

CloudPlay Game Hub: Your Smart Steam Game Recommender

Max Wong and Team
Spring 2025



Background & Business Use Cases



Problem Statement

Gamers struggle to decide if a game is worth buying due to:

1. **Fragmented** reviews across Steam and third-party sites.
2. **Lack of clear** sentiment insights (e.g., how positive or negative reviews are).
3. No easy way to get tailored game suggestions without **manually searching across multiple** third party game review platforms.

This often leads to **regretful** purchases or **missing out** on games they would actually enjoy.

Objective

Develop a cloud-based, interactive web platform that:

1. Analyzes user reviews using sentiment analysis.
2. Extracts and lists game titles with key metadata like genre, rating trends, and review sentiment.
3. Recommends games based on review patterns and emotional tone, not just popularity or pricing.

Mission

Help gamers make smarter game choices using **review-driven intelligence**, not guesswork.

Business Use Case

1. Deliver **holistic** recommendations using sentiment analysis across multiple review platforms.
2. Show **curated** game info (genre, price, platform) with sentiment breakdown in one place.
3. Surface **authentic** player insights, based on real experiences—not popularity or sales data.

Data Source & Procurement Details



Metacritic APIs

- Websites for gamers to rate & review games (user rating scaled from 0 to 10)
 - Retrieve user rating (scaled from 0 to 10) using provided APIs



SteamDB & Steam Web APIs

SteamDB

- Game names & app ID

Steam Web APIs

- Metadata for games (154.75MB, 300K rows)
 - User Reviews (JSONL file, 3.3 GB)
 - Examine key attributes and explore correlation between game and user ratings

Used to determine if the searched game is worth buying or not

Games on Steam

Metadata of Game

User Reviews of Games

JSONL File of User Reviews

Major Technologies Used



- Python Flask: Build web application
 - Amazon VPC: Network configuration (public and private)
 - Amazon EC2 Instance
 - Hosting webpage
 - Frontend and Backend API calls
 - Machine Learning Sentiment Analysis
 - Amazon S3: All code and data
 - Amazon DynamoDB: Store unstructured user reviews on games (JSONL file for faster query)
 - Amazon RDS: Store structured games raw metadata (CSV file)
 - Amazon Lambda: Data pipeline (for data processing)
 - Amazon EventBridge to trigger Lambda function



Design Overview

- Goal: Build a cloud-based system
 - Recommends games by analyzing real user reviews using sentiment analysis
 - Steam game data and reviews
 - Use machine learning (ML) & present curated recommendations through a web interface
- VPC: 2 public & 2 private subnets for secure network separation



7

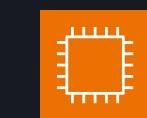
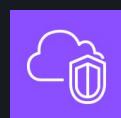
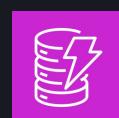
Design Overview



- Security
 - Security groups for EC2 instance
 - HTTP & SSH
 - Security groups for Aurora RDS
 - PostgreSQL



S3 Bucket

EC2
(Frontend +
Backend + ML)VPC
(2 Public,
2 Private
Subnets)Lambda
(Backup)EventBridge
(Triggers Lambda
Function)Aurora RDS
(game metadata
CSV File)DynamoDB
(JSONL File)

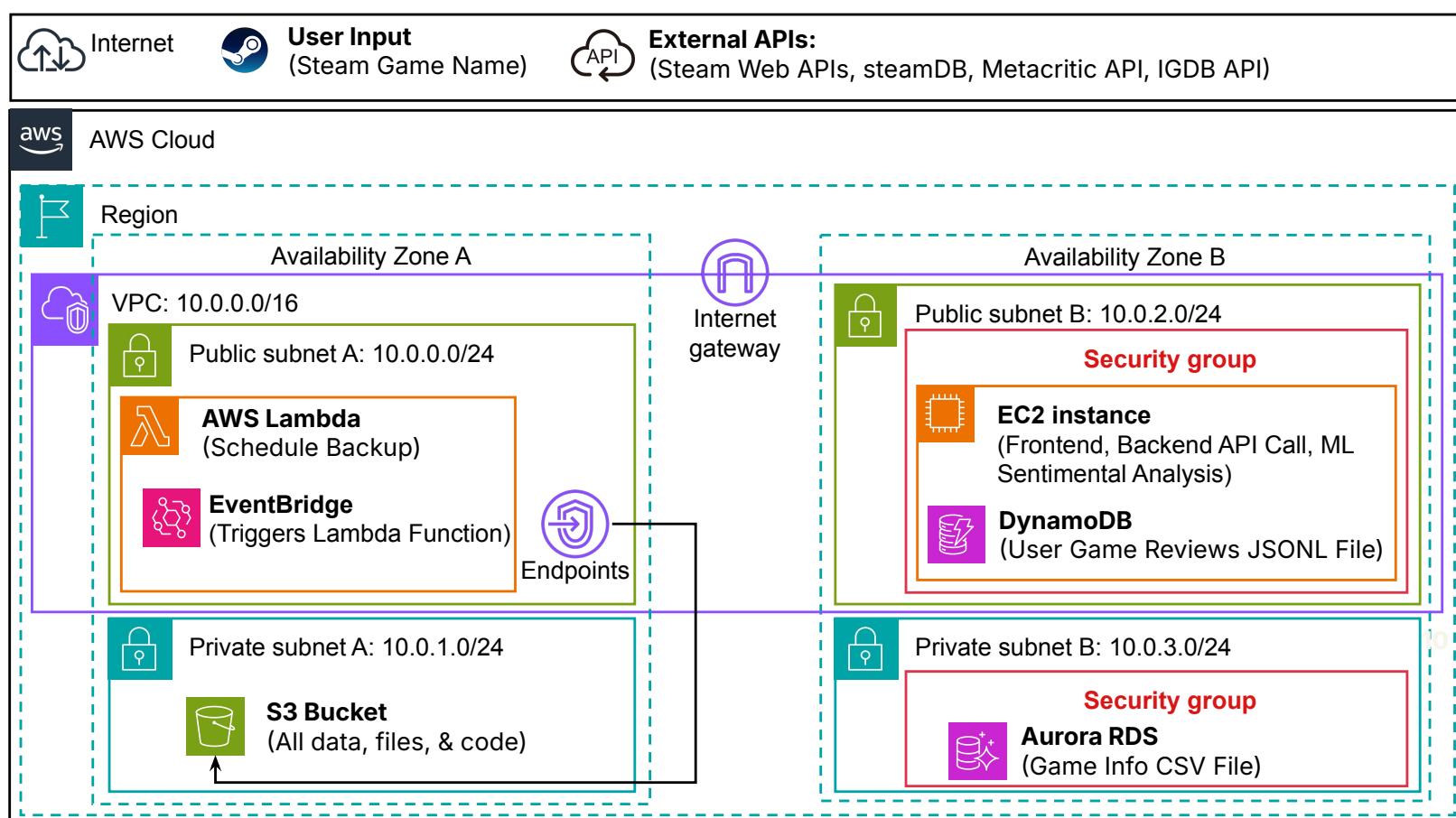
8

CloudPlay Game Hub

AWS Cloud Infrastructure (Student Account Version)



9





Virtual Private Cloud (VPC)

Your VPCs (1/2) Info					
Last updated 2 minutes ago Actions Create VPC					
EC2 Global View View Filter by VPC					
Virtual private cloud Your VPCs					
Name	VPC ID	State	Block Public...	IPv4 CIDR	
-	vpc-06074993616a5e25b	Available	Off	172.31.0.0/16	
<input checked="" type="checkbox"/> final-game-recommender-vpc-vpc	vpc-06982637718ef8326	Available	Off	10.0.0.0/16	

Subnets (10) Info					
Last updated 1 minute ago Actions Create subnet					
EC2 Global View View Filter by subnet					
Name	Subnet ID	State	VPC		
final-game-recommender-vpc-subnet-p...	subnet-035605662b415dfd1	Available	vpc-06982637718ef8326 final...		
final-game-recommender-vpc-subnet-p...	subnet-00b675c53c9bc5ad1	Available	vpc-06982637718ef8326 final...		
final-game-recommender-vpc-subnet-p...	subnet-0fcfcc343872267bb	Available	vpc-06982637718ef8326 final...		
-	subnet-0e10b01c6b518dcf7	Available	vpc-06074993616a5e25b		
final-game-recommender-vpc-subnet-p...	subnet-058018f09a0349fbf	Available	vpc-06982637718ef8326 final...		
-	subnet-030c505163f463a48	Available	vpc-06074993616a5e25b		
-	subnet-049dcf360e64abe73	Available	vpc-06074993616a5e25b		
-	subnet-0c75b707532b71abd	Available	vpc-06074993616a5e25b		
-	subnet-092818cac8c1d8958	Available	vpc-06074993616a5e25b		
-	subnet-0a497b20bf8a0e4c9	Available	vpc-06074993616a5e25b		

11



Data Storage

S3 (Backup Storage)

- Stores source code & ML model code
- User reviews & game metadata (e.g. title, genre, pricing)

S3 specs

- Standard storage class

DynamoDB (NoSQL Database)

- Holds JSONL-formatted user reviews scraped from Steam
- Ideal for flexible, scalable storage of unstructured data

DynamoDB

- Lives in EC2 instance

Aurora RDS (Relational Database)

- Stores structured relational data (CSV-based game info)
- Ideal for querying & joins across titles, genres, and developers

Aurora RDS specs

- Free tier
- Instance: db.t3.micro
- Storage 20 GiB General Purpose SSD gp2



S3 Bucket



Amazon EC2



Aurora RDS (CSV File)



DynamoDB (JSON File)

12



DynamoDB & S3

```
[ec2-user@ip-10-0-2-187 dynamodb_local]$ ls
DynamoDBLocal.jar  THIRD-PARTY-LICENSES.txt
DynamoDBLocal_lib  dbdata
LICENSE.txt        dynamodb-local-metadata.json
README.txt         dynamodb.log
[ec2-user@ip-10-0-2-187 dynamodb_local]$
```

The screenshot shows the AWS S3 console interface. On the left, there's a sidebar with navigation links like 'General purpose buckets', 'Storage Lens', and 'Feature spotlight'. The main area is titled 'Objects (12)' and contains a table listing 12 files. The table includes columns for Name, Type, Last modified, and Size. The objects listed are: apan5450.pem (pem), apan5450.pub (pub), app.py (py), background.jpg (jpg), base.html (html), fill_db.py (py), and results.html (html). All files were last modified on April 18, 2025, at 22:56:42 UTC-04:00.

13

Aurora RDS Setup



The screenshot shows the AWS Aurora RDS console. The left sidebar has links for 'Dashboard', 'Databases' (which is selected), 'Query Editor', etc. The main area shows the 'game-metadata' database details. Under the 'Summary' tab, it shows the DB identifier 'game-metadata', status 'Available', role 'Instance', engine 'PostgreSQL', and region 'us-east-1b'. The 'Connectivity & security' tab is active, showing the endpoint 'game-metadata.cdm2aaowkflq.us-east-1.rds.amazonaws.com', port '5432', and networking details like availability zone 'us-east-1b' and subnet group 'rds-subnet-group'. Security details include VPC security groups 'RDS-sg (sg-090f5415a7c98703a)' and public accessibility set to 'No'.

14



Service Delivery Through Web

Elastic Compute Cloud (EC2) instance hosts web server

- Frontend & backend API calls
- ML sentiment analysis
- Users can access our Steam game recommendation system directly through the web interface
- Fetch data from Aurora RDS & DynamoDB to serve dynamic content
- Locates in public subnet of VPC
- Security groups to control traffic that enters the instance

EC2 Configuration

- t2.large (2 vCPU, 8 GiB RAM)
- Amazon Linux 2023 OS
- Storage: 20 GiB General Purpose SSD gp2

Instances (1/1) Info

i-009aab6318b95500e (cloud-computing-project)

Details | Status and alarms | Monitoring | Security | Networking | Storage | Tags

Instance summary

Instance ID: i-009aab6318b95500e
IP address: 52.203.15.170
Hostname type: IP name: ip-10-0-2-187.ec2.internal
Answer private resource DNS name: ip-10-0-2-187.ec2.internal
Auto-assigned IP address: 52.203.15.170 [Public IP]
IAM Role: -
Instance state: Running
Instance type: t2.large
VPC ID: vpc-06982637718ef8326 (final-game-recommender-vpc-vpc)
Subnet ID: subnet-0fcfcc343872267bb (final-game-recommender-vpc-subnet-public2-us-east-1b)
Elastic IP addresses: 10.0.2.187
Private IPv4 address: 52.203.15.170 [open address]
Private IP DNS name (IPv4 only): ip-10-0-2-187.ec2.internal
Instance type: t2.large
VPC ID: vpc-06982637718ef8326 (final-game-recommender-vpc-vpc)
Subnet ID: subnet-0fcfcc343872267bb (final-game-recommender-vpc-subnet-public2-us-east-1b)
AWS Compute Optimizer finding: Opt-in to AWS Compute Optimizer for recommendations.
Auto Scaling Group name: -

15



Model Training

Using the VADAR sentiment analysis function to analyze users reviews

- VADAR can handle slang, punctuation, and contractions in reviews
- Calculate the compound sentiment score for each review
- Calculate the average compound sentiment score across all reviews
- Score ranges from -1 to 1
 - If no reviews, score = None
 - Positive Score = Recommend to purchase
 - Negative Score = Not recommend to purchase

16



Lambda

- Need automated & scheduled task that updates database info
- EventBridge
 - Trigger Lambda function
 - Setup cron schedule (quarterly)
 - Stops EC2 instance for database update

The screenshot shows the AWS Lambda console interface. At the top, there's a navigation bar with the AWS logo, search bar, and various icons. Below it, the main header says 'Lambda > Functions > stop_instance'. The main content area is titled 'stop_instance' and contains a 'Function overview' section. This section includes a 'Diagram' tab (which shows a single function block labeled 'stop_instance' with an 'EventBridge (CloudWatch Events)' trigger), a 'Template' tab, and buttons for '+ Add destination' and '+ Add trigger'. To the right of the diagram, there's a sidebar with sections for 'Description' (empty), 'Last modified' (3 days ago), 'Function ARN' (arn:aws:lambda:us-east-1:992382408616:function:stop_instance), and 'Function URL' (empty). There are also buttons for 'Throttle', 'Copy ARN', 'Actions', 'Export to Infrastructure Composer', and 'Download'.

17

Networking Configuration Challenges



- Limitations on EC2 instance
 - Best available option: t2.large
- VPC NAT gateway
 - Excessive cost → api calls cannot be placed in private subnets
 - EC2 instance and dynamoDB need to be in public subnets
- No permission to change DynamoDB IAM role
 - Create a local DynamoDB in EC2 instance (requires openJDK)
- Ensure DynamoDB listens to web page launched in EC2 instance
 - Data information is not shown on the website
- No permission to change Lambda IAM role
 - Cannot be used for machine learning

18



Security Configuration

Security Groups (4) [Info](#)

Actions Export security groups to CSV Create security group

Name	Security group ID	Security group name	VPC ID
-	sg-0d4cf7a4418feca90	default	vpc-06982637718ef8326 [2]
-	sg-090f5415a7c98703a	RDS-sg	vpc-06982637718ef8326 [2]
-	sg-0acc698862bc97a0f	default	vpc-06074995616a5e25b [2]
-	sg-030e24ef61385de73	game-recommender-sg	vpc-06982637718ef8326 [2]

sg-030e24ef61385de73 - game-recommender-sg

Details Inbound rules Outbound rules Sharing - new VPC associations - new Tags

Inbound rules (2)

Name	Security group rule ID	IP version	Type	Protocol	Port range
-	sgr-0700cae3b275259bd	IPv4	HTTP	TCP	80
-	sgr-0052603c5b6b7da2f	IPv4	SSH	TCP	22

sg-090f5415a7c98703a - RDS-sg

Details Inbound rules Outbound rules Sharing - new VPC associations - new Tags

Inbound rules (1)

Name	Security group rule ID	IP version	Type	Protocol	Port range
-	sgr-002673575482d378e	-	PostgreSQL	TCP	5432

Inbound Rules

- HTTP
 - Port 80, Anywhere IPv4
 - Web access (serve web page)
- SSH
 - Port 22, Anywhere IPv4
 - Remotely access & manage server
- PostgreSQL
 - Port 5432, any EC2 instance w/ the RDS-sg security group
 - Database access

19

CloudPlay Game Hub

Economics of Cloud Computing:
Cost Estimation



20

AWS Cost Under Current Scale				AWS Cost Projection			
Server	Cost Per Month	Descriptions	Total	Server	Cost Per Month	Descriptions	Total
EC2 (t2.large)	\$67.74	<ul style="list-style-type: none"> 2 vCPU, 8 GiB RAM. On-demand. \$0.0928/hr × 730 hrs (full month uptime) 	\$818.40 Per Year The Price Calculated by AWS Pricing Calculator	EC2 (c5.2xlarge) Aurora RDS (db.r6g.xlarge) S3 SageMaker (ml.m5.2xlarge)	\$156.22	<ul style="list-style-type: none"> 8 vCPU, 16 GiB RAM. Pricing strategy (EC2 Instance saving plans for 1 year). \$0.214/hr × 730 hrs (web server full month uptime) 	\$10,522.68 Per Year The Price Calculated by AWS Pricing Calculator
Aurora RDS (db.t3.micro)	\$0	<ul style="list-style-type: none"> Amazon Aurora PostgreSQL-Compatible DB instances 2 vCPU, 1 GiB RAM. Free-tier No charge for 730 hrs usage 			\$351.20	<ul style="list-style-type: none"> Amazon Aurora PostgreSQL-Compatible DB instances cost (Monthly): \$0.3401hr × 730 hrs 1 TB Storage (Monthly): \$102.93 4 vCPU, 32 GiB RAM. Pricing model (Reserved for 1 year). 	
S3	\$0.46	<ul style="list-style-type: none"> 20 GB storage 			\$23.55	<ul style="list-style-type: none"> 1 TB storage \$0.023 per GB 	
					\$345.92	<ul style="list-style-type: none"> Storage (General Purpose SSD (gp2), 100GB per month) Instance name (ml.m5.12xlarge: vCPU: 48, Memory: 192 GiB, Network Performance: 10 Gigabit) 	21

Google Cloud Cost Under Current Scale				Google Cloud Cost Projection			
Server	Cost Per Month	Descriptions	Total	Server	Cost Per Month	Descriptions	Total
Cloud Instance (n4-standard-2)	\$69.98	<ul style="list-style-type: none"> 2 vCPU, 8 GiB RAM On-demand 	\$1986.84 Per Year The Price Calculated by GCP Pricing Calculator	GCP Instance (N4-highcp u-8) Cloud PostgreSQL Cloud Storage Vertex AI Prediction (a2-highgp u-2g)	\$234.29	<ul style="list-style-type: none"> 8 vCPU, 16 GiB RAM Pricing strategy (Monthly Instance Plan for 1 year) 	\$14,539.62 Per Year The Price Calculated by GCP Pricing Calculator
Cloud PostgreSQL	\$95.29	<ul style="list-style-type: none"> GCP Cloud PostgreSQL 2 vCPU, 3.75 GiB RAM 100 GB database storage 			\$768.85	<ul style="list-style-type: none"> GCP Cloud PostgreSQL 10 vCPU, 64 GiB RAM 1 TB database storage 	
Cloud Storage	\$0.3	20 GiB storage			\$20.83	1 TB storage	
					\$170.08	<ul style="list-style-type: none"> 24 vCPU, 170GB RAM A100 40GB accelerator Average training job length 100 hours 	22

Windows Azure Cost Under Current Scale				Windows Azure Cost Projection			
Server	Cost Per Month	Descriptions	Total	Server	Cost Per Month	Descriptions	Total
Virtual Machines	\$240.90	<ul style="list-style-type: none"> • 2 vCPU, 14 GiB RAM. • On-demand. • \$0.33/hour 	\$4583.64 Per Year The Price Calculated by Azure Pricing Calculator	Virtual Machines	\$280.62	<ul style="list-style-type: none"> • 8 vCPU, 32 GB RAM. • Pricing strategy (Monthly Instance Plan). 	\$21,190.72 Per Year The Price Calculated by Azure Pricing Calculator
Azure Database for PostgreSQL	\$138.47	2 vCPU, 5 GiB RAM.		Azure Database for PostgreSQL	\$1023.17	<ul style="list-style-type: none"> • Azure Database for PostgreSQL • 16 vCPU • 1 TB database storage and 1 TB backup storage 	23
Azure Storage	\$2.60	32 GiB storage		Azure Storage	\$77	1 TB storage	
				Azure Machine Learning	\$228	<ul style="list-style-type: none"> • 24 vCPU, 224GB RAM and 1474GB temporary storage • NVv3 instance series: NV24s v3 • Average training job length 100 hours 	

Success Criteria

- Business Goals
 - Help users make smarter game purchases using sentiment analysis.
 - Reduce time spent searching for reviews by 30–50%
 - Surface valuable games that might otherwise be missed
- Cloud Goals
 - Deliver sub-second latency for Steam & Metacritic API fusion
 - Scale EC2 + Aurora with minimal cost spikes
 - Implement cost-optimized AWS architecture using reserved pricing





Six Pillars

- Reliability
 - Multiple AZs, reduce single point failure
 - S3 for backup
 - Lambda & EventBridge: Schedule database maintenance & update
- Availability
 - DynamoDB & Aurora: high availability & failover capabilities
- Performance
 - EC2 instance: customizable compute for ML process & API calls
 - DynamoDB: optimized, fast key-value reads of user reviews
 - Aurora RDS: structured game info with SQL capabilities
- Scalability
 - DynamoDB and S3: scale automatically
 - Improvement: use EBS for fault tolerance to store EC2 data & state
- Security
 - Private subnet: Aurora RDS & S3 (restricts public access)
 - Security groups control traffic
 - VPC Endpoints: secure EC2 to S3 communication (no internet gateway)
- Cost Optimization
 - S3, Lambda, EventBridge: serverless & pay-as-you-use
 - DynamoDB: on-demand (cost efficient for spiky workloads)

25

CloudPlay Game Hub

Website Design Screenshots



26

CloudPlay Game Hub

Find Your Perfect Game ;-)

Get Started

Enter Game Details

Game Name

Get Your Results

Home

Your Game Diagnostic Results

Game Details

AppID: 868230

Name: Mushroom Quest

Release Date: Jun 18, 2018

Developer: Flying Islands Team

Publisher: Flying Islands Team

Genres: Casual, Indie

Categories: Single-player, Steam Achievements, Steam Cloud, Family Sharing

Platforms: windows

Price: USD \$2.99

Metacritic Score: None

Description: Rethinking of the classic sokoban genre, slow and kind game with original gameplay elements. Mushroom Quest is a retro styled and addictive arcade, ideal to spend few unhurried evenings.

Recommendation: Purchase recommended

[Search Again](#)

[Home](#)

User Reviews

Review ID: 43052043

Author: 76561198040578523 (owned 299 games, 28 mins played)

Score: 1 (29 upvotes, 1 funny votes)

Review:

Mushroom Quest is a game that showed a lot of potential, but falls into 3 main traps that puzzle games often do. Note that it's not a bad game, it's just very easy, very short, and manages to have most of its difficulty be fake. I'll address each of these in turn along with what could be done differently. I'll also be comparing to similar games, like Abraxas Interactive PUSH and Chip's Challenge, the former of which I helped playtest, and the latter I've sunk thousands of hours into making and playing community content. First off: the difficulty. Mushroom Quest has very few mechanics. You can move, hold one key at a time, collect crystals, and push blocks. Blocks fall down and fill pits, and

Beyond Demo: Planned Improvements



Separated EC2 Instances

Deploy frontend and backend on separate EC2 instances to enhance data security



Cloud-Based ML with BERT

Upgrade RAM to support BERT model and migrate ML workflows to the cloud



Cost Analysis Update

Reevaluate and update cost analysis to reflect infrastructure upgrades



Enhanced Personalization

Incorporate Steam API to retrieve user behavior data for more personalized insights