

In collab when we run the code we get in this format

```
=====
PROCESSING: page_30_preprocessed.jpg
=====
→ Pre-processing: Starting...
• Correcting tilt...
• Removing noise...
• Enhancing contrast...
• Saved preprocessed: page_30_preprocessed_preprocessed.jpg
✓ Pre-processing: Complete
→ OCR: Extracting text with Groq Vision...
✓ OCR: Extraction complete
→ Text Cleaning: Processing...
✓ Text Cleaning: Complete
✓ Saved: output/page_30_preprocessed_structured.json
✓ Saved: output/page_30_preprocessed_report.txt
✓ Saved: output/page_30_preprocessed_preprocessed.jpg
→ Creating redacted image...
✓ Redaction complete
✓ Saved redacted image: output/page_30_preprocessed_redacted.jpg
=====
```

#### EXTRACTION SUMMARY

Patient: Santosh Pradhan (36/M)

Medical Notes: 2

Note 1: Progress Note

```
└ Doctor: Dr. Pruthia
└ Date: 11/9/24 3:30 PM
└ Medications: 4
• Inj THIAMINE (200) 100 ml NS TID
• Inj LOPEZ (1amp) slow IV TDS as - q sos
• Inj PAN (40) 1 vial IV BD
... and 1 more
```

Note 2: Progress Note

```
└ Doctor: Dr. Hiranmaya Mohanta
└ Date: 11/9/25 12:00 PM
└ Medications: 4
• Inj THIAMINE (200) 100 ml NS TID
• Inj LOPEZ (1amp) slow IV TDS as - q sos
• Inj PAN (40) 1 vial IV BD
... and 1 more
```

Total PII Detected: 7 items across 6 categories

The overall output of the model will be given as zip file where we need to download and extract it which consist of all this outputs

Summary text

```
D: > Download1 > results > page_14_pii_summary.txt
1 === PII REDACTION SUMMARY ===
2 |
3 [DOCTOR_NAME] Dr. H.M (Confidence: high)    No link definition found: 'DOCTOR_NAME'
4 [DOCTOR_NAME] Dr. Rakesh (Confidence: high)   No link definition found: 'DOCTOR_NAME'
5 [DOCTOR_NAME] Dr. A. Mahapatra (Confidence: high)  No link definition found: 'DOCTOR_NAME'
6 [DOCTOR_NAME] Dr. S. Dash (Confidence: high)   No link definition found: 'DOCTOR_NAME'
7 [DATE] 11/4/25 (Confidence: high)    No link definition found: 'DATE'
8 [DATE] 15.4.25 (Confidence: high)   No link definition found: 'DATE'
9 [DATE] 16.4.25 (Confidence: high)  No link definition found: 'DATE'
```

Extracted output in json format

```
1 "raw_text": "Drug (Approved Name) | Dose | Route | Other Direction | Date & Time
2 "pii_entities": [
3     {
4         "type": "DOCTOR_NAME",
5         "value": "Dr. H.M",
6         "confidence": "high"
7     },
8     {
9         "type": "DOCTOR_NAME",
10        "value": "Dr. Rakesh",      "Rakesh": Unknown word.
11        "confidence": "high"
12    },
13    {
14        "type": "DOCTOR_NAME",
15        "value": "Dr. A. Mahapatra",    "Mahapatra": Unknown word.
16        "confidence": "high"
17    },
18    {
19        "type": "DOCTOR_NAME",
20        "value": "Dr. S. Dash",
21        "confidence": "high"
22    },
23    {
24        "type": "DATE",
25        "value": "11/4/25",
26        "confidence": "high"
27    },
28    {
29        "type": "DATE",
30        "value": "15.4.25",
31        "confidence": "high"
32    }
```

PII REDACTED		INSTITUTION CODE NUMBER: 102																																																																					
<b>PLAN OF CARE</b>		<b>INSTITUTE OF MEDICAL SCIENCES &amp; SUM HOSPITAL</b> (Faculty of Medical Sciences) <b>SIKSHA 'O' ANUSANDHAN</b> (DEEMED TO BE UNIVERSITY) K-8, Kalindi Nagar, Bhubaneswar  																																																																					
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left; padding: 5px;">Residents Y N</th> <th style="text-align: left; padding: 5px;">Type</th> <th style="text-align: left; padding: 5px;">Reason</th> </tr> </thead> <tbody> <tr> <td style="padding: 5px;">Investigation Advised</td> <td style="padding: 5px;">Treatment Advised</td> <td style="padding: 5px;"></td> </tr> <tr> <td style="padding: 5px;">CBC</td> <td style="padding: 5px;">Rx +</td> <td style="padding: 5px;">1) INJ. THIAMINE (200) I.V. 100 ml NS TID</td> </tr> <tr> <td style="padding: 5px;">WFT</td> <td></td> <td style="padding: 5px;">2) INJ. LOPEZ (1amp) slow IV NS VSOS</td> </tr> <tr> <td style="padding: 5px;">LFT</td> <td></td> <td style="padding: 5px;">3) INJ. PAN (100) 1 vials IV OD B/F</td> </tr> <tr> <td style="padding: 5px;">S-electrolytes</td> <td></td> <td style="padding: 5px;">4) TAB. TAXIM-O (200) 1 tab BD</td> </tr> <tr> <td style="padding: 5px;">U/H</td> <td></td> <td></td> </tr> <tr> <td style="padding: 5px;">ECG</td> <td></td> <td></td> </tr> <tr> <td style="padding: 5px;">S- Amylase / S-Urinary</td> <td></td> <td></td> </tr> <tr> <td style="padding: 5px;">GGT</td> <td></td> <td></td> </tr> <tr> <td style="padding: 5px;">T<sub>3</sub>/T<sub>4</sub> ITSM</td> <td></td> <td></td> </tr> <tr> <td style="padding: 5px;">PPBS / FBG</td> <td></td> <td></td> </tr> <tr> <td colspan="3" style="padding: 5px;">→ NCCT Brain</td> </tr> <tr> <td colspan="3" style="padding: 5px;">UDS</td> </tr> <tr> <td colspan="2" style="padding: 5px;">           Diet Consistency (Normal / Soft / Liquid / Tube Feeds / Fasting)         </td> <td colspan="2" style="padding: 5px;"></td> </tr> <tr> <td colspan="2" style="padding: 5px;">           Type (Normal / Diabetic / Lowsalt / Other)         </td> <td colspan="2" style="padding: 5px;"></td> </tr> <tr> <td colspan="4" style="padding: 5px;">Preventive Care</td> </tr> <tr> <td colspan="4" style="padding: 5px;">           P.G. Resident <i>[Signature]</i>            Dept. of Psychiatry            I.M.S. &amp; S.O.U. Hospital         </td> </tr> <tr> <td colspan="4" style="padding: 5px;">References</td> </tr> <tr> <td colspan="4" style="padding: 5px;">           Resident Signature <i>[Signature]</i>            Full Name: DR. PRATEEK PRATAP            Mobile No.: 9843344556            Date:            Time:         </td> </tr> <tr> <td colspan="4" style="padding: 5px;">           Consultant Signature <i>[Signature]</i>            Full Name: DR. UJJWAL MISHRA            Mobile No.: 9843344556            Date:            Time:         </td> </tr> </tbody> </table>		Residents Y N	Type	Reason	Investigation Advised	Treatment Advised		CBC	Rx +	1) INJ. THIAMINE (200) I.V. 100 ml NS TID	WFT		2) INJ. LOPEZ (1amp) slow IV NS VSOS	LFT		3) INJ. PAN (100) 1 vials IV OD B/F	S-electrolytes		4) TAB. TAXIM-O (200) 1 tab BD	U/H			ECG			S- Amylase / S-Urinary			GGT			T <sub>3</sub> /T <sub>4</sub> ITSM			PPBS / FBG			→ NCCT Brain			UDS			Diet Consistency (Normal / Soft / Liquid / Tube Feeds / Fasting)				Type (Normal / Diabetic / Lowsalt / Other)				Preventive Care				P.G. Resident <i>[Signature]</i> Dept. of Psychiatry I.M.S. & S.O.U. Hospital				References				Resident Signature <i>[Signature]</i> Full Name: DR. PRATEEK PRATAP Mobile No.: 9843344556 Date: Time:				Consultant Signature <i>[Signature]</i> Full Name: DR. UJJWAL MISHRA Mobile No.: 9843344556 Date: Time:			
Residents Y N	Type	Reason																																																																					
Investigation Advised	Treatment Advised																																																																						
CBC	Rx +	1) INJ. THIAMINE (200) I.V. 100 ml NS TID																																																																					
WFT		2) INJ. LOPEZ (1amp) slow IV NS VSOS																																																																					
LFT		3) INJ. PAN (100) 1 vials IV OD B/F																																																																					
S-electrolytes		4) TAB. TAXIM-O (200) 1 tab BD																																																																					
U/H																																																																							
ECG																																																																							
S- Amylase / S-Urinary																																																																							
GGT																																																																							
T <sub>3</sub> /T <sub>4</sub> ITSM																																																																							
PPBS / FBG																																																																							
→ NCCT Brain																																																																							
UDS																																																																							
Diet Consistency (Normal / Soft / Liquid / Tube Feeds / Fasting)																																																																							
Type (Normal / Diabetic / Lowsalt / Other)																																																																							
Preventive Care																																																																							
P.G. Resident <i>[Signature]</i> Dept. of Psychiatry I.M.S. & S.O.U. Hospital																																																																							
References																																																																							
Resident Signature <i>[Signature]</i> Full Name: DR. PRATEEK PRATAP Mobile No.: 9843344556 Date: Time:																																																																							
Consultant Signature <i>[Signature]</i> Full Name: DR. UJJWAL MISHRA Mobile No.: 9843344556 Date: Time:																																																																							
<b>PROGRESS REPORT</b>		<b>DATE &amp; TIME</b> <b>PROGRESS NOTES</b> <b>TREATMENT/ADVICE</b>																																																																					
11/11/25 3:00 PM		patient received from OPD to Ward-1 on date → 11/11/25 : from → 12:00 PM (D1) MENTAL AND BEHAVIOURAL DISORDER DUE TO USE OF ALCOHOLIC DEPENDENCE SYNDROME CURRENTLY TAKING THE SUBSTANCE (F10.24+F12.1)																																																																					
<b>Rx</b> 1) Inj. THIAMINE (200) I.V. 100 ml NS TID 2) Inj. LOPEZ (1amp) slow IV TDS and SOS 3) Inj. PAN (100) 1 vials IV OD B/F 4) TAB. TAXIM-O (200) 1 tab BD		<b>Visits</b> BP → PR → RR → Temp →																																																																					
(Contd 3)																																																																							