Architekt Integrated Systems:

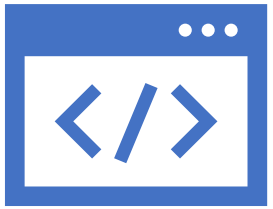
Module

Construction Extractor (MTO – Material Take Off)

Task

- Create 2 separate aggregated tables using the extraction methods, one table using OCR and one table using direct text/annotation.
- The output for each table should be the same
- Extraction areas need to be dynamic
  - Each project's drawing formats may be different, meaning different data extraction locations/areas.
- Implement Concatenate. Some data may need to be combined to create data, such as Drawing Numbers. i.e. "Process(0-78) + Medium(AG) + Sequence(26)" = Drawing (Output: "0-78-AG-26")

OCR vs Direct Text/Annotation Extraction

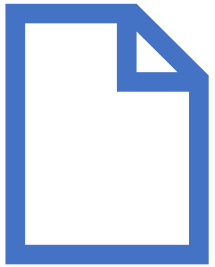


OCR (Optimized Character Recognition):

Extracts as image and attempts to convert image to text.

Less accurate than Direct method

To be used only as a backup if direct text/annotations are not found



Direct Text/Annotation Extraction:

Extracts text/annotations directly using the document's metadata.

- 100% Accurate: No conversion necessary
- Preferred over OCR
- Leverage leftover metadata left behind from software like AutoCAD

Direct Text/Annotation Extraction

See metadata

OCR Repos:

https://github.com/cmccall95/architekt_ocr

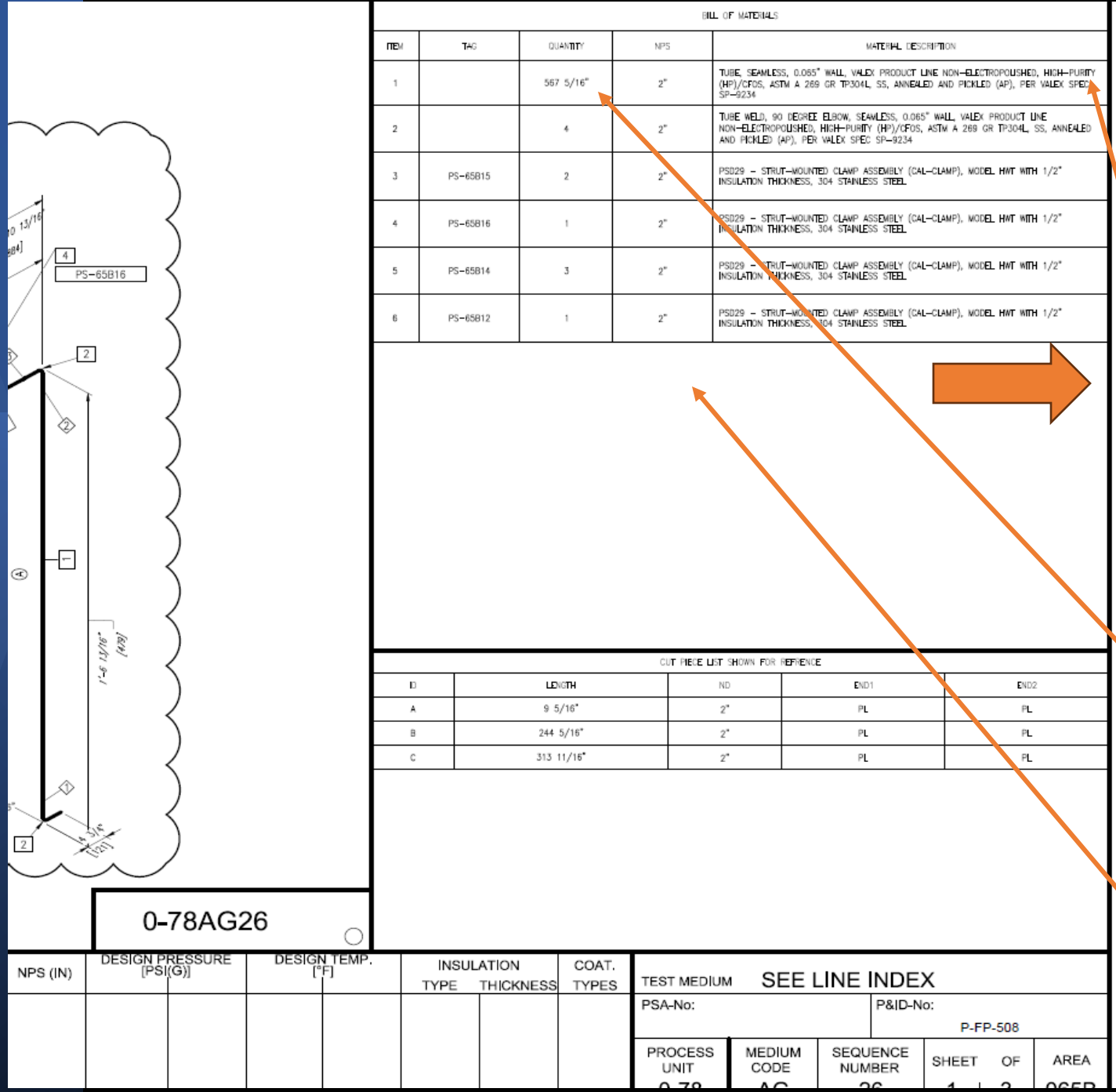
https://github.com/cmccall95/architekt

https://github.com/cmccall95/architekt-iss

Direct Text/Annots Repos:

https://github.com/cmccall95/Architekt-PDF-Extractor-Test

Click here to go to Drawings Folder

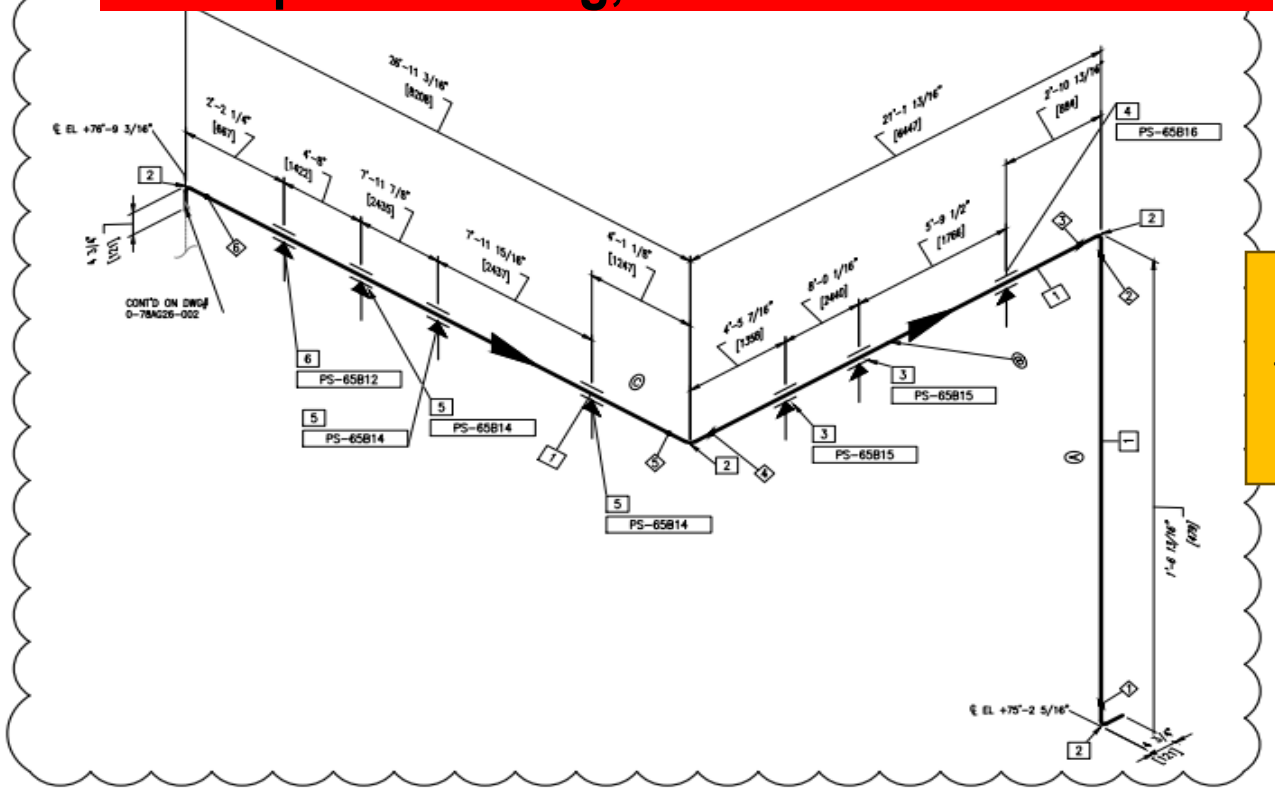


Material Take-Off Basic Table

This will be our basic table template

Drawing Information					BOM (Bill of Materials)				
Drawing #	Sheet	Pipe Spec	P&ID	Area	Item	Tag	Quantity	NPS	Material Description

Select Extraction Areas of first file. For each subsequent drawing, extract the same location



BILL OF MATERIALS

ITEM	TAG	QUANTITY	NPS	MATERIAL DESCRIPTION
1		567 5/16"	2"	TUBE, SEAMLESS, 0.065" WALL, VALEX PRODUCT LINE NON-ELECTROPOLISHED, HIGH-PURITY (HP)/CFOS, ASTM A 269 GR TP304L, SS, ANNEALED AND PICKLED (AP), PER VALEX SPEC SP-9234
2		4	2"	TUBE WELD, 90 DEGREE ELBOW, SEAMLESS, 0.065" WALL, VALEX PRODUCT LINE NON-ELECTROPOLISHED, HIGH-PURITY (HP)/CFOS, ASTM A 269 GR TP304L, SS, ANNEALED AND PICKLED (AP), PER VALEX SPEC SP-9234
3	PS-65B15	2	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL
4	PS-65B16	1	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL
5	PS-65B14	3	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL
6	PS-65B12	1	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL

6.

BOM will almost always be a table format. Need the option to extract and map table columns

CUT PIECE LIST SHOWN FOR REFRENCE

ID	LENGTH	ND	END1	END2
A	9 5/16"	2"	PL	PL
B	244 5/16"	2"	PL	PL
C	313 11/16"	2"	PL	PL

0-78AG26

Dynamically concatenate to create "0-78-AG-26

IFP	2.0	5/4/2022	AJS	SLL	TDW	ISSUED FOR AUTHORITY PERMIT
IFP	1.0	3/15/2022	BB	SLL	TDW	ISSUED FOR AUTHORITY PERMIT
IFQ	0.1	1/14/2021	BCZ			ISSUED FOR QUOTE
STATUS	ISSUE	DATE	DRAWN	CHECKED	APPR.	DESCRIPTION

CERTIFICATE PRE-FABRICATION AND ASSEMBLY (DATE / SIGNATURE)

	1. PRE-FABRICATION			2. ASSEMBLY					
	MATERIAL / FABRICATION	WELDING / FORMING / PWHT	NDE	MATERIAL / FABRICATION	WELDING / FORMING / PWHT	NDE	FLUSHING CLEANING	PRESSURE TEST	MECHANICAL COMPLETION
CONTRACTOR									
LINDE									

PIPE-SPEC	NPS (IN)	DESIGN PRESSURE [PSI(G)]		DESIGN TEMP. [°F]		INS TYPE		DIUM	SEE LINE INDEX
SS150P-AP									

1.

5.

P&ID-No: P-FP-508 2.

PROCESS UNIT	MEDIUM CODE	SEQUENCE NUMBER	SHEET	OF	AREA
0-78	AG	26	1	3	065B

GENERAL REMARKS

1) FIELD WELDS AND CUTTING LENGTH TO BE DETERMINED BY BUYER / CONTRACTOR / FABRICATOR.
2) ALLOWANCE FOR FABRICATION TOLERANCES ARE NOT CONSIDERED.
3) THE GASKET THICKNESS IS NOT SPECIFIED BUT IT IS INCLUDED IN THE TOTAL DIMENSION.
4) SEE ISOMETRIC PIPE SPOOL DETAIL SHEET FOR VCR PART NO.'S & OTHER DETAILS.

PROJECT #	RECEIVED	EXCEL
81062/81063	7/20/2022	

1410A642 / EXL00004 Transmittal-Letter (CUP Revision package 7/12/2022)

LINDE AMERICAS

3. 4.

Document Title: PIPING ISOMETRIC

Linde Doc. No.: 065BL-ZL (0-78) 0-78AG26

CODE: PHOENIX TSMC

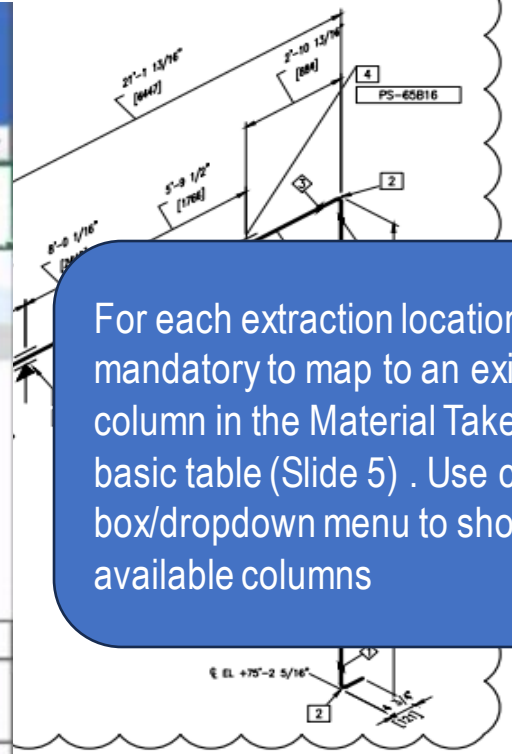
LINDE PROJECT No.: 1410A642

Select Extraction Areas of first file. For each subsequent drawing, extract the same location

Markup #	Map To Column
1	Pipe Spec
2	Drawing #
3	Sheet
4	Pipe Spec
5	P&ID
6	Area
6	Item
6	Tag
6	Quantity
6	NPS
6	Material Description

Drawing #
Sheet
Pipe Spec
P&ID
Area
Item
Tag
Quantity
NPS
Material Description

For each extraction location, it is mandatory to map to an existing column in the Material Take-off basic table (Slide 5) . Use combo box/dropdown menu to show the available columns



BILL OF MATERIALS

ITEM	TAG	QUANTITY	NPS	MATERIAL DESCRIPTION
1		567 5/16"	2"	TUBE, SEAMLESS, 0.065" WALL, VALEX PRODUCT LINE NON-ELECTROPOLISHED, HIGH-PURITY (HP)/CFOS, ASTM A 269 GR TP304L, SS, ANNEALED AND PICKLED (AP), PER VALEX SPEC SP-9234
2		4	2"	TUBE WELD, 90 DEGREE ELBOW, SEAMLESS, 0.065" WALL, VALEX PRODUCT LINE NON-ELECTROPOLISHED, HIGH-PURITY (HP)/CFOS, ASTM A 269 GR TP304L, SS, ANNEALED AND PICKLED (AP), PER VALEX SPEC SP-9234
3	PS-65815	2	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL
4	PS-65816	1	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL
5	PS-65814	3	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL
6	PS-65812	1	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL

6.

BOM will almost always be a table format. Need the option to extract and map table columns

CUT PIECE LIST SHOWN FOR REFRENCE

ID	LENGTH	ND	END1	END2
A	9 5/16"	2"	PL	PL
B	244 5/16"	2"	PL	PL
C	313 11/16"	2"	PL	PL

0-78AG26

Dynamically concatenate to create "0-78-AG-26

PIPE-SPEC	NPS (IN)	DESIGN PRESSURE [PSI(G)]	DESIGN TEMP. [°F]	TYPE	DIUM
SS150P-AP					SEE LINE INDEX

1.

P&ID-No: P-FP-508 2.

PROCESS UNIT	MEDIUM CODE	SEQUENCE NUMBER	SHEET	OF	AREA
0-78	AG	26	1	3	065B

5.

GENERAL REMARKS
1) FIELD WELDS AND CUTTING LENGTH TO BE DETERMINED BY BUYER / CONTRACTOR / FABRICATOR.
2) ALLOWANCE FOR FABRICATION TOLERANCES ARE NOT CONSIDERED.
3) THE GASKET THICKNESS IS NOT SPECIFIED BUT IT IS INCLUDED IN THE TOTAL DIMENSION.
4) SEE ISOMETRIC PIPE SPOOL DETAIL SHEET FOR VCR PART NO.'S & OTHER DETAILS.

PROJECT #	RECEIVED
81062/81063	7/20/2022

EXCEL
1410A642 / EXL00004 Transmittal-Letter (CUP Revision package 7/12/2022)

LINDE AMERICAS

4. 3.

Document Title: PIPING ISOMETRIC

Linde Doc. No.: 065BL-ZL (0-78) 0-78AG26

CODE: PHOENIX TSMC

LINDE PROJECT No.: 1410A642

	FABRICATION	FABRICATION	CLEANING	TEST	MECHANICAL COMPLETION
CONTRACTOR					
LINDE					

Material Take-Off Expected Output

Drawing Information					BOM (Bill of Materials)				
Drawing #	Sheet	Pipe Spec	P&ID	Area	Item	Tag	Quantity	NPS	Material Description
0-78-AG-26	1	SS150P-AP	P-FP-508	065B	1		567 5/16"	2"	TUBE, SEAMLESS, 0.065" WALL, VALEX PRODUCT LINE NON-ELECTROPOLISHED, HIGH-PURITY (HP)/CFOS, ASTM A 269 GR TP304L, SS, ANNEALED AND PICKLED (AP), PER VALEX SPEC SP-9234
0-78-AG-26	1	SS150P-AP	P-FP-508	065B	2		4	2"	TUBE WELD, 90 DEGREE ELBOW, SEAMLESS, 0.065" WALL, VALEX PRODUCT LINE NON-ELECTROPOLISHED, HIGH-PURITY (HP)/CFOS, ASTM A 269 GR TP304L, SS, ANNEALED AND PICKLED (AP), PER VALEX SPEC SP-9234
0-78-AG-26	1	SS150P-AP	P-FP-508	065B	3	PS-65B15	2	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL
0-78-AG-26	1	SS150P-AP	P-FP-508	065B	4	PS-65B16	1	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL
0-78-AG-26	1	SS150P-AP	P-FP-508	065B	5	PS-65B14	3	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL
0-78-AG-26	1	SS150P-AP	P-FP-508	065B	6	PS-65B12	1	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL

Material Take-Off Expected Output

- Now group the data and combine quantities (Only for BOM Data)
  - Group by ‘Material Description’ and Size(NPS)
  - Sum quantities

Drawing Information					BOM (Bill of Materials)				
Drawing #	Sheet	Pipe Spec	P&ID	Area	Item	Tag	Quantity	NPS	Material Description
0-78-AG-26	1	SS150P-AP	P-FP-508	065B	[Array]		567 5/16"	2"	TUBE, SEAMLESS, 0.065" WALL, VALEX PRODUCT LINE NON-ELECTROPOLISHED, HIGH-PURITY (HP)/CFOS, ASTM A 269 GR TP304L, SS, ANNEALED AND PICKLED (AP), PER VALEX SPEC SP-9234
0-78-AG-26	1	SS150P-AP	P-FP-508	065B	[Array]		4	2"	TUBE WELD, 90 DEGREE ELBOW, SEAMLESS, 0.065" WALL, VALEX PRODUCT LINE NON-ELECTROPOLISHED, HIGH-PURITY (HP)/CFOS, ASTM A 269 GR TP304L, SS, ANNEALED AND PICKLED (AP), PER VALEX SPEC SP-9234
Grouped	1	SS150P-AP	P-FP-508	065B	[Array]	[Array]	7	2"	PSD29 - STRUT-MOUNTED CLAMP ASSEMBLY (CAL-CLAMP), MODEL HWT WITH 1/2" INSULATION THICKNESS, 304 STAINLESS STEEL

Possible To Do Items

- Merge existing OCR methods with direct text extraction code.
- Choose font that is easily distinguishable e.g. “1” and “I”.
- Build Template Table (Dynamic as more columns will need to be added)
- Custom Themes
  - Dark Mode
  - Mac OS Layout
- Model table in a Excel Format
  - Add basic table functions (Sort, filter, find, custom font, fill colors, etc.)
- Add column mapping functionality
- Group raw data based on Material Description & Size
  - Sum quantities
  - After aggregation:
    - format quantity units for each unique item(e.g. in, cm, m, linear feet, etc). This will be a static list and selectable by drop down menu
    - Select material type from list(e.g. Pipe, Valve, 90, 45, etc).
    - Save material user inputs for future automation.
- Create link to extracted data and the original file
- Ability to create relationships between items and external files(PDFs, png, jpg, etc)
- Set up database
- What does it take to add machine learning/AI into this to analyze MTO data from contractor/clients?