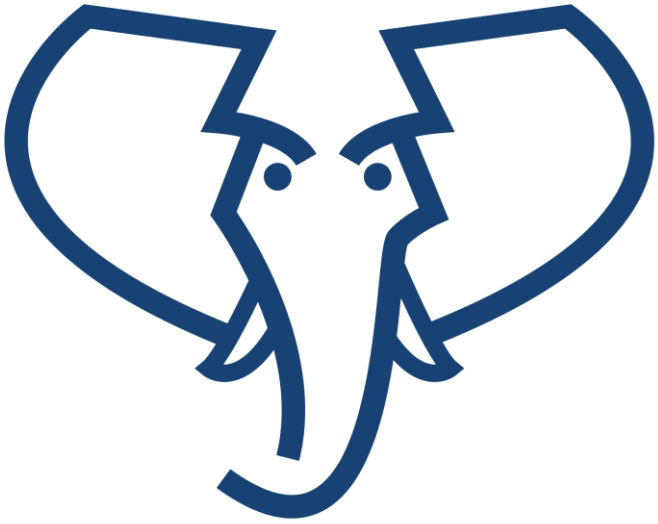
Forensic measures in Hansken



Hansken

The open digital forensic platform

Contents

Please note that these HANSKEN FORENSIC MEASURES provides general documentation, the actual forensic measures implemented by any organization depend on the actual settings and legal demands on using Hansken.

1. Introduction

The amount of data and data sources to be investigated in criminal cases is increasing at a rapid rate. To increase the effectiveness and speed of this investigation, the Netherlands Forensic Institute (NFI) has developed the forensic platform Hansken.

The purpose of Hansken is to make information from digital evidence such as telephones and computers quickly and easily accessible to users (with or without specific IT knowledge) who are authorized to examine such evidence. Hansken does this by structuring and indexing the data. Anything that may be relevant can then be searched for, for example words and names or properties of digital traces such as files, chat messages, emails or photos whether or not taken with a particular camera.

With Hansken, a user can continue to filter the search results until, out of those millions of traces, a manageable selection remains, the traces of which can be viewed one by one.

General information and explanations about Hansken are available in the Subject Annex Hansken[1].

In more and more criminal cases, traces found with Hansken are used as evidence. This requires a solid forensic foundation: the material must be processed transparently and the reported traces must be traceable to the source (e.g. the seized telephone or computer).

The use of Hansken includes several safeguards for the purpose of this foundation. This information sheet explains these safeguards based on the steps in the investigation process:

1. obtain source data
   - secure source data in evidence files
2. basic processing with Hansken
   - create a case
   - upload evidence files
   - extract traces
3. use for case investigation
   - trace search
   - follow-up processing of traces

Here a distinction is made between the processing of the evidence[2] itself (chain of evidence) and the recording of this process (chain of custody).

For the capture of user actions, Hansken offers the possibility to send audit information to an external audit system (e.g., a SIEM, a Security Information and Event Management). This function forwards information about all requests to Hansken, including the date and time, user and details of the respective request.

1.1. Hansken investigation environments

Investigative services generally run their own Hansken environments in their own data centers. Some investigation services use different Hansken environments serviced by the NFI or other organizations.

[1] The Subject Annex Hansken is available in Dutch at https://www.forensischinstituut.nl/vakbijlagen.
[2] Hansken extracts traces (e.g., emails and photos) from the evidence in several steps. This processing of evidence does not change the evidence itself.

Employees of their own organization manage their own Hansken environment. A distinction is made between system administrators (infrastructure, hardware, underlying software and Hansken platform) and functional administrators (use of Hansken for specific case investigations).

2. Acquisition of source data

The source data is the data collection that has been secured from the seized devices or has been demanded from, for example, providers.

2.1. Securing source data

The source data is secured in evidence files, also called *forensic images* or *images* for short. In addition to the secured source data, such an evidence file often contains information about the securing process, such as information about the secured device, the time of securing and the persons involved. To ensure the integrity of the source data, hashes or digests are calculated when the source data is secured. These file characteristics are reproducible and are often also recorded in the evidence file.

This securing is done by employees of the investigative services and is separate from (the use of) Hansken.

2.1.1. Processing of evidence

An evidence file is normally not changed during an investigation. However, an evidence file can be moved or copied. The file characteristics of the source data in the moved or copied evidence file must match the original file characteristics.

2.1.2. Recording of user actions

Securing and recovering data is done through established procedures and an official report is made by the investigating officer involved. This report includes the established file characteristics.

3. Basic processing with Hansken

3.1. Creating a case

To make the right data available to the right users, evidence files are grouped by case. For each case one trace collection is created. To do this, a functional administrator creates a case[3] in Hansken.

3.1.1. Processing the evidence

When creating a case in Hansken, no evidence is involved. Thus, Hansken does not process evidence in this step.

3.1.2. Recording of user actions

The actions performed by the functional administrators take place via fixed procedures and should be recorded in a separate case administration system.

[3] Because Hansken is also used for research and development, a case in Hansken is called a project.

Per case, Hansken records the following data[4]:

- unique project number (within Hansken environment[5]);
- name of the case;
- creation date and time and responsible administrator;
- linked evidence files.

If configured, the actions involving the creation of a case are recorded in an audit system outside of the Hansken environment.

3.2. Placing evidence files in Hansken

Evidence files normally reside in the digital investigation environment of the investigative agency. In order to process them with Hansken, Hansken must gain access to these files. Currently, this is done by copying the evidence files to a storage area specifically set up for Hansken within the Hansken environment. The option exists to upload the files to Hansken via the internal network.

3.2.1. Processing of evidence

When placing an evidence file in Hansken, the recorded source data in the evidence file is compressed and encrypted (for fast and secure processing) and stored in a Hansken-specific evidence file format (an *NFI image*). The source data itself does not change by this conversion. For each evidence file, Hansken records who placed it in Hansken and when that occurred.

If file attributes of the source data are available in the evidence file, they are copied into Hansken's records[6]. Hansken provides the ability to check the file digests of the copied data after it has been placed in Hansken.

3.2.2. Recording of user actions

Depending on the organization, different procedures are followed for placing evidence files in their Hansken environment. This is normally done by a digital expert or a functional administrator of Hansken. These actions performed take place via fixed procedures and are recorded per case in a separate case administration system.

For each evidence file, Hansken records the following data:

- unique ID for the evidence file (within the Hansken environment);
- start date and time and responsible administrator for placing in Hansken;
- acquisition file characteristics (digests);
- file characteristics (digests) after placing in Hansken.

If configured, the actions involving the uploading of evidence files are recorded in an audit system outside the Hansken environment.

3.3. Extraction of traces

Hansken makes source data in evidence files transparent by extracting traces (e.g. files, emails, chat messages, photos and documents). A trace consists of metadata (properties of the trace, e.g. sender of an email or name of a file), a keyword index (words that occur in the trace) and the data of the trace

[4] For investigative and management purposes, additional data can be recorded per case in Hansken.
[5] The functional administrators are responsible for recording the link between the unique project number within the Hansken environment and the organization-specific case numbers and case names.
[6] The so-called acquisition hashes are copied into the metadata of the evidence file in Hansken.

itself (the bytes in the source data). Traces can also contain other traces (e.g., a *picture* in an *attachment* to an *email* in an *email archive* in a *folder* in an *evidence file*).[7]

In the extraction process, Hansken repeatedly sends traces to dozens of forensic tools for processing (adding metadata, expanding the keyword index, and/or extracting new traces). Different forensic tools each have their own function in this process (e.g., extract emails from email databases, extract camera information from photos, determine keywords, or calculate file attributes). Hansken places the extracted data into the Hansken search engine.

The trace data is *not* included in its entirety in the search engine. For each trace, Hansken records where and how exactly the trace can be found in which evidence file.

The functional operator configures, starts and monitors this extraction process based on the request from the investigation team. Hansken uses a standard set of forensic tools for this. Based on the request, forensic tools are turned on or off.

If desired, an extraction process can be repeated. In doing so, other settings can be used. Such a re-extraction can replace (e.g. with new or different forensic tools) or supplement (e.g. with forensic tools that were previously off) the data extracted earlier. This feature is also used for filtering of privileged material.

When the extraction process is complete, the administrator checks whether the extraction went well and makes the results accessible for examination.

3.3.1. Processing of the evidence file

An evidence file is considered a trace by Hansken and is therefore offered to forensic tools. If a forensic tool[8] can process a trace, then this is recorded in the metadata of the corresponding trace.

Since millions of traces are often processed during an extraction process, several errors will generally occur in that process. Errors can have various causes, for example because data is incomplete (e.g., a part of an picture or PDF), because the structure of the source data does not match the structure expected by a forensic tool (e.g., an incomplete email database, a new version of a chat log), or due to environmental factors such as system overload. However, Hansken is robust, which means that such errors are caught and captured at the relevant traces.

For each trace, Hansken captures the following information:

- unique ID of the trace (within the case and extraction process[9]);
- unique ID of the superimposed trace;
- forensic tool that extracted the trace, including version details;
- forensic tool that processed the trace, including version details and order of invocation;
- for each property: forensic tool that determined the property, including version details;
- forensic tools that failed to process the trace;
- link to the source data in the evidence file;
- calculated digests.

[7] Hansken also considers and treats an evidence file itself as a trace.
[8] The Hansken platform provides insight into the use of forensic tools. This says nothing about the quality of these tools and reliability of the traces extracted by these tools. The forensic tools that are used with Hansken for the Dutch criminal justice chain have been developed and/or tested by the NFI.
[9] A re-extraction of an evidence file may not yield the same IDs for the same traces, for example, because different/new versions of forensic tools have been used or the order of processing the traces has changed.

The above data make the processing transparent and the traces traceable and reproducible.

3.3.2. Recording of user actions

The functional administrator configures and starts the extraction process.

For each extraction, Hansken records the following data:

- start date and time;
- settings used[10]:
  - general settings for the overall extraction;
  - overview of activated forensic tools;
- settings per activated tool;
- general statistics (e.g., number of traces and number of bytes read);
- statistics per forensic tool:
  - number of invocations, including profile (total computation time, number of bytes read);
  - number of traces created;
  - number of traces processed;
  - number of errors, with a summary of the errors;
  - unused forensic tools (tools that have been activated, but have not processed traces).

The above data allows the extraction to be performed again. The statistics support the control of a (re)extraction and give direction to the (further) development of Hansken and the forensic tools.

If configured, the actions involving the extraction of traces are recorded in an audit system outside the Hansken environment.

4. Use for case investigation

4.1. Searching the trace collection

The traces recorded in the Hansken search engine can be searched manually (via the user interfaces) or automatically (with a script).

Searching can be done in several ways:

- by filtering the metadata (e.g. only images, emails from a certain sender or documents printed in a certain period);
- by using the keyword index (e.g. traces where a certain word occurs).

These search functions can be combined, so that, for example, one can search for emails that contain certain words and were sent in a certain period.

For each trace, it is specified where exactly the trace can be found in which evidence file. If a Hansken user wants to view the trace data, then this trace data (e.g. the email, the office document or the picture) is copied from the evidence file at that moment.

Combining digital traces in Hansken with data outside Hansken takes place outside the Hansken environment.

[10] Settings are recorded when they differ from the default settings.

4.1.1. Processing of evidence

In this investigative step, only the already available trace collection is searched. The traces in this collection cannot be modified.

Users can, however, add notes and labels to traces that are used for investigative purposes (e.g., to record relevant traces or mark traces that require follow-up investigation).

4.1.2. Recording of user actions

For notes, a record is made of who placed the note with a trace and when. This does not apply to labels.

If configured, the actions concerning the searching of trace collections are recorded in an audit system outside the Hansken environment.

4.2. Trace Processing

It is not uncommon for traces made accessible by Hansken to be reprocessed, for example, when new formats are encountered for which Hansken does not (yet) support (e.g. a database of chat messages from a new chat app).

Hansken offers the possibility of placing external traces in Hansken. This typically takes place by digital experts. Traces placed in Hansken in this way are included in Hansken's search engine. These are then also searchable, as are the traces identified with Hansken by the forensic tools. The responsibility for these traces lies with the user who placed them in Hansken. Therefore, these traces can also be modified and overwritten.

4.2.1. Processing of the evidence

The subsequent processing of the evidence (and the recording of this processing) takes place outside of Hansken.

4.2.2. Capture of user actions

Traces that are placed and/or modified in Hansken in this way are marked. In the report of such a trace, it is recorded that this trace was added by a user.

For each trace placed in Hansken in this way, Hansken records the following information:

- user who added and/or modified the trace in Hansken;
- date and time the trace was added and/or edited in Hansken;
- description of the processing of the trace;
- the added, modified and/or deleted properties;
- the old and new value of properties (in case of modifications).

If configured, the actions involving the placement and/or later modification of traces are recorded in an audit system outside the Hansken environment.

Please note that these HANSKEN FORENSIC MEASURES provides general documentation, the actual forensic measures implemented by any organization depend on the actual settings and legal demands on using Hansken.