How Analysts Think: Think-steps as a Tool for Structuring Sensemaking in Criminal Intelligence Analysis

Nallini Selvaraj, Simon Attfield, Peter Passmore, William Wong

School of Science and Technology
Middlesex University
London, UK

Abstract— **Sensemaking has been described as a process involving information structuring. However, there are few detailed accounts of how this manifests in practice, particularly in relation to the creation and use of external representations such as data visualisations, and how such structuring aids sensemaking. To explore these questions in depth, we present an interview study of police crime analysts from which a model of their analysis process is developed. We describe the model focusing on the notion of 'think-steps', which for the analysts acted as a primary structuring concept. We describe how 'think-steps' propagate throughout the analysis process captured in the model. For the analysts, 'think-steps' are extensible templates that decompose a case into elements, provide a way of storing and visually structuring data, support generation of requests for information, focus research, simulate a case, and shape reporting. We reflect on the implications that our findings might have for design, including the possibility of a repertoire of evolving, sharable and reusable templates for sensemaking within a community of practice.**

Keywords—sensemaking; criminal intelligence analysis; data structuring and visualisation; interaction

I. INTRODUCTION

Sensemaking has been described as a process of putting data into some kind of framework which allows us "to comprehend, understand, explain, attribute, extrapolate and predict" [6]. There are various theories which account for this process. For instance, the Data Frame theory proposes that elements are explained when they are fitted into a structure that links them to other elements, and the notion of a 'frame' is used as an explanatory structure to define relationships between entities [4]. A frame can take the form of a story describing the chronology of events, a map showing where we are by showing distances and directions between landmarks, a script describing a job role, or a plan for describing a sequence of actions. Pirolli and Card [3] describe a process of intelligence analysis in which structuring (schematisation) forms a central part, and Russell, Pirolli and Card [8] show how structured templates evolve during the sensemaking process. Wong and Varga [9] suggest that structuring information during sensemaking helps the analyst by allowing them to see what they know and what they don't know.

Clearly, structuring is important but we need better understanding of how it takes place during work practices. We report a study of crime analysts which we conducted with an interest in developing a detailed understanding of the process that they follow, and how they create and use visual representations. We map out their process with an interest in how structuring is reflected in their activities and how this helps (or hinders) them. We focus in particular on the use of 'think-steps'. In the work that we studied 'think-steps' were a key structuring concept which supports analysts in decomposing a case into separate elements.

The paper is organized as follows: In the next section we review related work informing this research. The following section presents a description of the field study undertaken to conduct this research. The findings emerging from the data analysis are described next. The paper then concludes with a discussion of the findings and the theoretical and design implications.

II. RELATED WORK

A range of views on sensemaking have evolved from studies addressing different perspectives such as from an organizational view [1], theories arising in library and information studies [2], and models of intelligence analysis [3]. A review of these views on sensemaking is presented here.

A seminal view is that presented by the Data Frame theory of sensemaking [4][5]. It is a macrocognitive theory of sensemaking which provides an account of how comprehension is achieved in the context of complex, dynamic and evolving situations, where understanding aspects of the world must also evolve. The model distinguishes two kinds of entity that interact dynamically during sensemaking: data and frames. Data are aspects of the world that the sensemaker experiences as they interact with it. A frame is a (typically internal) representation that stands as an account for that situation which may take a number of forms including stories that may explain chronological and causal relationships between events, maps describing locations and directions and plans describing sequences of actions. The theory presents a process of continual framing and re-framing in the light of data. On encountering a new situation there are usually a few key elements or anchors that invoke a plausible frame as an interpretation of the situation. The frame can then guide active exploration to elaborate the frame by attempting to fill in gaps or incorporate new data for example which may challenge or

confirm the frame. Klein et al. [4] enumerate seven kinds of sensemaking activity in this process which include: mapping of data and frame, elaboration of a frame, questioning of frames, preserving a frame in the presence of conflicting data, comparison of multiple frames, reframing in order to accommodate new information, and seeking a frame when the data do not readily fit into existing frames.

Antecedent to Data Frame theory is the depiction of sensemaking by Weick [1] who employs organisational research as a platform for developing a perspective on how sensemaking operates in complex organisational settings. He was interested to show how individual and social sensemaking shaped organisational sensemaking which is distributed and embodied through phenomena such as shared understanding and routines and roles. In locating sensemaking in a social context, Weick sees organisational sensemaking as being shaped by the language that organisations and communities use every day in noticing, interpreting and communicating significance, maintained by ideologies, assumptions, paradigms and traditions. He draws on Starbuck and Milliken's idea that when people make sense of stimuli they do so by placing it into a framework which allows them "to comprehend, understand, explain, attribute, extrapolate and predict" [6]. Frameworks serve to categorise, fill in missing data, assign likelihoods to data, and filter and hide data. They are seen as abstractions which arise from past socialisation and impose structure on new experience. Weick [6] depicts the starting points for sensemaking as situations in which people notice salient, novel, unusual, unexpected cues which, in the context of a set of beliefs, have ambiguous or uncertain meaning. Similar to Klein, Weick emphasises the interaction between cues and frames during sensemaking. He argues that at the most basic level sensemaking has a frame, a cue, and a connection. "A cue in a frame is what makes sense, not a cue alone or the frame alone" [Weick, p110]. Similar to Data Frame theory, Weick sees sensemaking as a continuous alternation between information and explanations, with each cycle giving added form and substance to the other.

Sense Making Methodology (SMM) was developed by Dervin [2][7] in the area of library and information sciences following studies that addressed questions such as how people define information needs in different situations, how they present these to systems, and how they make use of what systems offer. This approach is based on a number of premises about human information interaction. The first is that information is created by a user at a particular point in time rather than it being external and objective. This emphasises that sensemaking arises not through passive receipt of information but through an active process of engagement with information. SSM also argues that specific situations give rise to information needs and serve as predictors for sensemaking behaviour. An emphasis is placed on understanding sensemaking from the user's perspective by considering what is 'real' to them. Another key concept of SMM is the 'discontinuity assumption' - that sensemaking events can be evoked by the occurrence of discontinuities or gaps, which hinder understanding and prompt the use of sensemaking strategies to resolve them. Three elements: situations in which gaps arise, gaps and the uses of information to resolve gaps,

form the basis of a conceptual framework which is used to guide research questions, elicitation methods and methods of analysis. A situation is seen as the context in which sensemaking occurs which includes the sensemaker's existing understanding, experience and expertise as well as their social and cultural context. A gap is seen as the question or confusion that arises out of a situation, a gap that needs to be bridged by obtaining answers, the construction of new ideas, values, intuitions, narratives or obtaining resources. The use is the use to which the sensemaker puts the newly created sense.

Another view is presented by Pirolli and Card [3], who report an analysis of sensemaking by intelligence analysts in which they characterise sensemaking as consisting of information gathering, representation of information in a scheme that aids analysis, development of insight through the manipulation of this representation, and creation for some knowledge or action as output. The representation could be internal or some form of external representation. Pirolli and Card suggest that through experience analysts develop domain specific schemas around aspects of their tasks, and they report some external representations such as entities and linkages between them. They describe two major activity loops, a 'foraging loop' which involves seeking information and searching, filtering it and possibly developing some schema from it; whilst the 'sensemaking loop' involves iterative development of a mental model or schema that best fit the evidence. Raw data is searched and filtered into a 'shoebox', which is then skimmed and processed into an evidence file along with schemas that organise information so that conclusions may be more easily drawn. These then feed into a hypotheses stage which represents tentative conclusions along with supporting arguments. This in turn feeds into a presentation stage which depicts the work. The model is not just data driven but consists of multiple loops that move both top-down from theory to data as well as bottom up.

The process of schematisation embodied in Pirolli and Card's model was based on earlier work on the Learning Loop Complex by Russell et al. [8] which considers how information can be structured in terms of some form of schema in order to support analysis. Their model outlines four sensemaking steps. The first is the search for representations or the generation loop where the user generates representations or schema to describe the data. The second step is to instantiate representations, also known as the data coverage loop, where the user iteratively identifies information of interest and encodes it in instantiated schemas called 'encodons'. The next step may be to shift representations to make changes to the schema to accommodate ill-fitting data. Finally, a task specific information processing step processes the schema to produce the outputs of the sensemaking activity. The authors also stress the interplay of top-down and bottom-up processes.

There are a number of common features to the models of sensemaking that have been discussed above. They tend to cover the processing of data and the fitting of results into some form of representational structure. In order to make sense, many definitions emphasize the role of structuring of information and its representation in schema or frames. Starbuck and Milliken [6] capture this common theme well in their description of sensemaking as a process of putting data

into some kind of framework which allows us "to comprehend, understand, explain, attribute, extrapolate and predict". Consequently, it may be useful to create structured externalised representations to help users make sense of both what they know and what they need to know [9]. Structuring, however, may take a number of forms both as internal and external representations. The motivation for this paper is to explore 'in-depth' how data structuring serves sensemaking. We undertake this through a study of police analysts undertaking criminal investigation analysis task which is described in the ensuing section. Based on the study, we present the role of 'think-steps' in structuring how analysts think about and structure criminal investigation cases, and how it cuts across the multiple steps entailing the analysis process.

III. CONCEPTUALISING SENSEMAKING IN CRIMINAL INTELLIGENCE ANALYSIS

A. Field Study Description

A field study was conducted at a European federal police station. Data collection took place primarily through interviewing. This involved in-depth, one-to-one interview sessions with four police analysts. Each session lasted about four and a half hours. The interviews were conducted at the analyst's workplace which allowed viewing the actual tools used and visualisations created during analysis. The interview probes consisted of open-ended questions formulated based on a previous study and from the information provided by the analysts. In addition, the standard operational process in crime analysis was provided by the police analysts to the interviewers prior to the field visit. This further guided the interview questions. In preparation for the interview sessions, the police analysts were asked to primed themselves with a case to talk about in depth and at length. They were also asked to bring printouts of visual representations they created during the case analysis.

During the interview sessions, the analysts were asked to talk about a case from the start of their involvement explaining what they did, step by step, drawing in other cases where relevant. They were first asked to talk about the data and briefing they were given. They were then asked what they did at each step in sequence. Where elaboration was needed the interviewer used neutral questions such as "tell me more about x". Occasionally the interviewer repeated their interpretation of what the interviewee has said, in order to provoke confirmation, elaboration or correction. The interviews were audio recorded. The visualisation printouts provided by the analysts were anonymized by the interviewer by redacting sensitive data. The field data also included drawings and field notes made during the interview sessions.

Data analysis involved verbatim transcription of the audio recorded interviews and coding the transcripts to model the process employed by police analysts during crime investigations. A model was initially developed from the steps in the standard operational analysis process provided by the police analyst prior to data collection. The coding scheme for further data analysis was then drawn from this model, with the model being open to revision in the light of data. In the next

section, we describe the model, discussing how 'think steps' propagated through the different activities.

B. 'Think-steps' as a Sensemaking Tool in Police Operational Intelligence Analysis

On the basis of the study presented above, a model of police operational crime analysis was developed. This is a descriptive model which provides a conceptual vocabulary to account for the analysis process (Fig.1).
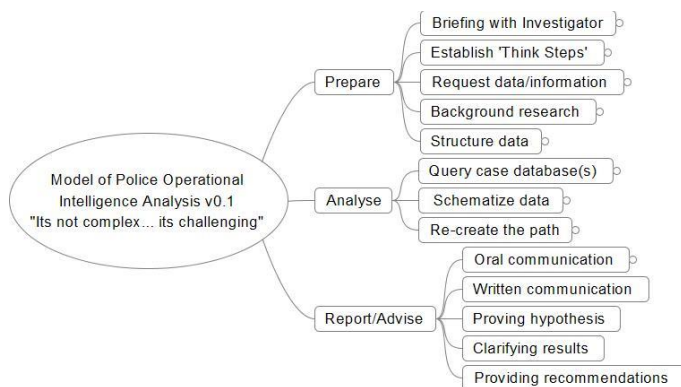

Fig. 1 Model of Police Crime Analysis Process

The model depicts the crime analysis process as a number of activities occurring in a number of phases. The phases are arranged according to a workflow 'logic', although in reality analysts passed through these phases many times during a case, sometimes quickly, sometimes less so, and sometimes stepped back rather than forwards. The activities within the phases are intended as a complete list of activities. Like the phases, they may not occur in the order shown and will not necessarily all be performed. We describe each of these phases and their constituent activities with an interest in how they related to 'think-steps'.

1) Prepare

a) Briefing: The prepare phase typically starts with a face-to-face *briefing* with an Investigating Officer. During *briefing* the Investigating Officer will explain what the investigating team knows about the case (objective data) and the theories that they may have about it. The investigator will provide any data (such as incident reports, telephone records, etc.) that had been gathered (there may have been some kind of analysis work already). During *briefing* the analyst and the investigator agree on goals for the analysis. When conducting the data collection, an interviewed analyst was asked what happened in the *briefing*, the participant says,

He (investigator) explained (the case), and he also talked about what he thought was going on. For example, he thought it was X perpetrators. It is very important because that turned out to be not right. That there was a big organization behind it because in the past you will see similar events - things get stolen from cars or cars get stolen – and then they end up somewhere in Y or where ever. So, they always think there is an organization behind it......So, he was fixed on that. And because of that of course he wanted to dismantle the organization. So, instead of going just after the thieves, he wanted to...well, not sit back but

he wanted to know more about them in order to get higher up in the organization. For me that was the starting point.

b) Establishing Think-Steps: The following transcript extract presents a description of 'think-steps' in the words of an interviewed analyst.

Think-steps can be a template for a generic crime but they can also be things that are drawn from the investigation about the case which I consider are specific areas to consider.

'Think-steps' provide a template that enables the analyst to "approach the case", decompose it into separate elements and classify associated data accordingly. One of the analysts illustrated 'think-steps' through a description of a case. The analyst said that there was a gang in country x who recruited girls to work in country y, under the illusion that they would be employed fetching potatoes. However, when the girls arrived in country y they were forced into prostitution. The gang members were subsequently involved in an incident with the police in country y, so the gang moved the girls to country z. The analyst considered what were the 'think-steps' that the analyst could 'hook' the case on in human trafficking, trading and prostitution. In this case, the analyst formulated the 'think-steps' as *recruiting, transport, housing, work, medical care and finance*.

The chosen 'think-steps' however depend upon the type of crime or major phenomenon. Different crime types have different 'think-steps'. In our interviews, the analysts discussed 'think-steps' for *murder, burglary, human trafficking, drug trafficking*, and *money laundering*. They may also differ in resolution. For instance, in the case of a murder or other single major incident, there will be a detailed moment-by-moment analysis of the time during and surrounding the event. In this case, the 'think-steps' would probably be formulated as what happened *before, during*, and *after* the event. Formulation of 'think-steps' begins at an early stage and are typically thought about during the first briefing. One analyst described this in the way they conducted an analysis of a mobile phone burglary case. In that case, a number of mobile phones had been found at two locations and the investigators were trying to identify whether there was a link. In the briefing, the investigators provided the analyst with telephone numbers and sim cards found in each location in the form of an Excel spreadsheet. When asked, at what point the analyst started drawing the 'think-steps' during case analysis, the analyst said:

The 'think-steps' are in my head from the beginning of the investigation. I know the numbers of the sim cards and telephone numbers are important. I know I have to visualise it that way later.

At any point in time, the chosen 'think-steps' represent a *theory* about the case at a given level of abstraction. One of the challenges is that investigation is a process of active discovery and new information may show that the 'think-steps' need to be adjusted or even changed entirely. In trafficking cases (e.g. drugs trafficking or human trafficking) for example, 'think-steps' might correspond to the stages in a transportation link (e.g. origin, transport, destination). Later, it might be discovered that there are multiple destinations (e.g. origin, transport 1, destination 1, transport 2, destination 2) or the template may change entirely. For instance, in a case of human trafficking, when the police investigators were interviewing victims, it was found that there had been a forced abortion and the baby had lived a couple of minutes before it died. A judge then decided that this was murder. So, the case then became a murder investigation leading to the analyst including more 'think-steps'. The analyst said "*it always changes*". The 'think-steps' start from a general template and as more information is collected they become more specific to the case. Sometimes, certain aspects of the case can become vital. For instance, in a murder case, someone passing a door can become important later in the investigation and this could become a 'think-step' later in the analysis.

c) Requesting data/information: For the analyst, one source of data is the Investigating Officer. They may request information at the initial briefing or they may do this as the investigation progresses. 'Think-steps' support the identification of information needs by indicating gaps and focusing further research. This is illustrated in a case described by one of the analysts. The crime was a series of car burglaries where the perpetrators stole GPS consoles from high-end cars in parking lots. The investigator provided the analyst with images of the perpetrators and number plates of the cars they were driving which were obtained from the surveillance cameras in the parking lots. The analyst said,

The first thing was to identify the thieves because they have pictures of two men. Identify them, locate them actually where they reside or where they are staying.

The data available about the crime led to the analyst formulating 'think-steps' about the men identified and this helped structure the data requests placed by the analyst. The analyst then received telephone and Automatic Number Plate Recognition (ANPR) data from the investigators through which the men were linked to the location of the crime. They also found out that the men had committed similar burglaries at another location but they did not have the data to relate the men and the crime. So, the analyst requested information from the police in this location as illustrated in the following interview transcript.

Of course we already knew of a series of burglaries. So, it was very quickly okay, there, there and there (pointing on the map) but also we saw them going to (I will show you later) to y and to x which is at the sea, at the coast. We didn't know of burglaries over there. So, we had to contact the police there in order to ask them - do you know of this kind of burglary. So, we could add more crimes to their records.

The analysts explained that structuring information visually according to 'think-steps' (e.g. in table form) can be particularly helpful for identifying information gaps and formulating requests for information. We discuss structuring in more detail below.

d) Background research: The analysis process typically requires the analyst to perform some form of *background research* which can include researching information about the

entities in the case (e.g. people, vehicles, locations etc.) using police databases and publically available data on the internet. The analyst may also perform *phenomena study* as part of *background research*. The phenomena can be for instance particular religious, cultural or community practices and beliefs. Where these are involved in a case it can be important for analysts to understand them in order to make interpretations of the information they have. The phenomena can also include the nature of a particular kind of crime and associated entities.

When asked to explain a *phenomena study*, one of the analysts produced a visualisation in Mind Map (Fig 2).
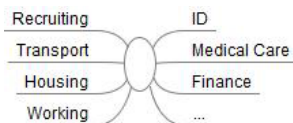
Recruiting	ID
Transport	Medical Care
Housing	Finance
Working	...

Fig. 2 Depiction of 'think-steps' created by an analyst as a result of *phenomenon study*

This was created by the analyst when handling the human trafficking case. The analyst was not familiar with this type of crime and so had to undertake a *phenomena study* to gain better understanding and obtain contextual information about such crimes. The analyst created a visual representation of 'think-steps' in this form (Fig 2) to help structure the research about the phenomenon. Elements on the left side represent stages in the crime and those on the right hand side are attributed to a particular person associated with it.

e) Structure data: As a part of the preparatory phase, the analysts structure their data before working with it. They had at least two different ways of doing this. The first approach involved manually extracting entities and events from case report documents and entering them into a database specific to the case ('object/entity' database). There were pre-prepared templates for doing this, although they were not referred to as 'think-steps'. For example, events had labels, locations, a start time, an end time as well as linked actors/entities an information source (e.g. interrogation x, document y etc.). Entities include people, objects and places, and are associated with each other and with events. Encoding the data in this way supported querying later.

'Think-steps' are an important conceptual tool for structuring data. In the previous example, the analyst said, "*I start by thinking what I should look for, then I collect the information and put it in the form of the 'think-steps' in a mind map. (Then, I question what do I have about a particular 'think-step'. For instance, what do I have about housing? What do I have about the financial stuff?).* In the case of mobile phone burglary, the analyst structured the data by categorising objects according to location. This was based on the 'think-steps' (*sim cards, location*) formulated by her at the time of the briefing.

As analysis progressed, the investigators obtained call records of the mobile phone and sim card numbers found at the two locations from mobile service providers. On receiving this data, the analyst structures it through the 'think-steps' as illustrated in the following interview transcript, *I have filtered the call data to show the links between the phones found at the two locations.* The analyst said they were looking for links between the two addresses or links that they could go further with in the analysis. In the words of the analyst: *Do I have a number which I can connect to the address here who made the phone call to a number I found in the other address?* One approach to data structuring using 'think-steps' was to 'code' textual data in much the same way that analysts code data during qualitative research. In this case, the 'think-steps' terms were used as codes or "keywords" and used to annotate sections of text. Essentially, this indexed the data in terms of 'think-steps', and was designed primarily for easy retrieval.

2) Analyse

a) Querying data: Much of the preparation phase, was focused on creating a resource for the analysis that would then permit the analyst to access data and generate ad-hoc queries in a way that would maintain the momentum of the analysis. In the case of a series of car burglaries, the 'think-steps' formulated by the analyst included the two men identified as perpetrators of the crime, car number plates, and location of burglary. In the course of analysis the investigators obtained related telephone data. Based on this, the men could be related to the location of the crime around the time that it was committed. The following transcript of interview data illustrates how the 'think-steps' formulated by the analyst structured the questions raised about the data.

The detectives, they are not really interested in who is he calling. They are interested in where is he and can we relate crimes to that, to them being there. But I saw – why is he calling 101, why is he calling that number? So, to me that was very important. Like I showed you there, the X number calling in and then he is coming from A to B. This was an anchor point for me because it is something you are sure of. You can confirm it by going to the telephone guys, the police and asking them – okay, retrieve the call. What was he saying? What was his name? Why was he calling?

In this scenario, the 'think-steps' lead to more questions about the activities of the two men. Also, the number plates were identified as belonging to country C. Based on this information the analyst raised the question with police from C as illustrated in the following transcript.

So, we have the number plates and we asked the C police – do you know these number plates. Did they show up in your database?

Just as 'think-steps' had provided an important tool for storing information, so they also provided an important resource for retrieving information through querying data files. One way was that the 'think-steps' provided ways of organizing questions raised during analysis that might be presented to a database as a query. Where an object/entity database was created for a case then the analyst would query it around different 'think-steps'. For instance, in the case of mobile phone burglary, the analyst uses an in-house database to store crime data. In the database, the analyst shows a dropdown list of incidents of crime reports. When asked how that list is created, the analyst says they manually type in

information from all the reports obtained from the investigator. The reports are provided in different forms such as in Word or Excel files. The analyst says *I don't have to have all the information but just the points that I need*. Then, the analyst explains that to have the list of all the things that are stolen at the two places the analyst can query the database and quickly get the list.

b) Schematize data: During analysis, the analysts schematize data by creating structured representations in the form of Excel spreadsheets, Word documents, or charts in IBM i2. For instance, in the mobile phone burglary case, the analyst started the analysis by creating a visualisation chart in IBM i2, in which the sim cards and phone numbers found in the two locations were placed as depicted in Fig 3. The red and blue boxes in Fig. 3 indicate the two locations from where the objects were confiscated, and the phone and sim card data found at each location have been placed within the corresponding boxes.
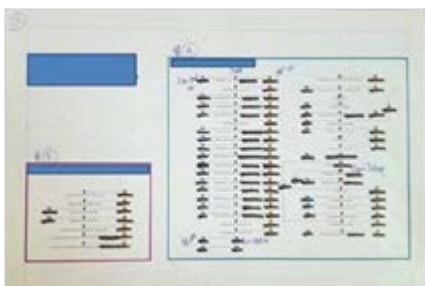

Fig. 3 Visual representation of objects found in each location

Here, the 'think-steps' of the analyst are derived from the data available about the crime - *location* and *objects*. When asked why the analyst created a visual representation of the data when the same information was presented to her by the investigator in a tabular form in an Excel spreadsheet, the analyst said,

This allows me to see what was found. This is where I start from. I try to gather more information but this is the starting point and this is where I go from. I always like to have something visual. The links can be seen here in the i2 chart better than in the table in Excel and this is useful. It is easier to get a sense of it visually.

The initial visualisation lays out the data, interpreted as fact, and provides a visually structured framework to act as a template over which subsequent data from sources yet to be accessed can be added. The spatial layout also acts as a planning resource allowing the analyst to quickly identify gaps and areas that can be fruitfully explored through new requests for information. For instance, during analysis the analyst tried to link the phone numbers found in the two locations. Once the telephone call record data from the mobile phone service providers had been obtained, the schematization evolves to the one depicted in Fig 4.

Here, the analyst has created a visualisation of the call data to show the links between the phone numbers found at the two locations - the hotel and house.
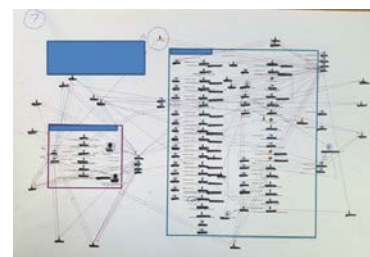

Fig 4 Visual representation of links between phone numbers found in each location

Re-create path: The analysts examined the visualisations they created to develop their understanding of what happened. This was referred to by one analyst as 're-creating the path' or developing a scenario addressing questions such as, *who did what?, where? and when?.* 'Recreating the path' seemed to have narrative orientation but to also involve staying close to the facts and not theorising too far. Competing accounts were also aligned against objective data to evaluate them.

This step often leads to the generation of follow-on questions. For example, it might involve hypothesising that a set of sim cards were used by a single person. This then becomes a question which can be addressed by querying the database to seeing records of the three sim cards aligned in order to consider whether there is support for the hypothesis in terms of time frames in which they were used. Or, it may result in questions that can only be answered with more data.

The following transcript of interview data illustrates the way the analyst recreates the path in the case of car burglary. The 'think-steps' formulated by the analyst included the two men identified which here structures the way the analyst recreates the path about the crime during analysis. In the following interview transcript, the analyst describes colour coding the representations of two men to draw out patterns in their roles with respect to a particular crime.

So, I am thinking – okay, the blue one or the red one? I don't know. The blue one has the car. He drives to the other one, picks him up and then they move. So, I am thinking, the blue one is the driver. Red one is the one sitting next to him and he is the one who actually gets out the car, gets into the victim's car, burgles the satnav and comes back. So, I can say more or less that number is the driver, that number is the thief. And they are always the same you know. It is always the same guy getting into the cars of the victims. Because when you go to court, it is important to determine who did what. Just it could be difference in punishment or whatever.... So, it was important for the detectives to determine which number was which guy.

3) Report/Advise

The analyst may present their findings to the investigator in emails, by talking through a schema or as a PowerPoint presentation. The report contains hypotheses/conclusions draw from data in a visual scheme which shows how one supports the other. Reports typically include recommendations to further explore hypotheses, which may involve acquiring more data (for example, applying to an inquiry judge for some phone data) and conducting more analysis. In the words of an analyst,

It can be in a lot of ways. Sometimes I just telephone the investigator and say have you noticed that? You can make a scheme. You can put it in a report. It depends – whom you work with, what kind of case, how important it is, and the time you have.

An instance of reporting done in the car burglary case was presented by the analyst. The analyst presents a report to the investigator which is structured on the 'think-steps' formulated in this case, specifically the two men identified as perpetrators of the crime. The analyst explains

Well, here I have an example of, basically it is the same report, but a different date. Is 7th of July....I am just saying these two guys are in X. Then I put, well basically type what is said in the telephone data, and I say where they are and when they are there. If it is important who they are calling because sometimes you know if they are calling the same number again it could be interesting. I am saying – okay, the red one (target A) is in X and classic base, that's where he is always staying, and the blue one (target B), ...he is also in X under his base station where he always sleeps. So, I can say – well basically we start the day on the 7th of July and they are probably at home and they are going to bed. And then they become active – at 10:53 , the red one is starts to call people, he is still under the same base station and then the blue one also gets active and then they start moving about in X.....So, I used the colours. So, they know the red one is target A and blue one is target B. So, I keep that throughout the reports. So, I made a report of everyday we saw him coming in.

IV. DISCUSSION AND CONCLUSION

In this paper we have reported results from a study conducted with a group of police crime analysts. We focus on the role that 'think-steps' play as structuring concepts in the way they do their work. We have done this by describing the analysis process in terms of a model involving three phases: *prepare*, *analyse* and *report/advise*, each with sub-activities. The model was developed through a qualitative study involving four police analysts in a single workplace. For these analysts, the idea of 'think-steps' was incorporated into their training programme. We assume from this that the ideas generalise well to other analysts who use the same or similar approach. It is, in effect, a structured analytic method which draws on a repertoire of 'frames' corresponding to 'types' of investigated phenomena. Other analytic domains will no doubt have their own frame repertoires and structuring tools in terms of these will have similar benefits. For example Attfield and Blandford [11] observed the way in which lawyers conducting regulatory investigations structured information around frames relevant to their enquiry, such as business activities (e.g. contracts), particular time periods surrounding major events within those activities, and protagonists or potential protagonists under investigation. Those authors argued that if systems can be configured around the concepts that analysts apply to data, then they are likely to provide a more suitable platform upon which investigators can apply their expertise in sifting through large volumes of data, and allow them to work to a higher conceptual level. The similarity of observations across the two domains helps to formalise these two substantive theories.

Structuring information appears in numerous theoretical accounts of sensemaking. Our account of the way that police analysts work fits with many of them. 'Think-steps' can be seen to correspond with elements of a 'frame' as described in the data/frame theory [4] in the sense that they act as a form of interpretative template/explanatory 'theory' about the data in terms of what it means. And just as frames, in the data/frame theory can be wrong, so the data can reveal that the 'think-steps' require evolutionary change or even complete revision. In the language of data/frame theory we think of this as arising through the process of 'questioning the frame' leading to 'elaborating the frame' in the case of minor revisions and 're-framing' in the case of a complete revision. We also see a notable link between these frame changes and what Russell et al. [8] call the 'representational shift loop' in their Learning Loop Complex model. Essentially, the Learning Loop Complex model describes the process of deriving a schema, just as our analysts did using 'think-steps' and then changing the schema when relevant data appears that doesn't fit the schema. These re-framings clearly present our analysts with a dilemma and our model can help us to elucidate this. The prepare phase involves defining analysis goals, constructing a preliminary theory, devising the 'think-steps', and then seeking and enriching (structuring) information in terms of that theory. Hence, there is a commitment of resource to a particular theory - a commitment which is not insignificant. If information is sought, structured and stored in terms of a theory, and the theory is changed or even rejected in favour of another theory then there may occur a resistance to change. In more general terms this might give rise to a form of confirmation bias in the extended system (i.e. the analysts and their external resources). Confirmation bias refers to the tendency of people to interpret data in a manner consistent with existing beliefs rather that change or update their beliefs. Chater and Lowenstein [10] argue that this is because finding alternate models is difficult. They describe confirmation bias as an inevitable consequence of sensemaking. An extension to this, perhaps, is that for any information system where effort is implied by organising information in terms of a prevailing theory, there may be resistance to change. Hence, reducing the viscosity of such changes might be an important priority.

From our interview data, we have attempted to draw out the uses and benefits of 'think-steps' for analysis. In summary, 'think-steps' provide a template that enables the analyst to approach the case and decompose it into separate elements. Formulation of 'think-steps' begins at an early stage based on an understanding of what kind of crime it is that is being investigated or major phenomena around. 'Think-steps' appear to provide a way of storing data. They also provide concepts that are used for visually structuring data in and support the generation of requests for information by indicating gaps and focussing further research. They can also help structure the way that an analyst internally simulates a case in 'recreating the path'.

Finally we consider some implications that our findings may have for the design of information systems to support the work of analysts. 'Think-step's appeared to provide the analysts we spoke to with the benefits of problem decomposition, asking questions systematically, retrieving data

quickly and structuring mental simulations of explanatory narratives. In order to do this it was necessary to be able classify or index data in terms of the 'think-step' concepts. Given the user-overhead in doing this, and particularly the overhead in reclassifying following frame-changes/extensions, there may be value to be had in reducing the time and cost to the analysts in performing this kind of function and making the most of the work once it is done. For example, we noted the analysts indexing textual data or 'coding' thematically in a similar way to that qualitative researchers code data as part of their analysis process. This suggests that analysts might benefit from the kinds of coding tools that can be found in Computer Assisted Qualitative Data Analysis Software (CAQDAS). CAQDAS software such as Nvivo and AtlasTi typically makes coding a relatively simple process and them allows the user to access data easily through the codes. Our findings also suggest that analysts might benefit from tools that could use existing coding to predict the 'think-steps' concepts that are applicable to unseen data. This might use machine learning algorithms to associate data with issue codes rather like the 'predictive coding' or 'computer assisted review' tools that are make an impact in the world of e-discovery. Such tools might be usefully augmented with information extraction technologies that can identify and associate references to individuals, places etc.

Our findings also show that analysts make good use of visualisation (in this case concept maps using IBM I2) and that 'think-steps' feature in the way that they organise these. The analysts appeared to benefit from lot of flexibility in the way that they created these visualisations, but since 'think-steps' were a significant organising concept these were used in creating visual organisation. In particular, visualisations organised in terms of 'think-steps' could show what data an analyst had about each element in the 'think-steps', and help suggest and make systematic the articulation of new information requirements. Visualisations organised by 'think-steps' as key elements in an investigation could also be used to help reveal how those elements interacted with each other, which could be important to know.

Lastly, our consideration relates to the 'think-steps' themselves and how these might be specialised and even shared within a community of practice. Our findings support the idea, proposed by [11], that investigators in given domains have a set of phenomena types with which they frequently have an interest. For the police investigators it was types of crime and other phenomena. What we saw was the application of these broad types as generic structuring concepts. This might be followed by some adaptation and specialisation (i.e. elaboration) depending on the specifics of the case. The extent to which this is a general finding in investigation domains is the extent to which such domains might benefit from tailored sets of structures like this. Not only this, as new phenomena emerge, specialisation/adaptation occurs and new structures develop to accommodate, it may be helpful to support the sharing of reusable templates perhaps with generic explanations. That way, investigators may be able to benefit from the research and experiences of others. We imagine such templates might play a role similar to reusable workflows. This raises the question of how such templates might be stored and accessed. Here, we imagine the possibility of structured catalogues of templates in which generic templates which are defined for fewer features, act as parents to more highly specified templates defined in terms of the same features and more.

In future research we plan develop a systematic cognitive task analysis approach for eliciting 'think-steps' templates and their variants to conduct further empirical studies that will enable us to explore in practice how 'think-steps' might be embedded within tools supporting different stages of crime analysis.

REFERENCES

[1] K. E. Weick, "The Nature of Sensemaking," Sensemaking Organ., pp. 1–62, 1995.

[2] B. Dervin, An overview of sense-making research: Concepts, methods, and results to date. The Author, 1983.

[3] P. Pirolli and S. Card, "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis," Proc. Int. Conf. Intell. Anal., vol. 2005, pp. 2–4, 2005.

[4] G. a. Klein, J. K. Phillips, E. L. Rall, and D. a. Peluso, "A Data-Frame Theory of Sensemaking," Expertise out of context. pp. 113–155, 2007.

[5] G. Klein and B. Moon, "Making sense of sensemaking 2: A macrocognitive model," IEEE Intelligent Systems, vol. 21, no. 5. pp. 88–92, 2006.

[6] W. H. Starbuck and F. J. Milliken, "Executives' Perceptual Filters: What They Notice and How They Make Sense," in The Executive Effect: Concepts and Methods for Studying Top Managers., Hambrick Donald, Ed. Greenwich. CT: JAI Press, pp. 35–65.1988.

[7] B. Dervin, "From the mind's eye of the user: The sense-making qualitative-quantitative methodology," in Qualitative research in information management, J. D. Glazier and R. R. Powell, Eds. Libraries Unlimited, pp. 61–84, 1992.

[8] D. M. Russell, M. J. Stefik, P. Pirolli, and S. K. Card, "The cost structure of sensemaking," in Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '93, pp. 269–276, 1993.

[9] B. L. W. Wong and M. Varga, "Black Holes, Keyholes And Brown Worms: Challenges In Sense Making," Proc. Hum. Factors Ergon. Soc. Annu. Meet., vol. 56, no. 1, pp. 287–291, 2012.

[10] Chater, N., & Loewenstein, G. The under-appreciated drive for sense-making. Journal of Economic Behavior & Organization, 126, 137–154, 2016.http://doi.org/10.1016/j.jebo.2015.10.016

[11] Attfield, S., & Blandford, A. (2008). E-Disclosure Viewed as Sensemaking with Computers: The Challenge of Frames. Digital Evidence & Elec. Signature L. Rev., 5, 62.