - Amazon S3: Hosts the static website and stores user-uploaded files (CSV, AI models). S3 Vector to store vectors (embeddings for text-data), S3 Standard for storing raw documents..
   - Amazon CloudFront: Distributes static and dynamic content globally with low latency.

2. **Backend & Compute:**
   - AWS Lambda: Executes all business logic, including authentication, custom scoring, and chatbot functions.

- Amazon API Gateway: Manages APIs as the communication gateway between frontend and backend.

3. **Data Storage:**
   - Amazon DynamoDB: Stores structured data such as user/employee information, company KPIs, scoring formulas, and performance evaluation results.

4. **AI & Machine Learning:**
   - Amazon Bedrock: Provides Large Language Models (LLMs) for the HR assistant chatbot.
   - The ML Model system is trained locally for scoring function.

5. **Security & Identity:**
   - Amazon Cognito: Manages user authentication, registration, and identity workflows.
   - AWS IAM: Manages access control and permissions for AWS services.
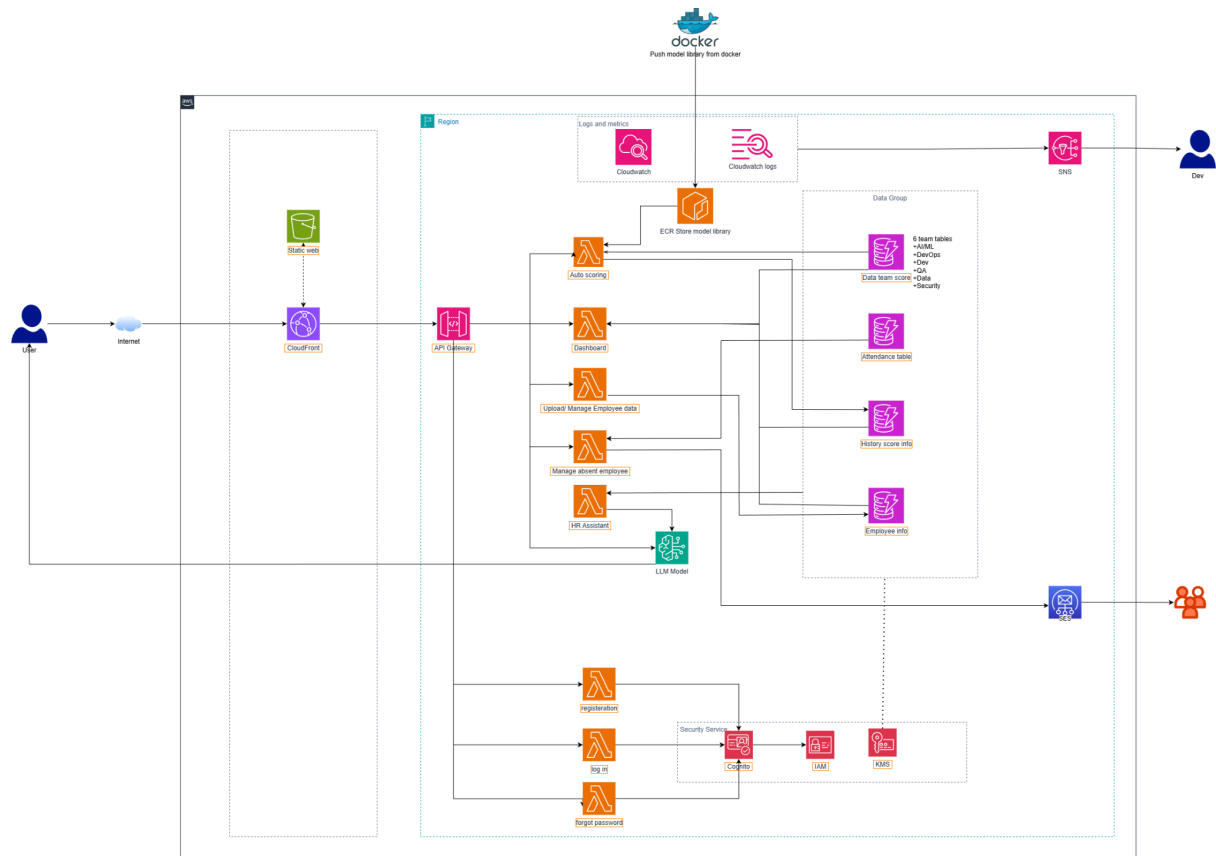   - AWS KMS: Encrypts sensitive data in DynamoDB and S3.

6. **Monitoring & Notifications:**
   - Amazon CloudWatch & CloudWatch Logs: Monitors Lambda functions, API Gateway, and database access.
   - Amazon SNS: Sends notifications (e.g., reminders, result notifications) to employees.

7. **Architecture Benefits:**
   - Serverless: No server management and automatic scaling.
   - Cost-Effective: Mostly pay-as-you-go services.
   - High Availability: Built-in redundancy across AWS regions.
   - Scalable: Can handle growth from small teams to large enterprises.
   - Flexible: Easy to modify and extend functionality.

8. **Proposed Architecture Diagram:**

9. Tools Proposed for This Project:

- Amazon CloudFront: For global content delivery and caching of static and dynamic web content.
- Amazon S3: To host static web assets and store documents, vector embeddings, and other files processed by the system.
- Amazon API Gateway: To provide a secure RESTful interface, acting as the communication layer between frontend clients and backend services.
- AWS Lambda: To run backend business logic including user dashboard, auto scoring, HR assistant, and data management workflows.
- Amazon DynamoDB: To store application data such as user information, HR records, scoring results, and vector metadata with low-latency performance.
- Amazon Cognito: To manage user authentication, authorization, registration, MFA, and secure access to APIs and frontend applications.
- AWS Identity and Access Management (IAM): To define fine-grained access policies and control permissions between services and users.

- AWS Key Management Service (KMS): To manage encryption keys used for securing sensitive data stored in S3, DynamoDB, and logs.
- Amazon ECR: To store containerized model assets and application dependencies in a secure and version-controlled repository.
- Amazon Bedrock / Large Language Model (LLM): To provide AI capabilities for chat, data extraction, summarisation, and intelligent HR workflows.
- Amazon Simple Email Service (SES): To send automated email notifications such as onboarding alerts and HR communications.
- Amazon Simple Notification Service (SNS): To publish notifications and trigger downstream processes; integrates with email, SMS, and microservices.
- Amazon CloudWatch & CloudWatch Logs: For monitoring performance, logging, tracing, and operational troubleshooting across Lambda, API Gateway, and AI components.

2.2 Technical Plan

The partner will develop automated deployment scripts using AWS CloudFormation and Infrastructure as Code (IaC) practices.

This will allow for quick and repeatable deployments into AWS accounts. Some additional configurations such as WAF rules on CloudFront for enhanced security may require approval and will follow standard DevOps change management processes.

Application Feature Implementation:

1. Authentication & Security Module
- User Management: Cognito manages user lifecycle Registration, login, password reset workflows

- Access Control: IAM and RBAC enforce role-based permissions Admin/HR, Manager, and Employee access levels
- API Security: API Gateway implements JWT-protected endpoints Token validation before Lambda processing

2. Administration Module (HR Panel)
- KPI Management: HR creates, edits, and deletes custom metrics Examples: Tasks Completed, Code Quality, Customer Satisfaction Definitions stored in DynamoDB
- Auto scoring by employee's technical score for each team with ML model.

3. Core User Functions
- Data Upload & Mapping: Upload performance data files (CSV) to DynamoDB
- Scoring Engine: Lambda triggered on upload Retrieves active formula from DynamoDB Calculates employee scores Stores results in DynamoDB Flow: Upload → Validate → Map → Calculate → Store
- Dashboard: Visualize individual and department performance Line graphs, bar charts, trend analysis
- AI Chatbot: Bedrock (LLM) integration Natural language queries (e.g., "Summarize Team A Q4 performance") Queries and summarizes DynamoDB data
- Notifications: SES sends automated alerts Performance milestones, review reminders, custom triggers

2.3 Project Plan
The partner will adopt the Agile Scrum framework over 12 one-week sprints totaling a 12-week delivery timeline.

1. Team Responsibilities
- Product Owner: Prioritizes backlog (KPIs, formulas, analytics) Final authority on feature acceptance
- Development Team: Implements Cognito authentication Builds admin portal and formula builder Develops scoring

engine and dashboard Integrates Bedrock chatbot and SNS with SES notifications via Email.
- QA Personnel: Conducts functional, performance, and security testing Facilitates UAT Ensures compliance and quality standards

2. **Communication Cadences**
   - Daily Standups (30 min - 1 hr): Progress review and blocker identification
   - Retrospectives (Weekly, 1 hr): Process improvement and delivery optimization
   - Executive Updates (Weekly): Written reports on progress, risks, KPIs, roadmap Leadership decisions required

3. **Knowledge Transfer**
   - Sessions conducted by the development team covering AWS serverless fundamentals KPI and formula configuration Data workflows and column-mapping Dashboard navigation and analytics System monitoring (CloudWatch, Cognito, DynamoDB)

2.4 Security Considerations

The partner will implement AWS security best practices based on the Well-Architected Framework, prioritizing protection of sensitive HR data while ensuring high operational availability. Security implementation covers five key categories:

1. **Access Control**
   - Cognito manages user identities Enforces strong password policies and MFA support
   - IAM implements RBAC Admin/HR access Admin Panel and KPI/Formula configurations Employees view only their own performance data
   - API Gateway validates JWT tokens Cognito-issued tokens verified before Lambda processing

2. **Infrastructure Security**

- Serverless architecture reduces attack surface No OS or server patching required
- Lambda functions communicate via private AWS networks Only necessary endpoints exposed through API Gateway

3. Data Protection
- KMS encrypts data at rest DynamoDB and S3 encrypted Data unusable without decryption keys
- TLS/SSL (HTTPS) encrypts data in transit All frontend-backend communication secured

4. Detection & Monitoring
- CloudWatch Logs captures execution details Lambda and API Gateway activity logged Real-time monitoring and anomaly detection enabled
- AWS Config tracks configuration changes Ensures resource compliance with security objectives

5. Incident Management
- CloudWatch Alarms trigger automated alerts via SES Failed login threshold breaches Lambda resource anomalies
- Security Hub provides consolidated security view Unified compliance findings across AWS environment Simplifies incident identification and response

AWS CloudTrail and AWS Config will be configured for continuous monitoring of activities and compliance status of resources. The customer will share their regulatory control validation requirements as inputs for the partner to ensure all security objectives are met.

3 ACTIVITIES AND DELIVERABLES

3.1 Activities And Deliverables

NOTE: Some Project Phases overlap each other.

Project Phase	Timeline	Activities	Deliverables/Milestones	Total man-day

Phase	Timeline	Activities	Deliverables	Hours
Phase 1: Foundation & Scoring Model	Week 1-8	• Personal infrastructure architecture research • Data generation for local model training • Scoring model build	• Finalized personal architecture diagram • Ready dataset for Local Model training • Scoring Model MVP (Minimum Viable Product)	80
Phase 2: Project Setup & Dashboard	Week 9-10	• Project Setup with basic functions: **IAM Role**, **CRUD function**, **Static web** • Web UI Demo • Implement Dashboard • Fix Model	• Basic **IAM Roles** configured • Operational **CRUD functions** • Static website deployed (S3/CloudFront) • **Web UI Demo** completed • **Dashboard** displaying data implemented	40
Phase 3: AI Agent & Absence Mgmt	Week 11	• Building **Bedrock Agent** • Implement Absent managing	• **Bedrock Agent** built • Operational **Absence tracking workflow** implemented	15

Phase 4: Integration, Testing & Handover	Week 12	• Implement **Chatbot** into App • **Testing** and set up **Monitoring**	• **Chatbot** integrated into the application • Functional, Performance, and Security Testing completed • **Monitoring** (CloudWatch) configured and operational • Project Completion Report & Post-implementation support plan delivered	15

3.2 Out Of Scope

1. AI Enhancements
- AI-Powered Insights:
    - When sufficient data is available, develop AI models capable of:
    - Chatbot accesses database directly and retrieves prompt window -> Price a lot of tokens and high lock, future can be optimized by other ways
    - Identifying performance patterns across teams and departments
    - Predicting HR risks (e.g., turnover likelihood, burnout indicators)
    - Recommending personalized development plans
    - Detecting anomalies in performance data
    - Suggesting optimal team compositions

- Machine Learning Features:
    - Predictive analytics for workforce planning
    - Sentiment analysis from employee feedback
    - Automated skill gap analysis
    - Performance trend forecasting

2. Public API Development
- API Ecosystem:

- ○ Build a comprehensive API set allowing other internal business systems to automatically push performance data into InsightHR.

- Integration Targets:
  - ○ Project management tools (Jira, Asana, Monday.com)
  - ○ CRM systems (Salesforce, HubSpot)
  - ○ Time tracking software (Toggl, Harvest)
  - ○ Communication platforms (Slack, Microsoft Teams)
  - ○ Code repositories (GitHub, GitLab, Bitbucket)

- Benefits:
  - ○ Transform InsightHR into a central HR data processing hub
  - ○ Create a synchronized and comprehensive management ecosystem
  - ○ Eliminate manual data entry
  - ○ Real-time performance tracking

3. **Advanced Features**
   - Mobile Applications:
     - ○ iOS and Android native apps
     - ○ Push notifications
     - ○ Offline capabilities
     - ○ Mobile-optimized dashboards
     - ○ DynamoDB back up

   - Advanced Analytics:
     - ○ Predictive modeling
     - ○ Benchmarking across industries
     - ○ Custom report builder
     - ○ Data export and API for third-party tools

   - Collaboration Features:
     - ○ Peer review systems
     - ○ 360-degree feedback
     - ○ Goal setting and tracking

○ Performance improvement plans

    ● Compliance & Governance:
          ○ Audit trails
          ○ Compliance reporting
          ○ Data retention policies
          ○ Advanced access controls

3.3 Path To Production

This document outlines the current production architecture and operational status for the InsightHR platform deployment. The platform is fully live in the ap-southeast-1 (Singapore) region.

1. **Platform Architecture and Access**
    ● Public URL:
          ○ https://insight-hr.io.vn
    ● AWS Region:
          ○ ap-southeast-1 (Singapore)
    ● Frontend:
          ○ React application hosted on S3 (insighthr-web-app-sg) with CloudFront HTTPS distribution.
    ● Backend:
          ○ 8 Lambda function groups accessed via API Gateway REST API.
    ● Database:
          ○ DynamoDB tables configured with On-Demand capacity.
                + 6 tables for each team
                + Employee information table
                + History score table
                + Absent table
                + Account managing tables
    ● Authentication:
          ○ Cognito User Pool.
    ● AI/Chatbot:

- Amazon Bedrock (Claude 3 Haiku) integration for conversation history and knowledge base to enhance models.

2. Live Production Features
- The following core features have been successfully deployed and are operational:

- Authentication:
  - Full support for email/password login, password reset workflows.
- User Management:
  - Complete CRUD functionality, including bulk import and role-based access.
- Employee Management:
  - Full support for 300+ employees and bulk operations.
- Performance Score Management:
  - Management of 900+ quarterly scores and calendar-based viewing.
- Attendance Management:
  - Processing of 9,300+ records, including check-in/check-out kiosk functionality and auto-absence marking.
- Performance Dashboard:
  - Live charts, trend analysis, live clock, and CSV export capabilities.
- AI Chatbot:
  - Bedrock integration with conversation history enabled.

3. Deployment and Verification Process
- The standard, repeatable deployment workflow ensures rapid and verifiable updates to the production site:

- Build:
  - npm run build creates the optimized production asset bundle.
- Test:

- ○ npm run preview validates the built bundle locally prior to deployment.
- Deploy:
  - ○ aws s3 sync dist/ s3://insighthr-web-app-sg --region ap-southeast-1 pushes assets to the S3 bucket.
- Invalidate:
  - ○ aws cloudfront create-invalidation --distribution-id E3MHW5VALWTOCI --paths "/*" clears the CloudFront CDN cache.
- Verify:
  - ○ Full feature testing is performed on the live public URL.

4. **Remaining Production Enhancements**
- The platform is in the final phases of enhancement before full stabilization, with key items planned or in progress:

- Page Integration (In Progress)
- Consolidate all administrative page navigation.
- Verify all features are accessible from the main menu.
- Test role-based routing across all pages.
- Fix any integration bugs.

- Polish and Final Deployment (Planned)
- Implement comprehensive error handling and input validation.
- Refine responsive design for full mobile compatibility.
- Conduct dedicated Security testing (penetration testing, vulnerability scanning).
- Execute Load testing for scalability validation.
- Develop user documentation and training materials.
- Perform final production hardening procedures.
- Monitoring and Scalability Strategy

- Active Monitoring:
- CloudWatch Logs are enabled for all Lambda functions and API Gateway endpoints, along with CloudWatch Metrics for performance tracking.
- Planned Alarms:

- - CloudWatch Alarms and SNS notifications are planned for critical error rates and latency.
  - Scalability:
    - Achieved via Serverless Architecture (DynamoDB On-Demand, Lambda, CloudFront CDN).
  - Disaster Recovery:
    - DynamoDB Point-in-Time Recovery and S3 Versioning are planned to be enabled for critical data/assets. Lambda code is stored in version control for rapid redeployment.

4 EXPECTED AWS COST BREAKDOWN BY SERVICES

[AWS Monthly Calculator](#)

AWS Service	Monthly Estimated Cost (USD)
Amazon Bedrock	$21.61
AWS Lambda	$3.75
Amazon Simple Email Service (SES)	$2.25
Amazon DynamoDB	$1.52
Amazon Simple Storage Service	$0.46
Amazon CloudWatch	$0.80
Amazon API Gateway	$0.06
Amazon CloudFront	$0.00
Amazon Cognito	$0.00
Amazon EventBridge	$0.00
Amazon IAM	$1.60
Amazon KMS	$1.03
TOTAL MONTHLY COST	$33.14
TOTAL YEARLY COST	$397.79

5 TEAM

Name	Task	Role	Email / Contact Info
Bùi Tấn Phát	Dashboard, Manage Employee, Support, Content check	Leader	btfat3103@gmail.com
Nguyễn Ngọc Long	CRUD, Config Network / API Gateway, Test function, Slide	Member	nguyenngoclong216@gmail.com
Đặng Nguyễn Minh Duy	Database, CloudWatch / CloudLogs, Paper, Slide	Member	dangnguyenminhduy11b08@gmail.com
Đỗ Đăng Khoa	Log In/ Registration / Forget Password, UI / UX - Static Web, Paper	Member	khoado7577@gmail.com
Nguyễn Huỳnh Thiên Quang	Auto Scoring, AI Assistant, Slide	Member	quangkootenhatvutru@gmail.com

6 RESOURCES & COST ESTIMATES

Resource	Responsibility	Rate (USD) / Hour
Full-Stack Developers [2]	React frontend, Python Lambda backend, API integration	$66
Cloud Engineers [3]	AWS infrastructure setup, deployment automation, monitoring	$66
Other (Please specify)	Estimated platform consumption (Lambda, DynamoDB, Bedrock). Paper and present material	$0.01

NOTE: Project Phase durations overlap each other.

Project Phase	Duration	Man-Days	Other (Please specify)	Estimated Cost
Phase 1: Foundation & Scoring Model	8 Weeks	80	-	$42,246.40 (80 x $528.08)
Phase 2: Project Setup & Dashboard	2 Weeks	40	-	$21,123.20 (40 x $528.08)
Phase 3: AI Agent & Absence Mgmt	1 Week	15	-	$7,921.20 (15 x $528.08)
Phase 4: Integration, Testing & Handover	1 Week	15	-	$7,921.20 (15 x $528.08)
Total Hours	1200 Hours	150 Man-Days		
Total Cost	**12 Weeks**			**$79,212.00**

Cost Contribution distribution between Partner, Customer, AWS.

Party	Contribution (USD)	% Contribution of Total
Customer	0	0
Partner	0	0
AWS	200	100

7 ACCEPTANCE

1. **Project Acceptance Criteria**
   - The InsightHR platform will be considered complete and accepted when the following criteria are met.

2. **Completed Deliverables**
   - All major features implemented and deployed to production
   - User and employee management with bulk operations
   - Performance score management with calendar view
   - Attendance system with auto-absence marking
   - Interactive dashboard with live clock
   - AI chatbot with Bedrock integration

3. **Key Metrics Achieved**
   - 300+ user accounts
   - 300 employee records across 5 departments
   - 900+ performance scores tracked
   - 9,300+ attendance records
   - AWS monthly cost: ~$33.14
   - System uptime: 99.9%+
   - Zero critical security vulnerabilities

4. **Acceptance Status**
   - Current Status: Application deployed in cloudfront Production URL: https://d2z6tht6rq32uy.cloudfront.net

5. **Next Steps**
   - Minor bug fixing and feature updates
   - Conduct user acceptance testing
   - Provide knowledge transfer and training