An Approach to Human-Machine Teaming in Legal Investigations Using Anchored Narrative Visualisation and Machine Learning

Simon Attfield[†]
Department of Computer Science
Middlesex University
London, UK
s.attfield@mdx.ac.uk

Bob Fields
Department of Computer Science
Middlesex University
London, UK
b.fields@mdx.ac.uk

David Windridge
Department of Computer Science
Middlesex University
London, UK
d.windridge@mdx.ac.uk

Kai Xu
Department of Computer Science
Middlesex University
London, UK
k.xu@mdx.ac.uk

ABSTRACT

During legal investigations, analysts typically create external representations of an investigated domain as resource for cognitive offloading, reflection and collaboration. For investigations involving very large numbers of documents as evidence, creating such representations can be slow and costly, but essential. We believe that software tools, including interactive visualisation and machine learning, can be transformative in this arena, but that design must be predicated on an understanding of how such tools might support and enhance investigator cognition and team-based collaboration. In this paper, we propose an approach to this problem by: (a) allowing users to visually externalise their evolving mental models of an investigation domain in the form of thematically organized Anchored Narratives; and (b) using such narratives as a (more or less) tacit interface to cooperative, mixed initiative machine learning. We elaborate our approach through a discussion of representational forms significant to legal investigations and discuss the idea of linking such representations to machine learning.

KEYWORDS

eDiscovery, TAR, anchored narratives, machine learning, sensemaking, distributed cognition.

1 Introduction

Legal investigations, particularly in regulatory and litigation contexts, tend to be characterised by the simultaneous challenge and opportunity of very large numbers of documents as a source of evidence. Given this complexity, investigators tend to create external representations or 'models' of the investigated domain as a means of cognitive offloading and creating structures for supporting reflection, insight and collaboration. Interactive Visualisation and Machine Learning have created interest as tools for supporting the identification of relevant documents as a prelude to such investigations. However, less attention perhaps has been paid to the potential for combining these technologies within the investigation process itself. We argue that such an approach might support more rapid convergence on investigatory narratives that matter by:

a) allowing users to visually externalise their evolving mental models of an investigation domain in the form of thematically organized Anchored Narratives;
b) using such narratives as a (more of less) tacit interface to cooperative, mixed initiative machine learning.

We argue that the effect of this can be cooperative human-machine teaming through an evolving symbiotic relationship between three distinct but interconnected elements: user cognition, external representation and machine learning. We develop our case by reviewing the role of external representations in investigatory sensemaking focussing on cognition and collaboration. We then consider harnessing machine learning as a tacit means of anticipating investigatory goals and enhancing access to relevant data.

1 Background - External Representations for Investigatory Sensemaking

The creation, augmentation and use of representations, whether internal (in the head) or external (in the world), are a central part of sensemaking. This idea is reflected in most significant theories and models of sensemaking. For example, Klein et al. [1] discuss the role of mental 'frames' in sensemaking, and Pirolli and Card [2] emphasise the way intelligence analysts externally structure

In: Proceedings of the First International Workshop on AI and Intelligent Assistance for Legal Professionals in the Digital Workplace (LegalAIIA 2019). June 17, 2019. Montreal, QC, Canada.

information into representations as part of a wider sensemaking process (referring to this step as ‘schematization’).

External representations, when created, can be intimately involved in the cognitive processes of sensemaking. The approach of Distributed Cognition is predicated on the idea that cognitive activities make use of external as well as internal representations, with external representations seen not only as sources of information, but as structures that transform the cognitive task itself [3]. Having an effective representation can lead to different and better strategies for carrying out a task, better performance, and lower mental effort. The form and properties of external representations can lead to changes in cognitive processes as these become integrated into and participate within these processes. Distributed Cognition aims to dissolve the traditional division of inside/outside the individual when analysing cognition in order to explore the complex relationships between people, artefacts and technology when accounting for how thinking gets done.

In an attempt to render the concepts of distributed cognition more useful and applicable to the design of human-computer interaction, Wright et al. [4] identified a collection of ‘abstract information resources’ that can form a part of the process of carrying out activities. Such abstract structures can be represented in a variety of forms, embodied in physical media (possibly as a result of the design of interactive technologies) or located in the minds of members of a distributed cognitive system. More recently, Attfield et al. [5] applied this idea to sensemaking, identifying a taxonomy of abstract information resources that can be represented internally or externally during sensemaking and which are transformed during the process of sensemaking. These resources include representations of the domain (specific or general), intents (high-level values to low-level and goals), and representations of action (possible, planned or performed). Actors involved in the sensemaking activity may make use of any or all of these, and the nature of their representation determine how they may do so.

Narrative

External representations can take many forms depending on the entities and relationships being represented. Faisal, Attfield and Blandford [6] proposed six basic types: spatial, sequential (including narrative), networks, hierarchical, argumentation structures and faceted. Here we discuss two types which are important for constructing domain representations during investigatory sensemaking: narrative and argument. Later we extend this with a discussion of thematic organisation.

For example, Attfield and Blandford [7] reported a study of the cognitive work of lawyers involved in some large corporate investigations. As part of their work, the lawyers represented their analyses in the form of sequences of connected events or chronologies, created around different themes of an investigation. These narrative representations, which were ultimately very large, played a central role in the way that the lawyers thought about and collaborated around the investigations and they were central in the generation of insights. The lawyers reported that this was a natural way for them to think about an investigation.

Research shows that narrative representations play a particularly important role in the way that people reason about evidence. For example, Pennington and Hastie [8] conducted a series of studies into the way that jurors mentally comprehend evidence in legal cases. They found that, irrespective of how evidence was presented, jurors structured it in terms of narratives that made sense to them. Not only that, they added information to make the stories make more sense. This finding is typical of studies into evidential reasoning and provided a basis for what Pennington and Hastie called their Story Model. According to the Story Model people find it easiest to make sense of legal evidence through narratives that they construct in order to explain the evidence. Importantly, the resulting narrative is constructed not just from the evidence, but by reasoning from evidence to explanation.

Argumentation

Investigatory sensemaking involves drawing conclusions from evidence using generalised beliefs about the way the world works [9]. For example, an investigator may infer from reading an email in which person *a* thanks person *b* for a gift, that a gift was exchanged, with this inference depending on both the text in the email and the more general belief that people don’t usually express gratitude in this way when in reality no gift has been exchanged. This is an example of an abductive inference (reasoning to the best possible explanation) which is characteristic of investigatory sensemaking. Many thousands of such inferences may be made during an investigation, and given their generally defeasible nature, it can be important that they are amenable to review. For example, Attfield and Blandford [7] reported on the way that lawyers maintained links from chronology entries to supporting documentary evidence and traversed them frequently.

Based on a study of how Dutch judges reasoned about cases, Wagenaar [9] observed their prominent use of narrative connections and argumentation links and developed from this the notion of Anchored Narratives. An anchored narrative is a hybrid representational form combining narrative with argumentational links to supporting evidence. Bex [10] has used this approach to develop a formal theory that combines stories with evidential arguments in a hybrid framework for structured argumentation.

Figure 1 shows an example of an Anchored Narrative in which events are represented as a connected narrative (from top to bottom in figure 1) attached to supporting evidence (where available).

Significantly, events are anchored, not only in evidence, but within the context of the unfolding story. The plausibility of each event is then judged not solely in virtue of its supporting evidence, but also by the support of plausibility afforded by its position in the surrounding narrative and how this relates to generalised beliefs about how the world words. Figure 1 also shows the representation of multiple competing narratives with a point of divergence based on evidence from interview 1 and interview 2. Explicitly representing such competing conclusions can be a helpful in a context of defeasible reasoning where multiple interpretations or claims may be explicitly considered.

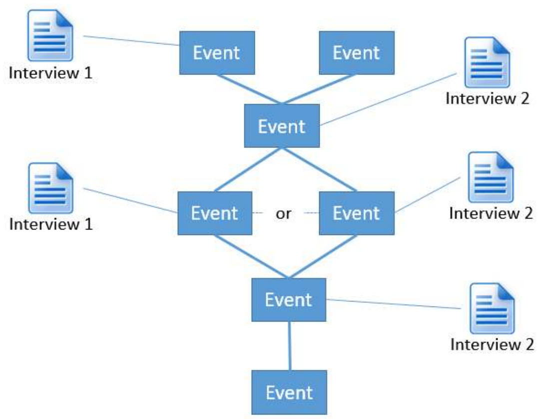

Figure 1 - An example of an Anchored Narrative

3. Interactive Visualisation

Data visualisation has a capability of supporting insight from abstract data by leveraging the power of the human perceptual system to convert cognitive problems into perceptual problems [11]. It can, reveal insights that are otherwise difficult to discover [12]. Interest has developed in extending data visualisation beyond the display of large datasets to support other aspects of sensemaking (including what Pirolli and Card [2] referred to as schematization) and also to enhance human sensemaking by coupling representations to computational components such as machine learning; this is an approach emphasised by Visual Analytics. Figure 2 shows Kohlhammer et al's [13] model of the Visual Analytics process. The main difference between this model and a data visualisation pipeline is the addition of the 'model' component (representing the product of automated data analysis such as machine learning) and its interactions with other components.

Visual Analytics tools can facilitate the process of constructing narratives from data and capturing the data and analysis that lead to them. Figure 3 shows a tool we have developed called SenseMap [14]. SenseMap provides the user with a freeform interactive space (right) which can be used for constructing anchored narratives from data. The user interacts with data and represents interesting discoveries as a boxes in the main panel (right) by a simple click. Discoveries can be moved freely to form thematic groups or evolving narratives. SenseMap also captures the provenance of the discovery such that clicking on a discovery will restore the original data source i.e. discoveries are anchored in source data.

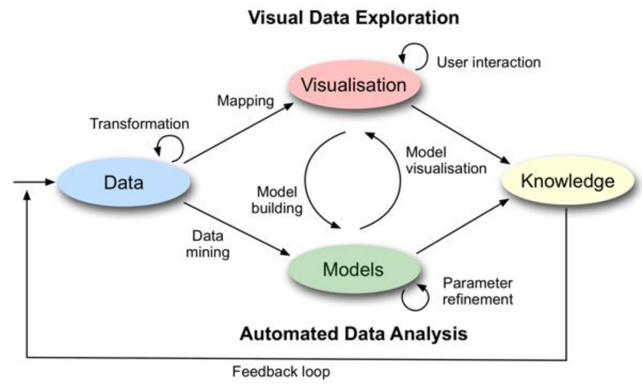

Figure 2 - Model of the Visual Analytics Process from Kohlhammer et al. (2011)

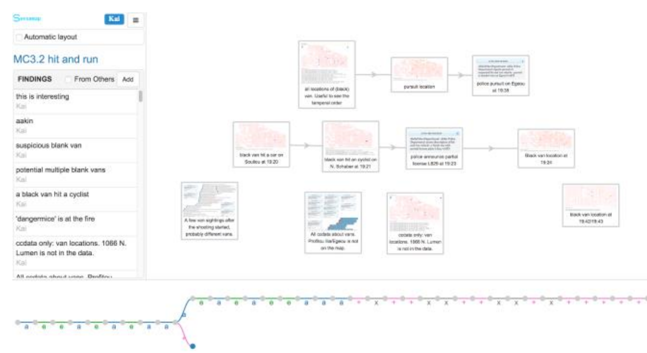

Figure 3 - The SenseMap allows interactive construction of episodes or narratives from discoveries. Each discovery (or event) is represented as a box, which can be grouped or connected to form a episode/narrative.

In addition to organizing discoveries into evolving narratives, we see value in organising narratives into identifiable episodes and themes. Investigations can be complex. Investigation teams have been shown to divide analyses along the lines of episodes and themes as these become apparent. This has the value of reducing cognitive complexity and supporting the division of labour [7]. Different episodes and themes will also have different theories of relevance, and we anticipate that such structuring can be exploited by machine learning for the (further) identification of relevant information in large evidential collections. Hence, we propose structuring events at the interface into discrete episodes and by hierarchical theme. Figure 4 shows a conceptual model of this idea in which connected events form episodes, which in turn become components in anchored narratives. Similarly, discoveries can be grouped as hierarchically organised themes.

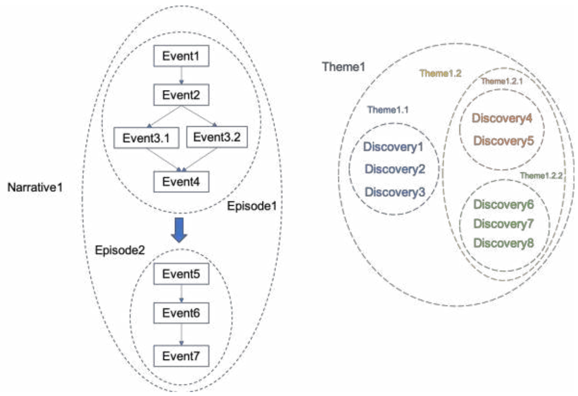

Figure 4 - Hierarchical structure of events and discovery based on time and theme.

Besides interfacing with users, there are many examples in which Visual Analytics can provide the interface between domain experts and machine learning algorithms [15]. Some of these allow users to provide feedback on the machine learning outcomes (such as classification or prediction), and improving the underlying machining learning model. These are often known as active/interactive learning. Other methods focus on exposing the inner workings of a machine learning model, i.e. how the model makes classification or prediction. This is known as explainable AI (XAI) and critical to the issues related to model transparency such as model bias and user trust. These issues are closely related to the discussions in the next section.

4. Coupling with Machine Learning

The nature of the problem as defined implicates a unique nexus between machine learning, human computer interfacing (HCI) and machine representation. While domain summarisation is a well-established aspect of machine learning-based textual and image analytics, it is necessarily a passive, feedforward process unless explicit human-in-the-loop considerations are incorporated. Our problem, when cast in machine learning terms, can be specified as the building of a recommender system for returning evidence in relation to significant, or user-salient, aspects of the chronological data stream at arbitrary levels of hierarchical aggregation/representation. The problem of relevance has both a 'vertical' (abstractive) as well as 'horizontal' (chronological) aspects, given that narrative sequences and events (evidence) exist in a subsumptive relationship.

Thus, we seek a system in which user and machine exist within a convergent hermeneutic feedback cycle, for which potentially supportive evidence is returned to the user on the basis of the current narrative representation at some appropriate level of hierarchical aggregation. In response, the user feeds back information on the utility of this evidence as part of the constructed narrative sequence (at its appropriate level of representation) in order to either to further develop an existing , or else initiate a novel representational frame.

The hierarchical aspect of the problem significantly multiplies the complexity of the machine learning methodology required to approach it. In particular, sequence-based recommender systems typically rely on query proximity within some appropriate metric (or quasi-metric) space. However, we here require that the proximal region to the user's query (anchor) within 'narrative space' takes into account arbitrary levels of aggregation (or narrative coarse-graining) in a way that both encompasses (potentially evolving) user preference and does not burden the user with excessive feedback requirements.

To this end, we propose to use active learning within the context of the querying of the sequential aggregation so as to achieve the optimal reduction in the bandwidth of user feedback required to obtain a convergent recommender platform for narrative construction. Active learning is a process by which machine learning hypotheses are fed back to the user (here via appropriate visualisation techniques) in a manner such that preference feedback to the machine learner is optimally exploited to improve learning performance. This typically provides a logarithmic improvement in user feedback requirements with respect to labelling effort/user load associated with classical machine learning approaches. Maximally rapid mutual convergence on hypotheses of interest to the user is thus ensured, such that human and machine mutually adapt to take advantage of their respective capabilities in the most synergistic fashion.

The proposed system would thus exploit feedback from the user in its learning-loop in order to develop a better tailored model of narrative and chronological salience via the use of active learning to pro-actively present representation alternatives to the user across the interface. Crucial to bootstrapping this process is an initial 'seed' set of domain-annotated data, constituting an initial extraction of salient descriptors from the narrative stream.

5. Discussion/Conclusion

We believe that there is a prospect of achieving high quality, synergistic relationships between human and machine cognition in which one supports the other to enable rapid convergence on significant and important narratives during investigatory sensemaking. An approach that we propose involves the use of interactive visualisation to allow users to construct structured external representations of the investigated domain, coupled to machine learning models that might exploit this structure to model and predict investigators' evolving interests around different parts of the investigation. This is essentially a mixed initiative approach to sensemaking in which computational and human agents establish common ground around investigatory goals through common access to a visualisation interface. In future work we seek to develop a prototype of this approach to provide proof-of-concept validation and to develop the techniques involved through iterative empirical trials.

REFERENCES

[1] Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. In *Expertise out of context: Proceedings of the sixth international conference on naturalistic decision making* (pp. 113-155). New York, NY, USA: Lawrence Erlbaum.

[2] Pirolli, P., & Card, S. (2005, May). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis* (Vol. 5, pp. 2-4).

[3] Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.

[4] Wright, P. C., Fields, R., & Harrison, M. D. (2000). Analyzing Human-Computer Interaction as Distributed Cognition: The Resources Model. *Human-Computer Interaction*, *15*(1), 1–41.

[5] Attfield, S., Fields, B., & Baber, C. (2018). A resources model for distributed sensemaking. *Cognition, Technology & Work*, *20*(4), 651–664.

[6] Faisal, S., Attfield, S., & Blandford, A. (2009). A classification of sensemaking representations.

[7] Attfield, S., & Blandford, A. (2011). Making sense of digital footprints in team-based legal investigations: The acquisition of focus. *Human–Computer Interaction*, *26*(1-2), 38-71.

[8] Pennington, N., & Hastie, R. (1991). A cognitive theory of juror decision making: The story model. *Cardozo L. Rev.*, *13*, 519.

[9] Wagenaar, W. A. (1995). Anchored narratives: A theory of judicial reasoning and its consequences. *Psychology, law and criminal justice: International developments in research and practice*, 267-285.

[10] Bex, F. (2015). An Integrated Theory of Causal Stories and Evidential Arguments. In *Proceedings of the 15th international conference on artificial intelligence and law* (pp. 13-22). ACM.

[11] Few (2013) Data Visualization for Human Perception, In: The Encyclopedia of Human-computer Interaction. Interaction Design Foundation.

[12] Card, S. K., Mackinlay, J., & Shneiderman, B. (Eds.). (1999). Readings in Information Visualization: Using *Vision to Think*. Morgan Kaufmann.

[13] Kohlhammer, J., Keim, D., Pohl, M., Santucci, G., & Andrienko, G. (2011). Solving problems with visual analytics. *Procedia Computer Science*, *7*, 117-120.

[14] Nguyen, P. H., Xu, K., Bardill, A., Salman, B., Herd, K., & Wong, B. W. (2016, October). Sensemap: Supporting browser-based online sensemaking through analytic provenance. In *2016 IEEE Conference on Visual Analytics Science and Technology (VAST)* (pp. 91-100). IEEE.

[15] Endert, A., Ribarsky, W., Turkay, C., Wong, B. W., Nabney, I., Blanco, I. D., & Rossi, F. (2017, December). The state of the art in integrating machine learning into visual analytics. In *Computer Graphics Forum* (Vol. 36, No. 8, pp. 458-486).