

## DIGITAL FORENSICS LAB

Exercise 4	
Name	S Shyam Sundaram
Registration Number	19BCE1560
Slot	L39+L40
Faculty	Dr. Seshu Babu Pulgara
Date	31 <sup>st</sup> August, 2021

### AIM

To investigate MS Word documents.

### PART 1

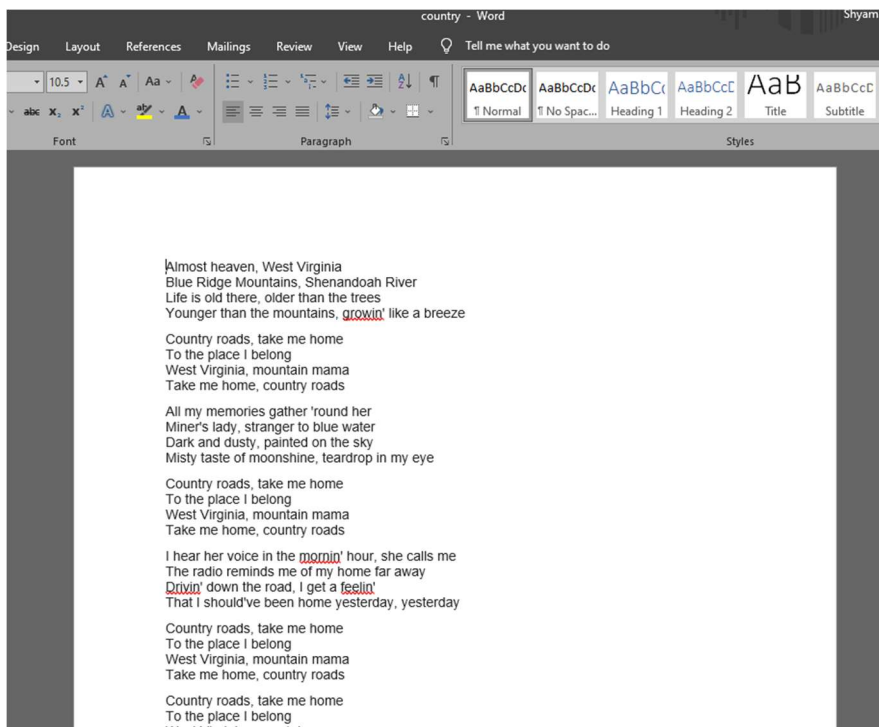
#### STRINGS

Strings just scans the file passed to it for UNICODE or ASCII strings of a default length of 3 or more characters.

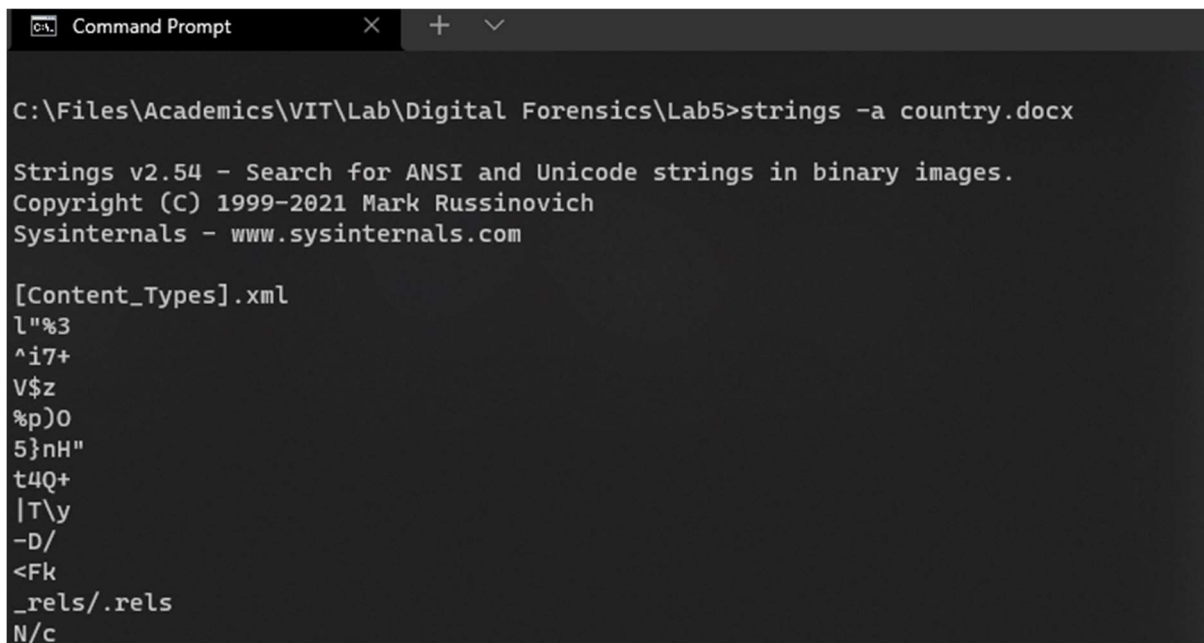
#### STEPS

1. Install strings and extract to a path.
2. Add the path to environment variable 'PATH'.
3. Open the command prompt and start working.

#### DOCUMENT USED



## ASCII only search



```
Command Prompt
C:\Files\Academics\VIT\Lab\Digital Forensics\Lab5>strings -a country.docx

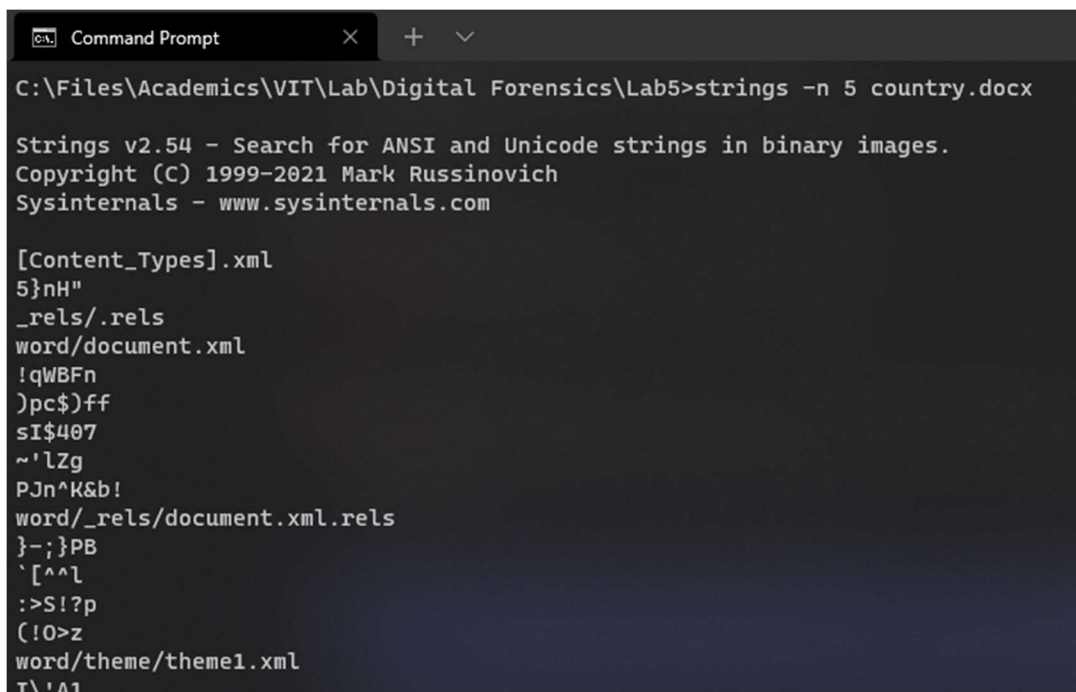
Strings v2.54 - Search for ANSI and Unicode strings in binary images.
Copyright (C) 1999-2021 Mark Russinovich
Sysinternals - www.sysinternals.com

[Content_Types].xml
l"%3
^i7+
V$z
%p)0
5}nH"
t4Q+
|T\y
-D/
<Fk
_rels/.rels
N/c
```

### **OBSERVATION**

This gives a list of three-character long ASCII strings, which is the default length.

## Search for strings that are at least 5 characters long



```
Command Prompt
C:\Files\Academics\VIT\Lab\Digital Forensics\Lab5>strings -n 5 country.docx

Strings v2.54 - Search for ANSI and Unicode strings in binary images.
Copyright (C) 1999-2021 Mark Russinovich
Sysinternals - www.sysinternals.com

[Content_Types].xml
5}nH"
_rels/.rels
word/document.xml
!qWBFn
)pc$)ff
sI$407
~'lZg
PJn^K&b!
word/_rels/document.xml.rels
}-; }PB
['^l
:>S! ?p
(!0>z
word/theme/theme1.xml
T\'A1
```

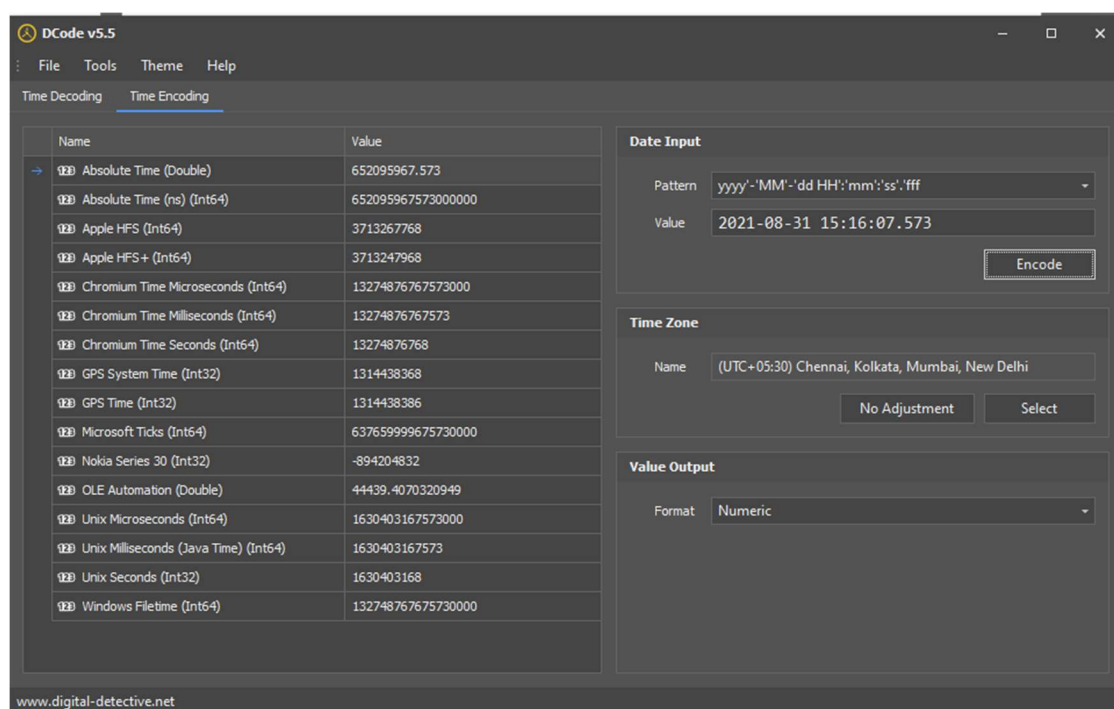
### **OBSERVATION**

The -n 5 tells the tool to search for strings that are at least 5 characters long.

## DCODE

Dcode is a utility designed to calculate date/time values from various timestamps that are found in a data file. It takes an input and gives out the decoded/encoded values.

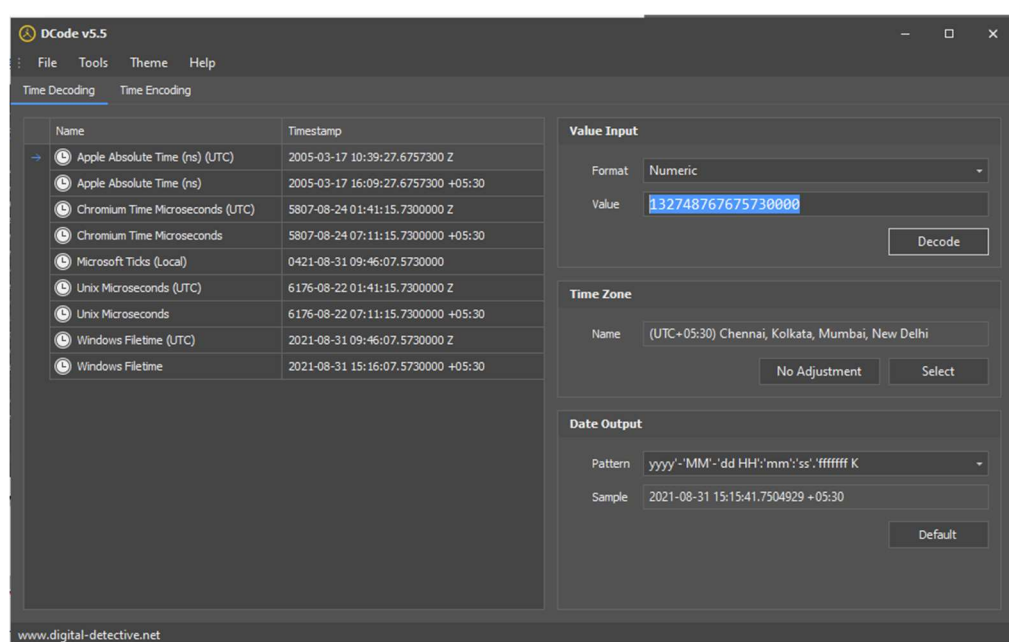
### Encode a date given in readable form



## OBSERVATION

We enter the date and time to encode in the Date Input dialog box, along with the pattern. Then, we click 'Encode' and get the list of encoded values on the left.

### Decodes a value given into readable timestamp



## OBSERVATION

We enter the value obtained in the 'Windows File Time' row of the encode section and enter here. Upon doing so, we get back nearly the same date and time in the Windows Filetime row.

## PART 2

DOCX contains a bunch of XML files that describes the document, has some settings, themes and even the content itself. In this section, we explore them.

### STEPS

1. Open WinRAR or WinZip.
2. Open the docx file and click 'Extract' to obtain the XML files.
3. All XML files can be seen in the folder.

List of files extracted from country.docx.

Name	Date modified	Type	Size
_rels	31-Aug-21 3:14 PM	File folder	
docProps	31-Aug-21 3:14 PM	File folder	
word	31-Aug-21 3:14 PM	File folder	
[Content_Types].xml		XML Document	2 KB

## WORD/DOCUMENT.XML

```
document.xml
File | C:/Files/Academics/VIT/Lab/Digital%20Forensics/Lab5/country/word/document.xml
<?xml version="1.0" encoding="UTF-16" standalone="yes" type="application/vnd.openxmlformats-officedocument.wordprocessingml.document" />
<w:document xmlns:w="http://schemas.microsoft.com/office/word/2012/wordml" />
  <w:body>
    <w:p w:id="1E674A89" w:14:taskId="77777777" w:rsidR="009467E1" w:rsidRPr="009467E1" w:rsidRDefault="009467E1" w:rsidP="009467E1">
      <w:pPr>
        <w:shd w:val="clear" w:color="auto" w:fill="FFFFFF" />
        <w:spacing w:line="240" w:lineRule="auto" />
      </w:pPr>
      <w:r>
        <w:rFonts w:ascii="Arial" w:eastAsia="Times New Roman" w:hAnsi="Arial" w:cs="Arial" />
        <w:color w:val="202124" />
        <w:sz w:val="21" />
        <w:s2Cs w:val="21" />
        <w:lang w:eastAsia="en-IN" />
      </w:r>
    </w:p>
    <w:r w:rsidRPr="009467E1">
      <w:rPr>
        <w:rFonts w:ascii="Arial" w:eastAsia="Times New Roman" w:hAnsi="Arial" w:cs="Arial" />
        <w:color w:val="202124" />
        <w:sz w:val="21" />
        <w:s2Cs w:val="21" />
        <w:lang w:eastAsia="en-IN" />
      </w:rPr>
      <w:t>Almost heaven, West Virginia</w:t>
    </w:r>
    <w:r w:rsidRPr="009467E1">
      <w:rPr>
        <w:rFonts w:ascii="Arial" w:eastAsia="Times New Roman" w:hAnsi="Arial" w:cs="Arial" />
        <w:color w:val="202124" />
        <w:sz w:val="21" />
        <w:s2Cs w:val="21" />
        <w:lang w:eastAsia="en-IN" />
      </w:rPr>
      <w:br />
      <w:t>Blue Ridge Mountains, Shenandoah River</w:t>
    </w:r>
    <w:r w:rsidRPr="009467E1">
      <w:rPr>
        <w:rFonts w:ascii="Arial" w:eastAsia="Times New Roman" w:hAnsi="Arial" w:cs="Arial" />
        <w:color w:val="202124" />
        <w:sz w:val="21" />
        <w:s2Cs w:val="21" />
        <w:lang w:eastAsia="en-IN" />
      </w:rPr>
    </w:r>
  </w:body>
```

## OBSERVATION

We see each line of the document within a <w:t> tag. They are also nested in tags that contain formatting information like font, size and color.

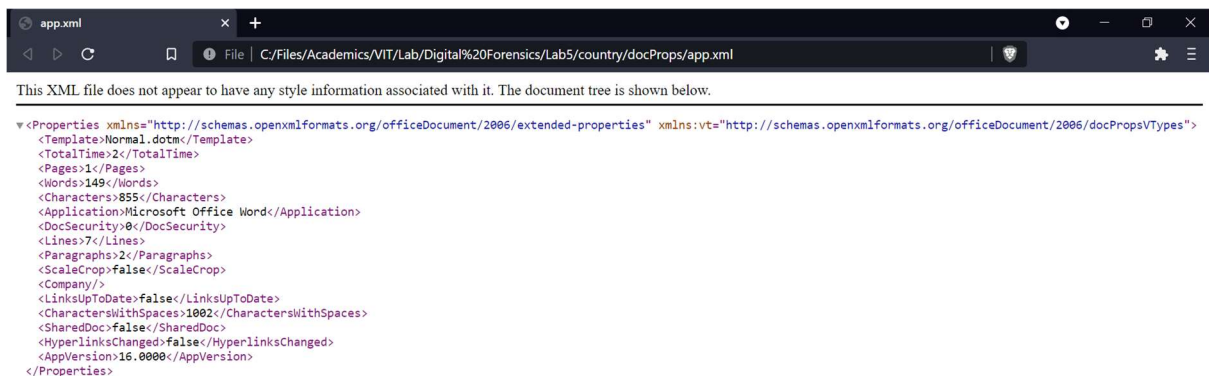
## DOCPROPS/CORE.XML



## OBSERVATION

We see details such as who created the document (Shyam Sundaram), who modified it last (Shyam Sundaram), when it was created and modified (in the <dcterms:created> and <dcterms:modified> tags) and also how many times the file was edited or revised (3 times, as shown by the <cp:revision> tag).

## DOCPROPS/APP.XML



## OBSERVATION

This XML file contains details such as the template the document follows, how many pages, characters and words it contains.

## CONCLUSION

We have thus applied file forensics on a DOCX file and fetched details from its XML files such as revision count, author, date and time of modifications etc.