Making sense of digital footprints in team-based legal investigations: The acquisition of focus

Simon Attfield

Middlesex University, UK

Ann Blandford

University College London, UK

RUNNING HEAD: MAKING SENSE OF DIGITAL FOOTPRINTS

Corresponding Author's Contact Information:

s.attfield@cs.ucl.ac.uk

Brief Authors' Biographies: (follow the template below)

Simon Attfield is a information interaction specialist with an interest in how people work with information and what this means for the design of human-centered interactive systems; he is a Senior Lecturer in the Interaction Design Centre of Middlesex University. **Ann Blandford** is a specialist in user-centred modelling of situated interaction, with an interest in amplifying human capabilities through design, including work on sensemaking; she is Professor of Human Computer Interaction and Director of UCL Interaction Centre at University College London.

(Additional note: Simon Attfield will take up the position above from the beginning of April 2010. Since the paper will appear after this date, these details are the most appropriate. Until April 2010 he is a Senior Research Associate at UCL Interaction Centre at University College London. The contact email address is a UCL address.)

See **hci-journal.com/editorial/final-guidelines.html** for explanations of the various parts of an article in *HCI* format.

ABSTRACT

Sensemaking occurs when people face the problem of forming an understanding of a situation. One scenario in which technology has a particularly significant impact on sensemaking and its success is in legal investigations. Legal investigations extend over time, are resource intensive, and require the sifting and re-representation of large collections of electronic evidence and close collaboration between multiple investigators. In this paper, we present an account of sensemaking in three corporate legal investigations. We summarise information interaction processes in the form of a model which conceptualises processes as resource transformations triggered and shaped by both bottom-up and top-down resources. The model both extends upon and validates aspects of a previous account of investigative sensemaking (Pirolli & Card, 2005) and brings to the fore two kinds of focusing. Data focusing involves identifying and structuring information to draw out facts relevant to a given set of investigation issues. Issue focusing involves revising the issues in the light of new insights. Both are essential in sensemaking. We draw this distinction through detailed accounts of two activities in the investigations: reviewing documents for relevance and the creation and use of external representations. This provides a basis for a number of requirements for sensemaking support systems, particularly in collaborative settings, including: document annotation, dynamically associating documents of a given type; interacting with documents in fluid ways; linking external representation elements to evidence; filtering external representations in flexible ways; and viewing external representations at different levels of scale and fidelity. Finally, we use our data to analyse the conceptual elements within a 'line of enquiry'. This provides a framework which can form the basis for partitioning information into hierarchically embedded enquiry 'contexts' within collaborative sensemaking systems.

CONTENTS

1. INTRODUCTION

When the legal conduct of a company is brought into question, the concerns that are raised can trigger an investigation on behalf of a regulatory authority, or as a prelude to possible litigation. These investigations are carried out by teams of lawyers who either present their findings to the regulator or, in the case of litigation, disclose the documents that are relevant to the matter and construct a case on behalf of their client. Such investigations typically involve making sense of large bodies of documentary evidence that record the day-to-day activities of corporate life. These can include many kinds of user-generated content, including emails and office documents, voicemail and instant messaging records. Once these documents have been obtained, a process begins whereby lawyers, working in teams, search, review, sort and re-represent information in order to make sense of facts relevant to the case—a process known as e-discovery, electronic data discovery, or e-disclosure. They also typically conduct interviews with key witnesses.

In recent years, the tractability of such investigations has been challenged by the exponential growth in the volume of electronically stored information within modern enterprises—a trend which is set to increase (Attfield & Blandford, 2008; Baron et al**.**, 2007). Electronic discovery requests for email alone can result in tens of millions of documents (Baron et al**.**, 2007). The challenge for lawyers working on such investigations is to identify and construct the narratives that matter from a very large collection of unstructured information. Whilst technological advances have created this challenge, it has been suggested that new technologies also offer an opportunity for addressing it (Baron et al**.**, 2007).

To effectively support investigative sensemaking, such technologies must be based on an understanding of the sensemaking processes of the people who perform them. Some research has already been done on investigative sensemaking by intelligence analysts based on collections of electronic evidence (e.g. Pirolli & Card, 2005; Bodnar, 2005). This work has provided an initial perspective on the human-centred processes underpinning such investigations. However, to date, this work has been relatively broad-brushed and has been restricted to the work of intelligence analysts, rather than lawyers or others who engage with similar processes. There is consequently a need to better understand sensemaking in electronic investigations, and to relate that understanding to findings from other sensemaking domains.

In this paper, we present a study of three large e-discovery investigations performed by lawyers and other staff within an international law firm. The study is based on fourteen in-depth interviews with lawyers, trainee lawyers and paralegals who worked on the investigations. The interviews aimed to provide detailed reconstructions of how the investigators approached these tasks and the problems they faced. We created and validated a generalised model to describe the overall processes of the investigations and, within this:

- how the investigators homed in on documents of interest using the technological tools at their disposal,

- how they generated and used external representations of the domain of investigation, and

- how they coordinated the team working necessary to conduct large investigations.

By addressing these questions we are able to identify specific challenges and needs, and hence how technology might provide additional support.

We start with a review of the related research that most directly informed our data gathering and analysis; we then present the case-studies, and relate our findings to the established literature. Our study has highlighted central issues to sensemaking with large bodies of data that have previously received little attention, namely how people achieve focus and coordinate their activities. This understanding, in turn, highlights requirements for technology design.

2. BACKGROUND

Sensemaking has been described as "the reciprocal interaction of information seeking, meaning ascription and action" (Thomas, Clark & Gioia, 1993, p240), and as "the deliberate effort to understand events". (Klein, Phillips, Rall & Peluso, 2007, p.114). It is a ubiquitous activity (Klein et al., 2007), and as a topic of research spans a number of disciplines. In addition to Human Computer Interaction, these include Naturalistic Decision Making (Klein, Moon & Hoffman, 2006; Klein et al., 2007), Organisational Studies (Weick, 1995; Starbuck & Milliken, 1988) and Information Science (Dervin, 1983; Savolainen, 2006). A common characteristic of sensemaking which has been identified in various studies and theories, and which could be said to be a signature phenomenon, is the interplay between top-down and bottom-up processing. Sensemaking operates as a bi-directional process between data on the one hand and representations that account for data on the other.

This dynamic is clearly described in Klein and colleagues' data frame theory of sensemaking, for example. The data frame theory (Klein et al, 2007; Klein et al, 2006), is concerned with comprehension as it occurs in the context of complex, dynamic and evolving situations. The theory presents sensemaking as a continual process of framing and re-framing in the light of data. As we encounter a new situation a few key elements, or anchors, invoke a plausible frame (internal representation) as an interpretation of that situation. Active exploration guided by the frame then elaborates it or challenges it by revealing inconsistent data. By extending further than the observed data, a frame offers an economy on the data required for understanding, but also sets up expectations for further data that might be available. Hence a frame can "direct" information seeking and in doing so reveal further data that changes the frame. Like Starbuck and Milliken (1988), Klein et al argue that an activated frame acts as an information filter, not only determining what information is subsequently sought, but also what aspects of a situation will subsequently be noticed.

Klein derives his theory from empirical studies in contexts such as command and control, intensive care and weather forecasting. The symbiotic interaction between data and frame also features prominently in Weick's account of sensemaking based on studies of organisations. He draws, for example, on Starbuck and Milliken's (1988) 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" (Starbuck & Milliken, 1988, p.51). Frameworks serve to "categorise data, assign likelihoods to data, hide data, and fill in missing data" (Starbuck & Milliken, 1988, p51).

Within Human Computer Interaction and Information Science, research has understandably focused on technologically mediated sensemaking. Users often interact with information systems in order to develop some 'picture' or 'model' of a domain (Dervin, 1993; Spence, 1999). Technologically mediated sensemaking often extends over time and involves searching for and integrating large amounts of information into a coherent understanding. Also, whereas, for both Klein and Weick, the representations considered are internal and cognitive, within Human Computer Interaction, there has been a particular interest in the role and design of technologically supported, user-generated externalisations of domain representations (see for example Russell, Stefik, Pirolli & Card, 1993; Pirolli & Card, 2005; Sereno, Buckingham Shum, & Motta, 2005; and Qu & Furnas, 2005)—something that, in principle, computer technology can support well.

Despite this difference, the same bi-directional process between data and representation is evident. For example, Russell et al. (1993) report on a study of course designers developing a course for laser printer technicians. The designers defined a set of schemas for capturing salient information about a range of printers (using a hypermedia structuring tool) in order to identify core concepts within the material. Once designed, the schemas provided a set of entity types with pre-defined slots for particular kinds of information. In using the schemas, however, they repeatedly discovered that they did not adequately capture salient information in an unambiguous way, so they adjusted the schemas throughout the process. This observation led Russell et al. (1993) to propose the learning loop complex model. This involves four sensemaking steps:

1. Search for representations - Generate representations (schemas) to capture salient features of the data (the generation loop).

2. Instantiate representations - Identify information of interest and encode it in instantiated schemas (encodons).

3. Shift representations – The discovery of ill-fitting data (residue) motivates changes to the representational schemas.

4. Consume encodons – A final task-specific information processing step is performed using the instantiated schemas.

Pirolli and Card (2005) report preliminary findings from a study of intelligence analysts, which similarly exemplifies the interplay between top-down and bottom-up processing in sensemaking. They provide a broad brush description of the process, and suggest some potential leverage points for developments in technology. Their model shows transformations that the analyst performs in converting multiple data sources into novel information. It consists of two major activity loops: a foraging loop and a sensemaking loop. Foraging involves seeking information, searching and filtering it, and reading and extracting information, possibly into some schema. The sensemaking loop involves the iterative development of a "mental model" or "conceptualisation" from the schema that best fits the evidence.

The foraging loop centres on a collection of raw evidence (external data sources). The analyst selects subsets of these for further processing (held in a shoebox). Snippets are then extracted from this data (stored in an evidence file). This information is then re-represented in a structured way (schemas) to support sensemaking. From this, tentative conclusions are generated (hypotheses) with supporting arguments, and ultimately the work product is communicated (presentation).

Significantly, the model is not committed to a single direction of processing. Rather, it is intentionally constituted of multiple loops which move both from the bottom up (data to theory) and from the top down (theory to data). Pirolli and Card report an opportunistic interplay between both kinds of process. From bottom to top, the analyst searches or monitors incoming information and sets aside relevant information as it is encountered, then nuggets are extracted and re-represented schematically, a theory develops and is ultimately presented to some audience. In the opposite direction, new theories suggest hypotheses to be considered and the schemas are re-considered in this light, collected evidence is re-examined, new information is extracted from the shoebox, and new raw data is sought.

Pirolli and Card note that a primary challenge for intelligence analysts is the need to cope with large amounts of information within limited time. They propose that technologies are needed that enable broader monitoring of an information space combined with support for narrowing in on key items and patterns. For example, they propose highlighting important information using pre-attentive codings or automatic summarization. In relation to the sensemaking part of the model, they identify the need to use external displays to represent multiple connections between data as well as support for generating, managing and evaluating multiple hypotheses.

The kind of sensemaking that people do, and the tools that might make it easier for them to converge on an understanding, depends on a number of factors. These include the domain they are trying to make sense of, their prior understanding of that domain, the sources of information that can provide information about that domain, and their motives for doing so. In this sense, the Pirolli and Card model provides particularly relevant context for the current paper. However, it is a relatively high level account that lacks detail on how people exploit external representations, how teams of people coordinate their activities, or how they achieve the focus that is essential when working with very large datasets. These questions form the focus for the study presented here.

3. METHOD

Data for the case-studies reported here were gathered at a large international law firm. Fourteen in-depth interviews were conducted with staff who had worked on three different cases (or 'matters', as they are referred to in legal firms). We interviewed one technical coordinator (with responsibility for supporting e-discovery operations), two trainees, six associate lawyers, one senior associate lawyer and three partners. A senior associate who managed one of the investigations was interviewed twice. Interviews lasted from 45 minutes to 1hr 40 minutes. Ten of the interviews (including the two with the senior associate) pertained to the identification of a suspected fraud; one pertained to an earlier suspected fraud (chosen to test the generality of findings within one kind of legal matter); and three pertained to a matter concerning the origin of anomalies within a set of legal contracts (to test the generality of findings across contrasting types of matter).

Participants were recruited through a combination of snowball (Johnson, 1990) and theoretical sampling (Glaser & Strauss, 1967). Theoretical sampling was used to focus in on emerging issues and explore similarities and contrasts between investigations. Following the practices of Grounded Theory (Glaser & Strauss, 1967), data gathering and analysis were interleaved.

For reasons of client confidentiality, it was not possible to gather real-time observational data. However, key sense-making artefacts were made available for inspection during interviews to help participants reconstruct specifics. In some cases these artefacts were also made available following interviews. They included software that participants had used, loaded with data that they had worked on, external representations created to support the investigation, and an evidence table from an investigator's final report. These artefacts provided a reference point for discussing and reconstructing specific aspects of the investigations. In addition, data from the first five interviews were analysed to develop a preliminary model describing the major sensemaking processes of the legal team. This model was then also used for reference during subsequent interviews in order to further support focused and systematic questioning, and also to validate and refine the model as a description that could generalise across all three investigations.

Interviews were conducted in an open and informal way and without the use of pre-defined interview scripts. Each participant was asked to provide a broad account of how the investigation had unfolded from the beginning of their involvement. During or after this account they were prompted to provide detail in relation to their interactions with evidential documents and external representations that the investigators created (either as hard-copy or mediated through software tools), and also how they coordinated their work with other team members. Participants were encouraged to contextualise these detailed descriptions in terms of their rationale, including the ongoing problems and questions of the respective investigation. In order to invite the participant to correct the researcher's understanding and provide additional detail, aspects of their accounts were summarised by the researcher at intervals during each interview.

Interviews were transcribed and analysed through open and axial coding (Strauss & Corbin, 1998) and used to refine a process model that would describe all three investigations. To derive the model, coding attended to capturing major areas of activity, the resources they used, and the products they produced. In this way it was possible to link activities through the fact that, during work, the product of one process typically provides a resource to be used by another (see for example, Attfield, Fegan & Blandford, 2009). In addition validating with participants, as the model developed it was verified through constant comparison against the data (Strauss & Corbin, 1998). However, qualitative, inductive data analysis is itself a sensemaking process, having many of the properties of sensemaking described in the literature and reviewed above. For clarity, it is not possible to convey the rich detail of the process whereby sense was made of the sensemaking of our study participants, beyond noting that it interleaved top-down reasoning, motivated by questions and informed by literature, with bottom-up analysis, driven by data.

We first present the model developed through the study which, while being informed by the literature, is data-driven. We illustrate the findings with representative extracts from the data (in these extracts, "[…]" means that words have been omitted for clarity and brevity). We relate our model to that of Pirolli and Card (2005) and then develop the key themes that extend the model beyond the themes that have previously been reported in the literature: we focus on focusing in relation to information interactions during document review, in relation to external representations, and we consider the structure of issue focusing.

4. FINDINGS: THE BASIC MODEL

4.1 Background to the investigations

Two of the investigations we examined were carried out under authority of a regulatory body. These focused on concerns about management conduct, triggered in one instance by specific allegations and in the other by apparent accounting discrepancies. In both cases, immediate action was taken to preserve information held by the companies in question. Data forensic techniques were then used to recover documents, which were predominantly electronic in their native form. In both cases, this resulted in collections consisting of millions of documents (hundreds of millions in one case). Both investigations were conducted in close consultation with regulatory bodies with the aim of discovering whether rules had been broken and, if so, by whom. A characteristic of these regulatory investigations was initial uncertainty about what issues might ultimately be brought to light. Despite being triggered by particular concerns, the job of the lawyers was additionally to discover any related impropriety. Consequently, the investigative foci were initially broad and subject to ongoing review.

The third investigation, in contrast to the other two, had a well defined focus. A review of a financial product by a client had led to the conclusion that the rules according to which it was administered had been drafted incorrectly. To address this, it was necessary for the client to demonstrate that the error misrepresented the original intention. The law firm was asked to investigate the history of the drafting of the rules

and represent the client's case in court. Another law firm acting as the 'other side' also assessed the evidence in order to challenge the misrepresentation theory where appropriate. The evidence, in this case, was a series of paper documents gathered from the client's offices, including memos, meeting records and draft rules, which were electronically scanned prior to review.

4.2 A model of the investigation process

Our data showed that each of the investigations involved a similar set of processes. These are summarised in model form in Figure 1. In this section we discuss some general features of the model. This provides orientation for detailed discussion of document review and the construction of external representations below (which are sub-processes in the model), and also of the conceptual structure of various lines of enquiry that the investigators pursued.

FIGURE 1 ABOUT HERE

In Figure 1, rectangular boxes represent information interaction processes. These may be supported by technology or they may not. Arrows in the model represent the flow of information between processes. This flow takes the form of different kinds of resource (marked against each arrow), with these being created or modified by one process and used by another. In some cases resources are external information objects, such as a database of evidential documents, search results, transcripts from witness interviews and claims (assertions about the investigated domain supported by evidence). These information objects had the role of providing the raw material for a subsequent process. Each process is also influenced by some characterisation of investigation issues (with the effect of structuring its goals) which are also shown as a kind of resource. 'Issues' was a term used by the investigators to refer to thematic lines of enquiry, of which there could be any number at any point in an investigation. These were typically based around one or more theories and associated questions (we explore the constituents of a line of enquiry in section 8).

The investigation model begins at the top with consultation with a client (1). Client consultation inevitably occurs throughout an investigation, but the process begins with objectives of the investigation being defined in collaboration with a client, and ends with the reporting of findings. Following initial consultation, information is gathered. The model shows two kinds of information seeking strategy which complement each other, represented by the two branches of the diagram in Figure 1.

In the right-hand branch, documents were recovered from source locations in the field (5) depending on the current formulation of the investigation issues (downward arrows) and used to populate a database (which investigators referred to as the 'document universe'). Initial document processing (e.g. metadata extraction, document de-duplication) was performed by litigation support staff prior to work by investigators; the details are outside the scope of the current paper. Again, based on the current investigation issues, keyword searches were submitted to the document universe (4) to create results sets for manual review. Many searches were conducted on an ongoing basis

throughout the investigations; in one of the investigations 200 searches were conducted each of which returned hundreds or even thousands of documents. These were then added to a document management system incorporating keyword search tools and tools for coding documents using metadata tags which could be created on a bespoke basis depending on the requirements of the individual investigation. Documents were further filtered through a manual review process (3) where they were individually read and metadata added to record (among other things) relevance to the investigation issues. In one investigation 130,000 documents were reviewed in all.

The other source of information (left-hand branch) came through interviewing witnesses (6). Interviews were recorded and transcribed.

Information from relevant documents and witness interviews were used as a source of data for the creation of external representations which summarised key findings as the investigators saw them. These representations included event chronologies, written narratives, social network diagrams showing communication behaviour, and organograms showing formal organisational structures. These representations were then used as a source of reflection and as a basis for reporting back to the client.

A key issue in the investigations was the acquisition of focus. Given the large amounts of evidence available, and the initial breadth of the issues under investigation, a significant challenge for the investigators was to be able to focus in on both the evidence and the questions that really mattered. This gives rise to two kinds of focusing which emerge from the processes represented by the model. We describe these under the labels: *data focusing* and *issue focusing*.

Data focusing propagates upwards through the model. Given a set of investigation issues (theories and questions) the investigators worked to identify, extract and structure information that would address the theories and answer their questions. From bottom to top, external information resources correspond to various stages of this analysis. At the lower ends of the model the volume of information held in external information objects is high, the average relevance is low, and the level of structuring is also low. At the higher end of the model, the volume of information in external information objects is lower, with higher relevance and greater structuring. In the most general of terms, the process is one of filtering information for relevance with respect to the investigation issues and integrating this into a representation of the investigated domain.

In contrast, *issue focusing* propagates downwards through the model. As the investigators worked with information, so insights arising from what they found changed their theories and questions and provided new foci for investigation. Typically, issues were re-specified into multiple sub-issues. This was particularly evident in the regulatory investigations in which the issues were initially broad. The following extract is from an interview with an associate lawyer, who worked on one of the regulatory investigations as a trainee. Here she discusses the effect of new information coming to light during document review:

P7 If you came across a new document, or email that showed someone was there. So then that would open up a whole new can of worms because then you think, right now this person is involved as well and this person definitely knew this much, you know, so that could sometimes open up new issues. […] You would then do more investigations focused on that person […] it was a really organic process where if someone discovered something that would then lead on to a whole new avenue of enquiry and that's how it grew really, so there was no sort of plan.

Similarly, a partner said,

P2 "I'm seeing a lot of this guy mentioned in relation to a search on the [issue name]. Do we know who he is? Some of the documents I saw, seem to suggest he was an underling, an unimportant." We were going, "Oh, but I saw that later he became this." Or, "I saw an email with him and so and so. And so and so is really important to me, we should keep an eye on this person." There'll be constant refinement of who and what we thought was important.

This illustrates how discoveries could cue new theories which then became a theme for further investigation. Changes to the issues were largely refinements on pre-existing issues rather than revisions to the overall scope of the investigations. We discuss the specifics of some of these changes in following sections. However, broadly, issues spawned sub-issues by depth-wise refinement with new findings providing the investigators with the domain language with which to articulate more focused interests. More focused interests could then give rise to new goals for processes such as the recovery of documents, searching, document review, interviewing etc. The net result of the two kinds of focusing is that there is no single path through the model. Rather, processing moves up or down depending on the goals that emerge at any particular time.

4.3 COMPARISON WITH THE MODEL OF PIROLLI AND CARD

Our data gathering was informed by earlier studies of sensemaking, but the model of legal sensemaking was derived from our data, not directly from any previous models. However, there are similarities between the model of legal sensemaking and Pirolli and Card's (2005) model of sensemaking by intelligence analysts. Pirolli and Card (2005) commented that intelligence analysis is an extremely variegated task, and cautioned against generalisation. Nevertheless, by considering points of comparison we can get an indication of features that do generalise.

Both models feature processes which act in sequence to filter and structure information into a representation of the investigated domain in support of a set of investigation concerns. They also show bottom-up and top-down influences acting between processes. Also, both lawyers and intelligence analysis make use of large collections of digital documents and represent information in structured ways, leading to some similarities at the level of individual processes (e.g. searching, filtering, and the creation of structured representations).

Comparisons between processes in both models are shown in figure 2. In this comparison, we use the terminology for "bottom-up" data gathering and analysis, while

recognising that in both models "bottom-up" analysis both informs and is informed by "top down" generation of theories, questions and issues.

FIGURE 2 ABOUT HERE

In the comparison some differences emerge. In the Pirolli and Card model, external sources of data are assumed, whereas in the legal investigations these were constructed as part of the investigation. Hence, decisions about what to recover responded to issue focusing and so formed part of the overall sensemaking process (Stage 5). Also, the legal investigations used additional data sources in the form of witness interviews and so these are incorporated as a parallel information channel. Decisions about which channel to use were made opportunistically depending on questions at the time. Finally, within the Pirolli and Card model hypothesis generation appears as a separate process at an upper stage of the model. However, and as described above, the data from lawyers showed that hypotheses (and associated theories and questions) could occur during any stage involving interaction with evidence or representations of evidence. Hence this is not a separate process but is distributed throughout the model.

Despite these differences, our data serves as a broad validation of the general form of the Pirolli and Card model based on data from a related but different domain. What we have demonstrated here is that a similar process occurs in the domain of corporate investigations, and so the form generalises well.

Pirolli and Card (2005) offered their model as a "broad brushed" characterisation without extensive elaboration of individual processes. In the following sections we focus in detail on some processes in order to "flesh out" the model, with a particular interest in issues that have implications for the design of technologies to support sensemaking within large team-based investigations of the kind we have studied. The areas we focus on are: document review; the creation and use of external representations; and the conceptual structure of issue focusing.

5. DOCUMENT REVIEW

In this section we look in more detail at document review (process 3 in figure 1). During document review, documents which have been selected in some way (typically by search) are individually read and coded with various kinds of metadata, ultimately in order to identify those that are relevant to various investigation issues. This is an extremely labour intensive process. During one of the investigations we studied, over 130,000 documents were individually read and coded in this way. Our data drew attention to a number of issues surrounding focusing during document review. These included: multistage reviewing, communication between review stages, dealing with emergent classes of irrelevant documents, identifying related relevant documents and the need for fluid document interactions.

5.1 Multistage reviewing

Manual review, and the metadata coding that results, has the effect of grouping documents into subsets considered relevant to particular defined issues of interest as a prelude to further analysis or future ad hoc retrieval. Depending on a number of factors, the documents reviewed might be all the documents recovered, or they might be a subset of these pre-filtered by searching.

Participants described two kinds of metadata that are applied during review. 'Objective' codes denote properties that are generally uncontroversial. These might include document date, title, author, recipient, etc., where these can be incontrovertibly identified from the document. 'Subjective' codes are used to denote properties that depend more on interpretation, including relevance to an investigation issue.

In two of the investigations we studied, the review was conducted in stages, with complementary review tasks performed by different personnel according to experience and expertise. For example, in one case, around twenty trainees were recruited from within the firm. Each was briefed and given reference material defining the issues. They were then each allocated folders of around 700 documents to review. The trainees' task was to read each document and code them in terms of relevance to the issues. This was done in order to filter documents prior to a second review. As one trainee explained,

> **P13** Your job is to filter it down to the ones that are relevant, which are then passed up to either be reviewed by associates, or trainees at another stage along the process. […] [the initial review] would make the reviewing task down the line easier because they'd know that all of the ones that had ticked [issue X] were responsive to a certain part of the review, all of the ones that had been ticked [issue Y] were responsive to a certain other part of it, and then you had non-responsive.

The trainees reported that, at this stage, irrelevant documents were by far in the majority. One trainee estimated that she had considered around 10 percent of the documents she reviewed to be potentially relevant to the investigation she was working on. But they recognised that less experience and less close involvement in the case made it harder to make judgements about relevance. When in doubt, their approach was to be inclusive—equivalent to a high recall information retrieval strategy at the possible cost of precision.

> P12 […] it was better to include too much than too little, because it would get filtered again at a higher level, rather than miss a vital document if it was in the, you know, if it was in the ballpark of what we were talking about […] because you don't want to miss stuff, and as I said, as so many of them were unresponsive, it wasn't as if we were firing hundreds a day through that were responsive to each of these things, it might have only been two or three a day, so to add another one wouldn't have placed a great burden on the people reviewing it.

The second stage review was performed by an associate lawyer who, given a deeper understanding of the case, was able to make more precise relevance discriminations. He estimated that he similarly reduced the documents to 10 percent of those coded as relevant in the first stage.

Multiple review stages form part of a sequence of filters (including recovery, search and the creation of external representations) which support data focusing by systematically reducing a very large document set to a set of key relevant documents. In this way each stage of review contributes to data focusing. The first stage provided a low-fidelity "cull" designed to eliminate clearly irrelevant documents. Later stages require more expertise. Notably, in the earlier stages when document numbers are high, filters are used which have low per-document costs. As the document numbers reduce, so higher precision, higher cost filters are used.

5.2 Communication between stages

Our data shows that the effectiveness of multistage reviewing can be enhanced by supporting communication about individual documents between review stages. In two of the investigations, the software used for reviewing allowed users to add comments to each document they reviewed. In one case it was possible to associate comments with specific sections of text.

These facilities were used by reviewers in two ways. First, they were used to explain relevance judgements and draw attention to any uncertainty to someone further down the line who may be more closely involved in the case and therefore better able to make a definitive judgement. As one trainee remarked,

> P12 […] you could say I think it's responsive because of blah blah… on the basis of X, but I may be wrong if I've misunderstood that.

Comments could also mitigate differences in knowledge between team members. A senior associate lawyer explained,

> P1 I know looking at this document it is quite important but I only know that because I went to a witness interview and this document now makes sense to me […] it is imperative that not only do I write this in as a key document, but I explain why it is a key document, so if someone else finds that document in the future, they don't change my coding saying, and say it's not relevant.

The second way this tool was used was to help a subsequent reviewer or analyst find key passages quickly. Documents could be hundreds of pages long. A trainee explained,

> P12 [we used it] to clarify, just to make sure, partly to make the reviewer's job, the next level reviewer's job a bit easier, so any issues with the document he can just go to straight away partly to make the reviewer's job, the next level reviewer's job a bit easier.

Supporting communication between review stages concerning individual documents, then, allows boundary cases at one level of review to be highlighted for detailed consideration at a subsequent stage, communicates the underlying rationale for review decision and can reduce the cost of subsequent reviewing by adding place holders to significant passages.

5.3 Identifying classes of irrelevant documents

Part of the review task is to eliminate irrelevant documents from further consideration. Further, identifiable classes of irrelevant documents can emerge to the reviewer as the review progresses. For example, as she worked through a "massive" folder, a trainee explained how she noticed that many documents she was reading significantly predated the events under investigations and therefore could be discarded. Identifying this class allowed her to adapt her reading strategy for each new document so they could be identified more quickly. She would first visually scan for a date (the documents did not have metadata denoting date and so she was unable to use search, and even if this had been possible, there was no facility for bulk tagging).

> **P13** I'd look and try and see if there's a date on it [referring to a particular document on the document review system]. The date that I can see so far is 1958, which strikes me as being fairly irrelevant [...]. I would have said this is almost certainly irrelevant, I've clicked it as non-responsive.

Another trainee explained how he discovered a set of invoices which, given the investigation he was working on, could similarly be discarded. He explains how he became accustomed to identifying their structural cues,

> **P12** … and then obviously, you just scan it, you don't have to read every line […] you just know that it's another one of those invoices […] you got to recognise the pattern of the document and just whizz through it […] you get to a stage where you can see a document and you can immediately go, no, it's another one of those […] you need to have read through a couple of them to say, OK they're all like this.

Another recognition cue he used was a pattern within the series of documents,

> **P12**… you would get the invoice followed by the cover letter, every time, and there was a whole series.

Another class of document that could be identified as irrelevant were duplicates. Given that in modern, networked organizations documents are easily copied from one machine to another, a given document could appear numerous times during a review. This was described as frustrating and as adding to the burden of review and communication within the investigation team. Although steps were taken to automate the removal of duplicates, this was technically a difficult and fallible process. One of the difficulties for the reviewers, however, was that identifying a document as a duplicate depended on them recognizing it as such,

> **P8** Sometimes you aren't sure whether to mark it as a duplicate or not because you've just seen so much information that you're in an information overload state that you are worried, have I seen it? haven't I seen it? I'm not going to mark it as a duplicate because I really just can't remember.

These examples show how reviewers came to recognise sets of similar, irrelevant documents through a process of induction arising from repeated exposure. This allowed the development of strategies for faster recognition. However, temporal separation

between instances of exposure and associated memory degradation could impede the recognition process.

5.4 Identifying related relevant documents

A similar situation arose in the identification of related relevant documents. By analysing documents and witness interviews the investigators aimed to re-construct narratives concerning events they were interested in (discussed in more detail in section 6). Very often they inferred these events by examining email exchanges between protagonists. A single email, however, would typically only provide partial and potentially inconclusive evidence for an event. A planned meeting may not have taken place or it may have been replaced by a telephone call. And so for each event they needed to review all the evidence relating to it. A partner explained,

> **P10** What you're then trying to do is to work out exactly what happened at that meeting or during that call and what you're then doing is enabling yourself, by putting the whole picture together, to work out what would have been most likely to have been said by A to B because you can put yourself then into the position of what was actually going on, what was in the mind of both people, what documents had flowed between.

Thus, encountering an email about an apparently significant event can result in a new low-level focus. This is issue focusing as discussed previously. The review systems used, however, did not provide a means by which reviewers could control their encounters with documents in order to maintain continuity around such an issue. Documents were presented in lists, and reviewers would start at the top and work down. Documents related to a given sub-theme, however, could be distributed throughout a collection, with no way of bringing them together. Another document relating to that sub-theme could be encountered an hour, a day or a week later. This had two effects. By disrupting cognitive momentum reviewers believed this increased the time it took to review any given document. A trainee responded to a question on this as follows,

> **Q** If you become interested in a particular issue […] you can't then go and find other documents?
>
> **P13** No, you just have to hope they're together.
>
> **Q** Why do you hope they're together?
>
> **P13** Because it's easier if you've just, say if you've done this over the course of three weeks, it's much easier if you've just read the document that related to it, to read the next one and it makes it quicker to read it because you don't have to go, what was that about again? Why did I think that was relevant? […] so it's helpful if then the next document that's relevant to that tricky point is next to it because then you can just use the same knowledge as opposed to having to reconstruct it two weeks later.

In addition to requiring reviewers to keep issues of interest in mind over extended periods, a consequence of this was a need to maintain multiple threads of interest simultaneously. We return to these themes in the discussion below.

5.5 Fluid document interaction

So far we have discussed the review process from the perspective of how document encounters relate to one another. In this section we consider some issues associated with individual document interactions.

Participants reported that some of the documents they reviewed were very long with only a very small part being relevant to the review. Although the systems they used supported search for constructing the sets for review searches within results and searches within documents were not possible. This was seen as a significant drawback. For example,

> **P5** Sometimes if you're in a big document, agreements, as you know, can run to many hundreds of pages, rather than having to scroll through each page looking for stuff to cross-refer to a clause or a trigger-word, it's much easier if you can just type it in, and it'll take you to page 30 or whatever, and then you see it in clause 8.4, and you can go, ah yes.

Also, long documents, such as draft legal agreements, are structured such that interpretation depends upon frequent cross-referencing, for example between clauses and definitions, and this was poorly supported. For example,

> **P5** Quite often a legal document, as you probably know, is structured with a set of definitions at the front which then feed into the rest of the document going forward, and to understand the substance of the document you need to be always cross-referring to your definitions, so it's much easier to be able to just go like that [demonstrates turning a page with a physical document], rather than, you know, going to page 14 and the back.

These examples highlight the need to be able to move around a document fluidly, homing in on specific areas of interest and cross-referring one section with another, in order to achieve focus within large, often structured, documents.

To summarise this section, document review clearly plays an important role in data focusing but at some cost. Multistage review provides a way of managing this but our findings suggest that it can be enhanced by other means, such as communication between stages about individual documents, and fluid document interaction. During document review, emergent classes of irrelevant documents can appear to a reviewer leading to adaptations in recognition strategy, but temporal separation between instances of exposure can result in memory degradation. Related to this, groups of related, relevant documents can emerge and become a focus of interest. Hence, document review also demonstrates issue focusing. However, temporal separation between instances of exposure can also present difficulties. We discuss these issues in more detail in section 8 below. In the next section we turn to the creation and use of external representations.

6. WORKING WITH EXTERNAL REPRESENTATIONS

A central process in the investigations was the creation, amendment and review of external representations (Figure 1, stage 2). These representations retained key facts and supported reflection and communication. In this section we consider the construction and use of these representations and how these reflected both data and issue focusing.

The investigator created a number of different kinds of external representation. Different representations were useful at different stages of investigation. For example, early in one investigation an investigator produced a large social network diagram (or 'link chart') which showed people and communication links between them based on some early information from a witness interview. This representation provided the investigators with an overview from which they could consider which people in the investigated domain to focus attention on, a decision which subsequently informed document recovery efforts.

The investigator also used company organograms for early issue focusing. As a partner explained,

> **P10** [...] organisations are difficult to get to grips with in terms of how they work […] we need to try and work out who reports to who for what purposes, then you often have a lot of dispute about it ((laughs)) and so an organogram which shows that is a useful tool […] it was also obviously useful to try to work your way around who you need to speak to.

Another representation used, only at a late stage of an investigation, were narratives which were written by trainees and junior lawyers as compilations of all that was known about particular characters or issues. These acted as briefings for more senior members of the investigation team in preparation for client meetings and witness interviews. One associate lawyer said,

> P7 If there were meetings or interviews then we would have to produce necessary analysis for that very discrete topic. So if there was an interview with a particular person for example we could probably just try to find everything that person was involved in, we would have to produce an analysis of what we thought that person knew at the time.

Within the regulatory investigations, however, the most significant representations were extensive chronologies that the investigators created to represent detailed events surrounding the issues they were investigating. These acted as comprehensive visual records of the investigators' understanding of various streams of narrative they considered significant to the investigation. In the following sections we describe the nature and construction of these chronologies, how they were used and how they supported sensemaking, focusing in particular on one investigation.

6.1 CONSTRUCTING AND AMENDING CHRONOLOGIES

> **P1** I think it's a very natural way for us to think here, we always use chronologies, our great organising basis. […] I had a team of five or six people and I allocated responsibilities to each of these people saying "Right you're going to become the

master of [issue a], I'm going to do [issue b], [issue c], [issue d]. Someone else is going to do [issue e]"

In the investigation in question, the investigators were split into teams, each assigned to one or more issue areas. As sets of results were returned from document searches, these were allocated to a relevant sub-team for review. Key documents (often emails) arising from this were then used to construct one chronology for each issue area.

Chronologies were created using Microsoft Excel spreadsheets. An entry might record a meeting between protagonists, the signing of a contract, a protagonist travelling, or simply the sending of an email with a significant message. A pre-defined schema was used for each entry (an anonymised example is shown in Figure 3). This included the date and time of the event, a text account (e.g. meeting between a and b at location c to discuss d; email from e to f asking for g), a field for recording the people involved and the location, and a field for recording references to supporting document(s).

FIGURE 3 ABOUT HERE

Maintaining references to supporting documents was important since reviewing these was a frequent activity. For example, when an investigator found a document that they thought warranted an entry in the chronology, they might find that an entry had already been created. They would then want to review the previous evidence to relate the new evidence to the previous documents. An associate lawyer said,

P6 And you would compare. And if it looked like you already had the document… I mean some of this was obvious, "yes, this is the document. We've already got it. It's in the file, it's in the chronology". Sometimes it was a little bit doubtful because sometimes the entry in the chronology wasn't complete, and then you would run to the paper file and you would double check.

The investigators also reviewed evidence underlying chronology entries in order to resolve ambiguities and errors in the raw documents. An associate lawyer said,

P2 And because sometimes you'd go, "Hang on, that doesn't make any sense." And you'd look back and you think, no, that's the wrong Mr. Jones.

And also,

P2 Sometimes you'd go […] "My god, that's a typo, that can't be right because we know from this, that's a much better source of evidence, that that person wasn't in the country." So you'd constantly be revising and reviewing the material.

Access to the raw evidence was also important for supporting team meetings. During meetings, members of the investigation teams presented findings. This provided an opportunity for other investigators to learn what had been found, but also to review their colleague's interpretations of the evidence.

As these individual issue chronologies were developed, so selected content from each was integrated into a single master chronology (again, Excel). Ultimately, this ran to

around 13,000 entries. Within this chronology, each entry was coded according to the issue coding scheme applied during document review. In addition, the spreadsheet was augmented with some underlying functionality which allowed it to be filtered (or collapsed) according to selections from the coding scheme. In this way, any combination of issue codes could be selected and used as the basis for selectively displaying some investigation narratives to the exclusion of others, and also for seeing two or more narratives interwoven sequentially so that one could be considered in the light of another.

6.2 Sensemaking with chronologies

Given the amount of information they had to deal with, it was essential for the investigators to focus in on key areas, and like the other representational forms, the chronologies also provided an important tool for issue focusing, which then fed down to inform and structure lower level processes (see Figure 1).

For example, a number of issue teams were tasked with looking at events surrounding specific contracts run by a company under investigation. They were concerned with the possibility of a particular kind of fraudulent activity, an indication of which would be communication between key people at a critical time in the contract lifecycle. By mapping out contract activities in broad terms they were able to identify these critical periods. A senior associate lawyer told us,

> **P1** We'd be thinking, well if we're right on this, this is a really important build up […] or, we think money must have been sucked out of this business around this time. […] And this is what we did. [Junior Partner] selected certain periods and posed certain questions in relation to those periods. And we would go back and interrogate the information further. Sometimes that would involve running brand new searches across this period.

By reflecting on the chronologies and narrowing their focus, the investigators were able to conduct far more thorough searches of the evidence relating to specific time-periods, often to the order of a couple of days. They searched by specific time periods and relaxed other search criteria and so significantly increased the recall of searches for short, well-defined periods. If the results set was not too large (in the order of 1000 documents was acceptable) then all the resulting documents would be reviewed. A partner said,

> **P3** […] and for key time-periods we just broke that down. It became more and more granular so you could see exactly what was happening. So some time-periods where it was absolutely critical to know… because you're following this through forensically trying to figure out what's going on… it's absolutely critical to know minute-by-minute the exact chain of events.

Ultimately the density of entries in the chronologies varied considerably depending on whether they formed part of a key period or provided the context for defining these periods.

Reflecting on the chronologies also helped the investigators identify unusual, unexplained or missing events. In one example, an issue team realised that they didn't have any record of communications between two people who they would have expected to have been communicating, given what they had found already. An associate lawyer said,

> **P2** […] have we seen emails from Joe Bloggs to Andy Smith? And we think, well, no we haven't, but surely there must be. […] Well, there is the head of procurement, and here's the CFO - surely they're reporting - one's reporting to each other, or they'd be on boards together. […] So we'd be constantly refining in that way.

Similarly, events that seemed odd or inexplicable in the light of a developing narrative could prompt new, highly specific questions. One issue team reconstructing the activities of a protagonist found something that "didn't make sense". A senior associate explained,

> **P1** Well you're kind of thinking why on earth in the middle of a really hectic […] schedule is this guy sending emails saying, "I've got to fly to [country name] tomorrow, but I'm only going for the day and then I'll be back", and you're scratching your head, why the hell was this guy going to [place] ?

Whilst questions such as these could motivate new, highly specific document searches, answers could also be found by aligning chronologies produced by different investigation teams. This was where the master chronology integrating all of the findings was particularly useful. The filtering that it supported allowed the investigators to select and view any combination of issue chronologies in a single, integrated form. The senior associate went on,

> **P1** And it's when you put that together with actually the chronology of this company 'ABC' that you realise that there's a big event involving that company on that date and hitherto you had no idea that this guy had had any dealings with that company and so then suddenly you're building another relationship that you would never have thought of before.

The value of filtering the master chronology in this way was that events occurring around the same time from different parts of the investigation could be aligned and considered in the light of each other. An important aspect of this was the ability to eliminate irrelevant events from the view. As the senior associate said,

> **P1** I think the biggest advantage of the collapsible chronology is […] you just want to be able to home in on five entries on a certain date, or on a event involving two or three people, so its really just the filtering of it just goes straight to what you want and because you just want the bare minimum that you need to get the answer […] One of the things you could do was just look at events involving two or three people.

In the investigation in which this form of collapsible chronology was used, however, there were some limitations in terms of filtering flexibility. Filtering options could only be expressed by defining combinations of the originating issue chronologies. This provided only a small number of relatively coarse filtering options. Consequently, 'noise'

could still occur between events being considered in the light of each other. This resulted in a need to scroll the representation backwards and forwards in order to consider relationships between certain events.

We can see that the creation and use of external representations reflects both data focusing and issue focusing. The social network diagrams, narratives and chronologies were created by selecting (and structuring) particular information for re-representation. The representational filtering functionality that the investigators used in the chronologies also offered a means of dynamic data focusing by narrowing the view on the data.

Issue focusing is evident in the ways in which these representations were used. The social network diagram and the organograms provided an early, low-fidelity view of the domain in a way that allowed specific areas of enquiry to be identified. The narratives supported interviewers in forming focused questions. Finally, the chronologies supported issue focusing by helping the investigators identify critical time periods and particular unusual, unexplained or missing events.

7. THE CONCEPTUAL STRUCTURE OF ISSUE FOCUSING

In relation to issue focusing, we have discussed how new discoveries enabled the investigators to develop more focused areas of enquiry and how these provided them with new questions. This was a particular feature of the regulatory investigations we studied. For example, in one investigation, the investigators were interested in exploring activities surrounding a particular class of contract held by the company in question. Initially, though they did not know what contracts of this type the company had. As a senior associate explained,

> **P1** Well actually what [class] contracts does the company have? And no one in the company knows or can tell you so you're then trying to piece that together.

Details of specific contracts emerged gradually through investigation. As this happened, each contract then provided a basis on which to define more focused areas of enquiry. As discussed above, investigation into specific contracts revealed particular time periods of interest which could be investigated more thoroughly. And ultimately, particular events were revealed which could become the subject of detailed, "forensic" investigation.

This shows how successive focusing occurred through the gradual definition of recursively embedded lines of enquiry, each triggered by discoveries. Each was independent insofar as it posed its own questions, had its own strategies (e.g. date-delimited search) and developed its own knowledge. But neither were they complete departures, but rather a discovery in one area of enquiry spawned a number of sub-problems, the results of which fed up to the broader questions and ultimately the investigation as a whole.

As new lower-level lines of enquiry were established, so responsibility for these was allocated to sub-teams of investigators and in some cases to individuals. But it was important that significant results propagated up to inform more senior investigators

responsible for larger chunks of the investigation. It was also important that findings could pass between investigators dealing with different but potentially related problems. This led to a need for a good deal of communication. As a senior associate explained,

> **P1** The amount of communication that has to go on in order to make that work is phenomenal. […] So what we would do is we'd have little huddles regularly through the day, like ten/fifteen minutes and you'd talk, particularly in the early stages […] At the beginning of the day what you'd do is each group would have ten minutes of telling everyone in the review room what the general discoveries they were making so that everyone was sort of aware generally of the whole…

In addition to informal huddles and review meetings, the investigators passed documents to each other, produced briefing notes for reviewers and had a senior investigator wandering between teams cross-pollinating them. These not only provided ways of communicating theories and questions around the team (issue focusing) but also allowed for the exchange of information in relation to these questions (data focusing).

The partitioning of lines of enquiry led us to consider how such structure might be usefully reflected within systems for supporting large-scale collaborative sensemaking in an effort to address challenges of scalability and collaboration. At the heart of this are the observations that:

- a line of enquiry establishes a context of elements which pertain to it

- investigators working on a line of enquiry want to focus on these elements to the exclusion of extraneous information (i.e. noise)

- however, investigators need to be able to exchange information and questions with investigators working on different problems

- lines of enquiry are recursively embedded and different investigators work at different levels of granularity defined by this structure

Of particular interest is to provide a means by which investigators could define a set of investigation 'contexts' that could support both decomposition and integration of large investigation problems. In order to explore this we returned to our data and used it to develop an account of the major, recurring conceptual entities that investigators associated with a line of enquiry at any level of granularity. To do this we performed a detailed Grounded Theory (Strauss & Corbin, 1998) analysis of discussions about various lines of enquiry within the interview data based around an approach described in more detail by Blandford, Green, Furniss and Makri (2008) and using the concept of a line of enquiry as a central category. This resulted in a framework which integrates elements of multiple lines of enquiry into a single structure. We describe this framework in the following section.

7.1 The Line of Enquiry Framework

According to the framework a line of enquiry is a primary object. The recurring conceptual elements within a line of enquiry are theories, questions, information seeking strategies, evidence (and evidence collections), knowledge, assigned investigators and lower-level lines of enquiry. Each line of enquiry included these elements. Significantly, it is the knowledge generated by work on a given line of enquiry that can give rise to one or more lower-level lines of enquiry, each with similar structure.

Theories

Our data shows that theories or conjectures were central to any line of enquiry.

P1 Well it's the theories that then define the issues you are coding for and looking for. […] we had lots of sub-issues and theories, well sub-theories that were helping to define the issues […]

A theory would be triggered by a cue. This could be an initial allegation, or knowledge arising from part of the investigation. Theories were systematically investigated and, if the evidence found was unsupportive, they were eliminated from further attention. When all the theories associated with a line-of-enquiry were eliminated then the issue, as an area of focus, would become inactive.

Questions

The investigators made a natural move from the theories they developed to research questions that would address them, and in many cases these were explicitly recorded. Research questions specified requirements for information that would test theories, or simply elaborate the focus that they defined. This elaboration could in turn provide cues for further decomposition, or could yield other unexpected findings.

Information seeking strategies

Given the questions, each line of enquiry would have associated with it a set of information seeking strategies that the investigators agreed upon to address the questions. These might include the recovery of documents from new sources, new keyword searches over an existing document collection, the examination of telephone records over key periods, or interviews with witnesses.

Evidence and evidence collections

Searches provided the investigators with collections of potentially relevant documents. A line of enquiry could have multiple associated searches, and these might be repeated periodically as new documents were added to the collection. Manual review resulted in collections of documents considered particularly relevant to different lines of enquiry.

Knowledge

The investigators used evidence to extract and re-represent facts using different forms of knowledge representation, including social network diagrams, written narratives and most importantly extensive chronologies. Even though chronologies from different lines of enquiry could be combined, it was still important to maintain separation according to the different lines of enquiry from which they originated. As discussed, knowledge representations provided an important resource for reviewing findings and developing new lines of enquiry and also maintained references to the supporting evidential documents.

Assigned investigators

Given the team setting, any line of enquiry could be allocated to one or more investigators. Hence, from the perspective of the investigating team, these assignments were associated with each line of enquiry.

Lower level lines of enquiry

Finally, knowledge associated with a line of enquiry could give rise to any number of more focused problems. These lower-level issues featured more focused theories, questions and information seeking strategies and gave rise to their own knowledge. They could be assigned to a smaller sub-set of investigators, or they could act as small scale deviations for a single investigator.

This framework provides an ontology of concepts associated with any given line of enquiry. We have found these elements to occur irrespective of granularity. In some cases, a line of enquiry might concern a single relationship or a single event, whilst the investigation as a whole can be considered a single line of enquiry. When instantiated, the framework gives rise to a hierarchy of enquiry nodes corresponding to successive levels of emergent issues focusing with relevant elements represented at each node. Each enquiry node establishes a 'context' of relevant elements for that line of enquiry. We discuss how the operationalistion of this framework might impact on the design of large-scale investigative sensemaking support systems in the next section.

8 DISCUSSION AND IMPLICATIONS

In this paper, we have described the activities of, and concepts used by, teams of investigators making sense of large bodies of data in collaborative legal investigations. We have summarised the investigation process, explored aspects of document review and the creation and use of external representations, and shown how investigation problems are decomposed into lower-level lines of enquiry. In this section we review our findings and discuss some implications.

The model of investigative sensemaking presented in Figure 1 summarises the processes of the three corporate investigations and the way that these processes interacted. This was through resources created or modified by one process and used by another. Two kinds of resource are significant here: information objects and the characterisation of the potentially multiple investigation issues.

It is useful to reflect on how the process reported here compares with sensemaking in other domains. The investigators had access to a large document collection as a source of evidence. They were also acting under constraints of 'due-diligence' arising from the fact that the findings mattered considerably to others. Similar to the case-study by Pirolli and Card (2005), this gives rise to an extended and considered process of searching, reading and extraction into external representations and ultimately to some form of third party presentation. As discussed in section 4.3, though, the availability of witnesses for interview added an additional prong to the lawyers' information channels. Different sensemaking domains may have different numbers and types of information channels and this will inevitably change the processes involved. Sensemaking processes are shaped as much by the resources that are available as they are by intended outcomes. One of the skills of the sensemaker is to strategically identify appropriate resource opportunities. Many of these features also hold with respect to Russell et al's (1993) study of course designers, giving rise to similar processes. However, in that study the kind of sense required corresponded to the associations between related objects within a domain. The lawyers, on the other hand, primarily needed narratives of activity based on the interpretation of episodic information. Constructing narratives allowed them to think causally and by doing so identify gaps and anomalies. Hence chronologies were a central representation.

The episodic (and therefore specific) nature of required understanding also points to distinctions between legal investigations and, for example, sensemaking in educational settings. The lawyers were experts in law and business practices applying their knowledge to the interpretation of sequences of events. Although abstract learning undoubtedly occurred, it wasn't the reason for the investigations. A student may similarly need to make sense of specific situations in the course of learning, but this typically acts as a vehicle for the acquisition of more abstract, semantic knowledge. Further, we expect that student learning is less collaborative than the legal investigations we have seen. This impacts on issues such as the distribution of labour, the associated need for the explicit articulation of process and problem structuring, the creation and use of shared external representations, and opportunities for collaboratively testing interpretations.

A key problem for large investigations is the acquisition of focus. Focusing is represented in the model as having two forms. Data focusing propagates upwards in the model and involves reducing the volume and structuring information to draw out information that addresses a given set of investigation issues. However, as new findings come to light so the formulation of the issues evolves, giving rise to issue focusing. This propagates downwards through the model re-initiating and changing the goals of lower level processes. We expect these two forms of focusing to be a generalisable feature of sensemaking. This unfamiliarity inevitably leads to a need for leveraging new knowledge in a process of continual problem re-structuring.

Data focusing and issue focusing are both necessary for effective sensemaking. Consequently, the problem of supporting sensemaking, particularly in contexts such as legal investigations in which very large amounts of information are involved, becomes a question of enabling both to occur as effectively as possible. In particular,

- Data focusing places a premium on being able to locate and extract information of interest.

- Issue focusing emphasises the ways in which this information is represented back to the sensemaker such that they are able to gain insight and reframe their interests.

We considered in detail two processes within the model: document review and the creation and use of external representations. From this we derive the following requirements for sensemaking support systems in this domain:

Document annotation - Multistage review is a data focusing activity involving the application of increasingly high-quality but expensive (per-document) filters on a systematically reducing document collection. This forms part of a longer sequence of filters which include document recovery (including database populating) and search. In terms of optimising costs and quality, the process is likely to be the most rational option, given the problem and the resources available. However, it also introduces the need to support collaboration. In particular, reviewers find it useful to be able to annotate documents to explain judgements and communicate any uncertainty about them through to the next stage. They also value tools for communicating key passages determining document decisions. This has the effect of reducing the costs associated with subsequent review filters in which judgements are reviewed.

Dynamically associating documents of a given type - During document review, the reviewer can become aware of classes of irrelevant documents within the collection, for which they develop recognition strategies, or relevant sub-issues through which some documents in a collection are related (a form of local issue focusing). A factor which mitigates against dealing with these effectively, however, is a lack of tools for drawing such groups together. Temporal separation and resulting interference effects become an apparent obstacle to maintaining cognitive momentum.

There are commercially available tools which address the problem of associating related documents within a collection in the e-discovery area. A leading example of this is the Document Mapper interface which is part of Attenex Patterns system (see McNee & Arnette, 2008). The Document Mapper uses term distributions to perform a cluster analysis over a document collection. Two-dimensional visual proximity is then used to show semantic associations between documents. McNee and Arnette (2008) report improvements in review productivity using the system in excess of one order of magnitude compared to traditional systems.

We believe that the study reported here argues in favour of solutions like Attenex Patterns since they can offer the capacity to associate documents in a way that supports more concurrent engagement with local foci and, potentially, bulk review decisions. What we see as an additional need, however, is for systems to respond to users' own characterisations of 'types' as these are inferred inductively from the ongoing engagement with evidence.

Fluid document interactions - Our data on document review also draws attention to a requirement for tools to support more dynamic document interactions. Evidential documents can be long and integrate various kinds of cross referencing. To better support more efficient data focusing, reviewers would benefit from tools to support the quick identification of sections of interest (e.g. within-document search) and the ability to quickly move to-and-fro between selected document sections in order to cross-refer.

The creation and use of external representations reflects the complementary nature of data and issue focusing, with data focusing apparent in their creation and issue focusing apparent in their use. This distinction highlights two general considerations in the design of external representation tools. The first, which relates to data focusing, is how easily external representations can be generated from data and manipulated to show different subsets of findings. The second relates to how a representation supports the identification of new questions and new areas of interest.

Linking external representation elements to evidence - Elements of a representation are based on evidence, and one kind of operation that was important in a legal investigation was the ability to easily relate entries back to supporting data. This allowed new evidence to be related to old, and supported the resolution of ambiguity and discussion at team meetings. This finding echoes a study in a different domain by Attfield, Blandford, Dowell and Cairns (2008). In that study, the authors evaluated a system designed to support journalists writing news articles using a digital library of past news articles to provide background information. The system, called Newsharvester, allowed the user to search for and collate extracts from past news stories and write the article. Links were maintained next to each extract to allow the review of an originating article at any time. Attfield et al compared this with the same system without the links and a condition in which users printed documents they were interested in. They found that users greatly preferred the linking option, and used the facility to relocate information not previously identified as useful, to better understand the context of information they had already extracted and as part of a more serendipitous search for information.

Findings from the current study similarly show the importance of allowing users to move flexibly between extracted representations that they create and source data, but in a new sensemaking domain. This suggests that it is a finding that generalises well across domains of digital information sensemaking. The effect of such a tool is that a representation becomes a structured index into the source data, supporting flexible access and review.

Filtering external representations in flexible ways - Whilst data focusing is reflected in the extraction and integration of information into an external representation, this may also extend to a need to dynamically filter a representation to provide views of selected aspects. Investigators in our study augmented an existing spreadsheet application to support this kind of filtering for a representation which integrated all strands of the

investigation. This allowed them to reflect on different aspects of the information they had gathered, including reviewing connections across separate lines of enquiry.

The use of a generic spreadsheet application (Excel) as a tool for constructing the chronologies warrants some discussion. We assume that part of the reason lay in the familiarity that the lawyers undoubtedly had with such a generic office suite application. Given this familiarity and the confidence it can engender, it is perhaps an understandable choice. However, the use of spreadsheets would have been less valuable if it wasn't for the adaptation of supporting filtered views. We have shown that the adaptation was not ideal, and this provides an important source of requirements, but the use of spreadsheets can perhaps be understood in terms of a trade-off between utility and risk.

The significance of filtering for the lawyers indicates a requirement particularly pertinent to sensemaking involving extensive amounts of information and also where the construction of a representation is distributed across multiple members of a team. Without such functionality, important relationships may otherwise remain hidden. However, we saw some limitations in the filtering that the investigators used since this was limited to relatively coarse filtering options which depended on pre-assigned codes. Greater flexibility for ad hoc filtering would be possible by allowing representational elements to be selected by search and individual manual selection.

Viewing external representations at different levels of scale and fidelity - The question of how well a representation supports issue focusing is a question about its representational form and how well this allows a sensemaker to discern more focused issues for enquiry. We saw a number of representational forms used which contributed to issue focusing at different stages in an investigation. Initially, low fidelity overviews showing broad communication patterns and reporting lines supported the identification of areas to focus on (people in this case). Once this was achieved, these representations had no further role and were set aside. Later, detailed and selective narrative accounts helped interviewers formulate specific questions to ask witnesses.

The most extensively used representations, however, were the chronologies. These supported issue focusing by mapping out broad time scales within which the investigators could identify periods of particular interest, and by allowing them to identify unusual, unexplained or missing events in the narrative. The combination of representing broad time scales and then using these to fill out detail in selected areas suggests that tools for supporting such representations should provide the capability for reflecting on data at different levels of scale and fidelity. Zoomed-out views would highlight major or landmark aspects of the data in overview, whilst zoomed-in views would show key areas of detail resulting from more targeted, forensic examination of the evidence.

Supporting recursive problem decomposition - Finally, successive focusing occurs through the gradual definition of recursively embedded lines of enquiry. The partitioning of lines of enquiry in the investigations led us to consider their structure in more depth. We were particularly interested in how this might be reflected within systems for

supporting large-scale collaborative sensemaking. The partitioning of elements of a legal 'matter' (or case) into a series of 'issues' is, in fact, an established way for both lawyers and the courts to think about a litigation, and as such it has been explicitly incorporated into some litigation support systems, such as LexisNexis CaseMap . CaseMap allows the user to record key entities in a case, such as people, organisations, documents, evidence, pleadings and events, and to link each to nodes within an issue hierarchy (Dale, 2008). The system shows a case ordered by its events, each linked to other relevant entities, with the option for filtering and creating reports by issue.

Industry commentators have reported advantages of this kind of issue structuring for both individual and collaborative sensemaking (see Dale, 2008). Indeed, CaseMap offers a form of sensemaking representation comparable to the external representations we observed being used in the investigations we studied. However, given its entities, CaseMap is suited to the representation of sensemaking outcomes (the upper stages of the model in figure 1) but less suited to its formative stages and the representation of earlier-stage sensemaking entities such as theories, questions and information seeking strategies. In relation to the process as a whole, we found that lines of enquiry in the investigations had seven distinguishable elements: theories, questions, information seeking strategies, evidence (and evidence collections), knowledge, assigned investigators and lower-level lines of enquiry; the combination of these establishes a context which defines a line of enquiry at any level of granularity.

The result is a framework that lends itself to the design of interactive systems for supporting the challenges of decomposition and integration within large, collaborative sensemaking exercises at both early and late stages. Implementation of the framework would partition lines of enquiry into work contexts established during successive levels of issue focusing. These would then allow investigators to eliminate extraneous information whilst accessing and developing these seven elements as they pertain to their local area of enquiry. By maintaining the hierarchical structuring implicit in issue focusing, however, senior investigators could view these elements from the perspective of higher-level lines of enquiry in order to maintain a view on how the lower-level lines of enquiry integrate into a bigger picture.

9. CONCLUSION

We have presented a qualitative study of three legal investigations that involved large quantities of data and teams of investigators. Our model provides support for, but also extends, the model proposed by Pirolli and Card (2005) of intelligence analysis in a number ways. The model emphasises two essential and complementary aspects of sensemaking: data focusing and issue focusing.

We have also focused in particular on document reviewing and the creation and use of external representations, considered how aspects of these activities reflect different forms of focusing and explored a number of requirements that they place on systems supporting large sensemaking exercises, particularly in a collaborative setting.

Finally, we have analysed the structural composition of the investigations and developed a framework which describes recurring elements associated with multiple, embedded lines of enquiry. We believe that understanding this can inform system design which allows users to reflect on and develop theories, questions, information seeking strategies, evidence and knowledge that are relevant to them at multiple levels of description.

NOTES

Authors' Present Addresses. UCL Interaction Centre, MPEB 8th floor, University College London, Gower Street, London, WC1E 6BT, UK

HCI Editorial Record. (supplied by Editor)

REFERENCES

Attfield, S., & Blandford, A. (2008). E-disclosure viewed as 'sensemaking' with computers: The challenge of 'frames', *Digital Evidence and Electronic Signature Law Review*, *5*, 62–67.

Attfield, S., Fegan, S., & Blandford, A. (2009). Idea generation and material consolidation: tool use and intermediate artefacts in journalistic writing, *Cognition, Technology and Work, 11*(3), 227–239.

Attfield S., Blandford A., Dowell J. & Cairns P. (2008) Uncertainty-tolerant design: Evaluating task performance and drag-and-link information gathering for a news writing task. *International Journal of Human-Computer Studies*. *66*, 410–424.

Baron, J., Braman, R., Withers, K., Allman, T., Daley, M., & Paul, G. (2007). The Sedona Conference best practice commentary on the use of search and information retrieval methods in e-discovery. *The Sedona Conference Journal*, *8*,189–223

Blandford, A., Green, T. R. G., Furniss, D., & Makri, S. (2008). Evaluating system utility and conceptual fit using CASSM. *International Journal of Human–Computer Studies*, *66*, 393–409.

Bodnar, J. W. (2005, May). *Making sense of massive data by hypothesis testing.* Paper presented at the International Conference on Intelligence Analysis . Retrieved from https://analysis.mitre.org/proceedings/index.html;

Dale, C. (2008, June). *CaseMap issue linking in UK civil proceedings.* Paper presented at DESI II, The Second International Workshop on Supporting Search and Sensemaking for Electronically Stored Information in Discovery Proceedings. Retrieved from http://www.cs.ucl.ac.uk/staff/S.Attfield/desi/DESI_II_agenda.html;

Dervin, B. (1993, May). *An overview of sense-making research: Concepts, methods and results.* Paper presented at the Annual Meeting of the International Communication Association. Retrieved from http://communication.sbs.ohio-state.edu/sense-making/art/artabsdervin83smoverview.html

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Klein, G., Moon, B., & Hoffman, R. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems*, *21*(5), 88–92.

Klein, G., Phillips J. K., Rall E. L., & Peluso D. A. (2007). A data-frame theory of sensemaking. In R. Hoffman (Ed.) *Expertise Out of Context: Proceedings of the*

Sixth International Conf. on Naturalistic Decision Making (pp. 113–155). USA: Lawrence Erlbaum.

Johnson, J. C. (1990). *Selecting ethnographic informants.* Newbury Park, CA: Sage.

McNee, S. M., & Arnette, B. (2008). *Productivity as a metric for visual analytics: Reflection on e-discovery*. Paper presented at BEyond time and errors: novel evaLuation methods for Information Visualization, A Workshop of the ACM CHI 2008 Conference. Retrieved from http://portal.acm.org/citation.cfm?id=1377966;

Pirolli, P., & Card, S. (2005, May). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. Paper presented at the International Conference on Intelligence Analysis . Retrieved from https://analysis.mitre.org/proceedings/index.html;

Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). The cost structure of sensemaking. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems* (pp. 269–276). New York: ACM Press.

Savolainen, R. (2006). Information use as gap-bridging: The viewpoint of sense-making methodology. *Journal of the American Society for Information Science and Technology*, *57*(8), 1116–1125.

Sereno, B., Buckingham Shum, S., & Motta, E. (2005). ClaimSpotter: An environment to support sensemaking with knowledge triples. In *Proceedings of the 10th International Conference on Intelligent User Interfaces* (pp. 199–206). New York, ACM Press.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. 2nd ed. London, England: Sage.

Spence, R. (1999). A framework for navigation. *International Journal of Human-Computer Studies,* 51, 919–945.

Starbuck, W. H., & Milliken, F. J. (1988). Executives' perceptual filters: What they notice and how they make sense. In: D. C. Hambrick (Ed.) *The Executive Effect: Concepts and Methods for Studying Top Managers* (pp. 35–65). Greenwich: JAI Press.

Thomas, J. B., Clark, S. M., & Gioia, D. A. (1993). Strategic sensemaking and organisational performance: Linkages among scanning, interpretation, action and outcomes. *Academy of Management Journal*, *36*, 239–270.

Qu, Y., & Furnas, G. W. (2005). Sources of structure in sensemaking. In Proceedings of the CHI '05 Conference on Human Factors in Computing Systems (pp. 1989–1992). New York: ACM Press.

Weick, K. (1995). *Sensemaking in organisations*. London, England: Sage.

FOOTNOTES

(Make a copy of all footnotes on a separate page here. This only has to be done for the final submission for production. During the review process, it is okay to just have footnotes at the bottom of pages.)

1. xxx

2. xxx

3. xxx

FIGURE CAPTIONS

Figure 1. **A model describing the process of the investigations from the case-studies.**

Figure 2. **Approximate mappings between processes in the model of intelligence analysis of Pirolli & Card (2005) and the model of legal sensemaking developed in the study presented here.**

Figure 3. **An anonymised event entry from one of the chronologies.**

FIGURES
(IT'S BEST TO PUT ONE FIGURE PER PAGE)

Figure 1. A model describing the process of the investigations from the case-studies.

Information volume *low*
Information relevance *high*
Information structure *high*

Report, discuss ①

Issue   *Claims*

Create/amend/review ②
external representation

Issue   *Transcripts*   *Issue*   *Coded collections*

Interview witnesses ⑥   Review & code ③
documents

Issue   *Results*

Search documents ④

Issue   *Database*

Recover documents and
populate database ⑤

Information volume *high*
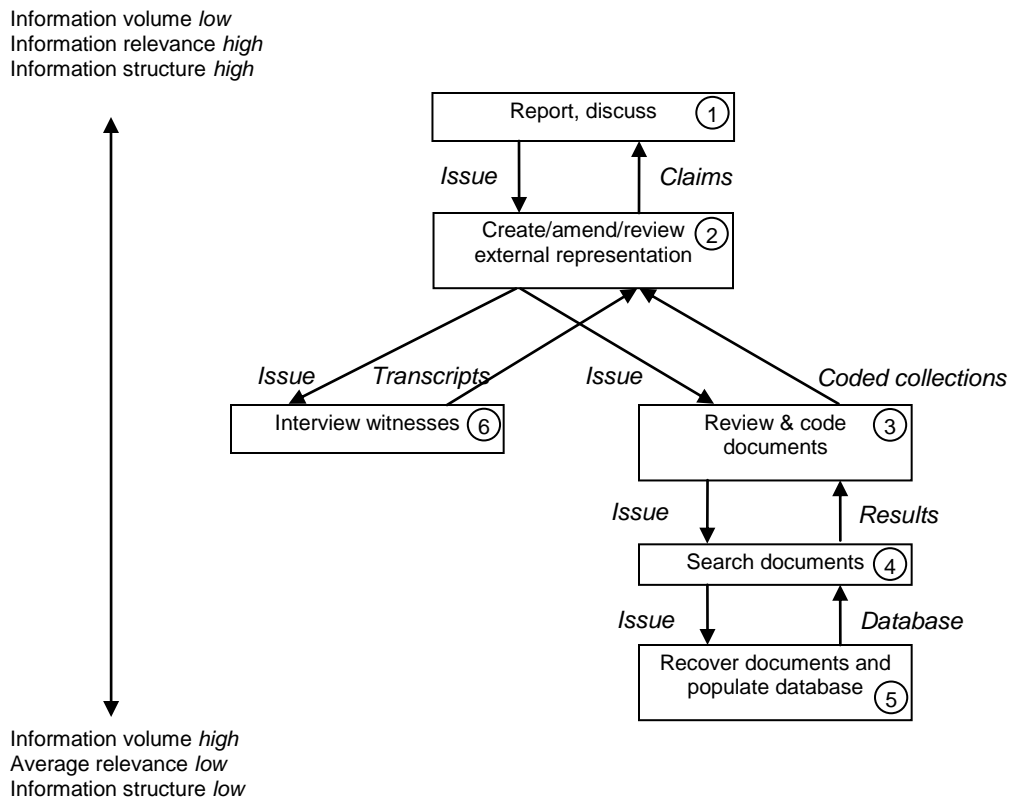Average relevance *low*
Information structure *low*

Figure 2. **Approximate mappings between processes in the model of intelligence analysis of Pirolli & Card (2005) and the model of legal sensemaking developed in the study presented here.**

Model of sensemaking in intelligence analysis (Pirolli & Card, 2005)	*Model of sensemaking in legal investigations*
(no equivalent)	Stage 5: Recover documents and populate database.
Search and Filter - Selection of a subset for further processing (held in a "shoebox").	Stage 4: Document searching, delivering results sets.
	Stage 3: Reviewing and coding documents, resulting in coded collections.
Read and extract ('nuggets' stored in an "evidence file").	Stage 2: Create / amend / review external representation.
Schematize - Represent information in a structured way (schemas).	
(no equivalent)	Stage 6: Interview witnesses
Build case - A theory or case is built by additional marshalling or evidence.	*(Incorporated throughout the model.)*
Tell Story – Presentation to some audience.	Stage 1: Report / discuss.

Figure 2.   **An anonymised event entry from one of the chronologies.**

Date	Time	Event/Document	People Involved/ Author/Recipient	Evidence / File Reference
8th Nov	7.45	{company A} Meeting in {country A} (time is {person B} flight departure from {location A} to {location B}) with return to {location A} for 12.55 on 9th Nov. {person I} to pick up {person B} at Airport	{person I}, {person B} and {person H} in {location C}	Email between {person I}, {person I}, {person H} and {person F}/Doc ID 169246