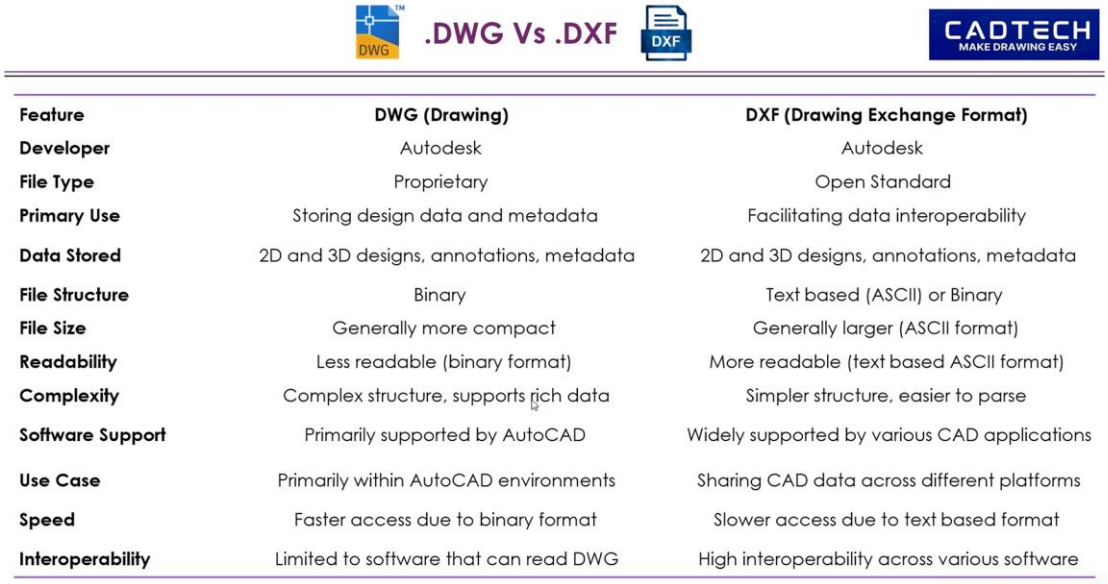
ANALYSIS OF THE DATA EXTRACTED FROM AN ENGINEERING DRAWING

PROBLEM STATEMENT:

CAD Analytics refers to the process of extracting, analyzing, and interpreting data from Computer-Aided Design (CAD) files (like DWG, DXF, or STEP formats). It leverages analytical tools and algorithms to gain insights from CAD drawings and models for better decision-making in engineering, manufacturing, architecture, and other related fields. CAD Analytics bridges the gap between design and data analysis, offering powerful insights to improve efficiency, cost, and accuracy in projects across various industries.



Feature	DWG (Drawing)	DXF (Drawing Exchange Format)
Developer	Autodesk	Autodesk
File Type	Proprietary	Open Standard
Primary Use	Storing design data and metadata	Facilitating data interoperability
Data Stored	2D and 3D designs, annotations, metadata	2D and 3D designs, annotations, metadata
File Structure	Binary	Text based (ASCII) or Binary
File Size	Generally more compact	Generally larger (ASCII format)
Readability	Less readable (binary format)	More readable (text based ASCII format)
Complexity	Complex structure, supports rich data	Simpler structure, easier to parse
Software Support	Primarily supported by AutoCAD	Widely supported by various CAD applications
Use Case	Primarily within AutoCAD environments	Sharing CAD data across different platforms
Speed	Faster access due to binary format	Slower access due to text based format
Interoperability	Limited to software that can read DWG	High interoperability across various software

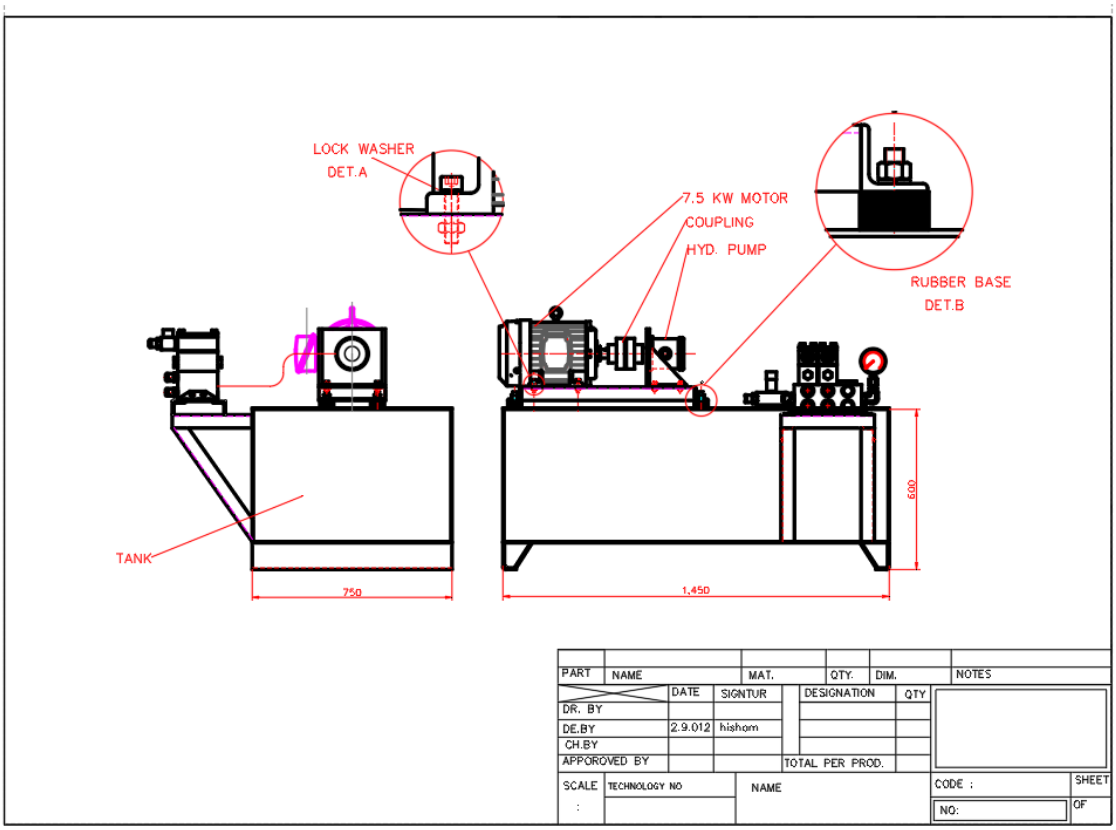
In this project we have predominantly worked with 2D CAD files (.dwg, .dxf and .pdf files) i.e. engineering drawings and tried to follow the methodology given below:

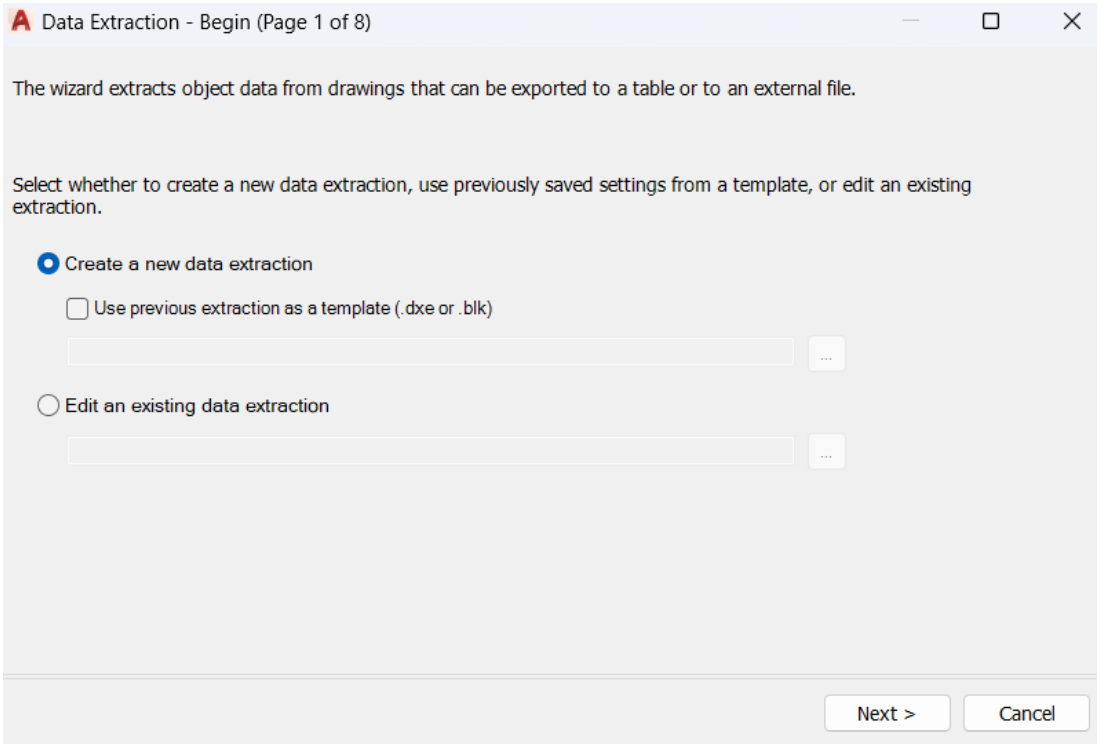
- **Data Extraction**: Extract geometric, annotation, and metadata information from CAD files (DWG, DXF, STEP, etc.).
- **Metadata Analysis**: Analyze layers, dimensions, units, materials, and object properties.
- **Dimensional Analysis**: Measure lengths, areas, volumes, and perform geometric calculations.
- **Bill of Materials (BOM) Extraction**: Extract part lists, quantities, and materials from CAD drawings.

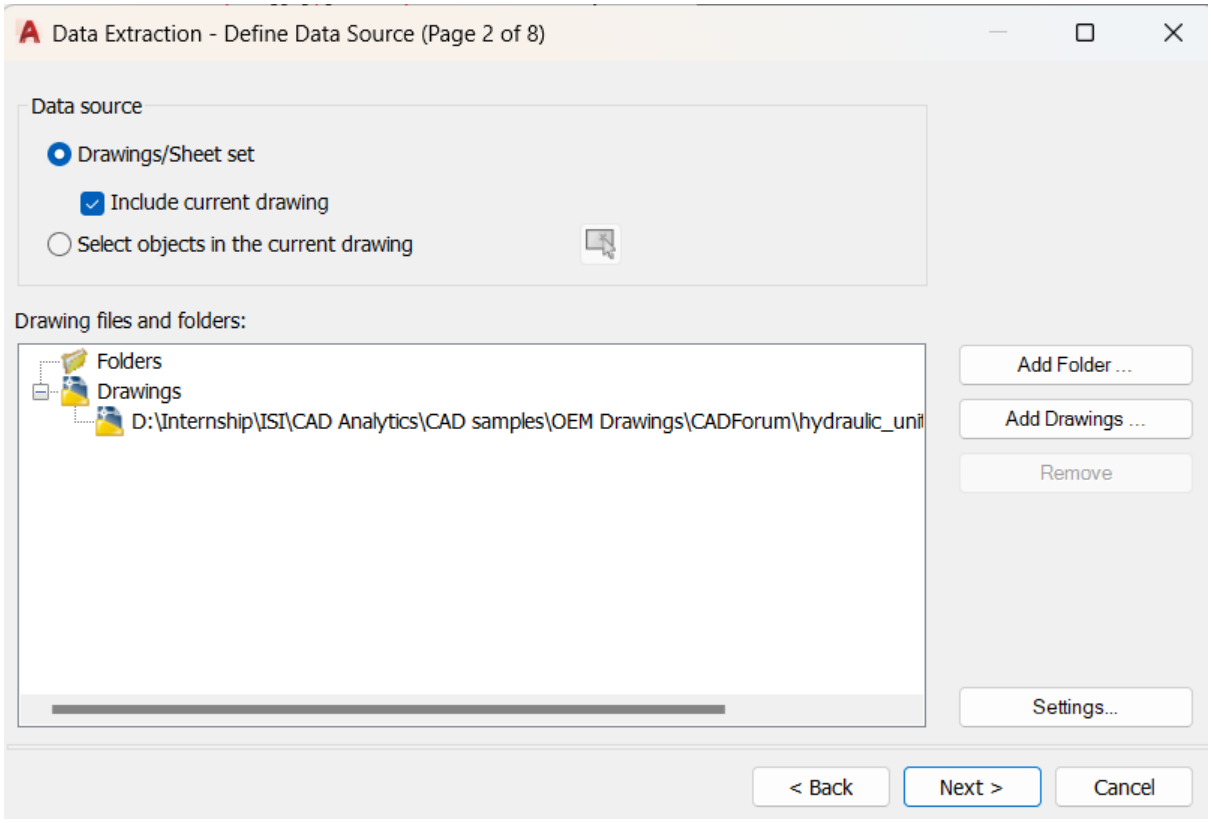
WORK APPROACH:

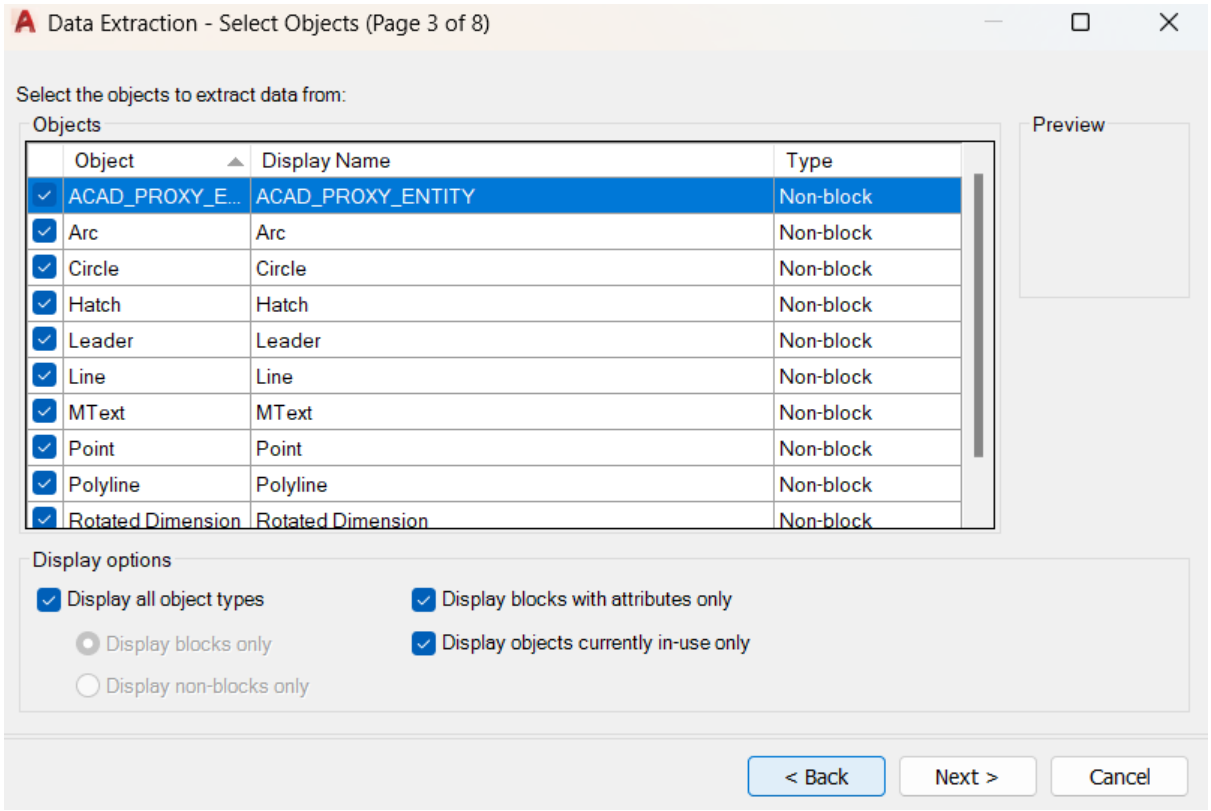
1. AutoCAD Data Extraction

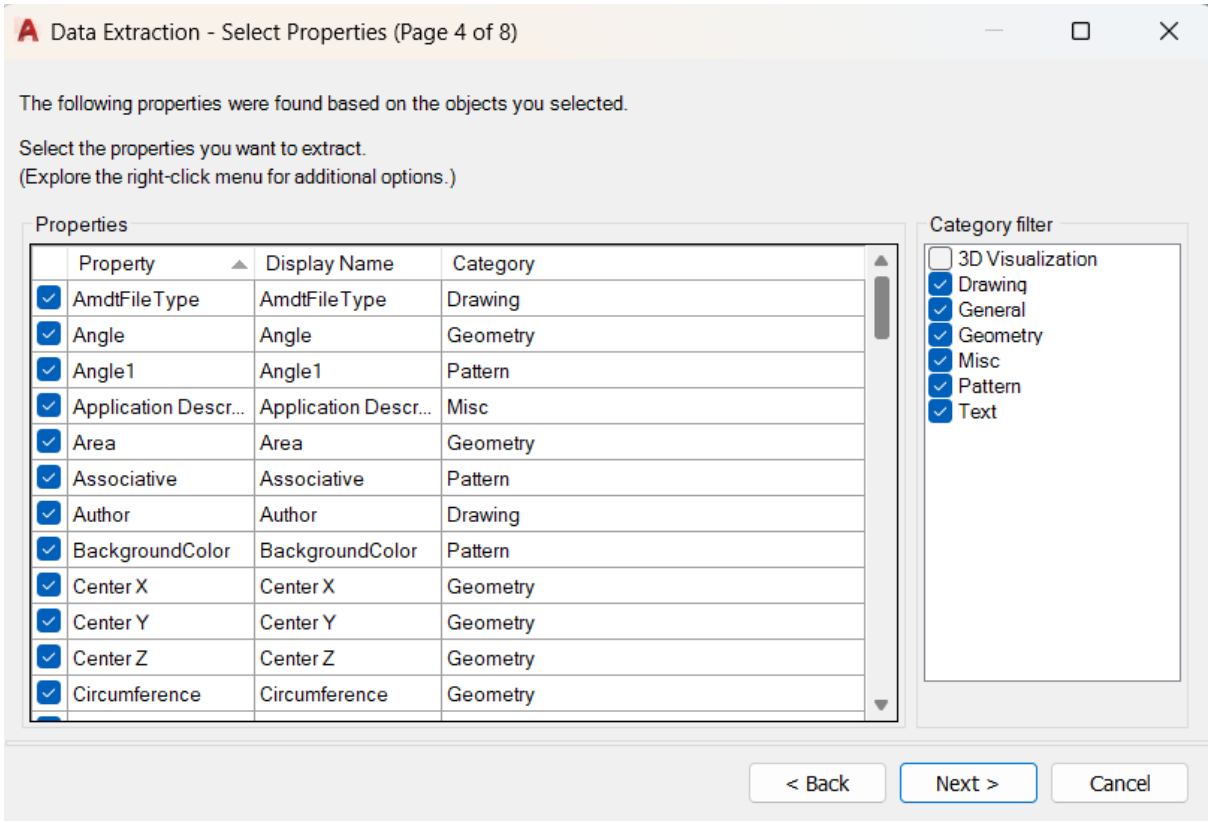
A file named hydraulic_unit.dwg is opened in AutoCAD and data is extracted from the drawing using the Data Extraction Wizard.

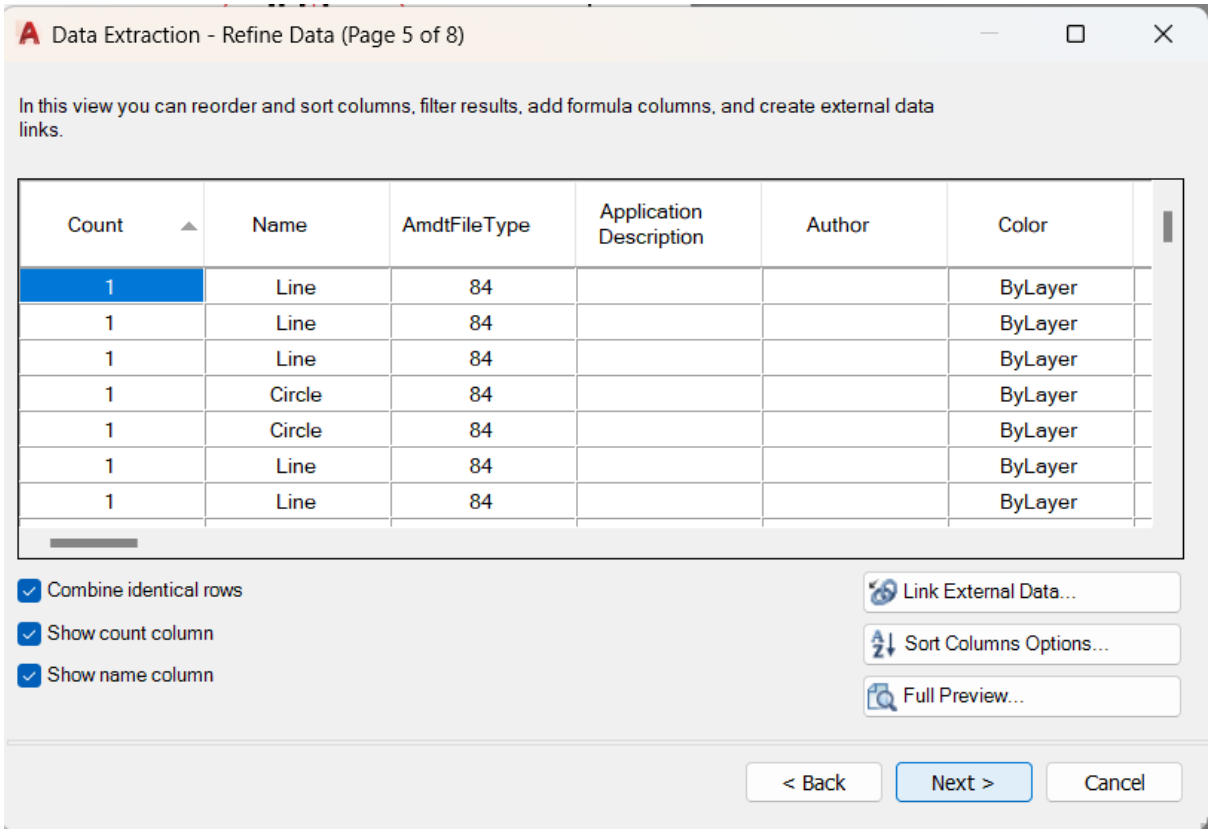





Data Extraction - Define Data Source (Page 2 of 8)
Data source
Drawings/Sheet set
Include current drawing
Select objects in the current drawing
Drawing files and folders:
Folders
Drawings
D:\Internship\ISI\CAD Analytics\CAD samples\OEM Drawings\CADForum\hydraulic_unit
Add Folder ...
Add Drawings ...
Remove
Settings...
< Back
Next >
Cancel


Data Extraction - Select Objects (Page 3 of 8)
Select the objects to extract data from:
Objects
Object
Display Name
Type
ACAD_PROXY_E...
ACAD_PROXY_ENTITY
Non-block
Arc
Arc
Non-block
Circle
Circle
Non-block
Hatch
Hatch
Non-block
Leader
Leader
Non-block
Line
Line
Non-block
MText
MText
Non-block
Point
Point
Non-block
Polyline
Polyline
Non-block
Rotated Dimension
Rotated Dimension
Non-block
Preview
Display options
Display all object types
Display blocks only
Display non-blocks only
Display blocks with attributes only
Display objects currently in-use only
< Back
Next >
Cancel



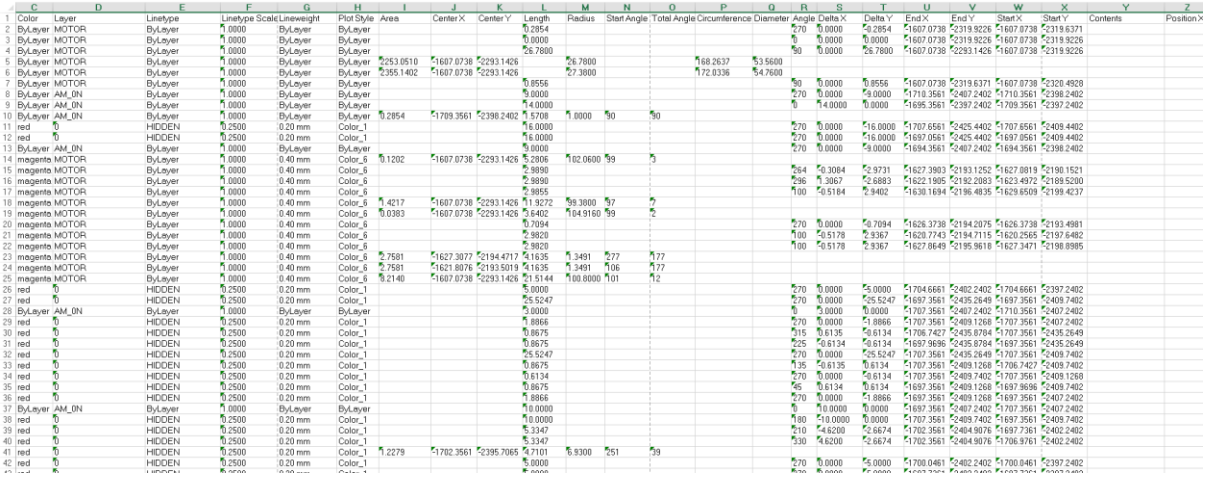


At last export the data as .xls file.

Eliminate unnecessary columns to show only the required attributes.

2. Analysis of the extracted data

After manually comparing the extracted data with the drawing, it has been observed that there has been some unnecessary data. After removing some empty columns, the data still looks complicated.

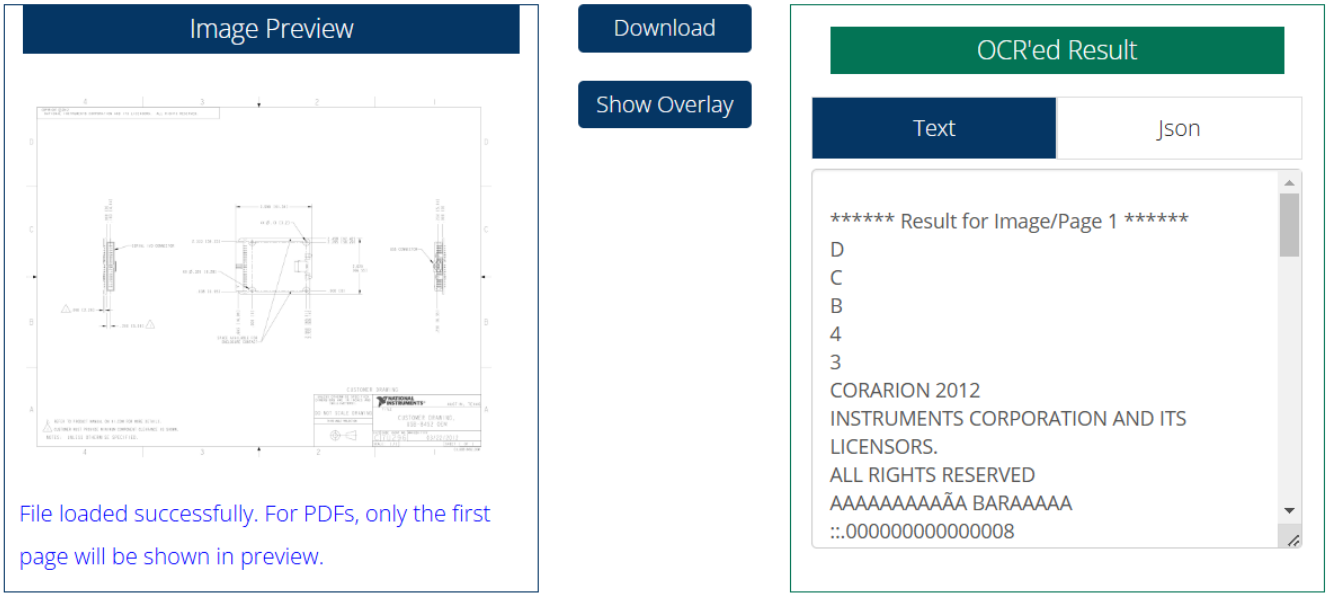


	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	Color	Layer	Linetype	Linetype Scale	Lineweight	Plot Style	Area	Center X	Center Y	Length	Radius	Start Angle	Total Angle	Circumference	Diameter	Angle	Delta X	Delta Y	End X	End Y	Start X	Start Y	Contents	Position X
2	ByLayer	MOTOR	ByLayer	1.0000	ByLayer	ByLayer				0.2854						270	0.0000	-0.2854	1607.0738	2319.9226	1607.0738	2319.6371		
3	ByLayer	MOTOR	ByLayer	1.0000	ByLayer	ByLayer				0.0000						0	0.0000	0.0000	1607.0738	2319.9226	1607.0738	2319.9226		
4	ByLayer	MOTOR	ByLayer	1.0000	ByLayer	ByLayer				26.7800						90	0.0000	26.7800	1607.0738	2293.1426	1607.0738	2319.9226		
5	ByLayer	MOTOR	ByLayer	1.0000	ByLayer	ByLayer	253.0510	1607.0738	2293.1426		26.7800			168.2637	53.5600									
6	ByLayer	MOTOR	ByLayer	1.0000	ByLayer	ByLayer	1355.1402	1607.0738	2293.1426		27.3800			172.0336	54.7600									
7	ByLayer	MOTOR	ByLayer	1.0000	ByLayer	ByLayer				0.8556						90	0.0000	0.8556	1607.0738	2319.6371	1607.0738	2320.4928		
8	ByLayer	AM_0N	ByLayer	1.0000	ByLayer	ByLayer				0.0000						270	0.0000	0.0000	1710.3561	2407.2402	1710.3561	2398.2402		
9	ByLayer	AM_0N	ByLayer	1.0000	ByLayer	ByLayer				14.0000						0	14.0000	0.0000	1695.3561	2397.2402	1709.3561	2397.2402		
10	ByLayer	AM_0N	ByLayer	1.0000	ByLayer	ByLayer	0.2854	1709.3561	2398.2402	1.5708	1.0000	90	90											
11	red	0	HIDDEN	0.2500	0.20 mm	Color_1				16.0000						270	0.0000	16.0000	1707.6561	2425.4402	1707.6561	2409.4402		
12	red	0	HIDDEN	0.2500	0.20 mm	Color_1				16.0000						270	0.0000	16.0000	1697.0561	2425.4402	1697.0561	2409.4402		
13	ByLayer	AM_0N	ByLayer	1.0000	ByLayer	ByLayer				0.0000						270	0.0000	0.0000	1694.3561	2407.2402	1694.3561	2398.2402		
14	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6	0.1202	1607.0738	2293.1426	5.2806	102.0600	99	3											
15	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6				2.9890						264	-0.3084	-2.9731	1627.3903	2193.1252	1627.8819	2190.1521		
16	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6				2.9890						296	1.3067	-2.6883	1622.1905	2192.2083	1623.4972	2189.5200		
17	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6				2.9855						100	-0.5184	2.9402	1630.1694	2196.4835	1629.6509	2199.4237		
18	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6	4217	1607.0738	2293.1426	11.9272	99.3800	97	7											
19	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6	0.0383	1607.0738	2293.1426	3.6402	104.9160	99	2											
20	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6				3.7094						270	0.0000	-3.7094	1626.3738	2194.2075	1626.3738	2193.4981		
21	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6				2.9820						100	-0.5178	2.9367	1628.7743	2194.7115	1620.2565	2197.6482		
22	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6				2.9820						100	-0.5178	2.9367	1627.8649	2195.9618	1627.3471	2198.8985		
23	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6	2.7581	1627.3077	2194.4717	4.1635	3.3491	277	77											
24	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6	2.7581	1621.8076	2193.5019	4.1635	3.3491	106	77											
25	magenta	MOTOR	ByLayer	1.0000	0.40 mm	Color_6	2.2140	1607.0738	2293.1426	51.5144	100.0000	101	12											
26	red	0	HIDDEN	0.2500	0.20 mm	Color_1				5.0000						270	0.0000	-5.0000	1704.6661	2402.2402	1704.6661	2397.2402		
27	red	0	HIDDEN	0.2500	0.20 mm	Color_1				25.5247						270	0.0000	-25.5247	1697.3561	2435.2649	1697.3561	2409.7402		
28	ByLayer	AM_0N	ByLayer	1.0000	ByLayer	ByLayer				5.0000						0	5.0000	0.0000	1707.3561	2407.2402	1710.3561	2407.2402		
29	red	0	HIDDEN	0.2500	0.20 mm	Color_1				1.8866						270	0.0000	-1.8866	1707.3561	2409.1268	1707.3561	2407.2402		
30	red	0	HIDDEN	0.2500	0.20 mm	Color_1				8.8675						315	-6.2695	-6.2695	1706.7427	2435.8784	1707.3561	2435.2649		
31	red	0	HIDDEN	0.2500	0.20 mm	Color_1				8.8675						225	-6.2695	6.2695	1697.9696	2435.8784	1697.3561	2435.2649		
32	red	0	HIDDEN	0.2500	0.20 mm	Color_1				25.5247						270	0.0000	-25.5247	1707.3561	2435.2649	1707.3561	2409.7402		
33	red	0	HIDDEN	0.2500	0.20 mm	Color_1				8.8675						135	-6.2695	-6.2695	1707.3561	2409.1268	1706.7427	2409.7402		
34	red	0	HIDDEN	0.2500	0.20 mm	Color_1				6.8134						270	0.0000	-6.8134	1707.3561	2409.7402	1707.3561	2409.1268		
35	red	0	HIDDEN	0.2500	0.20 mm	Color_1				8.8675						45	6.2695	6.2695	1697.3561	2409.1268	1697.9696	2409.7402		
36	red	0	HIDDEN	0.2500	0.20 mm	Color_1				1.8866						270	0.0000	-1.8866	1697.3561	2409.1268	1697.3561	2407.2402		
37	ByLayer	AM_0N	ByLayer	1.0000	ByLayer	ByLayer				10.0000						0	10.0000	0.0000	1697.3561	2407.2402	1707.3561	2407.2402		
38	red	0	HIDDEN	0.2500	0.20 mm	Color_1				10.0000						180	-10.0000	0.0000	1707.3561	2409.7402	1697.3561	2409.7402		
39	red	0	HIDDEN	0.2500	0.20 mm	Color_1				5.3347						210	-4.6200	-2.6674	1702.3561	2404.9076	1697.7361	2402.2402		
40	red	0	HIDDEN	0.2500	0.20 mm	Color_1				5.3347						330	4.6200	-2.6674	1702.3561	2404.9076	1706.9761	2402.2402		
41	red	0	HIDDEN	0.2500	0.20 mm	Color_1	2279	1702.3561	2395.7065	4.7101	6.9300	251	59											
42	red	0	HIDDEN	0.2500	0.20 mm	Color_1				5.0000						270	0.0000	-5.0000	1700.0461	2402.2402	1700.0461	2397.2402		

At the end, we will get to know about the coordinates and the contents of the drawing attributes like Line, Arc, MText, etc properly.

Difficulty faced: To recognize the text contents (MText) in such a big collection of data

Possible approach: Websites like https://ocr.space/ has been used to extract the text data from the drawing pdf which might be easily readable and user friendly.



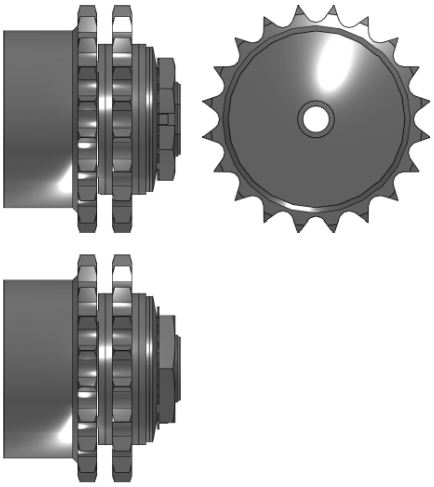
3. Emphasis on the technical aspects

Difficulty faced: Important technical information about the drawing is missing

Possible approach: Refer to OEM catalogues in https://www.3dfindit.com/en/ for better understanding of the technical data in the OEM drawings

Let us take an example of a Slip Clutch Hub

2D derivation

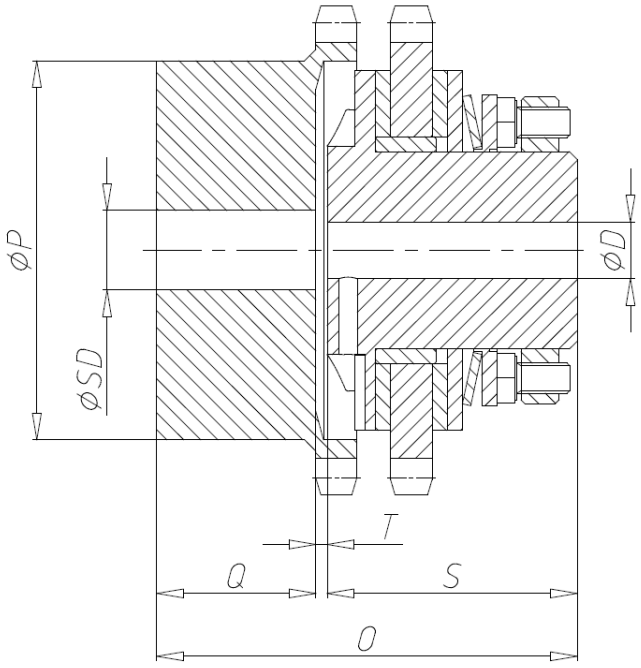


ISO view

Bill-Of-Materials

N°	Description	Amount
1	Hatorq-slip hub couplings KETTEN FUCHS Artikel-Nr. C 240/85	1

Technical drawings



Technical data

CNSORDERNO (Order number)	C 240/85
MTA (Size Mt × A)	240 × 85
MT (Transmission values max torque Mt. / Nm)	240
AT (Number plate spring)	2
DP (Parallel offset Δp)	0.35
AW (Axial offset αw / Grad)	30'
SD (Bore sprocket hub / mm)	21.0
SDB (sd min. / mm)	21.0
SDMAX (sd max. / mm)	65.0
D (Wheel hub bore chains / mm)	15.0
DB (D min / mm)	15.0
DMAX (D max / mm)	25.0
O (Dimensions / mm)	100.0
P (Dimensions / mm)	100.0
Q (Dimensions / mm)	42.0
S (Dimensions / mm)	55.0
T (Dimensions / mm)	3.0
Z (Number of Teeth)	20
KG (Chain size)	3/4 × 7/16"

The above table clearly shows the technical data about the OEM in an organized manner. Hence a similar approach has been undertaken for the AutoCAD data extraction of the above drawing file and the results are below.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	AT	AW	CATALOG	CNSORDERNO	Comments	COMPANY	D	DB	DMAX	DP	KG	Layer	MT	MTA	O	P	Q	S	SD	SDB	SDMAX	T	Z
2	2	30?	Ketten Fuchs	C 240/85	File was created by Cadenas	ketten_fuchs	15.0	15.0	25.0	0.35	3/4 × 7/16"	AM_6	240	240 × 85	100.0	100.0	42.0	55.0	21.0	21.0	65.0	3.0	20
3																							

After the thorough analysis of the CAD data, we arrive at the conclusion that predominantly we need to extract 2 types of data – Mtext (consists of Author, Drawing no., Scale, Date, Size, Organization Name, Units, etc which are present in the bottom right corner generally) and Technical data as shown in the above table.

Therefore the final solution is to develop a tool such that it can extract both these data simultaneously in an organized and user-friendly manner. The tool should support either .dwg, .dxf or .pdf files.

WORK PRODUCTS AND DELIVERABLES:

Software used for Data Extraction: AutoCAD 2019

IDE: Visual Studio Code

.DWG to .DXF Converter: TeighaFileConverter 4.3.2

Documentation: 1. https://ezdxf.readthedocs.io/en/stable/

2. https://pypi.org/project/ezdxf/

3. https://docs.streamlit.io/

Websites of Original Equipment Manufacturers (OEMs):

1. https://siemens.partcommunity.com/3d-cad-models/?cwid=0433
2. https://www.3dfindit.com/en/cad-bim-library/manufacturer/eatons-cad?path=eaton_cad
3. https://www.3dfindit.com/en/cad-bim-library/manufacturer/abb?path=abb
4. https://www.hermanmiller.com/resources/3d-models-and-planning-tools/planning-ideas/
5. https://www.kohler.com/en/for-professionals/technical-specifications
6. https://www.mitsubishielectric.com/fa/download/cad/
7. https://www.se.com/us/en/download/doc-group-type/120245927022-CAD,+Drawings,+&+Curves/
8. https://www.ni.com/en/support/documentation/dimensional-drawings.html
9. https://oem-bike-parts.com/en/
10. https://www.3dcontentcentral.com/parts/supplier/Saint-Gobain-Performance-Plastics.aspx

Websites of CAD (Computer Aided Design) Communities:

1. https://www.cadforum.cz/catalog_en/
2. https://www.3dfindit.com/en/cad-bim-library/
3. https://www.3dcontentcentral.com/

Websites used for Metadata Extraction:

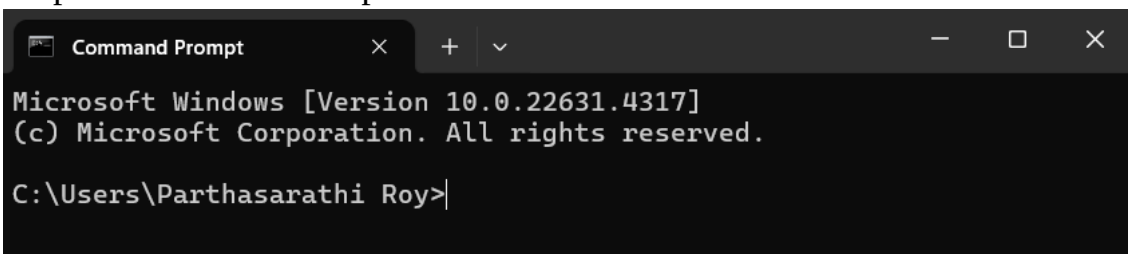
1. https://products.aspose.app/cad/metadata
2. https://products.groupdocs.app/metadata/total
3. https://ocr.space/

Github Repository Link:

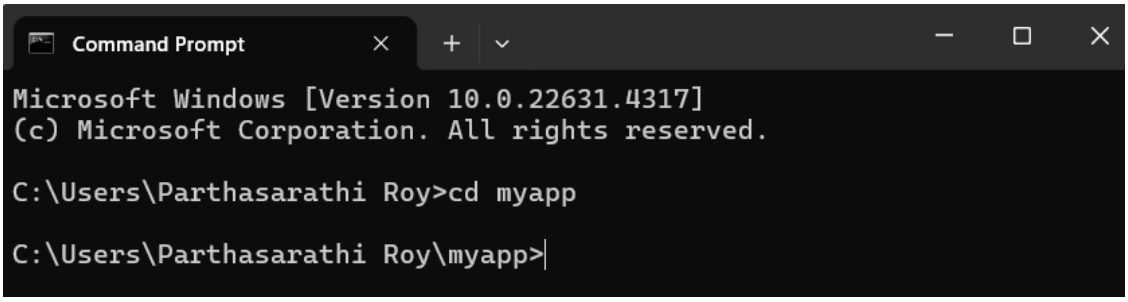
https://github.com/coding-with-parthasarathi/CAD-Analytics

USER MANUAL:

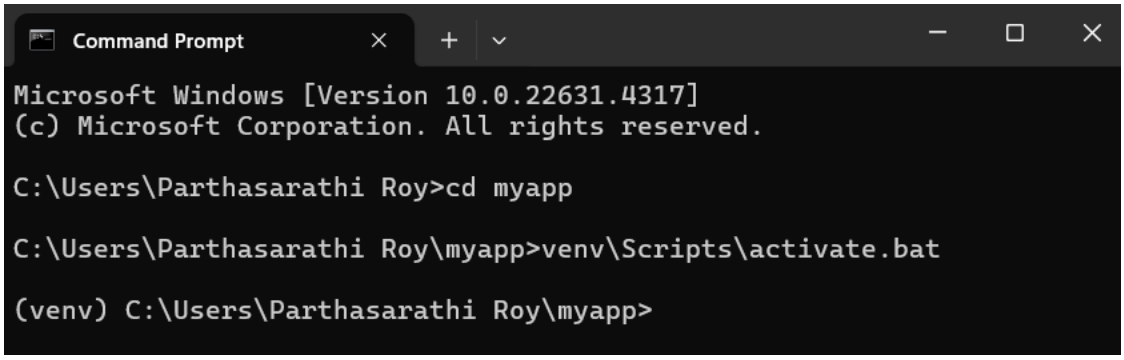
1. Open Command Prompt.



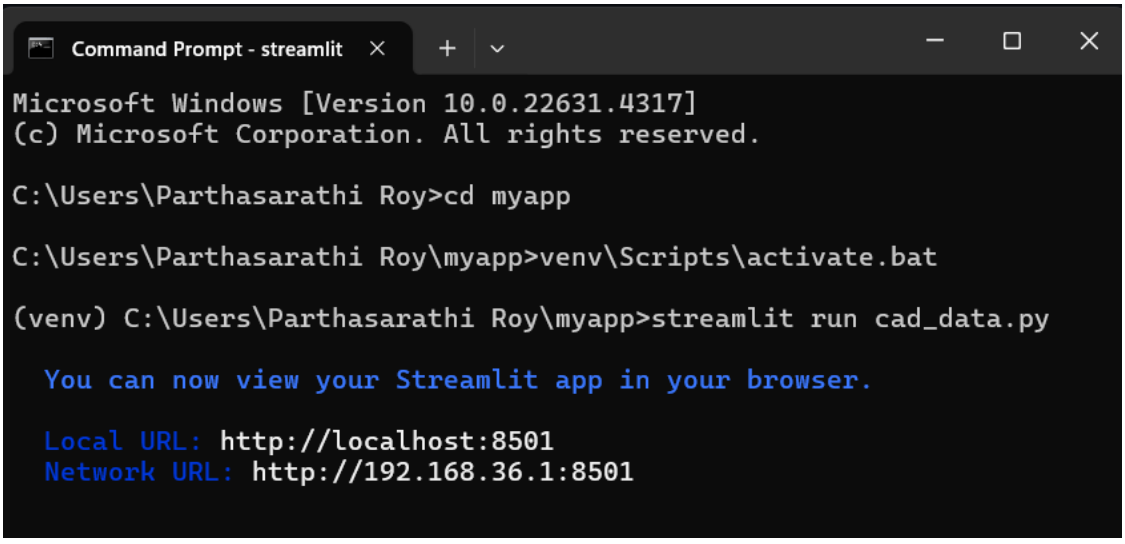
2. Navigate to the project directory by typing **cd myapp**.



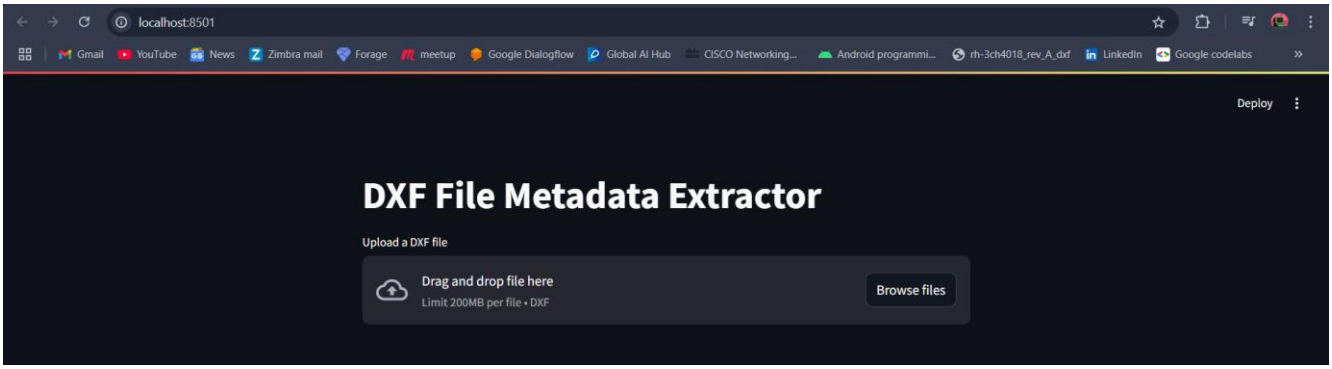
3. Activate the python virtual environment (venv) by typing **venv\Scripts\activate.bat**.



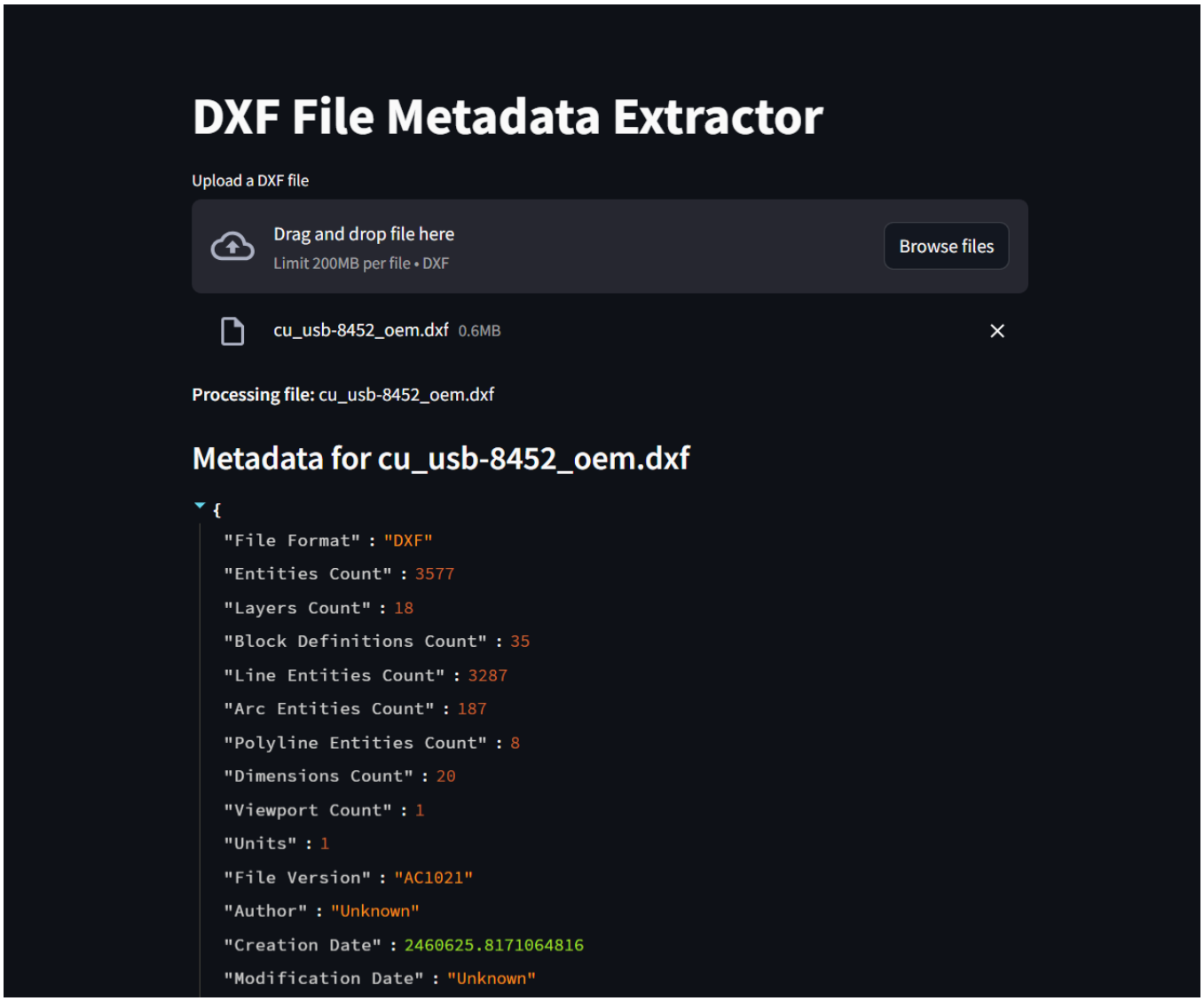
4. Run the **cad_data.py** file by typing **streamlit run cad_data.py**.



5. The Streamlit app will automatically open in the browser or you can follow the above URLs to open the app.



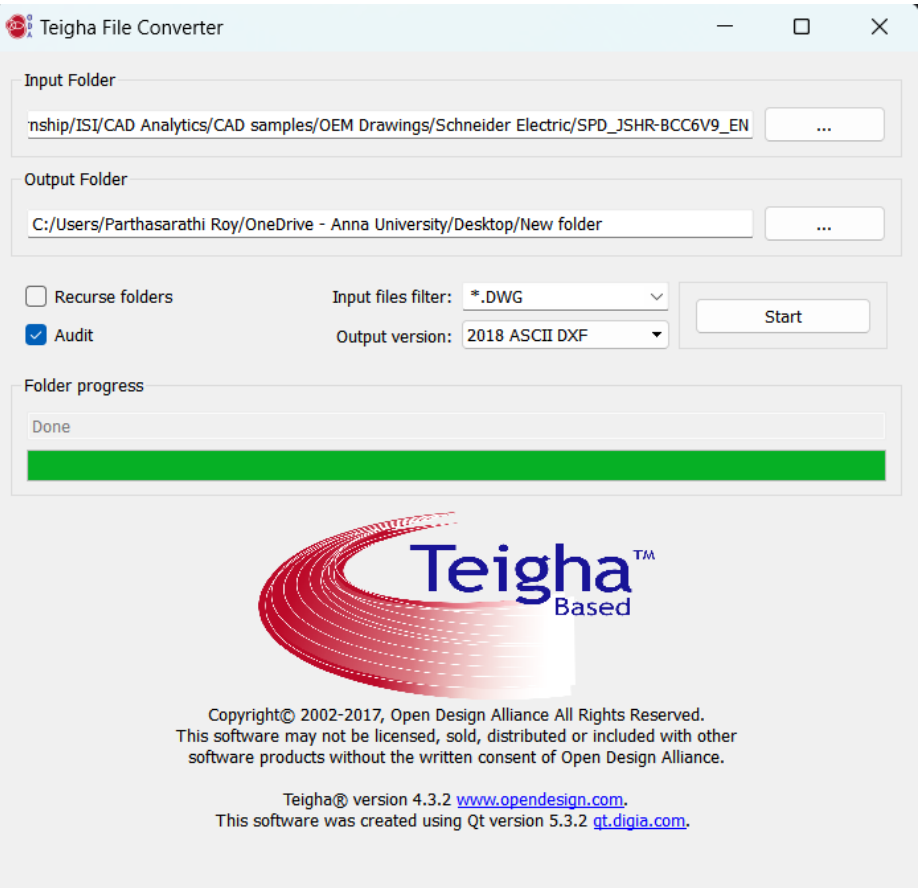
6. Next we have to upload a DXF file by clicking on Browse Files and the data extracted from the file will be shown below.



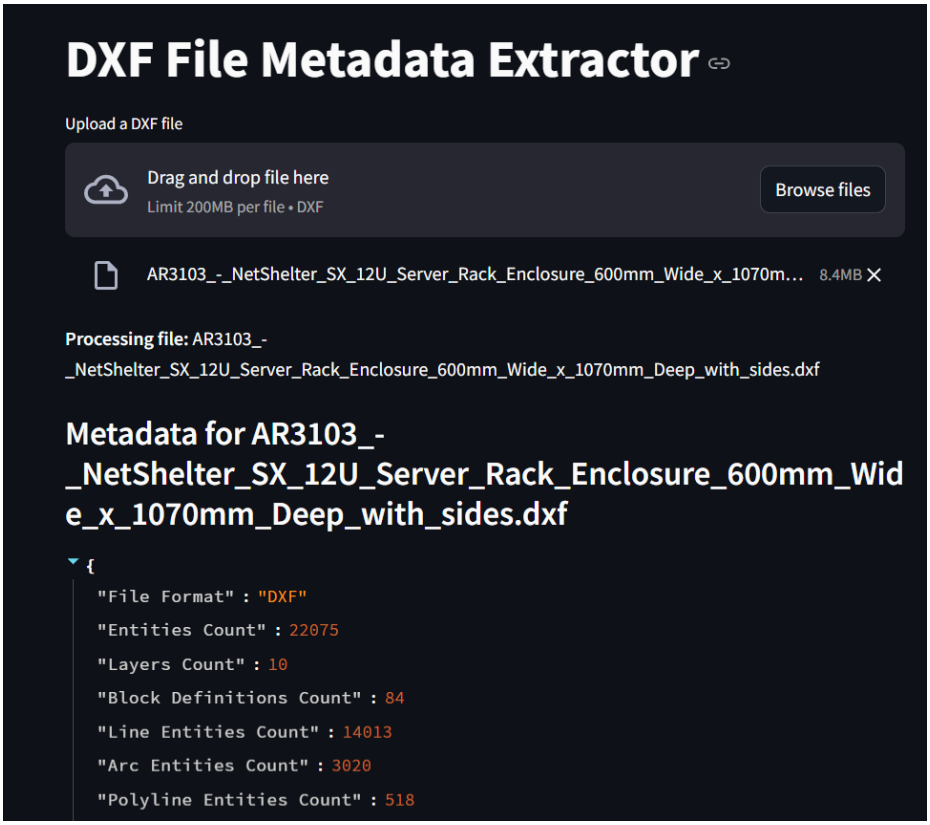
```
"Texts detected" :
"A
4
B
3
C
D
4
3
7U296
CODE IDENT NO.
2
DO NOT SCALE DRAWING
[MILLIMETERS].
DIMENSIONS ARE IN INCHES AND
UNLESS OTHERWISE SPECIFIED
SIZE
C
SCALE:
TITLE
1
MM/DD/YYYY
SHEET    OF
AUSTIN, TEXAS
A
2
C
1
D
R
THIRD ANGLE PROJECTION
B
CUSTOMER DRAWING
NATIONAL INSTRUMENTS CORPORATION AND ITS LICENSORS.  ALL RIGHTS RESERVED.
c
SPACE AVAILABLE FOR\PENCLOSURE CONTACT
USB CONNECTOR
SERIAL I/O CONNECTOR"
}
```


Note: If we need to extract data from a DWG file, it needs to be converted first to DXF file. The steps to do the same are as follows:

1. Open TeighaFileConverter 4.3.2.
2. Select the input and the output folders.
3. Select the input file filter as *.DWG.
4. Select the output version as 2018 ASCII DXF.
5. Check the Audit checkbox.
6. Click on Start to start the conversion.
7. The converted DXF file will be saved in the output folder.



The converted DXF file is then uploaded in the DXF File Metadata Extractor App and the necessary data is fetched as shown below.



```
"Dimensions Count" : 78
"Viewport Count" : 1
"Units" : 1
"File Version" : "AC1032"
"Author" : "SESA85085"
"Creation Date" : 2460625.9102199073
"Modification Date" : 2460625.907530775
"Texts detected" :
"{\LMAIN MECHANICAL-2}
THIS PANEL IS \PCONSTRUCTED\PWITH PERFORATED\PMATERIAL
{\H0.8689x;\LFRONT}
{\LLEFT SIDE VIEW}
{\LFRONT VIEW}
{\LREAR VIEW}
B
B
{\LMAIN MECHANICAL-1}
{\LMAIN MECHANICAL-2}
{\H0.6383x;\LSECTION B-B\P\H0.62994x;STANDARD POSITIONS FOR\PMOUNTING RAILS
AND\PTHEIR ADJUSTABILITY\P(LOOKING DOWN)}
{\H0.8144x;\LTOP VIEW \P\H0.6299x;CABLE OPENINGS\P(COVERS REMOVED)}
{\LTOP VIEW\P\H0.8202x;(DOORS FULLY OPEN)\l}
{\H0.8689x;\LFRONT}
171 [6.73] x 51 [2.02]\POPENING\PBRUSHES ARE INCLUDED\PAT OPENING.\P{\A1;FOR
COMMUNICATION/\PPOWER CABLES}
{\H0.5609x;\LDETAIL-A\P\H0.8571x;(RMPDU MOUNTING)}
45 X 30\P[1.77 X 1.18]\PCUTOUT
FRONT
REFER TO \PDETAIL-A\P(RMPDU MOUNTING)
(REFER TO\PSHEET-2)
THIS PANEL IS \PCONSTRUCTED\P WITH PERFORATED\P MATERIAL
{\H0.8689x;\LFRONT}
4
\pxql;PALLET BRACKET\P\pq*;(2-PLACES - 1 FRONT/1 BACK)
{\H0.8689x;\LBOTTOM VIEW - \H0.99683x;TOP ROOF NOT SHOWN\H1.0032x;\P\H0.7052x;
(WITH PALLET BRACKETS, CAN BE USED FOR ANCHORING\l)}
\pxqr;CASTER\P(4-PLACES)
VERTICAL\PMOUNTING\PFLANGE
ACCESSORY\PCHANNEL
{\H0.8689x;\LFRONT}
{\H0.78741x;\LBOTTOM VIEW \H0.99999x;-TOP ROOF NOT SHOWN\P\H0.77819x;(WITH
OPTIONAL BOLT-DOWN BRACKETS FOR ANCHORING)}
5
BOLT-DOWN BRACKET\P (4-PLACES - 2 FRONT/2 BACK)
\pxql;ADJUSTABLE\PLEVELING FEET\P(4-PLACES)
\pxql;9.5 [0.38]\PSQUARE\P9 PLACES
120 X 61\P[4.72 X 2.40]\PCUTOUT
{\H0.8689x;\LFRONT}"
}
```


Note: The data extraction might not be accurate but this app is able to extract some amount of data from the uploaded DXF file.

TECHNICAL MANUAL:

Datasets: 1. https://github.com/coding-with-parthasarathi/CAD-Analytics/tree/main/OEM%20Drawings

2. https://github.com/coding-with-parthasarathi/CAD-Analytics/tree/main/Extracted%20Metadata

Python libraries: Streamlit, ezdxf

Program documentation:

Program name: cad_data.py

Function Name	Purpose
def extract_metadata_from_dxf(file):	Extract metadata from a DXF file using ezdxf
def process_dxf(file):	To check if it is a .dxf file

Source code: https://github.com/coding-with-parthasarathi/CAD-Analytics/blob/main/cad_data.py

Prerequisites:

As with any programming tool, in order to install Streamlit you first need to make sure your computer is properly set up. More specifically, you'll need:

1. Python (version 3.8 to 3.12.)
2. A Python environment manager (recommended)

   Environment managers create virtual environments to isolate Python package installations between projects.

   For this guide, we'll be using **venv**, which comes with Python.
3. A Python package manager

   Package managers handle installing each of your Python packages, including Streamlit.

   For this guide, we'll be using **pip**, which comes with Python.
4. Only on MacOS: Xcode command line tools

   Download Xcode command line tools using these instructions in order to let the package manager install some of Streamlit's dependencies.

5. A code editor

   Most preferable editor is VS Code.

Installation and configuration guide:

1. Open Command Prompt and clone the given Github repository:
   `git clone https://github.com/coding-with-parthasarathi/CAD-Analytics.git`
2. Navigate to the project directory:
   `cd CAD-Analytics`
3. In the terminal type:
   `python -m venv .venv`
4. A folder named "venv" will appear in your project. This directory is where your virtual environment and its dependencies are installed.
5. Activate the virtual environment:
   `venv\Scripts\activate.bat`
6. Once activated, you will see your environment name in parentheses before your prompt. "(venv)"
7. Install streamlit:
   `pip install streamlit`
8. To deactivate the virtual environment type:
   `deactivate`
9. Install ezdxf which is a Python package to create new DXF files and read/modify/write existing DXF documents:
   `pip install ezdxf`
10. Repeat step 5.
11. Run the cad_data.py file by typing:
    `streamlit run cad_data.py`
12. Carry out the rest of the work in the Streamlit app by referring to the User Manual.