

Schematic AI Extractor

Angie Ng, See Tien;
Kiu, Siang Hui; Lee, Zhe Jin; Yen Ming, Siaw
Intel Corporation,
Cloud Based Remote Debug (CBRD),
Penang, Malaysia



Korea

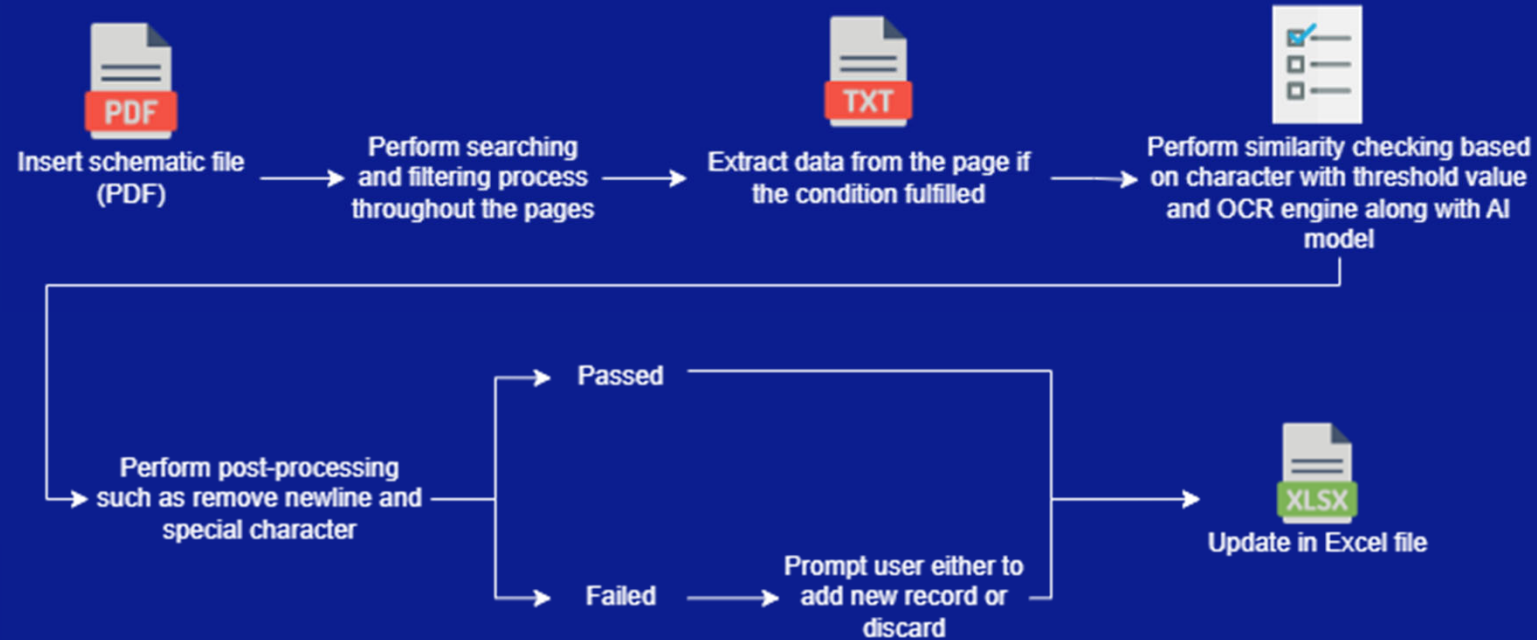
Incheon - November 7, 2023



Contents

- Workflow
- Challenges
- Solutions
- Efficiency
- Technique
- Demo Video
- Summary
- Future Plans

Workflow



Challenges

- Reviewer required to search and go through pages which is labor-intensive.
- Manually identify the component and its details leading to inaccuracies.
- Saving the information by copying and pasting one at a time is time-consuming.

Solutions

- A program to perform automation extraction, allows faster extraction and reduces the time required to process.
- Improve the accuracy by implement Natural Language Processing (NLP) AI model to replace the search and identification process .
- Increase efficiency by saving all the information in one go.

Efficiency

- Comparison on 30 schematic files

| | Manual Extraction | Schematic AI Extractor |
|--|-------------------|------------------------|
| Processing Time (approx. hours) | 5-6 | 1-2 |
| Time Reduced (%) | 60 | |

*extraction on one module

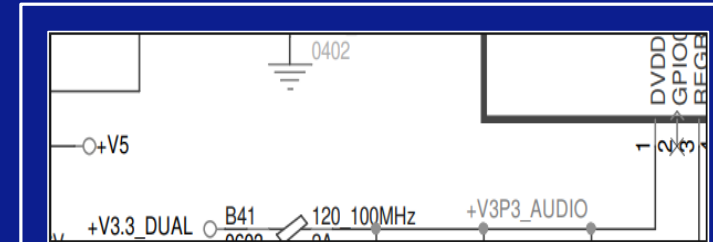
- Spelling error rate reduced as manual extraction for the reviewer might mistake on uppercase 'I' as lowercase 'L'.

Technique

- Working with Python programming language, a sentence-transformers (all-MiniLM-L6-v2) model, PyMuPDF library, and Optical Character Recognition (OCR) engine to perform the extraction.
- The all-MiniLM-L6-v2 model aim is to be used as a sentence and short paragraph encoder. It used the pretrained (nreimers/MiniLM-L6-H384-uncased) model and fine-tuned with 1B sentence pairs dataset.
 - Small model size 80MB.
 - Fast processing speed.
 - Above average performance.

Technique

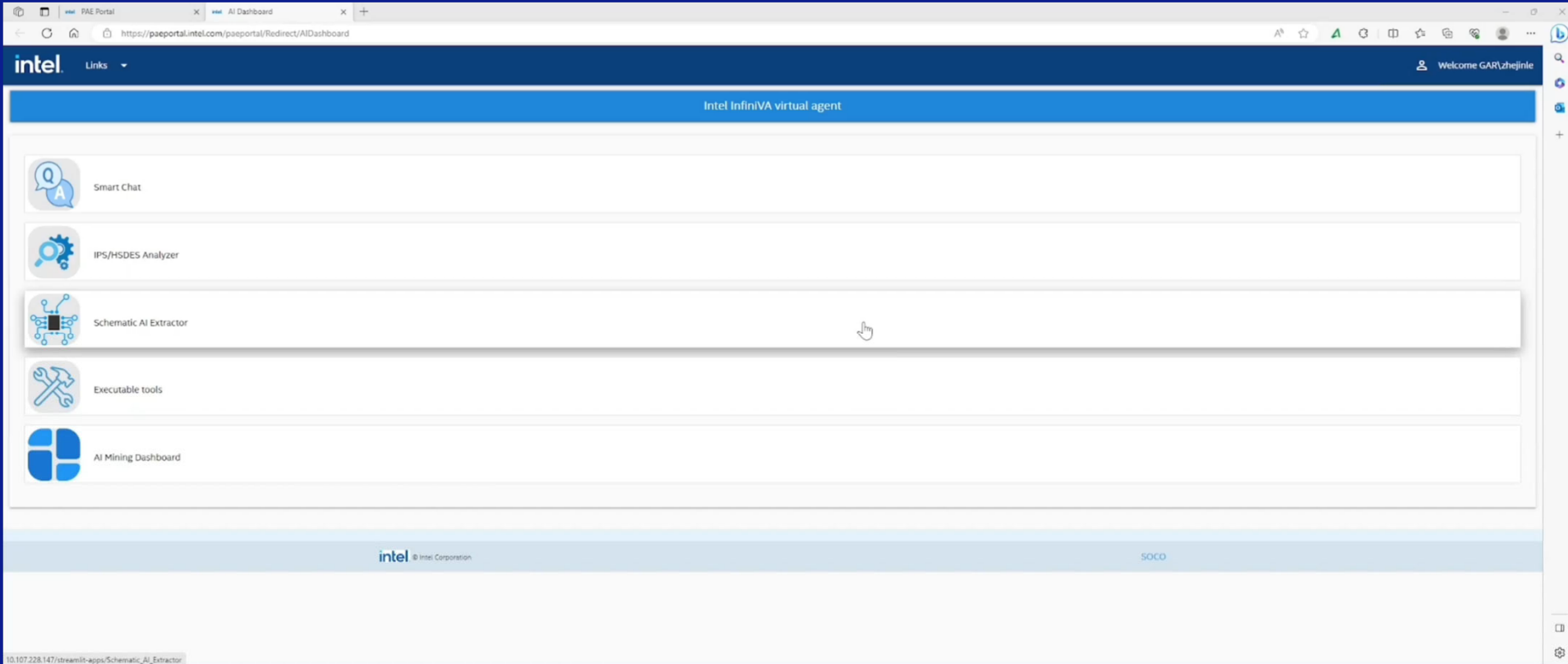
- PyMuPDF loads the schematic file and performs the search process.
- OCR engine recognizing the character patterns from image to extract into text.



```
result.txt - Notepad
File Edit Format View Help
0402
DVDD
PIO
P
O+V5
23
B41
120
100MHz
+V3P3
AUDIO
+V3.3
DUAL
```

The post-processing result extracted from the image to text using the OCR engine

Demo



Summary

- Schematic AI Extractor tool:
 - Reduces manual labor.
 - Improves extraction accuracy and processing time.
 - Other modules still work in progress.
- Current capabilities:
 - Automation extraction.
 - Saving extracted information automatically.
- Data usage:
 - Extracted data stored for future reference.
 - Enhances tool's performance.

Future Plans

- Able to perform automatic extraction with bulk schematic files.
- Add-on other module such as power, display, USB and PCIE.
- Possible to monetize this program to customer.
- Collaboration with other business unit or external customer to enhance this tool.

Presentation / Copyright Notice

- The presentations in this publication comprise the pre-workshop Proceedings of the TestConX Korea workshop. They reflect the authors' opinions and are reproduced here as they are planned to be presented at TestConX Korea. Updates from this version of the papers may occur in the version that is actually presented at TestConX Korea. The inclusion of the papers in this publication does not constitute an endorsement by TestConX or the sponsors.
- There is NO copyright protection claimed by this publication. However, each presentation is the work of the authors and their respective companies: as such, it is strongly encouraged that any use reflect proper acknowledgement to the appropriate source. Any questions regarding the use of any materials presented should be directed to the author/s or their companies.
- The TestConX logo, 'TestConX', and 'TestConX Korea are trademarks of TestConX.