

# Task Brief – Automated Numbered Marking & Description Extraction from Engineering Drawings

## 1. Objective

Develop a system that can automatically process engineering drawing PDFs, visually mark important text elements with numbered circles, and generate a corresponding output list where each number maps to the description of the marked text.

## 2. Input

- Engineering drawing in PDF format
- The drawing will be fed directly into the system
- A Google Drive link may be shared only for reference or examples

Drawings may contain:

- Dimensions
- Notes
- Tolerances / GD&T
- View labels (Top View, Section View, Auxiliary View, etc.)
- Feature descriptions
- Datum identifiers (e.g., X1, Z1)

Reference link of input image :

<https://drive.google.com/file/d/1q6PFy1BMgy2uB6Go4L8RIJLQ3hthIhHO/view?usp=sharing>

## 3. Core System Requirements

### 3.1 Visual Marking on Drawing

- Detect text elements present in the drawing PDF
- Identify relevant text based on configurable rules
- Draw a circle around the selected text
- Assign a unique number to each circle

Numbering Rules:

- Sequential numbering (1, 2, 3, ...)
- Each number must be unique within a drawing
- The same number must be reflected in the output

Annotation Rules:

- Circles must fully enclose the text
- Circles and numbers must be clearly visible
- Original drawing content must remain unchanged; markings must be overlaid

### **3.2 Number-to-Description Mapping (Output)**

For every numbered circle, the system must generate a corresponding output entry containing:

- Circle Number
- Description (exact text extracted from the drawing)

Example (for illustration only):

- 1 – Dimension 18 × 45 TYP
- 2 – Dimension 3.5
- 3 – Dimension 1.5
- 4 – Auxiliary View P
- 5 – Dimension 11
- 6 – 7.5 MIN WALL THICKNESS
- 7 – Flatness 0.2

Any table or format shown earlier is only for reference. The system is not constrained to a fixed column layout.

## **4. Critical Exclusion Rule – Datum Identifiers**

### **4.1 What Must NOT Be Marked**

- Datum identifiers such as X1, Z1
- Similar labels shown in black circles in the drawing
- Any datum-related symbols or references

These elements are part of the original drawing notation and must remain untouched.

### **4.2 Expected Behavior**

- Detect datum identifiers
- Explicitly exclude them from:
  - Circular marking
  - Numbering
  - Output description mapping

Visual intent:

- Black circled datums (X1, Z1, etc.) → Do NOT mark
- Dimensions, notes, tolerances, and feature descriptions → Mark and number

## **5. Output**

### **1. Annotated Drawing:**

- PDF containing numbered circles around selected text

### **2. Structured Output File:**

- Table / Excel / CSV / JSON

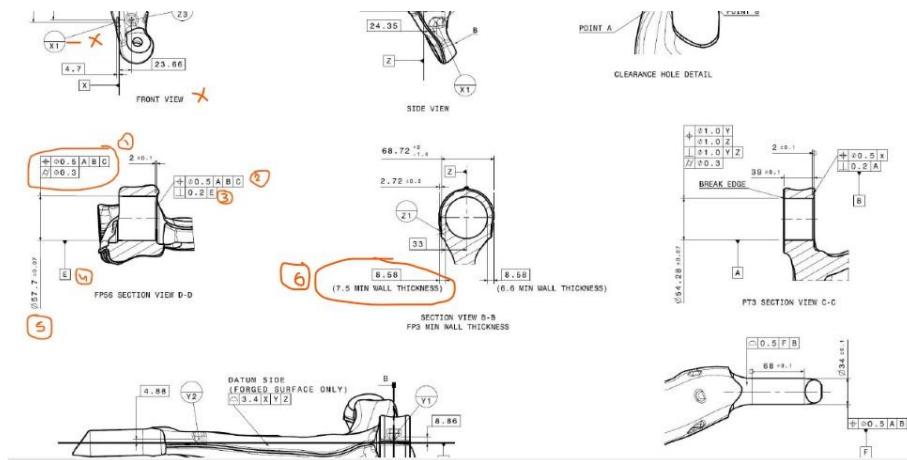
- Each entry includes:
  - Circle number
  - Corresponding text description

The numbering in the annotated drawing and the output must always match.

Output Annotated file sample :

(circle can be on whole text group like in 1 and can be divided like in 2 and 3)

(X should not be circled)



Output excel file sample:

ITEM - Technical Review	
Note: 1) Copy of marked drawing to be attached 2) Use additional	
Vendor Code: B03380	
Part No. / Mv: 548232453802_31/05/2024_Rev d	
S.No.	Parameters Critical to Fitment & Function (aesthetics, if applicable)
Y	Details Y
1	Dimension 1X45 TYP
2	Dimension 3.5
3	Dimension 1.5
P	Auxiliary View P
4	Dimension 11
5	Dimension 9
6	R13
7	Flatness 0.2
Top View	Top View1
8	Dimension 108.7
9	Eds Bracket 2
10	DIRT Shield 1
11	Eds Bracket 1