

# Reverse Engineering and Application Development

Iowa Code Camp 2024



# whoami

Nicholas Starke

Reverse Engineer at Hewlett Packard Enterprise

Focused on firmware security

Background in Application Development

# Agenda

- Why Reverse Engineering?
- Reverse Engineering Tools
- Reverse Engineering .NET
- Reverse Engineering JVM
- Reverse Engineering Native Machine Code Binaries

# Why Reverse Engineering?

- To understand the internals of a given piece of software
  - Find Security Vulnerabilities
  - Analyze Malware
- To develop data integrations for proprietary software
- Modify legacy applications to fix vulnerabilities or implement new features
  - Works even when the source code or compiler toolchain is unavailable

# Reverse Engineering Tools - .NET

- 1) dotpeek - <https://www.jetbrains.com/decompiler/>
- 2) DnSpyEx - <https://github.com/dnSpyEx/dnSpy>
- 3) ILSpy - <https://github.com/icsharpcode/ILSpy>
- 4) ildasm / ilasm

# Reverse Engineering Tools - JVM

- 1) JD-GUI - <https://java-decompiler.github.io/>
- 2) JADX-GUI (Android) - <https://github.com/skylot/jadx>
- 3) Jasper / Jasmin - <https://github.com/kohsuke/jasper> / <https://github.com/davidar/jasmin>

# Reverse Engineering Tools - Machine Code

- 1) Ghidra - <https://ghidra-sre.org/>
- 2) IDA Pro - <https://hex-rays.com/ida-pro>
- 3) Binary Ninja - <https://binary.ninja/>

# Reverse Engineering - .NET

- For applications that compile down to byte code (JVM / CLR, primarily) there are tools that can take a compiled dll, jar, war, exe and create a near-source code quality representation of the code.
- There are ways to modify a compiled application without source code.
  - Code signing helps mitigate the risk of this type of attack
- Obfuscation is usually enough of an impediment for Reverse Engineers



# Why reverse engineer server-side applications? - Security

- As an attacker, often compiled applications contain secrets like keys and passwords
- As an attacker, you might want to modify an application without the source code
  - This is possible using tools like ILASM.exe/ILDASM.exe for dotnet CLR

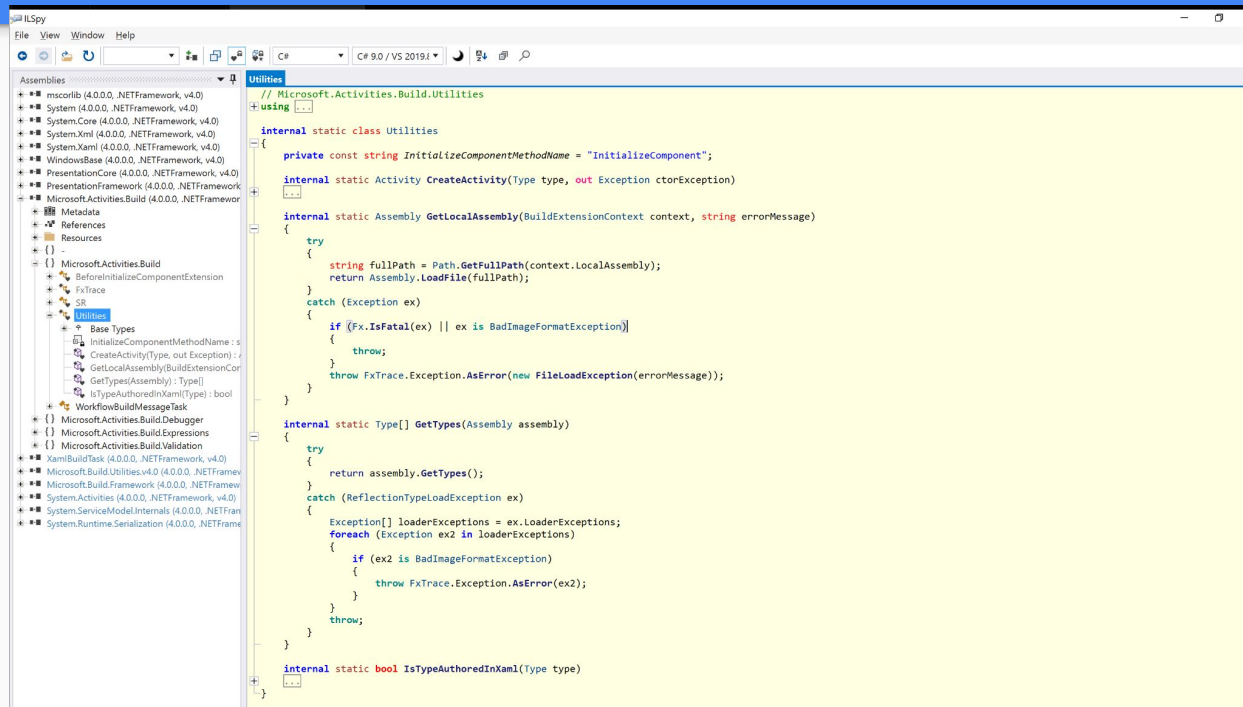
# Why reverse engineer server-side applications? - Dev

- Have you lost the source code? Data loss does happen :-)
- As a developer, you may need to integrate with a product that has no documentation (legacy code, anyone?)
- As a developer, you may want to analyze proprietary code to understand how it works
- As a developer, it is important to understand what an attacker can do with your production binaries from a security perspective

# .NET

- .cs files compile down to .dll or .exe files
- Based on MSIL bytecode for server side apps
  - The equivalent of Java's JVM Bytecode / Smali Bytecode
- Compiles down to MSIL (Microsoft Intermediate Language)
  - The .NET equivalent of JVM Bytecode
- This runs on the .NET CLR (Common language runtime)
- Source files are .cs files which compile to exe or dll
  - DLL's more common for web apps

# ILSpy



The screenshot shows the ILSpy application interface. On the left, the 'Assemblies' pane lists various .NET Framework assemblies, with 'Microsoft.Activities.Build (4.0.0.0, .NETFramework, v4.0)' selected. The main editor displays the source code for the 'Utilities' class, which is part of the 'Microsoft.Activities.Build.Utilities' namespace. The code includes several static methods: 'CreateActivity', 'GetLocalAssembly', 'GetTypes', and 'IsTypeAuthoredInXaml'. The 'CreateActivity' method takes a 'Type' and an 'Exception' as input. 'GetLocalAssembly' uses 'Path.GetFullPath' and 'Assembly.LoadFile'. 'GetTypes' uses 'ReflectionTypeLoadException' to load types. 'IsTypeAuthoredInXaml' is a simple boolean check.

```
// Microsoft.Activities.Build.Utilities
using System;

internal static class Utilities
{
    private const string InitializeComponentMethodName = "InitializeComponent";

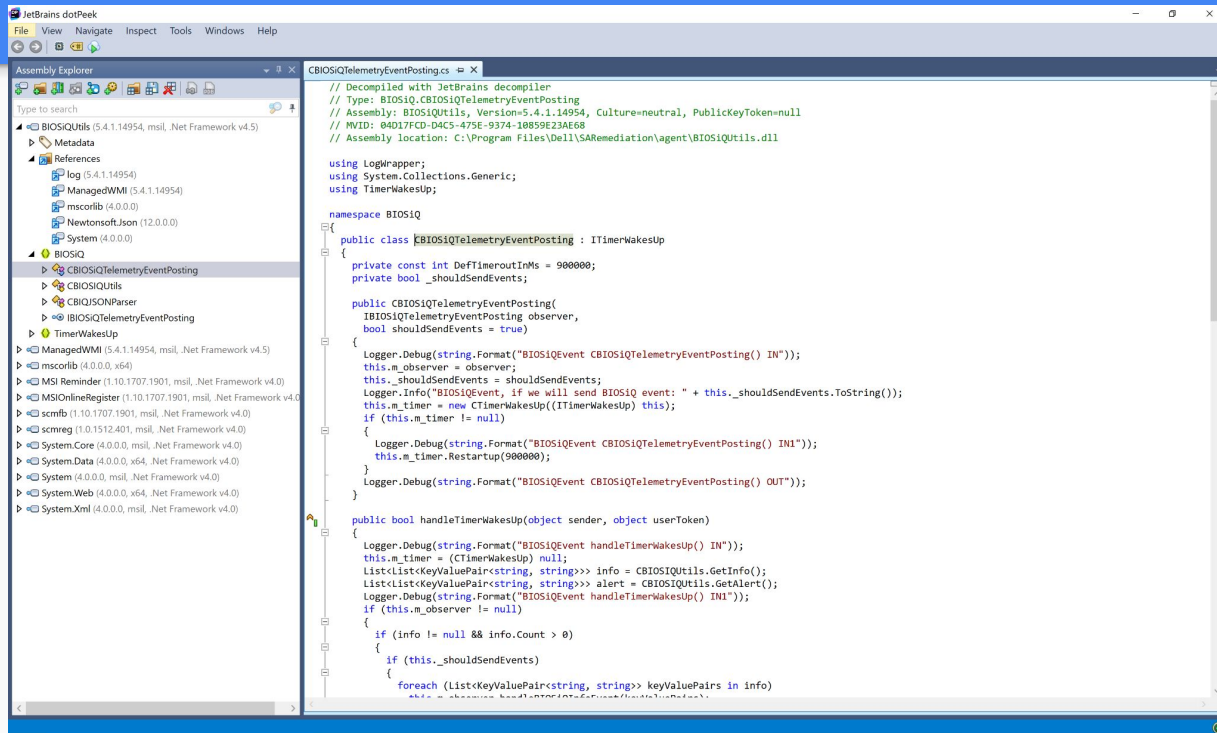
    internal static Activity CreateActivity(Type type, out Exception ctorException)
    {
        // ...
    }

    internal static Assembly GetLocalAssembly(BuildExtensionContext context, string errorMessage)
    {
        try
        {
            string fullPath = Path.GetFullPath(context.LocalAssembly);
            return Assembly.LoadFile(fullPath);
        }
        catch (Exception ex)
        {
            if (Fx.IsFatal(ex) || ex is BadImageFormatException)
            {
                throw;
            }
            throw FxTrace.Exception.AsError(new FileLoadException(errorMessage));
        }
    }

    internal static Type[] GetTypes(Assembly assembly)
    {
        try
        {
            return assembly.GetTypes();
        }
        catch (ReflectionTypeLoadException ex)
        {
            Exception[] loaderExceptions = ex.LoaderExceptions;
            foreach (Exception ex2 in loaderExceptions)
            {
                if (ex2 is BadImageFormatException)
                {
                    throw FxTrace.Exception.AsError(ex2);
                }
            }
            throw;
        }
    }

    internal static bool IsTypeAuthoredInXaml(Type type)
    {
        // ...
    }
}
```

# Dotpeek

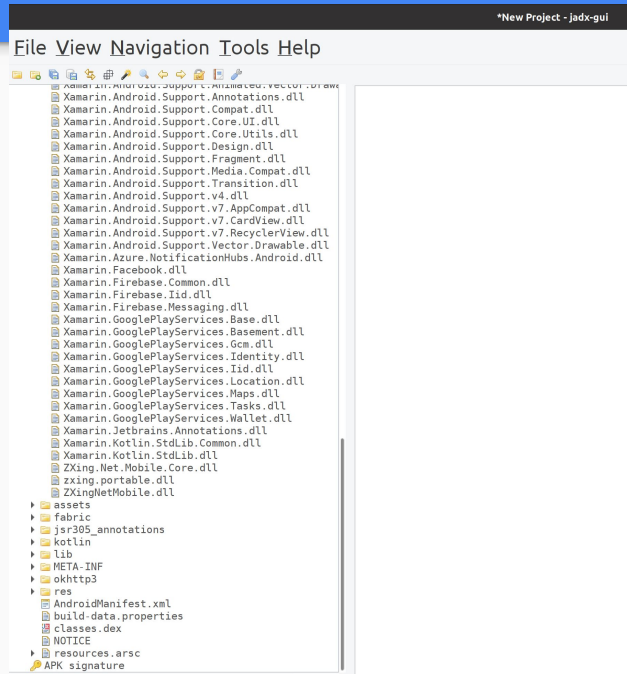


# Client-side .NET

Jadx-gui - <https://github.com/skylot/jadx>

- Useful for extracting Xamarin Assemblies
- Extracts the static content (res/) from the APK and presents it in a tree view

# Client-side .NET



# Modifying MSIL Bytecode without Source Code

The next few slides will focus on techniques for modifying MSIL Bytecode without access to the original source code.

We'll discuss:

- Why would anyone want to do this?
- Examples
- Process
- Tooling



# Why would anyone want to do this?

## Development:

- Modify an application when source code is lost

## Security:

- Patch an application to log out sensitive information

# Examples of Patching MSIL Bytecode - Security

## Server-side

- Server side: when a login request is received, log out the username and password to a file on the filesystem.

## Client-side

- Client side: make an HTTP request to an unauthorized remote server with authentication tokens received from a legitimate authentication request

# Process

- 1) Write out Dotnet code you wish to inject in a console application. Create a function that accepts the data you wish to operate on.
- 2) Disassemble this console application
- 3) Disassemble the source code you wish to inject code into
- 4) Modify the source code disassembly to include the console application disassembly and write integration disassembly to call the function you wrote in 1)
- 5) Reassemble source .cs file
- 6) Reassemble DLL / Drop on file system cache.

<https://starkeblog.com/backdooring/dotnet/2024/04/19/backdooring-dotnet-applications.html>

# Dotnet Disassembler / Assembler Duo

Dotnet MSIL Assembler: ILASM.exe -

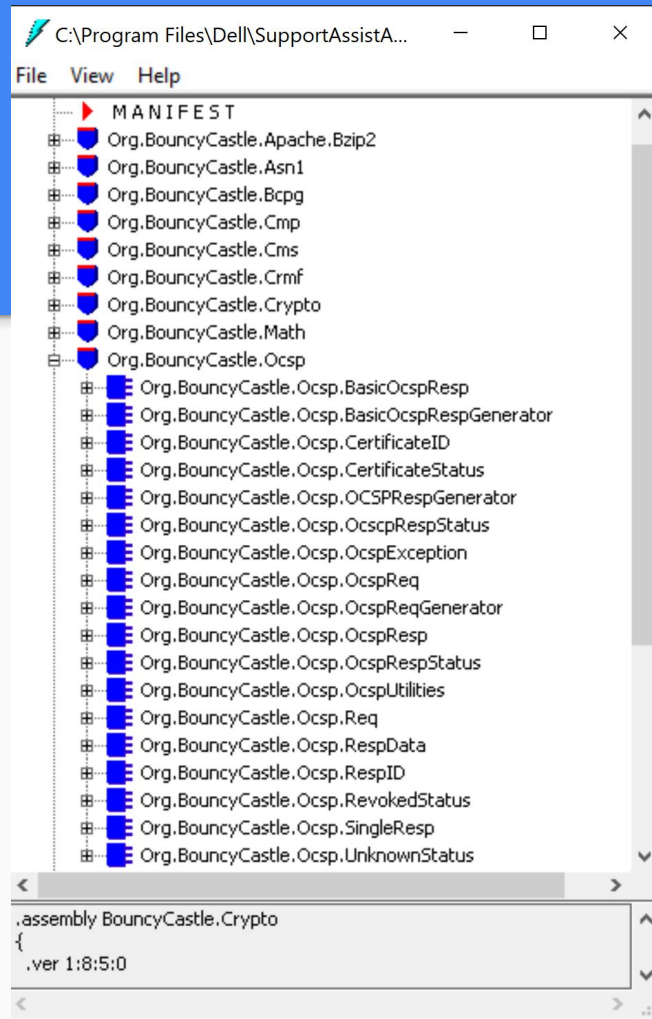
<https://docs.microsoft.com/en-us/dotnet/framework/tools/ilasm-exe-il-assembler>

Dotnet MSIL Disassembler: ILDASM.exe -

<https://docs.microsoft.com/en-us/dotnet/framework/tools/ildasm-exe-il-disassembler>

- These two tools are built to work with each other.
- Available through Visual Studio Developer Shell

ildasm



# Ildasm - Screenshot

```
Org.BouncyCastle.Apache.Bzip2.CRC::UpdateCRC : void(int32)
Find Find Next
.method assembly hidebysig instance void
    UpdateCRC(int32 inCh) cil managed
{
    // Code size          47 (0x2f)
    .maxstack 4
    .locals init (int32 U_0)
    IL_0000: ldarg.0
    IL_0001: ldfld      int32 Org.BouncyCastle.Apache.Bzip2.CRC::globalCrc
    IL_0006: ldc.i4.s   24
    IL_0008: shr
    IL_0009: ldarg.1
    IL_000a: xor
    IL_000b: stloc.0
    IL_000c: ldloc.0
    IL_000d: ldc.i4.0
    IL_000e: bge.s     IL_0018
    IL_0010: ldc.i4    0x100
    IL_0015: ldloc.0
    IL_0016: add
    IL_0017: stloc.0
    IL_0018: ldarg.0
    IL_0019: ldarg.0
    IL_001a: ldfld      int32 Org.BouncyCastle.Apache.Bzip2.CRC::globalCrc
    IL_001f: ldc.i4.8
    IL_0020: shl
    IL_0021: ldsfld     int32[] Org.BouncyCastle.Apache.Bzip2.CRC::crc32Table
    IL_0026: ldloc.0
    IL_0027: ldelem.i4
    IL_0028: xor
    IL_0029: stfld     int32 Org.BouncyCastle.Apache.Bzip2.CRC::globalCrc
    IL_002e: ret
} // end of method CRC::UpdateCRC
```

# How to mitigate this threat

- Strong name signing -  
<https://docs.microsoft.com/en-us/dotnet/standard/assembly/sign-strong-name>
- Read only file system for executable code

Android implements code signing by default - consider this for your production applications even when they are server-side.

# Anti-reverse engineering techniques

## Obfuscation!

- Dotfuscator - <https://www.preemptive.com/products/dotfuscator>

## Benefits:

- Makes code extremely difficult to reverse
- Makes code extremely difficult to modify

## Cons:

- Server-side: usually expensive in terms of \$ cost



# Goals of Obfuscation

Obfuscation can be used to deter attackers

Usually all you need to do is put up enough of a barrier to entry that it makes a potential attacker move on to the next target

Obfuscation alone is not sufficient to secure an application!

- Secrets should not be stored in source code
- Secrets should not be stored in source code
- **SECRETS SHOULD NOT BE STORED IN SOURCE CODE**

# Resources

- Managed Code Rootkits (Book):  
<https://www.amazon.com/Managed-Code-Rootkits-Hooking-Environments/dp/1597495743>

# Reverse Engineering Java / JVM

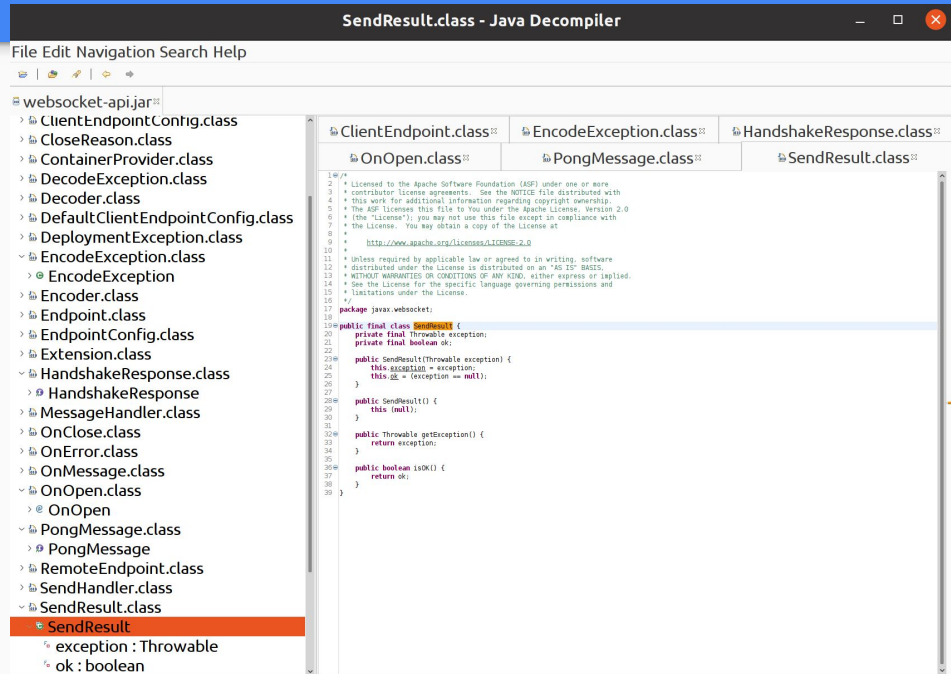
- .java files compile down to .class files
- Based on JVM bytecode for server side apps
  - The equivalent of .NET's MSIL

# Server-side Java - JD-GUI

Reverse engineering tools for Server-side Java applications

- JD-GUI (<https://github.com/java-decompiler/jd-gui>)
- `brew install jd-gui` on MacOS
- Install from github releases on Linux
- Requires JDK 1.8 specifically
- Has sufficient decompiler output
- Can output all java files in a jar

# JD-GUI Screenshot



# Java - Fernflower Decompiler

Fernflower is the JetBrains Java Decompiler

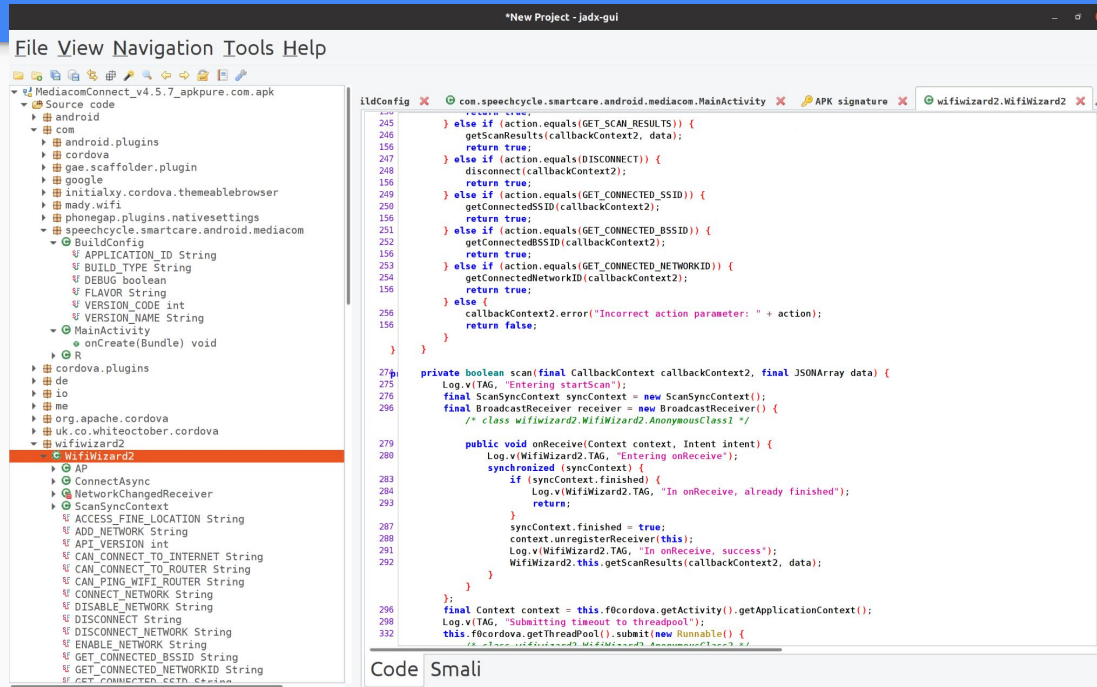
- Comes bundled with IntelliJ
- Can be run from the command line directly
- Has much clearer output than JD-GUI
- No UI, outputs .java files

# Client-side Java (Native Android)

Jadx-gui - <https://github.com/skylot/jadx>

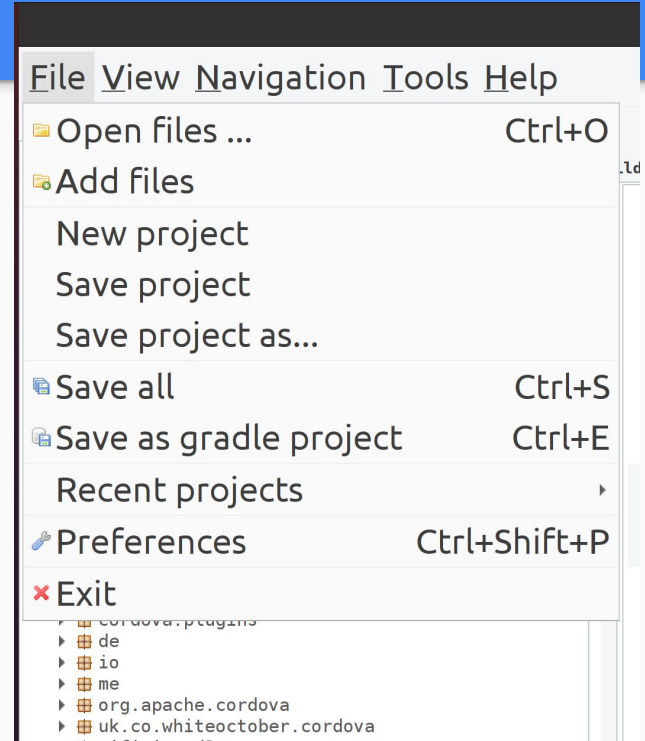
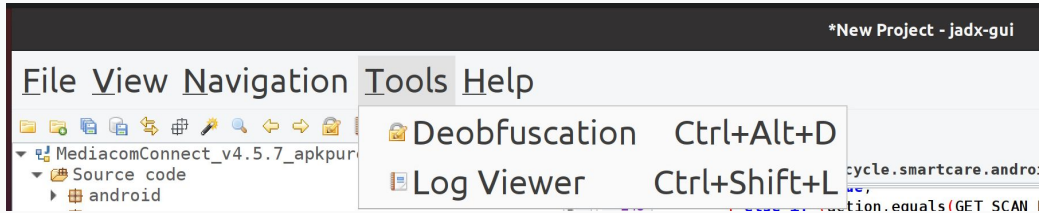
- A lot like JD-GUI
- Does all the manual work of extracting the APK then disassembling/decompiling the SMALI bytecode into Java classes
- Extracts the static content (res/) from the APK and presents it in a tree view

# JADX-GUI Screenshot





# More JADX-GUI Screenshots



# JD-GUI - Scala Decompiler Output

```
1  import scala.Predef$;
2
3  public final class Hello$ {
4      public static final Hello$ MODULE$;
5
6      public void main(String[] args) {
7          Predef$.MODULE$.println("Hello, world");
8      }
9
10     private Hello$() {
11         MODULE$ = this;
12     }
13 }
14
```

# Java Disassembler / Assembler Duo

Java Class Disassembler: Jasper - <https://github.com/kohsuke/jasper>

Java Class Assembler: Jasmin - <https://github.com/davidar/jasmin>

- These two tools are built to work with each other.
- Jasmin will not work with “javap -c”!
- Both tools were built between 2000-2004
- Modifications to source are necessary for both to compile with modern Java tooling.
- Jasper works with maven, Jasmin works with ant.

# How to mitigate this threat

- Jar signing (via `jarsigner` tool)
- Read only file system for executable code

Android implements jar signing by default - consider this for your production applications even when they are server-side.

# JVM: Anti-reverse engineering techniques

Obfuscation!

- Proguard - Java

Benefits:

- Makes code extremely difficult to reverse
- Makes code extremely difficult to modify

Cons:

- Server-side: usually expensive in terms of \$ cost

# JVM: More Resources

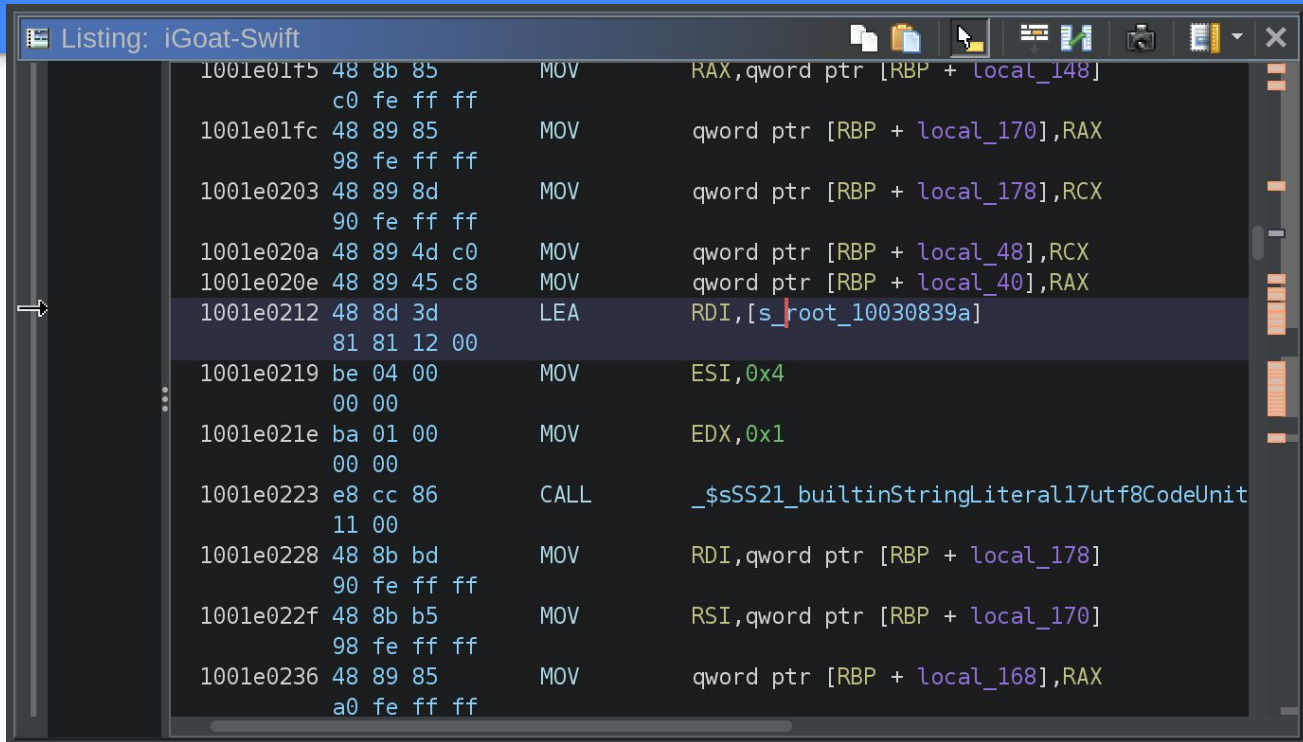
- Covert Java (Book): <https://www.amazon.com/dp/0672326388>

# Native Machine Code

Ghidra - <https://ghidra-sre.org/>

- Developed by the US National Security Agency
- Open-source

# Ghidra - Disassembler

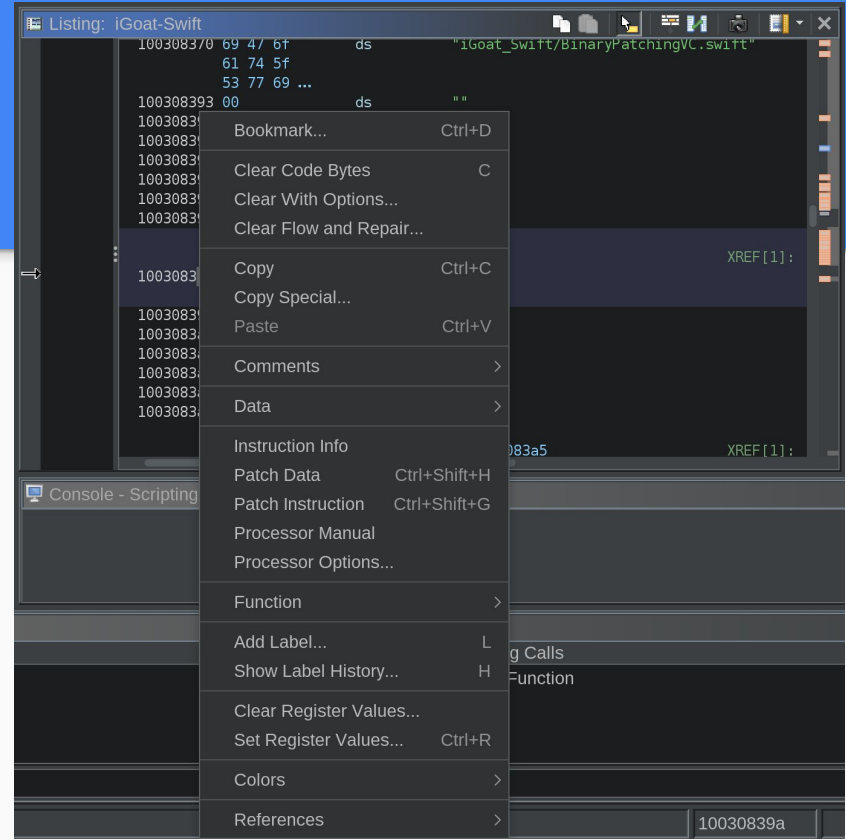


The screenshot shows a window titled "Listing: iGoat-Swift" with a dark background and light text. The window displays assembly code with columns for address, hex bytes, instruction mnemonic, and the instruction itself. A mouse cursor is pointing to the instruction at address 1001e0212.

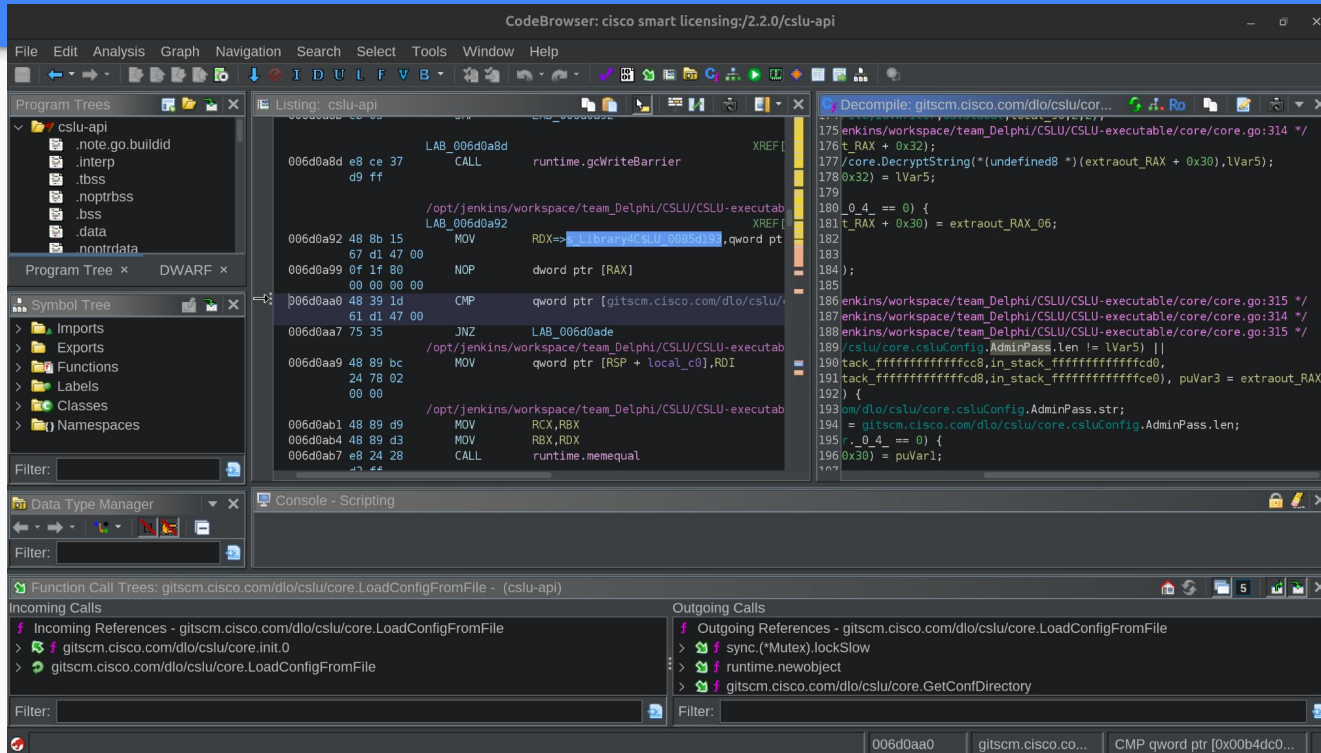
```
Listing: iGoat-Swift
1001e01f5 48 8b 85      MOV     RAX,qword ptr [RBP + local_148]
           c0 fe ff ff
1001e01fc 48 89 85      MOV     qword ptr [RBP + local_170],RAX
           98 fe ff ff
1001e0203 48 89 8d      MOV     qword ptr [RBP + local_178],RCX
           90 fe ff ff
1001e020a 48 89 4d c0   MOV     qword ptr [RBP + local_48],RCX
1001e020e 48 89 45 c8   MOV     qword ptr [RBP + local_40],RAX
1001e0212 48 8d 3d      LEA    RDI,[s_root_10030839a]
           81 81 12 00
1001e0219 be 04 00     MOV     ESI,0x4
           00 00
1001e021e ba 01 00     MOV     EDX,0x1
           00 00
1001e0223 e8 cc 86     CALL   _$sS21_builtinStringLiteral17utf8CodeUnit
           11 00
1001e0228 48 8b bd      MOV     RDI,qword ptr [RBP + local_178]
           90 fe ff ff
1001e022f 48 8b b5      MOV     RSI,qword ptr [RBP + local_170]
           98 fe ff ff
1001e0236 48 89 85      MOV     qword ptr [RBP + local_168],RAX
           a0 fe ff ff
```



# Ghidra - Patch Binary



# Ghidra - Full Screenshot



# Questions?

Thank you!