Brunnhilde Installation and User Guide

Niamh Murphy

Digital Preservation Librarian, University College Dublin

INTRODUCTION TO BRUNNHILDE 1

INSTALLING BRUNNHILDE 2

macOS Installation 2
Step 1: Install Homebrew 2
Step 2: Install Python Using Homebrew 2
Step 3: Install Siegfried 2
Step 4: Install Brunnhilde 3
Windows Installation 3
Step 1: Install Python 3
Step 2: Install Siegfried 4
Step 3: Install Brunnhilde 4
Installing Additional Dependencies (macOS) 4
Configuring ClamAV (macOS) 5
Installing Additional Dependencies (Windows) 6
Configuring ClamAV (Windows) 7

USING BRUNNHILDE 8

Basic Usage 8
Running Bulk Extractor 9

CONCLUSION 9

SPECIAL THANKS 9

Introduction to Brunnhilde

Brunnhilde is a valuable tool for digital preservation professionals, created by Tessa Walsh. It leverages Richard Lehane's Siegfried software to perform file format identification and produce comprehensive reports that support triage, arrangement, and description of digital collections.

The primary benefit of Brunnhilde is its ability to output user-friendly HTML reports summarising data in relation to file formats, dates, versions, errors, and duplicates, etc. These reports also include clickable links to PRONOM, an online resource for detailed information on file formats, from The National Archives (TNA).

In addition, Brunnhilde can perform virus scans with ClamAV and scan for sensitive information using Bulk Extractor.

This guide will take you through the installation process on both macOS and Windows, outline how to install the necessary and additional dependencies, and show you how to use Brunnhilde effectively.

Installing Brunnhilde

Before you can install Brunnhilde, you must ensure that Python and Siegfried are installed. This guide covers how to install Python and Siegfried for macOS and Windows followed by instructions for installing Brunnhilde and its additional dependencies.

macOS Installation

Step 1: Install Homebrew

Homebrew is a package manager for macOS that simplifies the installation of software. To install Homebrew, open your Terminal and enter:

```
/bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

Step 2: Install Python Using Homebrew

Once Homebrew is installed, use it to install Python. Enter the following command in the Terminal:

```
brew install python
```

Step 3: Install Siegfried

Install Siegfried using Homebrew:

```
brew install siegfried
```

Or you may have to run:

```
brew install richardlehane/digipres/siegfried
```

Step 4: Install Brunnhilde

Once Python and Siegfried are installed, install Brunnhilde via Python's package manager `pip`:

```
pip3 install brunnhilde
```

Due to a Python update, you may have to include the following flags:

```
pip3 install brunnhilde --user --break-system-packages
```

To upgrade Brunnhilde to the latest version use:

```
pip3 install brunnhilde --upgrade
```

Or you may have to include:

```
pip3 install brunnhilde --upgrade --user --break-system-packages
```

Windows Installation

Step 1: Install Python

1. Download the latest version of Python for Windows from https://www.python.org/downloads.
2. Run the installer, ensuring you select "Add Python to PATH" during installation.
3. After installation, verify Python is working by opening Command Prompt and typing:

   ```
   python --version
   ```

Step 2: Install Siegfried

1. Download Siegfried for Windows from https://www.itforarchivists.com/siegfried.
2. Once the download is complete, extract the ZIP file to a directory of your choice (e.g. C:\Siegfried).
3. Add Siegfried to PATH.
    a. Right-click on the Start Button.
    b. Select System.
    c. Click on Advanced System Settings and then Environment Variables.
    d. Find the Path variable, select it, and click Edit.
    e. Click New and add the path to the Siegfried folder (e.g. C:\Siegfried\siegfried_1-11-1_win64).
    f. Click OK.
    g. Verify Siegfried is working by opening the Command Prompt and running:

    ```
    sf -update
    ```

Step 3: Install Brunnhilde

Install Brunnhilde via Python's package manager `pip`:

```
pip install brunnhilde
```

To upgrade Brunnhilde to the latest version use:

```
pip install brunnhilde --upgrade
```

Installing Additional Dependencies (macOS)

To fully utilise Brunnhilde's capabilities for reporting of directories and disk images, including virus scanning and scanning for sensitive data, you will need to install the following additional tools:

1. Bulk Extractor: A digital forensics tool designed to extract sensitive information.
2. ClamAV: An open-source antivirus engine.
3. SleuthKit: A digital forensics tool designed to investigate disk images.
4. Tree: Reports on directory structure.

In macOS, these can all be installed using Homebrew:

```
brew install bulk_extractor
brew install clamav
brew install sleuthkit
brew install tree
```

Configuring ClamAV (macOS)

1. Create the necessary configuration file:

   ```
   mv /opt/homebrew/etc/clamav/freshclam.conf.sample
   /opt/homebrew/etc/clamav/freshclam.conf
   ```

2. Navigate to the `freshclam.conf` file using Finder:

   Press `Command + Shift + G` in Finder and enter:

   /opt/homebrew/etc/clamav/freshclam.conf

3. Open the `freshclam.conf` file using a text editor and edit from the following:

   ```
   # Comment or remove the line below.
   Example
   ```

   to:

   ```
   # Comment or remove the line below.
   # Example
   ```

4. Once the aforementioned is complete, update the database by running freshclam in the terminal:

   ```
   freshclam
   ```

ClamAV is now ready to be used with Brunnhilde.

Installing Additional Dependencies (Windows)

As Windows support is limited, the scanning of disk images is currently not available. However, reporting of directories, including virus scanning and running bulk-extractor, should work without issue. With this in mind, to utilise Brunnhilde to capacity within a Windows environment, you will need to install the following additional tools:

1. Bulk Extractor: A digital forensics tool designed to extract sensitive information.
2. ClamAV: An open-source antivirus engine.

To install Bulk Extractor on Windows, follow these steps:

Install Bulk Extractor V1.5.0

1. Go to the git repository for Bulk Extractor (https://github.com/simsong/bulk_extractor) and navigate to the Wiki installation notes.
2. Install V1.5.0 using the Windows GUI at: https://digitalcorpora.s3.amazonaws.com/downloads/bulk_extractor/bulk_extractor-1.5.0-windowsinstaller.exe
3. Ensure that 64-bit configuration and Add to path are both selected.
4. Click Install.
5. Once installation is complete, click Close.

Update to Bulk Extractor V.2.0.0

1. For V2.0.0 of Bulk Extractor, download the ZIP file at : https://digitalcorpora.s3.amazonaws.com/downloads/bulk_extractor/bulk_extractor-2.0.0-windows.zip
2. Once you have extracted the V2.0.0 bulk_extractor.exe file from the ZIP folder, you can replace the V1.5.0 bulk_extractor.exe file with it.

Note: According to Bulk Extractor's README, V2.0.0 is the latest version offering Windows support.

Test Bulk Extractor

1. Open Command Prompt and run:

   ```
   bulk_extractor -h
   ```

This should display help information for Bulk Extractor, confirming it's installed properly.

To install ClamAV on Windows, follow these steps:

Download ClamAV for Windows

1. Go to the official ClamAV downloads page (https://www.clamav.net/downloads).
2. Download the latest Windows release in a ZIP file.
3. Once the download is complete, extract the ZIP file to a directory of your choice (e.g. C:\ClamAV).

Add ClamAV to PATH

1. Right-click on the Start Button.
2. Select System.
3. Click on Advanced System Settings and then Environment Variables.
4. Find the Path variable, select it, and click Edit.
5. Click New and add the path to the ClamAV folder (e.g. C:\ClamAV\clam-1.4.1.win.x64).
6. Click OK.

Configuring ClamAV (Windows)

1. Create the necessary configuration file:

   ```
   move C:\ClamAV\clam-1.4.1.win.x64\conf_examples\freshclam.conf.sample C:\ClamAV\clam-1.4.1.win.x64\freshclam.conf
   ```

2. Navigate to the `freshclam.conf` file located within the ClamAV folder (e.g. C:\ClamAV\clam-1.4.1.win.x64).

3. Open the `freshclam.conf` file using a text editor and edit from the following:

   ```
   # Comment or remove the line below.
   Example
   ```

   to:

   ```
   # Comment or remove the line below.
   # Example
   ```

Update the database by running freshclam in the command line:

1. Open the Command Prompt.
2. Navigate to the ClamAV directory:

```
cd C:\ClamAV\clam-1.4.1.win.x64
```

3. Run the freshclam.exe to update the database:

```
freshclam.exe
```

ClamAV is now ready to be used with Brunnhilde.

Using Brunnhilde

Once Brunnhilde and its dependencies are installed, you can begin using it to scan directories or disk images, and generate reports.

Basic Usage

To run a basic scan with Brunnhilde, open your command line or terminal and enter:

```
brunnhilde.py /path/to/input_directory /path/to/output_directory
```

This will:

1. Run a virus scan using ClamAV.
2. Run Siegfried to scan the input directory or disk image.
3. Generate a set of reports in the specified output directory.

The output includes:

- A HTML report with information on file formats, versions, errors, duplicates, etc.
- A folder with CSV reports for detailed analysis (e.g. duplicate files, format versions, MIME types).
- A logs folder containing virus scan and other logs.
- A Siegfried report in .csv format.
- A file tree in .txt format.

Running Bulk Extractor

To include Bulk Extractor for scanning sensitive information like credit card numbers and social security numbers, use the `-b` flag:

```
brunnhilde.py -b /path/to/input_directory /path/to/output_directory
```

This will generate an additional folder containing reports on personally identifiable information (PII).

Conclusion

Brunnhilde is a valuable tool for anyone working in digital preservation.
Its ability to generate comprehensive and accessible reports aids the management and understanding of digital collections. The inclusion of virus scanning and sensitive data identification further enhances its value, providing both technical and security insights.

Following this guide, you should now have a fully operational installation of Brunnhilde on you macOS or Windows machine. I hope you've found it helpful.
If you need any further information or clarification, please feel free to contact me at niamh.e.murphy@ucd.ie.

Additionally, for a macOS installation guide using Python and Homebrew from 2022, please refer to this DPC blogpost, as the ClamAV configurations varies slightly.

Special Thanks

A special thank you to Tessa Walsh for creating and continuing to develop this fantastic piece of software, to Kieran O'Leary for his invaluable advice, and to Red Green Repeat: Adventures of a Spec Driven Junkie for their ClamAV installation guide .