Combing Through Crimes: Understanding and Enhancing Search in Hansken

Ko Schoemaker

Thesis submitted in partial fulfilment
of the requirements for the degree of
Master of Science
in
Human Computer Interaction.

Utrecht University

Graduate School of Natural Sciences

Supervision by
First supervisor: Dr. Eelco Herder
Second supervisor: Dr. Christof van Nimwegen

2025

Abstract

Hansken is a powerful digital forensic tool, but its search interface is often experienced as unintuitive and cognitively demanding, leading to less effective search behaviour. This study investigates how Hansken users search for information, how their behaviour aligns with theories of information seeking, and how interface design can better support their needs. An alternative prototype interface was developed through design thinking and a mixed-methods study was performed with Hansken users, combining interviews, think-aloud exercises and a survey. Findings show that Hansken users' information needs vary widely depending on role, organisation, and task. Users often indulge in exploratory search behaviour and rely on contextual information to guide new search directions. Data categorisation and overviews support this behaviour. As a query formulation style, natural language interaction are perceived as an enhancement but trust in AI agents remains low. Insights underline the need for search interfaces that support a broad range of information needs, facilitate categorical groupings of data, and reduce mental effort in query formulation. Design implications include implementing faceted search, offering exploratory and focused search modes, and cautiously integrating natural language interaction with mechanisms for building user trust.

Contents

Chapter 1

Introduction

As the amount of digital devices keeps increasing every year, digital evidence has become vital in the field of forensic science (Du, 2020). Both the volume and variability of data generated by electronic devices has rapidly increased. Whenever digital devices are seized following a criminal incident, investigators need to sift through ever-increasing amounts of digital evidence, leading to problems for both people and tools (Daubner et al., 2024). With these enormous amounts of data available, it is difficult to identify criminal activity and the malicious users behind it (Zawoad and Hasan, 2015).

Within the field of forensics, *Digital Forensic science* (DF) is concerned with collecting, processing and reporting on electronically stored information that constitutes digital evidence, carried out by law enforcement or intelligence agencies. Within the DF domain, specialized sub-domains like computer-, mobile- or network forensics exist which each have dedicated *Digital Forensic Tools* (DFTs) to assist investigators in their search through device data while preventing unintended modifications.

Since 2012, the Netherlands Forensic Institute (NFI) has been developing Hansken: a digital forensics search engine for processing and investigating seized digital material in huge quantities (van Beek et al., 2015). Hansken implements a *Digital Forensics as a Service* (DFaaS) model: a cloud service model for digital forensics to build a centralized evidence process system (Du, 2020, p.27; Lee and Hong, 2011). The benefits of cloud-based forensics can be found in moving storage and processing to data centres, thus improving performance and enabling data sharing and collaboration. Another benefit can be found in the lowered level of technical skills required to use it, which leads to a more accessible service without needing digital forensic tools knowledge.

Hansken has been used in hundreds of criminal cases in The Netherlands and abroad, and is able to process and store petabytes of different types of data (Henseler and van Beek, 2023; van Beek et al., 2015, p.20). To interact with the Hansken search engine, two main graphic interfaces are currently available: the *Hansken Tactical Interface* (HTI) and a technical user interface known as *ExpertUI*. HTI can be used for normal search tasks, while the technical interface contains more specialized tools for experts and project settings. In the graphic interfaces, multiple search modes exist for querying information, which are highlighted in section 2.2. Besides these graphic user interfaces, a programmatic interface exists to query information in Hansken by direct scripting using a Python library. The programmatic interface can be used for specialized custom applications

that are not possible through the graphic interfaces, which allows for complex and case-specific queries, or automatization.

While Hansken is seen as an incredibly powerful tool, users claim that certain functionalities are not intuitive enough for an average user to understand without substantial training. Writing the queries is said to take a lot of mental effort, which takes away from focus on the case itself. For novice users, this leads to creating unrefined queries, which yield too broad of a result set, leading to them spending additional time sifting though irrelevant results.

In an earlier unpublished usability study of the Hansken Tactical Interface, it was found that the search interface might not support users adequately in their search journey. Issues include filters that are too complex, the interface showing too much information, having inconsistent views, and not giving the user any information that might inform their next search direction - among other things. Participants of preliminary focus groups have stated that the Hansken training alone - currently through e-learning - is not sufficient for understanding what steps should be taken when starting an investigation using Hansken.

There is room for improvement in supporting users through their search in Hansken, but this can be extended to all forensic tools. It seems that many people believe that current digital forensic tools are limited in terms of ease of use and software engineering (Beebe, 2009, p.29). The interface of tools should be intuitive, not too technical and should be customizable for skilled users. The goal should not be to show data, but to provide information and knowledge.

To understand how users can best be helped in their goals in practice, we must first understand what general goals are and what users like to achieve when using digital forensic tools. While a lot of research has been performed developing methods, tools, processes, procedures and frameworks within the field of digital forensics, there is not a lot known about how digital forensic investigators search and analyse data in practice (Sunde, 2022, p.27; Beebe, 2009, p.25). The analysis process of DF-investigators needs to be understood so digital forensic tools can be adapted to fit their work process, making an investigation more efficient and effective.

An evaluation of forensic work processes in the Hansken Tactical Interface and a new Hansken prototype interface is conducted to answer the following research questions:

> RQ1: What is the information need of Hansken users interacting with the Hansken Tactical Interface?
> RQ2: How do Hansken users search for information in the Hansken Tactical Interface and how do their search behaviours and use of available search capabilities align with theory of information seeking?
> RQ3: How do specific interface features and interaction styles influence cognitive load, perceived usability, and effectiveness of information seeking in a Hansken prototype interface, and what design implications emerge from these findings?
> Sub-RQ3: How are natural language interactions with AI agents perceived in terms of usefulness and trustworthiness for information seeking and task decomposition?

Understanding the information need is done through semi-structured interviews with current users of Hansken. Furthermore, a new prototype interface for Hansken is made, incorporating possible

search enhancements, which will be evaluated with users of Hansken from several investigative agencies in The Netherlands. The results of this research can be used to enhance digital forensic tools to align more with the practical requirements of digital investigators.

In the Literature & Background (2), some basics of digital forensic science are discussed, together with theories of information seeking, search interfaces and design thinking. In Methods (3), the study design and used methods are explained. Additionally, creation of the prototype interface and its evolution through iterations of design thinking is described. Results from the study are shown in the section Results (4). Interpretations of the findings, as well as limitations and future work are discussed in Discussion (5). Lastly, the research questions are answered in Conclusion (6).

Chapter 2

Literature & Background

To understand how search in Hansken works and can be improved, several domains need to be explored and highlighted. The basics of the digital forensic process is highlighted in Digital Forensic Fundamentals (2.1). Next, current search capabilities of Hansken and its interface are discussed in the section Search in Hansken (2.2). Theories of Information Seeking (2.3) are expanded upon, including Exploratory Search (2.3.1), Query Formulation (2.3.2) and Search Goals & Tasks (2.3.3). Then, in Search Interfaces (2.4) different interaction methods are shown and related to Faceted Search (2.4.6), the Search Results Page (2.4.7) and Visual Interfaces (2.4.8). Additionally, Design Principles (2.4.9) and Theory Implications for Exploratory Search (2.4.10) are described, concluding on what interface elements can help users in their search. Information from these sections will ultimately inform the design of a new search interface prototype through design thinking, described in section Prototyping & Design Thinking (2.5).

2.1 Digital Forensic Fundamentals

In digital forensics, several different roles can be identified, among others: tactical investigators, digital experts and analysis. A *tactical investigator* is directly involved in investigation of criminal cases, collecting information in an attempt to bring them to a successful conclusion. They have has knowledge of a specific case they are working on, the persons involved and events as they occurred. A *digital expert* (or digital investigator) is an expert in investigating digital devices for finding digital evidence, but does not generally have detailed knowledge about a specific criminal case. An *analyst* seeks trends across multiple cases, for example to uncover networks of people or to investigate new trends in criminal methodologies.

A criminal investigation generally starts with some incident which is of interest to some investigative agency, like for example the police. An investigation can now be started with the goal of gathering sufficient information to determine if some person or entity should be indicted for the incident (Andersen, 2019, pp.5–7). The investigation consists of creating hypotheses about the incident, collecting data to inform these hypotheses and then evaluating them, until hopefully one hypothesis can be proven sufficiently to explain the incident beyond reasonable doubt.

It is important that professionals in the field of digital forensics also methodologically undergo a process of identifying, preserving, analysing and presenting digital evidence. To allow for this,

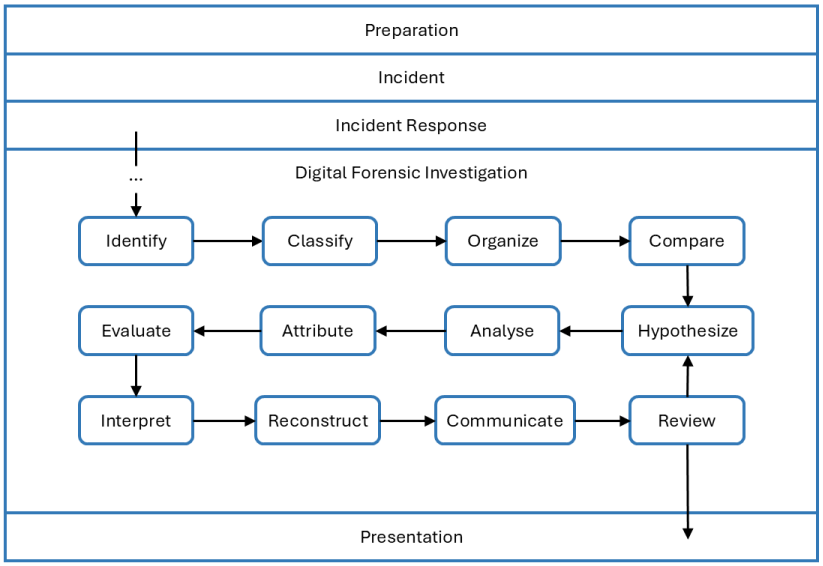


Figure 2.1: Overview of IDFPM with some steps pertaining digital forensic investigation written out. Steps regarding the extraction and data preparation phases in the digital forensic investigation section are omitted, but can be found in the original paper by Kohn et al. (Kohn et al., 2013)

many different investigation models have been proposed, formalizing the steps in a digital forensic process (Yusoff et al., 2011; Zareen et al., 2024). One such model is the *Integrated Digital Forensic Process Model* (IDFPM) has been applied to Hansken in previous research and is shown in figure 2.1 (Kohn et al., 2013; van Baar et al., 2014)

In broad terms, the work process can be divided into preparation, incident, incident response, digital forensic investigation and presentation (Kohn et al., 2013).

In the preparation phase, the investigative organisation prepares itself to deal with different types of incidents. When an incident occurs, investigations are started and (digital) evidence is secured. After that, the digital forensic investigation starts, the focus of IDFPM. A hypothesis is constructed in this phase. Findings from the investigator - together with a report of the investigation process, chain of evidence, chain of custody - can be presented in court, where a decision can be made regarding the person or entity to whom the incident was attributed.

The digital forensic investigation is the heart of IDFPM (Kohn et al., 2013, pp.11-12). This domain can be further subdivided as shown in figure 2.1. After extraction and data preparation, an investigator identifies the incident in the digital data, classifies similar digital evidence, organizes this evidence and compares classifications to similar past incidents.

As with any kind of investigation, hypotheses about the crux of the incident should be constructed and subsequently tested for validity, repeatedly if need be. A hypothesis is analysed to see if it might hold, attributed to an event or to persons, evaluated and interpreted to create meaningful statements in the legal context. Then, events as the investigator believes they occurred are reconstructed, which is communicated to other interested parties and lastly reviewed once more against the hypothesis. In this crucial last step, it is decided if the hypothesis passes or fails; if the incident can be explained with a valid hypothesis with sound, relevant, admissible digital evidence, we

proceed to the presentation phase. If not, a new hypothesis must be constructed.

The IDFPM is just one of many different investigation models that have been created over the years. It seems like groupings in other models explaining digital forensic investigations are generally rather similar, mostly changing in the level of detail they show (Yusoff et al., 2011; Zareen et al., 2024).

2.1.1 Digital Forensic Tools

When digital devices like computers, mobile phones or servers are used for criminal purposes, they constitute digital evidence and the contents of those devices might contain clues that might help prove how a crime happened. Since the 1990s, *Digital Forensic Tools* (DFTs) have been in development to capture disk copies of these devices, creating so called *forensic images* (Zareen et al., 2024, p.7). Since then, a lot has changed in variety, complexity and volume of digital evidence, and digital forensic tools have grown with these new demands.

Digital evidence encompasses many forms: files, file fragments, digital audio/video, messages, any electronic information stored or transmitted in binary form (Du, 2020, pp.8–9). Hansken is made to handle seized digital devices, while other DFTs might be more suited for cloud forensics, IoT forensics, network forensics. In a real-world setting, different DFTs for different types of evidence can be used concurrently for a single case because they all contain certain functionality needed for processing a specific type of evidence. In fact, among digital forensic professionals the best approach to solve this coverage problem is to "buy one of every tool on the market" (Garfinkel, 2010, p.S67).

2.2 Search in Hansken

Hansken is a digital forensic tool for law enforcement and intelligence agencies, designed to index, search and analyse digital evidence extracted from seized digital devices like phones or computers (Henseler and van Beek, 2023). Hansken is capable of processing and storing petabytes of different types of data from hundreds of different devices, while still ensuring data integrity and confidentiality. Hansken supports cases with more than a thousand seized devices, over 100 terabytes of data and over 100 million digital artifacts, called *traces* (e.g. pictures, browser history, emails and chat messages(van Beek and Henseler, 2023, p.104).

After physical digital devices are seized, forensic copies are made. These copies - called *forensic images* - contain all data and meta-data of the original device, and can subsequently be examined by recovering deleted files, file carving or other methods to create insights of the bytes in the image (van Baar et al., 2014, p.S56). Traces are created by applying forensic tools, and traces are processed, enriched and undergo a full-text index (van Beek and Henseler, 2023, p.104). After this, data will be made visible to investigators[1].

[1]Depending on the nature of the investigation, the exact processes or order may be different than described. Legal processes like marking of privileged communication could precede insight into the data by investigators.

The digital data is indexed in Elasticsearch[2] and can be searched through using the *Hansken Query Language* (HQL). More information on HQL can be found in section 2.2.1. Traces, extracted from forensic images, are part of the Hansken trace model: a definition of how digital artifacts should be described in order to be indexed by the search engine behind Hansken (van Beek and Henseler, 2023, p.105). In the Elasticsearch document store, these traces and their meta-data are saved as documents.

A trace contains mandatory properties of *name* and *id*, but can also have one or multiple *types* (e.g. email, file or picture) (van Beek and Henseler, 2023, p.105). In addition, a trace can have *contents*: one or more interpretations of the actual file or text contained in the document.

Hansken is designed as a service with an open interface, meaning any organisation that can access the Hansken engine can create their own interface according to their own specifications (van Beek et al., 2015, p.34). There are currently also two existing graphical user interfaces, designed for different user groups; a technical interface (ExpertUI), aimed at digital experts, and the Hansken Tactical Interface (HTI), designed for both tactical investigators and digital experts. The technical interface contains more powerful querying capabilities, configurable search result presentation and in-depth data overviews (van Beek et al., 2015, p.34). HTI is made to suffice for general search queries and for all types of users.

2.2.1 Hansken Query Language

The Hansken Query Language (HQL) is a powerful domain-specific full-text search query language specifically made for interacting with traces in Hansken. HQL supports powerful data query capabilities and does not offer data manipulation (as a language like SQL might). This is by design, because it is imperative data integrity should be guaranteed in forensic tooling and disk images should not be adjustable to guarantee integrity.

HQL is based on the *Lucene Query Language*[3] (LQL), a full-text query language (Rijksoverheid, 2020). Like LQL, HQL allows term search, pattern matching methods like wildcards, regular expressions and ranges, context queries like phrase and proximity queries, and fielded search, which can all be combined in boolean queries (Baeza-Yates and Ribeiro-Neto, 1999, pp.101-106). For Hansken, both full text searches can be done to match terms or phrases - like `Jennifer` for matching all places that name occurs - , and fielded data can be used for matching a specific field in meta-data -like `photo.createdOn:01-02-2025` for finding all photos that match this meta-data constraint. Besides text-based queries, HQL also allows for numeric, date and geolocation queries (van Beek and Henseler, 2023, pp.108–109)

Through implementations of boolean search it is possible to create relatively complex queries. Additional capabilities for data processing are nested queries (utilizing the hierarchy of traces that relate to other traces) and facets (showing aggregated amount of results for a selected field), which are used together with HQL queries to generate different standard data selections and views in the tactical user interface (see 2.2, View Selection). In the main search interface (figure 2.2, Search View), queries can be directly executed by the user but are also used as an intermediary format

[2] https://www.elastic.co/elasticsearch
[3] https://lucene.apache.org

when querying using a different interaction technique. Section 2.2.2 will highlight the different interaction techniques further.

2.2.2 Hansken Tactical Interface

The Hansken tactical interface is the GUI to the Hansken search engine aimed at tactical investigators but also suitable for digital experts. Different views can be selected from the panel on the left side (see figure 2.2), which show the user a predefined selection of documents (for example only pictures, only location data or only messages) in a predefined view (for example a gallery view for photos, a map view for locations or a conversation view for messages). The data selection in this view can be further refined using filters. In figure 2.3 some of these views are shown.

All document selections are made by executing HQL queries to the Hansken back-end, which are then shown to the user in a particular way. When a filter is added, a condition is added to the HQL query and the result set is updated.

A user may choose to execute a manual search from scratch in the search view (seen in figure 2.3a and 2.2). A user can choose from one of four search modes: searching for keywords, searching for documents in a user-defined period of time, using a form-interaction to search for a specific type of trace with field values that can be set, or to entering an HQL query directly. Multiple modes can be combined to specialize a query further (combining a keyword query with a time query lets the system only retrieve documents matching the keyword that are also from the specified time period).

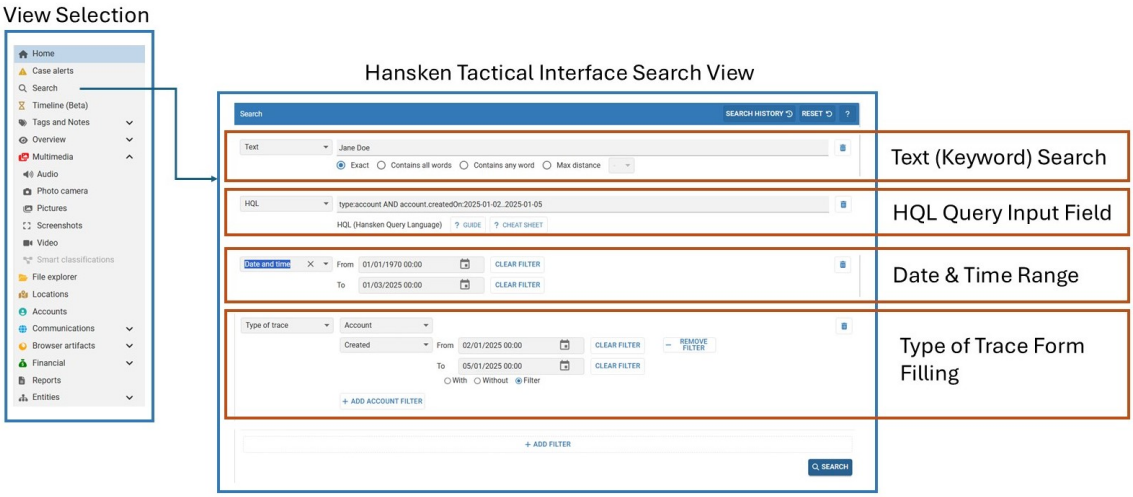
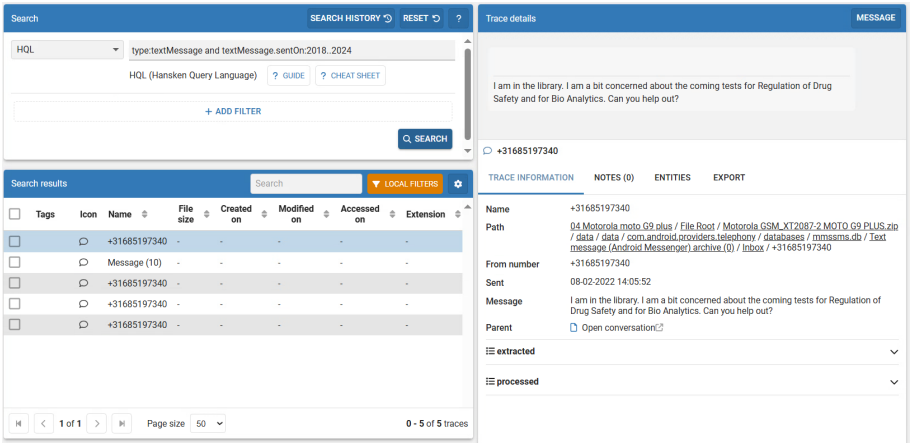
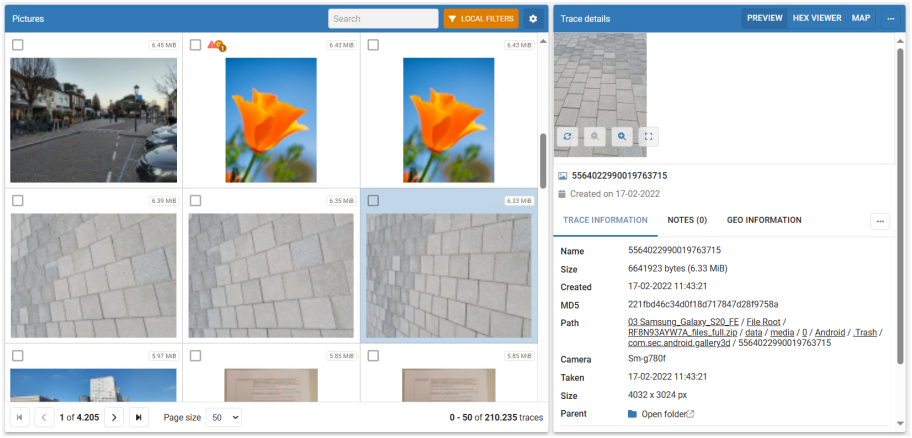


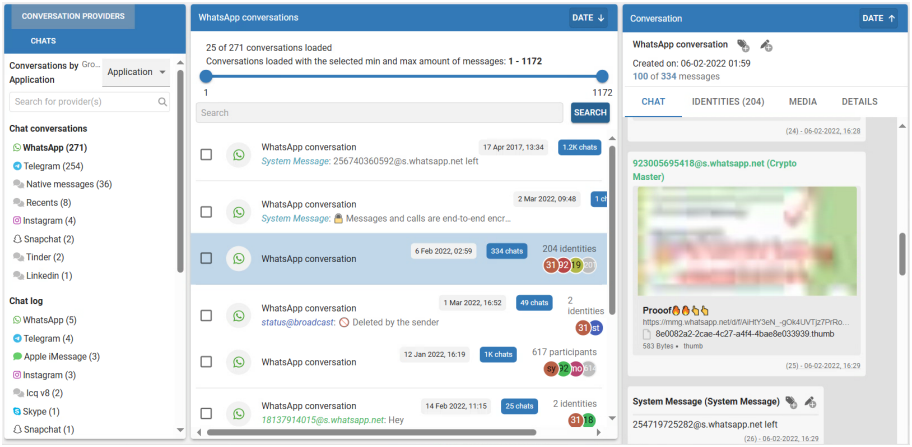
Figure 2.2: On the left, a panel with different views and different data type selections for users to choose from. The HTI search interface consists of four query modes: keyword search, HQL search, a date and time range with a calender picker. The modes can be combined to construct a full query. They can be added with the 'add filter' button at the bottom and removed with the bin-icon to their right. Upon clicking 'search', their full query is executed.



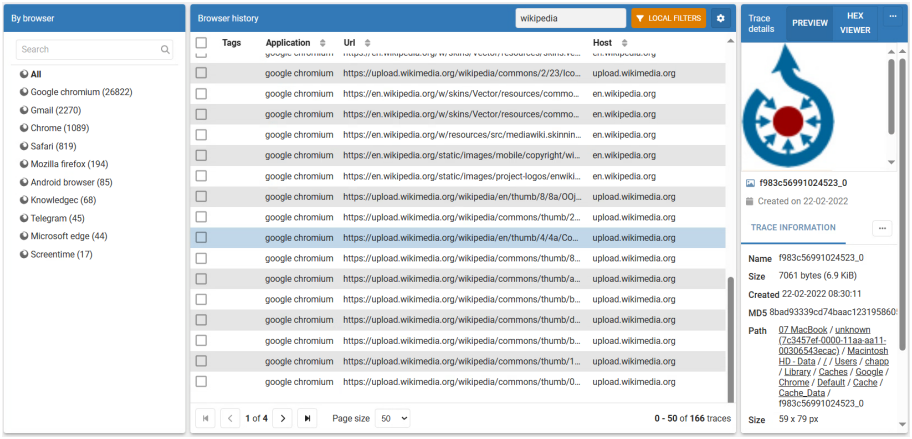
(a) Search view showing a list of traces as result. When an HQL query is executed and a trace is selected from the result list, it is shown on the right side.



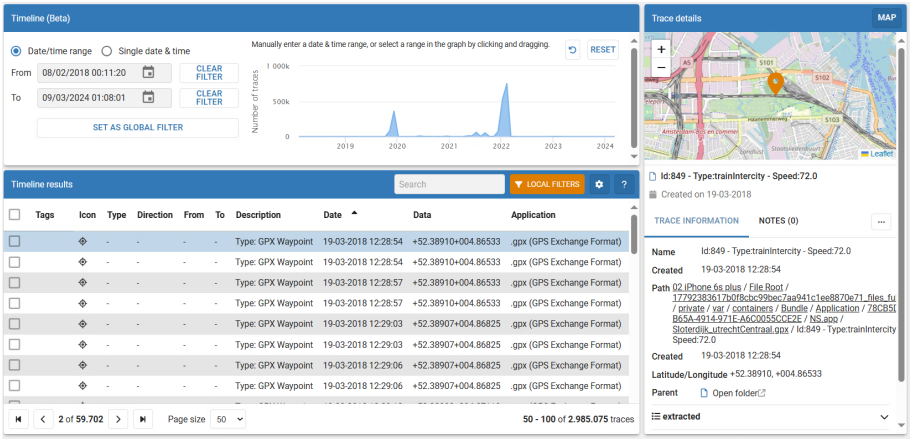
(b) Pictures view showing a gallery of images. On selection of a trace, its details are shown on the right.



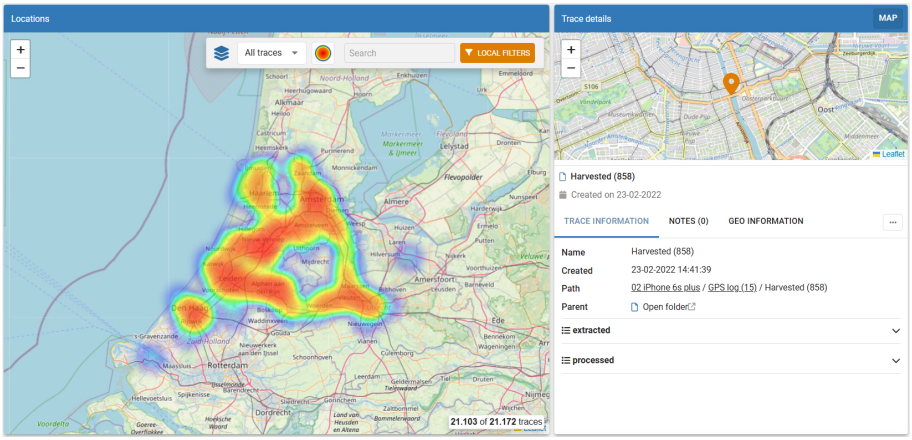
(c) Conversations view. On the left side, a selection must be made for the source application. Then, in the middle, a conversation should be selected. Lastly, the right panel shows all the messages in this conversation.



(d) Browser history view. On the left side, a browser must be selected. Then, in the middle, a trace relating to browser history can be selected, which will subsequently be shown on the right side.



(e) Timeline view showing an interactive timeline displaying all files containing any date field. Timeline can be interacted with using brushing and linking, updating the result set directly (also see 2.4.3 and 2.6). When a trace is selected, it is shown on the right side.



(f) Locations view showing a heatmap superpositioned over a geographical map. The heatmap represents the amount of traces containing geolocation data for that location. When zooming in further, the heatmap transforms into a cluster map, through interaction with the clusters or zooming in further individual traces can be see on the map. Clicking such a trace will reveal its details on the right side.

Figure 2.3: A selection of different preconfigured views in HTI. In general, trace selection is located at the left and middle of the page, and trace details will be shown on the right after one trace has been selected.

2.3 Information Seeking

The process of *information seeking* starts when a user encounters an information need or information problem (Baeza-Yates and Ribeiro-Neto, 1999, p.263). This is a psychological state in the mind of the user and not something directly observable (Cooper, 1971). The next step involves the user understanding the information problem and being able to define it. Only after a definition is made, can a user start thinking about how they will formulate the query through the interaction methods provided by the system, which is usually words which convey the semantics of the information need (Baeza-Yates and Ribeiro-Neto, 1999, p.4; Cooper, 1971). The user could then try to translate their information need into a query, which may or may not be an exact and complete representation of the information need.

The query can be entered into a search engine, which is a software system for retrieving information. When a user forms their information need, they can start a process of information seeking with the goal of reducing uncertainty considering the information encountered, or to increase uncertainty in the case of exploratory search (R. W. White, 2016, p.3). In the process of information seeking, the initial recognition and specification of the information need is followed by examination of the search results, which then inform the information need again, creating a cycle (R. W. White, 2016, p.129; Shneiderman and Plaisant, 2005, p.262). This interaction cycle of query specification and result examination only stops when the user finds a satisfactory result set. One simple process model depicting this information seeking behaviour can be found in figure 2.4.

This model shows a very simple overview of a search process. Other models have more refined process models. Marchionini proposes a model with more detailed stages and multiple points of transition, but does not take into account learning or understanding (Marchionini, 1995; R. W. White, 2016, p.134). The *Information Search Process* (ISP) model by Kuhlthau is based on a longitudinal study and contains searchers feelings, thoughts and actions that are experienced during the search process. The feelings evolve from uncertainty at the start of a search with vague thoughts, to clarity when a query can be formulated with focused thoughts, to satisfaction or disappointment at search completion.

The cognitive processes Kuhlthau described were derived from (among others) 'levels of specificity' from the *Anomalous States of Knowledge* hypothesis (ASK) by Belkin and 'levels of information need' by Taylor (Belkin, 1980; Taylor, 1968). ASK describes there is a gap between what a searcher knows and what they would like to know, and filling this gap will satisfy their information need (Belkin, 1980; Kuhlthau, 1991, p.362; R. W. White, 2016, p.129). In fact, this gap in knowledge drives the searcher to seek information. A user might not know how to formulate their information need and might not know specifically what information will fill the gap in knowledge. The user's state of knowledge is dynamic: it changes during the search process when the user learns more about the subject. With this, the users ability to articulate requests (make queries) to the system is expected to change as the level of understanding of the problem changes.

Belkin's ASK and Kuhlthau's ISP are based partly on Taylor's *four levels of need* theory (Taylor, 1968). This theory describes that after the user recognizes they have a need for information, their need can evolve through four different stages, from an initial vague sense of needing information

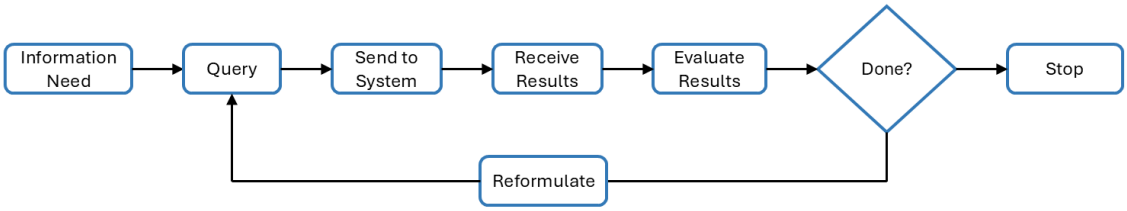


Figure 2.4: Simple overview of the information seeking process (adapted from Baeza-Yates and Ribeiro-Neto, 1999, p.263). The model assumes search is static (information need does not change) and a search is successful if the query is successively refined until the result set solves the original information need.

to a well-defined request (Taylor, 1968, p.182; R. W. White, 2016, p.42)

Taylor's four levels of need are as follows (Taylor, 1968, p.182):

1. Visceral: the initial unexpressed need for information
2. Conscious: the searcher possesses ambiguous mental description of the information need
3. Formalized: the need can be formalized as a statement or question
4. Compromised: the question is rendered in a form that can be submitted to the information system (a query)

While this stage classification was devised in 1968 in the context of searching books in libraries, it is applicable to formulation in modern information-seeking. The model describes how the users ability to describe their information need progresses as their understanding of the problem becomes more defined.

Theories like ASK and ISP assume the user learns during the searching process, which informs their ability to express themselves on the topic they are searching for, but might also change their dynamic information need. Interpreting the result set of an information request might influence the specificity of the original query and inform whether reformulation might be necessary to refine the result set towards a desired set. This is in contrast to most older research on information systems, which only focuses on a single query or a single question: a one-stop model (Bates, 2005, pp.59). A user forms the query, results are retrieved and the query is possibly reformulated. When this is done, the results are printed and the search is over (like depicted in the simple information seeking model in figure 2.4). This one-stop model does not reflect the information seeking behaviour of more complex information gathering.

Besides initial definition of the information need, it may also change over time and be influenced by intermediate insights like search results and suggestions: users learn during the search process and new information found during the search can facilitate new ideas and new directions (Baeza-Yates and Ribeiro-Neto, 1999, p.264; R. W. White, 2016, p.147). One of the models describing this behaviour is the *berrypicking* model (Bates, 1989). This model supposes information is scattered over an information space, and the information searcher is moving through this space gathering information and seeking clues to inform new navigation decisions. The concept of dynamic roaming between different sets of documents after learning more about the contents is based

on people collecting individual berries on bushes, where people picking berries move from bush to bush based on cues from local context.

The model assumes that each new piece of information gives the user a new idea and direction to follow, and - consequently - a new query (Bates, 1989). Both the search query but also the user's information need can constantly shift, as information encountered at one point may lead the search into a new, unexpected direction. In addition to this, the theory states the information need is not satisfied with one final set of results, but users gather new bits of information which may change their information needs and search queries (Baeza-Yates and Ribeiro-Neto, 1999, p.264; R. W. White, 2016, pp.147–149). Berrypicking as an information seeking strategy was already described as a strategy by Taylor, stating that the original question of a searcher may change due to adaption of feedback of the search process (Taylor, 1968, p.188).

Just like berrypicking, *information foraging* is also a metaphorical construct, which draws parallels between information seeking to organisms browsing for sustenance when hunting (Pirolli and Card, 1999; R. W. White, 2016, pp.143–147). Searchers are depicted as hunters or gatherers looking for potential prey or sustenance (information) that can be identified and accessed for use in an information environment (Savolainen, 2018, p.581). The theory covers how searchers will adapt to information environments and how these environments should look to support effective searching.

Information foraging supposes information is spread out in patches over the information space, and the searcher tries to find appropriate patches of information using information scent: their perception of value or cost obtained by proximal cues (R. W. White, 2016, pp.143–147). Since humans cannot explore all information that is available to them due to volume of the information and limited cognitive capacity, they need to make decisions about how to allocate their limited attentional resources. This leads to humans constantly making decisions about the sought information, whether to continue examination of the current information resource, to try to locate more information, or to move onto another resource.

Information is scattered in a patchy structure, with clusters of relevant information, and effort has to be made to move to other information patches when the searcher expects the utility of continuing to use a resource is lower than some personal threshold (R. W. White, 2016, pp.143–147). This threshold is made up of a searcher's previous experience and the information scent.

In web search, searchers navigate by information scent using proximal cues - textual or visual cues - to guide them to content of interest or solutions of information problems (R. W. White, 2016, pp.143–147). These cues can take many forms, among others titles, keywords, hyperlinks (and whether they were visited before, search engine auto completions and number of citations, but also the content of the summary or snippets, formatting of a page or credibility of an author can contribute.

A difference between the information foraging and berrypicking models be found in the inclusion of more detailed triggers driving information seeking in the information foraging model, describing both cognitive as contextual factors (information need and proximal cues) and the

exploitation-exploration trade-off, as opposed to berrypicking which considers only cognitive factors (Savolainen, 2018, p.589; R. W. White, 2016, p.147). Both models do have similarities as characterising search as a dynamic and sequentially evolving process.

Several theories that are discussed until now involve a user that has some information need or gap they desire to have fulfilled (ASK and ISP), while others assume this gap is dynamic and multi-faceted, possibly shifting over time (berrypicking and information foraging). If a user is not interested in posing a specific query, but might want to explore the document space looking for interesting documents, we can speak of *exploratory search* (Baeza-Yates and Ribeiro-Neto, 1999, p.65). When browsing, the user has a goal in mind but this might not be as refined as with searching, and the search tasks are open-ended and general rather than specific. Both the berrypicking and information foraging models explain exploratory search behaviour, berrypicking providing a prototypical approach of exploratory search, and information foraging extending this with concepts from outside information behaviour research (Savolainen, 2018, p.581). While exploratory search resembles both models in a behavioural sense, it is more likely to be driven by curiosity rather than a desire to fulfil an information need (R. W. White, 2016, p.158).

2.3.1 Exploratory Search

According to Marchionini, three kinds of search activities Lookup, Learn and Investigate exist (Marchionini, 2006, p.42). Lookup is the most basic search activity, and encompasses carefully specified queries and precise, discrete and well-structured response, with a minimal need for result set examination or result comparison. It is also known as known-item search, fact-finding, fact-retrieval or question answering, and contrasts Learn and Investigate activities.

Learning searches involve multiple iterations, and the result set might require cognitive processing, interpretation, comparing results and making qualitative judgements (Marchionini, 2006, p.43). Investigative search is comparable to learning search, but involves more in-depth analysis and critical assessment of results - possibly over a long period of time - before information is accepted as knowledge by a user. Examples include travel planning or academic research (Gao et al., 2023, p.6). This search activity allows for finding of new information, but may also be useful for finding gaps in knowledge. Together, Learning and investigative search activities constitute exploratory search.

Exploratory search usually appears in situations where users lack knowledge or context to formulate queries or navigate a complex information space, the search task requires browsing and exploration, or data indexing by the system is inadequate (Kules et al., 2009, p.314). A user might partake in exploratory searching when they experiences uncertainty or unfamiliarity (in their goals, terminology or the data in the information space) - which is often the case in complex search tasks as they are less well-structured - a searcher lacks knowledge of the topic or domain and is unable to conceptualise their problem (R. W. White and Roth, 2009, p.10-11).

Users engaging in exploratory search usually submits a tentative query to navigate through relevant documents in the collection, then explore the environment to better understand how to use it, seeking and passively obtaining cues about their next steps (R. W. White and Roth, 2009, p.6;

R. W. White et al., 2006, p.38). Two main activities can be discerned in exploratory search: exploratory browsing and focused searching, with more focus on the former (R. W. White and Roth, 2009, pp.16–21; R. W. White, 2016, p.159). In exploratory browsing, the user can better define their information needs and new ideas and cognition are promoted based on observed content in the system. Users will visit multiple documents to better understand what information is available, and familiarize themselves with the topic or domain. In contrast, focused search involves examining search results and documents that are in close proximity to other search results, and extract information relating to a user's goals. A user might have a clear sense of their information goals and how to meet them.

Orienteering and teleporting are two strategies relating to the focused searching activity in exploratory search (R. W. White and Roth, 2009, pp.17–18). *Teleporting* specifies a behaviour of looking for a specific document, a way to jump directly to the information target (Teevan et al., 2004, p.417). A user attempting to teleport is rarely successful, and would most likely still have to take intermediate steps to find their target. *Orienteering* describes users moving through the information space, relying on their recall and recognition skills and contextual information to narrow in on an information target, often in a series of steps. These steps can vary in size along the way, and methods chosen to take these steps can vary. Generally, a big step can be made towards the vicinity of an information target, but from there local exploration can be used to find the goal document. It is found that orienteering helps users maintain a sense of location during their search, helps them understand and trust the answer they found, helps giving context to search results and lessens the cognitive burden of finding information (Teevan et al., 2004, pp.419–421). Maintaining a sense of location helps users feel in control and prevents disorientation in the data, while attempting to teleport often elicits distrust and disorientation.

2.3.2 Query Formulation

Information seeking consists of different stages (see fig 2.4), one of which is query formulation. In this stage, the user needs to transform their information need into a query that can be entered into a search system. It has been found that of the different search stages, the formulation stage demands a significantly higher cognitive load from the user compared to the other stages in web search (Gwizdka, 2010). While these conclusions from the web may not extend to search in Hansken, query formulation remains an imperative part of the search process in any search engine, and should be explored.

Users trying to formulate their query can run into many different challenges if they have insufficient knowledge of the domain, the search system, which terms to use to create well-formed queries (vocabulary problem), which query operators to apply (boolean disjunction, boolean conjunction etc.) and how to modify the query to find more relevant documents (query modification) (Spoerri, 2004; R. W. White and Ruthven, 2006, p.934). Part of the vocabulary problem is term mismatch, where user queries and the search index are not based on the same set of terms, resulting in synonymous and polysemous words reducing relevance of search results (Furnas et al., 1987; Azad and Deepak, 2019, p.1699). Besides this, there is always a possibility for orthographic and typing errors or spelling mistakes (Mahdi et al., 2020, p.119576).

It is often difficult for users to express their information need into a comprehensive query in the query language of the search engine (Bruza and Dennis, 1999, p.489; R. W. White, 2016, p.42). When information needs are ill-defined, users may have difficulty transforming their need from the conscious state to the *compromised* state (Taylor, 1968; R. W. White and Ruthven, 2006, p.934). In the formulation stage, a user will make a change in their cognitive understanding of the search: from a vague problem to a more structured problem, which enables the user to gain focus on the issue at hand (Baeza-Yates and Ribeiro-Neto, 1999, p.263).

Queries can be divided into a class based on their function (Lau and Horvitz, 1999)[4]:

- New: a query for a topic not searched for previously
- Generalization: a query on the same topic as the preceding query, but seeking more general information. This is often done by subtracting terms from the query (term deletion) or by modifying existing terms (Bruza and Dennis, 1999, p.490).
- Specialization: a query on the same topic as the preceding query, but seeking more specific information. This is often done by adding more terms to the query (term addition) or by modifying existing terms (Bruza and Dennis, 1999, p.490).
- Reformulation: a query on the same topic as the preceding query, with the same level as specificity (no generalization or specialization)

To a certain extent, a system should be able to help with query formulation problems. For text search, this can be done through query and trail suggestions, spelling corrections, query expansion (automatic inclusion of additional meaningful terms, Azad and Deepak, 2019, p.1699), query validation (reviewing validity of boolean operators and syntax) (R. W. White, 2016, p.43).

Besides textual search using a text input field, queries can be constructed using many different interaction styles, including through direct manipulation or forms, in visual interfaces or through using facets. This is discussed in section 2.4.

2.3.3 Search Goals & Tasks

To understand a user's information seeking behaviour better, it is important to put into perspective why a search was initiated in the first place in relation to a user's *goal*. A person engaging in information seeking has one or more goals, and uses a search system to help achieve those goals (Baeza-Yates and Ribeiro-Neto, 1999, p.262). Usually, a goal is too big or complex to complete in one search action, meaning a goal first can to be decomposed into *tasks* (high level goals) and *sub-tasks* (specific components of those goals) and later into *actions* (steps taken by the searcher towards the completion of those components, like queries and clicks). This is an exemplification of the way people deal with large and complex problems, to decompose them into multiple smaller problems, hierarchically, until each sub-problem is manageable and actionable (Shneiderman and Plaisant, 2005, p.97). This task decomposition - defining sub-goals and sub-queries - is performed

[4]Several categories were omitted because they contain actions not directly related to formulation, namely requesting more results and interruption of search. Many different taxonomies exist for query reformulation. Some taxonomies define the specific syntactic modifications (Teevan et al., 2007), but a more general categorization like that by Lau and Horvitz, 1999 can be used for discerning overarching patterns.

by the user before they can start a search action, which imposes a burden on the user which aspects of their goal should be explored. This task-decomposition can be especially burdensome for novice users who are not familiar with the search-domain or lack technical knowledge of the search system.

While simple goals have a straightforward task compositions (often a lookup activity), complex search tasks require significant effort to organize (more like a learn or investigate search activity) (Marchionini, 2006, p.42). Complex search tasks are ill-defined & multi-step, spanning multiple queries, sessions and devices (R. W. White, 2024, p.56). They require deep engagement with the search engine involving many queries, backtracking and branching to complete them.

It is easy to imagine decomposition of goals into sub-goals in complex search might prove challenging. While traditional search systems do not support direct communication of user goals or intent to the system, communication over natural language - for example with generative AI - might shift the burden of task decomposition from user to the system. This is further discussed in section 2.4.5.

2.3.4 Information Seeking in Digital Forensics Search

A lot of research and theories discussed up to this point refers to the domains of web search or information seeking in general. While there are a lot of similarities compared to web search, search in digital forensic tools might include some key differences. The user base of DFTs consists of experts in the forensic field, and the search entails a specific closed domain, more akin to desktop search. Besides this, traces are enriched with possibly many types of meta-data, which can be searched through using fielded search. There are also differences in the extent (search until satisfaction vs exhaustive search) and breadth (simple goals vs complete event reconstruction). Similar to web search, in DFT search evolution in user search queries can be recognized based on interacting with the result set, much like theories like berrypicking describe (van Baar et al., 2014, p.S57). There is also a large diversity in user intent and diversity in goals that a user would like to achieve.

One major requirement of the relevance of search results in DFTs can be found in the precision and recall metrics, contrasting other types of search like web search. The effectiveness of an IR system can be determined by the extent to which retrieved information helps users to achieve their goal, but relevance of search result in the search engine can also quantitively be measured using precision and recall (Li and Belkin, 2008, p.1; R. W. White and Roth, 2009, p.18).

In information retrieval, precision is the fraction of retrieved documents that are relevant, and recall is the fraction of relevant documents that have been retrieved (R. W. White, 2016, p.360; Baeza-Yates and Ribeiro-Neto, 1999, p.75). In usual web search, a balance of both should exist (often measured using F_1-score); achieving a perfect recall can be done by returning all documents in response to a query, and increasing precision can then be done by decreasing the number of non-relevant documents.

While in web search there is big importance of precision, signified by intolerance for irrelevant results, and precision being modelled as a measure of user satisfaction. In contrast, recall is not

measurable by users because they do not have perfect knowledge of all documents in a system (Moffat and Zobel, pp.2:1–2:6). In forensics however, high recall is essential for allowing exhaustive search; an investigator would sometimes not like to miss any document that might be relevant to their query as to not miss any possible evidence, so recall must be maximized and no matches can be left out of the result set. An investigator's first focus is on making sure all possibly relevant documents are present in the result set, leading to a broadening of the query to maximize recall. Next, the investigator might notice a lot of documents in the result set that are irrelevant to their goal, which they would like to exclude. In this stage, they aim to change the query so the result set will improve in precision while not adapting recall by removing false positives of relevance. This step is done by refining the query, or filtering out results, which should be supported by the search interface.

Some research exists regarding search in Hansken, focused on information retrieval and technical characteristics of the search system (van Baar et al., 2014; van Beek et al., 2015; Du, 2020). This focus on the system instead of the users (information retrieval vs information seeking) can lead to a skewed view of effectiveness. A system might retrieve a perfect results set based on the entered query, but as described by Teevan et al., this might not result in 'perfect' information seeking behaviour and direct teleportation to an information target (Teevan et al., 2004). Besides this, influence of design of the search interfaces as part of the user experience needs to be considered. The next section will highlight theories of search interface design and show its current implementations in Hansken.

2.4 Search Interfaces

A lot of different theoretical models on search behaviour exist in literature - some superseding, others concurrent to other theories - competing for attention, being refined by promoters, extended by critics and applied by designers (Shneiderman and Plaisant, 2005, p.85). These theories have implications for how search interfaces should be designed as to support information seeking behaviour, especially when this concerns complex tasks or exploratory tasks (Liu et al., 2021, p.4).

Generally, there are two ways to find information on the web or in digital libraries, through queries - like typing words in a search box - and browsing and navigating using the architecture the system provides for these tasks (Capra et al., 2007, p.442). A good search system supports agile use of both methods. In designing a user interface for this, decisions must be made about how to arrange different kinds of information and how to structure sequences of interactions, which is daunting for a complex task like information retrieval (Baeza-Yates and Ribeiro-Neto, 1999, p.309).

While the search interface serves as a graphical user interface for the underlying search engine, it can also support the user in their query formulation, allowing them to properly translate their goal or tasks into actionable steps. When users approach an information system, they only have a vague understanding of how they can achieve their goals. An interface should aid in the understanding and expression of what a user wants to find (Baeza-Yates and Ribeiro-Neto, 1999, p.257).

Efficiency in information retrieval is especially important in the field of forensics, where the first

48 hours of an investigation are most important and pertinent information is most valuable at the earliest stages of investigation (van Beek et al., 2015, p.21; Du, 2020, p.26). If a search interface can increase user understanding of their own information need by offering search support, this might make a valuable impact on investigations.

Interfaces can make use of different *interaction styles* depending on what tasks and actions are expected, the proviciency level of users and technical constraints among other things. Interaction styles can also be combined to allow for more diverse tasks or users. Five different interaction styles exist: command language (direct input of a text-query), form filling (where a query can be constructed using form and input fields), menu selection (where pre-configured queries are shown and can be selected), direct manipulation (where a query can be constructed using a visual representation) and natural language (a query is constructed based on a natural language text prompt) (Baeza-Yates and Ribeiro-Neto, 1999, p.278; Shneiderman and Plaisant, 2005, pp.71-74).

This section will provide a small overview of the five different interaction styles, their applications in search interfaces and their relationship to Hansken Tactical Interface. Then, faceted search and the results page will be highlighted, after which general design principles for search interfaces will be discussed, along with implications of the theoretical models as discussed above.

2.4.1 Command Language (CL)

Command language is the oldest of the interaction styles, being the sole option in the 60s to mid-70s due to information technology constraints (Liu et al., 2021, pp.5-9). In CL, the searcher can enter a text query directly into a text area and submit their search, but interface techniques do exist to assist users, like suggestions and conversation-style formulation. For advanced and experienced users that are familiar with the syntax - also known as power-users - complex, powerful queries can be constructed rapidly. Drawbacks of this method of interaction is the need for substantial training and thus unfriendliness for novice users, high probability of syntax errors, poor error handling and heavy reliance on memorization (Shneiderman and Plaisant, 2005, p.72-73).

Problems in command language interaction exist. The user must remember command syntax, field names and possible field values, which can easily be forgotten between usages of the system (Baeza-Yates and Ribeiro-Neto, 1999, p.280). Moreover, existence of many different command languages with slight variations in syntax or functionality might introduce user errors. The lack of flexibility of a syntax can be attributed to it being designed for the system rather than the user of a system.

A command language interface is currently implemented in the Hansken Tactical Interface as the query input field for HQL in the search screen. Some search support methods are implemented as well; 1) suggestions for auto-completion of queries are given so users are aware of the possible field names and value syntax. 2) query syntax validation is provided when executing the search. 3) Search history and bookmarking queries enable getting back to previously searched queries. Other than this, a link to HQL-documentation is provided next to the text area for quick access.

In the text-field, keyword querying can be performed. In the basis, keyword-based searching retrieves documents that contain the keywords specified, but a more complex combination of

keyword-operations can also be specified (Baeza-Yates and Ribeiro-Neto, 1999, p.100). Keyword searches are intuitive and easy to express for users of all levels of system or domain expertise.

Besides keyword search, HQL also supports formulating boolean queries. This is not without problems, as many users have difficulty specifying queries in boolean format and often misjudge what the result will be (Baeza-Yates and Ribeiro-Neto, 1999, p.279). Inexperienced users believe the syntax is counter-intuitive, wrongly assuming the AND-keyword widens the scope of a query. This might be because the English 'and' can be used disjunctively (select all of A and all of B) while a logical 'and' - as used in a boolean query - is used conjunctively (select all that are both A and B), thus creating a more specified result set of documents. Moreover, users may be confused by operator preference and the use of parentheses for nested evaluations.

As all the other interaction methods are compiled to command language in the background, CL is by definition capable of all affordances of a search engine. This makes it very a powerful interaction method for more experienced users compared to the other interaction styles.

2.4.2 Form Filling (FF)

Form filling allows for a more graphical representation for constructing a query by offering all possible field options to a user and simplifying data entry, thereby offering assistance for users - especially novice users (Shneiderman and Plaisant, 2005, p.72, p.295; Liu et al., 2021, p.9). The full complement of information is visible, which limits possibility of errors but does not completely eliminates it if field values should be entered manually and should contain correct syntax. Using widgets like calenders and maps for selecting field values helps reduce user errors by further abstracting the compilation to the underlying query language.

Form filling is currently implemented in Hansken tactical interface as 'type of trace'-search and the 'date and time'-search (see figure 2.2). In this interaction style, the user is required to first select a trace type from a drop-down menu (browser history, email, picture or contact for example). After this, a new drop-down shows the possible fields available for that trace type (for type email the new fields comprise 'from', 'cc', 'subject' or 'timestamp' among others). Upon selection, another field appears for entering a field value, which can be a input text field, a (range) calender picker or a drop-down of possible values. It is also possible more nested sub-fields appear which can be searched on.

'Type of trace'-search is a hybrid of form filling (allowing a user to enter a value for a field) and menu selection (selecting a field from the dropdown list). Menu selection is highlighted in section 2.4.4.

2.4.3 Direct Manipulation (DM)

In direct manipulation, a visual representation of an object or action is used to signify possible user actions (Shneiderman and Plaisant, 2005, p.71). Interaction with the visual representations will lead to an immediate result, which gives the user direct feedback on their action. This style of interaction is flexible, provides control to the user and could be suitable for users of all experience

levels if implemented well (Liu et al., 2021, p.11). DM is characterized by rapid & reservable incremental actions by pointing actions on objects of interest (Shneiderman and Plaisant, 2005, p.71, pp.214-217).

Many visual interfaces have been designed for accommodating search or query formulation implementing direct manipulation techniques (Fishkin and Stone, 1995; Jones, 1998; Yi et al., 2005; Russell-Rose et al., 2019; Russell-Rose and Shokraneh, 2020). Some designs use metaphors of real-world physical objects (a magnet in Yi et al., 2005 or magnifying lenses in Fishkin and Stone, 1995) while others can be more abstract (Venn-like diagrams in Jones, 1998 and 'objects on a canvas' in Russell-Rose et al., 2019). Direct manipulation is a style often used when interacting with data visualisations and visual interfaces, which is discussed in more detail in section 2.4.9.

In the standard search interface of the Hansken tactical interface, there is no implementation of the direct manipulation interaction style. Some of the other views do have these functionalities however; a timeline, showing the amount of traces distributed over time, can be interacted with through brushing and linking to show the traces from a specific period in time (see figure 2.6 and 2.3e). The heatmap as shown in figure 2.3f can be interacted with through panning and zooming, and individual traces can be selected when zoomed in far enough. There are also several views that show data divided into facets, like the chat conversation view (see figure 2.5, 2.3c and 2.3d). Faceted search employs a combination of menu selection (selection from a list of items) and direct manipulation (immediate manipulation of the search query & encouragement of exploration) (Shneiderman and Plaisant, 2005, pp.71-73). This separation into facets is done in multiple views, including the browser history-view (grouping by URL), email-view (grouping by email server) and the photo camera-view (grouping by camera), but the general search view does not support it. Section 2.4.6 highlights faceted search in more detail. Visual search interfaces are very dependent on the direct manipulation interaction style, and are discussed in section 2.4.8.

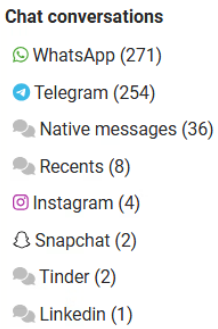


Figure 2.5: Excerpt from the chat conversations-view in HTI.

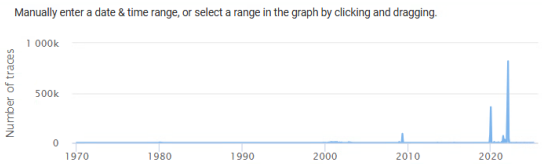


Figure 2.6: The timeline from the timeline-view in HTI.

2.4.4 Menu Selection (MS)

In menu selection, the user is given a list of items and they select the item that is most appropriate for them (Shneiderman and Plaisant, 2005, p.71). In the case of a search interface, this item could be a full pre-configured query or perhaps a part of a query. This method does not allow for data entry directly (like form filling) but does offer easy interaction with pre-defined queries, and a clear (possibly hierarchical) structure due to all possible options are presented at one time. The

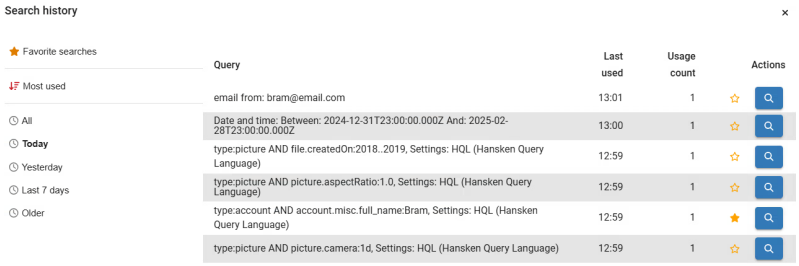


Figure 2.7: The window showing user search history and saved search queries. It shows the most recently executed queries during the chosen period in the left menu. On the right of a query, it can be marked as favourite (making it appear when selecting the 'Favorite searches' button) and a query can be executed again by clicking the magnifying glass icon.

style is suited for novices, but might be cumbersome or inadequate for more experienced users (Shneiderman and Plaisant, 2005, p.71; Liu et al., 2021, p.9). Due to the limited flexibility in customisation of the items, it is imperative the items are designed to actually reflect tasks the user would want to undertake.

Menu selection is implemented in HTI as part of 'type of trace'-search, by showing users a list (dropdown menu) of the possible fields (see figure 2.2). After selection of a field, users can fill in a value to construct their query (form filling interaction).

The menu selection interaction method is also implemented in the user history and saved query modal, as seen in figure 2.7. This interface shows all previously executed and bookmarked queries in a list, where selection of an item lets the user execute that query once more.

For general query construction, a menu style interaction on its own does not allow for the flexibility needed in a digital forensic tool. One possible application of the menu interaction style (combined with direct manipulation) is faceted search, which might be very valuable for allowing users to gaining insight into the data. Faceted search is highlighted in more details in section 2.4.6.

2.4.5 Natural Language (NL)

In natural language interaction, users can interact with a system using a familiar natural language (like English) to give instructions or give responses (Shneiderman and Plaisant, 2005, p.332). Natural ways of communication through computers - spoken and written natural language, speech, video and gestures - are preferable for humans because it allows them to think and move in a way that they use in their non-computing lives (Hearst, 2011). Another advantage of NL interaction is users do not have to learn command language syntax or use menus to fulfil their information need. System interaction with natural language is a powerful and expressive means of communication of user intentions (R. W. White, 2024, p.57).

A *Natural Language Interface* (NLI) implements NL interactions, which is usually through a simple input text area for entering a NL statement. A popular extension to this is the conversation style interaction - in a search application known as conversational search or a *Conversational Information Retrieval* (CIR) system - which allows users to interact with a system in multi-turn conversations in natural language (Gao et al., 2023, p.6). Conversational information retrieval might offer a more natural UI for information seeking compared to a single-turn search engine,

but as users are less familiar with them (in web search at least), it will take time before searchers adapt to them (R. W. White, 2024, p.59). A main objective of conversational search is to provide information interactively and similar to human-human interactions (Liu et al., 2021, p.104).

Comparing interaction styles, NL-interaction does not guide users the way for example direct manipulation or menu selection do. NLIs do not usually convey information about the available task actions - but rather are a simple input text area for entering a NL statement - which might make them more effective for intermittent users who are knowledgable about the possible tasks and interface concepts rather than novice users (Shneiderman and Plaisant, 2005, pp. 332–333). For expert users, command language or other interaction types might be preferable however, as they provide more control for the user and are more reliable. On the other hand, the NL-interaction style relieves the task of learning syntax and, which would make it convenient for novice users.

Besides considerations based on expertise, natural language communication has inherent benefits and drawbacks compared to other interaction styles (Hall et al., 1996, pp.2–3); benefits include familiarity with the language, concise and flexible formulation, requiring less formulation time. Possible drawbacks or challenges can be found in the system handling ambiguity of NL-statements, along with the need to handle linguistic and conceptual problems.

It is impossible to talk about NL-interactions without talking about the technology driving it. For an NL-query to be processed by a search system, it should be interpreted through *Natural Language Processing* (NLP), which is of itself an incredibly complex problem (Manaris, 1998, pp.2–4). With current significant advancements in Artificial Intelligence (AI) the natural language interaction style is very popular in model-based agents like *ChatGPT*, using Large Language Models (LLMs) and generative AI for question answering in a conversation style (Palta et al., 2025). Applications of AI to help people with complex cognitive tasks and empower them to amplify their cognitive capabilities are also named *AI agents* or *copilots* (R. W. White, 2024, p.57). They are designed to keep the human at the center of the task completion process, but help them to complete a broad range of tasks in less time and effort. In web search, usage of AI agents is emerging in the form of conversational assistance and integration of dynamic answers in result pages.

Another opportunity can be found in using natural language interactions to as a way to get started in search. A search interface should provide good ways to get started and avoid an empty screen or blank entry form which does not provide any cues to help a user decide in what way to start a search process, often called *blank page syndrome* (Baeza-Yates and Ribeiro-Neto, 1999, p.267). Users tend to start out with small queries to understand what kind of results are returned, which informs the way they will reformulate the query. Ways to support users starting out can be to provide an overview of the content of the collection, which can be interacted with to find items of interest (much as is described by Shneiderman and Plaisant in section 2.4.8. Other options include to show users examples of items, present wizards or show a faceted categorization of item attributes to help starting users (see section 2.4.6). In addition to this, conversational AI-agents line ChatGPT might help for avoiding blank page syndrome because it could be uses for generating initial queries based on your explained information need (Scanlon et al., 2023, p.8).

As described in section 2.3.3, complex searches in traditional search interfaces require users to

break up their goals into tasks and sub-tasks before actions (like individual search queries) can be formulated. Communication through natural language interfaces with AI agents enables users to communicate their intents and goals more directly (R. W. White, 2024, p.58). The searcher would not need to compose their goal into tasks, sub-tasks and actions as they can describe their goal in NL and the AI agent could be responsible for understanding intentions. While possibly posing challenges in terms of human control and learning, the responsibility for generating answers or creating a correctly structured query is delegated from the searcher to the system.

The behaviour above is comparable to strategies employed by a reference librarian even decades ago, working with the information inquirer to translate their formalized or conscious information need into relevant documents (Taylor, 1968, p.183). While the inquirer may try to put their need into actionable steps or a compromised question, the reference librarian will try to understand the information need (goal) and direct the user towards the relevant documents. In the case where a librarian - or AI-assistant - has more subject knowledge or expertise over the system contents compared to a user, they may be more capable of formulating the relevant actionable steps from an information need.

Using Hansken, Henseler and van Beek have explored the potential of AI-powered solutions through ChatGPT (Henseler and van Beek, 2023). Applications included writing structured queries in HQI, summarizing, evaluating and visualizing traces relating to conversations and analysis of search results. It featured a conversational search interface to interact with the system. The authors concluded that this application of ChatGPT shows great potential for a wide range of tasks aiding investigators, but that it is not (yet) ideal, mainly due to making mistakes and hallucinating facts.

AI agents in forensic tools could prove very useful. In fact, many opportunities for NL-interaction and AI exist; The main application is to use NLP for collection and analysis of digital evidence in the system (Ukwen and Karabatak, 2021). Other opportunities include trace understanding, evidence searching, code generation, anomaly detection, incident response, and education. (Scanlon et al., 2023). Major strengths can be found in creativity, reassurance and avoidance of blank page syndrome. Looking at the context of search systems, NLP could prove useful for writing queries, summarizing, evaluating, analysing and visualizing search results (Henseler and van Beek, 2023).

Expanding upon using LLMs for writing queries, translating natural language into structured search queries allows investigators to find the right evidence more efficiently without having to learn a sophisticated query language - like HQL in the case of Hansken (van Beek and Henseler, 2023, p.119). If interactions with forensic software become more natural, this enables non-technical investigators to perform queries without extensive training (Wickramasekara et al., 2024, p.8). Additionally, they can assist in human learning of the query language, later enabling users to formulate queries themselves with more power, with LLMs acting as interactive teacher to enhance comprehension (R. W. White, 2023, p.32; Wickramasekara et al., 2024, p.6)

Currently, artificial intelligence is implemented in several ways in Hansken. Two rule-based AI methods exist for extracting entities: recognizing unstructured data as a well defined pattern of symbols (like a licence plate, email address or credit card number), merging names of companies,

persons or places into a single identity (merging data containing a full name with data containing an email alias). Two deep learning methods are implemented: classification (predicting labels) for images and video frames, and face detection and comparison using feature vectors.

Some note can be made of the role of an AI agent in a digital forensic process. As their role is to work alongside humans and empower them, amplifying their cognitive capabilities, it could be argued that an AI system should aim to help humans and not be used to take over tasks (R. W. White, 2024, p.57; Baber and Alotaibi, 2024). Using AI as provocateur instead of servant improves critical thinking of the human in the loop and prevents using AI as a substitute for human expertise, (Sarkar, 2024, pp.18–19). It is argued that a collaborative approach, using AI as an argumentative tool, might lead to a more thorough and efficient investigation.

Challenges with AI agents may be plenty, especially in the context of forensics. While natural language interaction has its own challenges as described before in this section, NL-interactions through AI agents and LLMs pose their own challenges entirely.

General challenges of AI and LLMs include biasses, hallucinations, reasoning errors, logical errors, explainability, spelling & grammatical errors and prompt injection (Hadi et al., 2024). Additionally, AI agents may remove the need for searchers to learn from the search process (like in exploratory search) (R. W. White, 2024, p.60). Human control of the search process is delegated from the searcher to the AI, making it unclear for the user what information is included, what is not included and why, which would help with understanding and trusting the system output. *Grounding* may help with these issues of human learning and human control by providing reasoning traces along with their answers (R. W. White, 2023, p.41; R. W. White, 2024, p.61).

In the field of digital forensics, it is a precarious endeavour to implement LLMs and close scrutiny and caution should be important, especially towards hallucinations, also known by the alternative term of 'incorrect' information (Scanlon et al., 2023, p.9). Challenges specific to digital forensics can also be found in the need to train LLMs with forensic data to optimize results, which might lead to biased training data of questionable quality due to complexity and variation of cases (Wickramasekara et al., 2024, p.10). This is, if real evidence can even be uploaded to an LLM at all (Scanlon et al., 2023, p.8).

Alternatively highlighting explainability, investigators ought to be hesitant to trust information if it is unclear how a system came to its conclusion, and is required to understand information before making use of it as knowledge (Scanlon et al., 2023, pp.9–10). This also raises the issue of accountability from ethical and legal perspectives, as it is unavoidable for an LLM to sometimes generate incorrect information, but it is unclear if responsibility then lies with the developer for ensuring the models accuracy or the user for insufficient interpretation and validation of results.

2.4.6 Faceted Search

It has been established by now that users partaking in information seeking get informed by the result set of their query, which informs their subsequent search steps (Bruza and Dennis, 1999; Baeza-Yates and Ribeiro-Neto, 1999, p.264, p.281; Bates, 1989). There are several ways how this behaviour can be supported, one of which is *faceted search*. The goal of this type of search is to

enable a person to explore a domain via its attributes (R. W. White and Roth, 2009, pp.44-45). This is usually displayed by creating an overview of the result set through clickable categories of structured metadata from the search results, organising the results into meaningful groups in order to help make sense of the results and decide on actions (Kules et al., 2009, p.313). The document set is divided into bins of overarching characteristics, where each bin is called a facet. A key notion of faceted organization is multiple representation for each item, meaning a document is included in multiple facets if they contain the attribute that a facet comprises (Capra et al., 2007, p.443). This allows users to browse through the result set and narrow it down by refinement without explicitly reformulating their initial query. Faceted search offers flexible navigation, organized results and supports both query expansion and refinement, but is usually used to narrow down a result set (Kules et al., 2009, p.313). In addition, mental workload on the user is reduced by promoting recognition over recall, as all elements and options are visible without the user having to remember where to navigate (R. W. White, 2016, p.190; Nielsen, 1994)

Faceted search interfaces usually combine another mode of search - like keyword search - and browsing, allowing rapid and flexible information finding, avoiding users feeling like they are lost in the collection and making it easier to explore (R. W. White and Roth, 2009, pp.45). The goal of faceted search is to allow domain exploration via retrieved document properties by enabling both query expansion and refinement through hierarchical categorizations. Facets seem to play an important role in the exploratory search process (see section 2.3.1), informing users of how to proceed with a search and helping organise the view of a topic domain for further investigation (Kules et al., 2009, p.320).

A problem of keyword search or boolean queries is that result sets may be too large or empty because the user scopes their query too broadly or too narrowly (Baeza-Yates and Ribeiro-Neto, 1999, p.281). This may be due to the users not understanding or being familiar with the contents of the data they are searching through. This might be remedied with faceted search. Another issue with text-queries is the possibility of orthographic and typing errors. In faceted search, this is prevented by only allowing navigation through menu selection over existing item categories (Mahdi et al., 2020, p.119576).

In HTI, facets exist to a certain degree in filtering functionality after performing an initial search (in some result views, not for general search). In some views, this functionality can be found in the filtering menu (see figure 2.8), and in others it is part of the main part of the view (figure 2.3c and 2.3d), sometimes requiring mandatory selection of a facet before results are displayed. This functionality gives some insights into the data in the result set and allows refinement of the shown results, but lacks hierarchical facets and the implementation can be seen as inconsistent across different views. Similar functionality can be seen in figure 2.5, where clicking on an item will show only results that adhere to this characteristic.

Facets seem to play an important role in the exploratory search process (Kules et al., 2009, p.320). They can help users organize their view of the topic domain, and be used to select sub-topics that can be further investigated. In addition, facets can give users ideas on how to proceed a search. In other studies, it has been found that facets can both be used for understanding the structure of the data, gaining an overview of the data and allow the user to narrow down the search by

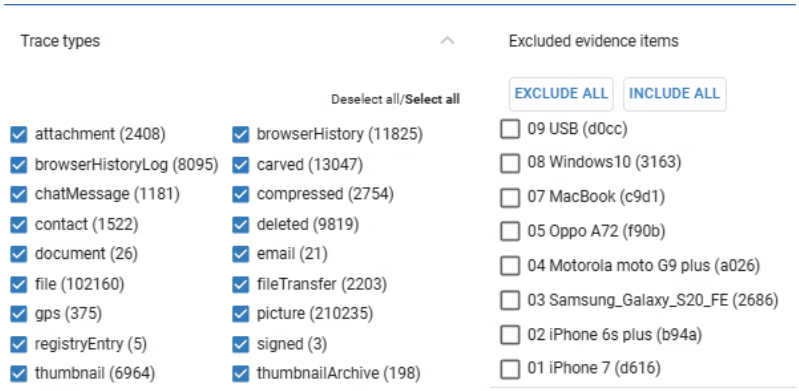


Figure 2.8: Two facets in HTI, showing different filter options a user can click to include or exclude items in their search results. The number next to a trace type filter item shows the amount of traces that comprise the bin.

selecting sub-topics in exploratory search (Capra et al., 2007, p.449; Wilson and schraefel m.c., 2008, p.55).

2.4.7 Search Results Page

When a search query has been entered and is submitted, the search engine will retrieve results which then have to be displayed to the user. This is done on a separate page - often also named *Search Engine Results Page* (SERP) - which displays the retrieved content in a certain way. Over the years, SERPs in web search have evolved in content, layout and navigation from a simple vertical list of hyperlinks to information-rich interfaces with complex presentation and results with variable granularities of information (Roy et al., 2022, p.2765). How and where content is displayed in the result page affects user interactions, and the addition of diverse new features in web search engines like Google lead to higher *good abandonment*, meaning a user's information need is entirely addressed to their satisfaction. On the other hand, continued growth of the range of SERP elements might also lead to clutter on the page, but this is not expected to halt its evolution (Oliveira and Teixeira Lopes, 2023).

Say there is only one result that needs to be shown to a user of a web search engine. Depending on the modality of the item, it might be displayed mainly textually (if the source is a document, webpage or news item for example) or mainly graphically (if the source is an image or video for example)[5] (Gritz et al., 2023, p.321). Because an item is generally only skimmed or scanned, it is important that the information displayed will allow users to make a rapid selection decision on whether an item is relevant for them (Feild et al., 2013, pp.2999–3000). Options for this include adding thumbnails, snippets (perhaps with query-term highlighting) and other relevant information depending on the modality of the source (like duration for video sources).

When a search engine returns more than one result, displaying these in the interface requires more considerations. The results could be displayed in a list layout - the classic interface from the days of the 'ten blue links' in web search - but also in a grid. While a grid can display items in a more condensed format compared to a list layout, allowing more items to be shown at once given the

[5] A thumbnail, favicon or image could be displayed for a textual source as well, just as textual information like a title or caption can be displayed for mainly graphical content.

screen space, users might end up exploring more in a list layout (Roy et al., 2022). On the other hand, the influence of item positioning on selecting trustworthy sources is drastically reduced in the grid layout compared to a list.

Even more considerations need to be taken when items from heterogeneous sources need to be displayed. While heterogeneous content can be displayed in a single list or grid layout, this might be a less effective and less easy-to-skim display method, and standard practice in web SERPs is to employ *aggregated search* (Zhou et al., 2012; Gritz et al., 2023, p.321). In aggregated search, sources in different modalities may be aggregated through vertical search based on their media type or genre into a so-called *vertical* (Zhou et al., 2012, p.115; Oliveira and Teixeira Lopes, 2023; Bron et al., 2013). A vertical then contains only one type or genre of source, in web search for example only images, videos, news items or stocks, which can then be displayed in a way that is optimal for that content. Aggregated search helps order the search results into meaningful groups, something that information seekers often express a desire for (R. W. White and Roth, 2009, p.44; Hearst, 2006, p.59). This categorized overview of search results can support exploration, understanding and discovery in complex search tasks (R. White et al., 2008).

Two general types of aggregated search interfaces exist: tabbed and blended. Tabbed interfaces provide access to each source in a separate tab, while a blended interfaces combines multiple sources into a single result page (Bron et al., 2013). A blended interface may be motivated by a user's exploratory information need, exploring the content of various sources simultaneously. This interface containing *blocks* of content is now a standard in many web search engines (Zhou et al., 2012, p.115). In the tabbed display, only results of one type are displayed in an interface, and tabs can be used to switch between these types (Gritz et al., 2023, p.322).

Modern search result interfaces have many more features and SERP element, offering featured snippets, direct answers, a knowledge panel, top stories, a people also ask-list and many more (Moran and Goray, 2019; Moran and Goray, 2020; Chilton and Teevan, 2011). Applying information foraging theory, these features all involve low interaction cost because they do not require extra clicks, lengthy reading or leaving the current view. At the same time, they could offer high expected benefit, providing an answer directly.

In the end, it is important to make use of familiar interaction styles and interfaces that users have become accustomed to for searching (Capra et al., 2007, p.450; Roy et al., 2022, p.2765). People are familiar with and influenced by web search engines in which they can type in search terms and quickly get a list of relevant documents. They are used to website UIs and faceted selection styles, and as more sites embrace interactive UIs, expectations and tactics of users will likely evolve (Capra et al., 2007, p.450). The influence of familiarity with web search on expectations in search systems might make concepts from web SERPs extendible to other search systems, like those of DFTs.

In the Hansken Tactical Interface, there are several different search result interfaces available. The result views are directly tied to the content type that is selected, meaning that the image view shows only image files in a gallery and the browser history view shows all browser entries in a vertical list. The Search view shows a vertical list with several properties pre-selected like filename, size,

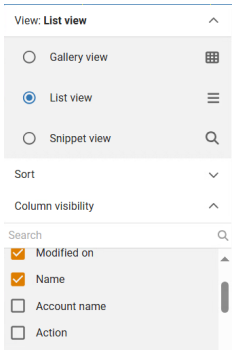


Figure 2.9: Settings for searching in the general Search view. The way of displaying results is by default a List view, but on selection can also be a Gallery view or Snippet view. Additionally, sorting order and the columns that are shown can be selected.

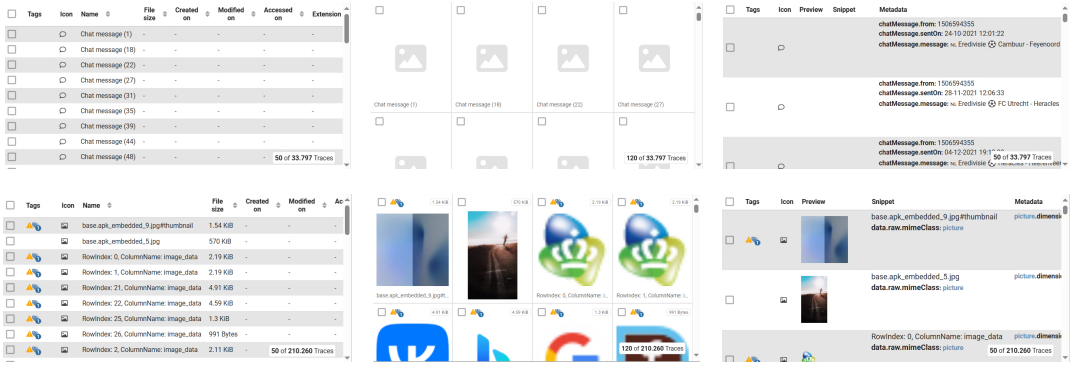


Figure 2.10: The different view types, as can be selected through settings shown in figure 2.9. Top row depicts a result set with only chat messages, bottom row shows only picture type results. From left to right, list view, gallery view and snippet view are selected. It can be seen that a snippet view might be more optimal for chat messages because it shows the contents of the message, but a gallery view could be better for pictures.

file extension and several dates relating to the file. These properties are often present, over all types of data, but more type-specific properties can be displayed by manual selection as shown in figure 2.9. Instead of a list, the user can also select they would like to see their results in a gallery or snippet view as well through the settings menu. Examples for chat message and picture type content of the three different views is shown in figure 2.10. It can be seen that depending on the type of content, the optimal way to view it can be different. Figure 2.3 shows some more ways of displaying data on search results pages in HTI, offering views specialized towards displaying specific data types an optimal way.

2.4.8 Visual Interfaces

With big information spaces and an increasing volume of possibly relevant information to navigate, exploring through the data becomes increasingly difficult, both for locating known items and for browsing in order to gain an overview (Shneiderman, 2003, p.1; Ahlberg and Shneiderman, 1994, p.313; Dörk et al., 2009). This difficulty in information exploration leads to users experiencing anxiety, information overload, feelings of falling further behind and inability to cope with the flood of information. One possible method for reducing user anxiety about the flood, for finding needles in haystacks, for supporting exploratory browsing, for finding patterns and exceptions and

for making browsing fun can be found in implementing UI design principles of *Visual Information Seeking* (VIS).

Visual Information Seeking is an application of information visualisation on information seeking, which helps users understand and analyse the data and is important for hypothesis generation (R. W. White and Roth, 2009, p.52, pp.259–265). Visualizations provides humans with the ability to take advantage of their physiological capability to process large amounts of information visually because humans are attuned to images and visual information (Altiero, 2015, p.35; Baeza-Yates and Ribeiro-Neto, 1999, p.260). Interactions with these visualisations are usually in direct manipulation interaction style (see 2.4.3) and allows users to understand the data and explore the space by zooming in on aspects of interest and filtering out irrelevant items. This can be done through for example brushing and linking, panning and zooming, focus-plus-context, magic lenses and the use of animation (Baeza-Yates and Ribeiro-Neto, 1999, pp.259–261). The best visualisations are highly intuitive, communicate their point clearly, stimulate viewer engagement and attention and does not require additional training to understand their meaning (R. W. White, 2016, p.261).

An application of data visualisation for information seeking is through *dynamic queries* (Ahlberg and Shneiderman, 1994; Shneiderman, 2003, p.5; R. W. White, 2016, pp.259–265; Shneiderman and Plaisant, 2005, pp.574–597). Using this, query parameters can be changed through a direct manipulation interaction style to filter and zoom in on relevant data, usually through sliders or buttons. Advantages include rapid, safe (from misformulation) and fun exploration of the information space, allowing discovery of trends in data like clusters, exceptions, gaps and outliers while engaging users and putting them in continuous control of the search process (Shneiderman, 1994; Marchionini, 2006, p.44). Visualisation in search query formulation can offer benefits as fewer zero-hit queries, improved query comprehension and better support for exploration of an unfamiliar information space (Russell-Rose et al., 2019, p.380). Dynamic queries are effective for exploration because they provide high-level overviews of a collection, and rapid previews of items help users understand data structures and infer relationships between concepts (Marchionini, 2006, p.44). In contrast, visual composition might be cumbersome and challenging to use or interpret, and it is more difficult to be precise with visual representations compared to text (Gatterbauer et al., 2022, p.4).

With dynamic queries, both the query components and the search results are presented visually and enforce *tight coupling*, where query components are interrelated and the impact on each selection can be seen in the other query components. Another aspect of tight coupling is using *output-is-input*, which eliminates the distinction between queries/input and results/output because every output is a candidate for input (Ahlberg and Shneiderman, 1994, p.315). It reduces screen clutter and offers a single location on the interface for both gathering information and applying an action to get more information.

An application of tight coupling in an interface can also be seen in *coordinated views*, where multiple views and interactive visualisations like timelines, maps and topics are shown in the same interface showing interdependencies in the information space (Wang Baldonado et al., 2000; Dörk et al., 2009). These views could then use interaction styles like brushing and linking - where highlighting items in one view will highlight corresponding items in another view.

Design principles of visual information seeking are hugely influenced by Shneiderman and his research, abiding by his visual information seeking mantra:

Overview first, zoom and filter, then details-on-demand

In more detail, Shneiderman created the *Type by Task Taxonomy* (TTT) of information visualizations (Shneiderman, 2003). The taxonomy assumes a collection of items containing multiple attributed is viewed, and the user wishes to perform a basic search task of selecting all items that satisfy values for a set of attributes. The seven tasks in the taxonomy are:

Overview: Gain an overview of the entire collection.
Zoom: Zoom in on items of interest.
Filter: Filter out uninteresting items.
Details-on-demand: Select an item or group and get details when needed.
Relate: View relationships among items.
History: Keep a history of actions to support undo, replay, and progressive refinement.
Extract: Allow extraction of sub-collections and of the query parameters.

In digital forensics forensic tools , data visualisations can just as well help users with finding relevant documents to answer an information need. Current tools employ a WIMP model (windows, icons, menus & pointing device), which are ill-suited for presenting large amounts of forensic data efficiently and intuitively (Garfinkel, 2010, p.S71). In these textual interfaces, it is difficult to find and correlate information - and find unknown and new hypotheses - when a fraction of the data is displayed (Altiero, 2015, p.28). Solutions to these issues can also be found in intuitive data visualisations, adopting Shneiderman's overview, zoom, filter, details-on-demand mantra to significantly increase the likelihood of discovering otherwise invisible data.

With advanced and fundamental changes in the computer industry, the field of digital forensics is facing many challenges (Garfinkel, 2010, p.S66, p.S71). Developing and adopting new approaches to visualisation and presentation of forensic targets, and integration of interactive visualisations with automated analysis tools into DFTs will allow investigators to interactively guide the investigation.

2.4.9 Design Principles

Design principles are guidelines for the design of a product, and understanding them can give useful insight into how an why designs work. One influential list of principles, derived from experience and refined over decades, are the *eight golden rules of interface design*. Shneiderman and Plaisant devised this list of eight principles which is applicable to most interactive systems (Shneiderman and Plaisant, 2005, pp.74–76):

Strive for Consistency: Have uniformity in actions, terminology, colours, layouts and fonts, among other things. Exceptions should be comprehensible and limited.

Cater to Universal Usability: Interfaces should accommodate diverse users, including novices, experts, individuals with disabilities, and different technologies. Provide features like explanations for beginners and shortcuts for advanced users.

Offer Informative Feedback: Every user action should produce a system response. Frequent and minor actions require subtle feedback, major actions should receive more substantial responses.

Design Dialogs to Yield Closure: Sequences of actions should be organized in groups with beginnings, middles, and ends. At completion of a group action, provide informative feedback to give users satisfaction and relief.

Prevent Errors: Minimize opportunities for users to make mistakes by disabling inappropriate options and guiding input. When errors do occur, provide simple, constructive and specific instructions for recovery.

Permit Easy Reversal of Actions: Allow users to reverse actions to reduce anxiety and encourage exploration of unfamiliar options. Reversibility should apply to single actions, data entries, or groups of actions.

Support Internal Locus of Control: Give users a sense of control over the interface. Interface actions should be predictable, desired actions should be producible, necessary information should be obtainable, and tedious data entry should be avoided.

Reduce Short-Term Memory Load: Keep displays simple and consolidated, minimize window-switching, and provide easy access to necessary information. There should be time for training and access to references relating to interface actions.

There is a big overlap in Shneiderman and Plaisant's golden rules and Nielsen's Usability Heuristics (Nielsen, 1994). Nielsen's heuristics have a focus on usability evaluation, while Shneiderman and Plaisant's golden rules highlight interface design. Because of the significant overlap and variation of focus, the usability heuristics will not be discussed here.

Adding to the golden rule of *universal usability*, having different interfaces for novice or expert users caters to the trade-off for simplicity vs. power (Baeza-Yates and Ribeiro-Neto, 1999, p.259). A simple interface may only offer basic search functionalities which are easy to learn and adequate for simple queries. A powerful interface may offer complex search functionalities which are time consuming to learn and remember, but adequate for more powerful, flexible and effective queries. Having an intuitive bridge between simple and advanced interfaces constitutes good interface design.

In supporting exploratory search systems, R. W. White and Roth devised eight features that should be present in systems (R. W. White and Roth, 2009, pp.41–59):

Support querying and rapid query refinement: Systems must help users formulate queries and adjust queries and views on search results in real time.

Offer facets and metadata-based result filtering: Systems must allow users to explore and filter results through the selection of facets and document metadata.

Leverage search context: Systems must leverage available information about their user, their situation, and the current exploratory search task.

Offer visualizations to support insight and decision making: Systems must present customizable visual representations of the collection being explored to support hypothesis generation and trend spotting.

Support learning and understanding: Systems must help users acquire both knowledge and skills by presenting information in ways amenable to learning given the user's current knowledge/skill level.

Facilitate collaboration: Systems must facilitate synchronous and asynchronous collaboration between users in support of task division and knowledge sharing.

Offer histories, workspaces, and progress updates: Systems must allow users to backtrack quickly, store and manipulate useful information fragments, and provide updates of progress toward an information goal.

Support task management: Systems must allow users to store, retrieve, and share search tasks in support of multisession and multiuser exploratory search scenarios.

2.4.10 Theory Implications for Exploratory Search

In section 2.3, many theories of information seeking behaviour are described and related. It is important that these behaviours are supported in the search interface, and that search engines should be designed to match user's search strategies rather than forcing users to accommodate to the current search interface (Chen and Chua, 2013, p.508). Theories should be central to research & practice and theories should be leading rather than lagging behind what has been produced already - guiding practitioners in designing new products (Shneiderman and Plaisant, 2005, p.85). As such, some theories related to exploratory search - as discussed in section 2.3 - and their applications on interface design are revisited.

If searching is seen as an evolving process, as is done in theories like Anomalous States of Knowledge and Berrypicking, the interface and functionalities should support this. This can be achieved by allowing users to reasses their goals and adjust their search strategy accordingly in the interface. (Baeza-Yates and Ribeiro-Neto, 1999, p.264). The interface should offer support for storing or marking intermediate results and navigating back through a user's search process, helping them refine queries and revisit previously explored information. The interface should support users modifying and refining their queries as their information need evolves (R. W. White, 2016, p.149; Bates, 1989; Belkin, 1980)

By extension, implementation of these features will also support the process of exploratory search.

High levels of interaction are integral to exploratory search and one way to achieve this is through categorized overviews of search results using result metadata in facets (see section 2.4.6) to support understanding, discovery and exploration (R. White et al., 2008). Providing users with continuous control and engaging them with the search process, dynamic query interfaces immediately update the result set and the display (see sections 2.4.8 and 2.4.3)(Marchionini, 2006). By applying the design principles for supporting exploratory search systems in section 2.4.9, the user is able to gain an overview of the information space and orientation within it by formulating multiple queries and considering search results (R. W. White and Roth, 2009, pp.41-59; Dörk et al., 2009).

Information foraging may be supported by features relating to navigation, item content and exploration. Leaving a patch of information in search of another can be supported by offering an overview of the data in various representations (supporting orientation), and offering navigational aids like breadcrumbs to make leaving a patch in favour of another easier. Users determining the information scent of a patch of information might be supported by clear cues and metadata in the information representation, like providing summaries, keywords, thumbnails and previews or snippets (Budiu, 2020). Providing context of search terms and whether users are on the right track can inform their perceived information scent.

2.5 Prototyping & Design Thinking

An unaddressed research themes for forensic tool development is the design of DFTs, specifically regarding ease of use and having powerful, intuitive interfaces that are not too technical (Beebe, 2009, p.29). As part of this research, a new prototype interface is made through the process of *design thinking* (Stanford d.school, 2010; Brown et al., 2008).

In the design thinking process, observations are used to fuel understanding of current problems, needs or desires. Then, ideas are generated, implemented and tested. Design thinking is not linear, and projects will loop back through the different stages as ideas are refined or new directions are taken (Brown et al., 2008, p.4). Moreover, the process is iterative as testing informs a next iteration of prototypes.

In this research, the five-stage design thinking process is used, comprising the phases (Stanford d.school, 2010):

1. Empathize: understanding user needs through observation and engagement
2. Define: framing the core problem based on insights
3. Ideate: generating a wide range of possible solutions
4. Prototype: creating low-fidelity versions of potential solutions
5. Test: evaluating prototypes with users to gather feedback and refine ideas

For implementation of this process and the iterations, see section 3.3.

Chapter 3

Methods

For this research, a distinction is made between the *main study* (3.1) and the *survey* (3.2), as these two methods can run in parallel and are independent. The main study is sequentially structured as an *interview*, a *think aloud exercise* and a *post-task interview*. These methods are discussed in sections 3.1.3, 3.1.4 and 3.1.5 respectively.

The choice of methods fits both the purpose of the research, but also considers the constraints and realistically available resources. Interviews were chosen to be able to get an in-depth perspective of individual information needs of investigators. The survey will yield less in-depth responses but will be distributed to a larger group of people, showing broader sentiment and adding quantitative insight. To show users alternative interaction methods, a design prototype combined with a think aloud exercise and a post-task interview will show how intuitive an interaction is and what elements of the interface support the search process.

Other methods have been considered. As there is no quantitative usage (log) data available about Hansken for security reasons, this could not be analysed for answering current usage of Hansken in RQ2. Understanding and respecting time constraints of investigators, a diary study of Hansken usage was not feasible. For confidentiality (case-specific details) and logistical (Hansken not always used in favour of other DFTs) reasons, observations were also infeasible. Congregating a focus group instead of doing interviews would allow for a greater breath of insights and experiences, but if information needs or experiences are not shared these cannot be explored, as interviews would allow.

Hansken is a licensed product (DFaaS, see chapter 1), used by partners with some investigative branch (hereafter referred to as *organizations*). NFI develops Hansken (partially), but generally has no users on premise. Therefore, several organizations were visited and the main study was conducted on location. This was both most respectful of investigators time and yielded more possible participants to the research. The research as described in this section was done in parallel to another user research into Hansken being led by NFI, and participants were shared between these two sessions.

For the main study, a *prototype* is made iteratively, using a design thinking process (see 2.5). This prototype is examined during user testing and refined using these insights. See 3.3 for the details on the design thinking iterations.

3.1 Main Study

To be able to answer the research questions, a mixed-method study was conducted. Interviews are held to understand the working processes, information need, query formulation methods and information seeking strategies used in Hansken (see section 3.1.3). This can answer RQ1 "*What is the information need of Hansken users interacting with the Hansken Tactical Interface?*" and RQ2 "*How do Hansken users search for information in the Hansken Tactical Interface and how do their search behaviours and use of available search capabilities align with theory of information seeking?*". Experiences can be compared to existing theories of information seeking discussed in 2.3 to see if they align. Think aloud exercises and subsequent post-task interviews (see sections 3.1.4 and 3.1.5) can give insight into what elements might improve participant's search experiences, thus answering RQ3 "*How do specific interface features and interaction styles influence cognitive load, perceived usability, and effectiveness of information seeking in a Hansken prototype interface, and what design implications emerge from these findings?*" and Sub-RQ3 "*How are natural language interactions with AI agents perceived in terms of usefulness and trustworthiness for information seeking and task decomposition?*".

3.1.1 Participants

Organizations where the research was to be conducted were selected by the NFI. Employees at these organizations were contacted and could participate if they were available and interested in contributing. To participating in the main study, possible participants needed to 1) have used Hansken at some point in the last year and 2) understand and speak Dutch. An employment role within an organisation did not matter for sampling, because a more diverse group of users is a better representation of the actual users of Hansken.

While it is interesting to include participants without experience of Hansken or HTI for RQ3 and Sub-RQ3, they do not have insights into RQ1 or RQ2. Due to constraints in the availability of participants, two participants without prior experience with Hansken (ID 41 and 89) participated in (part of) the main study.

A total of **12** participants across **5** investigative organizations were recruited for the main study. Participant details can be found in table 3.1. They were selected by a combination of purposive sampling and convenience sampling; participants were asked to join based on availability and interest, but the starting population consisted of individuals with interest and expertise in Hansken.

The large selection of organizations that were visited generates a broad view of working processes over different organizations. The diversity of employment roles in the participant group will additionally provide broader insights.

ID	Iter.	Role	Hansken Usage
7	1	Digital Expert	Frequently
46	1	Strategic Analyst OSB	Rarely
54	1	Tactical Investigator	Rarely
2	2	Tactical Investigator	Frequently
50	2	Tactical Investigator	Frequently
68	2	Analyst OSINT Intel	Occasionally
41**	2	Analyst-operator Intel	Never
5	3	Digital Expert	Frequently
35	3	Tactical Investigator	Occasionally
51	3	Operational Analyst OSA	Very frequently
62*	3	Operational Analyst OSA	Frequently
89	3	Strategic Analyst OSB	Never

Table 3.1: The participants that took part in the main study. *Iter.* denotes which iteration of the prototype they were given, *Role* is the role in their organisation and *Hansken usage* gives an indication of how often they use Hansken.
*: Participant performed main study online, **: participant only took part in think aloud session

3.1.2 Materials

To be able to conduct this research, several materials were required:

- A laptop with the prototype loaded on it. The device will also be used by the researcher for reading the questions in the interview.
- A phone with an internet hotspot (for the prototype) and to record the audio of the main session
- Surveys printed out on paper (see 3.2)

The organization that was visited was made aware of these materials in advance.

3.1.3 Interview

At the start of the session, an information sheet and consent form detailing data processing of participant responses and requesting permission for audio recording was presented (see D). After a participant signed for consent, they were told that any further questions could be answered at any time if they come to mind.

Of course, looking at the context of the research, it might be imperative some confidential data stays confidential. Participants were reminded not to share case details that are not supposed to be shared, and if they would like to rescind some comment, they can say so and this will be removed from the transcription.

The interviews were semi-structured. This was chosen to allow consistent insights into needs and practices of interest, but also to create space for exploring avenues that arise during the study (Blandford et al., 2016, p.4). The interview started with an icebreaker and diverted to the work of the investigator. The role of Hansken in their work was explored, and functionalities of HTI were evaluated based on the work tasks described by them. The focus was on the search process in HTI and the information need of the investigator.

The interview and subsequent think aloud exercise were held in a relatively quiet room at the organization. Audio recordings were made during the entire duration of the study (including the think aloud and post-task interview) to not overlook any data during later analysis, which could happen with note-taking (Blandford et al., 2016, pp.14–15).

The interview was followed with a think aloud study of the prototype interface.

3.1.4 Think Aloud

In a think aloud study, a participant is asked to interact with a system while articulating their thoughts as a 'stream of consciousness' (Blandford et al., 2016, p.39). The purpose of the think aloud exercise is to understand the strengths and limitations of the interface, as well as show how participants perform tasks. It also serves to get the participant acquainted with the prototype interface, about which some questions are asked in the post-task interview.

The participant was instructed on the concept of a think aloud exercise. In addition to the audio recording, a screen recording of the interaction with the prototype was made for later analysis. This recording captured the user interactions with the prototype, which is valuable insight next to the participant's thought process (Blandford et al., 2016, p.16). Of course, the participant was asked for their consent for recordings beforehand in the information sheet.

In the researcher's script, several prompts were included to encourage a participant to speak their thoughts throughout the entire process. This is important because the practice of thinking aloud often does not come naturally (Blandford et al., 2016, pp.39–40).

A small, simplified and non-specific scenario was created and handed to the participant on paper, along with general information and some exercises (see F and G). The participant was asked to read the information and start thinking aloud when they read the first exercise.

The exercises were made to have easy reference to the interface (using the same words as are present in the interface), while leaving some things to the user to figure out (e.g. how to interact with elements, what filters to use). The scenario is generic, to be applicable to all the different organizations that would be visited.

After the participant completed the tasks, they were asked about their experience in the post-task interview.

3.1.5 Post-task Interview

After the think aloud exercise, the participant was asked for their opinions of the prototype that they just worked with. The purpose of this is to gather more data of what they thought of various elements of the interface, and how it impacted their feelings and the effectiveness of search. A comparison was made to the current HTI to understand the difference in interactions and areas of improvement or worsening.

The post-task interview took the same semi-structured form as the interview at the start of the main study.

3.1.6 Data Analysis

After user studies had concluded, data processing was performed. From the audio recordings, transcriptions were made using Whisper[1], in compliance with guidelines from the Netherlands Forensic Institute. Transcriptions were refined by hand: correcting mistranscriptions, removing stop words and irrelevant or personally identifying passages, and indicating whether statements were given by the researcher of the participant. Screen recordings from the think aloud exercise were used as context for the audio transcriptions. Interesting interactions (for example quick understanding of an interaction or wrong interpretation of a task) were noted.

After transcription, data was coded using an emergent coding strategy to find patterns and build a narrative as codes are found. This was done in NVivo 15[2]. The practice of coding entails labelling data that is meaningful to the research in some way, in this case marking segments of text in the interview transcriptions. In emergent coding, the labels are not defined before coding but rather during the process of coding. Codes were iteratively re-arranged and renamed, and repetitions and redundant codes were removed.

Both open coding and axial coding were used on the qualitative data from the interviews. In open coding, interesting segments of text are assigned a code related to why it it interesting or relevant for the research. In axial coding, relationships between those codes are defined and hierarchy in codes is created. From this, a codebook was constructed with all discovered concepts from the interviews.

3.2 Survey

A survey, printed on paper, was handed out to participants when visiting an organization and at the NFI. If preferred, an online version was also available. Using a survey allows for a broader reach and a greater number of responses than merely conducting the main study. It gives a broader perspective, and show if the findings from the small set of interviews are generalizable across a bigger audience, and the interviews can conversely give in-depth insights into survey findings (Blandford et al., 2016, p.72). This way, the main study and the survey complement each other. The answers from the surveys can be used to complement the main study for RQ1 and RQ2, but also inform the final prototype, and thus RQ3

The questions of the survey are divided into three parts: 1) demographics, 2) statements on Hansken, information need and information seeking and 3) an Importance-Performance Analysis (IPA) of features in Hansken. For questions (except for the demographics), a 5-point Likert scale was used for for quantitative analysis (Likert, 1932). In addition, participants could optionally write a comment, for example to elaborate on why they chose a particular rating in a designated box. The survey can be found in appendix B.

[1] https://openai.com/index/whisper/
[2] https://lumivero.com/products/nvivo/

For part 2, a statement is presented and the participant can use the Likert scale for selecting the options:

Strongly Disagree - Disagree - Neutral - Agree - Strongly Agree

For part 3, there is a scale for if something is possible and one for if something is important:

To what extent is this currently possible in Hansken?

Not at all - Hardly - Partially - Well - Fully

How important is this to you?

Very Unimportant - Unimportant - Neutral - Important - Very Important

The survey had a small information sheet detailing what the gathered information will be used for and how it will be processed (see appendix A).

The online version of the survey was made in Qualtrics[3]. The contents of the online and paper survey were identical.

3.2.1 Participants

Participants for the survey were sampled at the NFI and at organizations that were visited for the main study. Just like for the main study (see 3.1.1), a combination of convenience sampling and purposive sampling ensured both qualified and available participants for the survey. A total of **14** survey's were fully filled in. Participants did not have to be fully acquainted with Hansken to be able to fill in the survey, and could choose not to answer Hansken-related questions if they felt they could not answer them. The survey may be filled by people that take part in the main study.

3.2.2 Data Analysis

After the surveys were filled in, paper surveys were made digital manually. The online survey results were exported and joined in this file.

The survey contains some Likert scale questions, which will allow for quantitative analysis to complement insights from the rest of the study. They were analysed with a simple frequency distribution, and standard deviation was calculated to understand if there is consensus between participants. Qualitative insights from the survey were coded as well, though they were kept separable from the interview answers.

Because the survey questions are in Dutch, translations are used in the later sections.

3.3 Development of the Prototype

Through theoretical insights and expert user feedback, a new search interface for Hansken was iteratively designed through the process of design thinking (see section 2.5). The interface was made to implement several search support methods that are aimed at increasing query result insights and improving usability of interactions, while also introducing novel functionalities. The

[3]https://www.qualtrics.com/ and https://survey.uu.nl

interface was made to be usable by both novice and expert users, much like HTI, but is styled differently to prevent comparisons to current interface functionality.

Theories described in section 2.3 were used to inform design choices in the new prototype. This applies the challenge in the human-computer interaction field, to let theories lead rather than lag behind practice (Shneiderman and Plaisant, 2005, p.85).

3.3.1 First Iterations: Sketches and Paper Prototypes

The define and empathise stage in the design thinking process were mostly based on challenges encountered in previous research, preliminary talks with colleagues and focus groups with users. Challenges - relevant to the domain of search - that were identified here included:

- Trouble getting an overview of the data present in Hansken.
- Difficulty formulating search queries for individuals with less technical insight or experience.
- Issues with not knowing how to start looking when working with a new data set.
- Difficulty identifying which results are relevant in large result sets.

The ideate phase consisted of making numerous sketches of interface elements, generating ideas from theory or known designs and ideating together with colleagues. A collection of sketches was made of possible interface designs, as sketches are very low fidelity and can easily be made or adapted. Some sketches can be found in figure 3.1.

For formulation of queries, several novel ideas were worked out, including node-link canvas, wizards and visual query interfaces. For showing trace results, hypergraphs and many visualizations were considered.

The prototype for this iteration consisted of a blended aggregated search view with several verticals of different data-representations, which serve as coordinated views - both showing output (showing results) and taking an input (filtering the result set). Examples of this is the map and timeline. Data-attribute facets (among others on data type or device) made it possible to specialize a result set. The goal of the interface is to provide an overview of the data and to be able to quickly filter though the results. Many other ideas from the sketches were also worked out so they could potentially be shown during testing.

The goal of this prototype was to support exploration in the data. Design choices were partially based on features described by R. W. White and Roth (R. W. White and Roth, 2009, pp.41–59). Specifically, querying and rapid refinement, visualisations and offering facet based filtering were implemented. A dynamic query interface further engages users with the search process by updating both result set and the display (Marchionini, 2006).

A paper prototype was made, which offers a flexible way of showing separate elements of an interface, but also allows participants to re-order elements, add them or remove them. This prototype can be seen in figure 3.2. Initial testing was done with several colleagues, and the main study was conducted with three participants subsequently.

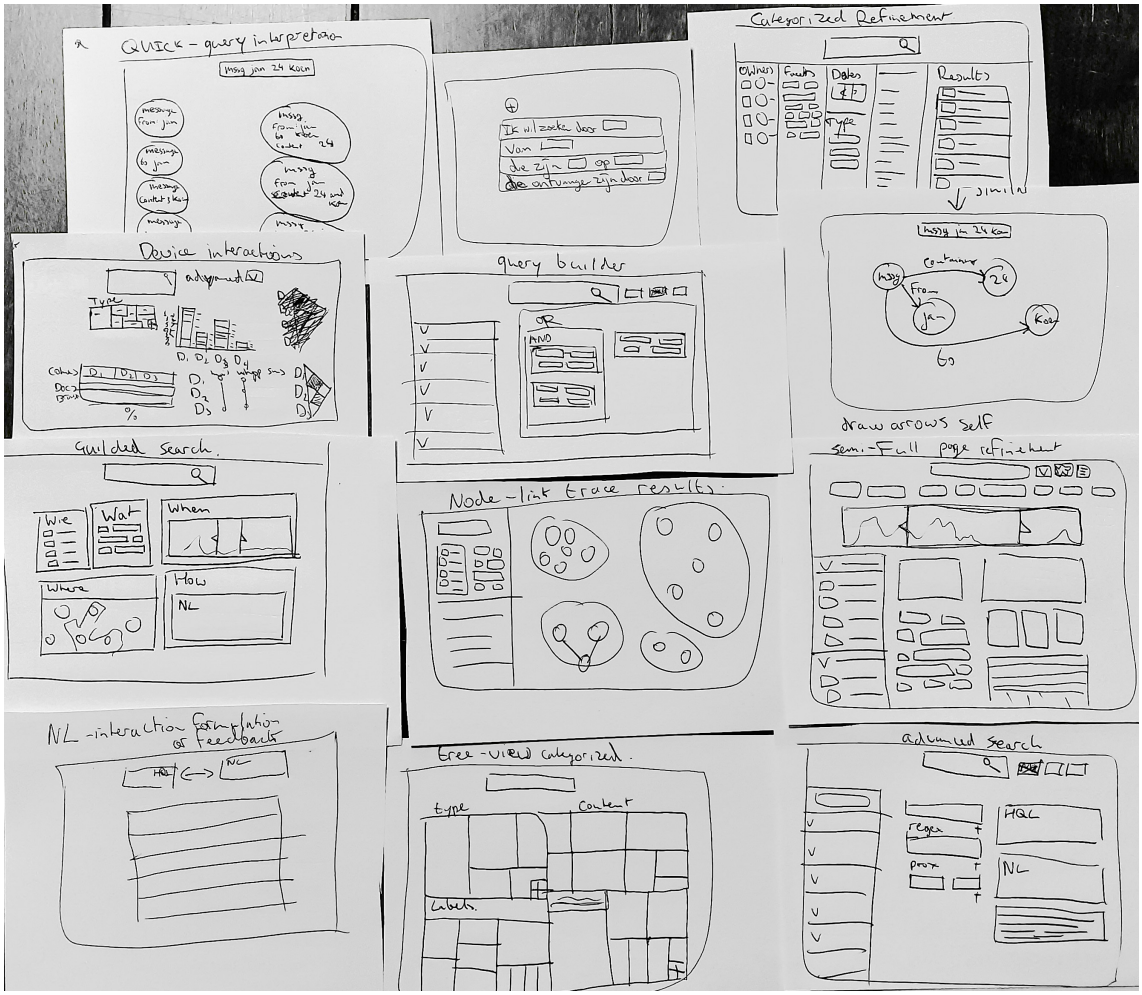


Figure 3.1: A collection of interface sketches made during the ideation and prototyping phases of the first iteration.

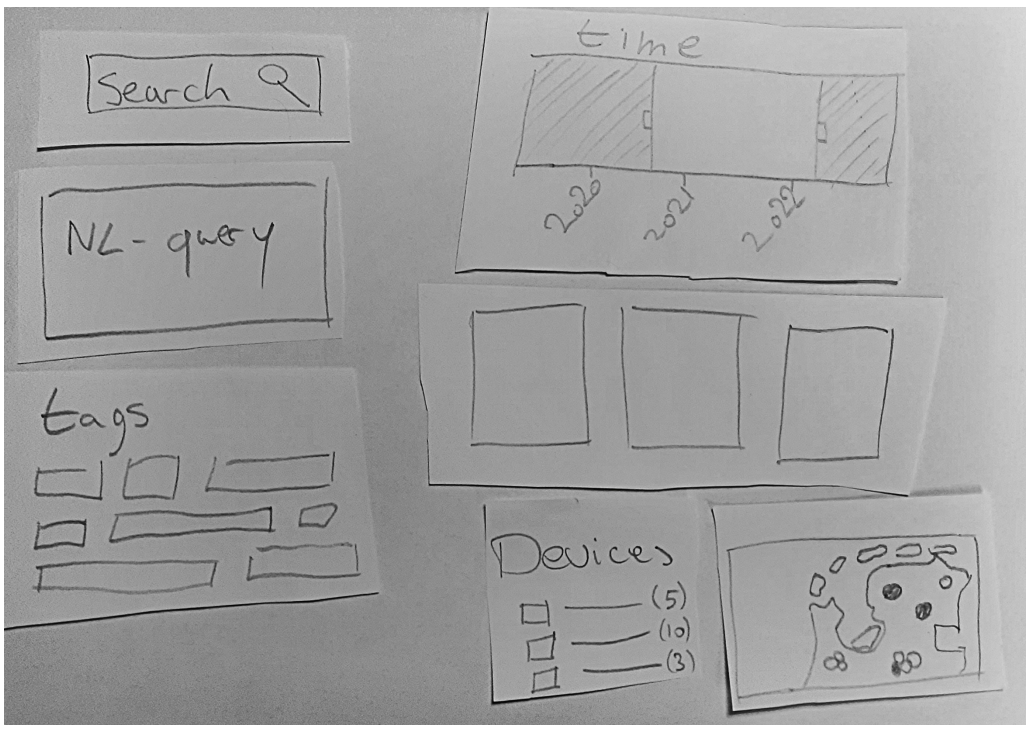


Figure 3.2: A paper prototype as laid out for the first iteration.

The three participants of the main study indicated they liked the having categorical divisions. Additionally, what is important to them is having an indication of relevancy of a device or set of devices.

3.3.2 Second Iteration: Initial Web Prototype

For the second iteration, feedback from the first three interviews and think aloud studies was used to understand users needs, their thoughts on Hansken and opportunities for improvements. Main problem definitions included trouble formulating with HQL, missing categorical overviews of data and missing context for communications. Opportunities can be found in having natural language as a formulation method and giving indication of relevant or interesting data.

In this next iteration, concepts from the paper prototype, combined with some new ideas based on definitions of the previous iteration are used to create a web-based prototype. This prototype is based on the same blended aggregated search interface as introduced in the first iteration, here called the *explore view* (see figure 3.3). It was chosen to make a web-based prototype with a functional filtering system to be able to show and test interaction methods and interplay between verticals, which is not possible in lower fidelity prototypes.

An overview of the verticals implemented in the interface can be seen in figure 3.5. Users could filter the data using the timeline, the device facet and the type facet, which then impacted all other verticals (input-is-output, as part of tight coupling). The other verticals are output only and cannot be used to filter the dataset.

Additionally, a keyword search can be made, filtering the dataset to only traces with content that match part of this term (see figure 3.4, top left). A copilot button brings the user to a mocked conversation AI interface (see figure 3.6). The natural language interaction is mocked, meaning the response and filters that are enabled are always the same regardless of the question.

Two ranked and sorted lists were created in this prototype as a response to participants in the previous iteration stating they have interest in a quick overview of data in a device. These verticals are 7 and 8 in figure 3.5.

Lastly, an advanced search panel is introduced. In this panel, users can make a complex keyword search, but more technically challenging query methods - like boolean expressions and wildcards - are abstracted away into a sentence. Through menu selection, the user can choose in what way keywords should match, see figure 3.7.

This prototype was tested with 4 participants in the think aloud exercise (see section 3.1.4 and appendix F). Participants had trouble interacting with the timeline and with deselecting certain filters. Otherwise, the exercises could be completed without much intervention. The aggregated search interface was met with mixed reactions, some explaining it might be useful and give oversight in exploratory tasks, others explaining they believe it is too cluttered and exclaiming preference of the more fragmented categories as they exist currently.

The copilot interaction was generally liked, and while trust in AI was low, the fact that the action of the AI was explained and they could see which filters were applied to the dataset made them more trusting.

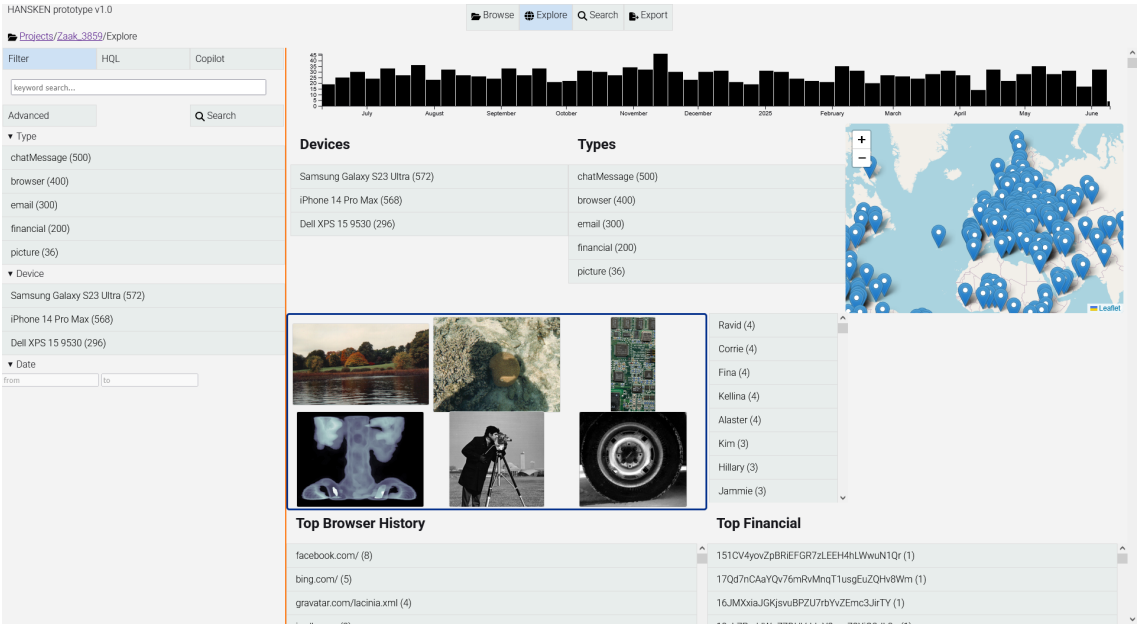


Figure 3.3: The screen you encounter when the prototype is first opened.

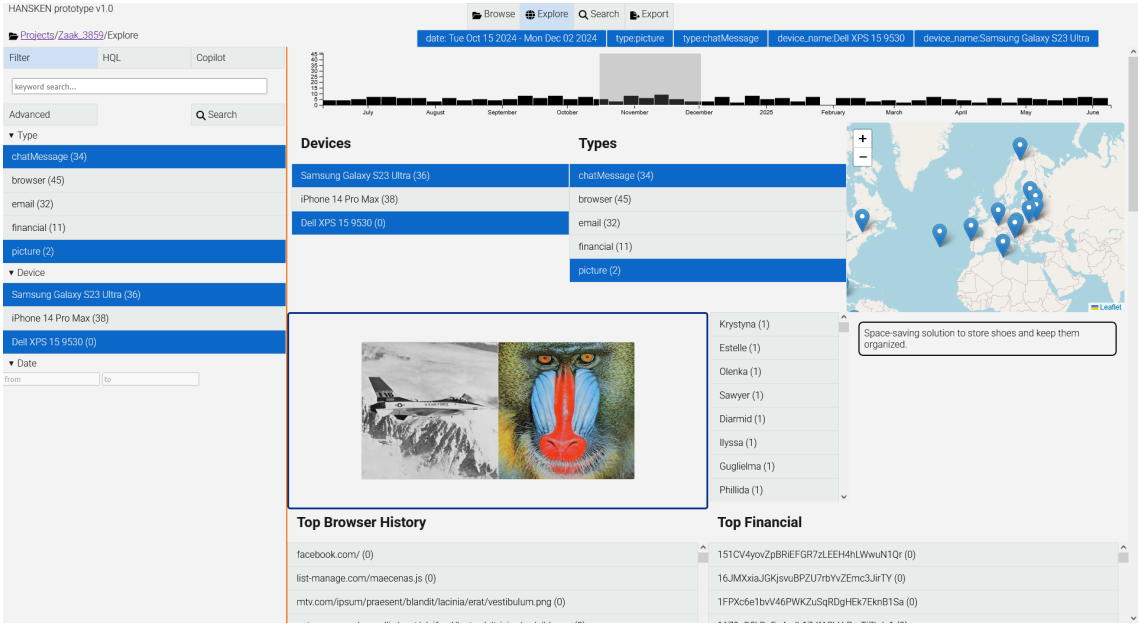


Figure 3.4: Zooming in or filtering out data can be done with the facets and the timeline. Through tight coupling, changes made in one coordinated view also changes the selection in another.

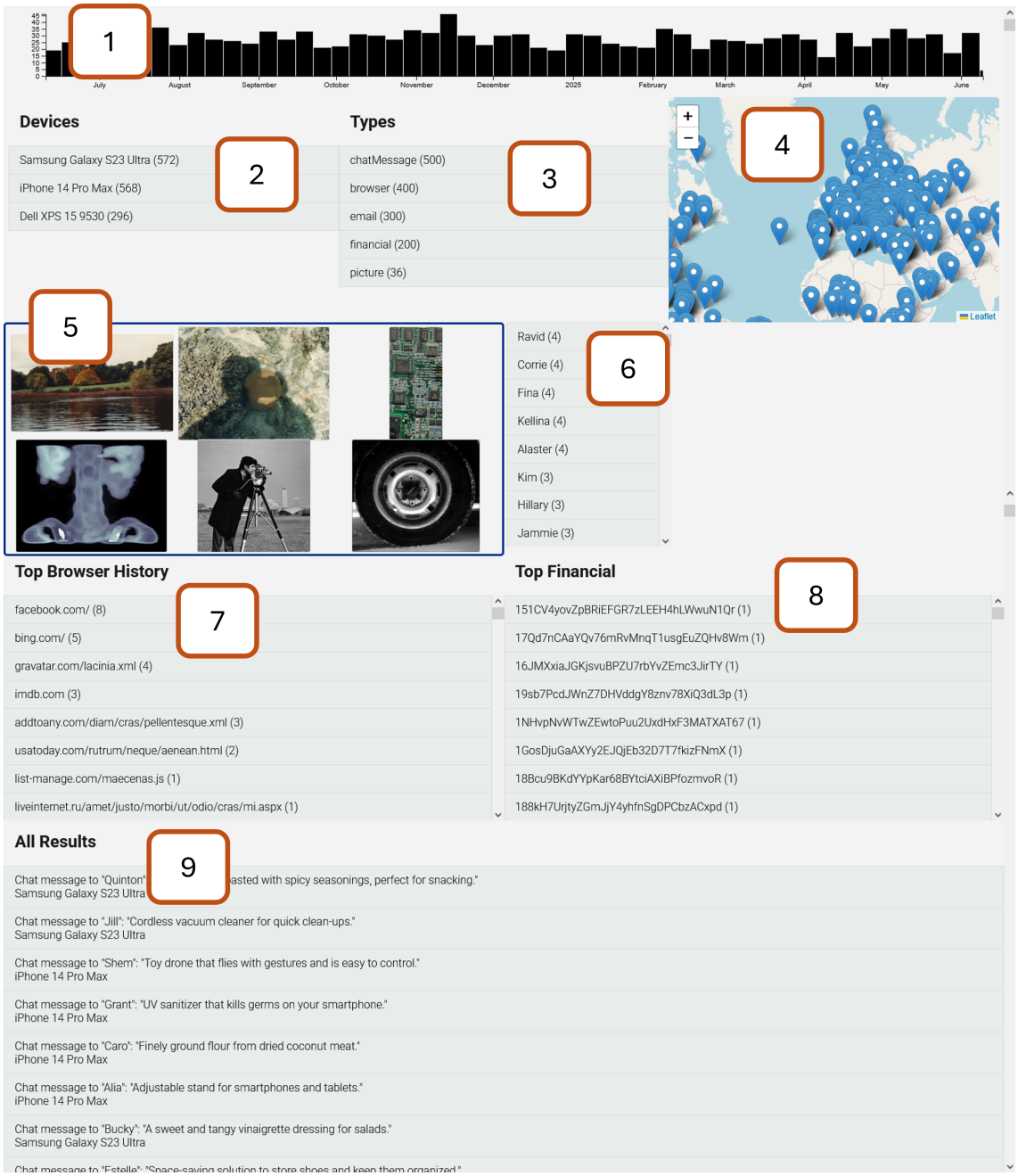


Figure 3.5: A complete look at the verticals in the prototype interface. The numbered elements represent one vertical. 1: Timeline, 2: Device facet, 3: type facet, 4: Trace map, 5: Pictures, 6: Chat conversations, 7: Top browser history, 8: Top financial, 9: All results

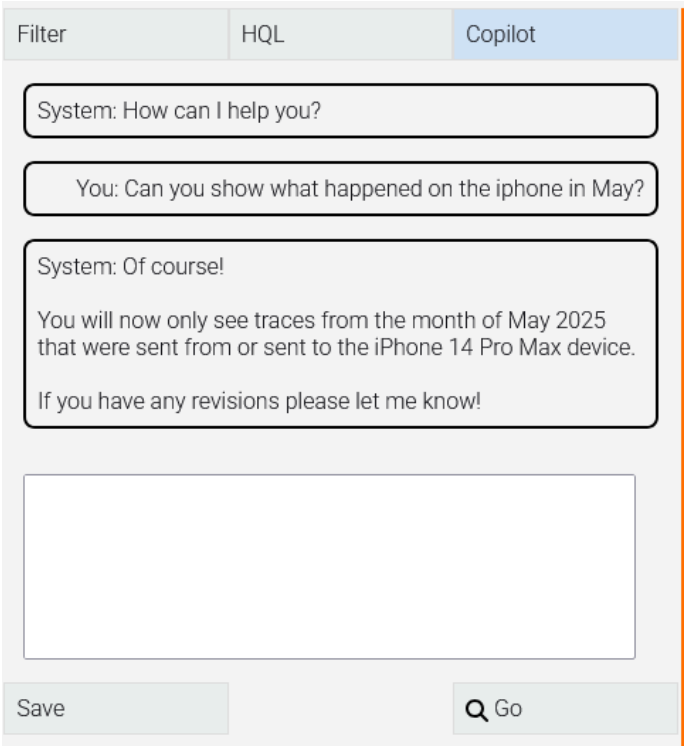


Figure 3.6: The conversational AI interface. A question had been asked by the user and the system replied with a per-defined answer.

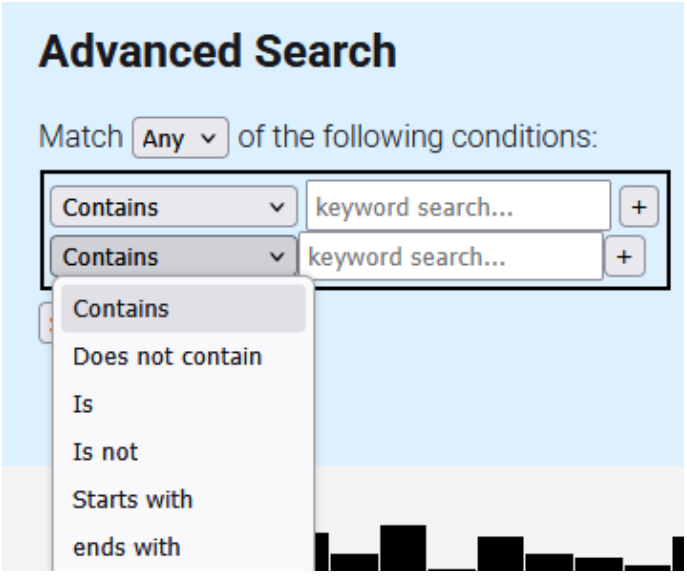


Figure 3.7: The advanced search panel. Users can select whether they want to perform match all conditions (AND-query) or any statements (OR-query), and how the documents should be matched with the keyword.

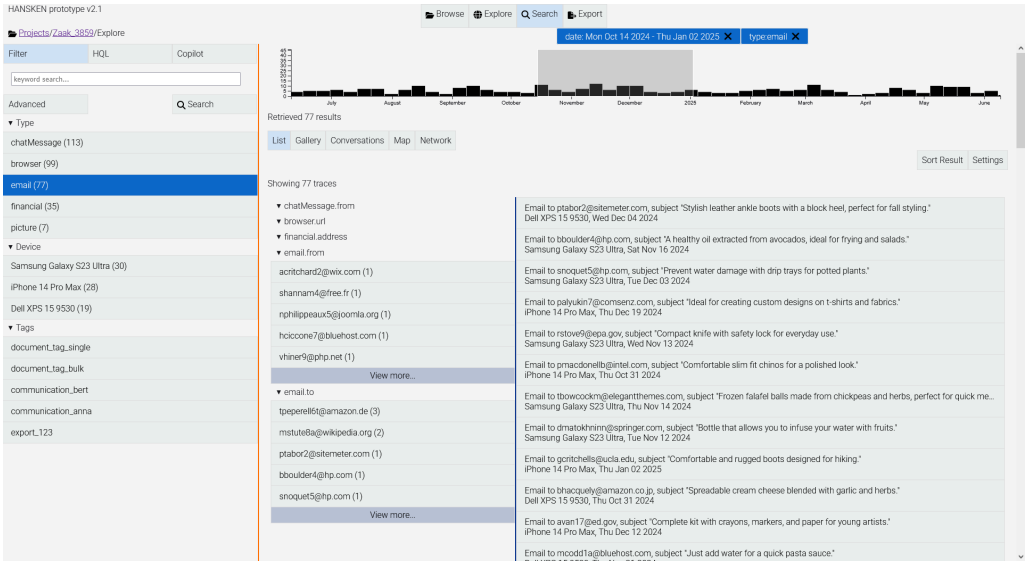


Figure 3.8: The tabbed aggregated search interface that was added in the third iteration. This view features a list of results with elaborate facet navigation, but also a gallery view, a chat conversation view and a map. This interface exists alongside the blended explore view, and users could switch between the two with buttons at the top.

3.3.3 Third Iteration: Refined Web Prototype

The prototype for the third iteration built upon the web-based prototype.

Feedback from the second prototype included participants claiming blended aggregated search is only useful for certain moments in an investigation, mainly at the start or when you might want to explore the data.

A tabbed aggregated search interface was added in addition to the existing explore view (see figure 3.8). In this tabbed view, users could choose between seeing a list, gallery, conversation view, or map of the results in the set. The list features more elaborate facet filters than existed before, making it possible to filter among other things on email sender & receiver or chat message sender. Chats in the conversation view were styled more familiarly.

Navigational cues were added in both the explore view as the tabbed view to help the user navigate between the two. Additionally, a small pop-up was introduced to tell the user how to navigate back and forth.

In the first prototype, users were distracted from the main exercises by missing or unexpectedly functioning interface elements. Several micro-interactions were added to prevent users from getting stuck. The prototype was tested similarly to the second iteration, but the exercises were extended to make use of the tabbed view (see appendix G).

3.3.4 Technology

The web interface is made using HTML5, CSS3 and Javascript ES6. A dataset was created containing 1436 dummy traces of 5 different trace types with unstructured attributes. Data wrangling was done through Python3.

Chapter 4

Results

This chapter is structured loosely around the research questions as described in 1, subdivided into topics. For each topic, quantitative results from the survey and qualitative results from the main study are shown. Quotes from qualitative research are translated from Dutch into English.

The number in front of a quote refers to the user ID of the participant. This user ID and other demographics of participants of the main study can be found in table 3.1.

4.1 Survey

The survey was fully filled in by a total of 14 respondents - 8 completed the paper survey and 6 fully completed the online version. Table 4.1 shows some demographics of the participants. An analysis of the full survey can be found in appendix C.

#	Role	Frequency	Digital skills
1	Tactical Investigator	Daily	Skilled
2	Tactical Investigator and Operator	Daily	Fairly Skilled
3	Digital Expert	Weekly	Very Skilled
4	Analyst	Rarely	Fairly Skilled
5	Analyst	Monthly	Fairly Skilled
6	Tactical Investigator	Weekly	Fairly Skilled
7	Digital Expert	Weekly	Skilled
8	Digital Expert	Daily	Very Skilled
9	Digital Expert	Daily	Fairly Skilled
10	Analyst	Daily	Limited Skills
11	Digital Expert	Weekly	Skilled
12	Tactical Investigator and Digital Expert	Weekly	Skilled
13	Operator, Developer or Administrator	Rarely	Skilled
14	Operator, Developer or Administrator	Daily	Very Skilled

Table 4.1: Demographics of respondents of the survey. *Role* refers to the role a respondent takes within their organisation, *Frequency* is how often respondents say they use Hansken and *Digital skills* refers to how skilled respondents believe themselves to be. All values were chosen from predefined list of choices (see appendix B).

4.2 Information Need

An information need refers to a gap in knowledge or understanding that the user wants to fill with information (see 2.3). In this section, results relating to information need or search goals are discussed.

All respondents from the survey know what they are searching for when doing research (Q6). However, it seems they might sometimes have difficulty actually finding it quickly in Hansken (Q7).

This is in line with qualitative responses from the main study. All participant indicated that at the start of their search, there was already some information available which functions as a starting point for a search. Examples of this information is details of a suspected entity or dates of an incident. Additionally, due to expertise and experience with a similar type of cases, a user might already have an idea of how to search. This information is then often used to further research in Hansken, for example by using it in keyword search (see 4.3.2).

> *2: "Before you start searching in Hansken, I think you already had a meeting with each other. So okay, what search word lists are we going to put together? What will that look like? But also, what are we even going to look for?"*
>
> *50: "If there is a case, you already have a direction of what you are looking for."*
>
> *35: "I usually start with certain search terms that seem logical in your investigation."*
>
> *89: "If you already have certain search terms like numbers or a date or so, then you first filter on that. If you still have way too many results, then you filter even more. Either make the time period smaller. Or only certain types of data."*
>
> *5: "You often already have an idea of what you're looking for. You have your search, you had your briefing, you know what it's about so you do have an idea, you know what the fraud is about, and you also know where you're going to search, because during a search and seizure you also want to take along good data. So then you must already know what you're looking for. And at the office it's then the translation into how do I find it."*
>
> *7: "Especially at the beginning you are still looking, do we find something about phishing, do we find something about bank helpdesk fraud, do we find something about false passports. And then we usually have in the back of our minds a bit which words we can use and in what kind of environment it occurs."*
>
> *46: "At a certain moment we had the moment that we were going to search for panel builders[1]. So then you just start with words that indicate that and then we get certain hits out of it. And then eventually you just go through such a chat."*
>
> *62: "We already had the action day. The modus operandi is often already known. We already have quite a lot of information. So you do know where you need to search specifically."*

[1] A person who develops and maintains phishing panels: the web pages and back-end used in phishing attacks.

Some participants stated that the ultimate goal with which Hansken is used is to find additional evidence that proves or disproves some suspicion. The actual actions to fulfil this goal completely depend on the specifics of the suspicion.

> *2: "We want to find evidence or exculpatory evidence. Then you search for that in different ways. So in that sense you need different search questions for it."*

> *35: "'That person or that person has stated this.' Can I find support for that in the digital data? Or can I verify or falsify it? And then you consult Hansken again for that."*

> *51: "Often an incident has happened, like a bombing we had recently. [...] There is often a suspicion of other suspects. Then they also want to look at the chats with those suspects. [...] Often there is a suspicion of who that is and actually it is then about looking for evidence."*

One of the participants - a strategic analyst - remarked that their goal is different. They are mostly looking for additional persons or case-transending information within the dataset.

> *46: "I think my task is actually always to look for new suspects. From the investigation itself they are looking for evidence within the ongoing case. So for the suspects that already exist they search for extra information that can contribute to or refute the story. But my role will indeed be much more like okay, but who else is interesting. And actually you would also want to know in what way they are working to see if we can put up some kind of barrier. [...] And from that selection you sometimes, depending on if it's not too many... you just start reading a bit like, well, what is this about."*

Other times, it is possible that a goal is already defined by others or is simple or clear directly.

> *54: "Look, in some investigations you don't need to do something super in-depth and then it's enough to just put a few simple findings on paper."*

> *62: "But we also sometimes get short operational requests, so for example if there's an action day and they want to check if the suspect is at home, then we do a telecom analysis to see if that person is at home at seven in the morning."*

Several participants stated that a specific information need has to be thought out carefully, depending on their current goal. The information need is determined by what evidence is needed for which claims, and is discussed in coordination with others.

> *7: "We often have an investigation leader and that person kind of directs like, well you know, this is what we want to get out of it, this is what we want to achieve with this investigation, and they also talk with the public prosecutor. [...] The public prosecutor therefore also determines what is needed for the investigation, because of course they ultimately have to bring it to court and they sometimes say, actually we need more of this. And then we look for a bit more of that, like for example a certain phishing panel we saw a photo of."*

> *35: "At a certain point you really start to put your file together and write it, and then you really look at what exactly do I have? What can I use for my file? What do I still need to find? What am I missing? And where might something interesting still be hidden?"*

> *89: "Especially now that Landeck[2] exists you need to have a goal. You must already indicate in advance to the examining magistrate, I am looking for this and I expect to find it there, there and there. In that time period. So that is in any case the data you get and that is the only thing you are going to see."*

Once a goal or information need is defined (consciously or unconsciously), the process of information seeking starts.

4.3 Information Seeking

Almost all respondents of the survey declare they often use multiple search attempts to reach their desired result (Q13), which might be through changing their query based on what new information they encounter throughout their search (Q12), specialization of a search query (Q11) or another reason.

Several participants indicated they generally start an information seeking process by globally reading through chats and submitting queries based on known information.

> *46: "And from that selection you just start reading a bit like, well, what is this about. And in addition you just throw those words through it."*

> *54: "In our tactics it's often really that we basically just enter a chat as it were and from that chat we just continue with what do we see?"*

> *62: "And then you can click through to extract the WhatsApp conversations, for example. So to go through the WhatsApp conversations a bit. Or to go through the contact list a little."*

> *35: "[The search in Hansken starts for me] always first with a general inventory. What digital data do I have and where is it located? And then you just try some search terms."*

> *68: "But if we already have a possible network in view, then we start interpreting the contacts and then you just go on to read the conversations. Scanning."*

Most participants believe that one search can then very rapidly lead to a new starting point: found information informs subsequent searches.

> *35: "I usually start with certain search terms that seem logical in your investigation. But then it often also depends on the results what you continue to search on. You can start at A and then at a certain point I end up somewhere completely different, that is*

[2] An arrest from the Dutch Supreme Court (Hoge Raad) in 2025. The judgement details that searching through digital devices is restricted due to concerns of infringements of privacy, unless permission is granted after assessment by a delegated judge: https://deeplink.rechtspraak.nl/uitspraak?id=ECLI:NL:HR:2025:409

the whole search process. It's not some fixed path, but you take kind of all sorts of... just wherever the results lead you."

5: "It usually starts broadly. And if you have a bit of an idea like here I need to search, then you look a bit further."

5: "At the moment you start investigating, it's actually a continuous process of: you start somewhere, you search a bit, you find something. Either you continue there, or you go back to the beginning because you see that it's different. So it's always a process of adjusting and possibly searching further in the results, focused on what you've found. That's actually a continuous process without thinking about it. What makes it very difficult in digital searching is that you very quickly get lost by jumping through too fast."

62: "Because interesting things often come by and you think, oh yes, this is also interesting, I'll keep reading a bit. The best thing would be to refine your search query of course, because then you get a more targeted result. But I just notice that I quickly get lost in clicking through and reading through. [...] the more you go through the data, the more you can pull out those needles. And on those needles you then continue searching."

7: "Then you've filtered somewhere or searched somewhere and then you look a bit further like, okay, in this folder there are interesting things. Well, then you go to that folder to see further what else is in there. So in that way you start to discover a bit what's inside."

62: "That depends on what comes out of the general search. If I find a hook in that, and that is for example a person. So suppose there is communication with a person, then my next search will be that person's name."

50: "And at that time we were just given specific points of this has to be investigated. And as you are working on that, you keep finding more like oh, we also have to look at this, we have to look at that. So gradually it keeps filling up more and more."

One participant remarks that they believe the current specifics of a case or subject might affect information seeking.

2: "I also think if you later have a completely different investigation, a completely different subject, that you will search differently again."

4.3.1 Context

All survey respondents believe it is important to be able to see a trace in its context (Q40.2), as well as to get extra information and details on a trace (Q39.2). While 67% of respondents feel they can currently see the details of a trace in Hansken (Q39.1), only 25% feel it is possible to see a trace in its context (Q40.1), and no respondents feel like it is fully possible currently. Respondents also seem to have trouble finding related or relevant traces around another found trace in Hansken (Q41.1), while almost all believe it is important for that to be possible (Q41.2).

In the main study, most participants stated that context is very important for determining the next steps in their search. One type of context is the context around a trace: finding the parent trace and/or finding which traces surround one trace[3] Another type of context is having information on one trace or a set of traces: meta-data on for example date/time, source device or file path. Both types of context are used in the information seeking process to get to new traces of interest.

> *46: "At the moment you've found interesting subjects who are criminally active then such a crypto trail often offers a very interesting possibility to trace them. [...] So they look at which crypto addresses there are. [...] And then you want to step from the addresses that exist to the context in which that crypto address was shared. So basically I always see those chats again. In the end you always come back to I want to read what it's about. You must be able to interpret somehow."*

> *35: "Sometimes you come across something very interesting that's on one specific USB stick. One item. Then for example you want to know... What else is on that specific USB stick? That's where I found that one interesting item. And sometimes it's in a certain search path, a certain part of a hard drive or whatever where you then find a very interesting document. And then you want to know, what else has someone stored in that place?"*

> *46: "I'm looking for panel builders, so I search on different variants of the word panel. Then you get a number of hits where that word occurs. Yes, then you want, you also want to read the whole context. Because maybe this one chat line is interesting. But how, in what context is that said?"*

> *54: "If that shows up in a chat, in which chat does it show up and make sure that if we click on it, that we immediately go to that chat and see who the participants are, from when to when the conversation runs, how many messages there are. [...] then it's also for example useful if it says where it comes from, so if it comes from a Snapchat conversation for example, and if you then click on the source, that you are automatically referred to that Snapchat conversation if that info is in the data."*

> *7: "Sometimes it's just difficult to know where a search result comes from. With that whole list and then you click somewhere and then you don't know where in a file the result is because that isn't shown."*

4.3.2 Filtering and Query Formulation

Respondents of the survey are generally successful in formulating a query in Hansken (Q27) and mostly believe it's easy to formulate (Q28). However, respondents are very divided on if they feel restricted in the ways you can search in the Hansken interface (Q16).

Nearly half of the respondents feel they need more help in formulating proper search queries, while the same amount of people do not feel this way (Q29). This does not necessarily correlate

[3]Traces can have a parent-child relation, for example when a document (trace) is also an attachment in an email (parent trace). 'Surrounding traces' can also mean -for example - other chat messages in the same chat conversation where one chat messages has been found.

to their respective technical expertise (technical skills for both groups have the same mean of 3.83 and median of 4 - skilled).

Most respondents spend little effort in understanding how to search or refine results (Q19). Moreover, most respondents believe it is possible in the Hansken interface to filter on attributes like time, type or source (Q33.1), use multiple filters simultaneously (Q34.1) and zoom in to a specific set of traces (Q36.1).

Most of the respondents have formulated search queries in HQL (Q18), and all but one participant believe it is possible in the Hansken interface (Q47.1)

All participants in the main study stated that electronic conversations - like chats and email - are often the most important type of data in investigations. Besides this, pictures or location data are often also of interest for some users.

> *51: "Chats are important and images are important. As a basis. Locations are also important by the way. Whether it's kidnappings, child pornography or a sexual offence case, the chats and the images are the first to be looked at."*
>
> *54: "For us, when we go searching in the data carriers it's mainly about checking if we come across lead lists, so victim data. Do we find call lists. Do we see that files are being sent via Telegram."*
>
> *62: "Yes, in my case it's really about communication. I'm looking for pieces of evidence, in which in the communication it is said like 'yes, we're going to do that' or 'yes, we did that,' those links."*
>
> *89: "You often see that you always end up with certain things, like the chats, location data for example, that is also often very important."*

When there is a large dataset, it is not feasible to examine all results. In these cases, participants indicated they would like to refine the data to a smaller and more specialized subset of results.

> *2: "You add another filter anyway or you just try another word. It really depends on what kind of investigation and what exactly you're looking for of course. But yes, you would definitely adjust. It's not the intention that we're going to look through 4000 results one by one."*
>
> *62: "But if you have a WhatsApp conversation of 800 chats, and you have to go through them all... It all takes quite a lot of time."*
>
> *35: "We really work with a lot of data in our investigation. [...] it really often comes down to terabytes of information and then it's very important that you can search really well, search efficiently, because you can't just say I'll quickly skim through it."*
>
> *50: "So certain search terms will already be defined in there anyway. [...] So you don't just immediately start looking through the chats. Because there really is a lot in there. That's impossible work."*

This seems to depend on the information need however, as several participants remark that it is sometimes important to look exhaustively through all the results of a query.

> *46: "And that's not, let's say, what we normally, what I don't normally do and also what Cyber Intel, which I mentioned earlier, the intelligence people who really sift through the whole seizure, they want much more context."*
>
> *50: "Yes, because you do want to conduct as complete an investigation as possible so you will have to go through all the results that your search query has produced."*
>
> *2: "It depends on how many results a search query yields. [...] But in general we check everything."*

Most participants stated that filtering a search happens through search words. Sometimes, a list of words relevant to a specific case and/or a type of crime exists.

> *50: "Sometimes. But mostly looking for a specific word. In this case what we are doing now. Or a specific number or..."*
>
> *54: "We just want to be able to do a targeted search query on what for us are known phishing panels. We want to be able to enter those in the search bar and from that we want results of does it appear in chats, does it appear in documents, is it stored on the device and if that shows up in a chat, in which chat does it show up."*
>
> *62: "In the investigation I am currently working on in Hansken, we actually work with keywords. We have certain companies and we then search for those companies in Hansken's data. Company names. And persons. [...] I must say that we ourselves also work a lot with keywords. Because of course we already know a bit of what we are looking for. Because we already had the action day so you know the modus operandi. And from there you continue looking."*

However, one participant noted that searching on keywords is not used for searching through chats due to the many ways things can be worded.

> *2: "For example we have deliberately chosen not to even use keywords within chats. Because, in WhatsApp you can really have a hundred ways of writing one word. So we mainly search for an email with keywords and basically that means searching for the general word and putting a wildcard after it. It's really just coming up with lots of words actually."*

Besides searching on keywords, filtering on date and time, on source device and on type of a trace is important for many participants as well. However, the types of filters that are used or are thought of as important are not the same for all participants or for all tasks.

> *51: "Often an incident has happened, like a bombing we had recently. Then we zoom in on the timeline at that time and check what happened there."*
>
> *2: "We don't do anything with location, date could be possible depending on the results. [...] What we mainly do is chats data. So a certain time period. Other filters if we're looking at email. But that's basically it. Occasionally some extra keywords."*
>
> *41: "Time span is always important. [...] That you really know in which devices you are working. Because that is actually the first question: where do I want to search my*

information? Yes and then the types."

7: "That date-time, that is really very important. Also on device, that you can check that. Which devices yes and no, that's something you also use a lot. Because we sometimes have four, five suspects in one investigation. And then you want to check, I only want the device of this suspect and then you switch that on or off."

46: "I think that in a case, that time is mainly relevant within one case. Because you are working there on one incident or multiple incidents that happen at a certain moment in time. When it comes to selecting new suspects, then that time is much less relevant. [...] you do want to be able to filter a lot on the type of chat program. That is still important. That you can quickly fix that. And in my case time is again not so important."

68: "And also for example in terms of location [...] that is of course quite interesting for us, the link with the Netherlands in many cases."

One participant stated they are unsure if they can properly translate their information need into a search query.

62: "I still find it difficult whether the system and I are speaking the same language, whether the system understands what I mean and the other way around. I have something in mind and I enter it and then it gives a result. And then I assume that that is the only result."

Filtering data using HQL is seen as difficult by some participants. It was said that using it is not intuitive and poses a barrier for entry.

68: "In those cases where you have to use that query language, I really find that still quite complicated. And also how you then... you can look it up somewhere else. Then I think well, never mind."

46: "But what you also don't want is that you first have to master some kind of query language. Because yes, new people join the team and you also just want them to be able to quickly work with a certain tool. If they first have to learn that language and you don't use it continuously... We don't have new data to dig into every week. If you only use that language a few times per year, so to speak, then you just lose it. Then you don't remember it anymore."

7: "But the HQL is not always intuitive to use. [...] The searching is very focused on the HQL. It's very HQL based. I think that for the everyday investigation teams that's just not convenient. That the most focus should really be on just plain text searching. Or that you keep HQL more as a kind of addition, without them having to learn HQL themselves through certain suggestions or checkboxes, that kind of thing."

54: "I think that for a lot of people from the tactical side, the HQL, that maybe at the beginning it kind of stands in the way of using Hansken."

On the other hand, some participants indicated they see value in this form of querying.

62: "Yes, I had the idea that with that HQL I could search a bit more specifically."

7: "You can usually figure out HQL if you know all those little things like oh I need this section, here I have to filter and then the HQL. So the option is certainly useful to keep in."

4.3.3 Search Goal Decomposition

No respondents of the survey stated that Hansken aids them in subdividing complex questions or search queries (Q45.1). However, 38% of respondents believe it is important that a system can aid users in this (Q45.2). 57% of respondents state they often face complex goals or hypotheses, which they themselves have to divide into smaller, actionable search tasks (Q26). Only 23% of respondents believe its important for the interface to aid in structuring their thinking process during search (Q46.2).

In the main study, several participant state that they only use simple functionalities and do not use Hansken to its full extent.

2: "We are now really searching very simply."

68: "And I also think that the search queries that I still use in Hansken are fairly simple."

50: "[...] we often discuss in advance the search queries that belong to a certain goal and then we just both go through them."

None of the participants indicated that they had trouble with decomposing their complex goals into actionable, smaller tasks like search queries.

5: "I don't really do many hypotheses myself. [...] I think that happens unconsciously. [...] I don't consciously think: I'm going to take smaller steps or something. It does happen, but unconsciously."

50: "With [colleague] I work very well together so we often discuss in advance the search queries that belong to a certain goal and then we just both go through them. So that basically always works out."

89: "Actually you always divide your investigation that way, at least I do. And that is purely based on if you want to prove a certain thing. Suppose you want to prove that someone used their computer at a certain time. Then you can prove that in various ways. I usually first look in the place where I expect like well, it's most likely to occur there and from there I work further out to other types of data where it could also possibly occur. [Do you find it difficult to break down that bigger goal into smaller goals?] No, generally not."

2: "Before you start searching in Hansken, I think you already had a meeting with each other. So okay, what search word lists are we going to put together? What will that look like? But also, what are we even going to look for?"

4.3.4 Overview of the Data

85% of survey respondents believe it is fully or well possible to gain a clear overview of all data in a project (Q32.1), which is also seen as something important by most participants (Q32.2). Respondents are divided but mostly positive on whether they feel they can currently gain a quick overview on the relevant data or search results in Hansken (Q30.1), while this is deemed important by all participant (Q30.2). A visual overview of search results is important to 62% or respondents (Q31.2), but only 25% believe it is currently possible in Hansken (Q31.1).

In the main study, some comments were made about the current Hansken dashboard, which contains functionality for selecting and deselecting devices as part of a search globally, and shows some bar graphs on the amount of different types of traces. Several participants stated they like the possibility to select and deselect devices to be included in search, but did not believe the visualizations were of much insight to them.

> *2: "It gives me an overview, but we don't do anything with it further. It is nice though that with that you can also switch devices on and off so that you clearly see it. It's not like we now think oh we see so much communication."*

> *46: "You get to see it all. You can switch it on, off. So you have the overview of what is all there. [...] That there are I don't know how many 100,000 images on it doesn't really tell me much."*

> *50: "I know which devices are in there for example. So I don't click through from there to something, but I do switch off certain devices if I'm specifically looking for something. But with all those little graphs I don't do anything."*

> *7: "What is normally on that start page now, that is really not important for most users. Those little things, they're in there with those nice tables but that's of course not really very directly clear, let's put it that way. [...] I don't think anyone cares how big or tall a table is, what's in it now."*

One participant remarked they find an overview useful if it was about the contents of the data, not the amount of data (as is reported on the Hansken dashboard currently).

> *7: "That you immediately have in one glance like, [this data] is in there. Getting a bit more out of the content of the data, instead of the amount of data."*

Some participants indicated that a lack of overview makes them doubt whether the results set contains the data that they seek.

> *62: "Eh, I still find that a bit difficult. Because just like when I enter a general keyword, that I then get such a huge list. Sometimes I wonder: am I really seeing everything? Or am I missing it because there's so much, and I don't have the overview?"*

> *7: "Because now it's sometimes just a bit of guessing like... I have 2000 results. What's in it? [...] The first 50 results are roughly this. Then I know I should discard this. [...] Then you think 2000 is too much. I need to try something else."*

Many participants say that categorizing data helps them. They highlight the importance of being

able to get certain overviews of all the data that exists in Hansken, but also the importance of an overview for results set of a single search.

> *54: "We see now for example in [another DFT] that this is just categorized. You have the image of the phone and then you have a number of categories: so you have media, under that fall videos, photos, audio files, or you have the chat, or you have the mail, or that kind of things. In the chats you then again have Snapchat, Telegram, that whole thing."*

> *68: "That that then is a results list with like these phone numbers appear in multiple phones. But that you then per phone again have to check, okay with whom does it appear? That you don't have that in one overview right away."*

> *62: "I do miss a little bit of the categorized aspect in Hansken. If I ask a question [...] then I get a whole range of thousands of answers, but I can hardly figure out if those are e-mails, or if those are WhatsApp conversations. It's just one long list, one long mush of answers. [...] The answers [in another DFT] are a bit more categorized. [...] I can see like it is mentioned 20 times in a WhatsApp conversation all at once. I get more out of that than getting a whole list with 10,000 results."*

> *7: "That [it is not intuitive] I also have with the search screen. I'm going to say it very bluntly, but that is completely a disaster. You see how many results there are, but not what those results consist of. [...] And actually you just want to see, this many results come from internet history. [...] Then it's nice to know, is [a type of data] in there at all? How much is in there? And if you already know there are so many chats in there, you can easily apply a filter like I don't want chat messages."*

The existing Hansken category views (see 2.2.2) are seen as useful and clear, and are used often among participants.

> *7: "That sidebar really does give a good one, and I could even show you with [another DFT], that also just has a lot of those main classifications with those sub-ones and such."*

> *35: "What I use a lot, is just the search thing and then searching in different ways. What you have in that left bar... For example that timeline, where you can search within a certain period. Or on a specific type of data. [...] I do that every now and then, and then it's also a bit of exploring, so I don't have the sort of fixed idea like... But sometimes you discover something like 'this is actually quite useful, I can do something with that' and then you start using it more often."*

> *54: "This is in itself I think a fairly clear. I think that they are logical... accounts, communication, chat conversations yeah I think these are pretty logical categories."*

4.3.5 Combining Filters and Views

As noted in 4.3.3, several participants noted they currently only have simple search tasks in Hansken.

Most respondents find they sometimes formulate complex searches, like using multiple filters simultaneously or using AND and OR statements in HQL (Q14). All respondents feel it is currently partially to fully possible to use multiple filters simultaneously in Hansken (Q34.1), and the majority believe it is important or very important to be able to use multiple filters to refine a query (Q34.2).

Participants of the main study often use combinations of both different search methods, filters and displays.

> *35: "The more specific you want to search, the more you want to funnel it, the more filters you want to put on it."*
>
> *62: "Yes for example a country with a person. Or a year with a company."*
>
> *68: "Yes, depending on how many results come out of that, that we then go and specify by date for example. That you then look at the most recent ones."*
>
> *68: "That you just know like okay, I want the results that contain this word, and this word, but not this one. And then you have certain combinations with that which you can then exclude. Which then give a lot of noise in your results."*
>
> *89: "If you already have certain search terms like numbers or a date or something, then you first filter on that. If you then still have way too many results, then you filter more. Either make the time period smaller. Or only certain types of data."*
>
> *50: "But I do find such a timeline really handy to have. And also if you could then from that timeline click on such a bar and that you then see which photos fall under that for example."*
>
> *51: "Sometimes you're up against that you want to combine image classification with photos in his camera roll for example. So that you then think like: hey how do I combine that. [...] And I can't combine those fields yet. So I do run into that sometimes."*

4.4 Interfaces and Interaction

4.4.1 The Hansken Interface

There were no strong opinions and much division on whether survey respondents believe the Hansken interface is user friendly (Q17). Respondents are also very divided on whether they feel restricted in their ways of searching in Hansken (Q2.16). Almost half of the respondents wish they had more support for the formulation of good queries, while an equal amount of respondents does not feel they need more support (Q29).

Respondents are similarly divided on whether they believe Hansken overloads them with unsolicited details (Q42.1), while they are of the very strong opinion that it is important for a system to not overload the user with unsolicited details (Q42.2). Respondents generally do not feel a high cognitive load while searching in Hansken (Q20).

Most respondents believe the current searching functionalities align well with their information need (Q8), but also believe that Hansken lacks certain search capabilities or filters to find the right

information (Q9).

From the main study, participants are generally very positive about Hansken, praising its ability to handle large quantities of data & devices.

> *5: "Hansken is, shortly said, good at searching fast. It is fast. Can handle large amounts of data."*
>
> *7: "The big strength is just that you can load a lot of data and that you can throw a lot of questions and words and that kind of things at it, and you get results quite quickly."*
>
> *89: "Well because Hansken is mainly nice to be able to view a lot of things at once. And also to be able to compare them with each other."*

Several participants stated some opinions on interacting with the different interfaces HTI and ExpertUI. They state the interfaces are not easy to learn or use.

> *62: "I must honestly say, I just don't find Hansken very user friendly."*
>
> *68: "Yes, in itself, this is not a very easy program to use. [...] I think if you were to use that [trace details window] for the first time without any prior knowledge you'd be like, well, forget it, I don't understand anything of this. You really have to think well about how you interpret the outcomes."*
>
> *35: "And that Expert interface, I still find that a bit difficult to use. So sometimes I try it a little bit. But then I also sometimes drop out because there... You need to know a bit more to use it properly."*
>
> *7: "I do find the expert UI somewhat better, because then you can better... Yeah, there's less around it. It's a bit more to the core."*

4.4.2 The Prototype Interface

Respondents of the survey are split in whether they would like to use their mouse to formulate a query by clicking (Q22).

In the think aloud study, a prototype was shown to the participant (see 3.1.4). This section contains results for all iterations of the prototype (see 3.3). The main prototype interface - the explore view - was tested with all participants while the tabbed view discussed later in this section is only available to participants in the last iteration.

This prototype contains several ideas and concepts that were tested in the interface. While some elements were seen as positive improvements, others were not understood or not seen as contributing.

Participants of the main study had some remarks about the concept of a categorical layout in the explore view. Some participants stated it helped them gain oversight, while others did not like it. When asked if a participant sees value in aspects of the explore view, many respond that there are use cases when it would be interesting - mainly at the start or orienting phase of research.

> *50: "I personally prefer it to be divided into categories."*

35: "Hansken doesn't have such a kind of screen as a whole. So that's then an addition. I think that data is there, but maybe more fragmented. And here it is all together. Does that give more insight? Yes, in principle I think so. [...] as a starting point from which you go further, I think it is helpful."

62: "Well yes, because you see the categories already clearly in the start screen. Actually you can at a glance see what data the program contains. Like I said, you then see the timeline, you see that map, so actually in one overview you already immediately see where for example hotspots are or which months might be interesting as it were."

51: "That you also immediately see which traces belong to what? And for example also what happened at a certain time? That would for example be very useful with that bombing for us. That you immediately see what all happened at that moment."

7: "But if you then really go to the specific points, that there is again a kind of separate screen, because otherwise you do get a lot in this dashboard. And otherwise you also get distraction from the rest,"

2: "For ourselves at this moment not, but later it could. But that's because we indeed always only look at one part. So either those chat messages... and not really use the combination much yet."

54: "I think at the start of an investigation I would use this, let's say. I think at a certain point as the investigation funnels more, that I would then just go search with a stricter filter. [...] So I do think it is indeed an added value for the interface."

89: "If you for example don't immediately know what to search for, then this is, or such a kind of screen is nice. To see what your top 5 people are you called with or sent chat messages with for example. Or the locations someone has been, or when and which month that phone was most active. So then those things are certainly useful."

68: "I think if you just start with your investigation and you want to know at all what is actually in that phone? That it then indeed gives a quick overview. [...] If you indeed want to search very specifically in a phone it is quite an easy way to search."

The explore view features different verticals. Three participants remarked on what verticals they believe are important to have, and which are not important for them.

41: "Timespan is always important. That it directly shows something and gets adjusted with filters right away. I find that very interesting. That you know well in which devices you are working. [...] Because that is actually the first question I think, where do I want to search my information? [...] What we work with, at this moment we are not at all interested in the photos. So if you say like I can also click away this image or a certain format, because for this investigation I am not interested in gallery photos. Then it's a waste that so much space is taken here which still draws your attention."

5: "But I do find that the things that are in it... Top browser, top financial, conversations, types, devices... Those are indeed the things that are often looked at."

> *35: "For example the location is not always relevant. Sometimes you might say, I don't really need to see that, I actually want to see something else. But in itself such an overview... Can be useful at any time, yes."*

Looking at the ways of interacting with the interface, participants had mixed reactions. Most liked the interaction of clicking on facets to form a query, believing its intuitive and easy to understand.

> *2: "I do think that filtering with just a simple click like, I just click all this on and I can do several at once. I find that really works fine, yes."*

> *5: "I do think that this is very easy to learn. [...] It is very easy to just select something from a device, one or two. Here you see some frequently occurring categories, so to say. You build here relatively intuitively, easily a query. You just click on something visible and results come up and you can refine them again. And that goes very fast, very easy. I think if you've done it for a few minutes, you easily find your way. [...] You don't have to invent anything yourself as a query. You click something and you get it."*

> *50: "I find this nice to work with. That you can immediately easily click, click again and it goes away again. And I do like that you immediately see a preview of a photo."*

> *89: "Yes, that was very intuitive. Just one click is on and the next click is off. With those bars at the top you immediately see which filters are on, so that is very clear."*

> *54: "What is mainly important for the tactic is that you don't have to ask the question in a sort of programming-language-like way. That it must be fairly simple, clear. And that there is a kind of layman's function and a deep function. I think that's the most important for the tactic."*

> *51: "Clicking is much better. Then clicking is really a priority."*

However, many participants stated they had trouble interacting with the timeline specifically. In the think aloud study, most participants did not understand how to interact with this element the first time they encountered it.

> *51: "Only with the timeline I did have to get used to it. Because normally that's something you kind of want to stay away from you know. The same as if you again start dragging here or that kind of things. But that is also a bit what I have in my head with a timeline. And because of those bars it also kind of looks a bit..."*

> *89: "And where I think I initially had the most difficulty with the timeline, was how you could indicate the period. [...] That dragging of the time bar, that could be a bit more intuitive."*

> *41: "Timespan is one I do find ok. Because you can't in everything... So if you have a certain timespan it is really nice to very quickly, and then with that mouse in that way – I didn't realize at first either, I first wanted to do it here you know – but that you can indicate it with the mouse. That you can really like zoom in to a month and then*

a week for example. And then maybe even a bit to hours."

51: "I was fiddling with that date filter. I was missing a bit of a filter icon. Because this looks very much like you can click on it huh. Like as if they are clickable buttons. But I think you actually have to drag a bit, so that it then selects a kind of range. Only, that you drag across a timeline, that is not really how I would use it... What I would expect. Because if you drag, well, think about in a Windows explorer, then you drag icons everywhere. And you actually don't want to mess around with your icons and that kind of things."

About the map interactions in the explore view, one participant stated they knew how to interact with it because they are used to working with maps.

51: "With the map anyway, because you are also used to that right, with Google Maps. Then you are used to doing it that way already."

None of the participants indicated that interacting with the prototype induced high cognitive load. They did however indicate that some interactions caused frustrations.

41: "No like you often do you have to read the questions well. I did make a mistake once. No mentally... I didn't really fall off my chair or anything..."

5: "That I have to think because I can't find some things, yes. If that's what is meant, yes. [...] Not irritated and not discouraged. It's maybe a disadvantage if you've used other search engines because you expect it to work in the same way a bit. And sometimes it just doesn't."

50: "It's not too bad. Actually most went smoothly once I figured out how to select a certain time span. But yes, the tasks weren't very complicated either. [...] Not insecure or discouraged. I think irritation comes quickly in small things like if you can't click something away."

62: "It's just figuring it out for now because you're looking at the program for the first time. But once it's explained once, then it speaks for itself."

89: "Well, it is of course a bit of searching, because you're working with it for the first time,"

51: "Well, more like I'll just make the best of it. And if I don't find it, then I don't. So kind of that."

2: "Insecure, yes, but I think that's just because I don't do this much."

35: "No. That goes too far as an emotion. Sometimes insecure, because sometimes I couldn't find things."

A concern of some users is scalability of the interface and if it can still provide the same level of insight with more and more diverse data.

2: "But [the interface provides overview] because here you are really only dealing with very few things. I think as soon as you do have thousands of files, it would really

give me an error."

5: "I don't know if this would work with a lot of data. [...] You always have to limit the number of [facets] of course. If there are again too many you also overshoot your goal. [...] Then you actually just get a category overview on the side again. But I do think that's handy."

51: "Because I do have the feeling with this layout, it can very quickly become a lot of course. If you then indeed really have a lot of chat messages, then it could possibly be a bit of a search."

7: "But if you want to put too much in that dashboard, I think it becomes a bit too busy, a bit too unclear and maybe sometimes you even end up with too small screens."

One participants notes that an explore view and a grouping as currently exists in Hansken - and was tested in the tabbed view for the last iteration of the prototype - would both be useful in different circumstances.

62: "I think I would use both. The one really for keywords. The other to get a global overview: This is what is generally there and with the other it is more like: This is a keyword and it occurs on this device. I do miss that a bit in Hansken as well, how I see it."

Similarly, another participant suggests having an explore view as an addition and not a replacement of the current categorical groupings.

51: "It's also a bit difficult to compare them with each other, I think. You could maybe add it as a tab."

The prototype interface contains an advanced search panel for creating more complex keyword queries. Many participants remarked that this is similar to how a query in Hansken can be made, but that it is better than creating a complex query in HQL. The fact that all possible options are presented and that interaction is simple is appreciated,

68: "Yes, I think this does make it easier. So that you don't have to think yourself. The bracket stands for this, the star stands for that. That you just know like okay, I want the results that contain this word, and this word, but not this. And then you have certain combinations with that which you can then exclude."

35: "But I find as it is here, contains, this or that and not, and I find that clear. [...] Because [in Hansken] it is a bit more complicated to use and I find this very clear. You can just click it. That is fairly easy so to say. So user-friendly so to say."

5: "This is kind of the way queries are made now. In itself it's not wrong, but we will have to explain what it means."

50: "I find the way it goes now, that looks a lot like this to me. [...] That is indeed a lot clearer. There really is such an overview form with which wildcards there are. But you forget that again and this is here. And they are just words instead of symbols so that is nice."

> *2: "Especially because we only later discovered that wildcard with that star. It can be really handy for us if for example something... That you have that starts with [in dropdown]."*
>
> *41: "Yes, because I think if you convert this to here you know, this is just, then you have to be a certain specialist to know that and this is actually quite simple where you can play with quotation marks or, and then I don't know how people start with this, but it's actually nice if you can already through a bit of brainstorming, which question do I have? to write that out. And then to see like, well can I gradually build that with this. This must actually be displayed very simplistically, so that you can ask more difficult questions that way. Or actually ask multiple questions in one. This is a good way for that."*

In the last iteration, a pop-up was shown to the user with a message describing how they can navigate the interface. All participants indicated that this pop-up did not provide them with any useful information.

> *35: "Yes, there was a little screen and there was some text in it. I looked at it briefly, but I don't remember what it said. I found it in the way, so I clicked it away. [...] pop-ups and such, they often get in the way. [...] So that there [a pop-up] appears and I then think: yeah, but I wasn't busy with that at all or I'm not interested in that."*
>
> *5: "I did see it. If you ask me: Did you find it useful information? I saw it and I think okay, click away. I often get error messages from applications and I just click them away."*
>
> *62: "I did see it, I also read it. I first like to check for myself actually."*
>
> *89: "I didn't read them all. [...] Sometimes I was more interested in answering the question than reading that window."*

4.4.3 Natural Language and AI Agents

Most respondents of the survey would find it useful to use natural language to form a search query in Hansken (Q2.24), but only 29% would trust suggestions or answers of an AI-assistant if it was built-into Hansken (Q2.25).

All but one participant of the main study stated that they have previously had interactions with AI-agents using natural language interaction. Of these participants, they all had positive views on the usage of these agents, while also recognizing its problems.

> *5: "If you ask that chatGPT and there was a point that didn't work... Asked the question, the answer wasn't correct for chatGPT. Asked again, still not quite right. And eventually at the third question the answer was right. Because of course... You ask a question and it goes further with the results and adjusts it. That is of course amazing."*
>
> *46: "Nice and easy. But my experiences so far are just not that good."*

When asked more specifically about the interaction style of communicating with natural language,

most participants like it but some are more comfortable using keywords.

> *54: "I do find it a nice way of asking a question."*
>
> *50: "Yes definitely because that is the most natural. That is how you also think in principle. So no I think that really would be a good idea."*
>
> *46: "I actually never really trust it to be honest. I always have the tendency to just put down keywords. Or I think those all have to come back in the answer or that's what I'm looking for. So I almost never ask it in a sentence or something indeed."*
>
> *5: "The answer comes out. Without you having to ask exactly the right questions. And that is of course great for investigators. And that is quite a problem now. If they search for something, but you use the wrong terms, even if you're close, you don't find what you're looking for. Whereas if it's chatGPT-like, AI-like, then despite not using the right search term, you already arrive at the right thing. That would be nice. I'm in favor."*

When talking about trust in the output of an AI-agent, the participants stated they are cautious to assume statements by this system are correct

> *89: "You should never trust it. You can ask the question, but you still have to go back yourself afterwards to check if it is actually correct. [...] I've had it multiple times already, with chatGPT for example. That it gave completely wrong answers. That nothing was correct. And then you say that afterwards to the program and then it says, oh yes, you are right, thanks. So you always have to check it yourself."*
>
> *51: "I do know that it's a data language model and that it pulls from different databases... And from that it makes a kind of prediction. So whatever you ask, it always gives you some kind of prediction. But it can give context. Of course, you have to fact-check whether it's actually correct. And never sensitive information."*
>
> *5: "You always have to check what you have anyway, but if chatGPT gives something as an answer, I would definitely check whether it's correct. Regardless of whether that's in an investigation where such an AI is used or outside of it, you have to check your answer. If I then come to the same conclusion, I would trust it."*

When asked about interest in using an AI-agent in Hansken, many participants were optimistic, but would not trust its output directly, even more so in the context of investigative work.

> *46: "If it demonstrably works, I'll definitely use it. [...] If someone builds something like that for a consumer like us, then you also want it to actually work properly. So in that sense, I would trust that something is produced and put into production that really works well. [...] Yes, and in the end you're of course just going to test it."*
>
> *5: "And that could be super in an investigation if I don't think about possible limitations on the amount of data it goes through or whether it's reliable and such. But if you ask me, would you want it? Well, I think everyone would want it. [...] You can ask a question easily and maybe get steered a bit in the right direction. If you ask me,*

does this actually check out? I don't know."

7: "For me, no, I would not use that at all. But I'm actually the one who sometimes uses HQL, and I think most colleagues wouldn't use that."

2: "You want to be able to rely on having seen everything you were supposed to see. And I've always thought: then I'd rather search with ten different words, and then it's my own responsibility. And then in an official report I can write: I searched these words and I found this. If I do an open question, I have no idea what it looked at."

35: "No, I wouldn't trust it at first. I'd always want to use both side by side, and then see what it yields. If at some point it turns out experimentally that it yields the same, then I'd start to trust it. But in the beginning, I don't know what it does."

50: "If I found out that I was missing something, I would no longer trust it. Then I wouldn't use it anymore either."

54: "That has to be proven, of course. A test has to be run first. And that again depends on the investigation."

62: "I would then have to investigate whether that trust is actually justified. But at first I do assume that the system understands my search query."

68: "I would always check that. So by asking it and then searching in Hansken myself."

Some participants expressed other concerns regarding AI or natural language interactions.

54: "For some things you just have to dig a bit deeper in the data. Then I don't know whether such a natural question is really the right one. Whether the system fully understands what you mean or what you're looking for, as it were."

89: "Of course it's nice to be able to ask questions in that way. The only question is, does someone know the right way to ask a question? So that you really get the answer to your question. With AI it's often that people ask the general question or not the right question to get their answer."

Several participants believe they would trust the system more if they receive feedback on its actions.

51: "Well, you saw that filter there of course. And the answer to it, so those two together, that gave me enough trust that it was indeed correct."

54: "That would give me more trust, and it also gives you insight in how I could ask the question differently, or how I could get more out of the question, so to speak. I think if you get a kind of feedback loop."

41: "I would personally, especially in investigations with evidence and so on, really always want to know that step: what was done, how did I get there, with those filters. To explain it."

89: "Yes, or to indicate: did you mean this? Or did you actually mean that? That the

system gives multiple options of what it should search for. Or what the intention of the question is."

In and after the think aloud exercise - which contained natural language interactions with a copilot - participants expressed several other thoughts.

51: "I think that Copilot is very useful, that you can ask questions in it."

2: "And especially that confirmation. That you know, okay, it didn't accidentally pick the wrong iPhone if I had two. [...] I think that's very nice, yes."

41: "So I would personally, especially in investigations with evidence and so on, really always want to know that step: what was done, how did I get there, with those filters. To explain it."

5: "In the Copilot I think it's funny. Whether I trust it, I don't know."

50: "Yes, I find it useful. I would definitely use it. [...] I do prefer to formulate my search queries myself."

68: "I think, with how... The question I just asked, I think, well, I could've done that myself just as fast. [...] That you can basically ask the program a normal human question, and then it turns that into programming language. Then I would definitely use it."

Chapter 5

Discussion

The goal of this research is to answer these research questions:

> RQ1: What is the information need of Hansken users interacting with the Hansken Tactical Interface?
> RQ2: How do Hansken users search for information in the Hansken Tactical Interface and how do their search behaviours and use of available search capabilities align with theory of information seeking?
> RQ3: How do specific interface features and interaction styles influence cognitive load, perceived usability, and effectiveness of information seeking in a Hansken prototype interface, and what design implications emerge from these findings?
> Sub-RQ3: How are natural language interactions with AI agents perceived in terms of usefulness and trustworthiness for information seeking and task decomposition?

In this chapter, the findings are discussed and related to existing literature. Then, limitations and contributions of this study are stated. Lastly, possible future work is discussed.

5.1 Findings

The aim of this study is to understand how users search in Hansken and how search in Hansken can be enhanced. The following sections address the different topics that were researched and relates them to literature described in chapter 2.

5.1.1 Information Need

The results from this study suggest that participants have a highly varying information need or goal based on their organization, their role within the organization and the current criminal case(es) they are investigating. All participants generally have a lot of information available to them to inform their need: professional expertise, previous experience with similar cases and known case details.

Some participants noted that the ultimate goal of why Hansken is used is to support or to disprove a claim using digital data: searching for evidence. This is also described in literature as a general goal of any investigation: to gather sufficient information to determine if some person should be indicted for an incident (Andersen, 2019, pp.5–7). The information need can be described as a

specific question that arises during the investigative process, for which digital traces can be used as an answer. This need is an attempt to reduce uncertainty surrounding some hypothesis, event or suspect.

It is possible that other users of Hansken - specifically analysts - have different, more emergent or explorative needs, where a new hypothesis is developed by analysis of the data. This is the case when data is analysed for finding for example new persons of interest or understanding a modus operandi. The information need here would be finding patterns or relationships that may reveal new aspects of this case, of this category of cases or of another case.

The information needs described above can be summarized as falling under either confirmatory or exploratory research, where confirmatory research is done to test and existing hypothesis and exploratory research is done to generate new hypotheses.

The IDFPM (see section 2.1 and figure 2.1) describes the process of digital forensic investigations (Kohn et al., 2013). While this framework is flexible enough to also describe the forensic processes of analysts, it does not explicitly take into account the exploratory research that seems to precede hypothesis generation for these users.

5.1.2 Information Seeking

Participants indicate they have some knowledge available which can be used to guide their search process. First of all, specific information related to details surrounding a case, incident or entities involved can be used: dates and times, names, phone numbers etc. Second, experience with similar or related incidents can be used in the search process. Searching for terms related to the type of incident or names of known resources are known to be used in these kind of incidents for example. Third, expertise of the investigator can inform where to look for different kinds of information. This entails a general understanding of forensic methods, criminal behaviour and digital devices.

Participants use this information for different search strategies in Hansken, depending on their current information need. Participants indicated that searches often start with actions like scanning through electronic communications or entering search terms. From there, new information is discovered which is then used as a stepping stone towards new search directions. This aligns with an idea brought up by some participants, that information seeking in Hansken is not a linear process but rather a continuous cycle of ‘hopping’ from one query to another, based on new traces and directions encountered in the data.

This process resonates with the berrypicking model, in which searchers move through an information space, learning information and continuously reformulate their queries as their information needs evolve (Bates, 1989). In Hansken, searchers similarly collect clues, such as names or fragments of conversations, that redirect the current focus of the search and information need.

Participants describe quickly jumping between different traces or contexts based on found information, and revisiting earlier starting points. Two participants also note that it is easy for them to jump too much, while others describe it as a choice of which scents to pick up on. This seems to imply there might be some economic function between the decision to explore one context

or to leave for another, resembling the exploration-exploitation trade-off in information foraging theory (Pirolli and Card, 1999). Searchers act as "foragers" navigating through an environment, continuously evaluating information "patches" and deciding whether to exploit or leave them in search of other contexts. However, participants describe that sometimes all results need to be exhaustively searched through, akin to systematic search. This is also described in section 2.3.4. Applying information foraging here, exhausting a patch even if the information scent is weak, the efficiency-logic of theory does not apply as well.

The interplay between exploratory browsing and focused search is also described by some participants (R. W. White and Roth, 2009, pp.16–21). At the beginning of an investigation, participants often engage in exploratory browsing - tentatively reading through chats or scanning for relevant leads. Once a potentially interesting trace is found, the search becomes more focused, looking at the direct context of that one trace: it's meta-data but especially traces surrounding it. From there, a new exploratory search could start again based on found information, or the searcher may return to their original search.

To this point, it has mostly been found that participants continuously acquire knowledge during the search process, which reshapes their information needs. Of course, when an information need only requires simple question-answering or known-item searches this process looks different. These simple search activities fall under lookup search, while more exploratory modes of search which evolve through learning can be described as learn or investigate modes of search (Marchionini, 2006).

As the discussions about information seeking behaviour clearly shows, the exact search methods or strategies that are employed are completely dependent on the information need and the goal with which the search is started. Berrypicking can be observed between exploratory browsing and focused search in case of a vague information need or unfamiliarity with the data. At the same time, lookup activities are done when the information need is well-defined and simple to answer. If an information need requires thorough review, a systematic search might be started.

When asked about a participants search goals, and decomposition of them into actionable tasks, participants do not seem to have any trouble doing this themselves or in collaboration with others. Less than half of respondents believe an interface should aid in task decomposition, while most face complex goals or hypotheses that need to be broken down into actionable search tasks. This is an interesting contrast with literature, which suggests that organizing task compositions for complex search requires significant effort (Marchionini, 2006, p.42).

5.1.3 Use of Search Capabilities

As forensic investigations often combine many pieces of data, the datasets can be very large. This is also true for Hansken, which is often used specifically when there is a lot of data. For making a smaller, more relevant set of results, participants often filter the data. There are multiple ways this is done.

It seems that keyword search is a vital method of zooming in on items of interest. HQL is another text-based way of filtering data, but is seen as unintuitive and complicated, though also powerful.

This is completely in line with literature on command language interaction (see section 2.4.1). On the one hand, keyword querying is intuitive and easy to express for all users but on the other hand, a command language is unfriendly for novice users, inflexible and has heavy reliance on remembering it (Baeza-Yates and Ribeiro-Neto, 1999, p.280).

One participant noted they do not use keyword search in chats, due to the many ways to write a word. This is one of the many challenges that can arise in query formulation - in this case the vocabulary problem, specifically term mismatch, where many terms can refer to the same concept (Spoerri, 2004). Another participant states they are sometimes unsure if they can properly translate their information need into a comprehensive query. This is a known challenge for users in case needs are ill-defined, where a transition from conscious level to compromised level is difficult (R. W. White and Ruthven, 2006, p.934; Taylor, 1968, p.182).

In Hansken, if there is interest in a certain type of data or representation of the data, a user can choose categorical views or use the search functionality to filter the data. Literature describes this as information seekers desiring organization of search results in meaningful groups (R. W. White and Roth, 2009, p.44) In the next section, this is further explored.

5.1.4 Interface Features and Interaction Styles

Many participants state they believe categorical division of the data helps them to quickly get insight of contents of a set of items. While Hansken has categorical views and facets in some of those views, a categorical division of items does not exist for search results in the search view. Several participants remarked they often have a very long list of results without oversight over the data.

A solution can be found in the usage of facets, offering both organization of results as query expansion and refinement (Kules et al., 2009, p.313). When testing facets in the prototype, participants were positive about both the categorization and interaction.

Interactions with other verticals was more cumbersome. Many participants remarked that interactions with the timeline - through brushing and linking - were unintuitive, and most had trouble interacting with it for the first time in the think aloud study. The existence of a timeline was however desirable, both as a filter and as a visualisation. Perhaps using more familiar interaction styles could mitigate this issue (Roy et al., 2022, p.2765).

The overall concept of the explore view was met with mixed reactions. The purpose of this interface is to provide a blended aggregated search interface with tightly coupled dynamic queries in coordinated views. In theory, this would allow rapid, safe exploration, reducing screen clutter and provide high-level overview (Ahlberg and Shneiderman, 1994, p.315; Marchionini, 2006, p.44).

In contrast, respondents of the survey were quite split on whether they would like to use the mouse in formulating a query. This might be due to respondents having a different idea of what 'clicking together a query' entails.

Some participants stated they would rather have categories separated, saying it looks cluttered or having the combination of representations is not useful for them. Others thought the blended view

offered an overview and it might be useful at the start of an investigation, if you are unfamiliar with the contents of the data or to see trends or hotspots. This might imply a use case specifically for exploratory search, where participants can explore contents of various sources at the same time (Bron et al., 2013). For more in-depth research, other methods of searching might be more applicable.

Concerns arose regarding scalability of the explore interface when a lot of (more diverse) data is present in the dataset. These concerns are valid, and are also voiced in literature for the context of DFTs (Garfinkel, 2010, p.S71). Perhaps that more use of data visualisations instead of textual information can help resolve these concerns (Altiero, 2015, p.28).

Respondents from the survey generally indicate they do not feel like they put in a lot of mental effort when searching in Hansken. For the prototype interface, participants felt similarly. Some remarked having experienced frustrations and insecurity while using the prototype.

5.1.5 Natural Language and AI-agents

Most participants are very positive about their experience with AI-agents and about communicating through natural language, but their trust levels are low. It is understandable that the interaction method is well-received, seeing it is a natural, flexible and concise way of communicating (Hearst, 2011; Hall et al., 1996, pp.2–3).

Participants were generally very aware of limitations of the AI-agents behind that mode of interaction, and were specifically worried whether the system would actually understand their intent. Some expressed doubt about whether the results returned after a natural language request would be complete. This might be related to explainability and to human control being delegated to the AI, making it unclear what information is included (R. W. White, 2024, p.60; Hadi et al., 2024). Users wanted to remain closely involved in shaping and refining their searches, and natural language interactions were seen as potentially taking away some of that agency.

For investigative work, this lack of transparency was perceived as a significant risk. Feedback given by the system, showing how results are derived or allowing users to verify and trace back outputs, would help in regaining the trust. Investigators ought to be hesitant to use information if it is unclear how the system came to its conclusion (Scanlon et al., 2023, pp. 9–10)

5.2 Limitations

This study comes with a few limitations. For the main study and the survey, only participants with prior experience with Hansken were recruited initially. Because of this specific criterium, the pool of possible participants is rather small and there was some trouble being able to recruit suited participants for the study. To have more participants for the main study, two persons without prior experience with Hansken and some with very limited experience were asked to participate as well. Similarly, some respondents of the survey also had little or no experience with Hansken. The result of this is that these individuals could not share their sentiment towards Hansken and the search capabilities it has. They could however participate in the think aloud exercise, and were still able to share insights about other DFTs and their information needs.

One more limitation regarding participation is the lack of respondents to the survey. While the goal of the survey is to generalize results from the main study, the number of valid responses was only slightly higher than the number of participants in the main study. The purpose of the survey is to have a much broader reach. It can be questioned if the survey can be used to generalize findings, taking into account some of the participants of the main study also were respondents in the survey. One of the reasons for the lack of responses can be found in wrong expectations of the number of available people at a visited location. The survey was adapted with an online version at a late moment in the research period and should have been distributed over electronic channels.

Another limitation can be identified in the ideation phase for the design thinking process in the prototype. While many novel ideas were generated during the first iteration of the prototype, not enough novel ideation took place in the later prototypes, testing generally the same ideas multiple times during all iterations without too much change. This can partially be attributed to the prototype being high fidelity starting at the second iteration, leading to alterations taking too much effort to implement in the available time before user testing. The result is less novel innovation, meaning that possibly better solutions were never found.

5.3 Future work

In future research, it might be valuable to perform a comparative study between Hansken and other digital forensic tools to understand strengths and weaknesses. Users state they often use many different DFTs, even during a single case, because of specific functionality present in that tool. If it is known in what ways these tools distinguish, development of Hansken could be directed towards these areas.

It might be interesting to also understand search in Hansken from even more diverse organizations. While five organisations were visited, it is known from unpublished previous research that some other organisation(s) using Hansken have completely different investigative processes. Its possible their needs require completely different insights into the data and thus different search functionalities in Hansken. Inclusion of intelligence and foreign users could provide completely new insights.

Looking at other research methods, performing a diary study with users of Hansken could uncover more rich data about user interactions with Hansken. The fact that a diary study can be conducted at a place of work while users are working with real data allows researchers to understand practical user behaviour and how it changes over time. Finding willing participants might however pose a challenge, and organizations might decline this option due to security concerns.

Other than this, it could be valuable to test different kinds of visualisations - also outside the search domain. Using visualisations in DFTs is a recommendation in literature, and visual interfaces can help users understand data and relations within that data (Beebe, 2009, p.25).

While the interaction style of natural language has existed for a long time, it has only become popular in practice quite recently. While technological advancements in AI can form interesting possibilities in the future, it might also be interesting to see how habituation and possible broader adoption of the natural interaction style in the future affects feeling towards it.

Chapter 6

Conclusion

In this research, the search process in Hansken was examined and an alternative search interface was developed and tested with Hansken users. The information need of users can vary wildly depending on organisation, role and work task. The information need can be confirmatory - searching for evidence to support or disprove a claim - or exploratory - where new hypotheses are developed by analysis of the data.

Search behaviour of Hansken users is very dependent of the information need. Prior information, experience and expertise are applied for electing terms used in keyword queries.

Users exhibit exploratory search behaviour when their information need is defined broadly or they are uncertain of where to find an answer. During periods of exploratory browsing, found traces can act as a source of information to inform new search directions, similar to the model of berrypicking. Once a potentially interesting item has been found, focused searching is employed to examine the context of this trace to understand its potential value. This context can be both about the trace itself (metadata) or relational (traces in proximity). Information about a trace can subsequently be used to shape a new search direction. In this way, the search is continuously evolving.

Depending on the information need, users might also undertake systematic searches - where all results of a search are examined exhaustively - or lookup activities - for simple and straight forward question answering.

When it comes to interface features, users desire categorization of data. Implementation of faceted search is appreciated due to its ability to provide an overview of the contents of the data and its easy and intuitive interaction style. Data visualisations, like a timeline or a map, can show large amounts of data. If visualisations act as dynamic queries, interactions should be familiar and intuitive.

Displaying traces in a blended aggregated search interface offers an overview, but is generally most effective at the start of a search journey. It could be a valuable addition for facilitating exploratory search, but non-aggregated search interfaces are more effective for focused searching.

Information seeking through natural language interaction is generally seen as an easy and lenient method of query formulation, but there is massive distrust in the correctness of the response of AI-agents. Some might trust the agent after personal testing and if it provides explanations of its

responses, but most say they would always stay cautious of errors. Users do not seem to need help with task decomposition, usually doing this alone without much trouble or together with colleagues.

Some design implications can be derived from these findings. There should be support for a diverse range of information needs: for exploration, but also for more focused search or lookup tasks. Showing and navigating through context of a trace, meaning both the metadata of one trace as its relations to other traces, helps users explore and estimate relevance of traces. Implementing more data categorization through faceted search to provide overview of data and reduce effort in filtering. Having an option for aggregated search to support exploration and orientation. Implement options for natural language interactions for its usability, but garner trust in AI-agents by providing transparent explanations of actions.

Bibliography

Ahlberg, C., & Shneiderman, B. (1994). Visual information seeking: Tight coupling of dynamic query filters with starfield displays. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 313–317.

Altiero, R. A. (2015). *Digital forensics tool interface visualization* [Doctoral dissertation]. Nova Southeastern University [Retrieved from NSUWorks, Graduate School of Computer and Information Sciences (24)]. https://nsuworks.nova.edu/gscis_etd/24

Andersen, S. (2019). Technical report: A preliminary process model for investigation. https://doi.org/10.31235/osf.io/z4wma

Azad, H. K., & Deepak, A. (2019). Query expansion techniques for information retrieval: A survey. *Information Processing & Management*, *56*(5), 1698–1735. https://doi.org/https://doi.org/10.1016/j.ipm.2019.05.009

Baber, C., & Alotaibi, A. (2024). Using chatgpt to support criminal investigations: A comparative study of ai and human query. https://doi.org/10.54941/ahfe1004661

Baeza-Yates, R. A., & Ribeiro-Neto, B. (1999). *Modern information retrieval*. Addison-Wesley Longman Publishing Co., Inc.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online review*, *13*(5), 407–424.

Bates, M. J. (2005). Berrypicking. In K. E. Fisher, S. Erdelez, & L. McKechnie (Eds.), *Theories of information behavior* (pp. 58–62). Information Today.

Beebe, N. (2009). Digital forensic research: The good, the bad and the unaddressed. In G. Peterson & S. Shenoi (Eds.), *Advances in digital forensics v* (pp. 17–36). Springer Berlin Heidelberg.

Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian journal of information science*, *5*(1), 133–143.

Blandford, A., Furniss, D., & Makri, S. (2016). *Qualitative hci research: Going behind the scenes*. Morgan & Claypool Publishers.

Bron, M., van Gorp, J., Nack, F., Baltussen, L. B., & de Rijke, M. (2013). Aggregated search interface preferences in multi-session search tasks. *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 123–132. https://doi.org/10.1145/2484028.2484050

Brown, T., et al. (2008). Design thinking. *Harvard business review*, *86*(6), 84.

Bruza, P., & Dennis, S. (1999). Query reformulation on the internet: Empirical data and the hyperindex search engine. *RIAO*, *97*.

Budiu, R. (2020). *Information scent: How users decide where to go next* [Accessed 30-March-2025]. https://www.nngroup.com/articles/information-scent/

Capra, R., Marchionini, G., Oh, J. S., Stutzman, F., & Zhang, Y. (2007). Effects of structure and interaction style on distinct search tasks. *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries*, 442–451. https://doi.org/10.1145/1255175.1255267

Chen, L., & Chua, C. (2013). Interactive interface for query formulation. *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, 507–510.

Chilton, L. B., & Teevan, J. (2011). Addressing people's information needs directly in a web search result page. *Proceedings of the 20th International Conference on World Wide Web*, 27–36. https://doi.org/10.1145/1963405.1963413

Cooper, W. (1971). A definition of relevance for information retrieval. *Information Storage and Retrieval*, *7*(1), 19–37. https://doi.org/https://doi.org/10.1016/0020-0271(71)90024-6

Daubner, L., Buhnova, B., & Pitner, T. (2024). Forensic experts' view of forensic-ready software systems: A qualitative study. *Journal of Software: Evolution and Process*, *36*(5), e2598. https://doi.org/https://doi.org/10.1002/smr.2598

Dörk, M., Williamson, C., & Carpendale, S. (2009). Towards visual web search: Interactive query formulation and search result visualization. *WSSP. Madrid, Spain*.

Du, X. (2020). *Alleviating the digital forensic backlog: A methodology for automated digital evidence processing* [Doctoral dissertation, University College Dublin].

Feild, H., White, R. W., & Fu, X. (2013). Supporting orientation during search result examination. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2999–3008. https://doi.org/10.1145/2470654.2481416

Fishkin, K., & Stone, M. C. (1995). Enhanced dynamic queries via movable filters. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 415–420. https://doi.org/10.1145/223904.223960

Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1987). The vocabulary problem in human-system communication. *Communications of the ACM*, *30*(11), 964–971.

Gao, J., Xiong, C., Bennett, P., & Craswell, N. (2023). *Neural approaches to conversational information retrieval* (Vol. 44). Springer.

Garfinkel, S. L. (2010). Digital forensics research: The next 10 years [The Proceedings of the Tenth Annual DFRWS Conference]. *Digital Investigation*, *7*, S64–S73. https://doi.org/https://doi.org/10.1016/j.diin.2010.05.009

Gatterbauer, W., Dunne, C., Jagadish, H., & Riedewald, M. (2022). Principles of query visualization. *arXiv preprint arXiv:2208.01613*.

Gritz, W., Otto, C., Hoppe, A., Pardi, G., Kammerer, Y., & Ewerth, R. (2023). Comparing interface layouts for the presentation of multimodal search results. *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*, 321–327. https://doi.org/10.1145/3576840.3578335

Gwizdka, J. (2010). Distribution of cognitive load in web search. *Journal of the American Society for Information Science and Technology*, *61*(11), 2167–2187. https://doi.org/https://doi.org/10.1002/asi.21385

Hadi, M. U., Al Tashi, Q., Shah, A., Qureshi, R., Muneer, A., Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Wu, J., et al. (2024). Large language models: A comprehensive survey of its applications, challenges, limitations, and future prospects. *Authorea Preprints*.

Hall, G., Popowich, F., & Fass, D. (1996). Natural language edit controls: Constrained natural language devices in user interfaces. *Proceedings Eighth IEEE International Conference on Tools with Artificial Intelligence*, 475–477.

Hearst, M. A. (2006). Clustering versus faceted categories for information exploration. *Commun. ACM*, *49*(4), 59–61. https://doi.org/10.1145/1121949.1121983

Hearst, M. A. (2011). 'natural' search user interfaces. *Commun. ACM*, *54*(11), 60–67. https://doi.org/10.1145/2018396.2018414

Henseler, H., & van Beek, H. (2023). Chatgpt as a copilot for investigating digital evidence. *LegalAIIA*, 58–69.

Jones, S. (1998). Graphical query specification and dynamic result previews for a digital library. *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology*, 143–151. https://doi.org/10.1145/288392.288595

Kohn, M., Eloff, M., & Eloff, J. (2013). Integrated digital forensic process model [Cybercrime in the Digital Economy]. *Computers & Security*, *38*, 103–115. https://doi.org/https://doi.org/10.1016/j.cose.2013.05.001

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science*, *42*(5), 361–371. https://doi.org/https://doi.org/10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-\#

Kules, B., Capra, R., Banta, M., & Sierra, T. (2009). What do exploratory searchers look at in a faceted search interface? *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries*, 313–322. https://doi.org/10.1145/1555400.1555452

Lau, T., & Horvitz, E. (1999). Patterns of search: Analyzing and modeling web query refinement. *UM99 User Modeling: Proceedings of the Seventh International Conference*, 119–128.

Lee, J., & Hong, D. (2011). Pervasive forensic analysis based on mobile cloud computing. *2011 Third International Conference on Multimedia Information Networking and Security*, 572–576. https://doi.org/10.1109/MINES.2011.77

Li, Y., & Belkin, N. J. (2008). A faceted approach to conceptualizing tasks in information seeking [Adaptive Information Retrieval]. *Information Processing & Management*, *44*(6), 1822–1837. https://doi.org/https://doi.org/10.1016/j.ipm.2008.07.005

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.

Liu, C., Liu, Y.-H., Liu, J., & Bierig, R. (2021). Search interface design and evaluation. *Foundations and Trends® in Information Retrieval*, *15*(3-4), 243–416. https://doi.org/10.1561/1500000073

Mahdi, M. N., Ahmad, A. R., Ismail, R., Natiq, H., & Mohammed, M. A. (2020). Solution for information overload using faceted search–a review. *IEEE Access*, *8*, 119554–119585. https://doi.org/10.1109/ACCESS.2020.3005536

Manaris, B. (1998). Natural language processing: A human-computer interaction perspective. In M. V. Zelkowitz (Ed.). Elsevier. https://doi.org/https://doi.org/10.1016/S0065-2458(08)60665-8

Marchionini, G. (1995). *Information seeking in electronic environments*. Cambridge university press.

Marchionini, G. (2006). Exploratory search: From finding to understanding. *Commun. ACM*, *49*(4), 41–46. https://doi.org/10.1145/1121949.1121979

Moffat, A., & Zobel, J. (2008). Rank-biased precision for measurement of retrieval effectiveness. *ACM Transactions on Information Systems (TOIS)*, *27*(1), 1–27.

Moran, K., & Goray, C. (2019). The anatomy of a search-results page [Accessed 27-March-2025]. https://www.nngroup.com/articles/anatomy-search-results-page/

Moran, K., & Goray, C. (2020). *Three key serp features: Featured snippets, people also ask, and knowledge panels* [Accessed 27-March-2025]. https://www.nngroup.com/articles/key-serp-features/

Nielsen, J. (1994). *10 usability heuristics for user interface design* [Accessed 30-March-2025]. https://www.nngroup.com/articles/ten-usability-heuristics/

Oliveira, B., & Teixeira Lopes, C. (2023). The evolution of web search user interfaces - an archaeological analysis of google search engine result pages. *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*, 55–68. https://doi.org/10.1145/3576840.3578320

Palta, S., Chandrasekaran, N., Rudinger, R., & Counts, S. (2025). Speaking the right language: The impact of expertise alignment in user-ai interactions. *arXiv preprint arXiv:2502.18685*. https://doi.org/10.48550/arXiv.2502.18685

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological review*, *106*(4), 643.

Rijksoverheid. (2020). *Vakbijlage hansken* [Accessed 06-Feburary-2025] (Versie 1.2). https://www.forensischinstituut.nl/publicaties/publicaties/2020/12/02/vakbijlage-hansken

Roy, N., Maxwell, D., & Hauff, C. (2022). Users and contemporary serps: A (re-)investigation. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2765–2775. https://doi.org/10.1145/3477495.3531719

Russell-Rose, T., Chamberlain, J., & Shokraneh, F. (2019). A visual approach to query formulation for systematic search. *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*, 379–383. https://doi.org/10.1145/3295750.3298919

Russell-Rose, T., & Shokraneh, F. (2020). Designing the structured search experience: Rethinking the query-builder paradigm. *Weave: Journal of Library User Experience*, *3*(1). https://doi.org/http://dx.doi.org/10.3998/weave.12535642.0003.102

Sarkar, A. (2024). Ai should challenge, not obey. *Commun. ACM*, *67*(10), 18–21. https://doi.org/10.1145/3649404

Savolainen, R. (2018). Berrypicking and information foraging: Comparison of two theoretical frameworks for studying exploratory search. *Journal of Information Science*, *44*(5), 580–593. https://doi.org/10.1177/0165551517713168

Scanlon, M., Breitinger, F., Hargreaves, C., Hilgert, J.-N., & Sheppard, J. (2023). Chatgpt for digital forensic investigation: The good, the bad, and the unknown. *Forensic Science International: Digital Investigation*, *46*, 301609. https://doi.org/https://doi.org/10.1016/j.fsidi.2023.301609

Shneiderman, B. (1994). Dynamic queries for visual information seeking. *IEEE software*, *11*(6), 70–77. https://doi.org/10.1109/52.329404

Shneiderman, B. (2003). The eyes have it: A task by data type taxonomy for information visualizations. In B. B. BEDERSON & B. SHNEIDERMAN (Eds.), *The craft of information visualization* (pp. 364–371). Morgan Kaufmann. https://doi.org/https://doi.org/10.1016/B978-155860915-0/50046-9

Shneiderman, B., & Plaisant, C. (2005). *Designing the user interface: Strategies for effective human-computer interaction* (4th ed.). Pearson Education.

Spoerri, A. (2004). How visual query tools can support users searching the internet. *Proceedings. Eighth International Conference on Information Visualisation, 2004. IV 2004.*, 329–334. https://doi.org/10.1109/IV.2004.1320165

Stanford d.school. (2010). *An introduction to design thinking process guide* [Accessed: 6 August 2025]. https://web.stanford.edu/~mshanks/MichaelShanks/files/509554.pdf

Sunde, N. (2022, November). *Constructing digital evidence - a study on how cognitive and human factors affect digital evidence* [Doctoral dissertation].

Taylor, R. S. (1968). Question-negotiation and information seeking in libraries. *College & research libraries*, *29*(3), 178–194.

Teevan, J., Adar, E., Jones, R., & Potts, M. A. (2007). Information re-retrieval: Repeat queries in yahoo's logs. *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, 151–158.

Teevan, J., Alvarado, C., Ackerman, M. S., & Karger, D. R. (2004). The perfect search engine is not enough: A study of orienteering behavior in directed search. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 415–422. https://doi.org/10.1145/985692.985745

Ukwen, D. O., & Karabatak, M. (2021). Review of nlp-based systems in digital forensics and cybersecurity. *2021 9th International symposium on digital forensics and security (ISDFS)*, 1–9.

van Beek, H., & Henseler, H. (2023). Servicing digital investigations with artificial intelligence. *Artificial Intelligence (AI) in Forensic Sciences*, 103.

van Baar, R., van Beek, H., & van Eijk, E. (2014). Digital forensics as a service: A game changer [Proceedings of the First Annual DFRWS Europe]. *Digital Investigation*, *11*, S54–S62. https://doi.org/10.1016/j.diin.2014.03.007

van Beek, H., van Eijk, E., van Baar, R., Ugen, M., Bodde, J., & Siemelink, A. (2015). Digital forensics as a service: Game on [Special Issue: Big Data and Intelligent Data Analysis]. *Digital Investigation*, *15*, 20–38. https://doi.org/10.1016/j.diin.2015.07.004

Wang Baldonado, M. Q., Woodruff, A., & Kuchinsky, A. (2000). Guidelines for using multiple views in information visualization. *Proceedings of the Working Conference on Advanced Visual Interfaces*, 110–119. https://doi.org/10.1145/345513.345271

White, R., Marchionini, G., & Muresan, G. (2008). Evaluating exploratory search systems: Introduction to special topic issue of information processing and management. *Inf. Process. Manage.*, *44*, 433–436.

White, R. W. (2016). *Interactions with search systems*. Cambridge University Press, New York. https://doi.org/10.1017/CBO9781139525305

White, R. W. (2023). Tasks, copilots, and the future of search. *Proceedings of the 46th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2023)*, 5–6. http://ryenwhite.com/talks/pptx/WhiteSIGIR2023.pptx

White, R. W. (2024). Advancing the search frontier with ai agents. *Communications of the ACM*, *67*(9), 54–65.

White, R. W., Kules, B., Drucker, S. M., et al. (2006). Supporting exploratory search, introduction, special issue, communications of the acm. *Communications of the ACM*, *49*(4), 36–39.

White, R. W., & Roth, R. A. (2009). *Exploratory search: Beyond the query-response paradigm*. Morgan & Claypool Publishers. https://doi.org/10.2200/S00174ED1V01Y200901ICR003

White, R. W., & Ruthven, I. (2006). A study of interface support mechanisms for interactive information retrieval. *Journal of the American Society for Information Science and Technology*, *57*(7), 933–948.

Wickramasekara, A., Breitinger, F., & Scanlon, M. (2024). Exploring the potential of large language models for improving digital forensic investigation efficiency. *arXiv preprint arXiv:2402.19366*.

Wilson, M. L., & schraefel m.c., m. (2008). A longitudinal study of exploratory and keyword search. *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries*, 52–56. https://doi.org/10.1145/1378889.1378899

Yi, J. S., Melton, R., Stasko, J., & Jacko, J. A. (2005). Dust & magnet: Multivariate information visualization using a magnet metaphor. *Information Visualization*, *4*(4), 239–256. https://doi.org/10.1057/palgrave.ivs.9500099

Yusoff, Y., Ismail, R., & Hassan, Z. (2011). Common phases of computer forensics investigation models. *International Journal of Computer Science & Information Technology (IJCSIT)*, *3*. https://doi.org/10.5121/ijcsit.2011.3302

Zareen, M. S., Aslam, B., Tahir, S., Rasheed, I., & Khan, F. (2024). Unveiling the dynamic landscape of digital forensics: The endless pursuit. *Computers*, *13*(12). https://doi.org/10.3390/computers13120333

Zawoad, S., & Hasan, R. (2015). Digital forensics in the age of big data: Challenges, approaches, and opportunities. *2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems*, 1320–1325. https://doi.org/10.1109/HPCC-CSS-ICESS.2015.305

Zhou, K., Cummins, R., Lalmas, M., & Jose, J. M. (2012). Evaluating aggregated search pages. *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 115–124. https://doi.org/10.1145/2348283.2348302

Appendix A

Survey Information Sheet

Universiteit Utrecht

Research participant questionnaire information sheet

Combing Through Crimes: Understanding and Enhancing Search in Hansken

- **Introduction**: You are invited to participate in the research project *"Combing Through Crimes: Understanding and Enhancing Search in Hansken"*. If you have any questions do not hesitate to ask the researcher. This information sheet only applies to the attached questionnaire.
- **Background**: The research project aims to understand the information need of Hansken users and users of other digital forensic tools. Specifically, information seeking behavior of users and interface elements that might enable desired behavior are studied, with the aim to provide possible insights for future Hansken interface improvements.
- **Researchers**: This study is carried out by Ko Schoemaker (**k.schoemaker@students.uu.nl**) as part of his master thesis under supervision of Eelco Herder (**e.herder@uu.nl**) at Utrecht University and Job van den Hoonaard (**j.van.den.hoonaard@nfi.nl**) at NFI.
- **Why**: The aim of this questionnaire is to give additional quantitative insights into search behavior in Hansken and inform possible improvements in its operation. This questionnaire is a stand-alone method of generating insights and does not require you to participate in other parts of the user study. The function of the questionnaire is to make it possible to receive responses from a broader audience than is logistically possible to conduct the full user study with.
- **The questionnaire**: In this questionnaire you will be asked to answer multiple-choice questions over three different parts. It is possible to elaborate answers or write remarks in the designated boxes. If you wish to not answer a question, it is possible to leave it blank. If you have any questions regarding part of the questionnaire, feel free to ask the researcher. The questions pertain to your experiences and opinions of Hansken or digital forensic tools, as a professional in the forensic field. The questionnaire is in Dutch. Filling in this questionnaire will take about ten minutes. You will receive no compensation for participating.
- **Data processing**: Multiple choice answers and written remarks are digitized for further processing and quantitative analysis. No personal data will be collected and your answers will be stored anonymously.
- **Risks and Benefits**: This study does not carry any foreseeable risks. Potential benefits include helping improve Hansken and the Hansken interface by sharing your professional opinions, which might directly positively impact you or your colleagues' future experience of the product.
- **Your rights**: Participation is voluntary. We are only allowed to collect your data for this study if you consent to this. If you decide not to participate, you do not have to take any further action. You do not need to sign anything, nor are you required to explain why you do not want to participate. If you decide to participate, you can always change your mind and stop participating at any time, including during the study. As this questionnaire does not carry any personally identifiable information, it is not possible to erase your recorded responses after the questionnaire has been handed in with the researcher.

If there are any further questions now or in the future, you can contact the researcher at ***k.schoemaker@students.uu.nl*** *or* ***k.schoemaker@nfi.nl****. If you would like contact after August 2025, please reach out to* ***koschoemaker@gmail.com****.*

If you have any complaints or questions about the processing of personal data, please send an email to the Faculty of Sciences Privacy Officer: ***privacy-beta@uu.nl****. The Privacy Officer will also be able to assist you in exercising the rights you have under the GDPR. For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information at* ***www.uu.nl/en/organisation/privacy****.*

Appendix B

Survey

Deel 1: Demografie

1. Hoe vaak gebruikt u Hansken?

O Dagelijks O Wekelijks O Maandelijks O Zelden O Nooit

2. Hoe vaak gebruikt u Hansken via de tactische gebruikersinterface (HUIB)?

O Dagelijks O Wekelijks O Maandelijks O Zelden O Nooit

3. Gebruikt u Hansken wel eens via een andere interface? (Meerdere antwoorden mogelijk)

O nee of bijna nooit O via ExpertUI O met Python O via de REST interface

O anders, nl:

4. Welke rol vervult u binnen uw organisatie? (Meerdere antwoorden mogelijk)

O Tactisch rechercheur

O Digitaal expert

O Analist

O Operator, ontwikkelaar of beheerder

O Anders, namelijk:

5. Hoe zou u uw eigen digitale vaardigheden inschatten?

O Helemaal niet vaardig O Beperkt vaardig O Redelijk vaardig O Vaardig O Zeer vaardig

Eventuele opmerkingen (ook voor opmerkingen over volgende pagina's te gebruiken):

	Deel 2: Statements	Schaal					
	Geef bij onderstaande statements aan in hoeverre u het ermee eens bent. Vul N.V.T. in als u het gevoel heeft dat u de vraag niet kunt beantwoorden. Bij het opmerkingenveld onderaan de pagina kunt u eventueel een opmerkingen achterlaten mocht u iets willen toelichten.	ZEER ONEENS	ONEENS	NEUTRAAL	EENS	ZEER EENS	N.V.T.
6	Ik weet meestal goed welke informatie ik zoek bij een onderzoek	O	O	O	O	O	O
7	Ik kan de informatie die ik zoek meestal snel vinden met Hansken	O	O	O	O	O	O
8	De huidige zoekfuncties in Hansken sluiten goed aan bij mijn informatiebehoefte	O	O	O	O	O	O
9	Ik mis bepaalde zoekmogelijkheden of filters in Hansken om de juiste informatie te vinden	O	O	O	O	O	O
10	Ik gebruik vaak alternatieve methoden (zoals andere software of vragen aan collega's) om informatie te vinden die ik niet zelf in Hansken vind	O	O	O	O	O	O
11	Ik begin meestal met een brede zoekopdracht en verfijn mijn zoekopdracht daarna stapsgewijs	O	O	O	O	O	O
12	Ik pas mijn zoekopdracht aan op basis van wat ik onderweg tegenkom	O	O	O	O	O	O
13	Ik gebruik vaak meerdere zoekpogingen om tot het gewenste resultaat te komen	O	O	O	O	O	O
14	Ik maak vaak een complexe zoekopdracht (zoeken met meerdere filters tegelijk of met AND of OR statements in HQL)	O	O	O	O	O	O
15	Ik weet goed hoe ik moet navigeren tussen verschillende categorie-overzichten, zoekresultaten en sporen in Hansken	O	O	O	O	O	O
16	Ik voel me beperkt in mijn manier van zoeken binnen de Hansken-interface	O	O	O	O	O	O
17	Ik ervaar de Hansken Interface als gebruiksvriendelijk	O	O	O	O	O	O
18	Ik formuleer mijn zoekopdracht weleens met de Hansken Query Language (HQL)	O	O	O	O	O	O
19	Het kost me weinig moeite om te begrijpen hoe ik moet zoeken of resultaten kan verfijnen in Hansken	O	O	O	O	O	O
20	Ik voel me mentaal belast tijdens het uitvoeren van zoekopdrachten in Hansken	O	O	O	O	O	O
21	De interface ondersteunt mij goed bij het vinden van relevante informatie	O	O	O	O	O	O
22	Ik wil graag mijn zoekopdracht met de muis aan elkaar kunnen klikken	O	O	O	O	O	O
23	Ik kan met Hansken efficiënt en snel een zoekopdracht uitvoeren	O	O	O	O	O	O
24	Ik zou het handig vinden om in natuurlijke taal (bijvoorbeeld gewone geschreven zinnen) met Hansken een zoekopdracht te formuleren	O	O	O	O	O	O
25	Ik zou suggesties of antwoorden van een AI-assistent vertrouwen, als deze assistent in Hansken ingebouwd is	O	O	O	O	O	O
26	Ik heb vaak te maken met complexe doelen of hypotheses, die ik zelf moet opdelen in kleinere uitvoerbare zoekopdrachten	O	O	O	O	O	O
27	Ik kan succesvol zoekopdrachten formuleren in Hansken	O	O	O	O	O	O
28	Ik vind het makkelijk om een zoekopdracht te formuleren in Hansken	O	O	O	O	O	O
29	Ik heb behoefte aan meer ondersteuning bij het formuleren van goede zoekopdrachten	O	O	O	O	O	O

Eventuele opmerkingen:

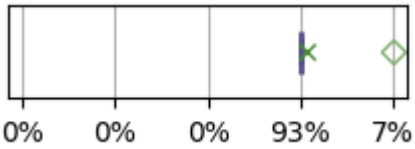
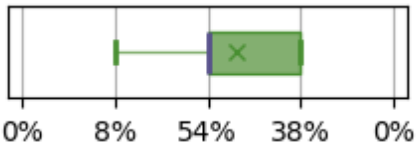
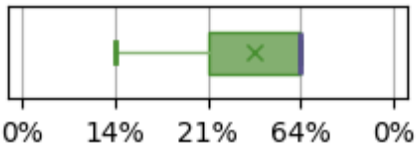
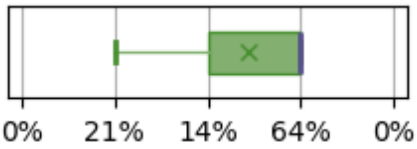
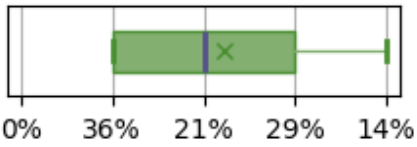
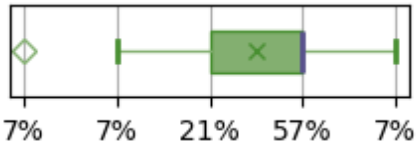
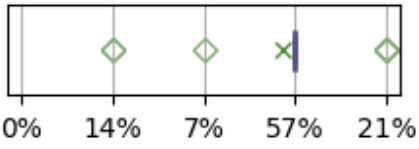
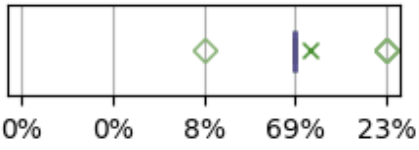
	Deel 3: IPA	In hoeverre is dit momenteel mogelijk in Hansken?					In hoeverre is dit voor u belangrijk?				
	Geef bij onderstaande statements aan in hoeverre u gelooft dat een functie ondersteund wordt door Hansken en de Hansken interface, en in hoeverre u deze functie belangrijk vindt. Bij het opmerkingenveld onderaan de pagina kunt u eventueel een opmerkingen achterlaten mocht u iets willen toelichten.	HELEMAAL NIET	NAUWELIJKS	GEDEELTELIJK	GOED	VOLLEDIG	ZEER ONBELANGRIJK	ONBELANGRIJK	NEUTRAAL	BELANGIJK	ZEER BELANGRIJK
30	Een snel overzicht kunnen krijgen van relevante data of zoekresultaten	O	O	O	O	O	O	O	O	O	O
31	Een nuttig visueel overzicht (bijvoorbeeld in dashboards of grafieken) krijgen van mijn zoekresultaten	O	O	O	O	O	O	O	O	O	O
32	Een duidelijk overzicht krijgen van alle data die aanwezig is in het project	O	O	O	O	O	O	O	O	O	O
33	Zoekresultaten kunnen filteren op eigenschappen zoals tijd, type of bron	O	O	O	O	O	O	O	O	O	O
34	Meerdere filters tegelijk kunnen toepassen om een opdracht te verfijnen	O	O	O	O	O	O	O	O	O	O
35	Duidelijk kunnen zien welke filters actief zijn	O	O	O	O	O	O	O	O	O	O
36	Kunnen inzoomen op een specifieke set sporen	O	O	O	O	O	O	O	O	O	O
37	Kunnen oriënteren waar in de data u op enig moment bent bij het vinden van een spoor of een set sporen	O	O	O	O	O	O	O	O	O	O
38	Eenvoudig navigeren tussen detail- en overzichtsniveaus	O	O	O	O	O	O	O	O	O	O
39	Extra details of informatie kunnen opvragen over een gevonden spoor	O	O	O	O	O	O	O	O	O	O
40	Een spoor in zijn context kunnen zien	O	O	O	O	O	O	O	O	O	O
41	Andere relevante of gerelateerde sporen rondom een gevonden spoor kunnen vinden	O	O	O	O	O	O	O	O	O	O
42	Niet overspoeld worden met ongevraagde details	O	O	O	O	O	O	O	O	O	O
43	Zelf controle houden over het zoekproces	O	O	O	O	O	O	O	O	O	O
44	De mogelijkheid hebben om eerdere handelingen, stappen of toepassing van filters ongedaan te maken	O	O	O	O	O	O	O	O	O	O
45	Hulp krijgen vanuit het systeem bij het opdelen van complexe vragen of zoekopdrachten	O	O	O	O	O	O	O	O	O	O
46	De interface helpt mij om mijn denkproces te structureren tijdens het zoeken	O	O	O	O	O	O	O	O	O	O
47	Een zoekopdracht kunnen formuleren met de Hansken Query Language (HQL)	O	O	O	O	O	O	O	O	O	O
48	Kunnen zien wat de impact van een zoekfilter is op het zoekresultaat	O	O	O	O	O	O	O	O	O	O

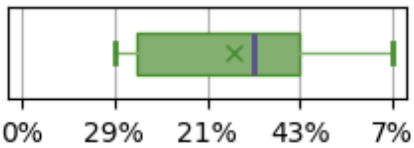
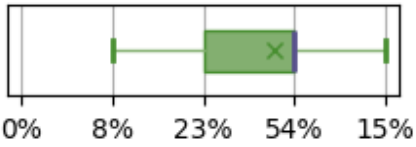
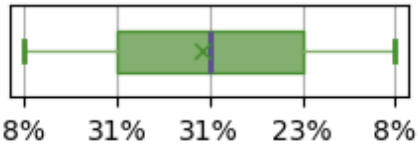
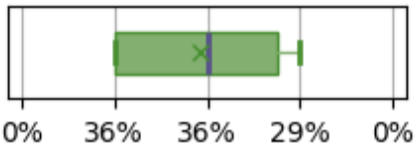
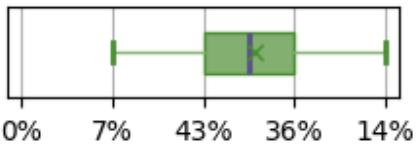
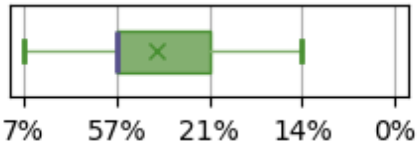
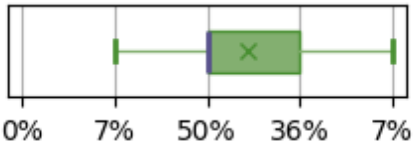
Eventuele opmerkingen:

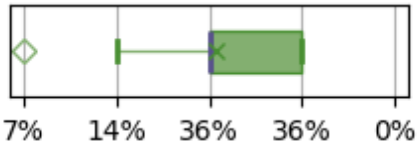
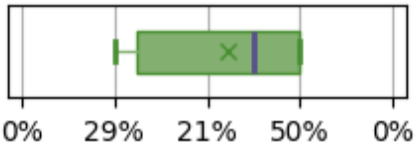
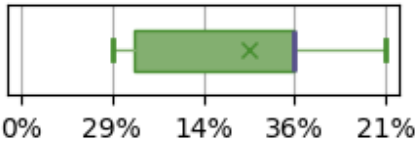
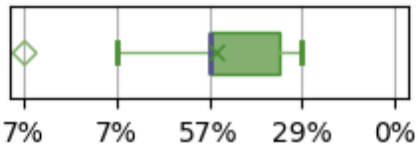
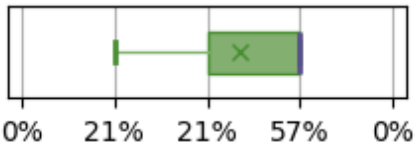
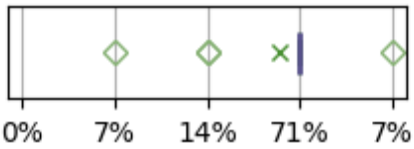
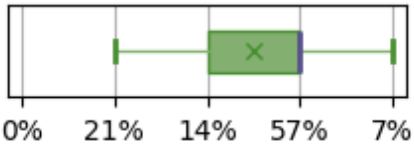
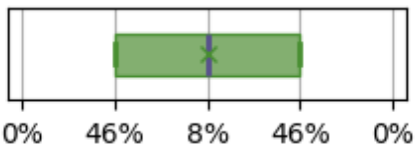
Appendix C

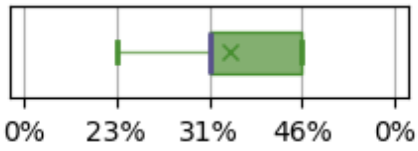
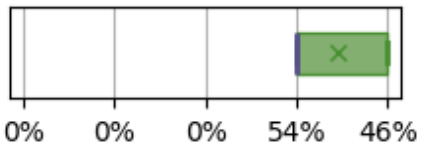
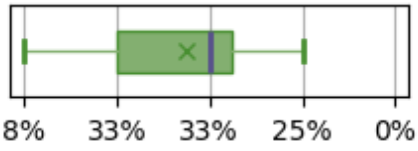
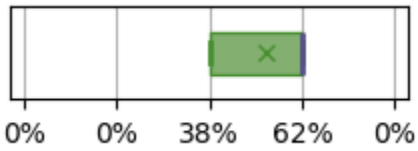
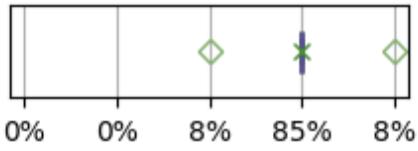
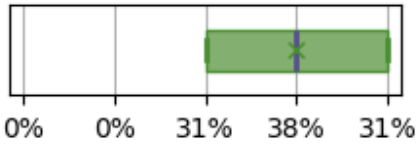
Survey Results

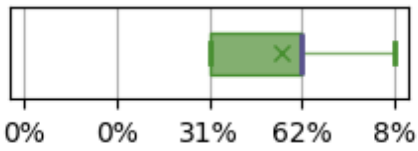
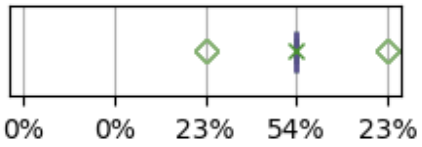
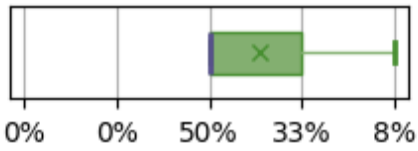
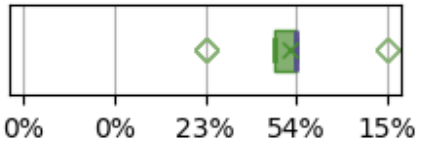
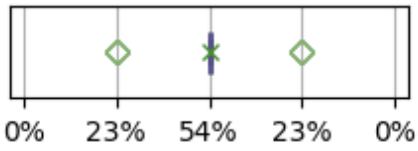
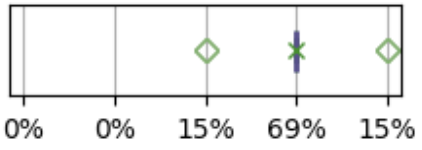
#	Role	Frequency	Digital skills	Frequency HTI	Other Interface
1	Tactical Investigator	Daily	Skilled	Daily	No or almost never
2	Tactical Investigator and Operator	Daily	Fairly Skilled	Daily	No or almost never
3	Digital Expert	Weekly	Very Skilled	Weekly	Via ExpertUI
4	Analyst	Rarely	Fairly Skilled	Rarely	No or almost never
5	Analyst	Monthly	Fairly Skilled	Monthly	Via ExpertUI
6	Tactical Investigator	Weekly	Fairly Skilled	Weekly	Via ExpertUI
7	Digital Expert	Weekly	Skilled	Weekly	Via ExpertUI
8	Digital Expert	Daily	Very Skilled	Daily	Via ExpertUI
9	Digital Expert	Daily	Fairly Skilled	Weekly	Via ExpertUI
10	Analyst	Daily	Limited Skills	Daily	No or almost never
11	Digital Expert	Weekly	Skilled	Weekly	Via ExpertUI
12	Tactical Investigator and Digital Expert	Weekly	Skilled	Weekly	No or almost never
13	Operator, Developer or Administrator	Rarely	Skilled	Rarely	Via ExpertUI
14	Operator, Developer or Administrator	Daily	Very Skilled	Daily	Via ExpertUI

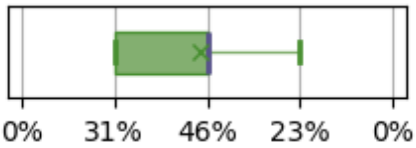
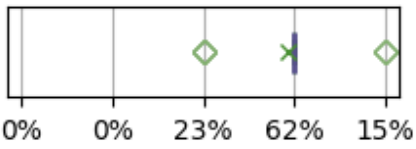
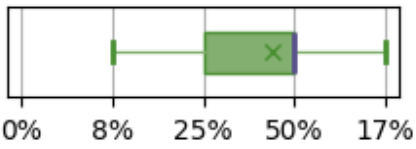
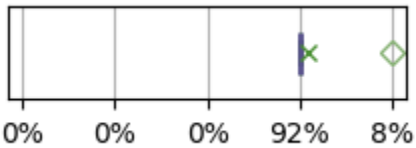
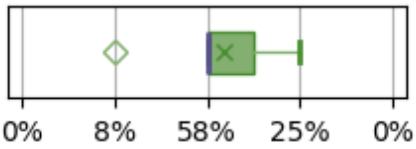
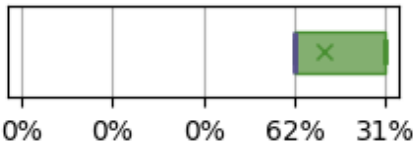
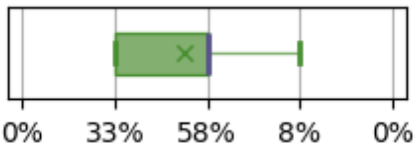
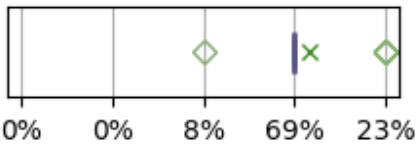
6	Ik weet meestal goed welke informatie ik zoek bij een onderzoek  N= 14 MEAN=4.07 SD=0.27 MEDIAN=4
7	Ik kan de informatie die ik zoek meestal snel vinden met Hansken  N= 13 MEAN=3.31 SD=0.63 MEDIAN=3
8	De huidige zoekfuncties in Hansken sluiten goed aan bij mijn informatiebehoefte  N= 14 MEAN=3.50 SD=0.76 MEDIAN=4
9	Ik mis bepaalde zoekmogelijkheden of filters in Hansken om de juiste informatie te vinden  N= 14 MEAN=3.43 SD=0.85 MEDIAN=4
10	Ik gebruik vaak alternatieve methoden (zoals andere software of vragen aan collega's) om informatie te vinden die ik niet zelf in Hansken vind  N= 14 MEAN=3.21 SD=1.12 MEDIAN=3
11	Ik begin meestal met een brede zoekopdracht en verfijn mijn zoekopdracht daarna stapsgewijs  N= 14 MEAN=3.50 SD=1.02 MEDIAN=4
12	Ik pas mijn zoekopdracht aan op basis van wat ik onderweg tegenkom  N= 14 MEAN=3.86 SD=0.95 MEDIAN=4
13	Ik gebruik vaak meerdere zoekpogingen om tot het gewenste resultaat te komen  N= 13 MEAN=4.15 SD=0.55 MEDIAN=4

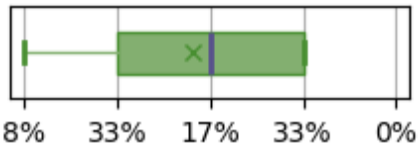
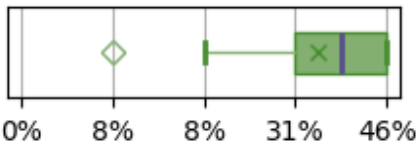
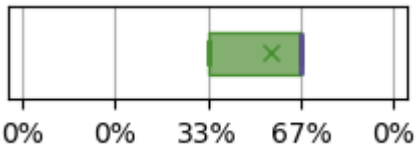
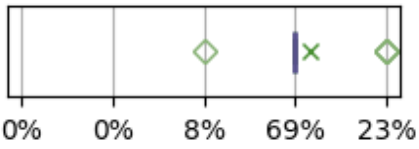
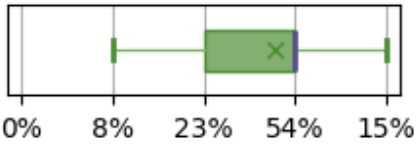
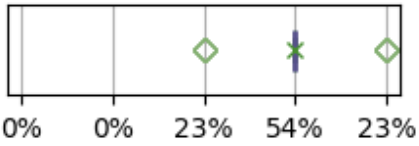
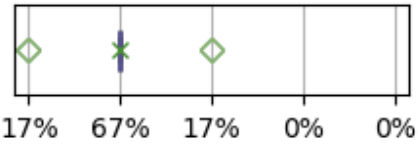
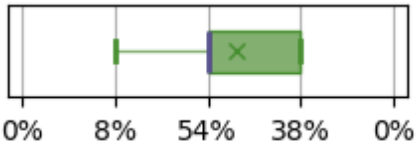
14	Ik maak vaak een complexe zoekopdracht (zoeken met meerdere filters tegelijk of met AND of OR statements in HQL)  N= 14 MEAN=3.29 SD=0.99 MEDIAN=3.5
15	Ik weet goed hoe ik moet navigeren tussen verschillende categorie-overzichten, zoekresultaten en sporen in Hansken  N= 13 MEAN=3.77 SD=0.83 MEDIAN=4
16	Ik voel me beperkt in mijn manier van zoeken binnen de Hansken-interface  N= 13 MEAN=2.92 SD=1.12 MEDIAN=3
17	Ik ervaar de Hansken Interface als gebruiksvriendelijk  N= 14 MEAN=2.93 SD=0.83 MEDIAN=3
18	Ik formuleer mijn zoekopdracht weleens met de Hansken Query Language (HQL)  N= 14 MEAN=3.86 SD=0.95 MEDIAN=4
19	Het kost me weinig moeite om te begrijpen hoe ik moet zoeken of resultaten kan verfijnen in Hansken  N= 14 MEAN=3.57 SD=0.85 MEDIAN=3.5
20	Ik voel me mentaal belast tijdens het uitvoeren van zoekopdrachten in Hansken  N= 14 MEAN=2.43 SD=0.85 MEDIAN=2
21	De interface ondersteunt mij goed bij het vinden van relevante informatie  N= 14 MEAN=3.43 SD=0.76 MEDIAN=3

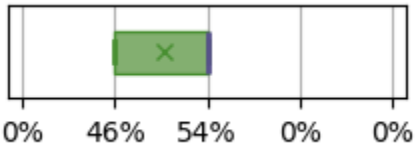
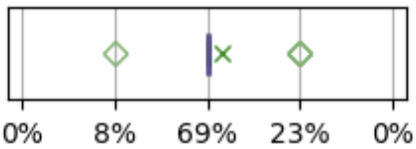
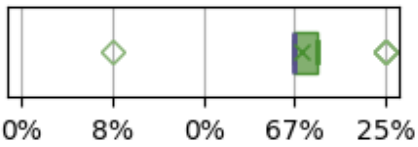
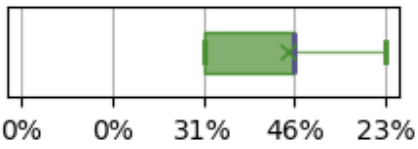
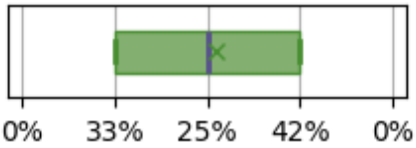
22	Ik wil graag mijn zoekopdracht met de muis aan elkaar kunnen klikken  N= 14 MEAN=3.08 SD=0.95 MEDIAN=3
23	Ik kan met Hansken efficiënt en snel een zoekopdracht uitvoeren  N= 14 MEAN=3.21 SD=0.89 MEDIAN=3.5
24	Ik zou het handig vinden om in natuurlijke taal (bijvoorbeeld gewone geschreven zinnen) met Hansken een zoekopdracht te formuleren  N= 14 MEAN=3.50 SD=1.16 MEDIAN=4
25	Ik zou suggesties of antwoorden van een AI-assistent vertrouwen, als deze assistent in Hansken ingebouwd is  N= 14 MEAN=3.07 SD=0.83 MEDIAN=3
26	Ik heb vaak te maken met complexe doelen of hypotheses, die ik zelf moet opdelen in kleinere uitvoerbare zoekopdrachten  N= 14 MEAN=3.36 SD=0.84 MEDIAN=4
27	Ik kan succesvol zoekopdrachten formuleren in Hansken  N= 14 MEAN=3.79 SD=0.70 MEDIAN=4
28	Ik vind het makkelijk om een zoekopdracht te formuleren in Hansken  N= 14 MEAN=3.50 SD=0.94 MEDIAN=4
29	Ik heb behoefte aan meer ondersteuning bij het formuleren van goede zoekopdrachten  N= 13 MEAN=3.00 SD=1.00 MEDIAN=3

	Een snel overzicht kunnen krijgen van relevante data of zoekresultaten
30.1	N= 13 MEAN=3.23 SD=0.83 MEDIAN=3 
30.2	N= 13 MEAN=4.46 SD=0.52 MEDIAN=4 
	Een nuttig visueel overzicht (bijvoorbeeld in dashboards of grafieken) krijgen van mijn zoekresultaten
31.1	N= 12 MEAN=2.75 SD=0.97 MEDIAN=3 
31.2	N= 13 MEAN=3.62 SD=0.51 MEDIAN=4 
	Een duidelijk overzicht krijgen van alle data die aanwezig is in het project
32.1	N= 13 MEAN=4.00 SD=0.58 MEDIAN=4 
32.2	N= 13 MEAN=4.00 SD=0.41 MEDIAN=4 
	Zoekresultaten kunnen filteren op eigenschappen zoals tijd, type of bron
33.1	N= 13 MEAN=3.85 SD=0.80 MEDIAN=4 
33.2	N= 13 MEAN=4.00 SD=0.82 MEDIAN=4 

	Meerdere filters tegelijk kunnen toepassen om een opdracht te verfijnen
34.1	N= 13 MEAN=3.77 SD=0.60 MEDIAN=4 
34.2	N= 13 MEAN=4.00 SD=0.71 MEDIAN=4 
	Duidelijk kunnen zien welke filters actief zijn
35.1	N= 13 MEAN=3.77 SD=0.44 MEDIAN=4 
35.2	N= 13 MEAN=4.08 SD=0.64 MEDIAN=4 
	Kunnen inzoomen op een specifieke set sporen
36.1	N= 12 MEAN=3.55 SD=0.69 MEDIAN=3 
36.2	N= 13 MEAN=3.92 SD=0.67 MEDIAN=4 
	Kunnen oriënteren waar in de data u op enig moment bent bij het vinden van een spoor of een set sporen
37.1	N= 13 MEAN=3.00 SD=0.71 MEDIAN=3 
37.2	N= 13 MEAN=4.00 SD=0.58 MEDIAN=4 

	Eenvoudig navigeren tussen detail- en overzichtsniveaus
38.1	N= 13 MEAN=2.92 SD=0.76 MEDIAN=3 
38.2	N= 13 MEAN=3.92 SD=0.64 MEDIAN=4 
	Extra details of informatie kunnen opvragen over een gevonden spoor
39.1	N= 12 MEAN=3.75 SD=0.87 MEDIAN=4 
39.2	N= 13 MEAN=4.08 SD=0.28 MEDIAN=4 
	Een spoor in zijn context kunnen zien
40.1	N= 12 MEAN=3.18 SD=0.60 MEDIAN=3 
40.2	N= 13 MEAN=4.33 SD=0.49 MEDIAN=4 
	Andere relevante of gerelateerde sporen rondom een gevonden spoor kunnen vinden
41.1	N= 12 MEAN=2.75 SD=0.62 MEDIAN=3 
41.2	N= 13 MEAN=4.15 SD=0.55 MEDIAN=4 

	Niet overspoeld worden met ongevraagde details
42.1	N= 12 MEAN=2.82 SD=1.08 MEDIAN=3 
42.2	N= 13 MEAN=4.25 SD=0.97 MEDIAN=4.5 
	Zelf controle houden over het zoekproces
43.1	N= 12 MEAN=3.67 SD=0.49 MEDIAN=4 
43.2	N= 13 MEAN=4.15 SD=0.55 MEDIAN=4 
	De mogelijkheid hebben om eerdere handelingen, stappen of toepassing van filters ongedaan te maken
44.1	N= 13 MEAN=3.77 SD=0.83 MEDIAN=4 
44.2	N= 13 MEAN=4.00 SD=0.71 MEDIAN=4 
	Hulp krijgen vanuit het systeem bij het opdelen van complexe vragen of zoekopdrachten
45.1	N= 12 MEAN=2.00 SD=0.60 MEDIAN=2 
45.2	N= 13 MEAN=3.31 SD=0.63 MEDIAN=3 

	De interface helpt mij om mijn denkproces te structureren tijdens het zoeken
46.1	N= 13 MEAN=2.54 SD=0.52 MEDIAN=3 
46.2	N= 13 MEAN=3.15 SD=0.55 MEDIAN=3 
	Een zoekopdracht kunnen formuleren met de Hansken Query Language (HQL)
47.1	N= 12 MEAN=4.08 SD=0.79 MEDIAN=4 
47.2	N= 13 MEAN=3.92 SD=0.76 MEDIAN=4 
	Kunnen zien wat de impact van een zoekfilter is op het zoekresultaat
48.1	N= 12 MEAN=3.08 SD=0.90 MEDIAN=3 
48.2	N= 13 MEAN=3.77 SD=0.44 MEDIAN=4 

Appendix D

Main Study Consent Form

Universiteit Utrecht

Consent form for participation in the research project

Combing Through Crimes: Understanding and Enhancing Search in Hansken

- I confirm that I am **18 years of age or over**.
- I confirm that the research project *"Combing Through Crimes: Understanding and Enhancing Search in Hansken"* **has been explained to me**. I have had the opportunity to ask questions about the project and have had these answered satisfactorily.
- I consent to the material I contribute being used to **generate insights for the research project** "*Combing Through Crimes: Understanding and Enhancing Search in Hansken*"**.**
- I consent to **audio and screen recordings** being used in this study. I understand that I can request to stop recordings at any time.
- I understand that if I give permission, the audio and screen **recordings will be held confidentially** so that only the researcher has access to the recordings. The recordings will be held on a secure laptop for **up to two months**, in which period they are transcribed in an anonymous form, redacted for sensitive information and the original recordings securely destroyed. From the screen recordings, de-identified **screenshots** can be made and might be published in the context of this research project. In accordance with the General Data Protection Regulation (GDPR) I can have **access to my recordings** and can request them to be deleted at any time during this period.
- I understand that my participation in this research is **voluntary** and that I may **withdraw** from the study at any time without providing a reason, and that if I withdraw any personal data already collected from me will be erased.
- I consent to the appropriately **de-identified data to be published** in appropriate data repositories **for verification** purposes.
- I consent to allow the de-identified data to be **used in future publications** and other scholarly means of disseminating the findings from the research project.
- I agree to take part in the above research project

If there are any further questions now or in the future, you would like to retract your consent or have insight of your personal data, you can contact the researcher at ***k.schoemaker@nfi.nl****. If you would like contact after August 2025, please reach out to* ***koschoemaker@gmail.com****.*

Name of participant	Date	Signature

Participant number (filled by researcher)

Appendix E

Main Study Protocol

Checklist

- Opname gestart?
- Docker + prototype gestart?
- Papieren kladblok + pen bij?
- Internet hotspot aan en verbonden met laptop?

Procedure

1. IJsbreker
2. Uitleg opzet onderzoek: Zoeken in Hansken. Interview, think aloud, interview. Uitleg think aloud
3. Consent form
4. Uitleg consent form: Opname van de sessie, opname twee maanden opgeslagen, getranscribeerd, anoniem

Interview

Ga in op interessante dingen die participant zegt. Volgorde van vragen aanpassen als we al eerder onderwerp aansnijden.

Demografie

- Wat is de naam van uw rol en hoe zou u die omschrijven?
    - Tactisch rechercheur, digitaal expert, operations, analist
- Wat voor soort zaken heeft u mee te maken?
- Hoe vaak gebruikt u hansken?
    - anders: andere forensische software?
- Gebruikt u de hansken interface HUIB?

Hansken

- Wat is een typische taak die je regelmatig uitvoert in Hansken?
    - of in andere digitale forensic tools
- Is dit goed uit te voeren in de interface?
    - Hoe voert u dit uit in de interface?
- Welke data is belangrijk voor jou?
- Op het moment dat u een nieuwe zaak begint, hoe begint u in een zaak in Hansken?

Zoeken en zoektocht

- Vergt zoeken in Hansken veel mentale inspanning?
- Voelt u weleens ontmoedigd, onzeker of geïrriteerd bij het zoeken of zoekopdracht opstellen met Hansken?
- Wanneer was de laatste keer dat je moeite had met het formuleren van je zoekvraag?
    - ook moeite met verfijnen?
- Past u vaak uw originele zoekopdracht aan? Bijvoorbeeld na het zien van de resultaten?
    - van categorie naar categorie springen
- Komt het vaak voor dat een zoekopdracht bij de eerste formulering geslaagd is of merk je dat de zoekopdracht meer iteratief veranderd?

Overzicht, filteren en zoektocht

- Vindt u het belangrijk om een overzicht van alle data, devices, personen oid te hebben?
    - hoe krijg je dit overzicht in hansken?
- Bent u tevreden over de filterfunctionaliteiten die bestaan in hansken
- Op het moment dat je een spoor vindt die interessant lijkt, wat is je vervolgstap?
- Wat mist u aan de zoekfunctionaliteiten of resultatenweergave?

    - denk aan inzoomen op interessante data, filteren van irrelevante data, zoeken op woorden, types. formuleren van zoekopdrachten.

Doelen

- Heeft u recentelijk een complex doel gehad die je op moest delen in meerdere vragen om te stellen?
    - [voorbeeld vakantie boeken?]
    - Bij het doorzoeken van digitaal bewijs heb je een groter doel of een hypothese die je wil bevestigen of ontkrachten. klopt dat?
- Is het voor u gemakkelijk om dit grotere doel om te zetten in kleinere stappen?

Natuurlijke taal

- Maak jij weleens gebruik van een systeem waarin je communiceert met natuurlijke taal, zoals recentelijk chatGPT of een soortgelijk programma?
    - hoe ervaar jij deze vorm van vragen stellen?
- Wat zou je denken als een degelijke vorm van vragen stellen in Hansken zou kunnen worden geïmplementeerd?

Think aloud

Neem aantekeningen van noemenswaardige zaken voor een volgende iteratie van prototype.

Aanmoedigingen:

- Wat probeer je te doen?
- Waar denk je aan?
- Wat gaat er door je heen?
- Weet ik niet, wat denk jij?
- Ik ben benieuwd naar je ervaring, vertel meer.

Post-task interview

- Hoe veel mentale inspanning heb je ervaren tijdens het uitvoeren van deze taken? (paas 1992)
- Was je onzeker, ontmoedigd of geïrriteerd tijdens een van de taken? (nasa)

Uitleg wat participant heeft gedaan, wat doel van interface was.

- Hoe verhoudt dit interface zich tot Hansken nu?
- Geeft dit een overzicht in de data?
- In welk stadium van een onderzoek zou een degelijk interface gebruikt worden?
- Wat vindt u van de manieren van interactie dat dit systeem biedt?

Advanced keyword

- Zijn geavanceerde zoekmethoden als dit iets wat u weleens gebruikt?
- Vind u het iets toevoegen dat hier een interface voor is?

Copilot

- Vond u de copilot een fijne manier van interactie?
- Vertrouwde u wat de copilot teruggaf?

Hulpvenster

- Heeft u het hulpvenster rechtsonder gezien?
- Zo ja:
    - Vind u het fijn als het systeem op deze manier met u meedenkt?

Appendix F

Think Aloud Exercises for Iteration 2

Think aloud

Bij deze opgave is het aangemoedigd op zoveel mogelijk **hardop na te denken** over wat u meemaakt; alle gedachten - hoe klein of irrelevant ook – die in u opkomen mag u uitspreken bij het doorlopen van de taken. Het is immens waardevol als u beschrijft **wat u ziet** en u uw **gedachten, acties, en frustraties** benoemt. Het is voor mensen heel onnatuurlijk om gedachten hardop uit te spreken, dus mogelijk moedig ik u soms aan om uw gedachten uit te spreken.

De onderzoeker zal tijdens deze opgave stil observeren. Op het moment dat u vastloopt in een opgave of u er niet uitkomt is het prima om naar de volgende opgave door te gaan. Onthoudt:

Wij testen u niet, wij testen het prototype

Het is niet erg als u ergens niet uitkomt, het is waardevol en zegt juist iets over het product.

Scenario

Er zijn twee verdachten opgepakt na een incident. Hierbij zijn twee telefoons en een laptop in beslag genomen. Gebruikerschap is vastgesteld als:

Verdachte 1	Anna	eigenaar **Samsung Galaxy S23 Ultra** en **Dell XPS 15 9530**
Verdachte 2	Bert	eigenaar **iPhone 14 Pro Max**

Taken

1. Bekijk op de tijdbalk **wanneer** er de meeste activiteiten waren op de **Samsung** telefoon
    a. In welke maand(en) of periode(s)?

    ➔ Het incident heeft plaatsgevonden in de maand **mei**

2. Schakel **het filter op devices en types** weer **uit** zodat u alle types data weer kunt zien
3. Gebruik de **tijdbalk** om alleen sporen van de maand **mei** te zien
4. Bekijk met wie **chatberichten** zijn gestuurd met de **iPhone van Bert** in de maand **mei**
    a. Met wie is het meeste chatcontact geweest?
5. Bekijk of er **foto's** zijn in de maand **mei** op een van de apparaten van **Anna**
    a. Wat staat er op de foto('s)?
6. Bekijk welke **websites** zijn bezocht op de apparaten van **Anna**
    a. Welke website is het meest bezocht?

    ➔ Het incident in mei vond plaats in **België**

7. Bekijk of er **chatberichten** in **mei** op de **Samsung of de iPhone** afkomstig zijn uit **België**
    a. Ja of nee?
8. **Schakel alle filters weer uit** totdat u de volledige dataset ziet
9. **Zoek op het woord** 'Bert' (let op hoofdletter B) en kijk of er **chatberichten** zijn met een **Bert**
    a. Wat staat er in de berichten?
10. Klik op **Advanced** onder de keyword zoekbalk, klik op de dropdowns en op het + icoon
    a. Wat verwacht je met dit nieuwe venster te kunnen doen?
11. Klik opnieuw op **Advanced** om het venster in te vouwen, en druk aan de linkerzijde op de knop **Copilot**
12. **Vraag in een volledige Nederlandse zin** of je zoekresultaten voor de maand **mei** van de **iphone** mag zien

Appendix G

Think Aloud Exercises for Iteration 3

Think aloud

Bij deze opgave is het aangemoedigd op zoveel mogelijk **hardop na te denken** over wat u meemaakt; alle gedachten - hoe klein of irrelevant ook – die in u opkomen mag u uitspreken bij het doorlopen van de taken. Het is immens waardevol als u beschrijft **wat u ziet** en u uw **gedachten, acties, en frustraties** benoemt. Het is voor mensen heel onnatuurlijk om gedachten hardop uit te spreken, dus mogelijk moedig ik u soms aan om uw gedachten uit te spreken.

De onderzoeker zal tijdens deze opgave stil observeren. Op het moment dat u vastloopt in een opgave of u er niet uitkomt is het prima om naar de volgende opgave door te gaan. Onthoudt:

Wij testen u niet, wij testen het prototype

Het is niet erg als u ergens niet uitkomt, het is waardevol en zegt juist iets over het product.

Scenario

Er zijn twee verdachten opgepakt na een incident. Hierbij zijn twee telefoons en een laptop in beslag genomen. Gebruikerschap is vastgesteld als:

Verdachte 1	Anna	eigenaar **Samsung Galaxy S23 Ultra** en **Dell XPS 15 9530**
Verdachte 2	Bert	eigenaar **iPhone 14 Pro Max**

Taken

1. Bekijk op de tijdbalk **wanneer** er de meeste activiteiten waren op de **Samsung** telefoon
    a. In welke maand?

    ➔ Info: het incident heeft plaatsgevonden in de maand **mei**

2. Schakel **het filter op devices** weer **uit** zodat u alle data weer kunt zien
3. Gebruik de **tijdbalk** om alleen sporen van de maand **mei** te zien
4. Bekijk met wie chatberichten zijn gestuurd met de **iPhone van Bert** in de maand **mei**
    a. Met wie is het meeste chatcontact geweest?
    b. Zijn er chatberichten met een **locatie** in Nederland?
    c. Is er financiële activiteit geweest?
5. Bekijk of er **foto's** zijn in de maand **mei** op een van de apparaten van **Anna**
6. Bekijk welke **websites** zijn bezocht op de apparaten van **Anna** in **mei**
    a. Welke website is het meest bezocht?

    ➔ Info: het incident in mei vond plaats in **België**

7. Bekijk of er **chatberichten** in **mei** op de **Samsung of de iPhone** afkomstig zijn uit **België**
8. **Schakel alle filters weer uit** totdat u de volledige dataset ziet
9. Druk bovenin beeld op de knop **Search**
    a. Vanaf welk **emailadres** zijn de meeste mails verstuurd?
    b. **Wanneer ongeveer** is een mail vanaf **dat** adres naar 'abc@xing.com' verstuurd?
10. Ga naar de tab 'Conversations'
11. **Zoek op het woord** 'Bert' (let op hoofdletter B) en kijk of er **chatberichten** zijn met een **Bert**
12. Ga bovenin het scherm terug naar '**explore**'.
13. Klik op **Advanced** onder de keyword zoekbalk, klik op de uitklapbare menu's en op het + icoon
    a. Wat verwacht je met dit nieuwe venster te kunnen doen?
14. Klik opnieuw op **Advanced** om het venster in te vouwen, en druk aan de linkerzijde op de knop **Copilot**
15. **Vraag in een volledige Nederlandse zin** of je zoekresultaten voor de maand **mei** van de **Iphone** mag zien