Watson - Extract forensic information from Linux desktop environment search software

Degree programme: BSc in Computer Science | Specialisation: IT Security
Thesis advisor: Prof. Dr. Bruce Nikkel
Expert: Dr.Wolfgang Joachim Kaltz (Camptocamp SA)

Forensic investigators face significant challenges due to increasing cybercrime and evolving technologies. Extracting data from damaged or destroyed devices, from multiple sources and locations is difficult. Watson aims to aquire another source of information for forensic investigations by providing a tool for extracting user data from Linux desktop environment search engines and their artifacts.

Overview

Timestamps play a crucial role in reconstructing past events and creating timelines for forensic investigations. These timelines help forensic investigators understand the sequence of actions taken on a digital device, such as a computer, a server, IoT devices like smartwatches or smart home gadgets, or any other digital device. It can be used to frame a situation during a cyber incident or determine the correct course of a crime. Additionally a timeline can be used as evidence in a legal case.

Shifting the focus to Linux and Unix-like system architectures, KDE and GNOME stand out as popular desktop environments that offer file indexing tools for a fast system wide search utility given by search engines. The extensive collection of user data by search engines and their indexing tools, which is essential to their efficient functioning, represents a valuable area for in-depth analysis and investigation.

Goal

GNOME employs tracker 2 and tracker 3, KDE employs Baloo and Akonadi as their search engines and indexing tools. Watson analyzes the indexed information collected by these tools, starting from the initial setup of the operating system. Watson evaluates various data points, including timestamps, file names, file paths, file sizes, plain text content, contact details, notes, mail and calendar content.

This thorough analysis enables Watson to generate a detailed timeline output, offering valuable insights for forensic investigations.

Result

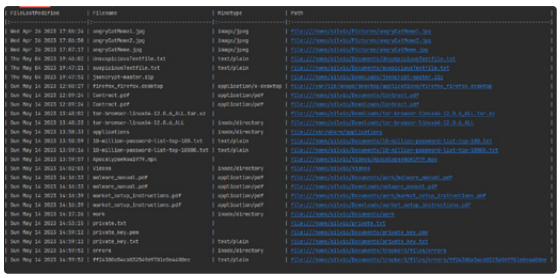
Watson generates an output in the form of a detailed timeline, encompassing timestamps, files, and paths. In addition to that Watson offers convenient features, including defining a time frame, searching for specific files or mime types, refining searches with keywords, generate a list of related contacts for associated files and displaying mails affiliated to the user. A notable feature is to display user activity statistics. Watson, developed with Python, is an open- source project that can be effortlessly installed with ‘pip install’ and is designed to analyze data from files not only from running systems, but also for instance images of dead systems or mounted drives.

Future

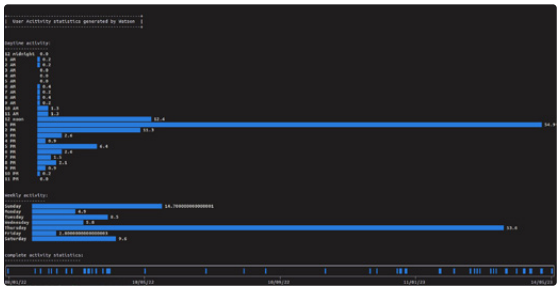
Having the functionality to recover deleted data that was also discarded by the indexing tools would offer further and deeper forensic opportunities. There are existing tools that seem promising, but do not yet completely provide the required functionality. An additional extension of Watson to analyze not only Linux and Unix-like systems, but also Windows and Mac desktop environments, which are more commonly used, would make Watson an even more comprehensive forensic tool.

Christina Lena Blumenthal

Silvio Hoppen



Timeline output analyzing the KDE desktop environment search engine



Statistics output about the user activity