FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

BITU 3923
WORKSHOP II
ANALYSIS OF PROPOSED SYSTEM

Student Name	MUHAMMAD PUTRA AZAM BIN MAZAH (Leader)	**Matric No.**	B032310414
	LOO WEN CHUAN		B032310334
	NUR HANIN AFRINA BINTI MOHAMMED SALLEH		B032310703
	MUHAMMAD AIMAN AFIQ BIN ZAINAL FITRI		B032420066
	YAP ZHAN HONG		B032310386
Programme	BITS	**Group No.**	12
Project Title	**Digital Automation for Transforming Equipment Drawing Data into Risk-Based Inspection (RBI) Inspection Plan**		
Supervisor Name	Dr Nor Hafeizah Hassan	**Authorized Signature and Stamp**	

3.2.1. System Overview

The proposed system is a digital automation platform designed to transform equipment data from General Arrangement (GA) drawings into structured asset databases and Risk-Based Inspection (RBI) plans. By integrating a Large Language Model (LLM), data validation, and automated report generation, the system eliminates manual data entry and reduces the time required to prepare inspection plans. It directly addresses the problems and limitations such as, inefficiency, data inconsistency, and delayed reporting in the current workflow.

3.2.2. Use of Large Language Model (LLM)

3.2.2.1 Introduction to LLM

A Large Language Model (LLM) is an advanced artificial intelligence model trained on massive amounts of text data to understand and generate human language. Using deep learning—especially the *transformer* architecture—LLMs can interpret context, extract meaning, and produce structured or summarized outputs from complex, unstructured text. According to Google Cloud AI (2024), LLMs enable applications to "analyze, reason, and generate language-based content from diverse inputs such as documents, images, or prompts," making them highly effective for data understanding tasks.

3.2.2.2 Usage of LLM in Proposed System

In this project, which automates the transformation of equipment GA drawing data into Risk-Based Inspection (RBI) plans, we integrate an LLM-based approach alongside OCR technology.
From the video, *"OCR vs LLM for Data Extraction"* by N8n School (2024) the presenter demonstrates that:
"LLMs with vision capabilities outperform traditional OCR when processing complex or multimodal documents, such as those containing charts, graphs, or tables, because they can infer meaning and context, not just extract text."
This capability directly benefits our system, which must interpret semi-structured and technical content from engineering drawings and templates.

3.2.2.3 Advantages of LLM Integration in Proposed System

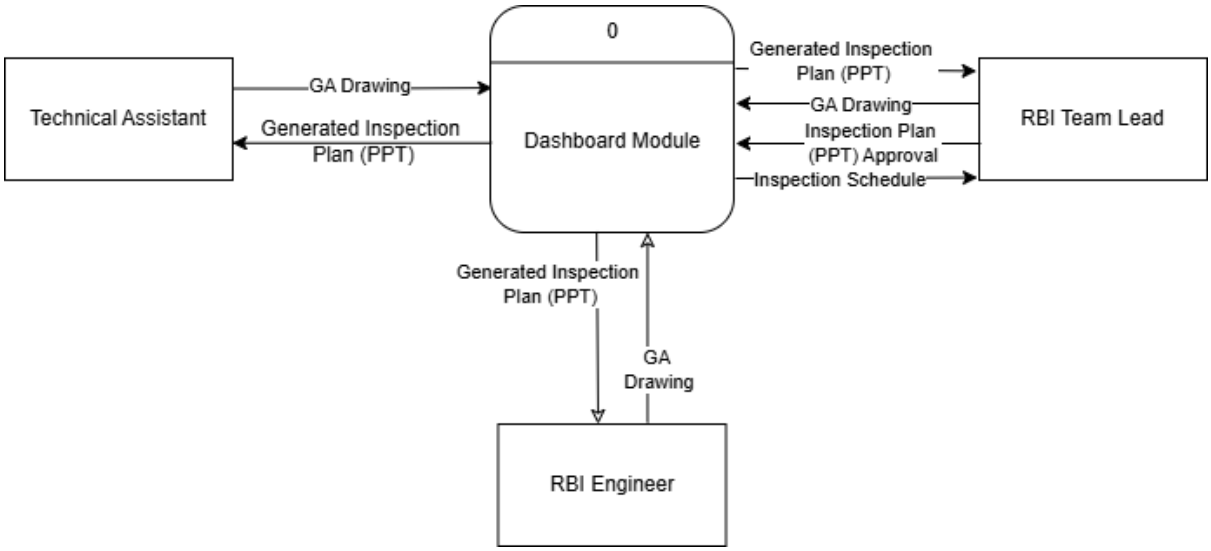
1. Improved Data Understanding
   Unlike traditional OCR that only reads visible text, an LLM can understand relationships, infer context, and interpret text near diagrams or symbols, helping accurately classify equipment information from GA drawings.
2. Flexible Extraction and Mapping
   The LLM supports natural-language reasoning, making it easier to map unstructured text into structured database fields.
3. Enhanced Report Generation
   In the Inspection Plan Generation Module, the LLM helps formulate descriptive and

consistent content within PowerPoint reports, ensuring alignment with API 580/581 wording and terminology.

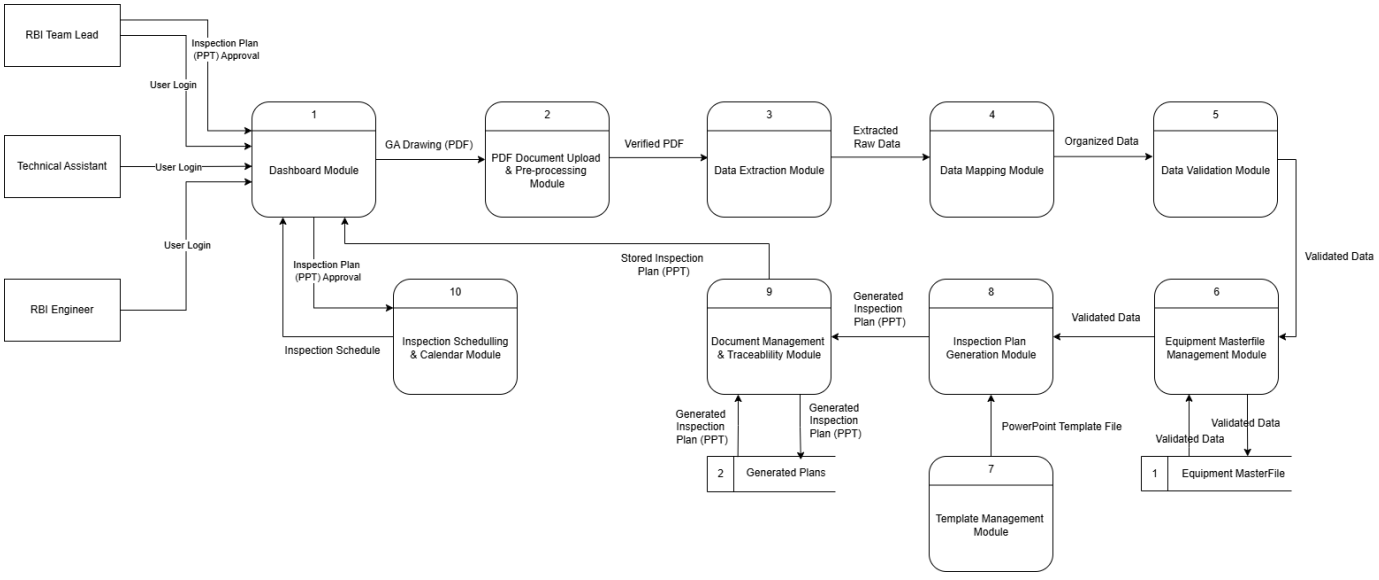
4. Contextual Inference
   As highlighted in the video, LLMs excel at "interpreting charts and long-form documents where OCR fails to capture relationships or meaning." This allows the system to extract insights from complex tables and engineering notes.
5. Scalability and Adaptability
   Because LLMs can be re-prompted or fine-tuned instead of manually re-programmed, the system can easily adapt to new drawing types or reporting requirements in future versions.

3.2.3 Data Flow Diagram

3.2.3.1 Context Diagram



3.2.3.2 Level 0 Diagram



3.2.4 Module Description

Table 3.2.4.1 Modules Description

Module Name	Description
PDF Document Upload & Pre-Processing Module	This module handles the secure upload and initial preparation of PDF drawings for processing. It validates file format and optimizes document quality for accurate data extraction.
Data Extraction Module	This module automatically scans and reads equipment GA drawings to extract raw text and data elements. Identifies key information blocks and prepares them for structured mapping.
Data Mapping Module	This module processes raw equipment data and maps it to standardized fields in the equipment master file. Converts unstructured text into organized, categorized equipment attribute.
Data Validation Module	This module systematically checks extracted data for accuracy, completeness, and logical consistency against engineering rules. Flags errors and anomalies for manual review before final storage.
Equipment Masterfile Management Module	This module serves as the centralized database for all validated equipment data and technical specifications. Enables CRUD operations and maintains version control for asset information.
Template Management Module	This module maintains and manages PowerPoint inspection plan templates with predefined placeholder fields. Ensures template integrity and provides validated templates for report generation.
Inspection Plan Generation Module	This module automatically populates template placeholders with equipment data to create customized inspection plans. Generates final PowerPoint outputs ready for review and approval.
Document Management and Traceability Module	This module provides a centralized interface where users can view all uploaded and generated files (PDFs, Excel, PPTs) with their source and output relationships clearly linked.
Inspection Scheduling & Calendar Module	This module converts approved inspection plans into scheduled tasks with assigned due dates and resources. Integrates with calendar systems and sends reminder notifications
Dashboard Module	This module serves as the central user interface and control centre for the entire system. It lets user to view, manage, and control every stage.

3.2.5 Module Input and Outputs

Table 3.2.5.1 Module Inputs and Outputs

Module Name	Inputs	Outputs
Dashboard Module	User Login	Dashboard visualization (charts, KPIs)
PDF Document Upload & Pre-Processing Module	GA Drawing (PDF)	Verified PDF file Error message if invalid format
Data Extraction Module	Verified PDF file	Extracted raw data
Data Mapping Module	Extracted raw data	Organized data
Data Validation Module	Organized data	Validated data Error log
Equipment Masterfile Management Module	Validated data	Updated equipment Masterfile (Excel) Stored version history
Template Management Module	None	PowerPoint template files
Inspection Plan Generation Module	Updated equipment Masterfile (Excel) PowerPoint template files	Generated inspection plan (PPT) Draft inspection report
Document Management & Traceability Module	Generated inspection plan (PPT) Draft inspection report	Stored Inspection Plan (PPT)
Inspection Scheduling & Calendar Module	Stored Inspection Plan (PPT)	Inspection schedule

3.2.6 Advantages of Proposed System

The proposed Digital Automation System for Transforming Equipment Drawing Data into Risk-Based Inspection (RBI) Plans directly addresses the five key problems identified in the current manual system at IPETRO Services Sdn. Bhd. It provides targeted solutions to improve accuracy, efficiency, and data consistency in the RBI planning workflow.

Problem	Solution using Proposed System	Stakeholders Affected
The extraction of equipment data from GA drawings is done manually.	The new system automates this task through the PDF Document Upload & Pre-Processing Module and the Data Extraction Module, which uses an LLM to scan GA drawings and extract relevant information automatically. This reduces manual effort and significantly shortens the time required to process each drawing.	1. Technical Assistant 2. RBI Engineer 3. RBI Lead
High risk of human error during data entry.	The Data Mapping Module and Data Validation Module ensure all extracted and manually entered values follow standardized	1. Technical Assistant

	formats, units, and required fields. Validation rules prevent inconsistent entries, missing attributes, and incorrect formats, reducing human error during transcription. Any incorrect or incomplete data is flagged and returned through the validation workflow before it is stored in the Equipment Masterfile Management Module.	2. RBI Engineer 3. RBI Lead
No integration between PDF, Excel, and PowerPoint workflows.	The system integrates both processes through the Equipment Masterfile Management and Inspection Plan Generation Modules, which automatically link validated Excel data with standardized PowerPoint templates. This ensures that all inspection plans are generated directly from the same verified data source, eliminating duplication and mismatch.	1. Technical Assistant 2. RBI Engineer 3. RBI Lead
Lack of proper version control for shared documents.	The Document Management and Traceability Module stores all extracted data, Masterfile updates, and generated reports with built-in version tracking. This creates full traceability and prevents the issues caused by multiple outdated versions stored in different team folders or personal laptops.	1. Technical Assistant 2. RBI Engineer 3. RBI Lead
No automation for applying standards and formatting.	Through the Template Management and Inspection Plan Generation Modules, the system standardizes inspection plan templates according to API 580 and API 581 guidelines. Automation ensures uniform report structure and formatting while reducing report preparation time, improving both consistency and efficiency.	1. Technical Assistant 2. RBI Engineer 3. RBI Lead
Must remember units, codes, formatting rules	The Data Validation Module enforces standardized units, naming rules, and formatting. The Template Management Module ensures all generated documents follow the required format, so staff do not rely on memory.	1. Technical Assistant
Inconsistent data entry across staff.	The Data Mapping Module and Data Validation Module standardize all mapped fields. The Equipment Masterfile Management Module ensures everyone uses the same validated dataset, eliminating differences between staff entries.	1. Technical Assistant
Works on local Excel files causes version conflicts.	The Equipment Masterfile Management Module centralizes data storage on the system instead of using local files. Everyone accesses one live Masterfile, preventing conflicts between multiple file versions.	1. Technical Assistant

Files overwritten when uploaded to Dropbox	The Document Management and Traceability Module performs secure document storage with automatic versioning. No files can be overwritten because each upload creates a new version with user and timestamp.	1. Technical Assistant
No tracking of which drawing already processed lead to duplicate work.	The Dashboard Module tracks processing status for every drawing.	1. Technical Assistant
Difficulty reading blurry/rotated drawings	The Pre-Processing Module (part of PDF Document Upload & Pre-Processing) automatically cleans, rotates, and enhances uploaded PDFs before extraction. This improves accuracy for low-quality drawings.	1. Technical Assistant
Asset register often contains missing or incorrect data (units, grades, pressure/temperature).	The Data Validation Module detects missing fields and incorrect units/grades. Validated data is then stored securely in the Equipment Masterfile Management Module, ensuring upstream errors are corrected before use.	1. RBI Engineer
Manual transfer of data to PowerPoint is time-consuming.	The Inspection Plan Generation Module automatically pulls data from the Equipment Masterfile and inserts it into the PowerPoint inspection plan templates managed by the Template Management Module. This removes all manual copying.	1. RBI Engineer
Relies on latest Excel updates but often receives outdated versions.	The Document Management and Traceability Module ensures that engineers always access the latest validated data version from the centralized Equipment Masterfile. No outdated local copies are used.	1. RBI Engineer
Difficult to coordinate when many engineers work on the same project	The Dashboard Module displays project-wide progress, assigned tasks, and status updates. The Inspection Scheduling & Calendar Module coordinates shared work timelines and prevents overlapping tasks.	1. RBI Engineer
Review delays due to multiple conflicting Excel/PPT versions.	The Document Management and Traceability Module eliminates multiple conflicting versions by storing Masterfile and inspection plans with version histories. The RBI Lead reviews only the latest approved document.	1. RBI Lead
No traceability on who updated what and when.	The Document Management and Traceability Module maintains a full audit log. This allows the RBI Lead to verify all updates quickly.	1. RBI Lead
Hard to maintain standard across large teams and many projects.	The Template Management Module enforces standard layouts, tables, and API 580/581 compliance across all projects. The Inspection Plan Generation Module ensures	1. RBI Lead

	that every inspection plan follows the same rules and structure, regardless of who is preparing it.	