COURSE TITLE: **CEF 506 - PYTHON/PERL PROGRAMMING DEVELOPMENT**

NAME	MATRICULATION NUMBER
FRU ISIDORE CHE	FE12A078
THEOPHILUS WABA NASALI	FE12A183
NKENG NEWTON	FE12A140
MOBA MELVIS RINGNYU	FE12A107
ENOMBE THIERRY EWANE	FE12A053
ALANGI DERRICK	FE12A113

Title: Programs to get the metadata of pdf and image files

Metadata literally means 'data about data'. Metadata provide additional information about a certain file, such as *its author, creation date, date modified, file size, possible copyright restrictions or the application used to create the file.*

Hardware Specification
This application was developed and run on a machine with the following hardware requirement. It could be run on a machine with better specification
> **RAM:** 3.00GB
> **Processor:** AMD Phenom™ II N620 Dual-Core Processor
> **Processor speed:** 2.80GHz
> **Hard drive:** 1TB
> **Mark:** Hewlett Packard (HP) ProBook 6455b

Software Specification
The following software used:
> **Operating System:** Ubuntu 14.04
> **Editor:** Sublime text editor 3
> **Command-line:** Ubuntu terminal
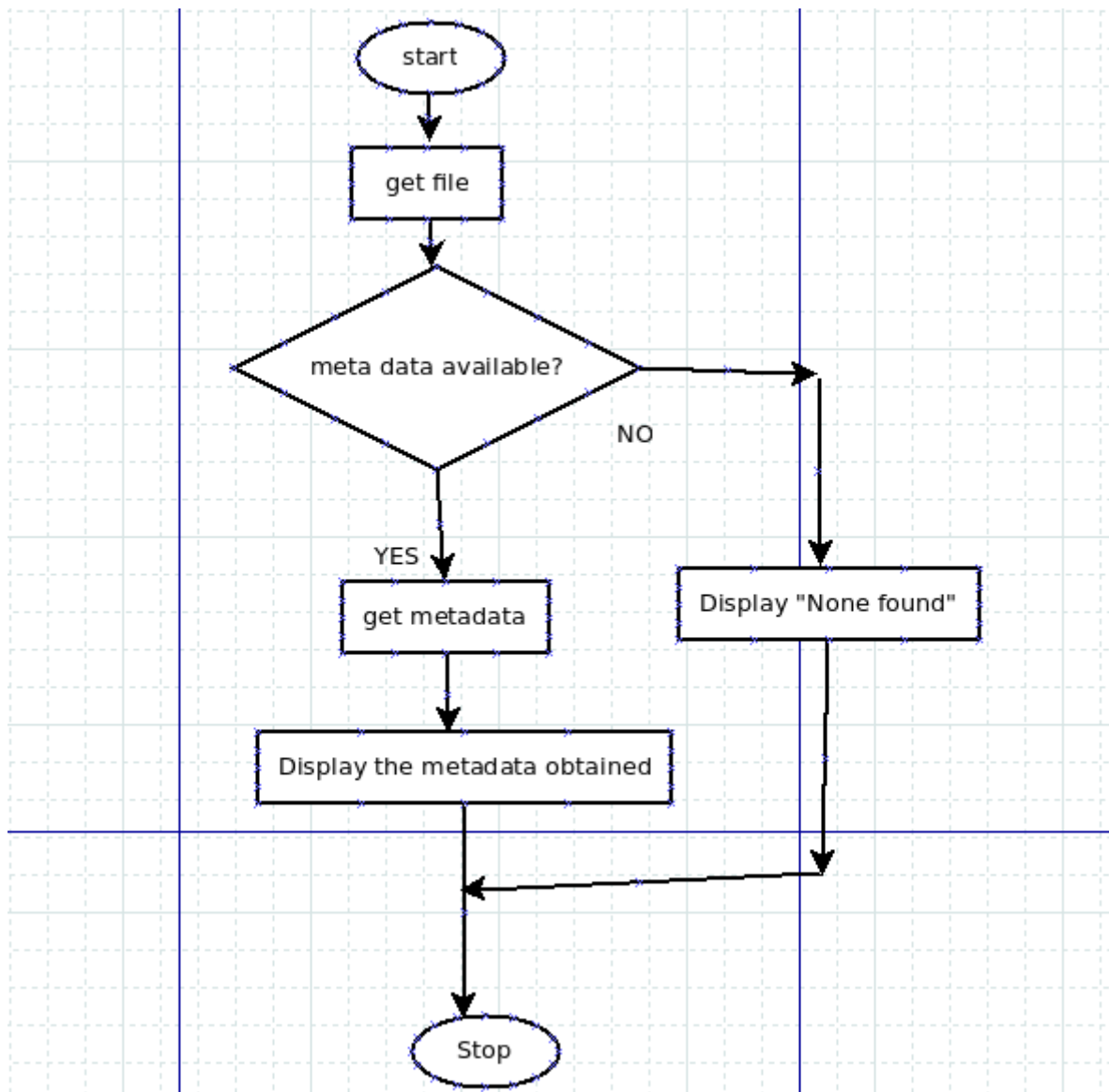
FLOWCHARTS

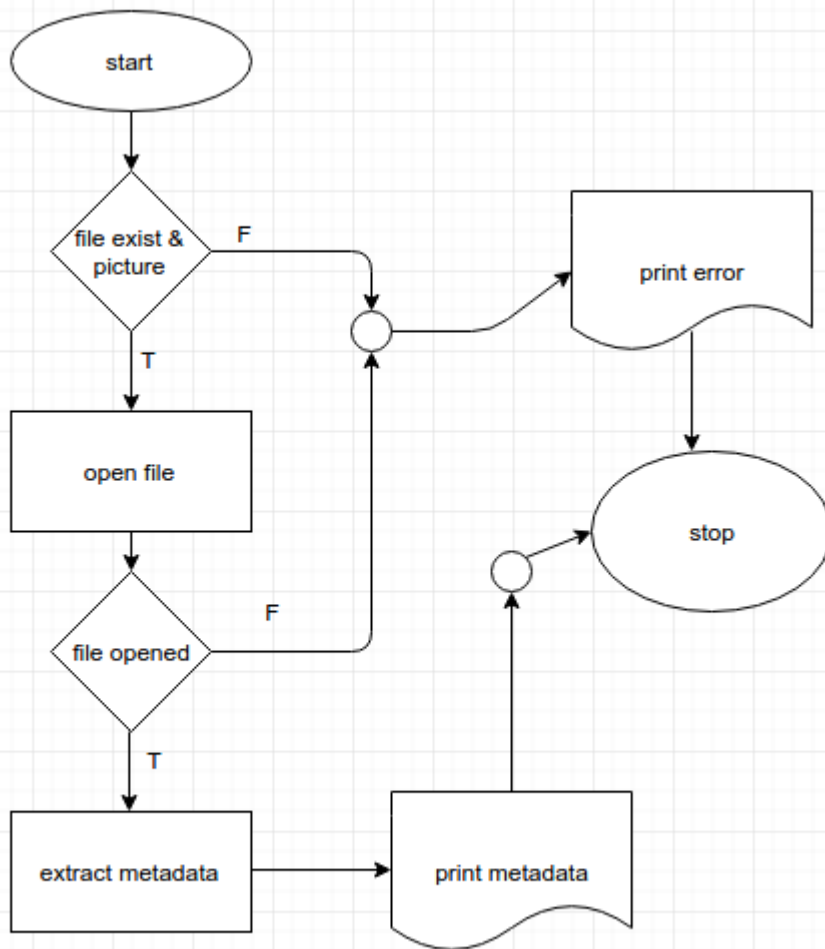*Figure 1: FlowChart of the Metadata of a pdf file*

Figure 2: FlowChart metadata of an image file

CODE

```python
#!/usr/bin/env python2.7

import pyPdf
import optparse
from pyPdf import PdfFileReader

def printMeta(fileName):
    pdfFile = PdfFileReader(file(fileName, 'rb'))
    docInfo = pdfFile.getDocumentInfo()
    #print docInfo.producer
    print '[*] PDF MetaData For:'  + str(fileName)
    for metaItem in docInfo:
        print '[+]'  + metaItem + ':' + docInfo[metaItem]

def main():
    parser = optparse.OptionParser('usage %prog "+\
    "-F ')
    parser.add_option('-F', dest='fileName', type='string',\
    help='specify PDF file name')

    (options, args) = parser.parse_args()
    fileName = options.fileName
    if fileName == None:
        print parser.usage
        exit(0)
    else:
        printMeta(fileName)

if __name__ == '__main__':
    main()
```

Figure 3: code to get Metadata of a pdf file

```python
from PIL import Image
from PIL.ExifTags import TAGS

def get_exif_data(fname):
    """Get embedded EXIF data from image file."""
    ret = {}
    try:
        # Open the file
        img = Image.open(fname)
        if hasattr( img, '_getexif' ):
            # extract the EXIF data
            exifinfo = img._getexif()
        if exifinfo != None:
            # Loop through the data extracted and decode
            for tag, value in exifinfo.iteritems():
                decoded = TAGS.get(tag, tag)
                ret[decoded] = value
    # Catch in case of I/O Error
    except IOError:
        print 'IOERROR ' + fname
    return ret

if __name__ == '__main__':
    fileName = 'food.jpg'
    exif = get_exif_data(fileName)
    print "\n\n"
    print exif
    print "\n\n"
```

Figure 4: Code to get metadata of an image

OUTPUT

Figure 5: Output of the metadata of a pdf file

Figure 6: Output of the metadata of an image file