ERD

USER
- user_id (PK)
- username
- email
- password
- role
- status
- created_at

COMMENT
- comment_id (PK)
- user_id (FK)
- content
- status
- created_at

MODERATOR_ACTION
- action_id (PK)
- moderator_id (FK)
- comment_id (FK)
- action_taken
- action_date

AI_ANALYSIS
- analysis_id (PK)
- comment_id (FK)
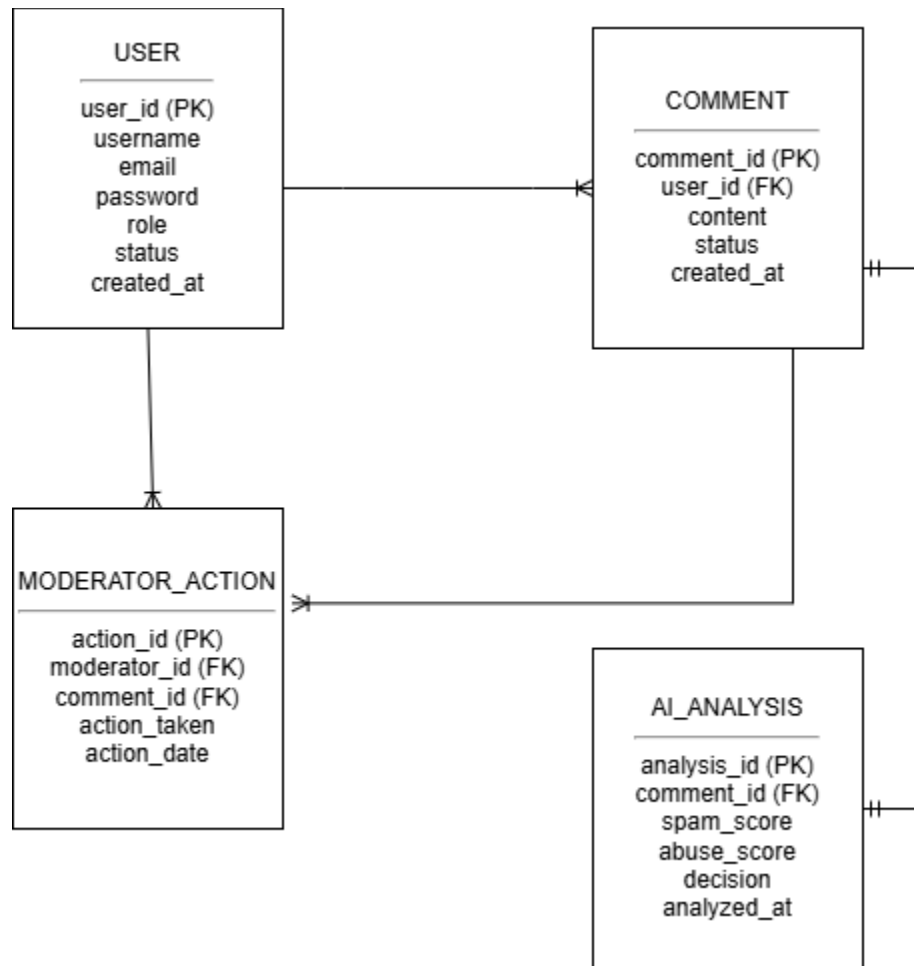- spam_score
- abuse_score
- decision
- analyzed_at

ACCEPTANCE CRITERIA

AI-Based Spam and Abuse Detection System

1. User Authentication Module

Acceptance Criteria:

- The system allows users to register with a valid username, email, and password.

- The system prevents registration with missing or invalid information.

- The system allows registered users to log in using valid credentials.

- The system denies access when invalid login credentials are entered.

- The system assigns a role (user, moderator, admin) to each account.

2. Comment Posting Module (User)

Acceptance Criteria:

- The system allows logged-in users to post text-based comments.

- The system displays posted comments in the comment feed.

- The system automatically submits comments to the AI analysis module upon posting.

- The system prevents empty comments from being submitted.

- The system records the comment with a timestamp and user ID.

3. AI Spam and Abuse Detection Module

Acceptance Criteria:

- The system analyzes each submitted comment using AI.

- The system generates a spam score for each comment.

- The system generates an abuse score for each comment.

- The system classifies comments as approved, flagged, or blocked.

- The system automatically blocks comments that exceed defined abuse thresholds.

- The system stores AI analysis results in the database.

4. Comment Status and Feedback Module

Acceptance Criteria:

- The system displays feedback to users when a comment is approved.

- The system notifies users when a comment is flagged for review.

- The system informs users when a comment is blocked due to policy violations.

- The system prevents blocked comments from being publicly visible.

5. Moderator Dashboard Module

Acceptance Criteria:

- The system allows moderators to view all flagged comments.

- The system displays spam and abuse scores for each flagged comment.

- The system allows moderators to approve flagged comments.

- The system allows moderators to delete abusive or spam comments.

- The system allows moderators to ban or suspend abusive users.

- The system records all moderator actions in the system logs.

6. User Management Module (Admin)

Acceptance Criteria:

- The system allows administrators to view a list of all users.

- The system displays user roles and account status.

- The system allows administrators to ban or unban users.

- The system prevents banned users from posting comments.

- The system logs administrative actions for auditing purposes.

7. Reports and Logs Module

Acceptance Criteria:

- The system generates reports on flagged and blocked comments.

- The system records AI analysis decisions for review.

- The system stores moderation history for accountability.

- The system allows administrators to view moderation statistics.

8. Non-Functional Requirements

Acceptance Criteria:

- The system responds to comment submissions in near real time.

- The system ensures secure storage of user credentials.

- The system maintains data consistency across all modules.

- The system supports multiple concurrent users.

- The system provides an intuitive and user-friendly interface.