Narrative Visualisation of Simulations: Finding the Stories Within the Data

James Walsh[a], **Andrew Cunningham**[a]**, Steven Wark**[b]**, Marcin Nowina-Krowicki**[b]** and Bruce H. Thomas**[a]

[a] *Australian Research Centre for Interactive and Virtual Environments, University of South Australia,* [b] *Joint and Operations Analysis Division, Defence Science and Technology Group*
Email: James.Walsh@unisa.edu.au

Abstract: Decision makers need to understand and draw key insights from large and complex data sets generated by simulations. This may be to answer a specified analytical problem addressed by a simulation or to support learning or mission objectives through after-action review of a virtual simulation. Geospatial displays are traditionally used to help a decision maker or analyst gain an understanding of the results of a force-on-force constructive simulation, or for after-action review of virtual simulations. They provide an evolving picture of the geospatial positions and spatial relationships between entities and key features in the environment, and can potentially incorporate other attributes such as entity state or key metrics to support the analyst. However, while this geospatial context is often important, sense-making also requires an understanding of the significant actors, their relationships, their interactions, and the consequences of these relationships and consequences. Narrative visualisation is an approach that focusses on the storylines of key actors, the events they are involved in, and the causal relationships between entities – i.e. the stories within the data. This paper describes initial work demonstrating a novel narrative visualisation approach for data from a large-scale force-on-force simulation, and the key features of this visualisation that supports a decision maker or analyst's understanding and insights on what occurred within the simulation, and why. To the authors' knowledge, this represents the first time simulation data has been applied to narrative visualisation. While still in its early stages, this approach shows promise and could potentially supplement, or indeed replace, geospatial-centric visualisations as the primary sense-making pathway for analysts in a wide range of operating domains, including real-time operating pictures. This is particularly apropos in domains where the geospatial context is less relevant.

Keywords: *Narrative visualisation, storytelling, analysis, after-action-review, sense-making*

1.   INTRODUCTION

Decision makers need to understand and draw key insights from large and complex data sets generated by simulations. This may be to answer a specified analytical problem addressed by a simulation or to support learning or mission objectives through after-action review of a virtual simulation. This understanding has traditionally been supported using geospatial-centric displays which convey the location and spatial relationships of entities and environmental features, and are often annotated with additional information such as entity states. These geospatial-centric displays inherently have difficulties representing the temporal aspects of the represented data (Walsh et al. 2018). The decision maker or analyst will often conceptualise this as a narrative incorporating the key actors, relationships, events within the simulation that contributed to the observed outcomes: i.e. as the stories within the data. Narrative visualisation approaches, as described by Gershom & Ward (2001), Wark & Lambert (2007), Segel & Heer (2011), and Figueiras (2014), can aid this process by extracting, compiling, and providing a visualisation that focuses on the key actors, events, and causal relationships within an evolving situation. This allows the decision maker or analyst to easily understand how the situation has evolved and what the decisive factors were. Narrative visualisation could potentially supplement, or indeed replace, geospatial-centric visualisations as the primary sense-making pathway for decision makers and analysts in a wide range of scenarios involving sequences of events. This is particularly apropos in domains where the geospatial context is less relevant. It is by exploring the events over time, rather than individually that we can gain an understanding of the context each unit is operating in.

In this paper, we describe a novel narrative visualisation approach, originally developed by Cunningham et al. (2018) to aid law enforcement investigators to understand the progress and outcomes of criminal investigations. This approach has been applied to visualisation of data generated by JFOrCE, a large-scale force-on-force simulation described by Au et al. (2018). We describe how an analyst can follow and understand the significant storylines within the simulation. While the initial prototype was restricted to the visualisation of a limited class of events and descriptions, this new direction promises to greatly assist and decision maker or analyst when coupled with AI techniques that identify and describe decisive events and causal relationships between these events.

2.   BACKGROUND

Narrative visualisation techniques can be applied to any data that contains sequences of events involving one or more actors. As the number of actors, events, and interactions increases, and complex relationships and goals become key, understanding of the stories within the data becomes increasingly important. This is particularly true for situations describing complex human-human interactions.

2.1.   Narrative Engine

In the system developed by Cunningham et al. (2018), the primary concern for law enforcement investigators is not the underlying data, but the higher level abstraction of who was involved, what happened, why and when. The narrative engine can identify the significant actors and events, summarise, filter, and query the narrative to allow the investigator to visualise different stories related to different actors, objects (e.g. evidence), locations, and events. The narrative data model used for this system is based on the situation model of Zwaan & Radvansky (1998), that allows it to distinguish between different narrative relationships and the active or passive roles of the actors involved in the events. Formally, a narrative N is described as a collection of events and entities, defined by:

$$N = (E, \varrho) \tag{1}$$

where E is the set of events and ϱ is the set of narrative entities. An event, e, is represented as a tuple:

$$e = (t_s, t_e, A), e \in E \tag{2}$$

where t_s and t_e are a timestamp for the start and end of the event respectively, and A is the set of subject-verb-object predicates describing that event. Given the set of all entities in the narrative, a predicate, p, is the tuple:

$$p = (S, O, v) \tag{3}$$

where S is the set of subject entities of the predicate (the entities taking action), O is the set of entities being acted upon, and v is a verb describing the action being performed. For example, the event 'Alice shot Bob' is represented by the single predicate *({Alice}, {Bob}, shot).*
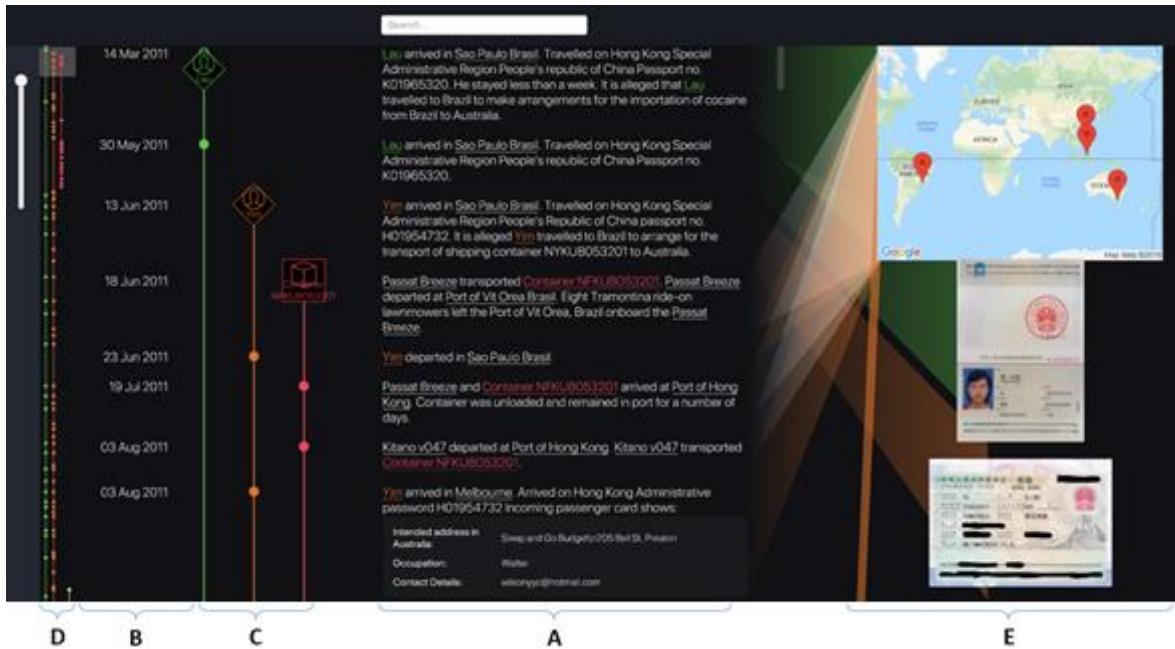
Figure 1. Sample screenshot of the Narrative Visualisation system showing its five main components.

2.2. Narrative Visualisation

The visualisation of the criminal investigation, see Figure 1, consists of five primary components:

A. The description of the events, which provides a textual account of the events.
B. The timestamps for the events.
C. The timelines associated with the entities of interest, and their involvement in the events of the investigation (indicated by pips). This provides a simple visualisation of the engagement of the entities in the events and significant interactions between many entities.
D. A compressed view of the timelines of interest, showing the timelines for selected entities across the whole duration of the investigation (starting at the top of the screen by default). This provides a simple visualisation of how significant an entity is across the duration of the investigation.
E. Supporting evidence and materials (maps, photos, videos, documents, etc.) associated with events.

Using the system, the investigator can explore the core narrative of the investigation involving the most significant entities, and the involvement of any particular entities with events of interest, along with the supporting evidence. The entity timelines are dynamic, allowing the user to display (and subsequently collapse) the timeline of any entity named in the narrative.

The investigator can filter the events displayed to only show events whose descriptions contain specified search terms. In addition, the events described in the investigation are weighted based on how important they are towards substantiating the possible charges that may be laid against persons of interest. The displayed narrative can be filtered using the slider on the left to only show those events above a threshold weighting, allowing it to be collapsed down to the top events defining the investigation (e.g. the top 5-10 key events). The selection of these events can involve not just those charged, and the events themselves, but potentially user-preferences for the types of events the user cares about most. When events are filtered, non-selected events are collapsed into a single icon showing the number of events hidden to cue the user that not all events are being shown.

These features allow the investigator to visualise the timelines of the events an entity is involved in, its interactions with other entities, and identify the key events. Through interaction and filtering, the investigator can explore the associations between actors and the significance of notionally peripheral events.

3. JFORCE

Large-scale force-on-force simulations allow an analyst to evaluate how effective a particular combination of forces may be at achieving desired outcomes from military scenarios. They can be used to evaluate courses of action during military planning or to inform defence procurement decisions that seek to identify the right

force mix needed. The Joint Future Operating Concept Explorer (JFOrCE) is an agent-based simulation that has been developed to study future force mixes and operating concepts as described by Au et al. (2018). JFOrCE scenarios typically contain hundreds of entities, and play out over tens-of-thousands of time steps, generating large data sets to understand. Furthermore, each scenario needs to be repeated hundreds of times to identify emerging trends and critical outliers, and hundreds of different scenarios may be required to explore different force mix options. The NATO data farming framework, as described in Horne & Seichter (2014), is being explored as a way of managing this problem. However, an analyst still needs to be able to assess: if mission objectives were achieved; what decisive points within a simulation that contributed to the final outcome, and; the causal factors that led to this outcome. This requires new ways of visualising the simulation data to allow an analyst to rapidly identify the decisive points and causal factors. We describe how narrative visualisation could be applied to simulations, limited to a single simulation run only.

3.1. Scenario

For this initial work, a simple fictitious scenario was used that captures the essential elements of the force-on-force simulations of interest. The following generic military entities were included in the scenario: tanks, submarines (sub), ground-based air defence (GBAD), air warfare destroyers (DDG), fighter jets (jet), and airborne early warning and control (AEW&C) aircraft. The same entity types were provided to two opposing sides, 'red' and 'blue', using uneven numbers of entities per side.

For this work, 2D entity locations, radar detection and jamming of entities, munitions firing, and kills data was collected from the simulation at each time step. In addition, metrics describing the total number of jets, tanks, submarines, GBAD, DDG, and AEW&C for each side, and the missiles types fired were also recorded for each tick. Time steps of 20 seconds were used, with the scenario completing after 20,000 time steps. We simplified the event data generated in our scenario by noting that it is primarily the onset and termination of detection and jamming interactions between entities that are of interest and that successful attacks will implicitly involve the termination of any detection and jamming interactions with them.

4. GENERATING STORYLINES

The narrative visualisation approach is likely to only be of benefit to an analyst if extended storylines involving interactions of one or more entities exist within the data. Narrative visualisation of largely isolated entities and sporadic events occurring between them are unlikely to convey additional insights to an analyst. Furthermore, when storylines do exist within the data, simply describing events in chronological order may not be the most effective way to convey those storylines, particularly when dealing with large numbers of entities across large time spans. To identify any storylines, and to explore different ways of conveying them, we have extended the Multiple Intersecting Chronologies (MIC) approach of Dall & Donnelly (2017), which seeks to trade off measures of *continuity*, representing the degree to which a narrative continues to follow an entity or group of entities, against *temporal tension*, representing the degree to which a narrative describes events out of chronological order.

4.1. Event Sequencing

The MIC algorithm generates a narrative sequence, G, by incrementally selecting an event, E_k, that minimises a cost function dependent on the last event added, E_i. In our revised version, the cost function, F_{ik} is:

$$F_{ik} = \lambda T_{ik} - (1-\lambda)N_{ik}, \quad \lambda \in [0,1] \tag{4}$$

where T_{ik} and N_{ik} are the temporal tension and continuity between E_i and E_k, respectively, and λ is a mixing parameter chosen to control the MIC sequencing. Both T_{ik} and N_{ik} are normalised by the maximum values calculated from the remaining available events E_k.

We use an implementation of T_{ik} and N_{ik} that includes the *centrality*, C_j, of an entity e_j, which is defined as the total number of entities it shares with any event, normalised by the total number of entities, M, in the scenario. This provides a measure of the level of influence of the entity within the scenario. For example, an entity that interacts with every entity in the scenario will have a centrality of *1*, while an entity that interacts with no other entities (except for itself) will have a centrality of *1/M*. Temporal tension is calculated using:

$$T_{ik} = \begin{cases} \sigma_k(t_i - t_k) \max_{e_j \in E_k} C_j, & when\ t_i > t_k, \\ 0, & when\ t_i \leq t_k \end{cases} \tag{5}$$

where σ_k is the *significance* assigned to E_k, and t_i and t_k are the times at which E_i and E_k occur respectively.

This revised definition of temporal tension places greater emphasis on the temporal tension associated with the entities with the greatest influence in the scenario than those with the least. The *relatedness*, R_{ij}, between two entities e_i and e_j is the total number of events e_i shares with e_j, normalised by the total number of events e_i participates in. This provides a measure of how closely coupled these two entities are in the scenario. Note that the relatedness is asymmetric: if e_i only interacts with e_j, $R_{ij} = 1$, but if e_j also interacts with other entities, $R_{ji} < 1$. This allows us to define the continuity between events E_i and E_k as:

$$N_{ik} = \sum_{e_j \, \epsilon \, E_k} R_{ij} C_j \tag{6}$$

This revised definition of continuity places greater emphasis on the entities with the highest centrality. The revised MIC algorithm will generate sequences of events that represent storylines between entities in a scenario. The length of the storylines generated depends on the mixing parameter, λ, chosen. When $\lambda = 1$, the MIC algorithm will only minimise the temporal tension between successive events, consequently producing a strict chronological order of events. When $\lambda = 0$, it will only maximise the continuity between successive events, generating storylines that follow all of the interactions (however unimportant) between entities before jumping to another storyline. For $0 < \lambda < 1$ the temporal tension will cause the sequencing to jump to new storylines when the continuity between events drops below a threshold determined by λ.

4.2. Storylines in JFOrCE

We applied our revised MIC algorithm to extract storylines from the simulated data generated by JFOrCE using our scenario described above, to determine if there were indeed any significant storylines to be found in the data that would warrant a narrative visualisation approach.
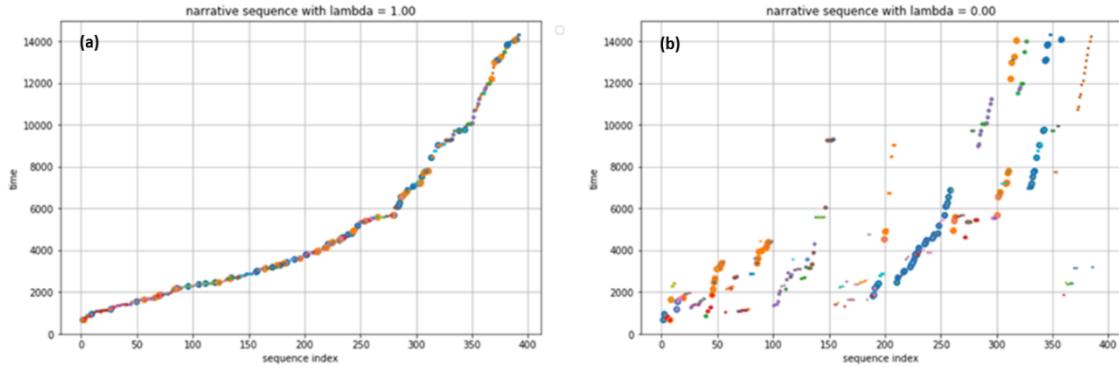

Figure 2. Narrative sequence generated by MIC algorithm with JFOrCE data with different mixing parameters. The different colours indicate events involving different entities, while the size of the data points is indicative of centrality. (a) Sequence determined only by temporal tension ($\lambda = 1$) giving a strict chronological order. (b) Sequence determined only by continuity ($\lambda = 0$), showing extended storylines within the data.

As shown in Figure 2(b), extended storylines exist within the data, suggesting that the use of narrative visualisation to explore these storylines may be applicable. An interesting feature to note is that in this case the MIC algorithm is, as expected, presenting the storylines of the entities with the highest centrality (larger data points) before the less central entities. These groupings form what we refer to as a *chapter*.

5. NARRATIVE VISUALISATION OF SIMULATIONS

Our initial approach to narrative visualisation of the JFOrCE simulation data re-used the core narrative engine and visualisation developed for visualisation of criminal investigations described earlier. In this simplified scenario, events only involve pairs of entities and are described as single subject-object-verb predicates. The set of predicates were *begin/end tracking, begin/end jamming, killed,* and *firing* of weapons.

The initial pilot visualisation developed incorporates many of the features of the original criminal investigation visualisation, as shown in Figure 3. The main features are:

A. The description of the events. Events related to the same set of entities within the same time step are aggregated. While currently brief, these descriptions could be easily expanded using the information provided by the simulation, or by additional analysis of the simulation data to help the analyst identify decisive points and causal relationships;

B. The timestamps for the events;

C. Storylines for entities of interest and their involvement in the events of the simulation. In addition to showing the most significant entities, they can also be selected at run time;

D. A compressed view of the storylines for the selected entities across the whole simulation;

E. A visualisation of the metrics of interest, showing how they change over the course of the simulation and how they are associated with the events in the simulation.
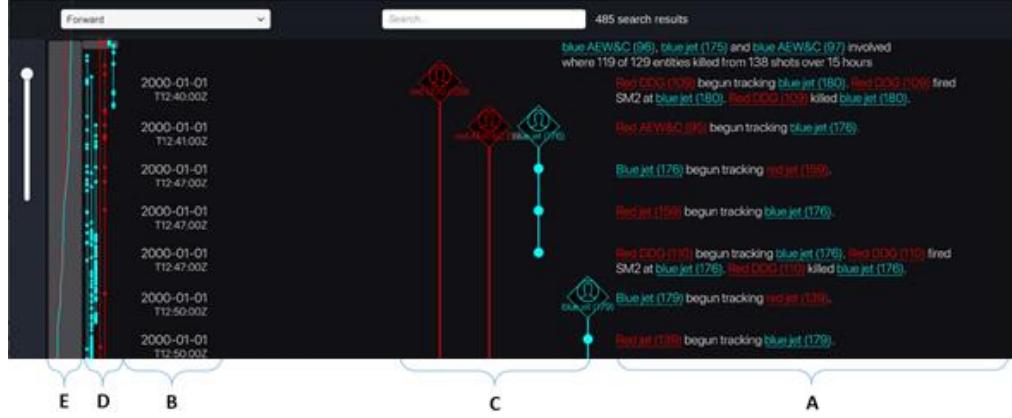
Figure 3. Cropped screenshot of pilot narrative visualisation system developed for simulation data

The visualisation by default shows the timelines of the most significant entities within the simulation, as determined by the entities with the highest centrality. The system selects entities with a centrality more than two standard deviations above the mean, otherwise just selecting the entities with the highest centrality.

Events are grouped into chapters, as described prior, with a generated summary of the entities and key events in those chapters appearing at the beginning of the chapter. By default it presents an initial summary of the most significant entities across the whole simulation. The MIC algorithm can be applied to generate a chapter for each of the storylines identified. The analyst can select the mixing parameter used to generate the chapters that best suits their comprehension. This allows the chapter grouping to be adjusted to place a higher precedence on maintaining the focus on the entities involved, versus the temporal tension between events.

In exploring individual events, the significance slider on the left-hand side of the visualisation (see Figure 3) can be used to filter events based on a level of significance. This allows the scenario to be presented as just the most significant events (as defined by the event type, but could incorporate weighting from the entities involved), through to showing more and more events of lesser detail, until all events are shown. This allows the analyst to begin by focusing on only the most significant events, before exploring the details.

6. DISCUSSION

The work described in this paper explored if narrative visualisation added value to an analyst, and used a very simple representation of events within JFOrCE. Much richer event information could be collected from the simulation, or produced by subsequent analysis, and used to enhance event descriptions and presentation of the storylines within the data. In its current form, the narrative visualisation has shown value by allowing the analyst to identify and visualise the storylines and interactions of the most significant entities within the simulation, and to expand and display the storylines of associated entities to explore their significance and consequent interactions. The associated visualisation of the metrics of interest mapped to the events within the storylines allows the analyst to identify those events that have a significant impact on the simulation outcomes. Together this allows the analyst to identify and explore causal relationships within the data.

The ability for analysts to search for entities or terms (as potentially annotated by an automated analyst) allows them to identify those events with particular features or interactions. The significance slider allows them to initially focus on the key events within the simulation. The automatic summarisation of the presented chapters could also allow an analyst to quickly identify storylines of interest that can be explored in more detail. By exploring different chapter groupings, the agent can try and resolve uncertainty within the dataset. The use of an automated analyst would address a core limitation for lack of descriptive content, as would support for an event to be present in multiple chapters.

A number of avenues remain for future work. Refinements to the storyline visualisation, including representative icons that map to the different entity types, and improved user control over the selection and presentation of the metrics of interest would assist the analyst in exploring the visualisation. The

incorporation of supplementary analysis products and geospatial visualisations associated with the events within the simulation, along with improved chapter visualisation and summarisation, and improved user control over the sequencing algorithm would allow the agent to explore which events are more likely to occur, whilst adjusting the chapter generation to focus on what they care about most.

7. CONCLUSION

In conclusion, we presented a narrative visualisation method to gain a deeper understanding of force-on-force simulations. In particular, we evaluated the use of JFOrCE as an agent-based simulation to supply data for our narrative visualisation techniques. The paper presents a new algorithm and associated implementation for generating storylines in a narrative visualisation system, to provide a graphical user interface to allow interactive manipulation and visualisation of the data.

Geospatial displays have limited abilities to convey the temporal aspect. They rely predominately on animating the entities during the replay of the simulations. The interactions between entities are often hidden when the data is presented in strict temporal ordering. We present a modified version of Dall and Donnelly's Multiple Intersecting Chronologies to group these interactions into interesting chapters to elicit the sub stories within the simulation. Our initial investigation found this new algorithm successfully clustered events into chapters. Future work will examine what information agents can extract versus existing techniques.

The narrative visualisation system provides a useful means of presenting an alternative to geospatial representations of the simulation data. The visualisation depicts the timelines of the most significant entities within the simulation. The events are grouped into chapters, with a generated summary of the entities and key events. The system provides an interactive mechanism for the analyst to filter out less important events.

REFERENCES

Au, T.A., Hoek, P.J. & Lo, E.H.S (2018), Combat Analysis of Joint Force Options using Agent-Based Simulation, *2018 Military Communications and Information Systems Conference (MilCIS)*, 13-15 November, Canberra, Australia.

Cunningham, A., Walsh, J. & Thomas, B.H. (2018), Immersive Visualisation of Geo-temporal Narratives In Law Enforcement, *2018 4th International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, 17-19 October, Konstanz, Germany.

Dall, I. & Donnelly, B. (2017), Event Sequencing for Situation Narratives, *DST Technical Report*, DST-Group-TR-3351.

Figueiras, A. (2014), Narrative Visualisation: A Case Study of How to Incorporate Narrative Elements in Existing Visualizations, *2014 18th International Conference on Information Visualisation (IV2014)*, 16-18 July, Paris, France.

Gershom, N. & Page, W. (2001), What Storytelling Can Do for Information Visualization, *Communications of the ACM*, 44 (8), pp 31-37.

Horne, G. & Seichter, S. (2014), Data Farming in Support of NATO Operations – Methodology and Proof-of-Concept, *Proceedings of the Winter Simulation Conference,* Savannah GA., USA.

Segel, E. & Heer, J. (2011), Narrative Visualization: Telling Stories with Data, *IEEE Transactions on Visualization and Computer Graphics*, 16(6), pp 1139-48.

Walsh, J., Cunningham, A., Smith, R. and Thomas, B. (2018), Tangible Braille Plot: Tangibly Exploring Geo-Temporal Data in Virtual Reality, *International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, 17-19 October, Konstanz, Germany.

Wark, S. & Lambert, D. (2007), Presenting the Story Behind the Data: Enhancing Situational Awareness using Multimedia Narrative, *2007 Military Communications Conference (MILCOM 2007)*, 29-31 October, Orlando FL, USA.

Zwaan, R. & Radvansky, G.A. (1998), Situation Models in Language Comprehension and Memory, *Psychological Bulletin*, 123(2), pp 162-185.