SCReeD Dataset Declaration Form

Dataset name:\* Peeler

Dataset version:\* 1

Dataset URL:\* https://github.com/CSCRC-SCREED/peeler

Creation date:\* 28/09/2022

Last update:\* 28/09/2023

Author(s):\* Muhammad Ejaz

Author(s) affiliation(s): Data61CSIRO

Author contact(s):\* ejaz.ahmed@data61.csiro.au

Keywords:\* ransomware, ransomware families, provenance data

Description/background:\* We collected active ransomware samples from more than 60 ransomware families. Each ransomware sample was executed in a VirtualBox and then manually labeled by each family type. We developled our own tool to collect provenance data which will be made available here. We ran each ransomware sample for ten minutes or until all user files were encrypted. It took more than 90 days to run all samples and collect data.

Dataset funding: Click or tap here to enter text.

Attribute details:\* provenance data

Intended target (if specified): Click or tap here to enter text.

Format:\* json

License:\* GNU v3

Standard compliance: Click or tap here to enter text.

Type:\* Malware

Size:\* 66 KB

Availability: Click or tap here to enter text.

Data status: Click or tap here to enter text.

Data provenance:\* VM

Source computing infrastructure:\* Simulated environment

Accompanying program(s)/script(s): Click or tap here to enter text.

Software installer or VM for replication: Click or tap here to enter text.

Generated or captured via:\* Software

Category/categories:\* ML training data

Published in: Click or tap here to enter text.

Open research question(s) (if any): Click or tap here to enter text.

Potential use case(s) or application area(s): Click or tap here to enter text.

Data access control:\* Global access

Data retention period:\* 5 years

Data validation/checksum:\* No

GDPR compliance:\* Yes

Consent: Yes

Ethics approval: N/A

Ethics considerations: Click or tap here to enter text.

\* *denotes mandatory field*

I confirm that I have read, understood, and agreed to the submission guidelines, policies, and submission declaration.

I confirm that the contributors of the dataset have no conflict of interest to declare.

I agree to take public responsibility for my dataset’s contents.



Signature of Corresponding Author Date: 28/09/2023

(signed on behalf of all contributors)