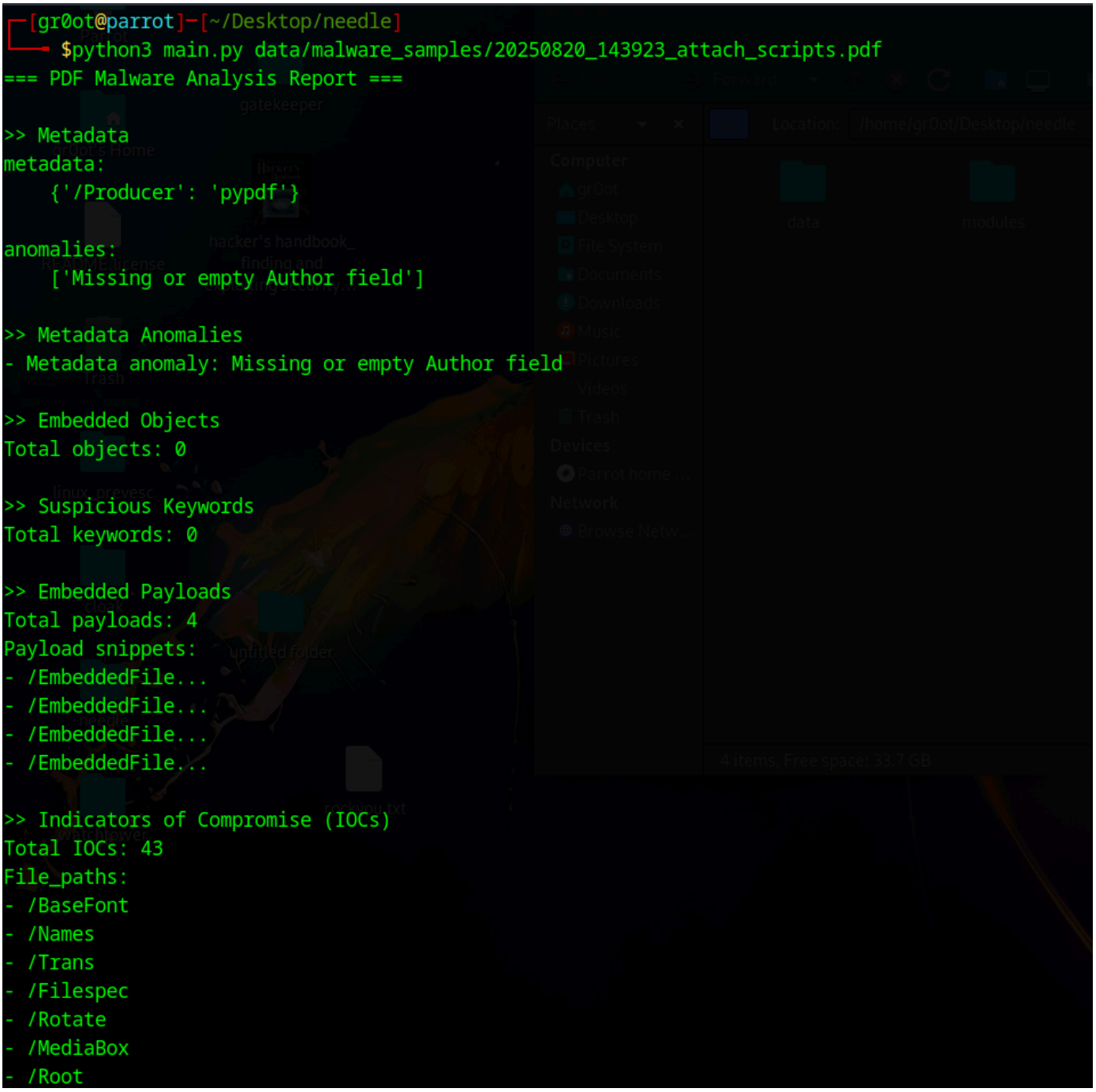
Needle_SS_File

Needle : PDF Malware Analyzer

1. **Test PDF Execution (Terminal Output)** :

   These screenshots show the toolkit analyzing a sample PDF directly from the terminal. The modules run sequentially and display extracted metadata, suspicious keywords, and indicators of compromise.


[gr0ot@parrot]-[~/Desktop/needle]
$python3 main.py data/malware_samples/20250820_143923_attach_scripts.pdf
=== PDF Malware Analysis Report ===
>> Metadata
metadata:
{'/Producer': 'pypdf'}
anomalies:
['Missing or empty Author field']
>> Metadata Anomalies
- Metadata anomaly: Missing or empty Author field
>> Embedded Objects
Total objects: 0
>> Suspicious Keywords
Total keywords: 0
>> Embedded Payloads
Total payloads: 4
Payload snippets:
- /EmbeddedFile...
- /EmbeddedFile...
- /EmbeddedFile...
- /EmbeddedFile...
>> Indicators of Compromise (IOCs)
Total IOCs: 43
File_paths:
- /BaseFont
- /Names
- /Trans
- /Filespec
- /Rotate
- /MediaBox
- /Root

```
>> Indicators of Compromise (IOCs)
Total IOCs: 43
File_paths:
- /BaseFont
- /Names
- /Trans
- /Filespec
- /Rotate
- /MediaBox
- /Root
- /Subtype
- /EmbeddedFiles
- /FlateDecode

>> Risk Assessment
Risk Score: 100
Severity Level: Critical
Reasons:
- Metadata anomaly: Missing or empty Author field
- Keyword /EmbeddedFile detected 4 times
- Keyword /Names detected 2 times
- Payload type EmbeddedFile extracted
- Payload type EmbeddedFile extracted
- Payload type EmbeddedFile extracted
- Payload type EmbeddedFile extracted
- file_paths found: /BaseFont, /Names, /Trans, /Filespec, /Rotate, /MediaBox, /Root, /Subtype, /EmbeddedFiles, /FlateDecode, /mBLNHos(37;mMNuqa+3=C@0,&.1Ng@ete'o^\m["iaILOTQWqYcj1rb)e.$rFA@u
9?U(@TK2qP]gjO=+T5HU.*)-L*%Y~>, /Pages, /ImageB, /Producer, /Size, /Helvetica, /ASCII85Decode, /Kids, /Length, /Text, /Font, /Parent, /Count, /Page, /F, /EmbeddedFile, /Type, /Filter, /EF, /
bin/sh, /Type1, /ImageC, /ImageI, /Info, /F1, /Name, /PDF, /Encoding, /Catalog, /Resources, /WinAnsiEncoding, /ProcSet, /Contents

=== End of Report ===

 Report saved to: results/20250820_143923_attach_scripts_report.txt
```

---

2. **Saved Analysis Report (Results)** :

   These screenshots highlight the automatically generated report saved in the `results/` directory. The report contains risk score, severity level, and reasons for detection, ensuring analysis is preserved for later review.

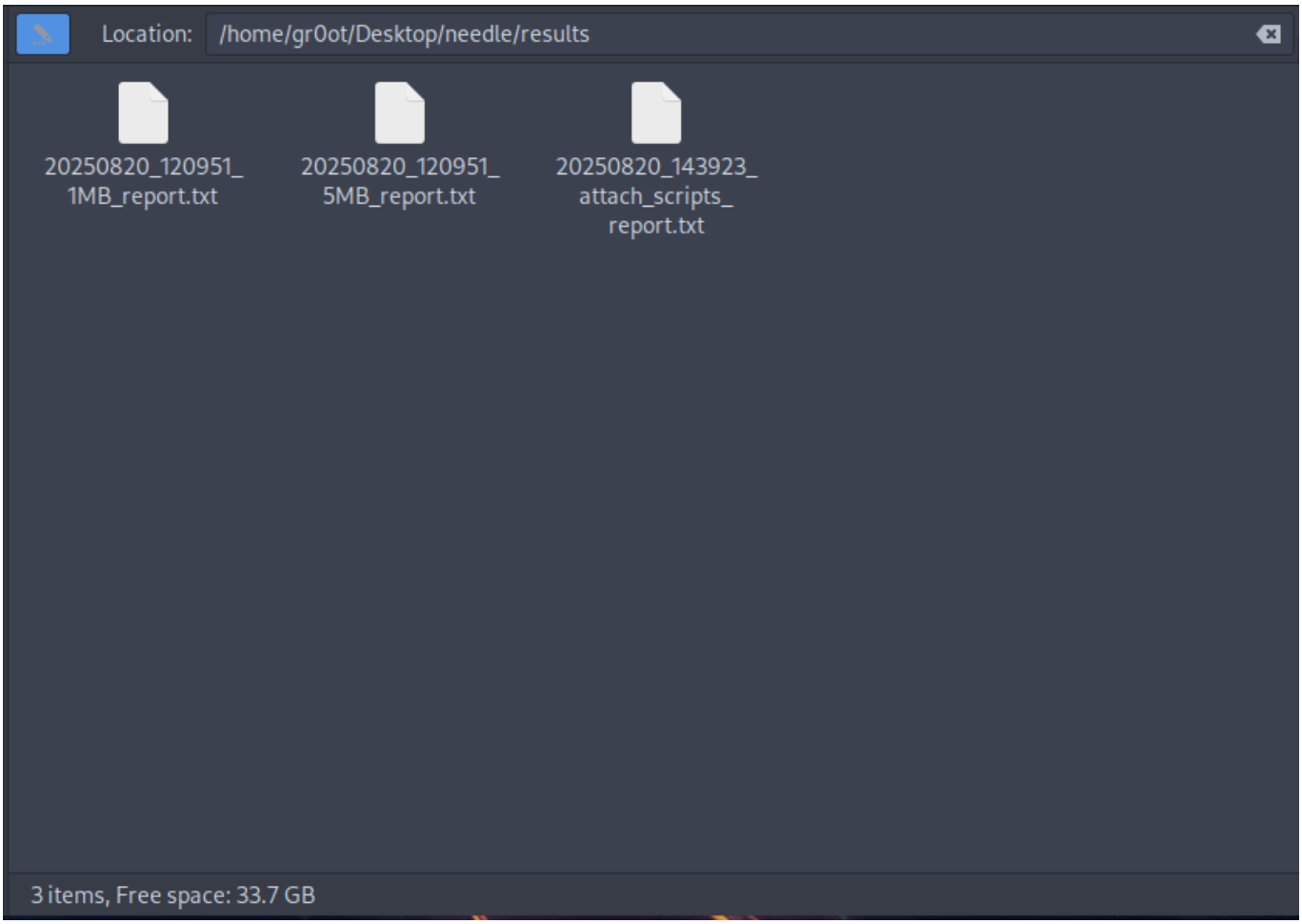
20250820_120951_1MB_report.txt

```
=== PDF Malware Analysis Report ===

>> Metadata
metadata:
    {'/Author': 'anonymous', '/CreationDate': "D:20250820120951+02'00'", '/Creator': 'ReportLab PDF Library - www.reportlab.com', '/Keywords': '', '/ModDate': "D:20250820120951+02'00'",
 '/Producer': 'ReportLab PDF Library - www.reportlab.com', '/Subject': 'unspecified', '/Title': 'untitled', '/Trapped': '/False'}

anomalies:
    ['CreationDate == ModDate (suspicious)', 'Suspicious /Creator value: ReportLab PDF Library - www.reportlab.com', 'Suspicious /Producer value: ReportLab PDF Library -
 www.reportlab.com']

>> Metadata Anomalies
- Metadata anomaly: CreationDate == ModDate (suspicious)
- Metadata anomaly: Suspicious /Creator value: ReportLab PDF Library - www.reportlab.com
- Metadata anomaly: Suspicious /Producer value: ReportLab PDF Library - www.reportlab.com

>> Embedded Objects
Total objects: 0

>> Suspicious Keywords
Total keywords: 0

>> Embedded Payloads
Total payloads: 0

>> Indicators of Compromise (IOCs)
Total IOCs: 51
Domains:
- www.reportlab.com
File_paths:
- /Text
- /Subtype
- /ImageC
- /ImageI
- /ImageI
- /MediaBox
- /Subject
- /Type
- /Title
- /WinAnsiEncoding
- /Page

>> Risk Assessment
Risk Score: 100
Severity Level: Critical
Reasons:
- Metadata anomaly: CreationDate == ModDate (suspicious)
- Metadata anomaly: Suspicious /Creator value: ReportLab PDF Library - www.reportlab.com
- Metadata anomaly: Suspicious /Producer value: ReportLab PDF Library - www.reportlab.com
- domains found: www.reportlab.com
- file_paths found: /Text, /Subtype, /ImageC, /ImageI, /MediaBox, /Subject, /Type, /Title, /WinAnsiEncoding, /Page, /Helvetica, /Size, /F1, /PageMode, /Producer, /ASCII85Decode, /Font, //
 www.reportlab.com, /Parent, /Author, /Trapped, /CreationDate, /Creator, /Root, /Trans, /BaseFont, /Resources, //www.reportlab.com), /UseNone, /Encoding, /PDF, /Name, /Catalog, /Rotate, /
 False, /Filter, /FlateDecode, /\_S8aimE0;[C[M9!n_8oQAS]#g#sMd/k2a^]a^DN[!W~>endstream, /Keywords, /ModDate, /ImageB, /Contents, /ProcSet, /Pages, /Kids, /Info, /Count, /Type1, /Length, /
 ID

=== End of Report ===
```

Plain Text Tab Width: 4 Ln 1, Col 1 INS

3. **Results Directory Overview** :

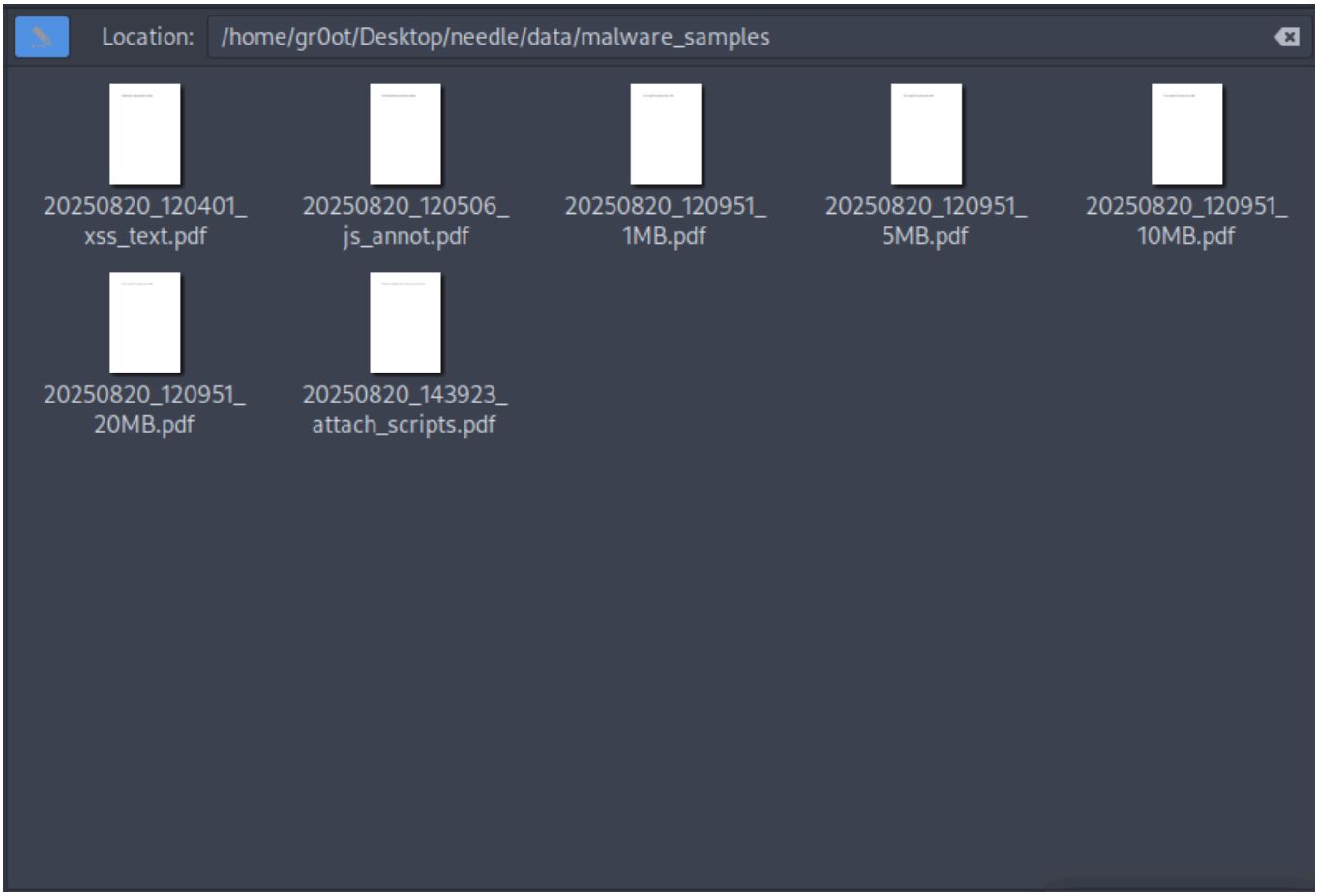
   This screenshot shows multiple saved reports in the results/ folder. It demonstrates that the toolkit can handle repeated analyses and maintain organized outputs for different test files.



4. **Test PDF Sample (Source File)** :

   This screenshot displays the test PDF file downloaded from GitHub. It serves as the input document for analysis and validates that the toolkit works on real-world samples.

   (I have downloaded these sample test pdf's from GitHub source : https://github.com/klausnitzer/pentest-pdf-collection/tree/main/pdf_files)


Location:
/home/gr0ot/Desktop/needle/data/malware_samples
20250820_120401_xss_text.pdf
20250820_120506_js_annot.pdf
20250820_120951_1MB.pdf
20250820_120951_5MB.pdf
20250820_120951_10MB.pdf
20250820_120951_20MB.pdf
20250820_143923_attach_scripts.pdf