Orchestrating Sensemaking: Investigating the Sensemaking Processes and Its Enablers at the Dutch National Police

Master Thesis Business Informatics

22 November 2023

Matthijs Blaauw

m.blaauw@uu.nl - 4925653

First Supervisor
Dr. Inge van de Weerd

Second Supervisor
Prof. dr. Floris Bex

Daily Supervisor
Joeri Peters

Universiteit Utrecht
Department of Information and Computing Sciences

Dutch National Police
National Police Lab AI

Abstract

Introduction. Intelligence analysis is the process of creating an understanding of information. Here the information can be of many types and can contain gaps and inconsistencies. Intelligence analysis is essentially a sensemaking task, of which the process has described in earlier studies. We reason that with the emergence of Big Data and AI, this process might have changed. Additionally, there is limited work available on the factors that enable sensemaking in organisations.

Method. In this study, we have conducted a case study at the Dutch Police to investigate its intelligence analysis operations. The case study consists of three embedded cases: *investigation analysis*, *security analysis*, and *strategic analysis*. A total of 17 interviews have been conducted for this research.

Findings. In the within-case analysis, we present the intelligence analysis processes found inside each case. In investigation analysis, the analyst works to schematise information that is collected by detectives and to derive insights from it. In security analysis, information on a specific criminal phenomenon within a unit is monitored and analysed. For strategic analysis, goal-driven analysis projects take place to provide leadership with information about a criminal phenomenon. The cross-case analysis resulted in a generalised description of the sensemaking process in the police organisation. Inside the Dutch Police, sensemaking is a combination of six activities: *filtering*, *searching*, *reading*, *converting/importing*, *schematising* and *analysis*. Products of sensemaking are *information needs*, *schemas*, and *insights*. We have furthermore looked at how the sensemaking processes interact to form a sensemaking ecosystem. Additionally, a set of 24 enablers are listed in the categories of data, software, organisation and process for sensemaking in the organisation.

Discussion. We have found schematisation to be a crucial part of sensemaking. Contrary to earlier literature, we found that sensemaking in our case study takes place through predefined structures, causing the search for good representations to be missing in our found sensemaking process. Additionally, the new concept of an *information flow* inside the sensemaking process is introduced. Also new to the process is a description of how analysis of structured non-textual data can be part of the sensemaking process. Lastly, based on the enablers a set of recommendations are presented for the Dutch Police organisation to improve its sensemaking operations.

Contents

1 Introduction

Intelligence and its related activity intelligence analysis are both terms with a broad definition. In its essence, it is the process of creating an understanding of information [33]. Here, information can be of many types, including raw data, events, and evidence. Yet, this information can contain gaps and inconsistencies. Intelligence is derived from information which is transformed in order to be fitting to the customer's unique environment [28]. It is an intelligence analyst's job to interpret the information in a process of "connecting the dots" and creating hypotheses [29]. In the process of interpreting the data, the analyst plays a crucial role, where a large part of the analysis is happening in the analyst's mind [60].

Intelligence analysis is essentially a sensemaking task, which process has been analysed in many studies. Pirolli and Card have described the process in an influential empirical study on intelligence analysis [40]. They found that sensemaking is a combination of two loops: The *Foraging-loop*, the collection and subselection of information into a so-called *shoebox*, and *sensemaking-loop* where information from the shoebox is analysed to create a hypothesis and eventually an understanding for the decision maker. However, there are limited descriptions of how the hypotheses formulation process works, and what tools are used [54].

Much has changed since the influential work by Pirolli and Card; with the emergence of Big Data and related AI technologies the need for an understanding of sensemaking is greater than ever [46]. AI can help with automatically interpreting large amounts of data but does not always justify its output, requiring the analyst to make assumptions for interpreting its output. The process of sensemaking is in itself susceptible to biases, when the interpretation of information happens solely in one analyst's mind [16]. The AI systems that analysts may use can likewise have challenges linked to biases and unfairness [34, 1, 36] since they can include black-box algorithms and do not provide a rationale for their output. The latter is actively being addressed with research into Explainable AI (XAI) (e.g. LIME [45]). Interestingly, there is little research on how analysts use new AI-like technologies, and on how this influences human-computer interaction during the sensemaking process.

With the rise of Big Data, the volume, variety, and, velocity of data have also increased [8]. This increase in available data would suggest that the process of sensemaking has also become more complex since there is more data to make sense of, which is not necessarily easy to interpret. Different technologies can help organisations to make full use of the opportunities provided by large amounts of data [38]. These technologies can aid in data collection and analysis, and possibly automate some tasks in the data analysis processes. This research is concerned with investigating how sensemaking within intelligence analysis has changed with the increase in the volume of data, and the emergence of AI technologies.

Lastly, while there is research available about sensemaking and its processes, new to the field is a study that looks at multiple sensemaking processes in one organisation and how these interact. This research will look at the different intelligence analysis processes inside the Dutch Police, to see how the organisation can utilise large amounts of data and how sensemaking plays a role in the analysis of this information. The analysis of the processes allows us to extract the factors that enable sensemaking in this sensemaking ecosystem.

1.1 Research Context

For this research, the intelligence analysis operations are studied at the Dutch National Police. Intelligence analysis occurs inside different units of the Police, with intelligence-led policing (ILP) being an important motivation within the organisation [13]. ILP's goal is to use intelligence to more effectively direct policing efforts and to combine intelligence with research to create evidence-based strategies [44].

The research will investigate the intelligence analysis inside multiple units of the Dutch Police. All ten regional units and one national unit have intelligence departments. All of these units have a sub-department that focuses on the analysis of information. This sub-department has specialised analysts who are attached to investigative detective teams focusing on the prosecution of a criminal or criminal group. Analysts can also work on security analysis, which is focused on a specific criminal phenomenon, such as synthetic drugs, cocaine, or terrorism. Lastly, there is also strategic analysis happening in each department helping leadership to set grounded strategies.

1.2 Research Objective

This research is concerned with investigating how sensemaking occurs within the Dutch Police, specifically how analysts at the Dutch Police do intelligence analysis. The research is motivated by a knowledge gap in the current literature on how sensemaking has changed with the increase in available data and the possible influence of AI technologies.

Additionally, this research forms a basis for future work. The design of computer-based systems for aiding an analyst during the different tasks within sensemaking is an active research area (e.g. [60, 54]). This work will describe the real-life processes of analysts. These process descriptions can provide future research with the basis for designing new improved sensemaking-aiding systems.

Lastly, this research provides the scientific community insights into how these found processes interfere, and combine to form a sensemaking ecosystem. Further analysis allows us to get insights into which factors enable sensemaking in the organisation. As of the writing of this paper, no other studies have been found that list the enablers of sensemaking in an information science context.

1.3 Research Questions

The following Main Research Question (MRQ) has been created for this research:

MRQ How does sensemaking take place inside the intelligence analysis operations at the Dutch Police, and what are the enablers of sensemaking in the organisation?

To answer this MRQ multiple Research Questions (RQs) have been constructed. First, we are interested in the current literature about sensemaking. The goal of this RQ is to gain information about what can be expected of the working method of analysts. The first research question is answered by a literature review.

RQ1 What does earlier research describe about the sensemaking process?

After a basis for sensemaking is established, the intelligence analysis operations inside the Dutch Police are investigated. This is accomplished by conducted a case study inside the organisation and interviewing analysts and other stakeholders.

RQ2 How is intelligence analysis conducted at the Dutch Police?

To answer RQ2 multiple Sub-Research Questions have been constructed (SRQ):

SRQ2.1 What are the roles in the intelligence operations at the Dutch National Police, and what are their main tasks?

SRQ2.2 How do the analysts perform their tasks in intelligence analysis at the Dutch Police?

SRQ2.3 What applications are used during intelligence analysis at the Dutch Police?

The collected information from the case study in RQ2 is generalised through a cross-case analysis in order to answer RQ3.

RQ3 How does sensemaking take place at the Dutch Police, and what are its enablers?

Finally, the answers from the research questions are combined to answer the MRQ.

2 Research Method

This research consists of two parts, a literature review and a case study. In this chapter, we will introduce both methods and provide a rationale for their design. The full research method is displayed in Figure 1.

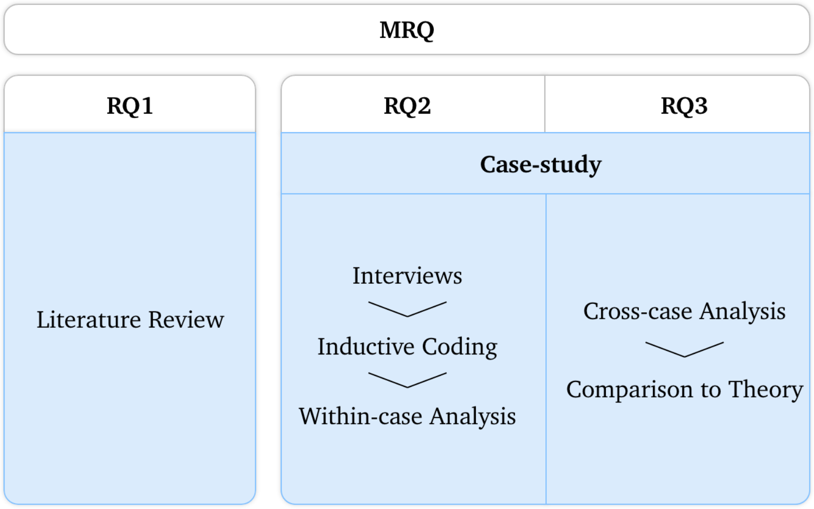


Figure 1: A diagram of the research method of this thesis.

2.1 Literature Review

In this study, the intelligence analysis operations at the Dutch Police are investigated. Before conducting this case study, it is useful to study existing work on sensemaking and intelligence analysis.

This literature review is conducted as a narrative literature review instead of a systematic review. The goal of the literature review is to get an overview of the current state of research, as well as to get an overview of what earlier research describes about the sensemaking process. Later, findings from studies found in the review can be compared to our findings, to further highlight the implications for the research community.

According to a study by Snyder, narrative literature reviews are best applied for overviewing a topic [49]. It is also effective for relaying how research progresses over time. A systematic literature review is best used if articles must be compared to quantify effect size. However, our research is of exploratory nature, where we do not quantify a specific phenomenon. The primary goal of our literature review is to answer RQ1 by creating an overview of what existing work describes about sensemaking and its process. These properties make a narrative literature review more fitting for this study. Snyder also mentions that narrative reviews are best suitable for a broad study of a phenomenon that crosses multiple fields [49]. This is the case for sensemaking since it crosses information science, human-computer interaction, and cognitive psychology.

2.2 Case Study

In order to answer the research questions RQ2 and RQ3, and their related sub-questions, the intelligence analysis operations at the Dutch National Police are investigated. This research is classified as an *embedded-single case study* [62]. The case study focuses on intelligence analysis and investigates who the main actors are, what their tasks are, what their work method is, and which systems they use. Specifics regarding the execution of the case study can be found in the Case Study Protocol in Appendix A.

2.2.1 Case Context

The case study is executed at the Dutch Police. The Dutch Police is a large and broad organisation. Since the reorganisation in 2012, intelligence has been implemented throughout the organisation, with Intelligence-led policing being an important design philosophy [15]. The Dutch Police is divided into ten regional units and one national unit. This case study focuses on intelligence analysis at the regional DRIOs (NL: Dienst Regionale Informatie Organisatie, EN: Regionale Information Unit) and the national DLIO (NL: Dienst Nationale Informatie Organisatie, EN: National Information Unit). Intelligence analysis happens at the A&O (NL: Analyse en Onderzoek, EN: Analysis and Research) departments of the regional and national intelligence departments. Analysis roughly happens on three levels: (I) *investigation analysis*, aiding a detective team in a specific crime investigation; (II) *security analysis*, looking into a specific criminal phenomenon; and (III) *strategic analysis*, which looks into a criminal phenomenon on a higher level, to aid leadership in setting strategies. These different levels of analysis also form the different embedded cases in this study. For investigation analysis, the analyst is attached to an investigation team that is part of either the DLR (NL: Dienst Landelijke Recherche, EN: National Detective Unit) or its regional counterpart the DRR (NL: Dienst Regionale Recherche, EN: Regional Detective Unit). Therefore these departments are also of interest for this case study. A high-level overview of the structure of the departments of interest within the Dutch Police is shown in Figure 2. The aim of this case study is to describe the typical instance in a *contemporary study* [62], which entails that we are looking at the current everyday work method of analysts.

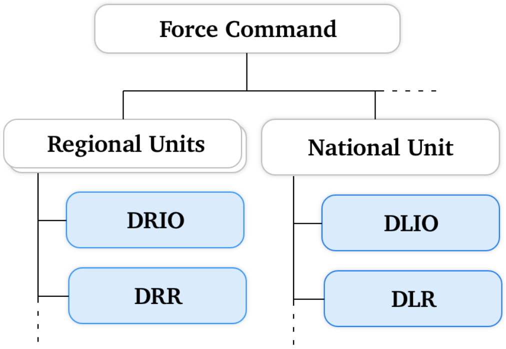


Figure 2: A diagram of the primary departments of the Dutch Police where intelligence analysis occurs. The departments of interest for this thesis are highlighted in blue.

2.2.2 Data Collection

Details about the intelligence analysis operation are collected through interviews with actors involved in the intelligence analysis process. The interviews are conducted as semi-structured interviews. These interviews allow for obtaining information about a subject using questions which are constructed beforehand, but also allow for more flexibility by providing the opportunity to ask follow-up questions about statements from the interviewee [37]. All interviews are recorded after informed consent is obtained.

The embedded cases for the case study are: (case I) investigation analysis, (case II) security analysis, and (case III) strategic analysis. For each embedded case multiple interviews are conducted. Interviewees will include employees who currently work as analysts, but also other related employees, such as managers, or people related to the IT that the analysts use. For the different levels of analysis, the interview protocol is slightly altered to adapt to the changes in general work methods.

Besides the data collected through the aforementioned interviews, manuals of related IT systems will also be looked at. This information will allow us to describe the functionality of the used applications in more detail.

2.2.3 Data Analysis

The qualitative data collected through the interviews in the case study is analysed in an inductive manner. The *general inductive approach* by Thomas is used, which can be applied to *"develop a model or theory about the underlying structure of experiences or processes which are evident in the text"* [53]. One of the goals of this research is to provide an accurate overview of how intelligence analysts operate at the Dutch Police. The inductive approach allows us to derive an accurate description of the current working method, with little influence of existing theory.

Coding. Transcribed interviews are coded using NVivo. The inductive coding process is performed based on the *general inductive approach* by Thomas and follows the following process [53]. First, the transcriptions and other documents are read to gain familiarity. The first themes are generated based on the understanding of combined texts and some are based on the research objective. One of the objectives of the study is to identify processes, actors, products, and systems. These objectives help identify general themes, for example, a general activity in the process can be a theme. Coding also starts here, while further reading of the text will derive lower-level codings with more detail. Some examples of our top-level themes are: *Analysis Duration*, *Analysis Goals*, and *Analysis Activities*. During the coding process, the category system is redefined further. This can include adding subtopics, adding counter-evidence or merging similar categories. In our study, for example, we add the individual activities found as sub-themes to the top-level *Analysis Activities* theme. The coding also consists of the *Enablers* and *Inhibitors* top-level themes. These two themes allow us to classify the factors that benefit sensemaking and the factors that hinder sensemaking. Note that coverage is not a goal, since not all information in the raw data is related to the research objectives. Additionally, multiple codes can also be assigned to one phrase. The resultant coding is described in the code book, which can be found in Appendix E.

Cross-case Analysis. After results from the data analysis are presented for each embedded case (i.e. investigation, security, and strategic analysis), a cross-case analysis is presented. A comparison between multiple cases can help improve robustness compared to analysing a single case [62]. The analysis processes that are presented in the individual cases are compared in this section. The comparison allows us to find similarities, and to better generalise the cases and findings. Comparing the embedded sensemaking processes also allows us to report on the shared context in which these processes operate. Combining the findings from all interviews also lets us find the enablers of the sensemaking processes analysed. The findings from the cross-case analysis can be used to answer RQ3.

Comparison to theory. After the cross-case analysis, the findings are compared with existing theories. In this comparison, we will reflect on how the found processes differ from what existing studies have described. We will furthermore compare the found enablers with what earlier literature has found.

2.2.4 Threats to Validity

Yin describes four threats to the validity of case studies [62]. In this section, we will describe how we try to mitigate these risks for our study.

Construct Validity. To limit the threat that the collected data does not accurately reflect the actual processes at the police, we try to improve construct validity. Multiple analysts are interviewed about the same processes, so we get information from multiple sources. Where possible a team leader/manager of the analysis is also interviewed to get information from a different viewpoint. These interviews will add up to a relatively high number of interviews for this research, mitigating at least partially, the construct validity threat.

Internal Validity. Internal validity issues are not applicable to this exploratory research, since no explicit causal relations are inferred.

External Validity. We partly accept the risk that generalisability might be limited in this research. The research looks at the intelligence analysis operations at the Dutch Police. A multiple-case study is chosen to get a good view of the actual operations of the Police, so our findings generalise to the Dutch Police. A subset of the found enablers are also generalisable to outside the police context. Lastly, general findings from the cross-case analysis and comparison theory are used to give direction to future research about sensemaking.

Reliability. The methodology of the case study is documented thoroughly including a Case Study Protocol (Appendix A) and an Interview Protocol (Appendix B) to facilitate replicability.

3 Related Work

Sensemaking is a research topic that touches multiple fields and lacks a single definition. In this chapter, we will briefly discuss sensemaking in organisational science and continue by describing the studies about sensemaking in the Information Science (IS) and Human-Computer Interaction (HCI) fields. Within the latter fields a distinction can be made between three different sub-areas [41]: (1) *representation-focused sensemaking*, (2) *cognitive-focused sensemaking*, and (3) *collaborative sensemaking*. We will continue the review by including the topics of sensemaking in information systems and intelligence analysts in practice. Lastly, the findings of our literature review will be summarised to answer RQ1.

Sensemaking in the research area of organisational science is a large and active field, with the theory by Weick from 1995 being a defining study [57, 14]. His theory describes how organisations and individuals make sense of events and their environment. Uncertainty and ambiguity trigger sensemaking, making an individual search for meaning. Weicks work and the related studies following it are part of the social psychology discipline. Topics addressed in this field of sensemaking research include politics, power, culture, identity, and organisational change [6]. These topics are relevant for organisational science but are not much relevant to the sensemaking researched in this paper. The following parts of the related work will therefore focus on sensemaking within the IS and HCI fields.

3.1 Representation-Focused Sensemaking

In this section, we will introduce the different models of interest in representation-focused sensemaking. This form of sensemaking primarily focuses on how analysts handle information in order to make conclusions about it.

3.1.1 Learning Loop

The basis for research on sensemaking in representation-focused sensemaking is the *Learning Loop* which was introduced in work by Russell et al. in 1993, where they generalised the different tasks in making sense of large amounts of information [47]. They define sensemaking as *"the process of searching for a representation and encoding data in that representation to answer task-specific questions"* [47]. One of the findings of the study is the *Learning Loop* as a part of sensemaking, which is illustrated in Figure 3.

The model contains the following data forms: A *Representation* is a way of representing selected data; *encodons* are instantiated representations containing the retrieved information that aid in answering a task-specific question; and *Residue* contains data that does not fit the representation and also contains unused representations. These data types are consumed or produced by three main processes: during the *search for representations* sensemakers search for a way to represent the data, this includes the creation of a method to transform data to the presentation; when *instantiating representation* the sensemaker searchers for related data and encodes it in a representation, which creates encodons; and *shift representations* where the sensemaker analyses residue to check if the presentations should be expended. The final results of the learning loop are encodons, which can be used for presenting the findings of an information processing task.

The learning loop has been used in work by Qu and Furnas to evaluate exploratory search [42]. In this research Qu and Furnas further extend the learning loop by incorporating a *structural information need* in the process, which can be answered using the results from search queries. The model is shown in Figure 4. The additions to the original learning loop are the result of three main findings: (1) generating a representation can use a form of inspiration other than existing knowledge, namely structure ideas or ready-made chunks of structure which both can be found by searching the web; (2) The current representation structure helps shape new searches, which in turn helps to get ideas for new structures and validate existing structures; (3) When a structural-information need arises, the need is especially for structure rather than information, making it non-trivial to select keywords for the search. In the structured information-seeking cycle, users query search engines with related words to get related web articles. They use the found information to not directly extract useful information but use it as pointers to structured information on the web. These patches of structures can

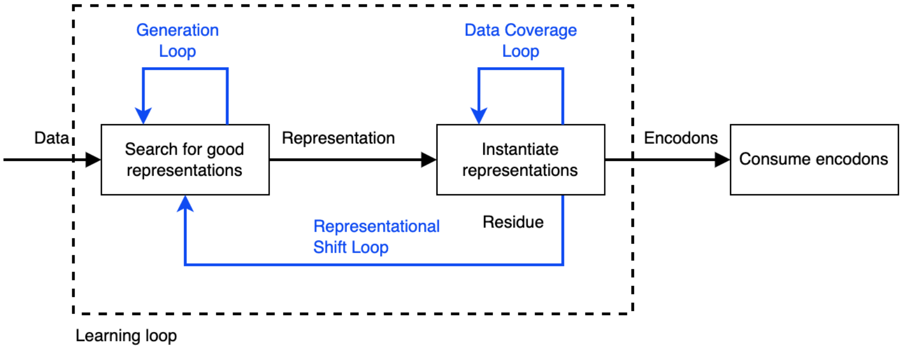


Figure 3: A diagram of the learning loop described by Russell et al. [47].

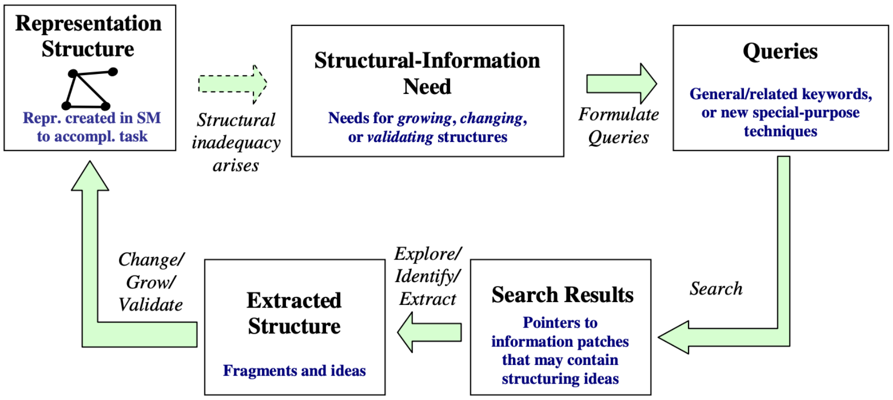

Representation Structure
Repr. created in SM to accompl. task
Structural inadequacy arises
Structural-Information Need
Needs for growing, changing, or validating structures
Formulate Queries
Queries
General/related keywords, or new special-purpose techniques
Search
Search Results
Pointers to information patches that may contain structuring ideas
Explore/ Identify/ Extract
Extracted Structure
Fragments and ideas
Change/ Grow/ Validate

Figure 4: A diagram of the structural-information seeking cycle by Qu and Furnas [42].

be used to validate and elaborate on the current representation structure. As the authors describe: *"Search is no longer a means to find information that directly satisfies users' information needs. Instead, the search leads people to an information patch where users can explore useful information structures."* [42]

3.1.2 Sensemaking Process by Pirolli and Card

One of the most important studies on sensemaking in information science is the research from Pirolli and Card from 2005 [40], where they continued on the research from Russell et al. which introduced the aforementioned learning loop [47]. They applied cognitive task analysis to the process of intelligence analysis, to further study the sensemaking process. The result of their analyses is a model of the sensemaking process and is displayed in Figure 5. They describe six forms of data that are transformed during the process (mentioned in increasing order of effort and structure): (1) *external data sources* contain raw data; (2) the *shoebox* is a smaller set of relevant data that should be processed; (3) the *evidence file* contains parts of information from the shoebox; (4) *schemas* are representations that aid in making conclusions; (5) *hypothesis* are unconfirmed conclusions with supporting arguments; (6) *presentation* is the end product which can be used to communicate findings.

The described process contains two important loops. First, a *foraging loop* which loops between external data sources and the extraction of information into a schema. It involves the filtering and extraction of relevant information from sources. Second, the *sensemaking loop* which was derived from the learning loop by Russell et al. and consists of the iterative creation of a mental model of the schema. This in turn includes an elaboration of the problem (hypothesis generation), reasoning (finding evidence to confirm or disprove

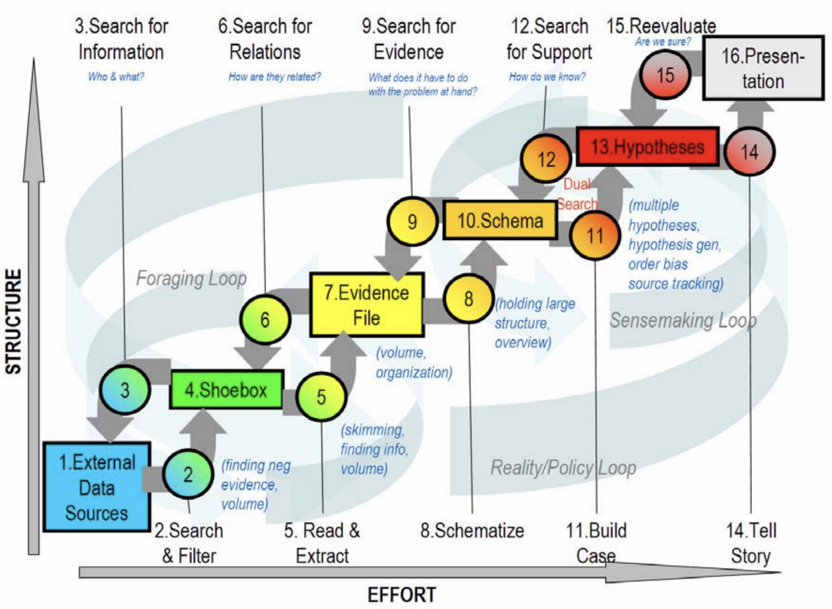


Figure 5: A diagram of the sensemaking process identified by Pirolli and Card. [40]

the hypothesis), and decision-making.

The foraging loop in this model can be further specified as a switch from exploratory search to focused search. With exploratory search low-precision queries are executed to find relevant documents, later specific high-precision sets of documents will be created. Eventually, these documents will be *exploited* to extract information by it using analytical techniques such as pattern recognition.

Pirolli and Card furthermore describe the sensemaking process as both bottom-up and top-down. The bottom-up process starts with external data sources and consists of the following tasks: *search and filter*, *read and extract*, *schematize*, *build case*, and *tell story*. And their top-down process consists of: *re-evaluate*, *search for support*, *search for evidence*, *search for relations*, and *search for information*.

Pirolli and Card's sensemaking model has been compared to sensemaking by real analysts. A field study with intelligence analysts found differences between the way they observed how analysts work, and how it is described in Pirolli and Card's model [21]. The main difference found is the non-linearity in the working method of the analysts. This non-linearity is expressed as having loops in the process, but also as a lack of sequence in the tasks. They also found a "parallelism" in the tasks from the analysts. To describe this "parallelism", the paper refers to an unpublished manuscript from Wheaton (2011) which describes the intelligence analysis process with a model displayed in Figure 6. The model shows that multiple activities occur concurrently where their prevalence changes over time. These differences are explained by the sensemaking model from Pirolli and Card having a focus on information transformation, rather than on the actual workflow of an information analyst. The model describes how information is transformed from raw data to information but is not fully in line with the workflow of an analyst.

The sensemaking model by Pirolli and Card has also been compared with the actual workflow of analysts in four different fields (biology, cyber security, intelligence analysis, and data science). For most of these fields, the researchers observed that not many analysts performed top-down reassessment steps (e.g. *re-evaluate* or *search for support*) [19]. They also found that the activities in Pirolli and Card's model preceding the information shoebox were not often present in their work method. In many cases, the analysts were already provided with a collection of filtered domain-specific data. Some differences were found between the work methods of analysts and their fields. Cyber and data science analysts appeared to mostly work on data

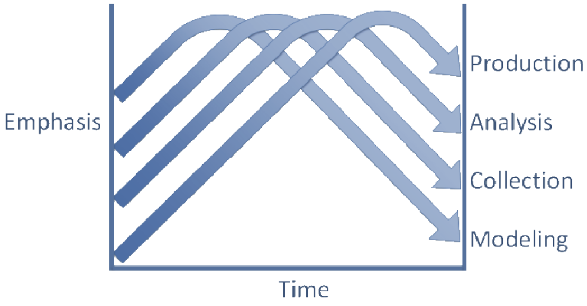


Figure 6: Wheaton's Multi-phasic Model of Intelligence Analysis.

foraging, where in some cases they only needed to transform unstructured data into a clean spreadsheet. For biology and intelligence analysts they found that their work primarily involves the sensemaking loop, specifically schematising, where the analyst creates a hypothesis and evaluates it on the (newly) found data.

3.1.3 Krizan's Cyclical Model of the Intelligence Process

Intelligence Analysis is essentially a sensemaking task. One such model that describes this process is the cyclical model presented by Krizan [28]. They present a model for intelligence analysis in a government and business context. The model is shown in Figure 7 (the model was originally created by the Department of Defense, but no further reference was given). Important to note with this model is that the actual process is not unidimensional and unidirectional like the model would indicate. This is an intended choice by the author to improve readability.

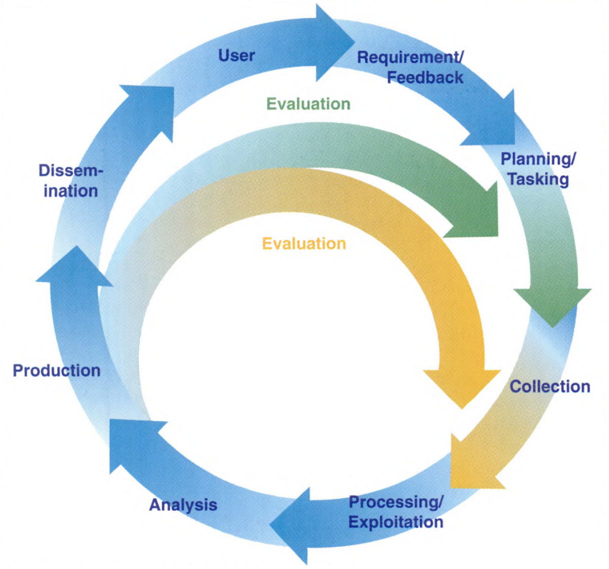


Figure 7: Cyclic Intelligence model.

The main can be described in five main steps [66]: The process starts with *Planning/ Tasking* which is driven by an intelligence need from an end-user; afterwards, information is *collected* for this need; this information is then *processed* where information from the collection is further subselected; this information is then *analysed* where analysts try to interpret the information and makes sense of it. Based on the found information, an information report is *produced*. Afterwards, the report is disseminated to the users, which results in (further) requirements and feedback.

3.2

Another form of individual sensemaking is cognitive-focused sensemaking, where existing studies focus on describing the high-level processes in an analyst's mind. The field is interested in how analysts go about creating an understanding of information. The following section will discuss existing work that describes these cognitive processes of a sensemaker.

3.2.1 Data/Frame Theory

One of the main studies on cognitive sensemaking is by Klein, Moon and Hoffman, where they introduce the Data/Frame Theory [24]. Klein later elaborated on this model in [26]. The Data/Frame Theory describes that sensemaking is the cyclic process of elaborating on an initial frame. Here the frame refers to a starting viewpoint or perspective. Note that this frame starts with an effort to make sense of data, while the frame actually also shapes the data. The authors give an example: "A house fire will be perceived differently by the homeowner, the firefighters, and the arson investigators" [24]. The Data/Frame theory is illustrated in Figure 8. It includes the following activities that manipulate the frame: (1) *Elaborating* on the current frame, by adding details and questioning it to strengthen its proof. (2) *Perserving* the frame by disproving conflicting data. (3) When *Reframing* an analyst reconsiders the previous initial frame and searches for better alternatives or creates a new frame to replace the first.

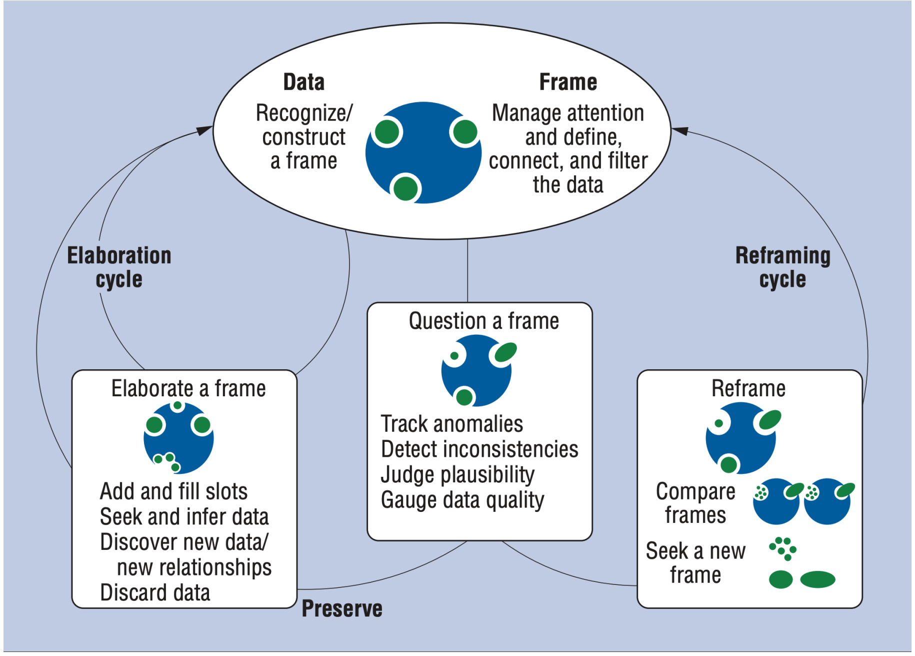


Figure 8: The Data/Frame Theory from Klein, Moon, and Hoffman [24]

Hypothesis generation and assessment seem to be crucial activities in sensemaking [19]. In the Data/Frame Theory hypothesis formulation already takes place during the initial phase of the analysis. Klein et al. state that this is in contrast to some earlier beliefs, where it is stated that this could lead to fixation errors [16]. Klein et al. however describe that this problem is limited for experienced analysts since they can more effectively analyse data by questioning the hypothesis in their mind. The same contradiction holds for a potential confirmation bias as a result of working with a frame from early on, which is also mentioned in the sensemaking work from Pirolli and Card [40]. Klein et al. claim that experts tend to focus on data that can disprove their current hypothesis rather than focusing only on confirming data, being unaffected by a confirmation bias.

There is some criticism of the use of the word "*frame*" by Klein et al., Attfield and Baber state that the "*frame*" can better be explained as an activated subset of an associative network, or collection of non-connected networks in an analyst's mind [4]. This network (or networks) will contain not only information about the specific situation that an analyst is trying to understand but also their prior experiences and generic beliefs. This is in contradiction to the interpretation that a frame is only a representation of the current situation, disregarding the prior beliefs of a sensemaker.

3.2.2 Applications of the Data/Frame Theory

The Data/Frame Theory has been compared to sensemaking activities in many fields. Hudson and Singh argue that the Data-Frame theory is the most appropriate model for medical decision-making. They state that the cyclic interaction between the practitioner and data about the patient contributes to sensemaking, where a frame is developed by the expert [17]. A similar conclusion was made by a naturalistic study that investigated sensemaking by US high-performance fighter jet pilots [22]. The authors mentioned that the Data/Frame Theory overlaps with the found pilot's activities, and state that *"this research supports the sensemaking model as a comprehensive conceptualisation of sensemaking"* [22].

The Data/Frame Theory has also been compared to analysis in the intelligence field. Moore and Hoffman found that the Data/Frame Theory overlaps with how analysts create an understanding of events [35]. The authors additionally mention that the Data/Frame Theory allows for an explanation of so-called "bias" during intelligence analysis. This work from the National Military Intelligence Foundation states that a "bias" is inherent when creating a frame in the initial phase of sensemaking since experiences from past events are also considered by the analyst. Additionally, it is a common misunderstanding that the data analysis part is the most complicated of intelligence analysis. However, based on interviews with Intelligence Analysts it was found that constructing a frame is a more complex and important activity for the process [21]. It is the initial process of getting an understanding of information and knowing what to investigate and how to do it that is difficult. The process of knowing what questions to ask relies on an understanding of the data.

Further analysis has been done regarding the Data/Frame theory and criminal investigations. Barret continued the Data/Frame model, by incorporating his findings from research into criminal investigations by UK detectives [5]. Based on these findings Barret created a model, which is not shown in this paper. However, important here are some assertions he made about the investigative sensemaking process. Investigative sensemaking overlaps with the Data/Frame model on hypothesis-generating, elaborating, and testing. It also describes to role of recalling of existing knowledge. Two findings are however new to sensemaking as described by Klein et al. First, Investigative sensemaking is goal-directed. Second, investigative sensemaking drives the choice of taking fitting investigative actions to fill in knowledge gaps.

In other case studies the Data-Frame model was applied for: abnormality detection by surgeons during cholecystectomy [50]; creating an understanding of a hurricane emergency by relief agencies [9]; extracting information from reports from unmanned aerial systems in the military [56]; studying how sensemaking occurs by air traffic controllers [32]; and how to create training for intelligent software tools [11].

Another group of studies focused on if insights into sensemaking activities from the Data/Frame Theory model can be used to analyse and possibly improve computer-based sensemaking systems. A general mythology was proposed to get more insights into the computational techniques of sensemaking [30]. The mythology would use Machine Learning to classify user interaction into sensemaking patterns. Another study used a comparable method to evaluate sensemaking computer programs, they used low-level logs of computer systems in combination with manually annotated labels from the Data/Frame Theory as predictor variables [27]. This resulted in a dataset of sensemaking activities and their corresponding interactions with sensemaking software. They trained a Machine Learning algorithm to classify interactions into sensemaking activities.

The resulting prediction accuracy depended on which software programs were used during the analysis. They found that interaction logs from using only a web browser and word processor were not successful predictors of sensemaking activities. However, the logs from the information program which allowed for information visualisation, called INVISQUE [59], were able to be used to predict the distinct sensemaking activities. The authors argue that this might be an indicator that information visualisation software is beneficial to sensemaking.

3.2.3 Cognitive Sensemaking Process

A model that has been constructed based on combining the different existing sensemaking models (including all those described above) is created by Zhang and Soergel [66]. In their work, they continue on a comprehensive model of the cognitive processes of individual sensemaking, which was earlier introduced in [67] and [64]. Their model is shown in Figure 9. One of the main features of the model is that most activities can be reached from most other activities. As described earlier, other research highlights that non-linearity is common in the working method of analysts and is a criticism of other models [21].

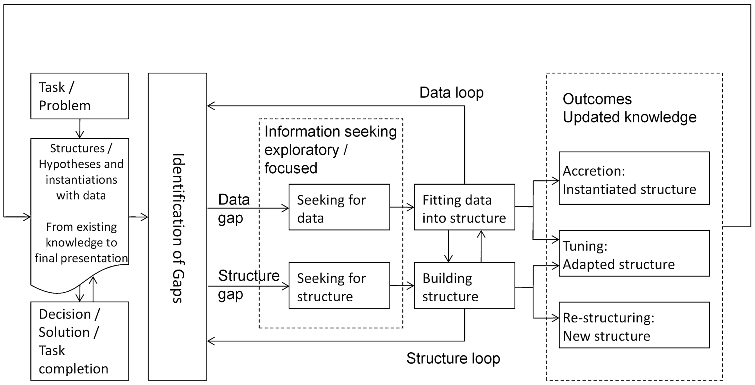

Task / Problem
Structures / Hypotheses and instantiations with data
From existing knowledge to final presentation
Decision / Solution / Task completion
Identification of Gaps
Data loop
Information seeking exploratory / focused
Data gap
Seeking for data
Fitting data into structure
Structure gap
Seeking for structure
Building structure
Structure loop
Outcomes Updated knowledge
Accretion: Instantiated structure
Tuning: Adapted structure
Re-structuring: New structure

Figure 9: Cognitive sensemaking process by Zhang and Soeergel [66].

Also new in this model is the differentiation between a *data gap* and a *structure gap*, where both are the result of the *Indentification of gaps* activity. A structure gap is handled by searching for relations and patterns in the found data. The data gap is solved by looking for specific data, and fitting it into the previously build structure. Both gaps are part of two loops, which are closely intertwined. Further differentiations can be made between the loops, namely exploratory search during a search for structure and a focused search in a search for data (searching *for* sources vs searching *in* sources). Like the Pirolli and Card model, the activities can be executed top-down or bottom-up, starting with structures or data respectively.

Results from a later empirical user study from the original authors of the model show that the cognitive sensemaking model also applies when combining it with search (search-sensemaking) [65]. The authors again found that the activities are non-linear and iterative. They furthermore found that most processes were ad hoc rather than planned. They also elaborated on the role of prior knowledge, by stating that information that is in line with existing knowledge was value-adding to sensemaking, but conflicting information to prior knowledge was challenging to sensemaking, leading to confusion.

Recently Zhang continued this work by further describing how actors prototypically interpret information [63]. The model shows the sequence of cognitive activities when making sense of data. This model is shown in Figure 10. Important to note that, like with other models mentioned earlier, the authors describe that this sensemaking process is non-linear. This model is the only discussed model with decision points, showing that activities like *Comparison* and *Classification* are only reached when the data is clear, otherwise, the activities *Definition* and *Specification* will be reached. The sequence is described in both a top-down (logic-driven) and a bottom-up (data-driven) manner.

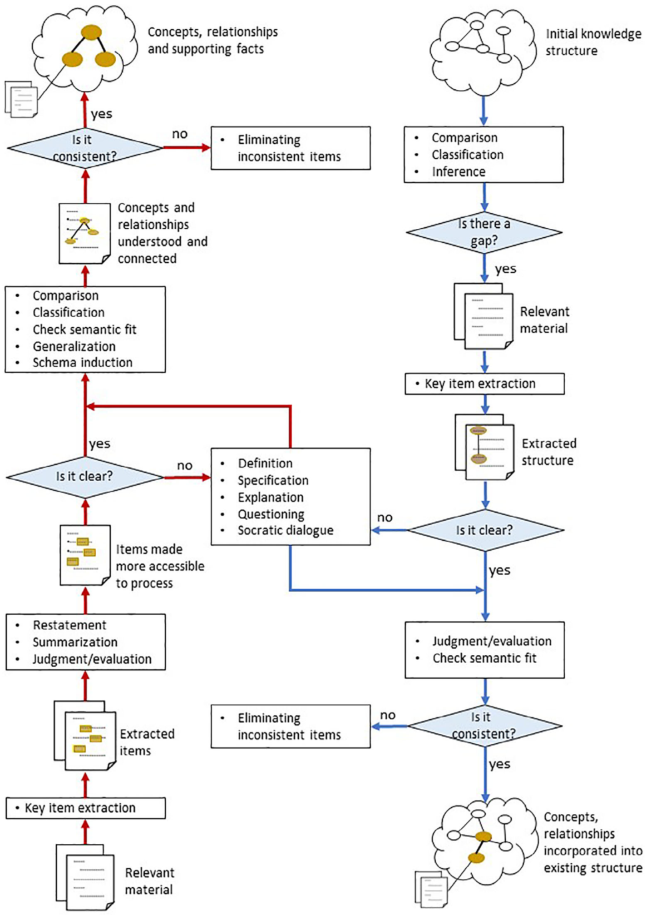


Figure 10: Prototypical activities when interpreting information [63].

3.3 Collaborative sensemaking

In the sensemaking research area, there is a distinction between work on individual sensemaking and collaborative sensemaking. The previous two sections have described sensemaking on an individual level. In this section, we will summarise the important studies on collaborative sensemaking.

Collaborative sensemaking goes paired with some challenges. For example, having diverse expertise among group members makes collaborating more complicated [39]. Another challenge is *teammate inaccuracy blindness*, this blindness is related to analysts falsely trusting other analysts' work, leading to less accurate results of a sensemaking process [20].

Three of the individual sensemaking studies described earlier have also been translated into team sensemaking. First, the Data/Frame Theory has also been analysed for teams [25]. The researchers describe that the general activities in individual sensemaking (e.g. *identifying a frame*, *elaborating a frame*, or *questioning a frame*), can also be translated to team sensemaking. They describe a set of strategies for each activity in the sensemaking model. For example, for teams, *identifying a frame* can be done by either collaboratively forming a frame or having a leader announce what the accepted frame is. The study ends with a call for future work on sensemaking in teams.

Second, the Learning Loop by Russell et al. has also been described in the context of team sensemaking.

They describe that the product from the learning loop (i.e. external representations) is only a subset of the mental understanding (i.e. an internal representation) of an individual sensemaker [43]. This creates an issue when sharing an external representation with others since the resulting shared understanding is a smaller subset of the shared representation; i.e. the group's understanding is smaller than all shared external representations combined. They describe that the goal of collaborative sensemaking should thus not be agreeing on a representation, but also on the meaning of that representation. This can be aided by collaboratively working on creating the representation.

Third, the cognitive sensemaking model by Zhang et al. [67] (see Figure 9), has also been compared to collaborative sensemaking. A study found that most of the activities in the cognitive sensemaking model are also present during collaboration [52]. Four main activities were found during collaborative sensemaking: (1) *Structuring the task* creates a deeper understanding of the tasks, and allows for the sub-dividing of tasks and organising information, these activities are comparable to the activities *building structure*, *identification of data gaps* and *identification of structure gaps* in the cognitive sensemaking model, where they form a loop; (2) *searching for information*, also for collaborative sensemaking this is an individual activity, where search queries are constructed and results are interpreted. The related activities from the cognitive sensemaking model are *search for data* and *search for structure*; (3) *sharing information*, where information is shared about the task topic (e.g. links and snippets for data or representations for structure), or information about the process. There are no activities in the individual cognitive sensemaking model that overlap with this collaborative activity. (4) *Synthesising Information* is the process of converting the information to a structured and meaningful representation, which can happen collaboratively. In the individual cognitive sensemaking model, this is the process of *Building structure* or *Instantiating process*. Collaborative sensemaking is the process of collaborating to produce an external representation as the result of the sensemaking task. Here individual representations can be combined, authors can collaboratively write in a shared document, or one person can synthesise shared information.

3.4 Information Systems and Information Visualisations

Sensemaking is largely an HCI discipline, making the computer systems that are involved also a subject of interest. Existing studies have described how sensemaking is paired with computer systems, highlighting challenges and opportunities. For sensemaking supporting information systems, three different workspaces can be distinguished [58]: (1) a *data Space* provides general information about the data available; (2) the *analysis Space* is where analysis techniques are executed on the data; (3) the *hypothesis space* is used to form frames and hypotheses and where they evaluate or elaborate on these. These spaces should ideally work together to create a *reasoning workspace* which promotes sensemaking.

3.4.1 Challenges with Information Systems in Sensemaking

When using computer systems to aid with the sensemaking of large pieces of information, two main challenges can be identified [58]. The *keyhole problem* is the issue where a sensemaker only sees a smaller subset of the to-be-analysed data due to a lack of screen real estate or computational constraints. This leads the sensemaker to create visualisations of an incomplete or out-of-date dataset. The *black hole problem* relates to issues with visualisation when it comes to missing data. It is hard to visualise missing data, or sensemakers cannot be aware that there is data missing. For example, it can be unclear whether an event did not happen or if there is no evidence of the event.

The possible influence of automated systems on cognitive sensemaking abilities has also been described. In a preceding article, the authors describe that computer systems can limit the sensemaking abilities of analysts [23]. They state that experts' analysis can suffer from hidden data, which could be the result of automatic data processing. They furthermore mention that it can negatively impact sensemaking if humans cannot comprehend how algorithms do data processing. Trust in the data processing from the analyst is also mentioned as an important factor.

In practice, it also appears that sensemaking systems are fragmented. Many different systems and methods are used in practice which do not interface with each other [21]. Many systems only aid with a subset of activities in Pirolli and Card's sensemaking model (see Figure 5) [40]. Since the workflow of analysts is non-linear, and analysts often switch between different activities it is also important that these systems can interface with each other to allow for easy switching between tasks.

3.4.2 Recommendations for Information Systems That Aid Sensemaking

Earlier studies provide some suggestions of features that could help in sensemaking. One of the answers to the *keyhole problem* is increasing screen real estate. It has been shown that using multiple large high-resolution monitors can aid with sensemaking [3]. It functions as a form of rapid access to external memory, and the physical position of information on the monitors can also be used to encode meaning and relationships [3].

Another feature that could promote sensemaking is a classification algorithm that can prioritise well-structured information [65]. The search for structure is essential in sensemaking, especially in the early phases. Having access to documents which structure information could therefore be beneficial for the process. Another recommendation made from the same research is the automatic extraction of entities and relationships from the raw data. This search for relationships is an important aspect of sensemaking, so automation could be beneficial.

For collaborative sensemaking, some straightforward recommendations have been made [39]. The creation of timelines of the sensemaking or investigative activities from other collaborators can be beneficial for other sensemakers to understand the process. Systems should also provide functionality for comments. These comments should be date stamped and the author should be recorded.

3.5 Intelligence Analysis

The previous sections described the works on the generalised process of sensemaking. In this section however, we will describe in more detail the work method of analysts during intelligence analysis in a policing context. This will be done by discussing an important similar study, which describes the analytical work method of intelligence analysts in detail. There are few papers which describe the analytical methods of analysis at the same level of detail. The paper in question described the intelligence process, without explicitly naming sensemaking, and is created by Chin et al. [7]. The study looked at how analysts handled the two different artificial scenarios created by the authors. One scenario focused on an individual's analysis methods while to other looked at how teams of information analysts made sense of information. Some different analysis strategies were found to be used by analysts: (1) *competing hypotheses*, is a strategy where an analyst creates a set of all possible hypotheses and maps data to support or disprove these hypotheses. (2) *Data-based analysis*, where no preconceived hypothesis is used to avoid any form of confirmation bias. (3) *Specific-factors* where an analyst looked at information from a set of specific factors (in this case access, intent, motivation, and capability).

It is concluded that intelligence analysis is indefinite and self-perpetuating, so if more information is analysed it becomes clear what information is missing regarding the current analysis. Eventually, the number of questions will reduce as more evidence is analysed until a coherent story is formed. Another important finding regarding their working method is time pressure. A report is never fully-complete so a time-constrained deadline forms the end of a project, where in the end their findings contain all the essential evidence that is needed. This time pressure also prevents intelligence analysts from having a systematic approach.

It is furthermore described how information is filtered in the first phase. All intelligence analysts were observed to print out the evidence even though they received all documents in a digital format. Some laid it out on the floor, some labelled them or created graphs to represent the found documents. Some also used physical folders to organise the document, later reproducing this sorting in digital folders. The printed-out documents were highlighted with highlighters pens, and some also noted down important facts in graphs or spreadsheets. On blank sheets of paper, analysts created annotated maps, crime timelines, and graphs.

It is worth noting that not every piece of information provided has the same credibility, so evaluating information credibility plays an important role. Most information is mistrusted by default until it could be confirmed by other sources. Information that is deemed irrelevant or unreliable is eliminated to limit the amount of information. Analysts also actively try to protect their own reputation, which is actively monitored by other analysts in order to evaluate the credibility of an analyst's report. It is also subjective which information is deemed relevant for the case, in their research they found that different analysts focused on different parts of the evidence, this could be dependent on how much background knowledge an analyst has on the subject of the evidence.

Pattern recognition is another important part of this analysis. Pattern recognition is done by discovering concepts, and finding similar or orthogonal ones. Pattern matching heavily relies on existing knowledge and experience. For example, an experienced analyst knew that criminals often committed crimes in their own proximity, leading the analysts to further look into local criminals.

Lastly, when it comes to collaborative sensemaking, its prevalence was limited in this study. Analysts tended to work alone since trust in other analysts was low by default. However, the information analysts did view collaborative analysis as useful and necessary. The study found an increase in effectiveness at resolving discrepancies in data by working as a team, compared to working individually.

3.6 Conclusion

A summary of the discussed related studies will allow us to answer **RQ1**: What does earlier research describe about the sensemaking process? This in turn helps us to get an understanding of what we can expect from the sensemaking process in our case study. The results from this review are later used for the comparison to theory after the generalisation provided by the cross-case analysis from the case study.

First of all, sensemaking is a non-linear ad hoc process, meaning that the activities of an analyst will not follow a straightforward sequence [21, 28, 63, 7]. However, Pirolli and Card's model can be used to describe the transformation of information through the process (see Figure 5) [40]. One important data structure is *schemas* (i.e. *representations*), which are a way of representing filtered data in a structured manner that can be used to more easily draw conclusions [47]. The process of creating representations and drawing conclusions from them is the *sensemaking loop*. The process of filtering, extraction and transforming is data the *foraging loop*. This loop contains a shift from *exploratory search* to *exploitation* of the relevant documents. For the actual workflow of intelligence analysts, we can expect some to only work on some parts of the sensemaking process model (e.g. only the foraging loop) [19].

Second, the studies on cognitive sensemaking allow us to conclude that hypothesis formulation and testing are important processes in an analyst's mind [19]. Earlier knowledge is an important factor in this process [4, 7]. In order to get an understanding of the data the analyst iteratively develops and tests a frame of the data [26, 24]. Even though these activities are cognitive, different patterns of usage with computer systems can be observed depending on which activity in the Data/Frame Theory the analyst is performing [27]. Another concept found in the reviewed studies is a knowledge gap that can lead to more data being analysed, or to more data being searched. A distinction can be made between a data-gap and a structure-gap, where the latter includes relations and patterns in the found data [66]. Exploratory search is a separate process, which can not only help with finding data but also with finding a structure that can be used to elaborate on the current understanding [42, 65].

Third, sensemaking can be an individual or collaborative effort. Three of the individual sensemaking models translate to collaborative sensemaking (Data/Frame Theory [25], Learning Loop [43], and the Cognitive Sensemaking Model [52]). One challenge is *team inaccuracy blindness*, where analysts mistakenly trust other analysts [20]. The contrary has also been found where it was found that analysts mistrust others by default during intelligence analysis, limiting collaboration [7]. Another challenge with collaborative sensemaking is different analysts having diverse expertise [39].

Lastly, we can expect systems to be used by the sensemakers to be fragmented [21]. The systems can be separated into three separate workspaces: *data space*, *analysis space*, *hypothesis space* [58]. Earlier research also highlighted problems that can occur with systems related to sensemaking, such as: not being able to view all available information (*keyhole problem*) and improperly representing missing data (*blackhole problem*) [58].

4 Findings

In this section, the findings of the case study will be presented. For the case study, a total of 17 interviews have been held in order to gain insight into the intelligence operations of the Dutch Police. A full list of interviewees can be found in table 4. In order to get a better view of intelligence analysis itself and how it's positioned in the organisation we have interviewed employees who are themselves analysts, team leaders, and intelligence employees who work closely together with analysts.

A total of three cases can be distinguished, namely: (case I) **Investigation Analysis**, where the analyst supports a criminal investigation; (case II) **security analysis**, where the analysis happens looks into a criminal phenomenon, results of the analysis can be used prioritise investigations; and (case III) **strategic analysis**, which can be paired with scientific research, the results of these analyses can contribute to determining strategy or to directing policing efforts. Each interview is assigned one or more cases.

All three cases which are part of this case study operate in the same shared environment. This section therefore starts by introducing the shared context of the Dutch Police, after which the applications that are used by most analysts will be introduced. A within-case analysis follows for each of the three cases. These sections will describe the working method of that type of analyst. The cross-case analysis section after that combines and compares the findings from the cases. In this last section we will also list the enablers of sensemaking that are found throughout the case study.

Identifier	Unit	Role	Case
1	DLIO	Case Analyst	Investigation Analysis
2	DRIO	Information Coordinator TGO	Investigation Analysis
3	DLIO	Case Analyst	Investigation Analysis
4	DRIO	Tactical Analyst TGO	Investigation Analysis
5	DRIO	Case Analyst TGO	Investigation Analysis
6	DRIO	Team leader	Investigation Analysis
7	DRIO	ICT Administrator	Investigation Analysis
8	DRIO	Former Team Leader Analysis and Research, currently Product Owner	Investigation Analysis and Security Analysis
9	DLIO	Former Security Analyst, currently Product Owner	Security Analysis
10	DLIO	Security Analyst	Security Analysis
11	DRIO	Security Analyst	Security Analysis
12	DRIO	Tactical and Strategic Analyst	Security Analysis
13	DRIO	Senior Intelligence	Security Analysis
14	DRIO	Manager Implementation New Analysis Workflow	Security Analysis
15	DRIO	Security Analyst	Security Analyst
16	DLIO	Researcher	Strategic Analysis
17	DRIO	Researcher and Innovation Manager	Strategic Analysis

Table 1: Interview participants with their identifier, unit, role, and the cases which use the data from the interview.

4.1 Context

Analysis happens at different levels within the Police organisation. The Dutch Police consists of ten regional units and one national unit. The departments of interest of the national unit for this research are the DLIO (NL: Dienst Landelijke Informatie Organisatie, EN: National Information Unit), and the DLR (NL: Dienst Landelijke Recherche, EN: National Investigation Unit). These have a regional equivalent of DRIO

(NL: Dienst Regionale Informatie Organisatie, EN: Regional Information Unit), and DRR (NL: Dienst Regionale Recherche, EN: Regional Investigation Unit). The DLIO and DRIO both have a sub-department called A&O (NL: Analyse en Onderzoek, EN: Analysis and Research). Strategic and security analysis takes place within these departments of the organisation. Security analysis typically happens in groups that focus on one theme. The DLR and DRR focus on criminal investigations and concentrate on individual cases. The DLR and DRR consist of multiple teams that investigate individual cases, the DLIO or DRIO supplies analysts to these departments. A simplified organogram is displayed in Figure 11, where the departments where analysis is studied are highlighted in blue.

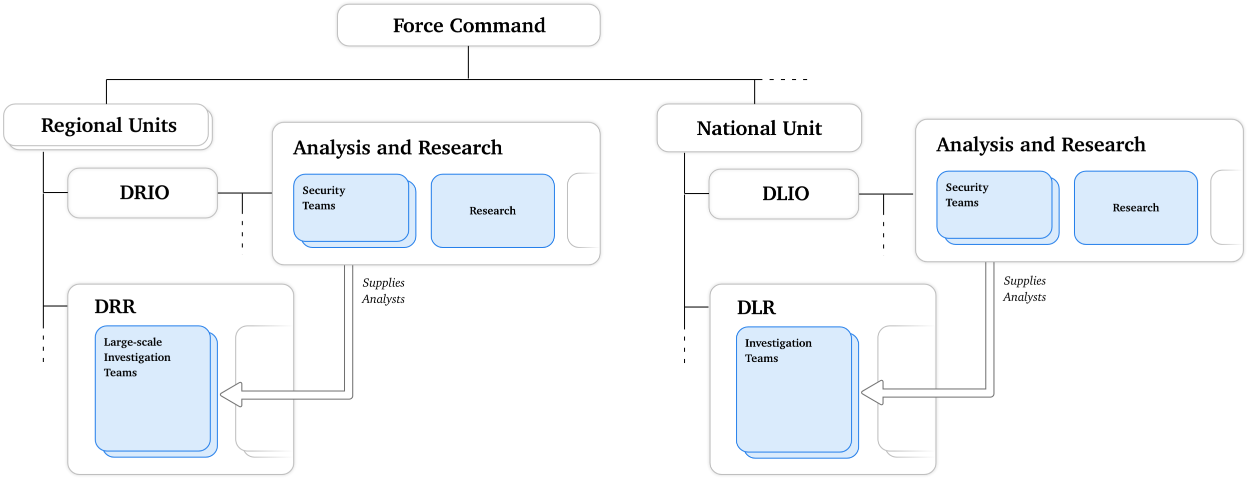


Figure 11: A diagram of the departments where analysis operations are investigated.

4.1.1 Applications

The Dutch Police has many systems available for diverse use cases. As one analyst describes: *"There are a lot of systems, and new ones are created constantly"* (IV05). Even though there is a wide range of systems, there is a set of primary systems that almost all analysts use on a daily basis. Other systems have specialised use-cases. In this section, we will describe the systems the analysts that were interviewed used the most:

SummIT is an important system in investigation teams, which is used by all relevant actors in an investigation such as the detectives and the analyst. The application functions as a journaling system, containing entries (also called mutations) about findings from the active investigation. The main data in SummIT is findings from detectives, which can be written as official reports. An analyst can also add their findings as an entry to this system. For each investigation a new workspace is created. These entries are mostly text-based but can contain files. The entries can also be coupled to entities, *"Everything that is being recorded in SummIT, is also linked to entities such as persons, addresses, and phone numbers"*. (IV01).

Analyst Notebook is a tool that allows users to create entities and relationships visually. The program also allows for the creation of timelines. Analysts often work with their own environment and do not share the created schema directly. Users who are not an analyst typically do not use Analyst Notebook. The application also allows for a visual way to query the schemas that an analyst has built.

iBase is the program which contains the database behind the program Analyst Notebook. Everything that is added to the schema from Analyst Notebook is added to the database in iBase. The application allows analysts to populate and query the database. One project manager explains the functions of iBase: *"iBase is essentially a SQL Database, but with a very simple interface. The idea is that you can create and edit properties* *of entities. [. . .] You can link people together, which Analyst Notebook can visualise, and show in a neat graph."* (IV09).

BVH contains all the reports made by police officers that work on the street. Incidents that agents report end up in this system. Official reports (NL: Process Verbalen), and charges pressed by citizens (NL: aangiften) are also included in this system. Typically detectives nor analysts input data into these systems directly. However for analysts this system can be an important source of information.

BlueView is a search engine and that can find more information about persons, addresses, cars, or other entities. The search engine can look in to different sub-systems used by the police. In addition to looking through BVH reports it can also look through entries made in SummIT workspaces from earlier investigations. Other systems from partners are also included like the FIOD (NL: Fiscale inlichtingen- en opsporingsdienst, EN: Fiscal Information and Investigation Service) and KMAR (NL: Koninklijke Marechaussee, EN: Royal Military Police).

Information from BlueView can be imported into an iBase database through the tool BlueBase. As one analyst (IV09) describes: *"You were never able to directly import it to iBase so somebody made a Microsoft Access tool, to completely reformat the csv files so it would fit into iBase. That was quite the intermediate step that was built in between the systems and was called BlueBase"*

Refinery (NL: Raffinaderij) is an integrated analysis environment that allows for easier analysis of data from a wide range of sources. *"The refinery is built to aid analysis within an operational context with highly modern tools"* (IV09). The system allows for analyst to easily import data, or automatically loads relevant data. "*. . . pretty much all police systems are coupled to it, which you can easily access. But also, trackers placed under cars, tapped phone lines, or requested <mobile traffic histories logs> . . . '* (IV01). The main use case of the application is to visualise, analyse, and search combined information. It does not store any data directly. It also allows for geo-based analysis, by displaying objects such as trackers on a map. These geo-based visualisations can be combined with a temporal dimension so moving objects can be displayed over time. It also has functionalities which overlap with Analyst Notebook, such as the creation of timelines and relation schemas. One analyst (IV03) states: *"In fact, it replaces all systems in one, I find that more convenient"*.

4.2 Within-Case Analysis

In this section, the findings from the three cases from the case study are presented. In each section we will describe the working method of the analysts, the team they interact with and what applications they work with. These findings will answer our **RQ2**: How is intelligence analysis conducted at the Dutch Police?

4.2.1 Case I: Investigation Analysis Process

The first case that is presented is investigation analysis. This concerns analysis that is directly linked to prosecuting crime. The investigation is targeted directly towards an individual or a group of individuals, that should be found and stopped or arrested.

Team Structure
Investigation analysts are attached to a team of detectives. Within investigations in the National Unit at the DLR there is only one analyst in a team of detectives, this analyst usually has the title of *Senior Intelligence* or *Operational Specialist A*. Topics of investigation at the National Unit are usually more prominent than those covered by the regional unit. IV4 explains characteristics of the cases: *"<cases at the National Unit> are always serious criminality, but also always directed at criminal associations, and to work towards the leaders at the ends"* (IV03). The role of the analyst here is to provide insights and overviews of the investigation.

The Regional Investigation Unit (DRR) can have multiple types of investigations, for this research, we looked at the larger subtype, called TGOs (NL: Team grootschalige opsporing, EN: large-scale investigation team). TGO teams contain about 15 detectives. For these teams, there are two types of analysts attached, one tactical analyst and one operational analyst. The operational analyst, also referred to their job title as *Senior*

Intelligence, has a comparable role to the one at the DLR. The tactical analyst has the role of creating and managing hypotheses and scenarios. The official job title of the tactical analyst is *Operational Specialist A*.

Activities

A model of the working method of analysts is displayed in Figure 12. This model shows how the analyst interacts with the team, and what products they provide. In this section we will walk through the activities found in the working method of the analyst.

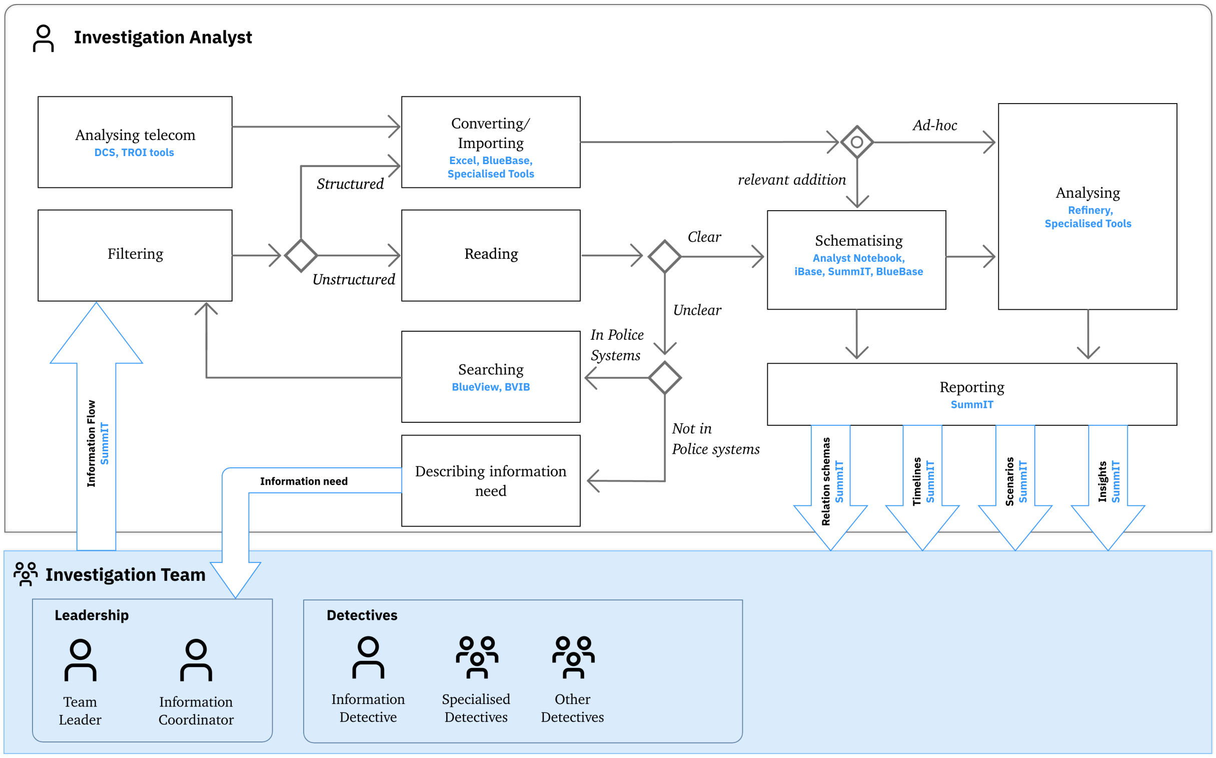


Figure 12: A model of the investigation analysis at the National Unit of the Dutch Police. Squares indicate activities, diamonds decisions, and arrows information flows. It is recommended to start reading from the information flow arrow on the left.

One analyst describes that cases can have multiple phases where their activities shift over time: *"At the start it's actually all unclear to everyone. We start based on an official report from the TCI [Team Criminal Intelligence]. They say <Example Name> is importing large amounts of cocaine. At that point you know not much, that's the only thing you know. Then you are going to check, who is <Example Name>? Who are his friends? What is he doing? What car does he drive? What phone does he have? You are investigating very broadly, to look who he is. As an analyst, you have a pretty big role in creating an overview of that information."* (IV16)

To provide this overview of information to the team the information needs to be analysed. This starts with **reading** information that is in the SummIT system, which contains textual information about the case such as reports and statements. One manager explains the task: *"You will have to go through all texts if you want to extract the information. That takes the most time. And that also applies to telephone calls and the like."* (IV02).

While the analyst is reading the information they **filter** information and decide whether or not to **schematise** the information and store this into a schema. The schema is typically created and managed in the application Analyst Notebook and the linked database in iBase. An analyst can automatically add entities and their relationships that are included in SummIT by **converting** them. Other entities and relationships are added manually to Analyst Notebook. Here an analyst filters which information is relevant, and includes a subselection in the schema. One analyst (IV02) described: *"..., and especially reducing it. Sometimes you already have quite a lot of data, which is one big mess, and then you just have to follow the aim of the investigation and try to refine it"*.

In some cases the team leader can also ask the analyst to **analyse telecom**. In this process the analysts looks at the raw data that is collected related to telephony, and performs analysis on this information. The telephone data can then be converted and optionally added to the schema. One analysts describes what kind of questions new data from an IMSI-catcher can bring up: "*we will look in to what phones they are using. We had an idea which kinds of phones they were using, because four phones were observed. Later, i will receive the data from the catcher, and I have to look at which phones were intercepted, and which ones were used by the subject.*" (IV03)

The activity of schematising not only includes the building of a relational schema but also that of a timeline. The latter is often created in Excel. Hence the activity also includes putting events with corresponding dates and times in Excel. This timeline can also be constructed in Analyst Notebook, but all operational analysts interviewed prefer Excel for its ability to smoothly edit, filter, and sort the timeline. One analyst explains why they prefer Excel: *"You can also create a timeline in Analyst Notebook, it then contains all kinds of events which are coupled to each other. That all looks very pretty, but your timeline can get very long when you are looking at a longer timeframe. If you print it, you have a whole wall full with the timeline, that just is not practical to work with"* (IV01). Timelines typically include (but are not limited to): sent messages, phone calls, location history, information from CCTV footage, and official statements. Telecom data is often provided in a structured format, and can be converted to be added to the timeline. This telecom information usually forms the skeleton of the timeline.

Analysts also can have the task of looking further into the entities that are described in SummIT entries. For example, a SummIT entry can contain the names of other persons. An analyst can **search** for information about that other person in BlueView or a similar program called BVIB. This information can then also be included in the schema. Throughout the investigation, an analyst can also **describe information needs** if they find that information is missing in the investigation. In this case an information coordinator can make sure more information is collected.

Analysts have read most if not all information that has been collected for the investigation. Using the explicit knowledge in the created schema, and the implicit knowledge built up in the process, the analyst can also generate hypotheses or scenarios and provide them to the team. These hypotheses and scenarios can be a small part of the whole investigation, for example stating whether or not a subject is likely to have been at location x, or have been inside car y at time z. Hypothesis generation is not an explicit task of an analyst but is done by most analysts that have been interviewed. One analyst also describes this more as informal activity, in the sense that they do not systematically produce them, but rather discuss them informally with the team. The fact that the analysts has read through most information in the investigation, also means the analyst can answer questions from the detectives in an informal setting.

During the process further **analysis** occurs on the data that is available to the analyst. This is the most ad hoc activity in the process, in the sense that it does not follow a standard method. The product of the analysis can be general insights about the investigation, or specific answers to questions asked by leadership. Refinery is the software that can help with these different types of analysis, by combining the information that has been collected in the investigation. One analysts gives the example of ad hoc analysis: "*At some point the question was asked, what time of day can we best deploy an observation team? For that, I made a graph showing which days had the most criminal activity. So we found that is was best to deploy on a Wednesday, because that is when he is most criminally active. Those kinds of things are part of the investigation.*" (IV03)

The products of the process are diverse. The timelines and relational schemas, are products which are maintained throughout the investigation. Scenarios and insights are more ad hoc products. All products can be created by a question or assignment from the leadership, or be initiated by the analyst themself. These findings can be shared by the analyst by *reporting* a documented finding, which can be shared in SummIT. Note that analysts are allowed to make assumptions to propose hypothetical scenarios, while in detectives have to be strictly factual in other official reports. These insights can also be shared in a more informal way, for example in the morning briefings.

In a later phase of the investigation, the information flow continues. An analyst monitors this information flow from SummIT and continues to work on expanding the schema and timeline. More information that requires analysis can also come directly from other persons. One analyst (IV03) calls this an *"Ad Hoc Phase"* and describes it as follows: *"I once had an analyst in training sitting next to me, and <they> said, 'Isn't this driving you crazy? You are trying to work, but the one after the other comes to your desk with questions'. And that is, in my opinion, the job of an analyst. You are the all-knowing person inside the team where they all come and ask 'Can you look into this for me?' "*

Later in the investigation, there is also more time to look into things more broadly. Analysts can look into the systems, try to get more insights regarding incoming information from SummIT and try to, for example, back the incoming information based on earlier analysed data.

In the last phase, the subject or subjects are arrested. A file is built against them, which is where an analyst can help, but this help is limited since most do not write official reports. An analyst can help with additional ad hoc products such as interrogation plans. This last phase is usually less time-intensive for an analyst. Analysts can work on reflecting and evaluating the investigation, or on building an overview of things that came by in the investigation but are not included in the prosecution.

Investigation analysis at a regional level

Overall the activities of an analyst in TGOs at the regional units are similar. However, these teams also have a dedicated tactical analyst who is concerned with creating and managing hypotheses. The tactical analyst works closely together with the leadership of the investigation team, called VKL (EN: Fixed Core Management Team, NL: Vaste Kern Leidinggevende) and can help with directing the inspective actions within an investigation. One tactical analyst describes their role as: *"Your role is to keep an overview of the investigation, and to really think based on hypotheses and scenarios, and to also be the critical note within the VKL, and to prevent tunnel vision"* (IV04). The processes of investigation analysis at a regional level are shown in Figure 13.

The process of creating hypotheses & scenarios is aided by a mind-mapping application called MindManager. This application essentially allows for the creation of a deep tree where nodes can have properties. The mind-map starts based on a template that is uniform across the police organisation. There are templates for different types of crimes. These templates contain an initial set of hypotheses. In the case of murder, these are natural death, lethal accident, suicide, or murder. Information from SummIT, containing collected information for the investigation, is read and is included in the mind-map as supporting or disproving facts for a particular scenario. This linking process is called **Adding and subtracting** (NL: Plussen en minnen). The mind-map is further expanded with more sub-hypotheses, which are again linked with evidence. The mind-map functions as a local database, where information can be stored in a structured manner. Therefore, the template also contains some subtopics that should be filled in by the analyst, such as a rough timeline of events. An example of such a mind-map is shown in Figure 14.

In the process of evaluating hypotheses the assessment of the trustworthiness of the information is also important. An example given is that the quantity of anonymous crime reports does not entail that the scenario is more likely to have happened, since the anonymous reports can be made by one person. The same holds for reports resulting from interrogations, which can contain false information.

Actually searching for information in police systems is an activity that does not occur often for tactical analysis. A tactical analyst can **Describe an information need** which the investigation team should look into. This information need can either be solved by searching in the Police systems, or can require the investigators to retrieve more information. While in principle the tactical analyst does not have to search for information themselves, in practice the analyst does have access to the internal search engines, and they sometimes also use them to look up information themselves. This can be a more uncomplicated and quicker way to obtain information.

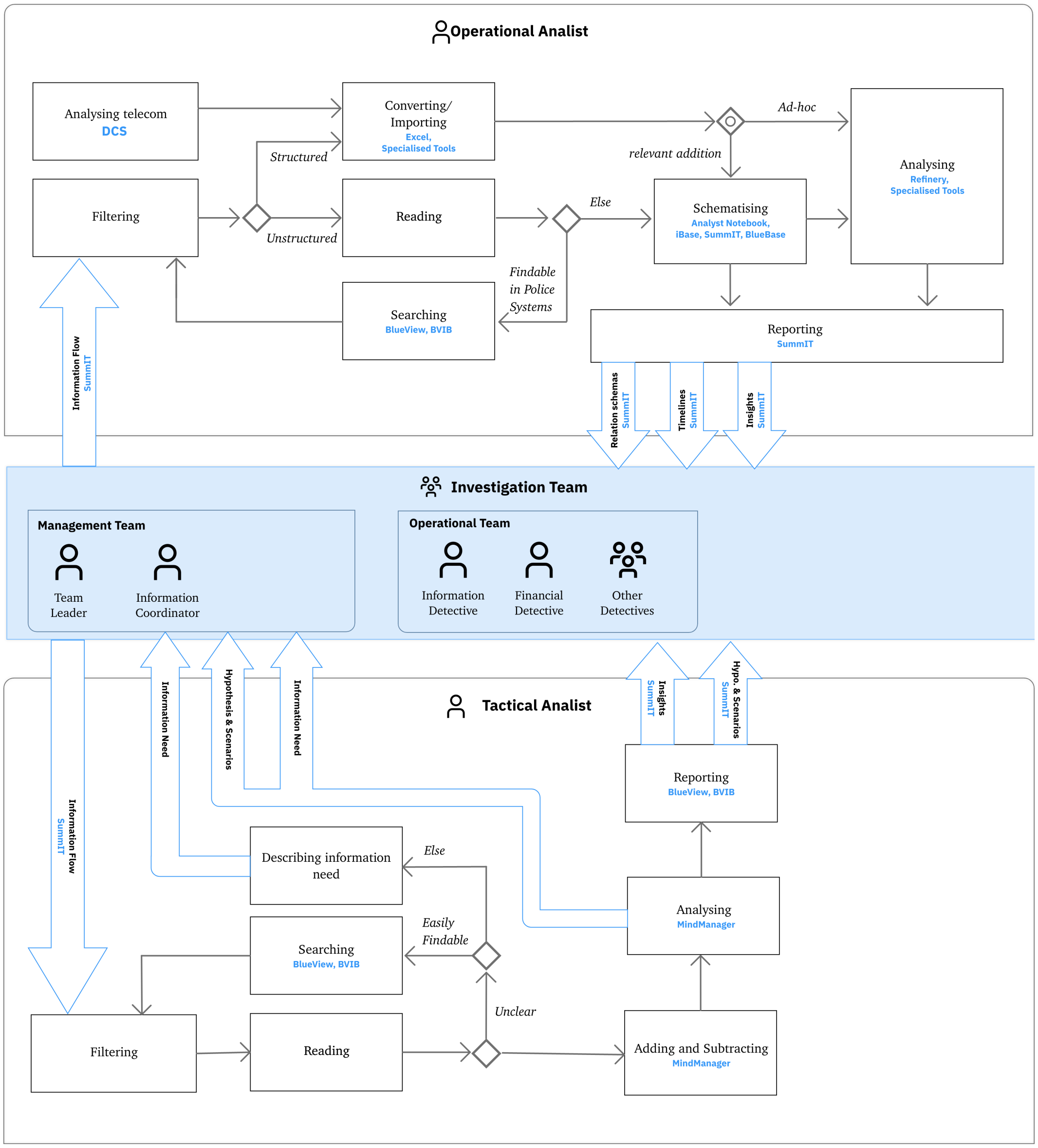


Figure 13: A model of the investigation analysis at TGOs at the regional units of the Dutch Police. Squares indicate activities, diamonds decisions, and arrows information flows. It is recommended to start reading from the information flow arrows on the left.

The interviewees described that collaboration between the tactical and operational analysts is uncommon and is also dependent on the persons who fulfil the roles. The operational analyst is also positioned within the operational team rather than the management team.

Not included in the diagram is that tactical analysts are also instructed to keep an information collection plan, where they track the questions or information needs they find, and monitor their status (e.g. if they are open or have received answers already). This plan is created in Excel. Not all questions are assigned to detectives, and some are only for the analyst themself. One analyst describes the benefit of revisiting unresolved questions in light of newly emerged knowledge during the investigation.

All analysts and coordinators on a regional level and national level that were interviewed for this case described that their task descriptions are clear. In the sense that they know what analysis products the team leader expects, and how to create these products.

4.2.2 Case II: Security Analysis Process

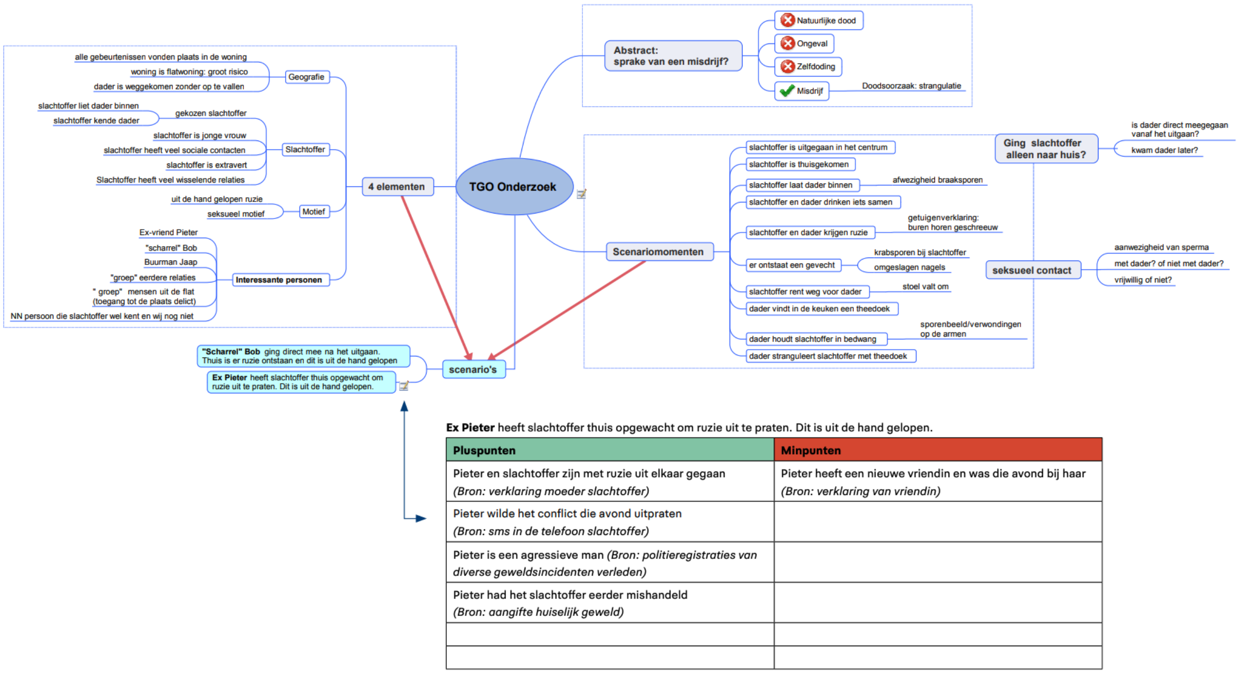
The second type of intelligence analysis is security analysis, which happens on a higher level. While investigation analysis is concerned with a specific investigation into one or more subjects, the security teams look at criminal phenomena. These teams exist at a regional and national level, correspondingly scoping their analysis. Each team consists of an analyst, intelligence employees, and sometimes detectives. The teams are focused on a specific theme. Examples of such themes are synthetic drugs, cocaine, human trafficking, criminal organisation structures and CTER (Counter-Terrorism, -Extremism and -Radicalisation).

These teams monitor the crime within high-impact themes. The general goal of these teams is to provide insights that help direct policing efforts. The result of security analysis can support starting a new criminal investigation or current investigations. Analysis can also be directed to provide an overview of the prevalence of that phenomenon or an overview of major persons active in that phenomenon. The analysis can also take place with the goal of answering a question from leadership who requests specific information about a phenomenon.

The processes of security analysis differ between regions and themes, but in general, we can distinguish between two types of processes in security analysis, the monitoring process and the analysis process.

Monitoring process
The first is the monitoring process, which is displayed in Figure 15. This process consists of monitoring information that is brought up by an automatic reporting tool. This tool differs per theme, but COGNOS and Excel were most often mentioned. It also differs per team who is monitoring this information. Within some themes, the monitoring is completely handled by other intelligence employees, while in other teams it is also the task of the analyst. The incoming information is then **read** and reviewed to check if it is relevant. It is then **filtered**, if relevant but some specifics are unclear the employee can look up more information in the police systems. An analyst might for example **search** for more details about the subject of interest. Sometimes it is also useful to **contact the source**, such as the police agents on the streets who encountered the situation. Lastly, the information is **structured** and stored in SummIT. This structuring activity is comparable to a schematisation step, however, the goal of this step is not to create a schema as a product, but to structurally store the data so that it can be more easily used for further analysis. A security analyst describes how they monitor this information: "*Me, as an analyst, also keeps up our workflow, which means that I am also monitoring the incoming reports in SummIT. [...] This includes a report from a neighbourhood police officer about a small street dealer who is arrested with 300 grams of cocaine under the saddle of his scooter, as well as a lab that exploded on some farm and were a person got killed. [...] We look through them and decide on which are interesting for us*" (IV10)

Most security teams use automatic scripts that provide relevant theme-specific information. This information originates either or both from relevant ongoing crime investigations reported in SummIT, or reports filed by police officers on the street which are stored in BVH. This automatic system is maintained by another department and aims to have high recall and minimise false positives. This is an active process between the


TGO Onderzoek
Abstract:
sprake van een misdrijf?
Natuurlijke dood
Ongeval
Zelfdoding
Misdrijf
Doodsoorzaak: strangulatie
4 elementen
Geografie
alle gebeurtenissen vonden plaats in de woning
woning is flatwoning: groot risico
dader is weggekomen zonder op te vallen
Slachtoffer
gekozen slachtoffer
slachtoffer liet dader binnen
slachtoffer kende dader
slachtoffer is jonge vrouw
slachtoffer heeft veel sociale contacten
slachtoffer is extravert
Slachtoffer heeft veel wisselende relaties
Motief
uit de hand gelopen ruzie
seksueel motief
Interessante personen
Ex-vriend Pieter
"scharrel" Bob
Buurman Jaap
"groep" eerdere relaties
" groep" mensen uit de flat
(toegang tot de plaats delict)
NN persoon die slachtoffer wel kent en wij nog niet
Scenariomomenten
slachtoffer is uitgegaan in het centrum
slachtoffer is thuisgekomen
slachtoffer laat dader binnen
afwezigheid braaksporen
slachtoffer en dader drinken iets samen
slachtoffer en dader krijgen ruzie
getuigenverklaring:
buren horen geschreeuw
er ontstaat een gevecht
krabsporen bij slachtoffer
omgeslagen nagels
slachtoffer rent weg voor dader
stoel valt om
dader vindt in de keuken een theedoek
dader houdt slachtoffer in bedwang
sporenbeeld/verwondingen
op de armen
dader stranguleert slachtoffer met theedoek
Ging slachtoffer
alleen naar huis?
is dader direct meegegaan
vanaf het uitgaan?
kwam dader later?
seksueel contact
aanwezigheid van sperma
met dader? of niet met dader?
vrijwillig of niet?
scenario's

"Scharrel" Bob ging direct mee na het uitgaan. Thuis is er ruzie ontstaan en dit is uit de hand gelopen

Ex Pieter heeft slachtoffer thuis opgewacht om ruzie uit te praten. Dit is uit de hand gelopen.

Ex Pieter heeft slachtoffer thuis opgewacht om ruzie uit te praten. Dit is uit de hand gelopen.

Pluspunten	Minpunten
Pieter en slachtoffer zijn met ruzie uit elkaar gegaan *(Bron: verklaring moeder slachtoffer)*	Pieter heeft een nieuwe vriendin en was die avond bij haar *(Bron: verklaring van vriendin)*
Pieter wilde het conflict die avond uitpraten *(Bron: sms in de telefoon slachtoffer)*	
Pieter is een agressieve man *(Bron: politieregistraties van diverse geweldsincidenten verleden)*	
Pieter had het slachtoffer eerder mishandeld *(Bron: aangifte huiselijk geweld)*	

Figure 14: And example of a mind-map created by a tactical analyst in a TGO.

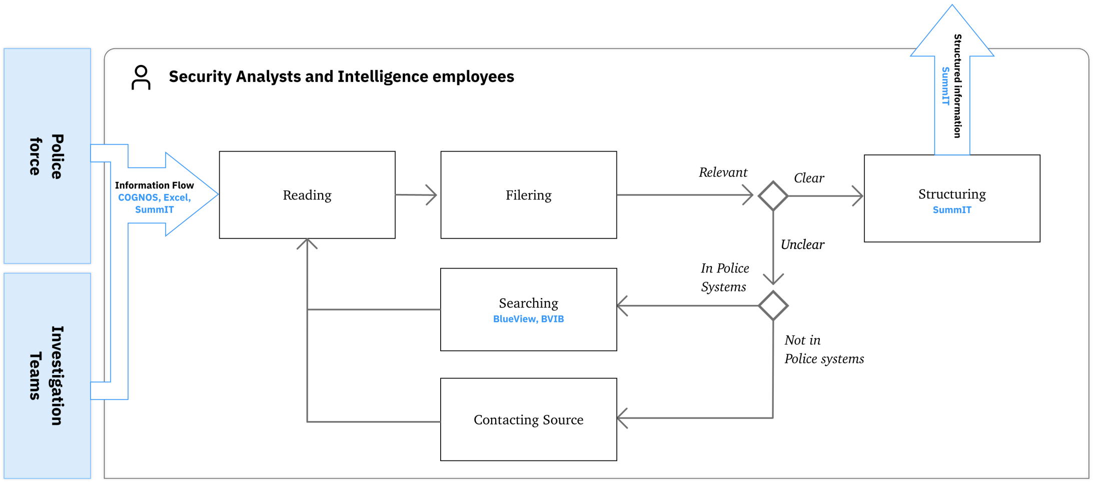


Figure 15: A model of the the monitoring process in security analysis. Squares indicate activities, diamonds decisions, and arrows information flows.

maintainer and the people monitoring the information flow. These information feeds are largely executed by queries containing word lists with complex logic with negation, AND, and OR operations. In addition to the automatic information flow, for some themes, police officers can directly link a report to a specific theme, so the security analysts can look at the incident. One manager describes how the automated search can help find hidden types of crime: "*We have also developed a proactive scan, so we can proactively query al our systems on a daily basis were we check for possible signals of human trafficking. We found that human trafficking can be hidden in many other situations, where it is not always evident that it concerns human trafficking... We found that many signals of human trafficking are really hidden in our own systems*" (IV14)

Some thematic teams have described their monitoring process in detail. Here they have clearly defined how they handle incoming reports, and what to do with them. For example, the human trafficking teams also have scoring forms that scores how prevalent signals of human trafficking are in a report. This form then indicates if further action is needed. The results of the form are also stored in a systematic way to allow for analysis later on.

Teams may have specialised methods for structuring the information that is monitored. The structuring methods are created with the intention of fighting relevant criminal organisations as effectively as possible. One example of such a method is *crimescripting*. Crimescript is a method of labelling the different actors in a criminal organisation. Crimescripts have been developed for different themes. The creators of crimescripting found that, for example, the process of creating synthetic drugs is, to a certain extent, uniform across different criminal organisations. All synthetic drug producers for example need base ingredients, transport of the base ingredients to the lab, storage of these ingredients, a laboratory operator, and a stash location. Crimescripting contains sets of roles that exist in the different phases of creating the synthetic drugs. A security analyst gives an example of the assignment of such role: "*so if somebody was cooking with a kettle warning a gas mask, we call them a lab technician, and they are in phase five, the production phase*". This assignment allows for a more complex social network analysis. Combining the network with the results of earlier interventions, the targets can be found that have the most potential to disrupt the criminal organisation.

The information that is structured is stored in the system SummIT. SummIT was originally created to help operational crime investigations, however, most thematic security teams now also use this system to log relevant mutations. SummIT can act as a source system, meaning that the analysis programs interface with it. Complex theme-specific input guides have thus been set, to store as much structured information as possible in the system. An example of such a rule is how to format the subject line of a mutation, to make it include information in a structured manner. It also differs per theme to what extent this SummIT environment is shared with other units. Some themes have a SummIT environment per region, where the national unit looks at all eleven, while others work with one shared one across all regions. The latter does require a strict input specification to keep information organised.

Analysis process

The second relevant process in security analysis is the analysis process which is displayed in Figure 16. This process contains little schematisation because, in theory, the monitoring process already includes storing the data in a structured fashion. This structured information from the monitoring process is available in SummIT and the analysts iBase. This is usually done through an automated import from the information in SummIT. In that case, the relevant information for the analysis can be **queried**. One analysts describes that a small network analysis is relatively straightforward because of this: "*If we created a good basis, and collected and structured information already, well than of course it doesn't take that long. But it is also dependant of how in-depth the analysis needs to be. If it is a flat analysis looking for leader in an organisation, and you want to know what the relations are, and what they are currently doing, these kind of things, well that doesn't have to take longer than half a day to get the first insights.*" (IV15)

Analysis can start with questions from leadership, by interest of the analysts themselves, or by needing to do planned tasks. When the tasks are given by leadership, further **clarification** might be needed on the specifics of the analysis product. Some of the analyst's task are complex, and requires more information than that is available in their iBase environment, so **exploiting other data sources** can also be part of the process. Later,

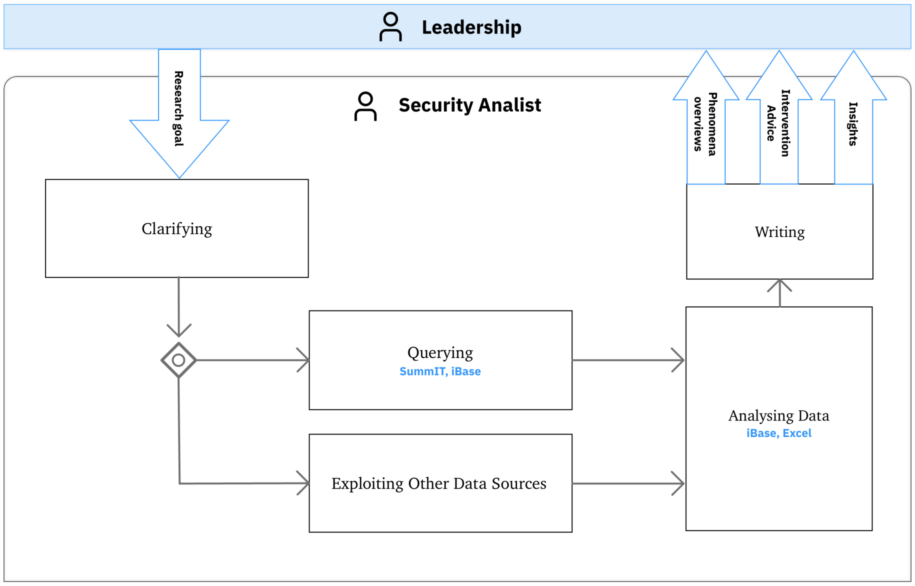


Figure 16: A model of the the analysis process in security analysis. Squares indicate activities, diamonds decisions, and arrows information flows.

the (combined) **data is analysed** in order the make conclusions about the analyst's task. Lastly, the findings are **written** into a report.

Products of security analysis differ widely. Sometimes products can be relatively straightforward and easy to execute. Other products can be more complex and include details about the specifics of the analysed crimes. For example, for a phenomena overview about synthetic drugs, the report includes findings about the sources of the chemicals used, which criminal organisations are active, which chemical processes are used, and if the markets are intertwined with other criminal markets. Analysts can also work on advice for the fighting of the type of crime they are analysing, a security analyst mentions: "*For example, if there is a problem with break-ins in a neighbourhood, well then there will be a security analyst who will look at how can we best approach this? They can investigate the problem, and think of scenarios in order to find the best ways to tackle the problem. And then to actually make a proposal for this.*" (IV10)

The analysis process can differ, and for some complex analysis tasks the information available in the theme's SummIT environment does not suffice, because the information that is provided through the monitoring workflow does not cover the matter at hand. An example of this is an analysis within the CTER theme that looked at the upcoming sovereign-movement. This movement involves people who are of the opinion that they are not a citizen of the country, and think they are not required to adhere to the local laws. This analysis requires domain knowledge, as one analyst describes: "*You have to read up on other sources, so absolutely not only the police systems. You have to understand where the movement originated and why.*" (IV12) This example illustrates the challenges of discovering new trends among movements or modi operandi. Officers who have encounters with people following the sovereign moment are not classified as such because they are either not recognised or because there is no option to classify it in a certain predefined list. This creates the problem that these new developments are not automatically shown to the right analyst. Therefore, the media still plays an important role in discovering such new trends. In this example the analyst needed to be creative with the search query to find relevant incidents in police systems, she describes: "*sovereign people tend to say: 'I am a human of meat and blood', 'I am not a person' or 'My name is from the family', so they don't have a last name or first name. You would just have to try these words as queries, so new cases come forward*" (IV12)

4.2.3 Case III: Strategic Analysis Process

Another form of analysis that happens at the Dutch Police is strategic analysis. This strategic analysis can also be of the form of scientific research. This type of analysis is often executed by criminologists and occurs in multiple units. The national DLIO and ten regional DRIOs all have scientific researchers. Strategic analysis and scientific research have a lot overlap in the Police, but not all strategic analysis is executed by scientific researchers, in this section it is therefore labeled as strategic analysis. Comparable to security analysis, strategic analysts are subdivided into themes on which they focus. Examples of these themes are: organised crime, terrorism, and motor gangs. Strategic analysis provides leadership with insights into the current criminal environment. The process of strategic analysis within the Dutch Police is shown in Figure 17.

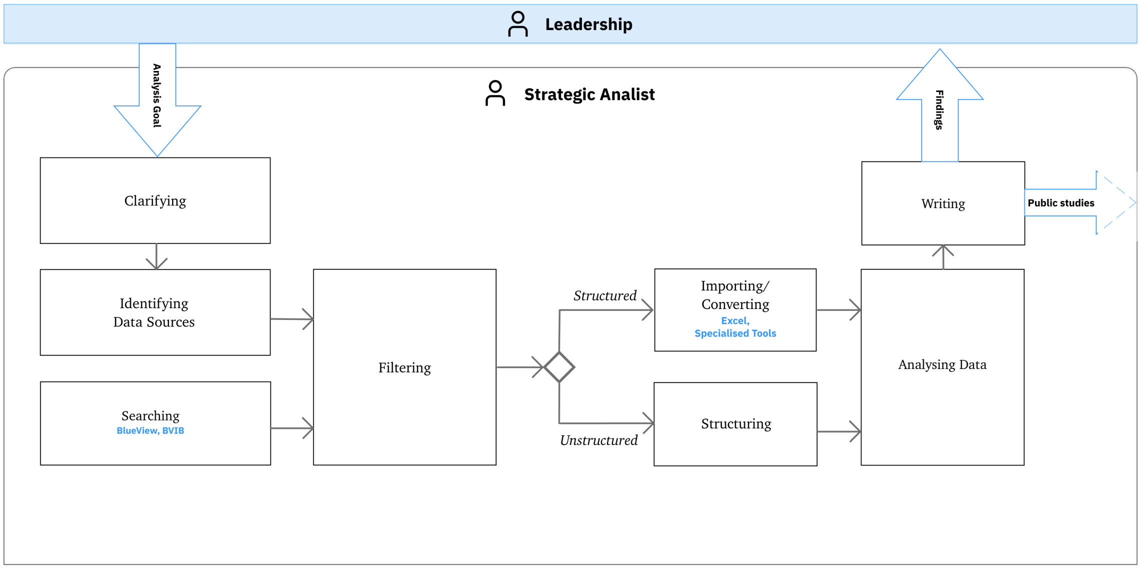


Figure 17: A model of the strategic analysis process at the regional and national units of the Dutch Police. Squares indicate activities, diamonds decisions, and arrows information flows. It is recommended to start reading from the analysis goal arrow on the top left.

A strategic analysis project starts based on an idea from the analysts or based on an information need by leadership. The information needed by leadership, is however not a concise analysis goal, so further **clarification** is needed. One researcher describes this process as follows: *“The big problem is that basically, our clients are portfolio holders. It can be the Ministry of Justice or the Public Prosecution Service, but often they are Police Chiefs. And they are never, or let me put it better, not always able to articulate their question or problem very well. They will say: ‘Yes, just look at everything of everything’. Well, that is not always possible. So, you often must take them by hand, and that is often a process in which you must talk to each other several times. And that is quite difficult because they are very busy and don’t think that’s quite their job.*" (IV16)

Depending on the analysis goal, the project starts with **identifying data sources** that are suitable for the project. One researcher describes that experience is helpful for knowing which sources are available, since it is not clear to all analysts which information they can use. There are also legal restrictions on which data can be used and if it can be combined with other data. Note that this kind of analysis often involves the anonymisation of data, since the analysis is not focused on individuals but rather on criminal organisations or phenomena. Sources for this analysis can be diverse, examples are: decrypted criminal messages; interviews with experts inside the police organisation; internal questionnaires; expert meetings; and social media analysis. Partner organisations from the police can furthermore provide data for strategic analysis, examples of such organisations are: Europol, Interpol, municipalities, and the military police.

Often for strategic analysis projects data stored in the BVH system is used. This data has some structured points which can be used in the analysis. The BVH registrations can for example include affiliation with gang members. Information from earlier investigations can also be used, this is derived from SummIT mutations. These mutations, or the parent investigation, contain some structure but mainly consist of textual findings. Finding specifics about the investigation is therefore highly time-consuming. For example, one investigation can cover multiple illegal acts, a structured property of this investigation can be that it concerns a set of illegal acts. However, if an analyst wants to find more details about one specific illegal act, they have to read through the many mutations inside the SummIT environment. One researcher describes the challenges with using SummIT data: "*[...] we have a system with which you can search in SummIT for specific criminal codes, called CA codes, and then you can also see the relevant investigations. So you look in the investigation to find these crimes back... You can really spend hours on it. Because for every investigation it lists all mutations, all investigative powers used, all permissions, all decisions... And that is a gigantic amount of work, so for your strategic analysis it actually ends very quickly.*" (IV16)

If the analyst wants to use the aforementioned structured information from BVH or SummIT, they have to **search** for the relevant data. Results from this search may then be **filtered**. Already structured data, for example found in BlueView, can be **Imported/Converted** to add to a local storage of information, such as Excel or iBase. Unstructured textual data may be manually **structured** and added to the local structured storage.

How the **data analysis** takes place is highly dependent on the type of strategic analysis that is being conducted. Some analysis goals ask for complex tools, such in the case of geo-analysis where mapping tools are essential. Others can be more straightforward, where the analysis is a numerical analysis. For quantitive analysis, statistics are not always used. An researcher describes the challenges with statistics: "*Usually, I do it as little as possible, because the data... as i said... if the input is already flawed, I can still apply a very nice statistical analysis to it, but then, yes... you're really taking a kind of double risk.*" (IV16)

Another challenge with using statistics can be explained by the iceberg analogy that the police uses to represent the real-world criminal environment. The police are aware of only a subset of the real-world crime, comparable to the small part of the iceberg that is above the water. However, they have little knowledge about what lies beneath the waterline. The police do not know the full scale of criminal phenomena and organisations. Because the size of the full population is unknown, some analysts argue that statistics are not fitting. So in order to draw conclusions about the magnitude of crime, they often rely on the assumption that the ratio between discovered and undiscovered crime remains constant.

The results of strategic analysis are presented in different manners, but often a report is **written** wit the findings. PowerPoint presentations can be created to present information directly to relevant stakeholders. The results of the research can also be presented as scientific papers. Sometimes the researcher also authors or co-authors a book that will include the findings from the research. One researcher describes that approximately half of their work can be published, while the other half is for internal use only. Within the police organisation there is a desire for reports to be short and concise. In practice however, this is a challenge for analysts, because shortening reports can remove relevant information. Stakeholders often call for the findings of strategic analysis to be presented as infographics to be more easily digested.

4.3 Cross-case Analysis

In this section, we will present the findings from our cross-analysis. In the cross-case analysis we generalise the findings from the case study, which allow us to derive a generalised sensemaking model. Analysing the different cases also lets us reveal how the processes interact with each other to form a sensemaking ecosystem. Lastly, we will discussing the similarities and differences between the three cases.

4.3.1 Sensemaking and the Sensemaking Ecosystem

Comparing the modelled processes in each of the cases allows for a generalisation of the analysis process at the Dutch Police. This generalised model is shown in Figure 18. Six common activities can be found:

(1) *Filtering*, analysts either get data from the information flow, or find it themselves. Not all information is relevant for the analysis, so information is filtered. (2) *Searching*, if an analysts needs more information for their analysis, they can search for more information. When the information is not in the existing systems, the *Information Need* product is produced, which can lead to more information being provided to the analyst. (3) *Reading*, if the information is unstructured, the analyst will read through the information in order to create an understanding. (4) *Importing/ Converting*, if the data is structured, the analyst may need to convert the data before it can be analysed. If relevant to other analysis it can also be added to the schema, if it is an *Ad hoc* analysis, the data will be analysed directly. (5) *Schematising*, the structured or unstructured information can be added to a schema the analyst is managing, the schemas are a form of structured storage, for example timelines and relational schemas. These *schemas* can already be a product of the analysis. (6) *Analysing* can happen on the information that has been collected in the project. The result of this activity are insights that are relevant to the specific analysis task.

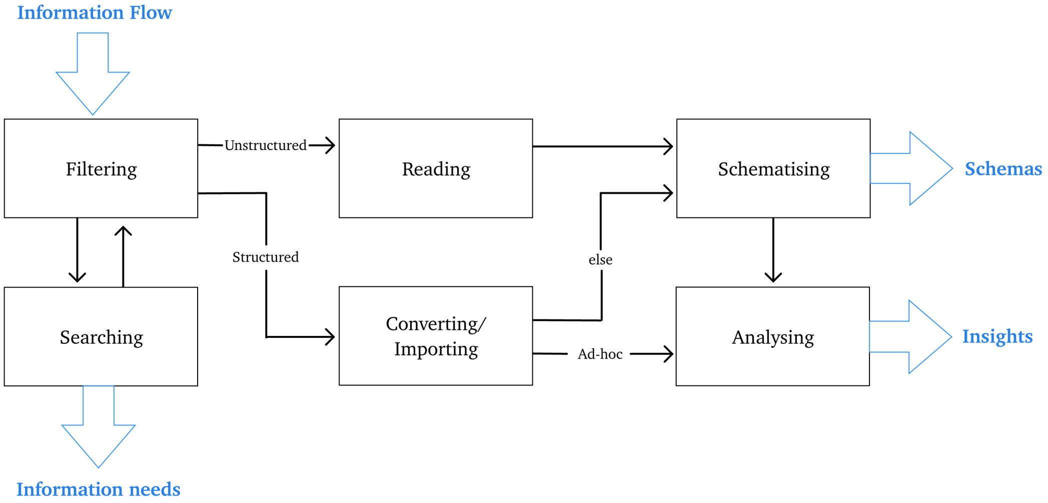


Figure 18: A generalisation of the sensemaking processes found.

Interesting for this case study is that the embedded-cases are not fully isolated and interact with each other. The different sensemaking processes inside the Dutch Police work together to form a sensemaking ecosystem. This sensemaking ecosystem should help the organisation to derive value from their source data. The results the data flow in this sensemaking ecosystem is shown in Figure 19. The figure shows the different systems that are used, and how data is extracted or added to these systems. The systems with a blue border are systems that are accessible for most teams in the organisation. In the ecosystem, investigation teams report in their SummIT environment. They can also lookup information in BVH. The analyst in the team schematises the relevant data and stores it in iBase. The security teams monitor multiple investigations through an information flow that is automatically extracted. The system used for this information flow differs, but COGNOS or Excel was most often mentioned to be part of this information flow. The security analysis team in turn structure and schematise this information. The strategic analysis teams, do not use the systems used by the security analysis teams, but have to search in the source systems, including the SummIT environments used by investigation teams. These strategic analysts also structure the information they have found for their individual analysis projects.

Interesting is that the schematisation or structuring activity is happening multiple times on the same source data. Investigation analysts structure the SummIT data for their own analysis. Security and strategic analysis can come by a subset of information at which an investigation analysts already looked at. The investigation analysts does not share its schema, so the information needs to be schematised/structured again. Important to note here is that schematisation of the information can be goal specific. This means that the schema that the investigation analyst is creating, might not be relevant for the security and strategic teams.

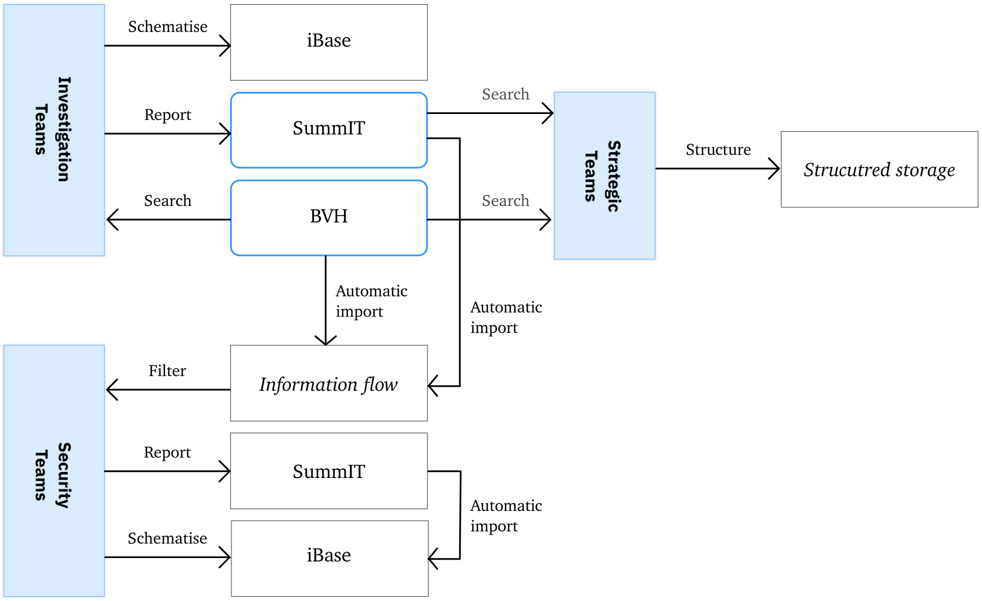


Figure 19: An overview of how information flows between the different systems and teams in the sensemaking ecosystem

Figure 20, shows a graph of how the different source systems, and the structured storage (for example in iBase) relate to each other when it comes to the structuredness of information, and the amount of information the system contains about a specific case. On the top left we see the source systems which are in most cases SummIT and BVH. During investigation analysis this information is schematised and filtered after which it is stored in the analysts iBase, following line A in the figure. In this transformation the amount of information is reduced and the structuredness increases. For security analysis a subset of relevant information is part of the information flow, that requires analysis (line B). This information is first filtered and then stored in a structured manner (line C). For strategic analysis, the analysis is question-driven, meaning that the relevant data has to be searched for, this is represented with line D. Note that security and strategic analysis all use data from multiple investigations in their analysis, explaining why the amount of information about a specific case is decreasing, while the total amount of information is not necessarily less than in a criminal investigation.

4.3.2 Similarities and Differences

Comparing the sensemaking process found in the analysed cases, allows us the discuss the similarities and differences between these. Activities that occur in all the three analysed cases are filtering, reading, schematisation, and analysis. Together, these activities form the basis of sensemaking. Another common occurrence among the two of our cases is the presence of an information flow. Analysts often work on monitoring this information flow. In investigative analysis, this information flow is a simple feed of information that is produced by the team. But for security analysis this includes a complex automated querying system that brings up potential reports that are of interest and require further analysis.

Some differences can also be found between the studied cases. The analyses cases have an increasing order in analysis project complexity and analysis project size. Analyst projects from investigation analysis are

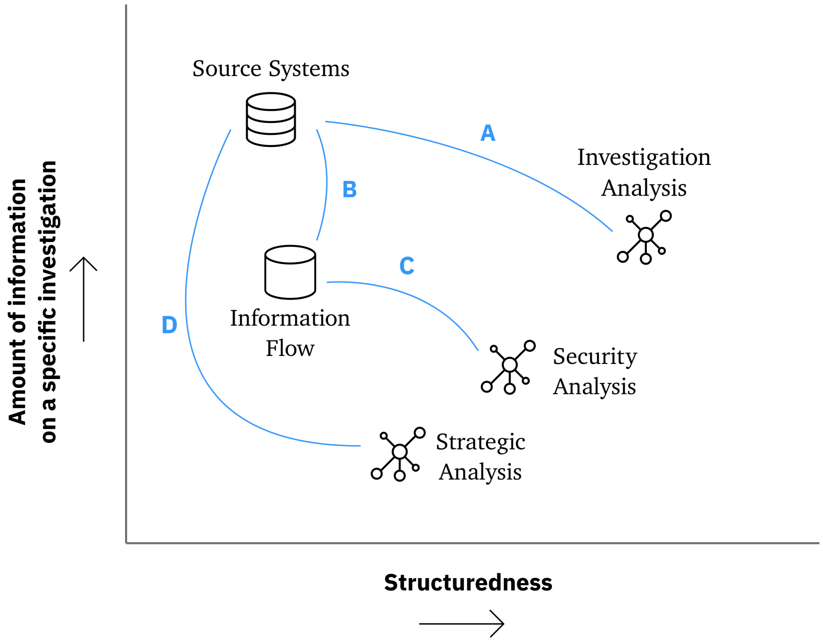


Figure 20: An overview of how the information is structured and reduced across the different systems in the sensemaking ecosystem

relatively short, while security and strategic analysis have analysis projects that take multiple days or months to finish. Also, in Investigation Analysis the small projects are more ad hoc, while security and strategic Analysis projects are more formal, and include the activity of clarifying the analysis task they will work on.

Another difference is the initiator of the analysis, differing between being driven by data (bottom-up) or by questions (top-down). The analysis on the level of investigation analysis is driven by data that is supplied by the detectives. In earlier phases the analyst is primarily working on maintaining this information flow, in the later phases an analyst has more time to initiate their own analysis. In investigation analysis, even when the task is initiated by leadership, it can have a data-centric viewpoint. In such case the question can be to ask the analyst to look at the data from an intercepted phone. For strategic analysis, the analysis is often driven by questions of the leadership, searching for the relevant data in the process. Security Analysis is a combination of both, because there is a constant monitoring flow being bottom-up (data-driven). However, the found schema is then used for top-down questions.

4.4 Sensemaking enablers

In our case study, we have spoken with many analysts and investigated how sensemaking takes place in the police organisation. With the information obtained from the interviews, we can furthermore derive the factors that enable sensemaking. We identified these enablers using the thematic coding of the transcripts of interviews. Some enablers were explicitly mentioned by the interviewees, and others were formulated by us by describing the inverse of the inhibitors mentioned by interviewees. The codebook in Appendix E describes which enablers were initially found as inhibitors. Table 4.4 shows in which interviews the enablers and inhibitors were coded. In this section the enablers are presented in four categories: (1) *data-based*

Enabler	Coded in interviews
Data-based Enablers	
High data quality	IV02, IV07, IV08, IV09, IV10, IV12, IV13, IV14, IV15, IV16
Easy data integration	IV01, IV07, IV08
Clear input rules	IV13, IV14, IV15, IV17
Automated searching tools	IV08, IV13, IV14
Easy data sharing	IV01, IV02, IV14
Availability of rich data sources	IV01, IV02, IV11, IV16
Availability of data catalogues	IV04, IV16
Data source mutability	IV07, IV14, IV15, IV17
Software-based Enablers	
Continuous IT Maintenance	IV01, IV07
Fast software speed	IV03, IV07, IV17
Availability of software catalogues	IV05, IV11
Complex visualisation and analytical tools	IV01, IV03
Software intuitiveness	IV03, IV05, IV08, IV10
Regularly using systems	IV04, IV05, IV07
Organisational Enablers	
Collaboration between sensemakers	IV03, IV12, IV15, IV16
Collaboration between information producers	IV01, IV04, IV05, IV11, IV14, IV15
Domain knowledge and specialisation	IV12, IV13, IV14
Analysis skills	IV02, IV08, IV12, IV13
Leadership that facilitates	IV03, IV08, IV13
Technical skills	IV03, IV05, IV07, IV08, IV11, IV17
Data Knowledge	IV04, IV16
Process-based Enablers	
Clear task descriptions	IV09, IV12, IV15, IV16
Standardisation of processes	IV12, IV13, IV14
Systematic hypothesis generation	IV04

Table 2: An overview of the found enablers linked with the interviews in which they were brought up.

enablers relate to the to-be-analysed data; (2) *software-based enablers* relate to the systems that analysts use to analyse the data; (3) *organisational enablers* relate to the qualities of analysts themselves, and how analysis is organisationally structured; and (4) *process-based enablers* relate to the sensemaking process. This section concludes by comparing our found enablers with those found in earlier studies.

4.4.1 Data-based Enablers

High Data Quality

For sensemaking, it is important that the to-be-analysed information is of high quality. In the policing context, there exists an interesting problem; the source data of the police, such as official reports by police officers or detectives, is textual, which is complex to analyse. The reporter, however, also has the option to create entities and assign relations alongside the reports. Entities can include persons, goods and locations. Properties of these entities can be labels, such as someone's phone number, or an address of a location. Relations can also be created between entities and can have properties. This classification creates a model of reality. In the policing context the police have to do truthful reporting, meaning that the text in an official police report has to be true. Nevertheless, in practice this does not always hold for the entities and relations that are accompanied by the textual report of the officer.

Faults in this structural reporting can include missing or false links between entities. These errors can have significant consequences in higher-level analysis, such as social network analysis. An example is giving of how such data can lead to problems in social network analysis: "*So a car has been pulled over, which is part of a set of multiple traffic stops done by a team, where they created a collective entry for 30 cars and 90 persons. They then all have a relation with each other. Then you think in your analysis: 'Well look at that! These two criminals have been in contact with each other!'* "(IV15). Therefore, it is important in such social network analyses to validate these links between persons, which can be a time-consuming task.

Another occurrence during the structuring of information is that certain properties of entities or relations have to be abstracted. This abstraction can be the categorisation of properties, which can facilitate quantitative analysis. An example of the complications that can occur during categorisation is the labelling of the type of crime that has been committed, one former analyst describes how to categorise theft: *"... colloquially we say theft, so that is what they have added in BVH, a title per case. We call this title social category, so basically what you would call such a case colloquially. It's a good idea, but super complicated in practice. Because in a lot of cases, many different things happen. A break-in can occur first, after which the person gets robbed and after that he gets shots. Which title do you give this case? All three? That gives all kinds of complications."* (IV09)

Easy Data Integration
An important part of the sensemaking process is the schematisation step. The schematisation process is time-consuming for unstructured textual information, while it should be as straightforward as possible for structured information. In the current situation, this is not always the case. For analysis, a tool was created to transfer entities, relations, and properties that are in SummIT to the analyst's iBase environment. One analyst describes the difference between systems where the imports have to be done manually and systems where it is done automatically: "*It just costs time if you want to properly import data. In other systems like Refinery, it is handled automatically, it gets imported and you can also easily export it. That allows you to have a lot more information than you would have in your own iBase environment, so that saves a lot of time.*" (IV01)

Structured data can also be used to build timelines more easily. Messages and phone calls from intercepted phones form an important basis for the timeline and should be available to the analyst in a structured format. Logs of these messages and calls contain precise timestamps, which should make it straightforward to put in a timeline.

Clear Input Rules
As described in the previous enablers, it is beneficial to be able to use structured data for sensemaking. This comes with the caveat that the structured data should accurately represent the real-world situation. In order to do so rules must be created so not only the structured data is entered in a consistent format, but also that different analysts interpret the structured data in the same manner. Input rules can also make unstructured data, such as regular text fields, be structured. Some thematic security teams work with a strict format for subject lines of mutations in SummIT, so the field can also convey structured information.

Having well-structured data aids in faster analysis because less data cleaning is necessary, and better analysis because the input data contains fewer errors. It can even help do automated analysis of structured information, such as those resulting in dashboards. One analysts describe the benefit of input rules: "*At this point in time we process everything in the SummIT environment, so that all information comes together. We also enter it according to fixed input rules, so that we can also run dashboard on it, with the validated data, so to speak*" (IV15)

Automated Searching Tools
The amount of data that is available to the police is high, and only a small piece of information is relevant for a specific team or analyst. For security teams, the analysis also contains the monitoring of an information flow. This information flow is powered by a search tool. This can be a simple or complex query, or may

include machine learning technologies to find reports that are relevant for the thematic security team.

Sometimes the to-be-analysed data also contains a large set of messages, for example originating from hacked encrypted messaging service providers that criminals use. These include millions of messages that cannot all be read by police officers. TROI, a team whose goal is to help analysts, created a tool which can search in the messages to find the ones that relate to criminal activities. This search includes computer vision to find images that contain illegal goods, such as weapons and drugs. These search tools can bring up potentially interesting information.

Easy Data Sharing
Each criminal investigation produces a large amount of data. Some time ago, investigation teams did not share details about an investigation with other teams. This has changed over time, one analyst describes it as follows: "*Almost nothing got shared, but since a few years the new vision is, you have to share as much information as possible with colleagues across the country, unless you have a very good reason not to do so.*" (IV02)

Software got developed to utilise this shared information. Two examples are: a system that automatically checks if a person that is entered in SummIT is already mentioned in another investigation. If a match is found the analyst can look up more information about this subject in the other investigation. Second is a tool called Full Contact, which allows analysts to search in a large database containing contact lists of confiscated phones. This allows the analyst to find subjects from earlier investigation that might have ties with the person of interest.

Availability of Rich Data Sources
When it comes to sensemaking and criminal investigations in general, it is beneficial to have as much data as possible about the subject or subjects you are investigating. The more data that is available the more the police can get to know about a subject. Data sources from the police can be diverse and range from findings of infiltrators to trackers placed under cars. Combining this information allows the police to know the details of a specific subject. Sensemaking is thus aided by having not only a high quantity of data but also a wide set of different data sources.

Availability of Data Catalogues
the types of analysis that occur within strategic analysis teams differ between projects. This means that for each new project, the analysis starts with identifying the data sources that can be used for that project. One researcher describes the process: ".... *what did the earlier literature find, and what are the most important variables that I want to look into? I will then look for sources that can get a view on that.*" (IV16). It is beneficial for sensemaking if the analyst knows which data sources are available for their analysis. This can help them find the best data possible for their research question (top-down), as well as to know which data is valuable for analysis (bottom-up).

Data Source Mutability
In some cases, the analyst might do an analysis of information where the structured part of the information is incorrectly labelled. This information can be part of an import from source systems such as BVH and SummIT. In this case, the analyst can spot the mistake and fix it in their local schema, but ideally, the source systems need to be corrected as well. One iBase expert user describes an example: "*You have collected information in iBase and you are working with it, then you see, oh, I see a duplicate name... I will merge it. Or, I see 5 companies but they are all the same company, so I will combine them into one... I am missing a link, I will link them myself, otherwise I cannot continue with my analysis. But I will still pass it along, so it will be entered into the source systems*" (IV07).

4.4.2 Software-based Enablers

Continuous IT Maintenance
The police provide the analyst with a wide range of tools. These tools support analysis directly or help in its subprocesses such as searching, converting, and schematising. Sometimes tools are also created for specific

methods of regional units or thematic teams. These tools must be properly maintained, so that functionality is ensured when operating systems or programs with which the tool interfaces change. An analyst described the following about a tool that stopped working after an update: "*... the colleague who maintained it... it was not properly guaranteed... It was the responsibility of one colleague, while many analysts used the program. So when we moved to Office 365, the tool stopped working.*" (IV01)

Fast Software Speed
The speed of software is also a factor that can benefit sensemaking. If a computer spends a long time loading, it can have the analyst waiting for the program. This is time that could be spent better. It can furthermore be bothersome for the sensemaker if this process takes too long. One analyst describes this problem: "*But we also work with outdated systems. Sometimes there is a modern shell over it, which makes it look very hip, but in one case there is even still an old MS-DOS system underneath. So that does annoy, because it just slows down your work a lot, which does not suit modern times. I find that part annoying or at least a waste of my time*" (IV03)

Availability of Software Catalogues
Comparable to the data catalogues enabler, the analyst needs to know which IT systems are available to them. In multiple interviews, analysts mentioned that there are a lot of systems available to them and that new ones are regularly being introduced. If the analyst knows which tools are available to them, they can choose a fitting one to use for their analysis, as well as to know which kinds of analysis are possible on the data that is available. An analyst describes the challenge: "*But there is so much in development at the moment and that is a challenge as an intelligence employee. Everyone tries to do all kinds of smart things, but with that you lose an overview of which tools are actually available, and how I can use them to improve my work with them?* (IV03)

Complex Visualisation and Analytical Tools
The police also have an increasing amount of structural information. One of the data types is locations. Trackers, phones and other devices can share precise locations of subjects that are of interest to the police. Analysing these found coordinates across the time dimension requires specialised tools. These tools can help the analyst interpret the raw data for geospatial analysis. Another use-case of such tools is the analysis of social networks. Tools can compute metrics of networks, such as the centrality of a node in the network.

Software Intuitiveness
Multiple analysts mentioned in the interviews that the software that they use is hard to understand. This can sometimes be a reason why an analyst does not use a specific system. Many systems require training before an analyst can use them. For sensemaking, it is beneficial if the barrier to entry for software that aids in the process is as low as possible. Increasing intuitiveness can therefore increase the amount of analysts that can perform more complex analysis. An analyst described the challenges they were facing with the analysis program Refinery: "*I do not use it that often, that has some reasons. It is not an easy program, Refinery, it costs quite some time to learn it. It is not like you would say, I will go do a training for a week and then I know it.*" (IV10)

Regularly Using Systems
Continuing on the complexity of systems, is that interviewees brought up that for some applications is crucial to keep using the system. This will keep the analyst proficient with the application, and let them know what possibilities the system offers for their analysis. In the continued explanation the analyst describes: "*because you have to continue using Refinery to be skilled at it.*" (IV10)

4.4.3 Organisational Enablers

Collaboration Between Analysts
Most analysts spoke to, mention that collaboration can be a crucial part of performing analysis. However, in practice there is a shortage of analysts on projects. Some regions are struggling with assigning a single analyst to an investigation team, so assigning multiple to investigation or security teams is a challenge. Still,

one analyst states the following about collaboration: "*We have also discovered that for analysis you actually need a minimum of two analysts to be able to regularly reflect, to discuss: am I seeing this right? Am I not falling into a tunnel vision? Which other possibilities are there for my analysis and tooling? Or to look at, can we get another view or a better view on something?*" (IV10). This collaboration can also be on a cross-regional level. Regional security analysts can discuss with other analysts what they are working on and how they approach it.

Collaboration Between Information Producers
In investigation teams, the analyst is often placed close to the investigators. This allows for close ties with the information producers in the investigation, which are in this case the detectives. This relation goes both ways. The analyst can ask questions about information produced by the detectives and can highlight information needs which require further investigation. But also, the analyst can informally receive questions from the detectives and provide answers. The following was mentioned about collaboration between detectives and analysts: "*That is why, inside the DLR it was chosen that the case analysts sit together with the investigation team themselves, to have constant contact with the other members of the team, the detectives. So if they have something they can come to me and if I have a specific question, I can ask them. Then you have very short lines, which is actually the intention*" (IV01)

Domain Knowledge and Specialisation
Specialised knowledge is required to be an analyst at the Police. Not only is knowledge needed about the police organisation and operations, such as what investigative resources are available in a typical investigation, but also knowledge about the criminal phenomena. For example, the human trafficking team worked together with Leiden University to derive signals of human trafficking. Analysts and other intelligence employees are trained to recognise these signals. This allows thematic security teams, to specialise in the type of crime they are monitoring.

The same holds for the other types of specialisations mentioned in the security analysis case study; requiring up-to-date knowledge about radicalisation across the world to spot trends inside within the CTER theme in The Netherlands. Likewise, for the analysis regarding the production of synthetic drugs, the specialised crimescripting methodology is used which is adapted for the specific team.

Analysis Skills
Analysis was brought up to be an important part of the skill set of an analyst. Analysts have to question the results they are getting. The results they are getting might not represent reality or indicate results that are initially thought. For example, when mapping the addresses of bank accounts that have committed fraud, one analyst found an address that came forward a lot. This can lead to a false conclusion that somebody lives there who committed a lot of fraud. However, in this case, the actual address was a homeless shelter, where many homeless people were registered. These homeless people sold their bank access to criminals, which allowed them to commit fraud in their name.

Leadership That Facilitates
In the interviews it was also mentioned that the team leaders can play an important role, in what analyses the analyst performs. Some team leaders, limit the analyst capabilities by expecting relatively simple products from them. This problem was described by one analyst: "*There is a really big difference between analysts and what a team leader from an investigation expect of analysts. Some think: 'nah, they only make simple overviews'. No, we are not of the simple overviews, we do it too, but we are there for analysis. We want to give insights, so you can do smart things as a tactical team.*" (IV03)

Technical Skills
The complex tools that aid in sensemaking, require technological skills from the analyst. One former team manager describes that this goes paired with challenges about keeping analysts on the same level: "... *the Refinery workflow is pretty complex, and that requires a completely different skill set. Not everybody can go along with that. That is a big challenge for a lot of A&O departments, and many will get stuck in using somewhat simpler tools. iBase and Analyst Notebook are simply easier to understand. Indeed, you will also stay behind in the analysis and the possibilities of the products. That will then be a ceiling. I would say, it is the challenge to go*

beyond that." (IV08)

Other general or miscellanies technical challenges can occur during the analysis process, such as the converting or retrieval of data that can be aided by computer dexterity. This increases the kinds of analysis an analyst can perform and can improve speed. One analyst describes how other colleagues with programming skills do other tasks: *“There also exists a digital specialist. They like to work on large sets of data and structure it. Yes, you know, he then writes some python script which he then executes, yeah that goes a lot quicker then if I spit through the information old school.”* (IV11)

Data Knowledge
When data is analysed, analysts need to know how the data was collected, and what limitations it can have. Some sources can have important limitations. Especially in the policing domain, not all information might be truthful. One analyst gives an example: "*It’s also included in the education, be cautious with all your sources. . . . If you hear scenario A 6 times, but scenario B only once. . . The number of times does not matter more, for example, report crime anonymously, someone can call up to 6 times. . .* " (IV04)

4.4.4 Process-based enablers

Clear Task Descriptions
Security analysis and strategic analysis both involve some degree of discussing the project’s goal with the client. This is an important predecessor of sensemaking. One researcher describes the importance of clarifying the goal of the strategic analysis project clearly: "*That is a very difficult one, but it is the most central step because otherwise, you will always end up with hassle afterwards, saying ’you should have done that too” because only when you deliver something, they suddenly know retroactively what they would have wanted, so that remains one of the big challenges.*" (IV16)

Standardisation of Processes
Different analysis levels, teams, and regions have adopted standardised processes to perform analysis more effectively. These standardisations can come paired with the specialisations, that are also mentioned in the *Domain Knowledge and Specialisation* enabler, such as crimescript. Another form of standardisation is that certain types of analysis can have valuable results, and should be executed more often either at different times or on different data. For example, threat assessments are performed in a systematic way that is teached at the police academy. Here the reason for the standardisation is the importance of the accuracy of the result and the impact of the result itself. Another example of standardisation is the documented procedure that was created to perform a social network analysis on criminal groups.

Systematic Hypothesis Generation
In investigation analysis, where hypothesis formulation has a clearly defined process, there is an analyst whose primary role is to create and consider hypotheses and scenarios. The analysts systematically write down evidence in a mind map, and link sub-hypotheses with confirming and disproving evidence. This method overlaps with the method of *analysis of competing hypotheses* described in literature [16, 10], and is also found in other intelligence analysis processes [7]. This method allows for a systematic way to evaluated created scenarios based on evidence.

4.4.5 Comparison to Theory

The enablers of sensemaking in an information context have not been described in earlier literature. Even though this area has gone overlooked in the literature, there are several studies on the factors that are beneficial for Big Data (Analytics). We can compare our findings with enablers in this field. The comparison allows us to highlight the new contributions to the literature. As well as to identify the overlap and differences between the kind of enablers in the two fields.

In Table 3 the enablers that are found are mapped tofive studies that looked at the enablers, challenges and critical success factors of Big Data: (1) A case study looking at the enablers and inhibitors of effectively using

Big Data found inside a case study [48]; (2) A case study that looked at challenges implementing Big Data Analytics [31]; (3 & 4) Two literature reviews that looked at critical success factors of Big Data [55, 2]; (5) A study which combined data from 60 case studies to find critical success factors of Big Data categorised per step in the process [12]. Note that because the mapping is made between enablers in two different fields, the relationship entails similarity, not equality.

Table 3: A mapping of the enablers found in our case study to factors that have been found to be important for Big Data (analysis) in earlier studies.

Enabler	**Sejahtera et al.** [48]	**Malaka and Brown** [31]	**Walls and Barnard** [55]	**Al-Sai, Abdullah and Husin** [2]	**Gao, Koronios and Selle** [12]
High data quality	Poor data quality (inhibitor)	Data quality; Data integrity	Data	Data sources	High data quality
Easy data integration	Data silos (inhibitor)	Data integration		Access to sources	Combine different data sets
Clear input rules				Data standardisation	
Automated searching tools					
Easy data sharing			Information sharing	Data sharing	
Availability of rich data sources			Data	Data sources	
Availability of data catalogues					
Data source mutability					
Continuous IT maintenance					
Fast software speed	Adequate system capabilities			Flexibility and scalability of software applications	
Availability of software catalogues					
Complex visualisation and analytical tools			Infrastructure and analytics platform	Technology, infrastructure, and applications	Visualisation
Software intuitiveness					
Regularly using systems					
Collaboration between sensemakers					
Collaboration between information producers					
Domain knowledge and specialisation					

Continued on next page

Table 3: A mapping of the enablers found in our case study to factors that have been found to be important for Big Data (analysis) in earlier studies. (Continued)

Enabler	**Sejahtera et al.** [48]	**Malaka and Brown** [31]	**Walls and Barnard** [55]	**Al-Sai, Abdullah and Husin** [2]	**Gao, Koronios and Selle** [12]
Analysis Skills		Skills shortage	Analytical skills of the employees	Human capability	Analytical skillset
Leadership that facilitates	Champions		Managerial skills		
Technical skills	Lack of technical skills (inhibitor)	Skills shortage	Technical knowledge	Human capability	Technical skillset
Data knowledge				Data documentation	
Clear task descriptions					
Standardisation of processes					
Systematic hypothesis generation					

5 Discussion

In this section, we reflect on the research we have done and discuss its implications. First the results of this study are compared to those found in earlier literature in the scientific implications section. After that the practical implications of the study are discussed. The findings of the study also allow us to list recommendations for the police organisation, which are presented in the recommendations section. The next section discusses the limitations of the study. Lastly, we call for future works on specific subjects in the future directions section.

5.1 Scientific Implications

In this study, we have analysed the intelligence analysis processes inside the police and looked at how sensemaking takes place. This study is the first we could find that looks at multiple sensemaking processes in one organisation, and how they form a sensemaking ecosystem together. In this section, we will look at how the sensemaking processes found compare to those described in theory. We will conclude this section by reflecting on the found enablers of sensemaking.

5.1.1 Top-down, bottom-up, and ad hoc sensemaking

The sensemaking processes found in our case study are executed as a mixture of ad hoc, top-down (goal-driven), and bottom-up (data-driven). Many studies highlighted the ad hoc nature of sensemaking [21, 28, 63, 7]. Pirolli and Card describe sensemaking as an “opportunistic mix” between bottom-up and top-down processes, which is consistent with our findings [40]. Each embedded case we analysed has a predominant direction: investigation analysis involves the schematisation and analysis of information provided by the investigation team, which is a bottom-up process. Security analysis consists of two processes, a bottom-up monitoring process where relevant information is structured and a top-down process where the structured data is analysed. Strategic analysis uses top-down processes that are driven by questions from leadership or the analysts themselves. Note that this classification as bottom-up or top-down regards the primary direction of analysis. In our case study we found that, for the bottom-up processes, the created schemas can lead to questions which can call for the search for more information, reversing the direction of the process.

Comparing our found process to the process described by Pirolli and Card [40], we notice that the *shoebox* and *evidence file* are not actually containers of information, while information is filtered, read, and schematised the information itself is not necessarily transformed nor moved between these steps. The *shoebox* and *evidence file* missing from the actual work method is in line with earlier findings [19].

Furthermore, we found that the process on the investigative level consisted of a smaller project cycle, making the projects more ad hoc. Comparing this to he *cyclical intelligence model* [28], we found the intelligence analytics process includes less *planning/tasking* and *requirement/feedback* for bottom-up (data-driven) analysis compared to top-down analysis (goal-directed). Overall, we found that investigation analysis mainly involves just a sub-set of activities, since their individual analysis products are smaller and more informal. In contrast, the security and strategic analysis products are more time-consuming and include more activities mentioned in the cycle.

5.1.2 Schematisation through predefined structures

In our study, we found schematisation to be a crucial part of the sensemaking process which is in line with findings from [19]. Pirolli and Card have described schematisation as being part of their sensemaking process [40]. The found schematisation activity is in line with the *learning loop* proposed by Russell et al. [47] and the *structural information seeking cycle* of Qu and Furnas [42]. In both models, a representation is built, which is similar to building a schema. The two models include a concept related to structuring information, namely *searching for good representations* in the learning loop, and *structural-information need* in the structural information seeking cycle. These concepts are focused on the need for structuring ideas. In the schematisation activity we identified in our case study, the search for ideas for a structure was not mentioned as being part of the process. This can be explained by the fact that, in general, investigation analysis projects overlap in the types of entities, relations, and properties that are available. For example, most investigations have persons, vehicles, goods, and locations as entities, which all have a standard set of properties. This suggests the analyst already knows an adequate representation to schematise the information in, and does not need to develop or find this structure.

5.1.3 Handling the information flow

In the sensemaking process that we revealed in our case study, we identified several important concepts that were not described in the literature on sensemaking. The first concept is *information flow* in bottom-up (data-driven) sensemaking. This is a flow of information that is ought to be analysed. This flow can be provided by other people in a team, or by an automated system searching for relevant data. We suggest two reasons why this has not been described in earlier studies. First, is that technological innovations have enabled these real-time information feeds with complex searching algorithms. These technologies might not have been available during earlier sensemaking research. Second, the information feed is the result of a sensemaking process that is highly-integrated into an investigative organisation. In contrast to an isolated project-based sensemaking process where this infrastructure may not be in place, the Police is an organisation where sensemaking is essential and takes place throughout the organisation. In our case study, interactions between the sensemaking process and the environment in which it operates are reflected in the *information flow*.

Also new in our sensemaking process description, is that we explicitly describe *schemas* and *information needs* as being a product of the process. Earlier literature has described the role of schematisation to be an intermediate step in the process [40], only aiding in the creation of subsequent insights from sensemaking. However, within the context of the investigative sensemaking process, we observed that one of the main tasks of the analysts was to maintain schematisation products. The schematisation products took the form of a timeline of events and a relational schema of the subjects involved. The same holds for the *information need* product, the Cognitive Sensemaking model does mention a *data gap* as an intermediate concept in the process [66]. Yet, in the sensemaking process we found through our case study, we find that this information need is not only an intermediate step, but a tangible product of the process. With this information need, the team will try to resolve this need by retrieving additional relevant information.

Another addition to our sensemaking process description compared to earlier studies, is the integration of structured data analysis into the process. Because of innovations related to Big Data, the variety of data has increased. The information that is available to an analyst is not only textual; in our case study, we found that investigation analysts also work with structured data from, for example, trackers and intercepted phones. This information can also play a crucial role in sensemaking, because it acts as another unique source of information. In the sensemaking model described by Pirolli and Card [66], it is not mentioned how structured information fits into the process. In contrast, our sensemaking process describes how structured information is handled. We added the activity of *converting/ importing* the information to facilitate the analysis. The structured information may also be added to the schema, through the *schematisation* activity.

In the analysed processes of our case study, we have only looked at what happens outside the analyst's mind, i.e., what activities they are performing and what systems they are using. Nevertheless, we can highlight a difference found between our findings and the cognitive sensemaking model [66]. The structure that is build in their described processes, represents either an internal or external representation of information. The structure is grown by the *identification of gaps* and *seeking for data* activity. However, in our case study, we found that in the sensemaking process there is also a feed of data that is ought to be analysed, without requiring the analysts to seek additional data. Even when the model regards an internal structure, which according to the authors can be a narrative in criminal intelligence analysis, the same holds; the story in the analyst's mind is not created by only *seeking for data* and *seeking for structure* that will fill in *data gaps* or *structure gaps* respectively, but also by data that is provided by the detectives that must also be considered.

We argue that there is a missing activity that tests the structure in the analyst's mind on new data, without it being motivated by a *data gap*.

5.1.4 Enablers

Based on our case study, we have also presented a list of enablers for the sensemaking found at the Dutch National Police. We did not find any papers listing enablers that link to sensemaking in an information science context. One of our found enablers is confirmed by statements made in papers that looked at sensemaking. Chin et al. state that *multivariate visualisations* can be beneficial, comparable to our *Complex visualisation and analytical tools* enabler [7]. We have furthermore not found papers that look at the enablers of sensemaking in the larger context, including the sensemaking ecosystem.

In Section 4.4.5, we have compared our findings with enablers found in a Big Data research context. We found that especially our enablers relating to data quality, integration, and availability are relevant in Big Data. The same holds for software capabilities, analytical skills, technical knowledge, and management being important factors for Big Data.

The enablers that could not be mapped to those described in the selected literature, allow us to highlight the unique challenges that come paired with sensemaking, especially in a sensemaking ecosystem. The *Automated searching tools*, *Data source mutability*, and *Collaboration between information producers* are all resultant of the challenges that arise in a complex sensemaking ecosystem with a large amount of unstructured data.

5.2 Recommendations

The Dutch Police organisation is particularly complex. Implementing changes in the organisation is highly challenging organisationally, legally, and culturally. In this section we will list a set of recommendations that the police can implement which we believe, based on the findings of this study, can improve the sensemaking operations. When creating these recommendations we have tried to make them realistically implementable. While these recommendations are created for the Dutch police organisation, some of them can be applied to other organisations wanting to promote the sensemaking of both structured and unstructured information.

5.2.1 Organisational recommendations

Encourage collaboration Earlier literature [7] and the analysts in our study, described that collaboration can be beneficial in intelligence analysis. We recommend facilitating this collaboration between analysts so they can reflect on their findings. This reflection can also bring up insights about their methods which can be helpful for the other, as well as challenges with their own workflow, with which the other analyst can help.

Keep critically reflecting on processes and learn from others. We recommend managers to reflect on their present processes to check if they have any issues or opportunities for innovation. Inside the police, analysis happens at a lot of different places within the organisation. One type of analysis, security analysis, occurs in multiple teams per region each focusing on one specific crime theme. In our study, we found that these work methods differ per theme and also sometimes per region. We see that the thematic teams individually innovate and set standard processes to improve their workflow. If successful, these process optimisations can be a useful insight to others. We therefore recommend other teams to compare their processes to the ones from other teams, and to see which parts of that new process may be beneficial for their own work methods.

Make and enforce input guidelines. An issue that was often mentioned by the interviewees is the data quality of the inputted entries inside the BVH and SummIT systems. This information is textual but contains the option to add structured elements. In practice, however, this information can be improperly labelled, inconsistent, incomplete, or incorrect. The quality of this information is crucial for the analysis that happens on all of the three analysis levels. We recommend continuing the efforts to improve this data quality. Further defining input guidelines can help to increase the amount of structured information that fits into the current systems, aiding analysis. Analysts also mentioned in the interviews that it can be beneficial to contact the person who incorrectly inputted something into the system, to explain to them not only what was incorrectly inputted, but also what impact this incorrect input has.

Be aware that standardisation is a double-edged sword. Standardisation inside security teams within a specific theme has allowed some teams to monitor incoming signals more effectively, by prescribing the way incoming signals are monitored. It also allowed for better analysis since information is stored in a more structured manner. Standardisation can deliver promising results to intelligence analysis in the form of specialisation and better data quality. When designing a standardised software or process one should however not overlook that some security teams can be highly specialised. Forcing standard processes or standardised software inside these teams can have challenges with continuing to enable these specialised analyses, negatively contributing towards their analysis operations.

Have technically capable leadership. Earlier work has described that managers who promote IT solutions, called champions, can be enablers for Big Data in organisations [48]. We argue that having a leader in teams who knows about the value and opportunities for complex analysis, can in turn help facilitate this analysis. Whilst the analysts themselves also have to be technologically capable to do this complex analysis, it is important that the direct environment is also facilitating. Multiple interviewees mentioned that in practice some team leaders under which the analysts operate do not facilitate these more complex analyses.

Increase and maintain technical knowledge. In our study, we found technical knowledge to be an important enabler of sensemaking. We recommend the Police organisation to investigate how to increase the technical knowledge of current analysts. This can for example be done using training and courses. These courses must adapt to the real-world work of the analysts, so they can implement the learnings in their own flow. Working with these systems in the real world is also a crucial part of maintaining technical knowledge. We therefore also argue that it is important that analysts keep using systems, so they maintain their knowledge about it. This can for example be done in coordination with management, so the analyst is offered work that utilises the systems.

5.2.2 Technical recommendations

Standardise data formats. Interviewees from our case study described that the process of using already structured data in their analysis can have some miscellaneous technical challenges. This can occur if they receive information from other units inside the police in an unconventional format. In some cases, they can then not import this information automatically into their schemas. We therefore recommend to set output format guidelines for information that is shared. For example, when sharing intercepted messages, they should be in a predefined format with a predefined set of properties which all have defined names. Other systems that can import this information should then be able to accept these files in this structured format.

Make importing information easier. Besides the aforementioned standardisation of data formats, it is also beneficial if structured information can be transferred between systems as easily as possible. Where possible it is also beneficial if these imports are not one-time only, but track changes over time. This would entail that if information in the source system has changed, you can choose to synchronise the local copy in another system so the information remains up-to-date. These imports should be as straightforward and fast as possible.

Source system mutability. In the analysis ecosystem within the police, it can occur that multiple analysts will look at the same source data. For example, an investigation analyst can schematise reports from a detective to create a timeline, while a security analyst schematises it to build a relational schema. It can happen that the SummIT mutation of a detective does not contain the appropriate entities and relations in the source system. The investigation analyst might have encountered this issue and changed it in their own schema. However, they might have not changed it in the source system. This will cause another analyst to run into the same issue. This problem can also occur with data in the BVH system. We therefore recommend making it easy for people to correct the mutations in these source systems, and to also motivate others to correct mistakes they encounter. This will lead to higher-quality data inside the source systems.

Look into how AI can help schematisation or structuring. With the recent advances made in the ability of AI to understand natural language by using Large Language Models (LLMs) [51], we reason these technologies can aid in the structuring or schematisation of text. Reports from police officers, such as those on the streets and detectives, contain a large amount of text. With these reports there are structured fields, where the reporter can assign properties to the report, allowing for a classification. It is also possible to create and link entities to the report. In practice, this classification can be incomplete or incorrect. We therefore recommend looking for a solution that can help in this process. An interesting experiment would be to implement an AI algorithm that can monitor incoming reports and look for missing or incorrect structuring. If any were found, the program could email this to the reporter. The reporter can then look at the provided feedback, and check if they want to accept the corrections.

Increase intuitiveness. As described in the enablers, the intuitiveness of systems enables sensemaking. We recommend having the intuitiveness of a solution as a requirement for the development of new software. It is beneficial if sensemakers can use new systems without the need for dedicated training.

Large screens. Earlier literature has described that using one or more large high-resolution screens can facilitate sensemaking [7]. The literature describes the additional screen can function as rapid-access memory [3]. We recommend that analysts should be provided with such displays.

5.3 Practical implications

Our research gives a high-level overview of the processes that are related to intelligence analysis at three levels within the Dutch National Police. Unique to the policing context is the large amount of unstructured information that is available. We have described what the current processes are of analysts who extract insights from this information. This research can be used to get an understanding of these analysis processes. Additionally, this research can help with creating an understanding of the policing context in which these processes operate. In turn, the process descriptions can be used to find opportunities for improvement inside the police organisation.

Our found enablers furthermore describe the factors that are beneficial for intelligence analysis. The police, and other companies relying on analysis of unstructured information, can learn from these enablers. These enablers can be implemented to aid in the value extraction from unstructured information that is available to an organisation. The Dutch Police, and other organisations like it, can furthermore learn from the recommendations that have been made in section 5.2.

5.4 Limitations

In this section, we will reflect on the limitations of this research. First, the participants in our case study have been selected through convenience sampling. Analysts, former analysts, and managers spoke to have been recruited for the case study because they were people in the organisation who were known, or colleagues of these individuals we know. Furthermore, participation to the research was voluntary. This could possibly lead to a bias where only the working methods of analysts were studied of those who are okay with others looking in to it. The findings of the case study are furthermore limited by the honesty of the interviewees and the accuracy of their answers. We have tried to mitigate this limitation by interviewing multiple persons within each case, as well as different roles.

In this research there are two levels of generalisability, the first being generalisability to the Dutch police, and the second being generalisability towards sensemaking in general. For the first the limitation regarding generalisability is that we have only looked at three main varieties of intelligence analysis. In the organisation there are, however, more places where intelligence analysis can take place. Also, inside these three levels, there are a high number of teams of which many use different working methods. We argue that our cross-case analysis can generalise the findings to the Dutch police since it takes into account the three separate analysis levels. The second type of generalisability regarding whether the findings can be generalised is however more limited. The police context is highly complex, with a unique sensemaking ecosystem. This context is part of all three embedded cases of our case study, which makes generalising the results outside of this context complex. We do argue that our results may be generalised to other policing organisations worldwide, as well as, other investigative organisations where the analysis of large amounts of textual information takes place

Another limitation of this research is the inductive coding that has been performed. The coding process is performed by one researcher. Even though the coding has been done in multiple iterations, the results of the coding are still limited because of the subjective nature of (inductive) coding. To counter this limitation the resulting code has also been discussed with the other researchers. The resulting code book is described in Appendix E.

Lastly, no validation has been performed on the results. Even though this was initially planned, it was found that there were large differences between the working methods of the analysts in one group. Having validation interviews with a subset of the initial interviewees would therefore falsely make their responses more influential in the results. The validation interviews were therefore substituted by increasing the number of interviewees.

5.5 Future directions

Lastly, we call for future studies into some specific areas of the sensemaking field. First, our research did not include a validation of the found generalised process and the enablers. A repeating study can validate the enablers and test their generalisability in another context.

Additionally, we call, in general, towards more research on sensemaking within organisations, within the IS context. Specifically interesting is how the individual sensemaking operations combine to form a sensemaking ecosystem. Our found enablers and descriptions of real-world sensemaking processes can help design a future study that will look into such ecosystem.

In our study we found the schematisation activity to sometimes happen multiple times on the same information. We call for research on how to the reuse of schemas can be improved, so other users can use it in their sensemaking processes. Here an interesting challenge is the structural storage of this information, in a way that it can easily be reused, but also queried and searched through. We argue that the field of Knowledge Graphs[18] may provide insights for improving the schematisation in sensemaking, especially when it comes to the sensemaking we found where predefined structures are used.

In terms of information systems that aid sensemaking, we suggest that it can be interesting to look in to the possibilities of automating schematisation efforts in sensemaking. NLP techniques can possibly aid in automated schematisation. Especially recent advancements in natural text understanding using LLMs [51], may be beneficial for such implementation. Also for the field of Knowledge Graphs, the Automatic Knowledge Graph Construction research area (e.g. [68]) may have interesting insights for automated schematisation. This automatic construction can also be aided by LLMs [61].

In our research we have focused on how sensemaking takes place externally, looking at the activities of sensemakers, their products, the systems they use, and the greater context in which sensemaking takes place. We have however not investigated the processes that occur in an analyst's mind. Because of this, we were unable to validate the claim by earlier literature stating that the *data/frame theory* [24] is the *"best model for intelligence"* [35]. We therefore call for future works on the internal sensemaking processes. This research area is specifically interesting with the further rise of AI and other black-box algorithms.

6 Conclusion

In this study, we have investigated the intelligence analysis operations inside the Dutch Police to study how sensemaking takes place. This section will conclude our study by answering the research questions.

RQ1 What does earlier research describe about the sensemaking process?

Our first research question has been answered in the conclusion section of the related works (Section 3.6). In this section, we have discussed the processes that have been described in earlier studies. We have looked at sensemaking from three different viewpoints: *Representation-focused sensemaking*, *Cognitive-Focused sensemaking* and *Collaborative sensemaking*.

RQ2 How is intelligence analysis conducted at the Dutch Police?

The intelligence analysis operations are described in our findings section. In our within-case analysis (Section 4.2) we have described three analysis processes in detail: investigation analysis, security analysis, and strategic analysis. For each case, we have presented a process diagram of the working method of analysts, what systems they use for the activities, what products they deliver, and with whom they interact.

In our cross-case analysis, we have also described how these intelligence analysis operations interact with each other. We have generalised how the different analysis levels interact with the same set of systems. We have also described how information is transformed between these systems.

RQ3 How does sensemaking take place at the Dutch Police and what are its enablers?

By generalising the results of the three individual cases, we were able to present a generalised sensemaking process in the cross-case analysis section (Section 4.3). The model describes how we found sensemaking takes place in the organisation. We have also described how sensemaking can vary between the higher and lower levels of analysis. The answer to this research question allows us to answer the main research question:

MRQ How does sensemaking take place inside the intelligence analysis operations at the Dutch Police, and what are the enablers of sensemaking in the organisation?

Our generalised description of the sensemaking process found in the case study allows us to answer the research question. Inside the police sensemaking is a combination of six activities: *filtering*, *searching*, *reading*, *converting/importing*, *schematising* and *analysing*. Products of sensemaking are *information needs*, *schemas*, and *insights*. For sensemakers there can also be an incoming *Information feed*. The generalised model is described in Section 4.3.1 and shown in Figure 18.

The data collected in the case study also allowed us to look into the enablers of sensemaking in the organisation. We have listed a set of 24 enablers, which are presented in Section 4.4. The enablers have been categorised into four categories: (1) *data-based enablers* relate to the to-be-analysed data; (2) *software-based enablers* relate to the systems that analysts use to analyse the data; (3) *organisational enablers* relate to the qualities of analysts themselves, and how analysis is organisationally structured; (4) *process-based enablers* relate to the sensemaking process.

The contribution of this research is, not only, detailed insights into individual intelligence analysis processes of our case study, but also a generalised sensemaking process description with new additions compared to earlier literature. Also new to the literature is a study that lists enablers of sensemaking in an information science context. Practically these enablers, together with the recommendations we have made, can help improve sensemaking at the Dutch Police, as well as at other similar organisations.

References

[1] Shahriar Akter et al. *Algorithmic bias in data-driven innovation in the age of AI*. 2021.

[2] Zaher Ali Al-Sai, Rosni Abdullah, and Mohd Heikal Husin. “Critical success factors for big data: a systematic literature review”. In: *IEEE Access* 8 (2020), pp. 118940–118956.

[3] Christopher Andrews, Alex Endert, and Chris North. “Space to think: large high-resolution displays for sensemaking”. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2010, pp. 55–64.

[4] Simon Attfield and Chris Baber. “Elaborating the frames of data-frame theory”. In: *13th International conference on Naturalistic Decision Making*. The University of Bath. 2017, pp. 25–32.

[5] Emma Caroline Barrett. “The interpretation and exploitation of information in criminal investigations”. PhD thesis. University of Birmingham, 2009.

[6] Andrew D Brown, Ian Colville, and Annie Pye. “Making sense of sensemaking in organization studies”. In: *Organization studies* 36.2 (2015), pp. 265–277.

[7] George Chin Jr, Olga A Kuchar, and Katherine E Wolf. “Exploring the analytical processes of intelligence analysts”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2009, pp. 11–20.

[8] Tsan-Ming Choi, Stein W Wallace, and Yulan Wang. “Big data analytics in operations management”. In: *Production and Operations Management* 27.10 (2018), pp. 1868–1883.

[9] Emma Afua Ameng Codjoe, Celestine Ntuen, and Jules Chenou. “A case study in sensemaking using Data/Frame Model”. In: *IIE Annual Conference. Proceedings*. Institute of Industrial and Systems Engineers (IISE). 2010, p. 1.

[10] Mandeep K Dhami, Ian K Belton, and David R Mandel. “The “analysis of competing hypotheses” in intelligence analysis”. In: *Applied Cognitive Psychology* 33.6 (2019), pp. 1080–1090.

[11] Corey K Fallon et al. “The calibration of trust in an automated system: A sensemaking process”. In: *2010 International Symposium on Collaborative Technologies and Systems*. IEEE. 2010, pp. 390–395.

[12] Jing Gao, Andy Koronios, and Sven Selle. “Towards a process view on critical success factors in big data analytics projects”. In: (2015).

[13] Pascal Gemke et al. “Towards a maturity model for intelligence-led policing A case study research on the investigation of drugs crime and on football and safety in the Dutch police”. In: *Police practice and research* 22.1 (2021), pp. 190–207.

[14] Jean Helms Mills, Amy Thurlow, and Albert J Mills. “Making sense of sensemaking: the critical sensemaking approach”. In: *Qualitative research in organizations and management: An international journal* 5.2 (2010), pp. 182–195.

[15] Marielle den Hengst-Bruggeling, Bart De Graaf, and Peter Van Scheepstal. “Modelling intelligence-led policing to identify its potential.” In: *Journal of Police Studies/Cahiers Politiestudies* 1.3 (2013).

[16] Richards J Heuer. *Psychology of intelligence analysis*. Center for the Study of Intelligence, 1999.

[17] Darren Hudson and Gurmeet Singh. “Expert medical decision-making: how the data-frame theory can explain physician sense-making.” In: *ITCH*. 2017, pp. 167–171.

[18] Shaoxiong Ji et al. “A survey on knowledge graphs: Representation, acquisition, and applications”. In: *IEEE transactions on neural networks and learning systems* 33.2 (2021), pp. 494–514.

[19] Sheriff Jolaoso, Russ Burtner, and Alex Endert. “Toward a deeper understanding of data analysis, sensemaking, and signature discovery”. In: *Human-Computer Interaction–INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings, Part II 15*. Springer. 2015, pp. 463–478.

[20] Ruogu Kang, Aimee Kane, and Sara Kiesler. “Teammate inaccuracy blindness: when information sharing tools hinder collaborative analysis”. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 2014, pp. 797–806.

[21] Youn-ah Kang and John Stasko. “Characterizing the intelligence analysis process: Informing visual analytics design through a longitudinal field study”. In: *2011 IEEE conference on visual analytics science and technology (VAST)*. IEEE. 2011, pp. 21–30.

[22] Katherine P Kaste. “Naturalistic Study Examining the Data/Frame Model of Sensemaking by Assessing Experts in Complex, Time-Pressured Aviation Domains”. In: (2012).

[23] Gary Klein, Brian Moon, and Robert R Hoffman. “Making sense of sensemaking 1: Alternative perspectives”. In: *IEEE intelligent systems* 21.4 (2006), pp. 70–73.

[24] Gary Klein, Brian Moon, and Robert R Hoffman. “Making sense of sensemaking 2: A macrocognitive model”. In: *IEEE Intelligent systems* 21.5 (2006), pp. 88–92.

[25] Gary Klein, Sterling Wiggins, and Cynthia O Dominguez. “Team sensemaking”. In: *Theoretical Issues in Ergonomics Science* 11.4 (2010), pp. 304–320.

[26] Gary Klein et al. “A data–frame theory of sensemaking”. In: *Expertise out of context*. Psychology Press, 2007, pp. 118–160.

[27] Neesha Kodagoda et al. “Using machine learning to infer reasoning provenance from user interaction log data: based on the data/frame theory of sensemaking”. In: *Journal of Cognitive Engineering and Decision Making* 11.1 (2017), pp. 23–41.

[28] Lisa Krizan. *Intelligence essentials for everyone*. 6. Joint Military Intelligence College, 1999.

[29] Stéphane Lefebvre. “A look at intelligence analysis”. In: *International Journal of Intelligence and CounterI* 17.2 (2004), pp. 231–264.

[30] Mark Lycett and Alaa Marshan. “Capturing sensemaking pattern during data analysis: A conceptual framework”. In: (2016).

[31] Iman Malaka and Irwin Brown. “Challenges to the organisational adoption of big data analytics: A case study in the South African telecommunications industry”. In: *Proceedings of the 2015 annual research conference on South African institute of computer scientists and information technologists*. 2015, pp. 1–9.

[32] Stathis Malakis and Tom Kontogiannis. “A sensemaking perspective on framing the mental picture of air traffic controllers”. In: *Applied ergonomics* 44.2 (2013), pp. 327–339.

[33] Charles A Mangio and Bonnie J Wilkinson. “Intelligence analysis: Once again”. In: (2008).

[34] Margaret Mitchell et al. “Diversity and inclusion metrics in subset selection”. In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. 2020, pp. 117–123.

[35] David T Moore and Robert R Hoffman. “Data-frame theory of sensemaking as a best model for intelligence”. In: *American Intelligence Journal* 29.2 (2011), pp. 145–158.

[36] Eirini Ntoutsi et al. “Bias in data-driven artificial intelligence systems—An introductory survey”. In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10.3 (2020), e1356.

[37] Briony J Oates, Marie Griffiths, and Rachel McLean. *Researching information systems and computing*. Sage, 2022.

[38] Ahmed Oussous et al. “Big Data technologies: A survey”. In: *Journal of King Saud University-Computer and Information Sciences* 30.4 (2018), pp. 431–448.

[39] Sharoda A Paul and Madhu C Reddy. “Understanding together: sensemaking in collaborative information seeking”. In: *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 2010, pp. 321–330.

[40] Peter Pirolli and Stuart Card. “The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis”. In: *Proceedings of international conference on intelligence analysis*. Vol. 5. McLean, VA, USA. 2005, pp. 2–4.

[41] Peter Pirolli and Daniel M Russell. *Introduction to this special issue on sensemaking*. 2011.

[42] Yan Qu and George W Furnas. “Model-driven formative evaluation of exploratory search: A study under a sensemaking framework”. In: *Information Processing & Management* 44.2 (2008), pp. 534–555.

[43] Yan Qu and Derek L Hansen. “Building shared understanding in collaborative sensemaking”. In: *Proceedings of CHI 2008 Sensemaking Workshop*. 2008.

[44] Jerry Ratcliffe. “Intelligence-led policing”. In: (2003).

[45] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. “" Why should i trust you?" Explaining the predictions of any classifier”. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 2016, pp. 1135–1144.

[46] Daniel M Russell et al. “Sensemaking in a senseless world: 2018 workshop abstract”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, pp. 1–7.

[47] Daniel M Russell et al. “The cost structure of sensemaking”. In: *Proceedings of the INTERACT’93 and CHI’93 conference on Human factors in computing systems*. 1993, pp. 269–276.

[48] Feliks Prasepta Sejahtera et al. “Enablers and inhibitors of effective use of big data: insights from a case study”. In: Association for Information Systems. 2018.

[49] Hannah Snyder. “Literature review as a research methodology: An overview and guidelines”. In: *Journal of business research* 104 (2019), pp. 333–339.

[50] Lygia Stewart, Cynthia O Dominguez, and Lawrence W Way. “A data-frame sensemaking analysis of operative reports: Bile duct injuries associated with laparoscopic cholecystectomy”. In: *Informed by Knowledge*. Psychology Press, 2011, pp. 343–352.

[51] Alex Tamkin et al. “Understanding the capabilities, limitations, and societal impact of large language models”. In: *arXiv preprint arXiv:2102.02503* (2021).

[52] Yihan Tao and Anastasios Tombros. “How collaborators make sense of tasks together: A comparative analysis of collaborative sensemaking behavior in collaborative information-seeking tasks”. In: *Journal of the Association for Information Science and Technology* 68.3 (2017), pp. 609–622.

[53] David R Thomas. “A general inductive approach for qualitative data analysis”. In: (2003).

[54] Alice Toniolo et al. “Human-machine collaboration in intelligence analysis: An expert evaluation”. In: *Intelligent Systems with Applications* 17 (2023), p. 200151.

[55] Candice Walls and Brian Barnard. “Success factors of Big Data to achieve organisational performance: Theoretical perspectives”. In: *Expert Journal of Business and Management* 8.1 (2020).

[56] QR Waraich. “Application of Sensemaking: Data/Frame Model, to UAS AIB reports can increase UAS GCS resilience to Human Factor and Ergonomics (HF/E) shortfalls”. In: *Sensemaking in Safety Critical and Complex Situations*. CRC Press, 2021, pp. 209–234.

[57] Karl E Weick. *Sensemaking in organizations*. Vol. 3. Sage, 1995.

[58] BL William Wong and Margaret Varga. “Black holes, keyholes and brown worms: Challenges in sense making”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 56. 1. SAGE Publications Sage CA: Los Angeles, CA. 2012, pp. 287–291.

[59] William Wong et al. “INVISQUE: intuitive information exploration through interactive visualization”. In: *CHI’11 Extended Abstracts on Human Factors in Computing Systems*. 2011, pp. 311–316.

[60] William Wright et al. “The Sandbox for analysis: concepts and methods”. In: *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 2006, pp. 801–810.

[61] Liang Yao et al. “Exploring large language models for knowledge graph completion”. In: *arXiv preprint arXiv:2308.13916* (2023).

[62] Robert K Yin. *Case study research: Design and methods*. Sage Publications, 2013.

[63] Pengyi Zhang and Dagobert Soergel. "Cognitive mechanisms in sensemaking: A qualitative user study". In: *Journal of the Association for Information Science and Technology* 71.2 (2020), pp. 158–171.

[64] Pengyi Zhang and Dagobert Soergel. "Examining a comprehensive sensemaking model with user studies of computer-assisted sensemaking". In: *Sensemaking Workshop at CHI*. Vol. 2009. 2009.

[65] Pengyi Zhang and Dagobert Soergel. "Process patterns and conceptual changes in knowledge representations during information seeking and sensemaking: A qualitative user study". In: *Journal of Information Science* 42.1 (2016), pp. 59–78.

[66] Pengyi Zhang and Dagobert Soergel. "Towards a comprehensive model of the cognitive process and mechanisms of individual sensemaking". In: *Journal of the Association for Information Science and Technology* 65.9 (2014), pp. 1733–1756.

[67] Pengyi Zhang et al. "Extending sense-making models with ideas from cognition and learning theories". In: *Proceedings of the American Society for Information Science and Technology* 45.1 (2008), pp. 23–23.

[68] Lingfeng Zhong et al. "A comprehensive survey on automatic knowledge graph construction". In: *arXiv preprint arXiv:2302.05019* (2023).

A Case Study Protocol

Overview The purpose of the case study is to get insight into the intelligence analysis operation at the Dutch National Police. We are interested in getting information about the working method of the analysts at the Dutch Police. The case study is of exploratory nature, where its results are used to model the relevant processes.

The objective is to get insight into the following aspects of intelligence analysis:

- Which actors are relevant in the process of intelligence analysis?
- What kind of analysis tasks does an actor do?
- What are the activities in an analysis task?
- What is the starting and end point for a single analysis task?
- What systems are used for the analysis tasks?
- With who do actors interact in intelligence analysis?

Procedures. Together with my supervisor at the Dutch Police, a set of interviewees with experience will be selected. These potential participants will be sent an invitation to the case study. Interviews will take about one hour and can be conducted through online meeting software (Teams) or on-site. Before the interview, the participant is emailed with some final details about the interview. This includes an information sheet about the research and a consent form. The consent form and information sheet might be slightly personalised for the receiver but uses Appendix C and D as templates respectively.

The audio of the interview will be recorded, in order to transcribe it later. The interview protocol is included in Appendix B. The transcription will occur within seven months after which the recordings will be deleted. The transcription will also be anonymised at that time. Parts of the recording that are clearly not relevant to this research will be omitted from the transcription process.

Coding Inductive coding will be used for the transcripts, which is described in the method section. The coding will follow the *general inductive approach* by Thomas [53]. NVivo 12 will be used for coding. The themes will be induced and inputted into NVivo. During the process, more themes will emerge, or be further refined. Some themes can also be merged.

Case Study Database The transcripts and recordings are stored on Utrecht Universities secured cloud storage, in a file directory structure. The interviewee will be numbered, and the related files will use a naming scheme including this identifier to order the resulting raw data. Later findings from the research can be linked to the raw-data in this case study database.

B Interview Protocol

This section will describe how the interviews are executed. Note that the interviews are classified as *semi-structured interviews* [37]. This method allows for flexibility and continuing with unscripted questions when further elaboration is wished. The interview is divided into three sections, and is conducted as follows:

B.1 Introduction

Approx. 5 min
In this section we introduce the topic of the research, and describe the research method, thus explaining why the interview is being conducted. The goals of the research are also told to the participant. It is furthermore confirmed that informed consent has been given, so the recording can be initiated. Most information described in the introduction is already provided in an information sheet that has been emailed to the participant beforehand (see Appendix D).

B.2 Informing About Job

Approx. 15 min
In order to get a better understanding of intelligence in general, and the job of the interviewee, some basic questions are asked about their work for the Dutch Police. Another goal is to get to know at which team and department to interviewee works, and how the team is positioned in the organisation.

- What is your job title at the Police, and what are your main activities in that role?
  - Are these also your day-to-day activities?
- Of which team, unit, and department are you part?
  Ask participants to further clarify the team and department and ask about this team compares to earlier found departments. Also, ask if the interviewee collaborates with these other departments.
- What kind of intelligence analysis tasks do you work on?

B.3 Information About Analysis Tasks

Approx. 35 min
In the next and primary phase, questions are asked regarding intelligence analysis. As described in the related works (see Section 3) we can expect the process to be ad hoc and non-linear [21, 28, 63, 7]. The strategy to still get a description from the interviewee with distinct activities is to first ask them for describing the main phases and continuing with more descriptive questions about the resulting activities. Comparable to our motivation to use an inductive coding method for the analysis of the resulting transcriptions from the interviews, limited literature is used for the interview questions. Our goal is to get an accurate process description of the current intelligence analysis processes at the Dutch National Police. Therefore for questions related to the process itself, we do not use a basis from the processes described in earlier work.

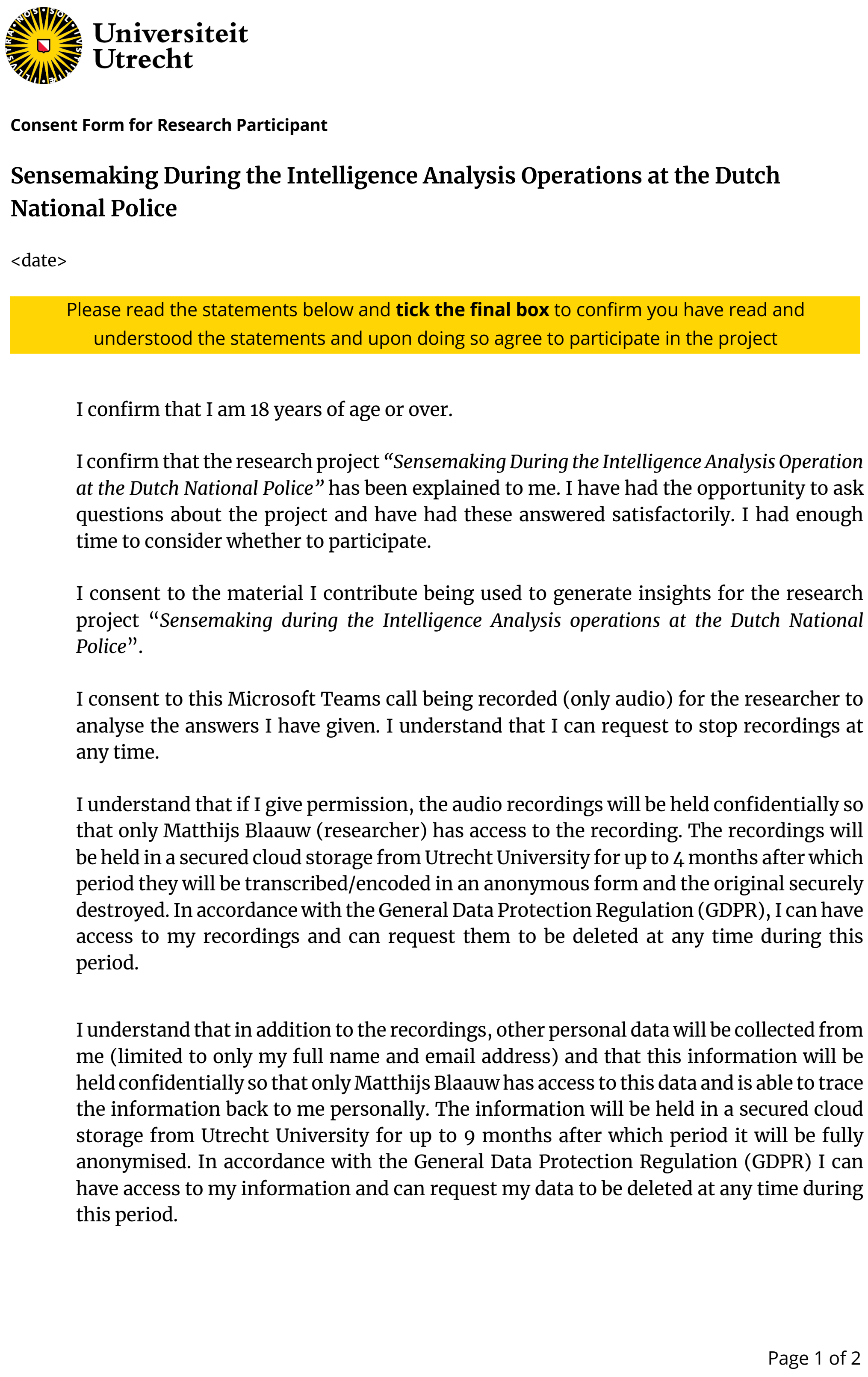
- What marks the start and end of an analysis project?
  - How long does this process take?
- What are your tasks at the start of the intelligence project?
- When you start with an analysis task, how complete is the description of the task?
  Earlier research describes that intelligence analysts have to work with ill-defined tasks descriptions, making the understanding of the task a crucial part of sensemaking [67].
  - How do you go about creating an understanding of the analysis task?
  - How long does this process take?

- What are the main activities in intelligence analysis?
    - Can you elaborate on *activity x*? What do you do during *activity x*?
    - What computer software do you use for *activity x*?
    - What do you do in *software program y*?
    - Do you collaborate with other colleagues during *activity x*?
      These questions are asked for each of the activities described by the analyst.
- How do you make sense of information during intelligence analysis?
  This question is asked to get more information about the general processes in an analysts mind. These can for example be *hypothesis Generation* or *structured Information Search* [66, 24].
    - What software tool do you use to aid with *cognitive process x*?
- When searching for information during intelligence operations is it clear what you are searching for? Are some searches more of exploratory nature? How does this differentiate in your activities?
  This question is derived from the distinction in existing literature between exploratory search and focussed search [66, 40].

B.4 Concluding Remarks

Approx. 5 min
Lastly, the interviewee is thanked for their time. A sub-set of interviewees will also be asked if they are available for a validation interview. In this validation interview, the interviewee is shown the BPMN process model, created based on the information provided in the interview with them and with comparable analysts.

C Consent Form



Universiteit Utrecht

Consent Form for Research Participant

Sensemaking During the Intelligence Analysis Operations at the Dutch National Police

<date>

Please read the statements below and **tick the final box** to confirm you have read and understood the statements and upon doing so agree to participate in the project

I confirm that I am 18 years of age or over.

I confirm that the research project “*Sensemaking During the Intelligence Analysis Operation at the Dutch National Police*” has been explained to me. I have had the opportunity to ask questions about the project and have had these answered satisfactorily. I had enough time to consider whether to participate.

I consent to the material I contribute being used to generate insights for the research project “*Sensemaking during the Intelligence Analysis operations at the Dutch National Police*”.

I consent to this Microsoft Teams call being recorded (only audio) for the researcher to analyse the answers I have given. I understand that I can request to stop recordings at any time.

I understand that if I give permission, the audio recordings will be held confidentially so that only Matthijs Blaauw (researcher) has access to the recording. The recordings will be held in a secured cloud storage from Utrecht University for up to 4 months after which period they will be transcribed/encoded in an anonymous form and the original securely destroyed. 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.

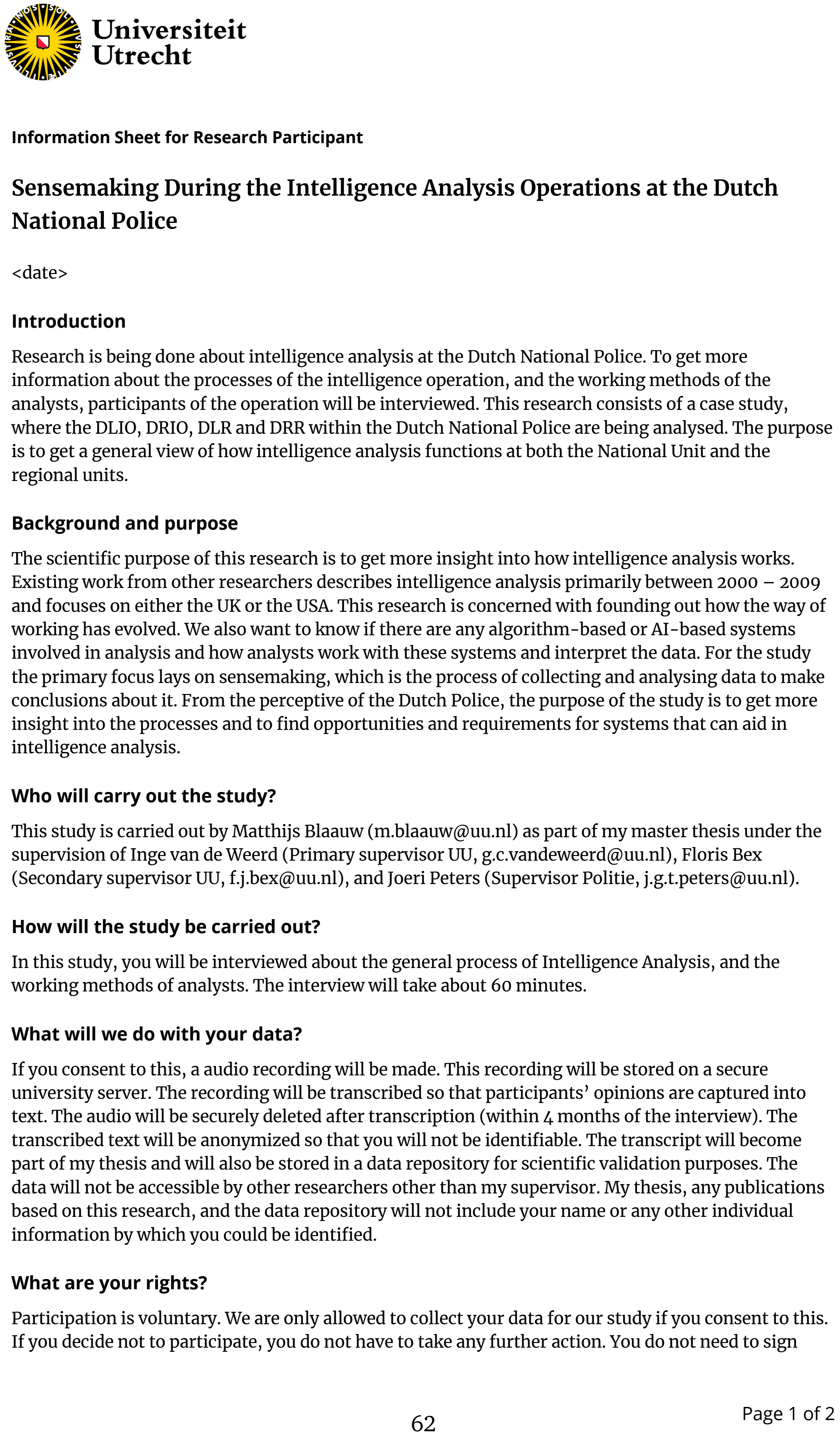
I understand that in addition to the recordings, other personal data will be collected from me (limited to only my full name and email address) and that this information will be held confidentially so that only Matthijs Blaauw has access to this data and is able to trace the information back to me personally. The information will be held in a secured cloud storage from Utrecht University for up to 9 months after which period it will be fully anonymised. In accordance with the General Data Protection Regulation (GDPR) I can have access to my information and can request my data 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 allow the fully anonymized data to be used in future publications and other scholarly means of disseminating the findings from the research project.

□ I confirm that I have read and understood the above statements and agree to participate in the study (Check the box).

D Information Sheet



Universiteit
Utrecht

Information Sheet for Research Participant

Sensemaking During the Intelligence Analysis Operations at the Dutch National Police

<date>

Introduction

Research is being done about intelligence analysis at the Dutch National Police. To get more information about the processes of the intelligence operation, and the working methods of the analysts, participants of the operation will be interviewed. This research consists of a case study, where the DLIO, DRIO, DLR and DRR within the Dutch National Police are being analysed. The purpose is to get a general view of how intelligence analysis functions at both the National Unit and the regional units.

Background and purpose

The scientific purpose of this research is to get more insight into how intelligence analysis works. Existing work from other researchers describes intelligence analysis primarily between 2000 – 2009 and focuses on either the UK or the USA. This research is concerned with founding out how the way of working has evolved. We also want to know if there are any algorithm-based or AI-based systems involved in analysis and how analysts work with these systems and interpret the data. For the study the primary focus lays on sensemaking, which is the process of collecting and analysing data to make conclusions about it. From the perceptive of the Dutch Police, the purpose of the study is to get more insight into the processes and to find opportunities and requirements for systems that can aid in intelligence analysis.

Who will carry out the study?

This study is carried out by Matthijs Blaauw (m.blaauw@uu.nl) as part of my master thesis under the supervision of Inge van de Weerd (Primary supervisor UU, g.c.vandeweerd@uu.nl), Floris Bex (Secondary supervisor UU, f.j.bex@uu.nl), and Joeri Peters (Supervisor Politie, j.g.t.peters@uu.nl).

How will the study be carried out?

In this study, you will be interviewed about the general process of Intelligence Analysis, and the working methods of analysts. The interview will take about 60 minutes.

What will we do with your data?

If you consent to this, a audio recording will be made. This recording will be stored on a secure university server. The recording will be transcribed so that participants' opinions are captured into text. The audio will be securely deleted after transcription (within 4 months of the interview). The transcribed text will be anonymized so that you will not be identifiable. The transcript will become part of my thesis and will also be stored in a data repository for scientific validation purposes. The data will not be accessible by other researchers other than my supervisor. My thesis, any publications based on this research, and the data repository will not include your name or any other individual information by which you could be identified.

What are your rights?

Participation is voluntary. We are only allowed to collect your data for our 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. You will even be able to withdraw your consent after you have participated. However, if you choose to do so, we will not be required to undo the processing of your data that has taken place up until that time. The personal data we have obtained from you up until the time when you withdraw your consent will be erased (where personal data is any data that can be linked to you, so this excludes any already anonymized data).

More information about this study?

If you have any questions or concerns about this research, please contact me (Matthijs Blaauw) at m.blaauw@uu.nl or my first supervisor Inge van de Weerd at g.c.vandeweerd@uu.nl.

E Code Book

Inductive coding has been conducted on the transcripts of the 17 interviews, following the *general inductive approach* by Thomas [53]. The resulting code is discussed in this section. The themes and it's sub-themes are presented in Table 4, each with a description and example of the coded text.

The coding is a maximum of three levels deep, of which one can be the embedded case. For example, the top-level theme of *analysis products* has three 2nd-level sub-themes, that represent our embedded cases *investigation analysis*, *security analysis*, and *strategic Analysis*. The 3rd-level sub-theme represent each analysis product found for the case. To improve conciseness of the Code Book only one example is included for themes that are three levels deep. This devision into three levels, with the embedded case a subdivider, allowed us to more easily overview which themes were found for each embedded process during our within-case analysis. Lastly, when the theme is two levels deep, this 2nd layer may be the embedded cases, in that case, the embedded cases are not included in the code book as sub-themes.

Theme	Description	Example
Analysis Duration	The duration of an (typical) analysis project, or the greater project in which the analysis takes place.	"Yes, that highly differs. It can be two years, it can also be half a year or a few months. It depends on what it is" (IV01)
Analysis Goals	A concise description of the primary goal of the analysis level.	"Because strategic analysis has the goal of obtaining an overview of the nature and extend of a certain criminal phenomenon, with the goal to make policy choices" (IV16)
Analysis Phases Analysis projects, or the greater project the analysis is assisting in can have multiple phases. For each of the cases we were able to differentiate one, two, or three phases in a project.		
Investigation Analysis Phases	Found phases are: *Starting phase*, *middle phase*, and *end phase*.	"you will see that in the last phase, and then i am talking about when the suspects have been arrested and it comes to the finalising the file, then it also ends a bit for the analyst. This applies to both analysts." (IV02) [coded as *end phase*]
Security Analysis Phases	Found phases are: *Starting phase*.	"When you're doing an analysis It is of course important that you have the purpose of the analysis clearly in mind. So what is the question, and based on the question you will look what the analysis requires and what kind of systems you will use for that." (IV15) [coded as *starting phase*]
Strategic Analysis Phases	Found phases are: *Starting phase* and *main phase*.	"...this resulted in a number of questions, and we subsequently set up a process that we deployed with a project team, a small research team. And every 6 weeks we provided feedback to the steering group." (IV17) [coded as *main phase*]
Analysis Products The products that the analyst creates with their analysis process. Each product is coded as 3rd-level sub-themes.		
Investigation Analysis Products	Found products are: *Hypothesis and Scenarios*, *Information Need*, *Insights*, *Overviews*, *Relational Schemas*, *Timelines*.	"… that you will also come up with hypothesis, that you not only record what you have seen, but also do the next step and come with a hypothesis of what could be going on." (IV02) [Coded as *hypothesis and Scenarios*]

Security Analysis Products	Found products are: *Phenomena Overviews*, *Network analysis*, *Trends Analysis*, *Targets*, *Knowledge Documents*.	"… i just happen to have it here … The phenomenon overview drugs. Fortunately I don't make these kinds of reports all by myself. But it's about, gee, what is the nature and size of the drug market? Specific to me was the synthetic drugs market, and how it is divided into certain main groups." (IV15) [Coded as *phenomena overviews*]
Strategic Analysis Products	Found products are: *Phenomena Overviews* and *Scientific Papers*	"… i wrote an article and i think it is in Global Crime, the English version, and also in the Dutch version, about what organised crime in The Netherlands looks like." (IV16) [Coded as *Scientific Papers*]
Software Used by Analysts	The software used by analysts for their analysis processes. Included are statements if they use the system, what functionality the systems provide, and how the analysts use the systems. Found software, coded as sub-themes, are: *BlueBase*, *BlueSpot Monitor*, *BlueView*, *COGNOS*, *Data Warehouse*, *DCS*, *Excel*, *Full Contact*, *iBase and Analyst Notebook*, *MindManager*, *Raffinaderij*, *PowerPoint*, *SummIT*, and *Topic View*.	"Well, we don't work with BVH. We work with SummIT. SummIT is the detective environment of the police systems. That is where I get my information from and the underlying documents that have been registered. So that is a lot of reading, and that is where I get the data from." (IV02) [Coded as *BVH* and *SummIT*]
Initiation of Analysis	This theme contains descriptions of how an analysis task starts.	"If I get the question from <name>, can you look into the crypto data […], can you do an analysis on people who order drugs in Gibraltar?" (IV15)
Analysis Activities This theme contains descriptions of the activities performed by the analyst that relate to the analysis process. Each of the activities found are coded as 3rd level sub-themes.		
Investigation Analysis Activities	Found activities are: *Analysing*, *Hypothesis Generation*, *Monitoring Information Flow*, *Prioritising*, *Reading*, *Reporting*, *Schematising*, *Searching*. and *Telecom Analysis*.	"Wile you are reading, I will try to make it more clear. I often start making a diagram in Analist Notebook, so that you can immediately plot the people and the phone numbers and addresses, so you know what belongs together." (IV02) [coded as *Schematising*, and as iBase and Analyst Notebook (sub-theme of *software used by analysts*)]
Security Analysis Activities	Found activities are: *Analysing Data*, *Clarifying*, *Exploiting other sources*, and *Querying*.	"Also, the tactical analyst has networks, so network analysis, where you map the network surrounding the subject, and then also look at what is the network doing? What are the opportunities within tapped conversations, what are the greatest opportunities to tackle a network, to frustrate a network" (IV10) [Coded as *Analysing Data*]
Strategic Analysis Activities	Found activities are: *Clarifying*, *Identifying Sources*, *Searching*, *Filtering*, *Importing*, *Structuring*, and *Analysing Data*.	"You just see that the amount of data is increasing, so you have to somehow ensure that reading all the information is redundant to a certain extent. You can make a sharp selection in a way and say, this is the relevant piece and you are only really going to read that." (IV17) [Coded as *filtering* and *reading*]
Investigation Analysis Specifics This theme consists of sub-themes that are important details, related to the investigation analysis process.		
The analyst is all-knowing about investigation	Descriptions of the all-knowing role of the analysts in the team.	"… they come to you straight away because you are one of the few, besides the team leader, who has all the detailed knowledge of the investigation" (IV04)

<table>
<tr><td>Attachment of analysts to cases</td><td>How investigation analysts (operational or tactical) are assigned to investigation teams</td><td>"Well, as long as the investigation runs from the beginning, I am on the team… Until the end of the investigation." (IV01)</td></tr>
<tr><td colspan="3"><b>Collaboration</b>
The following sub-themes have been coded regarding the collaboration of analysts with others</td></tr>
<tr><td>Collaboration between analysts</td><td>If and how analysts collaborate with each other for their analysis.</td><td>"I like it myself, because your reliability just increases a little when you do things together. If you do it all on your own, there is always a risk that you will miss things or make mistakes." (IV16)</td></tr>
<tr><td>Collaboration with information producers</td><td>The theme includes mentions of if and how the analyst interacts with the information producers in their analysis or investigation team</td><td>"Plus you share a room together, so the lines of communication are very short, so every time you have an idea, you throw it out straight away." (IV04)</td></tr>
<tr><td colspan="3"><b>Security Analysis Specifics</b>
This theme consists of sub-themes that are important details related to the security analysis process.</td></tr>
<tr><td>Monitoring Roles</td><td>The roles who monitor the incoming information flow.</td><td>"As an analyst, I also keep track of the workflow here, which means that I also have to minitor what gets added to SummIT on a daily basis. We have certain scripts running so that we receive an Excel file from BVH every day containing the most common drug-related incidents of the past 24 hours." (IV10) [Also coded as <i>Information Flow</i>]</td></tr>
<tr><td>Information Flow</td><td>The functions and design of the automated information flow within security analysis team</td><td>"… so we can actually proactively query our systems, which we do every day to see which possible signals of human trafficking we have missed. Because human trafficking can also be hidden in many other situations that does not always point out human trafficking at the front" (IV14)</td></tr>
<tr><td colspan="3"><b>Data Sources</b>
This theme contains the sources that are used for analysis on each analysis level. Each of the sources found are coded as 3rd level themes.</td></tr>
<tr><td>Investigation Analysis Sources</td><td>Found sources are: <i>BVIB</i>, <i>Data from detectives</i>, <i>BVH</i>, <i>Crypto</i>, and <i>BlueView</i></td><td>"Yes the crypto messages, those are millions of messages. Yes, that is a lot of information that I have at my disposal. And as a investigation analyst you don't yet have the tools to understand search through these easily" (IV01) [Coded as <i>Crypto</i>]</td></tr>
<tr><td>Security Analysis</td><td>Found sources are: <i>BlueView</i>, <i>BVH</i>, <i>Crypto</i>, <i>GBA</i>, <i>KVK</i>, <i>OSINT</i>, and <i>SummIT</i></td><td>"We use BVH, SummIT, TopicView, OSINT. Basically everything that is available within the police from the source systems." (IV13) [Coded as <i>BVH</i>, <i>SummIT</i>, <i>TopicView</i>, and <i>OSINT</i>]</td></tr>
<tr><td>Strategic Analysis</td><td>Found sources are: <i>BVH</i>, <i>Interviews</i>, <i>Partners</i>, <i>SummIT</i>, <i>OSINT</i>, and <i>Crypto</i></td><td>"We also often conduct interviews with team leaders, but also with others, because not everything is in the systems" (IV16) [Coded as <i>Interviews</i>]</td></tr>
<tr><td colspan="3"><b>Organisational Specifics</b>
These themes contain specifics on how the analysts are positioned in the organisation, and how to parent analysis teams are organised at the Dutch Police. It can also contain information on how the regional units differ from the national units.</td></tr>
<tr><td>Investigation Analysis Organisation</td><td>Explanations of how investigations and their analysts are positioned in the organisation. The sub-theme also includes explanations of differences between national an regional investigations</td><td>"Yes, that is actually the case when we are deployed as analysts. In general you are not tied to a team. It often happens automatically, because it is logical that you stay seated. But officially we are flexible." (IV03)</td></tr>
</table>

Security Analysis	Explanations of how security analysis and their analysts are positioned in the organisation. The sub-theme also includes explanations of differences between different thematic-teams and differences between national an regional investigations	"We're now at intelligence, we are in the <street name> building. But intelligence works together with the tactical investigation department of CTER. They are at <street name> and they purely do investigations" (IV12)
Strategic Analysis	Explanations of how strategic analysis and their analysts are positioned in the organisation.	"We are with about 15 researchers, we work at the National Unit, but there is a network, and I don't know exactly how many researchers work in scientific research within the Dutch police, but there are 100." (IV16)
Other roles analysts interact with The analysis projects often occur in teams of people or occurs in collaboration with others. In this theme, the other roles of actors an analyst interacts with have been coded. Each of the roles are coded as a 3rd-level theme. Included are descriptions of these roles, and how the analysts interact with these actors..		
Investigation Analysis	Found other roles are: *Detective*, *Specialists*, *Team Leader*, *Information Detective*, and *Tactical Coördinator*.	"So In the collecting phase I mainly work together with the information coordinator and the tactical coordinator. The Information Coordinator then ensures that the correct information is available, and he also determines the course." (IV03) [coded as *information coordinator* and *tactical coordinator*]
Security Analysis	Found other roles are: *Intelligence Employees*, *Information Detectives*, and *Team leader*.	"The only thing I could do, for example, is ask a question to an information detective, to for example, walk through a certain family tree and tell me, well, how big is that family, what is it like, who are the people within that family, so that as an analyst I don't have to worry about that. So I can actually only focus on the people." (IV10) [coded as *Information Detectives*]
Strategic Analysis	Found other roles are: *Portfolio Holders*, *Research Coordinator*, *Steering Committee*.	"Yes, we have a head of analysis and research and there must also be a research coordinator. But we actually work very independently. Because it is generally too difficult for them to make choices about what you should or should not do. So they are there to help you if you have too much work or too little work. But you usually arrange that yourself." (IV16) [coded as *Research Coordinator*]
Security Analysis Monitoring Process	Descriptions of the activities in the security analysis monitoring process. Found activities are: *Contacting Source*, *Searching*, *Filtering*, *Reading*, and *Structuring*.	""We then process that in SummIT. So, for example, of the 80 registrations that come in, we have 4 that meet the criteria, then those 4 are added to the SummIt environment as four separate mutations" (IV10) [coded as *filtering* and *structuring*]"
Freedom in Analysis	Descriptions of the freedom an analyst has in determining what they will be analysing, as well the freedom in the method of performing an analysis.	"So we as Rotterdam have the financial sub-theme. So yes, together with the working group we decide what needs to happen. And that is on our own initiative, where I play a bit of a driving role from the analysis point of view. Because I see what certain information can offer to our analysis. [...] So yes, partially own initiative, but also what sub-theme has been decided on a nation level " (IV11)
Enablers This theme includes the enablers found in our case-study. Each of the enablers are coded as their own sub-theme.		

Clear input rules	If and why input rules are experienced as beneficial for analysis.	"So, for example, such a subject line in SummIT. You can enter a lot of free text there, which everyone also does. But we said, no, we use that space to have a kind of property line generated by our scoring form, whereby we can derive all kinds of information based on that subject line" (IV14)
Automated searching Tools	If and why automated searching tools are experienced as beneficial for analysis.	"We have those 70+ queries that are still running today and that already provided hidden signals of human trafficking that serve as input for our process in which, yes, investigations are also regularly initiated because a missed signal comes from such a query, which is so distressing that immediate action must be taken." (IV14)
Availability of rich data sources	If and why having more data, or a wider range of data sources is experienced as beneficial for analysis.	"A lot more is becoming possible. Regardless of whether it is on investigation level or strategic. But if you look at, say, the data that can be made available to you and how current it is, it's really beautiful. So in that sense it becomes easier, or easier, I don't think it's the right word... But more is possible so that is really beneficial." (IV11)
Easy data sharing	If and why data sharing is experienced as beneficial for analysis.	"Also because you have much more data at your disposal. Because you work together more, all investigations are actually open because you have to share information, so get more data." (IV01)
Data source mutability	If and why it is important that corrections which were found in local databases are put back in the source systems.	"So you would have to continue with your analysis in iBase, so you can indeed add that relationship yourself. You will put it in iBase and not the source, so you can continue with your analysis. Of course you would have to make sure that it will also be put back in to the source. But in practice, a large part is not reflected in the relationships. Entities will generally be made, but you will really miss those relationships." (IV07)
Regularly using systems	If and why it is important to use systems regularly to keep dexterity with analytical software.	"I had a Refinery training two weeks ago. After using it twice, i haven't used it. If I will use it now, i wouldn't know how it works, and that is how it starts. If you are 3 or 4 months further, then you've actually lost everything they told you back then." (IV07)
Collaboration between sensemakers	If and why collaboration between analysts is experienced as beneficial for analysis.	"and what is important as an analyst, you really have to actively approach other analysts to ask them would you like to think along with me about my method? Are you by any chance already working on that? and how do you approach it? Then perhaps we should work together on this." (IV12)
Collaboration between information producers	If and why collaboration between the analyst and information producers is experienced as beneficial for analysis.	"That is why it was chosen within the national investigation unit, that analysts are also placed in the team itself. So that you also have constant contact with the other members of the team, the detectives. So if they have something they come to me. And if I have a specific question, I can ask them. Then you have very short lines, that is actually the intention." (IV01)

Domain knowledge and specialisation	If and why specialisation is experienced as beneficial for analysis.	"I come from the CTER domain, where I see how difficult it is to get people to classify things right. You certainly shouldn't ask people to do the same for 5 different markets. And above all, do not ask to create relationships during the intake of the report, for example, is it a friendly or a hostile relationship? What are the cross-connections? That is far too much to ask of the people who will do the work in the real world." (IV13)
Availability of data catalogues	If and why knowing the which sources are available is experienced as beneficial for analysis.	"Yes, that is the difficult part, because I looked for a source document for this interview, but there is none, so I have made a list here very quickly with all the sources I used in the past." (IV16)
Technical Skills	If and why technical skills is experienced as beneficial for analysis.	"I find it het biggest challenge to keep up with all the tools and systems that are involved with everything. The same holds for other systems we work with. This is also a big obstacle." (IV05)
Analysis Skills	If and why analysis skills is experienced as beneficial for analysis.	"We also call the process refining, where we can check really difficult information and verify or falsify it. You can think of, gee, for example, we see that a suspect has taken a certain route, which persons, vehicles or phones did all take a similar route? That is a completely different analysis [...] , so also the ability to ask yourself these kind of questions, and to understand the limitations. That is getting more important" (IV08)
Data knowledge	If and why knowing the context and the limitations of the data is experienced as beneficial for analysis.	"you also have criminal intelligence from Team Criminal Intelligence, but the disadvantage of criminal intelligence is that you cannot estimate its reliability in any way. They do provide a very nice schema of the reliability, representing if the informant has been reliable in the past. But yeah, it's more in the background. So I never really use them in my analysis. The same holds for anonymous crime reports, you can often not validate them" (IV16)
Leadership that facilitates	If and why suitable leadership is experienced as beneficial for analysis	"... because their team leader simply expects you to look for data imports. They can try to say no, but this is not always easy. [..] There are big differences of what is expected of analysts" (IV03)
Complex visualisation and analytical tools	If and why complex visualisation and analysis tools are experienced as beneficial for analysis	"Yes, the refinery, as I just said, that is a useful tool, it really contains an awful lot. Even if, for example, there are trackers under the car, you can immediately run them over a map. You combine data from tapped phones which included locations. So that is a nice tool." (IV01)
Standardisation of processes	If and why standardised processes are experienced as beneficial for analysis	"But in itself the CTER process, yes, I think is an example of how it should work, on paper it looks absolutely fine." (IV13)

Inhibitors

This theme includes the inhibitors found in our case-study. Each of the inhibitors are coded as their own sub-theme. In our Findings section these are inverted and presented as enablers

Bad Data Quality	If and why bad data quality is experienced as hindering analysis (inverted to high data quality enabler)	"That's really difficult. And often mistakes are made there. So we as analysts are often busy structuring the data. So then in principle, you spend 70% of your time properly organising and structuring your data, after which only 20% remains for the actual analysis and then another 10% for the preparation of your report. Yes, and that is not enough for a good analysis." (IV15)
Bad Data Integration	If and why bad data integration is experienced as hindering analysis (inverted to easy data integration enabler)	"Sometimes I think, it should actually be automated, but it isn't. Yes, then I think yes. I cannot do anything. And well, that's a shame, now i really have to put some effort in it, spending a lot of time to make it happen." (IV01)
Bad IT Maintenance	If and why bad IT Maintenance is experienced as hindering analysis (inverted to continuous IT maintenance enabler)	"... the colleague who maintained it... it was not properly guaranteed... It was the responsibility of one colleague, while many analysts used the program. So when we moved to Office 365, the tool stopped working." (IV01)
Unclear which systems are available	If and why not knowing which systems are available is experienced as hindering analysis (inverted to availability of software catalogues enabler)	"Yes, there are so many systems out there and every time something new is invented. Then there is the Amsterdam Police Academy, which has a lot in development, and sometimes you can't see the forest for the trees" (IV05)
Difficult Systems	If and why difficult systems are experienced as hindering analysis (inverted to software intuitiveness enabler)	"It's too complex for me to use it there. I know colleagues who use it, but then they solely use that, because they don't get around to other things. It would be nice if, but that is difficult, if it was more user-friendly. It's not a simple program, let me put it that way. You also have to keep at it so you don't forget, just knowing where all kinds of buttons are to do thinks is already difficult." (IV10)
Slow Systems	If and why slow systems are experienced as hindering analysis (inverted to software speed enabler)	"Things have been going well lately, but sometimes you have to wait a long time for the system to be ready, even though you know that it can be done much faster." (IV03)
Unclear Task Descriptions	If and why unclear task descriptions are experienced as hindering analysis (inverted to clear task descriptions enabler)	"And that is very important to discuss with the client or with the requester, to clarify it clearly so that the question, and later the outcomes, are aligned. Otherwise you will get an analysis that people are not interested in. Then as an analyst you can be very satisfied with your analysis; but if that does not meet your client's expectations, then you have done the analysis for nothing" (IV15)

Table 4: The codebook, containing the themes and sub-themes used in this research, each with a description and example. Top-level themes are printed bold.