**Data Flow Diagrams & User Stories – SmartSDLC**

**Date:** 25 JUNE 2025  
**Team ID:** LTVIP2025TMID33777  
**Project Name:** SmartSDLC  
**Maximum Marks:** 4

**Data Flow Diagram – Level 1 (Text Description):**

This outlines the high-level flow of information within the SmartSDLC system.

**User Input**

* **Document Upload:** Users upload software requirement documents (PDF format) via the frontend interface.
* **Query Input:** Users can input specific queries about SDLC phase classification or ask AI for suggestions.
* **User Authentication:** Admin and Users log in securely to access the system.

**Processing**

* **Backend Controller (Python + FastAPI):** Receives and processes incoming documents and user queries.
* **PDF Parsing:** The uploaded documents are parsed using PyMuPDF or PDFplumber to extract raw text.
* **Prompt Engineering:** The extracted content and queries are converted into structured prompts for the language model.
* **AI Model Interaction:** Prompts are sent to IBM Granite/GPT-4 via Hugging Face for classification into SDLC phases.
* **Phase Classification:** Responses are parsed to identify categorized content under SDLC phases (Requirements, Design, Testing, etc.).
* **Logging:** The system logs user interactions, document metadata, and AI responses.

**Visualization**

* **Phase-wise Breakdown:** Results are displayed in a clear, tabulated format mapping sections of the text to corresponding SDLC phases.
* **Document Summary:** A brief summary of the uploaded document is shown with percentage breakdown per phase.
* **Interactive Interface:** Users can interactively view/edit suggestions for document refinement.

**Storage**

* **Session Memory:** User sessions and classified data are temporarily stored in the app session (Streamlit or in-memory caching).
* **Database (Future):** Future enhancements may use Firebase/PostgreSQL to persist uploads and results.

**Output**

* **AI-Categorized Content:** Clear classification and suggestions per SDLC phase shown on the interface.
* **Recommendations:** Additional improvement suggestions provided to align documents with ideal SDLC structure.
* **Download Options:** (Planned) Users can export results as PDF or DOCX.

**User Stories**

These user stories define the core system functionalities for SmartSDLC across multiple roles.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **User** | **Epic** | **USN** | **Story** | **Acceptance Criteria** | **Priority** | **Release** |
| Web User | Document Upload | USN-1 | As a user, I can upload a software requirement PDF for classification. | The document is accepted and parsed successfully. | High | Sprint-1 |
| Web User | AI Classification | USN-2 | As a user, I get SDLC phase classification from uploaded content. | The content is categorized under correct SDLC phases. | High | Sprint-1 |
| Web User | User Login | USN-3 | As a user, I can log in securely and access my dashboard. | Auth system verifies identity and grants access. | High | Sprint-1 |
| Web User | Results View | USN-4 | As a user, I can view the breakdown of content per SDLC phase. | Tabulated results appear with proper phase-wise mapping. | High | Sprint-2 |
| Web User | Query Assistant | USN-5 | As a user, I can ask queries related to SDLC suggestions. | AI provides context-aware, relevant answers. | Medium | Sprint-2 |
| Web User | Recommendation Feedback | USN-6 | As a user, I can get suggestions to improve the uploaded document. | Clear tips on missing or weak SDLC phases shown. | High | Sprint-2 |
| Web User | Export Results | USN-7 | As a user, I can download the AI-processed results as PDF/DOCX. | User can generate export file from results. | Medium | Sprint-3 |
| Admin | Dashboard Analytics | USN-8 | As an admin, I can see upload logs and AI performance stats. | Admin panel shows usage metrics and accuracy logs. | Medium | Sprint-4 |
| Developer | API Integration | USN-9 | As a dev, I can connect backend to Hugging Face securely. | API key is stored securely, and rate limits managed. | High | Sprint-1 |

**Summary:**  
This document describes the structured data flow for SmartSDLC and its key user-driven functionalities. The DFD outlines how user documents are handled, processed with AI, and delivered with categorized results. The user stories help track agile sprints, ensuring incremental delivery of all high-priority features like AI classification, document upload, and improvement feedback.