# **Fawzooz AI - AI Incident Response Records**

Version: 1.0

Effective Date: [Date]

Document Owner: [e.g., Chief Information Security Officer (CISO)]

Maintained By: [e.g., AI Operations Team / Security Operations Center (SOC)]

### 1. Purpose and Scope

This document serves as the official log for all incidents related to Artificial Intelligence (AI) systems at Fawzooz AI. An AI incident is defined as any adverse event, security breach, or unexpected system behavior that results in, or has the potential to result in, harm to users, data breaches, service disruption, or reputational damage.

The purpose of this record is to:

* Ensure that all AI incidents are systematically tracked, investigated, and resolved.
* Provide a clear and auditable trail of actions taken in response to an incident.
* Facilitate root cause analysis to identify underlying issues and prevent future occurrences.
* Fulfill regulatory and compliance obligations for incident management and reporting.

This scope covers all AI systems within the Fawzooz AI Management System (AIMS) and all personnel involved in the incident response process.

### 2. Incident Management Process

1. **Detection & Reporting:** Incidents can be detected through automated monitoring or reported by users or staff. All potential incidents must be immediately reported to the AI Operations Team.
2. **Triage & Classification:** The incident is given an initial assessment to determine its severity and impact.
3. **Containment, Eradication & Recovery:** Immediate steps are taken to contain the incident and limit the damage. Once contained, the root cause is eradicated, and systems are restored to normal operation.
4. **Post-Incident Analysis:** A thorough analysis is conducted to determine the root cause, document lessons learned, and identify improvements to prevent recurrence.
5. **Logging:** All stages of this process must be meticulously documented in this Incident Response Record.

### 3. AI Incident Log

#### **Incident ID: [Unique Incident ID]**

| **Field** | **Description / Instructions** | **Entry** |
| --- | --- | --- |
| **Incident ID** | A unique identifier for the incident (e.g., AI-INC-2025-001). |  |
| **Status** | The current status of the incident (e.g., Active, Contained, Resolved, Closed). |  |
| **Severity Level** | The severity of the incident (e.g., Critical, High, Medium, Low). |  |
| **Date/Time Detected** | The exact date and time the incident was first detected or reported. |  |
| **Date/Time Resolved** | The date and time the incident was fully resolved and the case closed. |  |
| **Incident Reporter** | The name/team that reported the incident (e.g., User Report, Monitoring Alert). |  |
| **AI System(s) Affected** | The specific AI system(s) and version(s) involved in the incident. |  |
| **Incident Summary** | A brief, clear summary of the incident. What happened? |  |
| **Impact Assessment** | A detailed description of the actual or potential impact on users, data, operations, and the organization. |  |
| **Containment Actions** | A log of immediate actions taken to contain the incident and the date/time they were performed. |  |
| **Investigation Summary** | A summary of the investigation's findings. |  |
| **Root Cause Analysis** | A detailed analysis of the fundamental cause(s) of the incident. Why did it happen? |  |
| **Resolution Actions** | A log of actions taken to eradicate the cause and recover from the incident. |  |
| **Lessons Learned** | Key takeaways from the incident. What can be improved? |  |
| **Follow-up Actions** | Specific, actionable tasks to be completed to prevent recurrence (e.g., update risk register, modify SDLC, create new training). Include owner and due date. |  |
| **Lead Investigator** | The individual responsible for managing the incident response. |  |
| **Communication Log** | A summary of communications sent to stakeholders (e.g., users, regulators, leadership). |  |

### 4. Sample Entry

#### **Incident ID: AI-INC-2025-001**

| **Field** | **Entry** |
| --- | --- |
| **Incident ID** | AI-INC-2025-001 |
| **Status** | Closed |
| **Severity Level** | High |
| **Date/Time Detected** | 2025-07-22 14:30 EEST |
| **Date/Time Resolved** | 2025-07-23 11:00 EEST |
| **Incident Reporter** | User Report via In-App Feedback |
| **AI System(s) Affected** | Fawzooz AI Mindful Moments Recommender (v3.2) |
| **Incident Summary** | The AI recommender system began suggesting distressing or triggering content (related to grief) to users who had explicitly set preferences to avoid such topics. |
| **Impact Assessment** | Caused significant psychological distress to approximately 250 users who received inappropriate recommendations. Minor reputational damage from social media mentions. No data breach occurred. |
| **Containment Actions** | 1. Recommender service immediately disabled for all users (2025-07-22 14:45). 2. A static, non-personalized list of general wellness content was displayed as a fallback. |
| **Investigation Summary** | An update to the content tagging algorithm introduced a bug, causing it to misclassify a new set of articles on "processing loss" and ignore user preference filters. |
| **Root Cause Analysis** | The pre-deployment validation tests for the algorithm update did not include a specific test case to verify the exclusion of filtered topics. This was a gap in the V&V phase of the AI SDLC. |
| **Resolution Actions** | 1. Rolled back the recommender algorithm to the previous stable version (v3.1). 2. Re-enabled the service (2025-07-23 11:00). |
| **Lessons Learned** | Our V&V testing protocol must be expanded to include negative test cases that explicitly verify user-defined content filters are respected after any algorithm update. |
| **Follow-up Actions** | 1. (Owner: Head of QA) Update the AI SDLC V&V checklist to include mandatory filter verification tests. Due: 2025-07-30. 2. (Owner: CISO) Update AI Risk Register (AI-RISK-011) to reflect this new testing control. Due: 2025-07-30. |
| **Lead Investigator** | Lead, AI Operations |
| **Communication Log** | In-app banner and email sent to all affected users with an apology and explanation (2025-07-23 12:00). |