

# IBM AI AR techdoc

Ani Bitri

October 8, 2025

## Contents

<b>1</b>	<b>Problem Statement</b>	<b>2</b>
1.1	Introduction and Context . . . . .	2
1.2	Existing Technologies . . . . .	2
1.3	Gaps in Current Solutions . . . . .	2
1.4	Problem Definition . . . . .	3
<b>2</b>	<b>Project Overview</b>	<b>3</b>
2.1	Purpose . . . . .	3
2.2	Objectives within the scope . . . . .	3
2.3	Boundaries and Out-of-Scope Elements . . . . .	4
2.4	Expected Deliverables . . . . .	4
2.5	Target Outcomes . . . . .	4
<b>3</b>	<b>Requirements</b>	<b>4</b>
3.1	Functional Requirements . . . . .	4
3.2	Non-Functional Requirements . . . . .	4
<b>4</b>	<b>System Architecture</b>	<b>4</b>
4.1	Component Diagram . . . . .	5
4.2	Frontend Design . . . . .	5
4.3	Backend Design . . . . .	5
4.4	Technology Stack . . . . .	5
<b>5</b>	<b>Development Plan and Project Philosophy</b>	<b>5</b>
5.1	Methodology . . . . .	5
5.2	Time Management and Milestones . . . . .	5
5.3	Resource Management . . . . .	5
5.4	Risk Management . . . . .	5
<b>6</b>	<b>Conclusion</b>	<b>5</b>

# 1 Problem Statement

## 1.1 Introduction and Context

Technical documentation plays a critical role in helping developers understand, implement, operate and maintain software systems. In enterprise environments, documentation is the primary interface between developers and complex systems, making clarity and accessibility essential. However, despite its importance, technical documentation remains predominantly static, text-based and difficult to navigate. As systems grow in complexity, dynamic processes such as data flows, dependencies and interactions are often poorly represented in traditional documentation formats. Consequently, many users often struggle to find the information they need, leading to frustration, errors and inefficiencies. This raises the need for innovative solutions that can enhance the user experience and improve comprehension of technical documentation.

## 1.2 Existing Technologies

There are several existing technologies which address parts of the problem but do not yet offer a comprehensive solution. These include:

- **Static and Web-Based Documentation Systems:** Different Platforms like GitBook, ReadTheDocs and Markdown-based tools offer well-structured and accessible documentation but lack interactivity and visualization capabilities. Information is conveyed effectively through text and images but remains static and fails to envision the dynamic nature of modern software systems.
- **AI Assistants:** In the last few years, the rise of AI-powered assistants like ChatGPT and Gemini have been transformative. These tools can understand and interpret technical text, summarize content and answer user queries. Yet, these assistants operate in isolation from the documentation itself, requiring users to switch contexts and manually input queries, disconnecting them from visual elements and offering limited support for the understanding of system architecture and interactions.
- **Augmented Reality Applications:** Industrial AR tools, such as PTC Vuuforia, Microsoft Dynamics 365 Guides and Scope AR, have been used to overlay digital information onto real-world objects. These solutions enhance spatial understanding and are particularly useful in maintenance and training contexts. However, they typically rely on specific and expensive hardware, such as the HoloLens, or complex 3D models, thus limiting their accessibility for general users.

## 1.3 Gaps in Current Solutions

Despite the advancements in AR and AI technologies, the current solutions exhibit several limitations:

- **Lack of Integration between AR and AI:** Visualization (AR) and comprehension (AI) are treated as separate domains. There is limited integration or implementation that combines the strengths of both to deliver a seamless user experience.
- **Limited Visual Comprehension:** Text and 2D diagrams are insufficient for conveying complex interactions and dynamic processes. Current documentation formats do not effectively utilize visual aids to enhance understanding.
- **Hardware and Platform Barriers:** Most AR implementations are dependent on specialized hardware, such as the HoloLens or AR glasses, which are costly or not widely available. This restricts the accessibility of AR-enhanced documentation to a broader audience.

## 1.4 Problem Definition

Given the limitations of existing documentation systems and the potential of emerging technologies, there is a need for an innovative solution which can be easily accessible, interactive and capable of enhancing the understanding of technical documentation. In order to address this need, the project aims to develop an AI-powered mobile application that uses AR technology to augment technical documentation. The proposed system must allow users to point their mobile devices at a technical diagram and view interactive AR overlays that explain components, relationships and interactions. At the same time, an integrated AI assistant, IBM Watson/Granite, will be available to answer user queries and provide additional context.

Therefore, this project aims to bridge the gap between visual comprehension and intelligent interaction in technical documentation, ultimately demonstrating how static documents can become interactive, explainable and user-friendly.

## 2 Project Overview

### 2.1 Purpose

Many developers face challenges while reading and understanding technical documentation. This project aims to create an AI-powered mobile application which, supported by AR technology, will enhance the user experience by providing interactive and immersive documentation. Powered by IBM Watson/Granite, the app will offer several features to assist developers in navigating and comprehending complex technical documents, including text recognition, interactive AR overlays, chatbot assistance and more.

### 2.2 Objectives within the scope

The primary objectives of this project are to:

1. Implement AR diagram augmentation
  - Detect and track diagrams in printed and digital forms.
  - Overlay interactive elements on diagrams to provide additional context and explanations.
2. Integrate an AI assistant
  - Employ IBM Watson/Granite to interpret the scanned documentation text and answer user queries.
  - Support natural language questions such as "What is the purpose of this diagram?" or "Explain this concept in simpler terms."
3. Preserve accessibility and compliance
  - Keep the core document unchanged and externalize enhancements through AR overlays to comply with accessibility and legal requirements.
4. Develop a functional mobile prototype
  - Deliver a working mobile application prototype that demonstrates the key features and functionalities.
  - Conduct user testing to gather feedback and refine the application.
  - Provide a short demonstration.

## 2.3 Boundaries and Out-of-Scope Elements

To keep the project achievable within the given timeframe, the following elements are considered out of scope:

- Full production deployment or enterprise-level integration.
- Hardware-specific AR is excluded; the focus is on mobile devices.
- Cross-platform optimization beyond the primary target platform (e.g., iOS or Android).

## 2.4 Expected Deliverables

- A prototype mobile application demonstrating real-time recognition and overlay of technical documentation.
- Integrated AI assistant interface capable of answering user queries based on the documentation content.
- A comprehensive project report detailing the design, implementation, testing processes and evaluation results.
- VIVA presentation and demonstration.

## 2.5 Target Outcomes

- Enhanced user experience for developers interacting with technical documentation.
- Improved comprehension of complex technical concepts through interactive AR elements and AI assistance.
- A foundation for future development and potential commercialization of the application.

# 3 Requirements

Each requirement will be described in the format R<sub>n</sub>C/D, where n is the requirement number and C or D indicates whether the requirement is customer (C) or developer (D) oriented. For example, R<sub>1</sub>C refers to the first customer requirement, while R<sub>2</sub>D refers to the second developer requirement. Each requirement will be detailed with its description, priority, verification method and traceability.

## 3.1 Functional Requirements

List the key functionalities the software system must support. Provide clear and concise descriptions of features and interactions.

### 1. Augmented Reality System

## 3.2 Non-Functional Requirements

Outline performance, usability, reliability, and other quality attributes expected from the system.

# 4 System Architecture

Provide a high-level overview of the system architecture. Include diagrams where appropriate to illustrate the system components and their interactions.

## **4.1 Component Diagram**

## **4.2 Frontend Design**

## **4.3 Backend Design**

## **4.4 Technology Stack**

List the technologies, programming languages, frameworks, and tools that will be used in the project.

# **5 Development Plan and Project Philosophy**

## **5.1 Methodology**

Given the dynamic nature of the project and the need for iterative development, the Scrum framework within the Agile methodologies will be adopted. The approach will allow for flexibility, continuous feedback and incremental delivery of features. Weekly stand-up meetings will be held to discuss progress and address any challenges. The work will be organized into sprints, each lasting approximately one to two weeks, with clearly defined goals. At the end of each sprint, a review session will be held with the assessors to demonstrate progress and gather feedback.

## **5.2 Time Management and Milestones**

Detail the project timeline with key milestones and deliverables.

## **5.3 Resource Management**

Outline the resources (e.g., personnel, equipment) required for the project and how they will be allocated.

## **5.4 Risk Management**

Identify potential risks and outline strategies for mitigating them.

# **6 Conclusion**

Summarize the key points of the specification document and outline next steps.