Conversational Services for Multi-Agency Situational Understanding

Alun Preece
Cime and Security Research Institute
Cardiff University
Cardiff, UK
Email: PreeceAD@cardiff.ac.uk

Dave Braines
Emerging Technology
IBM United Kingdom
Hursley Park, Winchester, UK
Email: dave_braines@uk.ibm.com

Abstract

Recent advances in cognitive computing technology, mobile platforms, and context-aware user interfaces have made it possible to envision multi-agency situational understanding as a 'conversational' process involving human and machine agents. This paper presents an integrated approach to information collection, fusion and sense-making founded on the use of natural language (NL) and controlled natural language (CNL) to enable agile human-machine interaction and knowledge management. Examples are drawn mainly from our work in the security and public safety sectors, but the approaches are broadly applicable to other governmental and public sector domains. Key use cases for the approach are highlighted: rapid acquisition of actionable information, low training overhead for non-technical users, and inbuilt support for the generation of explanations of machine-generated outputs.

Introduction

Decision-making in the governmental and public service sectors commonly involves multiple agencies working together. Decision cycles often need to be relatively rapid, and processes must be agile, based on a best-possible understanding of the current, evolving situation. The term *situational understanding* is commonly used in a security and public safety context to refer to the 'product of applying analysis and judgment to… determine the relationships of the factors present and form logical conclusions concerning threats… or mission accomplishment, opportunities for mission accomplishment, and gaps in information' (Dostal 2007). In this paper, we focus on the problem of achieving *multi-agency situational understanding* (MASU) by means of approaches from cognitive computing (Kelly and Hamm 2013), chiefly natural language processing, human-computer interaction, and knowledge representation and reasoning.

We assume MASU domains where data and sources of data may be plentiful, but where it may be difficult to assemble the right set of data and analytic services to enable decisions to be made in an effective and timely manner. MASU emphasises the collection and fusion of actionable information, to provide a clear picture of options, threats, and consequences (Broome 2012). Individual decision-making actors may be at various levels in an organisation, from high level commanders located near the centre of an information network, to lower level operatives at or near the edge of the network. Recent thinking in the field of command and control emphasises the empowerment of individuals at the network edge who, prior to the widespread availability of mobile information and communication platforms, have traditionally been unable to exploit the best-available actionable information (Alberts and Hayes 2003). Empowering such individuals in domains such as emergency response, policing, and military operations is viewed as highly desirable since they are capable of directly affecting the evolving situation through their decisions and actions. Because the situation unfolds quickly, the information architecture that supports MASU must be highly responsive to changes in the decision-maker's requirements and the availability of relevant sources.

Information needs to flow in two directions in an agile MASU service-oriented architecture:

- A *forward chain* from data to decision: data is collected by sensors (for example, imagery or audio data) or retrieved from other sources (e.g., media feeds, eyewitness reports), processed by analytics services, and delivered to a decision maker according to their information requirements.
- A *backward chain* from a decision maker's requirements to relevant analytics services able to provide the needed information, and to data sources that can 'feed' those services.

Traditionally, prior research and development work in this area has tended to focus on the data-driven forward chain; the backward direction has received less attention. Nevertheless, rapid construction of these backward chains has been evident in recent well-publicised emergency responses. For example, during the Fukushima Daiichi nuclear disaster in the wake of the 2011 earthquake in Japan, it became necessary urgently to track the spread of radiation, resulting in the rapid construction of geospatial visualisation services fed by networked Geiger counters — including private devices shared via early Internet of Things (IoT) technology[1].

[1] http://www.wired.com/opinion/2012/12/20-12-st_thompson/

This example is an instance of the general problem: how to rapidly construct pipelines by working backwards from an intended decision (or hypothesis or query), identifying useful data analytics services and underlying data sources that can meet the decision maker's requirements.

Recent advances in cognitive, mobile, and context-aware technologies enable an even more flexible and agile kind of MASU system. Data sources are becoming increasingly self-describing and communicative. Autonomous robotic systems, together with increasingly computationally-capable IoT devices operating in decentralised networks open up greater potential for collective intelligence and self-organisation at or near the edge of the network, close to where data sources are often situated (Fraga-Lamas et al. 2016). Moreover, increasingly capable mobile devices have freed decision-makers to operate effectively in contexts much nearer the 'front line'. Widespread user familiarity with commercial products such as Amazon's Alexa[2], Google Now[3], Apple's Siri[4], and IBM's Watson[5] have raised user expectations in terms of what MASU systems should be able to deliver. In this context, MASU can be viewed (Figure 1) in terms of decentralised 'conversational' exchanges between agents with different specialisms: the data sources, analytic services and decision-makers. In this perspective, chains of interaction can start anywhere in the network and flow in any direction, backwards and forwards.

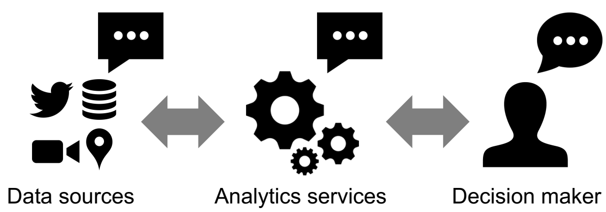

Figure 1: MASU viewed as conversational interaction

This paper reports on an approach to MASU intended to prioritise agile decision-making via human-machine collaboration across multiple agencies and teams. We draw on examples from work in the security and public safety sectors, but the approaches are broadly applicable to other governmental and public sector domains. The next section outlines a conceptual architecture for MASU systems. We then describe the core technical approach: a combination of natural language-based knowledge management and human-machine interaction. The latter sections explain how the approach supports sensemaking activities to connect data with hypotheses and decisions, and we conclude the paper with pointers to future work.

[2]https://developer.amazon.com/alexa

[3]https://en.wikipedia.org/wiki/Google_Now

[4]http://www.apple.com/ios/siri/

[5]http://www.ibm.com/watson/

Conceptual Architecture for MASU Systems

Figure 2 depicts a layered conceptual architecture for a MASU system. The bottom layer comprises a collection of data sources, accessible across the multiple partner agencies. These sources include 'hard' physical sensor data and 'soft' human-originated content. At each layer above the data layer, the figure illustrates the primary cognitive computing techniques applicable to processing products of the layer(s) below:

HCC human-computer collaboration

KRR knowledge representation and reasoning

MAS multi-agent systems

ML machine learning

NLP natural language processing

VSP vision and speech processing

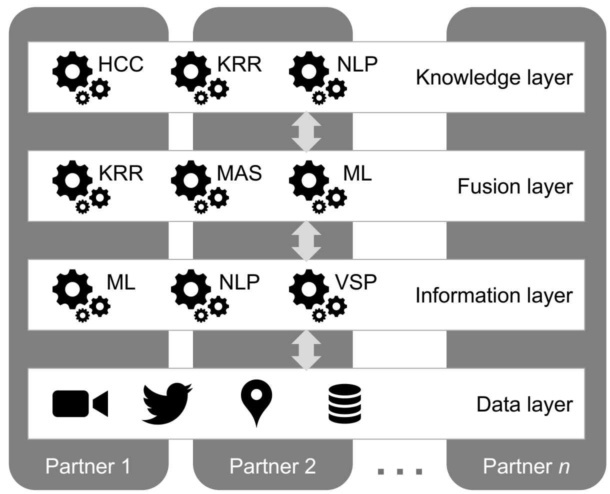

Figure 2: MASU layered conceptual model distributed across multiple partners

The *information layer* uses incoming data streams to identify semantic entities together with their relationships at multiple scales of granularity. The resulting information representation explicitly or implicitly encodes a history of past observations. ML, NLP, and VSP are the predominant cognitive computing techniques used in this layer. MAS techniques play an important role in communication and coordination among distributed processing services across the multiple agencies. The *fusion layer* utilises algorithms and techniques to estimate the current state of the world from information derived at the lower layer; KRR and MAS have a significant role here in terms of world-modelling and interagency communication, though ML, NLP, and VSP techniques are applicable at this level also. The *knowledge layer* then uses the current world model and histories of past observations to explicate likely future states, with KRR again playing a key role in reasoning about the world; HCC and NLP approaches handle the necessarily-rich interaction between this layer and the human user.

The upper layers in Figure 2 need to be open to humans to (i) provide expert knowledge for reasoning and (ii) be capable of generating explanations of the reasoning performed by the system. Information flows in two directions between the layers: in the upward (forward) direction, inferences made at the lower layer act as input for the higher layer; in the downward (backward) direction, information is used to adjust the model and algorithm parameters and change the tasking/querying of the data sources. The requirements to create more agile systems to support MASU necessitates developing mature models and algorithms that can, over time, reduce the need for human intervention and increase machine autonomy, without entirely replacing human engagement and oversight.

The architecture emphasises the systems view of cognitive computing as requiring a hybrid set of computational techniques. A key issue is how to combine KRR and ML/NLP/VSP techniques. Figure 3 illustrates these as separate subsystems with points of semantic articulation. For example, classes forming the output of an NLP or VSP classifier form part of a KRR ontology, allowing classified instances to be fed upward (i.e., forward in Figure 1) to become part of a represented model and used in reasoning processes. Generally, the output of a ML/NLP/VSP classifier will have some measure of associated uncertainty, that also needs to be fed upwards.

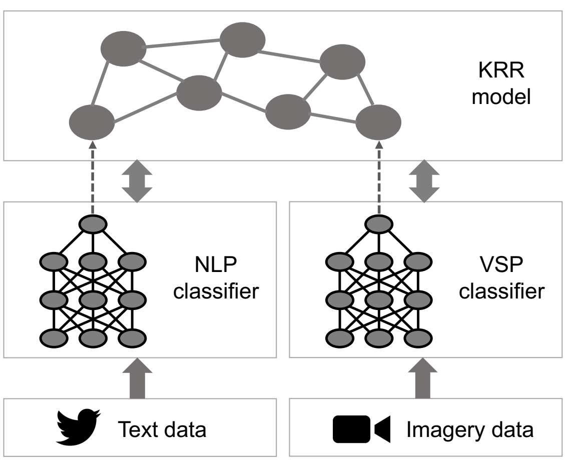

Figure 3: Linking ML/NLP/VSP and KRR approaches in a MASU system

Natural Language Interaction and Knowledge Management

The main elements of our conversational model (Preece et al. 2014) are shown in Figure 4. The model supports both human-machine and machine-machine dialogues, differentiating between natural language (NL) and controlled natural language (CNL) content. NL content is represented as text but its origin can be speech, directly-typed input, or material gathered from external sources including documents, short text messages, or social media. It can also be originated by machine agents. A CNL is a uniform information representation that is both machine-processable and human-consumable (Kuhn 2014) and therefore provides a means to express information *models* as well as structured *instance data*. In this section, we first briefly introduce the CNL used in this work, before discussing the conversational protocol. ITA Controlled English (CE) is available in both Java (full[6]) and JavaScript (lightweight[7]) open source implementations. An example CE model definition is shown below.

```
conceptualise an ˜ event ˜ E that
 has the time ST as ˜ start time ˜ and
 has the time ET as ˜ end time ˜ and
 ˜ involves ˜ the agent A and
 ˜ is located at ˜ the place P.

conceptualise a ˜ protest ˜ P that
 is an event.
```

A `conceptualise` CE sentence introduces a new concept in a model. New terms in the model appear between the tilde (˜) symbols. The example defines two concepts, `event` and `protest`, the latter being a child of the former. The concept `event` is defined as having properties `start time` and `end time` and relationships to other concepts: the relationship `involves` relates an `event` to an `agent` (a human, machine agent, or organisation) and the relationship `is located at` relates an `event` to a `place`. The example below shows the CE that creates an instance of the concept `protest`.

```
there is a protest
 named 'Main Plaza protest' that
 has the time '2017-11-09T15:07:44Z'
   as start time
 and involves the group 'Violet Group' and
 is located at the place 'Main Plaza'.
```

This instance, named `Main Plaza protest`, has UTC time `2017-11-09T15:07:44Z` as the value of its `start time` property; it has an `involves` relationship with a `group` instance named `Violet Group` and an `is located at` relationship with a `place` instance named `Main Plaza`.

CE is intended to be human-readable though generally we expect users to input information in NL and we use NLP to derive CE from the NL input. The simplest approach to interpreting NL as CE uses a bag-of-words technique to map elements of NL sentences to CE models. This approach has proven effective in both laboratory and field studies (see (Preece et al. 2017) and the next section).

The conversational protocol illustrated in Figure 4 consists of four distinct types of interaction: *confirm*, *ask/tell*, *why*, and *gist/expand*. A conversation can begin with any of these except *why*. Contextual state is maintained during an interaction and between interactions via the CE knowledge base (KB) constructed as a result of the interactions.

Confirm interactions: These handle mapping of NL input to CE. When the NL input originates from a human user, this interaction typically involves showing the user a piece of CE generated via NLP from their input, and asking them

[6]https://github.com/ce-store/

[7]http://cenode.io

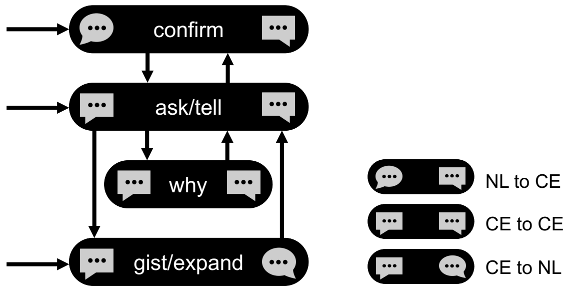

Figure 4: Model for human-machine conversational interactions

explicitly to confirm or edit it. When the interpretation has low ambiguity or does not originate directly from a user — for example, if it comes from an external social media source — the explicit confirmation step may be skipped. In either case, the system retains the original NL in case it needs to be revisited later.

Ask/tell interactions: NL can be ambiguous to the point where it is not always possible to determine whether a sentence is a query or states a fact. Mapping NL to CE removes such ambiguity and, from this point, agents can engage in query-response exchanges shown in Figure 4 as *ask/tell* interactions. Such an exchange can be as simple as a single *tell* — it does not need to involve an *ask*.

Why interactions:A key feature of the model is to allow any agent to obtain an explanation, justification or provenance for a piece of information in the form of rationale. For example, if told a fact, an agent may seek an explanation of how that fact was obtained or inferred. This is the purpose of the *why* interaction. CE has a specific syntax for rationale sentences, beginning with the keyword `because`.

Gist/expand interactions:CE is intended to be human-readable but it is often rather verbose and can be especially difficult to comprehend on mobile devices where smaller screens favour shorter messages. A machine agent may generate NL purely for the convenience of human users, e.g., to make some output more easily readable when the human user is engaged in tasks that require digestible information; we refer to such generated NL as 'gist' and the conversational protocol requires any agent that issues gist to be able to provide a full CE expansion of this if required by the recipient: this is shown in Figure 4 as the *gist/expand* interaction. For the generation of gist from CE we use a simple template-based approach that preserves the mapping from the original CE to the gist form, to allow expansion back into full CE if requested by the recipient.

Sensemaking for MASU

We view decision-making in the MASU context in terms of the sensemaking process, commonly defined as a set of interconnected loops (Figure 5). In the *foraging loop*, data is gathered from the external environment and assembled into a body of evidence. Then, in the *sensemaking loop*, schematised evidence is connected to hypotheses and cases are built to support decisions. Feedback loops exist between each pair of successive steps in the process. The progression of the process from bottom to top and from left to right reflects increasing structure in the information artefacts created as well as increasing effort on the part of the humans and machine agents involved.

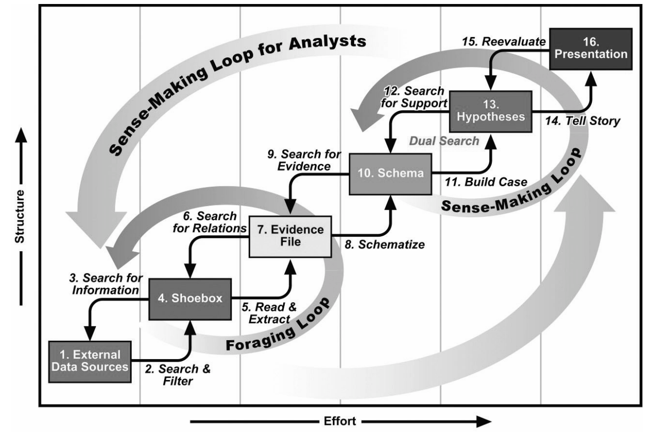

Figure 5: The sensemaking process for intelligence analysis (Pirolli and Card 2005)

The conversational model in Figure 1 and conceptual architecture in Figure 2 are a means to operationalise the sensemaking process for MASU, specifically by increasing the level of automation in setting up data analytics pipelines to improve operational agility and empower human actors at the network edge (Alberts and Hayes 2003) as explained in the introduction. We hypothesise that, a result of this increased automation and agility, 'edge' users become more active participants in the sensemaking process, shaping information products by means of ask/tell interactions. Users, especially those at the network edge, are often themselves sources of information — they are human sensors (Srivastava, Abdelzaher, and Szymanski 2012). Our previous work examines a number of use cases including spot reporting, crowdsourcing, asset tasking, and fusion of soft and hard information (Preece et al. 2014).

We observe that the backward transitions between each successive pair of steps — detailed along the top of Figure 5 — take the form of queries (i.e., *asks*), while the forward transitions — detailed along the bottom of the figure — impart information (i.e., *tells*). This is consistent with the *speech act* view of conversational exchanges in both linguistics and software agent communication (Austin and Urmson 1975; Labrou and Finin 1998). Viewed at a macro scale, the process of adding structure that comes with increasing effort, moving upwards and rightwards in the figure, corresponds to the view of conversation as co-constructing shared informational artefacts (Pask 1972). As the conversation progresses, the potential for increased structure, increased awareness of the different conversational partners' worldviews (conceptual models), and increased alignment of the agents to the context of the current task can occur.

Our approach has been evaluated and refined via a series of field exercises (2014–present). The primary goal of each exercise was to monitor a large-scale event through a combination of real-time social media analysis and ethnographic fieldwork; testing the information architecture and tools described in the previous sections were secondary aims (Innes et al. 2016; Roberts et al. 2017).

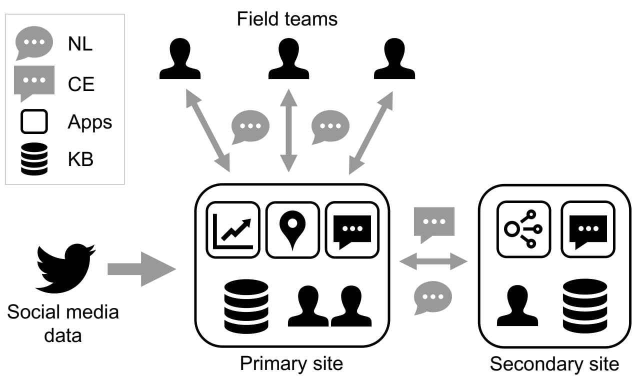

Figure 6: Typical field exercise set-up

Figure 6 shows a typical field exercise configuration. Two sites are used to assess the capability to support real-time analysis of a situation involving analysts working at different locations and across multiple agencies: one site is designated as primary and the other is designated secondary. Non-identical CE KBs are maintained at both sites, with each incorporating local agency-specific information as well as shared information. Different analytical apps are used at each site (e.g., spatial visualisations, trending tools, or social network analysis algorithms). Information and knowledge was communicated between both sites in NL and CE. Audio and video communication (e.g., via Skype) was also used. Social media data collection focussed on Twitter, due to its properties as an effective carrier of real-time and situational information, incorporating links to other media. In the set-up shown in Figure 6, social media data collection and processing is based at the primary site; information thus acquired is added to the KB and shared as appropriate with the secondary site.

A key feature of the exercise set-up is the ability to task field teams to visit specific locations and gather spot reports, e.g., to supplement or verify data collected from social media. Field team members are equipped with mobile devices and communicate with the primary site via NL messages and, where appropriate, attached imagery. The entire team uses a collaboration platform[8] supporting (i) real-time messaging organised into a collection of thematic 'channels', (ii) posting of documents and other media, including items foraged from social media, and (iii) a persistent timeline-based record of discussions after the event. Our conversational agents (including the Moira agent described below) are connected as virtual users to the collaboration platform (i.e., as chatbots).

[8]We have used both Slack (http://slack.com) and Mattermost (http://mattermost.org) for this.

Foraging for 'Fast Data'

Information foraging is supported by the ability to interpret NL statements into CE to support subsequent machine analysis and fusion. A key use case here is enabling direct submission of *in situ* reports via the conversational protocol *confirm* interactions. This facility is available to users through a chatbot called Moira (Mobile Intelligence Reporting Agent) (Preece et al. 2014). We have also experimented with allowing a user to post brief reports via social media, specifically Twitter, with the Moira bot configured to 'follow' specific accounts — public or private — and/or to process data collected via the Twitter API. This kind of data obtained from social media or *in situ* reporting has been termed 'fast data' (as opposed to 'big data') (Roberts et al. 2017).

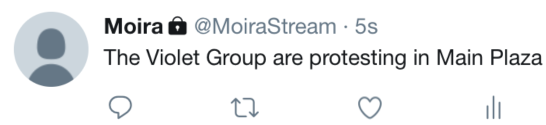

Figure 7: Tweeted spot report as NL input to the Moira chatbot

Figure 7 shows an example spot report posted on Twitter from a private account used in our exercises. The equivalent CE form of this report was shown in the previous section as the example `protest` instance. Key entities mentioned here are an instance of an organisation (Violet Group), a location (Main Plaza) and a kind of event (protest). As shown above, the CE model of this domain includes the entity classes (concepts) `organisation`, `location`, `event`, and `protest`, together with the information that a `protest` is a kind of `event`. Instances include `Violet Group` and `Main Plaza`. In principle these can be discovered through named entity recognition in natural language processing (NLP)[9] though in this case both entities would probably already be known *a priori*.

When the Moira chatbot is used, the *confirm* interaction provides users with an opportunity to manually confirm (or edit) the generated CE. Experiments indicate good usability of the CE-based bot by untrained users in a crowdsourcing context (Preece et al. 2017). However, for input via Twitter it doesn't make sense to do this as the user may be a member of the public who would find such an interaction very confusing; moreover, typical CE sentences tend to overrun the 140 characters of a tweet.

We have explored the use of rapid fact acquisition to build CE KBs of background information such as people, organisations, and places, useful as a means of connecting or contextualising other information in support of crowdsourcing and *in situ* reporting. This background knowledge is valuable in NLP, performing an important role in named entity recognition as illustrated by the examples 'Main Plaza' and 'Violet Group' in Figure 7. In support of our field exercises, focussed on monitoring community reactions to disruptive

[9]For example, using http://nlp.stanford.edu/software/CRF-NER.shtml

events in the UK, we constructed models and fact sets for notable places, organisations, and public individuals including politicians and journalists.[10] This allowed our Moira bot to perform question answering via the conversational interface on this body of rapidly-sourced background data, as illustrated in Figure 8.

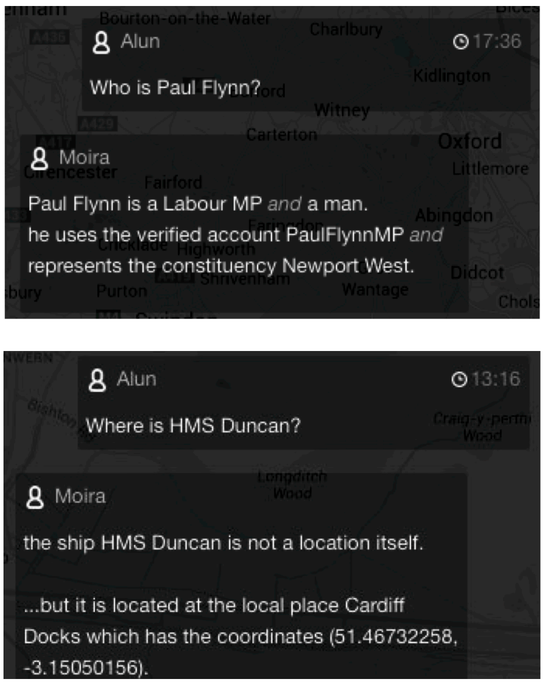

Figure 8: An example of question answering from foraged data using the Moira chatbot

All of these activities demonstrate how the conversational approach can support the foraging loop in sensemaking. We regard part of our CE KB as the *shoebox* shown in Figure 5. This contains both NL and CE input from external sources, including crowdsourced data, tweets, and background knowledge as discussed above. It is slightly different from the traditional sensemaking shoebox in that the use of CE for interpreted data and metadata gives the information a degree of structure from the moment it arrives in the system, whereas traditionally such structure is added during the process of ‘schematising’ the information. However, note that we are not attempting to ‘make sense’ of the data as it arrives into the shoebox: we are adding only low level context corresponding to the information layer in Figure 2; i.e., our ‘CE shoebox’ covers basic facts like *who*, *what*, *when* and *where*, but does not yet attempt to capture anything about *how* or *why*.

In the context of a sensemaking process, automatically collecting large volumes of tweets into a foraging shoebox would be unmanageable. CE models offer various ways to help filter relevant social media fragments, e.g., by using NLP to associate the input media with related model elements. In the example above, this can be done by detecting that the tweet refers to `Violet Group` (an `organisation`), `Main Plaza` (a `location`) and a `protest` (an `event`). This information, together with tweet metadata (e.g., the poster, together with any knowledge of their reliability, the time, GPS coordinates if the tweet is geotagged, and so on) provides a significant amount of context allowing the tweet to form part of a larger situational picture (for example, activities of the Violet Group, currently-known protests, or disruption in the Main Plaza area). Part of assembling the situational picture also involves seeking corroboration unless the tweeter is a trusted source, and this can be achieved in important cases by tasking the field teams.

The contents of the shoebox are not limited to NL data such as that acquired from humans or documents but can also include multimedia data as illustrated in Figure 3. The progression from shoebox to evidence file involves linking, summarisation and inference. An advantage of collecting and contextualising acquired information in a CE KB is that we automatically build a graph of linked data, around common entities such as people, organisations, and places, and the relationships defined in the CE models between these and other entities. Note that the CE models are not static: they evolve as new relationships and concepts are found. This is done semi-automatically using techniques like automatic term recognition to propose new concept names (Spasić et al. 2013); it can also be done manually by querying and filtering the data.

Sensemaking and ‘Storytelling’

The initial steps in the sensemaking loop (following on from the end of the foraging loop in Figure 5) involve schematising evidence by adding further higher-level structure and interpretation — corresponding to the information fusion and knowledge layers in our MASU architecture (Figure 2). As noted above, the CE-based approach already introduces a degree of schematising at the lower levels of the process, and at the end of the previous section we discussed how the approach allows progressive enrichment of the model, to better contextualise evidential information. Thus, we do not see a hard boundary between the two loops, but rather a gradual KB refinement and enhancement.

The foraging process is framed by the human decision makers’ hypotheses and intents; e.g., our field exercises were framed by social science motivations to understand the impacts of large-scale events involving major local disruption due to increased security, and threats of significant protest by a broad spectrum of groups. The requirement to monitor and make sense of such unfolding situations framed the foraging activity primarily around event detection — especially protests and crowd mobilisation — and directed us to gather background knowledge on key actors (including politicians and journalists) and significant public locations.

The backward chain discussed in the introduction pushes this contextual framing down into the foraging loop, resulting in the attachment of metadata to collected data that is

[10]For this we used public sources such as Tweetminister, http://tweetminster.co.uk, Muckrack, http://muckrack.com, and Wikipedia/Dbpedia.

useful in the later stages of interpretation. In the case of social media data acquired from Twitter, we are interested in the context as well as the content: who is saying it, when and where. Whether a tweet originates from a politician's account, an activist group, a journalist, or a member of the public is important in making sense of the signals available from open source media. Our approach represents all of this information — data and metadata — in a uniform CE KB (shareable across partner agencies) and information architecture, processable by a diverse set of decision-support apps.

The end-goal in Figure 5 is that of presenting a case: using the connected data, evidence and hypotheses chain to tell a story to inform making a decision. The conversational CE-based approach emphasises HCC and human-consumable KRR, including the key ability for a user to ask *why* to uncover rationale for any inference, statement or connection in the KB. Building on these features, we can apply narrative framings to the assembled KB, and have trialled the use of techniques including comic strips and multiple-act story structures modelled via CE metadata (Braines et al. 2015). The CE `story` concept exists at a more abstract level to the domain of interest (people, places, events, etc) and permits domain-level information to be organised into a sequence of episodes (`preface`, `act one`, `act two`, etc) with associated key events and actors drawn from the domain-level model and instances. The conversational ability to ask questions, including *why?*, permits the consumer of the narrative to reveal detail that the higher-level story intentionally omits. This may reveal unanswerable questions or prompt the user to tell the system something that has not yet been taken into account, causing feedback to flow back down the sensemaking process, triggering further machine-machine and human-machine *ask/tell* interactions.

In terms of higher-level situational understanding, including modelling of intents and threats, we have ongoing work looking at polarisation and conflict in relation to major crime events (Innes et al. 2016; Roberts et al. 2017). Our information modelling focus w.r.t. this social science research is to gain a better understanding of the analytic processes with a view to encoding these into our information architecture with a view to assisting the analysts with their higher-level sensemaking tasks. Concurrent research in applying CE-based KRR techniques to intelligence analysis has highlighted promising use cases where HCC can assist human cognitive processes (Mott et al. 2015).

Discussion and Conclusion

Open source information provides MASU processes in governmental and public sector domains with an important type of external data source. However, the vast amount of available data, especially in terms of social media, presents enormous challenges in foraging and sensemaking. While there have been significant advances made in techniques for event detection using social media streams (Aiello et al. 2013; Roberts et al. 2017; Wang et al. 2014), the collection and processing of open source data to derive higher-level information products beyond the shoebox is fundamentally one of HCC, relying as it does on the relative strengths of human and machine agents (Crouser and Chang 2012). Analysts are increasingly well versed in modern team collaboration environments and in the exploitation of social media, and systems are emerging that seek to combine the benefits of these approaches with existing software tools and processes for structuring and supporting intelligence analysis (Wollocko, Farry, and Stark 2013). Multi-agency and collective intelligence approaches are seen as particularly important in this context (Hall and Jordan 2010) since, not only is collaboration essential within the same analyst team, but the outcome of analysis process can be greatly improved when collaboration is extended to the crowd and mediated by an intelligent software agent (Brantingham and Hossain 2013).

In the light of current concerns regarding transparency in big data approaches (Lazer, King, and Vespignani 2014), a significant issue in the use of all open source intelligence concerns the potential for bias and misinformation (Jin et al. 2014), and mitigating these risks is a very active are of current research (Wang, Abdelzaher, and Kaplan 2015). However, patterns of (mis)information flow are often extremely valuable in terms of situational understanding, e.g., rumouring is often a form of coordinated activity which needs to be countered (Roberts et al. 2017). A key feature aimed at promoting transparency in our approach is the *why* interaction, allowing a user to seek an explanation for any inference made by the system.

In addition to text-based open source media, there is considerable value in social media imagery data, e.g., attached to tweets on Twitter and on imagery-centric platforms such as Instagram and YouTube. In this sense, social media is a source of both hard and soft data, leading to significant challenges in information fusion (Hall and Jordan 2010). The application of ML and VSP to extract key features from such sources — particularly common symbols and objects, as well as face recognition — offers considerable potential but remains a hard problem. In the near term, processing such data effectively is another aspect of MASU that requires HCC. An issue that arises here is the generation of *why* explanations for inferences drawn by VSP and ML: these approaches are often regarded as black boxes, but there is considerable ongoing work to make them more interpretable (Lipton 2017).

In summary, we have presented an approach to MASU founded on natural language interaction and knowledge management that supports three key use cases:

- integration of cognitive approaches to exploiting 'big data' with support for humans to input 'fast data';
- usability by people with relatively little technical training and where, especially in field settings, personnel will be primarily focused on tasks other than operating software;
- inbuilt support for users to seek explanations from cognitive systems via *why* interactions.

Going forward, our immediate agenda is to focus on (i) improved explanation generation and interaction, particularly with ML-based cognitive services, (ii) improved handling of misinformation including rumours and propaganda, and (iii) integration of KRR and ML cognitive services for hypothesis exploration.

Acknowledgement

This research was sponsored by the U.S. Army Research Laboratory and the UK Ministry of Defence under Agreement Number W911NF-16-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the UK Ministry of Defence or the UK Government. The U.S. and UK Governments are authorised to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

Aiello, L. M.; Petkos, G.; Martin, C.; Corney, D.; Papadopoulos, S.; Skraba, R.; Gker, A.; Kompatsiaris, I.; and Jaimes, A. 2013. Sensing trending topics in Twitter. *IEEE Transactions on Multimedia* 15(6):1268–1282.

Alberts, D. S., and Hayes, R. E. 2003. *Power to the Edge: Command and Control in the Information Age*. CCRP.

Austin, J., and Urmson, J. 1975. *How to Do Things With Words*. Harvard University Press.

Braines, D.; Ibbotson, J.; Shaw, D.; and Preece, A. 2015. Building a 'living database' for human-machine intelligence analysis. In *Proc 18th International Conference on Information Fusion*.

Brantingham, R., and Hossain, A. 2013. Crowded: a crowd-sourced perspective of events as they happen. In *Proc Next-Generation Analyst (SPIE Vol 8758)*. SPIE.

Broome, B. 2012. Data-to-decisions: a transdisciplinary approach to decision support efforts at ARL. In *Proc Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR III (SPIE Vol 8389)*. SPIE.

Crouser, R., and Chang, R. 2012. An affordance-based framework for human computation and human-computer collaboration. *IEEE Transactions on Visualization and Computer Graphics* 18(12):2859–2868.

Dostal, B. C. 2007. Enhancing situational understanding through the employment of unmanned aerial vehicles. *Army Transformation Taking Shape…Interim Brigade Combat Team Newsletter* (01–18).

Fraga-Lamas, P.; Fernández-Caramés, T. M.; Suárez-Albela, M.; Castedo, L.; and González-López, M. 2016. A review on internet of things for defense and public safety. *Sensors* 16(1644).

Hall, D. L., and Jordan, J. M. 2010. *Human-Centered Information Fusion*. Artech House.

Innes, M.; Roberts, C.; Preece, A.; and Rogers, D. 2016. Ten 'Rs' of social reaction: using social media to analyse the 'post-event' impacts of the murder of Lee Rigby. *Terrorism and Political Violence* in press.

Jin, F.; Wang, W.; Zhao, L.; Dougherty, E.; Cao, Y.; Lu, C.-T.; and Ramakrishnan, N. 2014. Misinformation propagation in the age of Twitter. *IEEE Computer* 47(12):90–94.

Kelly, J., and Hamm, S. 2013. *Smart Machines: IBM's Watson and the Era of Cognitive Computing*. Columbia Business School Publishing.

Kuhn, T. 2014. A survey and classification of controlled natural languages. *Computational Linguistics* 40:121–170.

Labrou, Y., and Finin, T. 1998. Semantics and conversations for an agent communication language. In Huhns, M. N., and Singh, M. P., eds., *Readings in agents*. Morgan Kaufman. 235–242.

Lazer, D.; King, R. K. G.; and Vespignani, A. 2014. The parable of Google Flu: Traps in big data analysis. *Science* 343:1203–1205.

Lipton, Z. C. 2017. The mythos of model interpretability. In *2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016)*.

Mott, D.; Shemanski, D. R.; Giammanco, C.; and Braines, D. 2015. Collaborative human-machine analysis using a controlled natural language. In *Proc Next-Generation Analyst III (SPIE Vol 9499)*. SPIE.

Pask, G. 1972. A fresh look at cognition and the individual. *International Journal of Man-Machine Studies* 4:211–216.

Pirolli, P., and Card, S. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of International Conference on Intelligence Analysis*.

Preece, A.; Braines, D.; Pizzocaro, D.; and Parizas, C. 2014. Human-machine conversations to support multi-agency missions. *ACM SIGMOBILE Mobile Computing and Communications Review* 18(1):75–84.

Preece, A.; Webberley, W.; Braines, D.; Zaroukian, E. G.; and Bakdash, J. Z. 2017. SHERLOCK: Experimental evaluation of a conversational agent for mobile information tasks. *IEEE Transactions on Human-Machine Systems* in press.

Roberts, C.; Innes, M.; Preece, A.; and Rogers, D. 2017. After Woolwich: analyzing open source communications to understand the interactive and multi-polar dynamics of the arc of conflict. *British Journal of Criminology* in press.

Spasić, I.; Greenwood, M.; Preece, A.; Francis, N.; and Elwyn, G. 2013. FlexiTerm: a flexible term recognition method. *Biomedical Semantics* 4(27).

Srivastava, M.; Abdelzaher, T.; and Szymanski, B. 2012. Human-centric sensing. *Phil. Trans. R. Soc. A* 370(1958):176–197.

Wang, D.; Abdelzaher, T.; and Kaplan, L. 2015. *Social Sensing: Building Reliable Systems on Unreliable Data*. Morgan Kaufmann.

Wang, D.; Amin, M. T.; Li, S.; Abdelzaher, T.; Kaplan, L.; Gu, S.; Pan, C.; Liu, H.; Aggarwal, C. C.; Ganti, R.; Wang, X.; Mohapatra, P.; Szymanski, B.; and Le, H. 2014. Using humans as sensors: An estimation-theoretic perspective. In *Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*, IPSN '14, 35–46.

Wollocko, A.; Farry, M.; and Stark, R. 2013. Supporting tactical intelligence using collaborative environments and social networking. In *Proc Next-Generation Analyst (SPIE Vol 8758)*. SPIE.