bulk_extractor 2.2 Current Operating Guide

bulk_extractor project

July 2026

Scope

This guide describes the supported 2.2 command-line workflow.

Build from source

Clone the repository or download a release from `https://github.com/simsong/bulk_extractor/releases`. Install the platform dependencies described by the repository README, then run:

```
bash bootstrap.sh
./configure
make
make check
sudo make install
```

The README is the canonical build and platform-support reference. Native Windows builds are not supported. The repository's Windows MinGW workflow cross-compiles a standalone `bulk_extractor64.exe` on Ubuntu for each pull request. Download it from the Windows workflow artifact; it is not a signed installer or release distribution. The workflow verifies the PE imports and runs the executable on a Windows runner against a Unicode-named input file.

Run a scan

Choose a new output directory and an input image or file:

```
bulk_extractor -o output mydisk.raw
```

Use `bulk_extractor --help` on the installed binary as the authoritative reference for options and scanner settings. A scan produces feature files, histograms where enabled, carved content where applicable, and DFXML run metadata in the output directory.

Develop scanners

The maintained scanner contract is `https://github.com/simsong/bulk_extractor/blob/main/doc/scanner_api.md`. Read it before changing a built-in scanner or starting a loadable scanner, and begin from `doc/scanner_template.cpp`. It documents the 2.x lifecycle, concurrency requirements, feature-recorder rules, and module factory.