Interpreting Text and Image Relations in Violent Extremist Discourse: A Mixed Methods Approach for Big Data Analytics

Kay L. O'Halloran
Curtin University, Australia

Sabine Tan
Curtin University, Australia

Peter Wignell
Curtin University, Australia

John Bateman
University of Bremen, Germany

Duc-Son Pham
Curtin University, Australia

Michele Grossman
Victoria University, Australia

Andrew Vande Moere
KU Leuven, Belgium

Abstract

The paper presents a mixed methods approach for analysing text and image relations in violent extremist discourse. The approach involves integrating multimodal discourse analysis with data mining and information visualisation, resulting in theoretically informed empirical techniques for automated analysis of text and image relations in large datasets. The approach is illustrated by a study which aims to analyse how violent extremist groups use language and images to legitimise their views, incite violence and influence recruits in online propaganda materials, and how the images from these materials are re-used in different media platforms in ways that support and resist violent extremism. The approach developed in this paper contributes to what promises to be one of the key areas of research in the coming decades: namely the interdisciplinary study of big (digital) data sets of human discourse, and the implications of this for terrorism analysis and research.

Keywords: mixed methods, multimodal discourse analysis, data mining, information visualisation, big data analytics, violent extremist discourse, terrorism analysis, Islamic State (ISIS and ISIL), *Dabiq*

Introduction

Although multimodal discourse analysis, involving the study of the interaction of language, images and other resources in multimodal texts, interactions and events,[1] provides a rich array of theoretical tools for investigating text and image relations[2] in violent extremist discourse, the latest challenge is the move from manual analysis and discursive interpretation of a limited number of multimodal texts by mostly humanities-based experts towards automated recognition of multimodal meanings in large data sets to cope with the sheer volume of data being generated today.[3] This paper addresses this challenge by presenting a mixed methods approach that integrates qualitative methods of multimodal discourse analysis with quantitative methods of data mining and information visualisation for the analysis of text and image relations in large datasets. The mixed methods approach resolves the gap between highly-detailed, contextualised analyses of small samples of multimodal texts on the one hand, with highly-aggregated, decontextualised big data approaches (e.g. reductive content analysis) on the other.

The mixed methods approach is illustrated by means of a study which aims to analyse how violent extremist groups use language and images to legitimise their views, incite violence and influence recruits in online propaganda materials, and how the images from these materials are re-used in different media platforms in ways that support and resist violent extremism. Significantly, the approach developed in this paper addresses the key issue of the meanings arising from the integration of language and images in these texts. As Kovács explains, the real issue is not the use of text or image in terrorist discourse, but "the interconnection of texts and images" where "the text is often an integral part of the image and the same applies vice versa".[4]

Why is it important to understand the relationship between texts and images in the analysis of terrorist discourse, and how can multimodal analytical techniques help us better understand the

contemporary intersection of social media-driven terrorist communication and social influence? It is by now a truism to note that on-line narrative and messaging by violent extremist individuals and groups has grown exponentially over the last decade, as violent extremists and their supporters capitalise on the strategic and organisational 'gift to terrorists'[5] provided by new networked opportunities for reaching and influencing transnational audiences[6]. In this sense, violent extremists, like many others, are participating in wider shifts in social movements initiated by the rise and spread of Web 2.0[7], a trend in which web-based and wireless spaces have become highly interactive, dynamic and volatile. Yet much analysis of the intersection between terrorist influence and the virtual world of social influence continues to be grounded in the persistence of content- and/or platform-focused analysis of on-line violent extremist messaging and interactions (see for example[8]). These approaches employ 'big data'-based methodologies such as social network analysis, data mining and other tools for analysis of very large-scale datasets, and they offer important knowledge and insights about the rise and spread of violent extremist interactions in virtual environments and about shifts in the morphology of social influence. However, they remain insufficient when attempting to explain two key factors: 1) the representational and communication strategies by which terrorist social influence is constructed and perpetuated in particular contexts, and its rapid adaptation to new environments and imperatives, and 2) the experiential as opposed to the agentic features of particular messaging strategies as a vital dimension of social transformations toward violent action.

Content- and platform-based analyses may be able to tell us *who* is influential, and *what* they are saying, but not *how* their representational messaging strategies – strategies that rely increasingly on sophisticated deployment of multimodal communication elements and a broad range of cultural semiotic resources – achieve their intended effects and influence at the crossroads of cognition and emotion. Social media communication practices that deploy multimodal combinations of the visual, the aural and the textual are reconfiguring the relationship between distance and nearness, public and private[9]. Such reconfigurations play key roles in the ways in which social influence is now performed, distributed and experienced, and the 'media effects' paradigm of cause and effect that dominated social science and cultural research in the age of mass media is no longer sufficient to understand these trends and developments[10]. Multimodal analysis, by contrast, 'approaches representation, communication and interaction as something

more than language'[11], extending the study of the social interpretation of language to the whole domain of meanings which are made through visual, sonic and other semiotic resources (e.g. image and symbolism, gesture, gaze, proxemics, sounds) and their interactions in multimodal texts. Together, multimodal choices form recognisable clusters or configurations of meanings, constituting domains of cultural activity – including social influence that encourages or reinforces violent extremist ideology and action. While important insights in understanding how social influence and digital technologies converge in the 21st century are now gathering pace, especially in social psychology (see for example[12]), "virtual worlds don't just tell us something about the general nature of human social behavior, they also create new social phenomena"[13]. These emerging social phenomena are intimately bound up with dimensions of multimodality: new and complex ways of re-presenting, processing and experiencing the convergence of information and influence that are remapping how meanings can be made.

These transformations in communicative and representational capacity are sometimes termed *affordances* by social semioticians (see for example[14]: that is, the 'meaning potential'[15] and limitations of what can be expressed within a mode based on its historical, social, cultural and technical histories, uses and capacities[16]. In the case of the internet and social media as an instance of multimodality, we have yet to fully understand the impact of its emergent virtual affordances in relation to online violent extremist discourse. Nor do we yet have the kinds of analytical and applied tools that will help us gain deep interpretive insight into such convergences and then operationalise this knowledge in defense and security contexts. However, once we shift toward the analysis of social media as a multimodally grounded discursive and sensory space[17] in which violent agendas and moral persuasion are encapsulated and reconfigured, then social media data can help us understand how violence is experienced through specific moral sentiments that get played out online[18]. Engaging critically not only with *human social interactivity* but also with *intermodal interactivity*[19] can help develop new knowledge and insights about how social media messaging uses multimodal strategies to achieve a range of social semiotic effects that lie beyond surface analysis of 'content' or 'platform' but that nevertheless exert enormous influence for individuals at the experiential level of meaning-making in situated contexts[20].

The data for this study comes from the English language edition of *Dabiq,[21]* the online magazine of the organisation which refers to itself as Islamic State (also known as ISIS and ISIL), given the prominence of this extremist group and their prolific media output.[22] Although all ISIS media content is published in Arabic, much is also published in other languages to reach non-Arabic speaking audiences, particularly Western audiences. The study focuses on materials published in English, the language other than Arabic most frequently used by ISIS. The English language version of *Dabiq*, ISIS's official online internet magazine, is chosen as it is "one of the few original sources of data that directly comes from ISIS".[23] The aim of this paper is to present the mixed methods approach and research design for analysing text and image relations in *Dabiq* and online media texts with images from the magazine, where the objective is:

- to analyse how the text and images in the ISIS online magazine *Dabiq* function to promote extremist views, incite violence, and recruit jihadists.
- to use data mining techniques to analyse how these images are re-contextualised multimodally in social media, online news and other websites to promote and resist violent extremism.

The mixed methods approach contributes to the interdisciplinary study of big (digital) data sets of human discourse, by providing (a) evidence-based assessment of how ISIS use language and images for propaganda purposes and how these images are re-used in ways which function to support and resist terrorism in different media platforms; and (b) theoretically informed empirical techniques for automated analysis of text and image relations in large datasets of multimodal texts. From here, the aim is to develop interactive information visualisations for displaying the resulting discourse patterns as a realisable aspiration resulting from this approach.

Background

Multimodal discourse analysis, the study of the meaning arising from the integration of language, images and other resources in texts, interactions and events[24] has emerged as an interdisciplinary field of research, providing powerful analytic frameworks for analysing how language and images combine to communicate meaning.[25] However, large-scale analysis[26] and corpus-based empirical grounding and testing of insights[27] are required to map discourse patterns and trends. Such research is already underway: for example, O'Halloran, Chua and Podlasov[28]

use large-scale analysis of text-based and visual social media data in an interactive map to investigate urban life in Singapore. Their study demonstrates how vast quantities of different social media data can be amassed, clustered, analysed and converted into interactive visualisations to provide rich resources for mapping socio-cultural trends, in this case to investigate variations in language and image use according to social context. In another study, Podlasov and O'Halloran[29] integrate automated analysis of photographs, data visualisation techniques using self-organising maps and topology learning algorithms with multimodal social semiotic analysis to map patterns and trends in Japanese street fashion. These studies demonstrate how different levels of analysis for multimodal texts, context, and culture can be integrated using multimodal social semiotic theory and digital approaches.

The mixed methods approach developed in this paper builds upon this innovative program of research to investigate how ISIS use text and images for propaganda purposes and how these images are re-used in different media platforms for various purposes. Extremist and terrorist groups have long used the internet "for a wide range of purposes including community and operational communication and propaganda, technical information sharing and intelligence gathering, recruitment, training, financing and equipment procurement".[30] Terrorist groups' use of the internet has rapidly evolved "as a result of new technological opportunities, the proliferation of social media platforms, developments in online policing and a number of other factors".[31] The focus of this study is how ISIS uses language and images in propaganda materials which are freely available on the internet (in this case, *Dabiq* magazine) and how images from these sources are re-contextualised multimodally in ways that support and resist terrorism. ISIS makes extensive use of images (including images from elsewhere) because they are cheap and easy to produce compared to videos and other materials (e.g. news reports, statements and so forth).[32] Also, the impact of images is immediate, unlike long stretches of text which take more time to absorb.

The mixed methods approach involves qualitative, social semiotic analyses of the text and image relations in *Dabiq* which are transformed into quantitative data using multimodal annotation software. These quantitative results inform machine learning techniques that are applied for data mining of text and image relations in large multimodal data sets which contain the same or edited

versions of the images from *Dabiq*. The outcomes are a multilevel contextual model with empirical techniques for mapping discourse patterns in large databases of multimodal texts. The approach complements and extends existing research in terrorism studies which apply discourse analytical perspectives, mixed methods methodologies, quantitative computational approaches, and sociological and cultural studies approaches to the study of violent extremist propaganda materials, as described below.

In the first case, the approach extends existing approaches from discourse analytical perspectives, as represented, for example, by Al-Hejin's[33] corpus-based analysis of the semantic macrostructures and discursive strategies employed in the representation of Muslim women; Bhui & Ibrahim's[34] review of the rhetorical patterns, images and symbolisms manifested in text, videos, and interactive discourse formats deployed as persuasive techniques on jihadist websites; Jan's[35] study of how extremist news discourse conceptualises the world, the Self and the Other, conducted from a poststructuralist discourse perspective; and Wyszomierski's[36] critical discourse analysis of the structural relations of dominance in depictions of Boko Haram's evolving video propaganda capabilities. In the proposed approach the social semiotic frameworks for analysis of text and image relations are implemented as data mining and visualisation techniques to study discourse patterns in large multimodal datasets. In this respect, the approach also extends research that applies qualitative and quantitative (mixed) methods to the analysis of jihadist multimedia propaganda material, such as represented, for example, by the work of Allendorfer and Herring[37] who conduct a comparative content analysis of the themes and language used in the scenes and utterances of ISIS videos and videos posted in response by the U.S. Department of State; and the work of Wright and Bachmann[38] who perform a content analysis of the persuasive devices used in jihadist and Islamic extremist digital media products. The proposed approach involves data analysis of multimodal content, together with analysis of the evaluative stance towards those materials in large multimodal datasets, representing a new approach in the field of terrorism studies.

Secondly, the approach complements quantitative computational approaches to text analysis of Al Qaeda, ISIS, and Taliban propaganda material, as represented, for example, by Skillicorn's[39] empirical assessment of the language used in English language magazines produced by jihadist

groups; and the work of Vergani and Bliuc[40] who use the computerised text analysis program LIWC (Linguistic Inquiry and Word Count) to investigate the evolution of the language employed in *Dabiq*; as well as qualitative approaches to image analysis, as represented, for example, by Kovács'[41] investigation of visual representations in propaganda material produced by various Sunni Islamic radical and reformist movements. In this case, the focus is the text and images relations in multimodal texts, rather than the analysis of language and images as separate resources.

Lastly, the approach complements research which draws upon sociological and cultural studies approaches to the study of online violent extremist discourse, as represented, for example, by Aly, Weimann-Saks & Weimann's[42] and Archetti's[43] research on socially constructed narratives in communicative approaches to countering terrorism online; Grossman & Tahiri's[44] qualitative investigation of grassroots community and government perspectives about what radicalisation and violent extremism mean for Australian communities; Grossman's[45] study of narrative and discourse in countering violent extremist propaganda; and O'Hara & Stevens'[46] assessment of the various uses of the Internet for propaganda purposes, such as recruitment, radicalisation, and incitement, by radical extremist groups. In this case, the study is concerned with the language and images used by violent extremists, and how images are re-contextualised multimodally to both resist and support terrorist activities.

Significance of the Mixed Methods Approach

The approach offers a new research paradigm to complement the state-of-the art approaches to social media intelligence for counter-terrorism, as described by Bartlett and Reynolds.[47] That is, existing studies examine the language and images of violent extremist discourse from a range of perspectives as described above, but none of the approaches involve *the study of the meanings arising from the text and image relations in large datasets*. In this case, the approach leads to enhanced understanding of the multimodal strategies employed by violent extremist groups, and how the associated images are re-used for different purposes. This includes 'counter-speech', a strategy designed to confront 'hate speech', given that censorship and filtering initiatives have been judged to be largely ineffective because terrorist supporters rapidly reconfigure themselves

to avoid moderation, and in doing so, find new ways to circulate their messages.[48] While counter-speech has advantages (e.g. speed, flexibility, responsiveness and maintenance of public space for debate), its forms are varied, as are the effects that can be counter-productive. For this reason, "there is growing interest in a more rigorous and evidence-led approach to counter-speech",[49] as developed here. The outcomes of the approach include empirical techniques for identifying text and image relations that conform to certain profiles such as legitimising, supporting, rejecting and resisting terrorism.

In the proposed approach, multimodal social semiotic theory provides the foundation for developing a multi-level contextual model for analysing text and image relations. The model draws upon knowledge about sociocultural life from Wikipedia to provide contextual information for interpreting keywords in the text and key objects in the images, resulting in hierarchical layers of semantic information that allow connections to be made between the import of words and images and their contextual relations. Theoretically and analytically, the approach will result in an enhanced capacity for analysing, visualising and tracking significant patterns of meaning in large data sets. The approach is described below.

The Theoretical Framework

The mixed methods approach integrates qualitative methods of multimodal discourse analysis with quantitative methods of data mining and information visualisation to develop and implement a multilevel contextual model for analysing multimodal texts involving language and images. The theoretical approach is based on social semiotics,[50] which studies how sign systems are used to create meaning contextually. In this case, multimodal social semiotics is concerned with the meaning arising from the interaction of language with other resources in multimodal texts.[51] Specifically, the approach is based on Michael Halliday's[52] systemic functional theory, where language, images and other sign systems (i.e. semiotic resources) are viewed as resources for making meaning. The functions which semiotic resources have evolved to serve are reflected in their underlying organisation. As such, semiotic resources are conceptualised as 'systems of meaning' with networks of options from which choices are made in multimodal texts. Following Halliday, the systems are organised according to (a) experiential and logical meaning: to

structure experience of the world; (b) interpersonal meaning: to enact social relations and create a stance towards happenings; and (c) textual meaning: to organise the messages into coherent forms. The messages in any communicative situation are characterised in terms of options selected from these systems.

Table 1: Examples of Text and Image Systems[53]

Metafunction	Rank	System	Description
Text			
Experiential	Clause	Processes; Participant Roles; Circumstance	Happenings, actions and relations
Interpersonal	Discourse	Appraisal	Evaluation in terms of attitude, emotion and judgment
	Clause	Speech Function	Exchange of information (e.g. statements and questions) and goods & services (e.g. commands and offers)
Textual	Clause Discourse	Information Focus	Organisation of information, with points of departure for what follows
Pictures			
Experiential	Work	Narrative Theme; Representation; Setting	Nature of the scene
	Episode	Processes; Participant Roles; and Circumstance	Visual happenings, actions and relations
	Figure	Posture; Dress	Characteristics of the participants
Interpersonal	Work	Angle; Camera Distance; Lighting	Visual effects
	Episode	Proportion in Relation to the Whole Image: Focus; Perspective	Happenings, actions and relations with respect to the whole image
	Figure	Gaze-Visual Address	Direction of participant's gaze as internal to image or external to viewer

Textual	Work	Compositional Vectors; Framing	The organisation of the parts as a whole, with the visual marking (e.g. framing) of certain parts
	Episode	Relative Placement of Episode; Framing	Position of the happenings, actions and relations in relation to the whole image, and the visual marking of certain aspects
	Figure	Relative Placement of the Figure within the Episode; Arrangement; Framing	Position of figures in relation to happenings, actions or relations, and the visual marking of certain aspects of those figures

Halliday's systemic functional theory provides the foundation for modelling language and images as inter-related systems of meaning, and analysing how choices from those systems work together to create meaning in multimodal texts. The metafunctionally-based systems are organised according to different ranks of constituency, as illustrated in Table 1. These are based on Halliday's[54] and Martin's[55] systems for language and O'Toole's[56] framework for images. In many respects this is analogous to approaches within visual studies and communication studies that attempt to provide standardised (and hence quantifiable) methods for qualitative visual analysis:[57] here, however, the categories adopted have been related to broader schemes of functional interpretation both within and across modalities and artifacts.
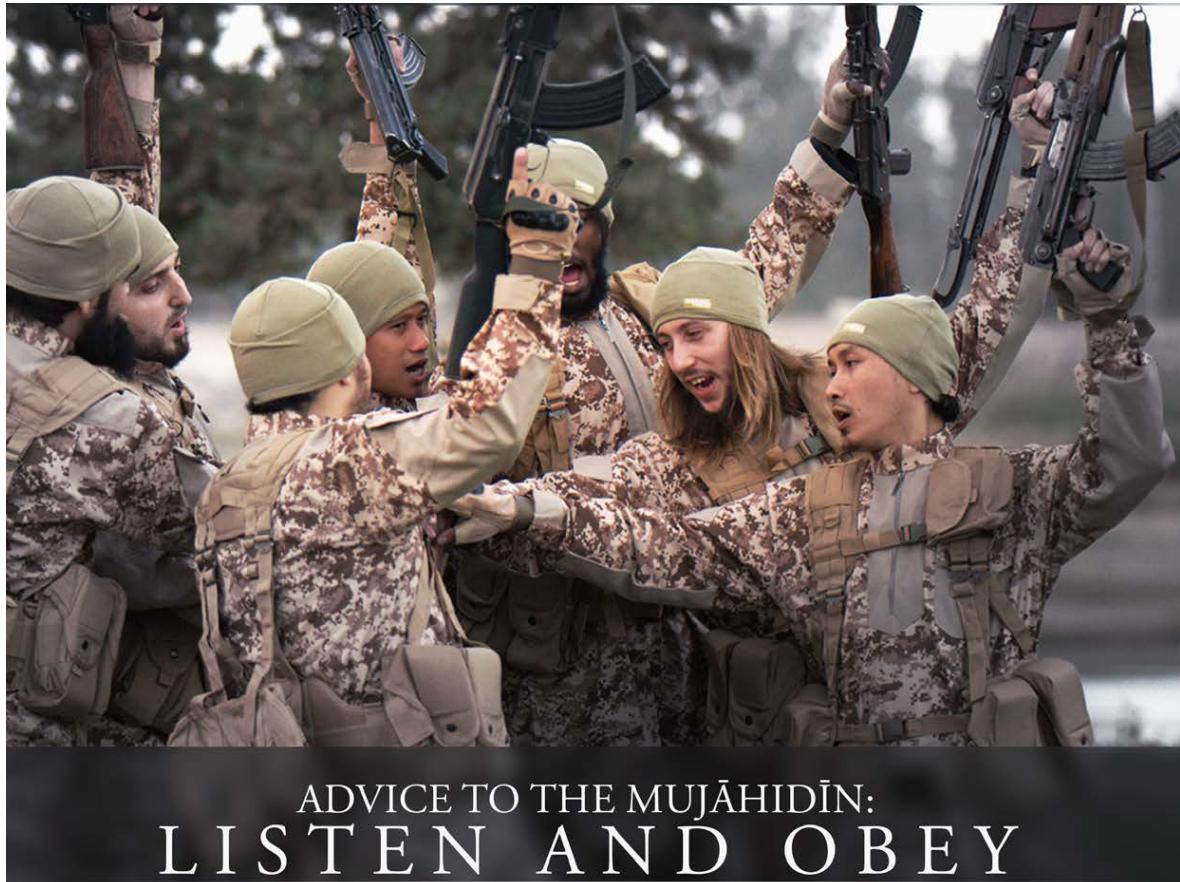
Figure 1: ISIS's *Dabiq* Magazine (Issue 12, p. 9)

As an example, the framework in Table 1 is used to analyse images in *Dabiq*, such as the one in Figure 1 which appears in the article "Advice to the Mujāhidīn: Listen and Obey" (*Dabiq*, Issue 12, p. 9). The distinctions in Table 1, although not exhaustive, provide a robust scaffold for drawing out of the image the interpretations that viewers will typically make, at the same time making explicit the qualities of the image that support those interpretations. Furthermore, the framework (with systems indicated by initial capital letters) can be used to analyse how these meanings are co-contextualised or re-contextualised by the linguistic text (and vice versa), as explained below.

For example, in Figure 1, there is a relative placement of the figures in the foreground in terms of composition, and the foreground is also in focus. Compositional Vectors (oblique) are formed by the seven figures and their raised arms and guns, making the scene dynamic, yet all members of the group are closely united through the stacked hands, a point emphasised by the Gaze of all

figures towards their hands. The effect of the raised AK47 rifles is highly interpersonal, as this is one of the most popular and widely used assault rifles that has become a symbol of resistance. In terms of experiential meaning, the figures are identifiable as mujahideen, given the choices made for Dress (military fatigues). They are engaged in processes of celebration, given their facial expression and open mouths (indicating a song or chant). Interpersonal meaning then makes a particularly significant contribution: even though the figure (second right) is partially obscured by the figure on the far right, the clear access granted to his face, together with his contrasting fair Caucasian features and long hair makes him critical to the interpretation of the photograph. In this case, the figure (second right) assumes a central Participant Role in what appears to be a culturally diverse, closely-knit team of army fighters. Interpersonal meanings are further contributed by the choices for the systems of Angle (approximately eye level), Camera Distance (medium-close), Lighting (naturalistic), Perspective (foreground) and the Gaze-Visual Address (internal to the image itself). These function to position the viewer as a close observer of the scene. Combining the meanings described, the reading for the image consists of a Narrative Theme where ISIS is legitimised as a proper state through the 'army' with standard battle fatigues. The multi-ethnic character of the army demonstrates a global support base in which Caucasian fighters are playing a key, active role. While the tone is celebratory in the image in Figure 1, the Speech Functions of statement and command in "Advice to the Mujāhidīn: Listen and Obey" (with large font for the commands) sets the moral tone for the information which follows. The linguistic analysis of this text (see Figure 2) reveals that ISIS legitimise their views using references to Allah and extracts from the Qur'an: for example, "Listening and obeying is from unity, cooperation and strength, in contrast with quarrelling, disunity, and disagreement. Quarrelling and disagreement in jihad is considered a means of failure and defeat" are followed by the Qur'anic reference Al-Anfal: 46. As illustrated here, the visual and linguistic strategies may differ – i.e. the celebrations of a close-knit 'legitimate' army are re-contextualised in terms of the strict behavioural codes dictated by the Qur'an – but they nonetheless function together to promote the agenda of ISIS. As this example shows, images play a key role, evidenced by the fact that ISIS media releases are primarily visual (i.e. pictures, videos and graphics).[58] Therefore the text and image relations in official media productions (i.e. *Dabiq*) are analysed in order to track how these images are re-contextualised in other media for different purposes.

There are various approaches to conceptualising text and image relations as described by Bateman,[59] but these are generally formulated in terms of logical relations derived from language. While such frameworks provide templates for multimodal relations, "it then remains a challenging and exciting research question to find out more precisely just whether there are further relations that are supported in other semiotic modes and to determine which particular subset of linguistic conjunctive relations may be employed and which not".[60]

In the approach developed below, text and image relations are conceptualised in terms of the co-contextualising and re-contextualising relations where similar meanings are reinforced (i.e. co-contextualised) and/or new meanings are made (i.e. re-contextualised). The nature of these contextualising relations are considered in relation to two strands of meaning: (a) experiential meanings, which are concerned with processes, participants and circumstances (see Martin and Rose's[61] discourse systems of ideation) and (b) interpersonal meanings which are concerned with evaluations of the content (i.e. attitude, emotion and judgment) (see appraisal theory[62]), These two strand of meanings are enabled by the compositional organisation of the multimodal text. The approach involves mapping the relations between choices from linguistic and visual systems for experiential and interpersonal meaning, as displayed in Table 1. Significantly, the approach permits semantic expansions to be mapped within and across ranks: for example, when key words function to re-contextualise the theme of an image, as discussed below.

The Research Design and Methods

The mixed methods approach is based on the premise that multimodal social semiotic theory offers a rich theoretical platform for manually analysing the meanings arising from text and image relations in *Dabiq*; the text and image relations in other media containing the same (or edited) images can be automatically interpreted using contextual information, machine learning and data mining techniques; and the resulting discourse patterns can be revealed using information visualisation techniques. The research design involves specific tasks for data collection, multimodal analysis, data mining, and information visualisation. The four stages in the research design are explained below.

Stage 1 – Data Collection: The multimodal data set consists of two components: (a) all issues of the English language edition of *Dabiq*, and (b) social media messages (e.g. Twitter, Instagram, Flickr), online news and other websites with images from *Dabiq*, collected over the duration of the study. The issues of *Dabiq* are readily available for download.[63] The large multimodal dataset with images from *Dabiq* are collected using automated Python scripts based on public APIs of these major data sources, including Google custom search,[64] Twitter[65] and Instagram.[66] For automated collection of generic website content, we use two Python libraries, including Scrapy[67] which is a popular Python library for large-scale web crawling and Beautiful Soup 4[68] which is a powerful library for parsing HTML and extracting specific content. The developed Python scripts will be deployed to collect selected websites that are related to *Dabiq*, and their contents will be extracted and stored in MongoDB databases. Note that this is in addition to the original *Dabiq* magazines which can be easily obtained from the Clarion project[69] and their PDF-formatted contents can be parsed with Python tools such as PDFMiner[70]. For searching and collecting *Dabiq*-related tweets from Twitter, we use Teepy[71] which is a powerful Python library for accessing the Twitter API. For automated collection of *Dabiq*-related images via Google, we use the Python tool GoogleScraper[72]. With Instagram, there is some restriction on automated collection imposed by the public API, and hence we need to develop custom tools based on our previous experience with crawling content from Weibo with similar restrictions. All collected data will be parsed, indexed, and tagged suitably and stored in MongoDB databases for subsequent high-level analysis.

The multimodal dataset will not be exhaustive; for instance, terrorist accounts are constantly being closed down and re-opened under new names.[73] However, the aim is not to track ISIS's use of social media and smart phone apps. Rather, the aim is to analyse how images in *Dabiq* are re-contextualised multimodally in different media platforms in various ways that support and resist the group's agenda. The collection of freely available multimodal texts with images from *Dabiq* will result in a large dataset suitable for this purpose.

Stage 2 – Multimodal Analysis: The qualitative analysis of *Dabiq* is undertaken manually using Multimodal Analysis Image software.[74] Catalogues of system networks can be freely defined in the software for different research purposes.[75] In this case, pages from *Dabiq* are imported and

analysed using sets of analytical categories (e.g. the systems in Table 1), represented as system networks of options (e.g. the system of 'Gaze-Visual Address' has the options of 'Direct', 'Indirect' and 'None').
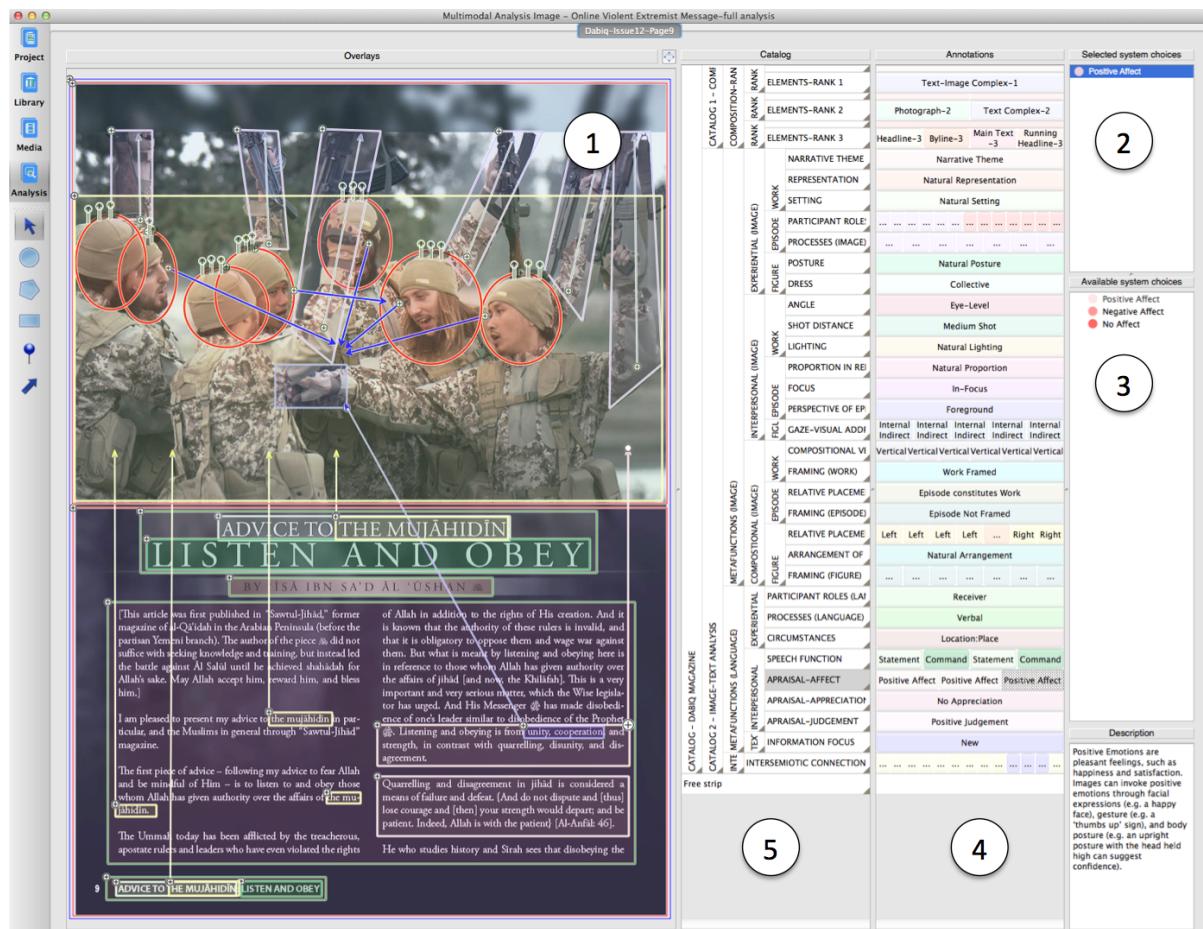

Figure 2: Screenshot of Sample Analysis in Multimodal Analysis Image: Overlays [1]; Selected system choice [2]; List of available system choices [3]; Annotation nodes [4]; System strips [5]

In Figure 2 for example, the image and text from Figure 1 are annotated by creating overlays (in the form of rectangles, circles, polygons, lines or pins) on the image [1], and assigning a system choice [2] from a list of available system choices [3] from the system frameworks (which contain the various options) to those overlays. The annotations of assigned system choices and overlay coordinates are stored in the database for machine learning purposes. The structured data consists of the text identification code, the system catalogue identification code, the system name, the set of options in the system network organised according to hierarchical order, the selected system

choice, the overlay type for the system choice (i.e. rectangle, circle, polygon, line or pin), the coordinates of the overlay, and the textual descriptors: i.e.

Text/Image Name, System Catalogue, System Name, System Network, Selected System Choice, Overlay Type, Points X-Y, Description.

The structured data set for each annotation is stored in the database. This includes the spatial co-ordinates of the overlay which has been placed over the image (e.g. the x and y co-ordinates for a rectangular overlay) and the system choice which corresponds to those co-ordinates. As the spatial co-ordinates of the overlay correspond exactly to words and/or parts of the image in the original media file, the results of the analysis can be machine learnt (i.e. this part of the image corresponds to this system choice) to inform the subsequent development of data mining algorithms.

Stage 3 – Data Mining: The manual analysis of the *Dabiq* dataset using Multimodal Analysis Image software provides the basis for machine learning to develop data mining algorithms for analysis of the large dataset of multimodal texts with the images from *Dabiq*. However, this large dataset contains linguistic text, and in some case images, which extend beyond the *Dabiq* database. For this reason, the automated analysis of the large database is undertaken using contextual information from Wikipedia, which provides an extensive classification system for deriving structured information about human socio-cultural life, and by default, what people are communicating via linguistic and visual resources. Wikipedia categorisation structures provide reasonably robust semantic approximations of higher order semantic categories for keywords associated with a Wikipedia page, as shown in O'Halloran, Chua et al.'s[76] study where Wikipedia categorisations of Foursquare venues provided semantic descriptions of social activities that revealed how linguistic and visual selections in social media vary across different social contexts in Singapore.[77]

In this case, Wikipedia classifications are used to provide higher order semantic categories for keywords in the linguistic text and, where relevant, key objects in the images. For example, "AK-47" (the word or image) is easily identifiable as an 'assault rifle', 'firearm' and 'weapon'

using Wikipedia classifications, as seen in the different category levels for this term in Table 2. Also, key words such as "financial crisis" (see The Guardian article in Figure 3) are related to semantic concepts such as 'risk', 'economic problems', 'money', 'financial system' and 'behavioral and social facets of systemic risk' using Wikipedia categories, as displayed in Table 2. In this way, keywords in the linguistic text and key objects in the images, identified by natural language processing and image processing algorithms, are linked to higher order semantic information, providing the basis for interpreting the experiential meaning of linguistic and visual elements. DBPedia[78] is used to organise the structured representation of Wikipedia classification trees (i.e. article categories and categories of category mappings). DBPedia is a community project that structures information from different multilingual versions of Wikipedia. According to Lehman et al.[79], there are more than 13 millions things in DBPedia, each "thing" refers to a Wikipedia concept. They are organised in an ontology consisting of 320 classes and 1,650 properties. The use of DBPedia in the research community has recently gained particular interest as it is an additional source for mapping semantic concepts, which is crucial for improving the conventional shallow approach in data mining that simply relies on the input data. Using the knowledge bases provided by DBPedia, one may discover possible links between different observed concepts/topics (extracted by multimodal analysis and data mining technologies) by analysing the semantic hierarchy. An example of such an application in the context of clustering scientific articles is presented in Szczuka et al.[80], where DBPedia is used as an additional dictionary to enable a more accurate, semantically-oriented distance calculation between semantically-related text items. In this project, each discovered topic from a given social source will be mapped to the DBPedia semantic hierarchy. This will then allow us to subsequently analyse the semantic distance to their discovered topics and thus establish the social links between different multimodal sources. From here, key linguistic and visual selections associated with a Wikipedia page are mapped onto higher-level semantic categories, providing the basis for formulating the experiential meaning of the text and new images in the large multimodal database. The interpersonal meanings arise from automated sentiment analysis[81] and analysis of non-core words[82] which carry positive or negative connotations.

Table 2: Sample 3-level Categorisation Tree for "AK-47" and "Financial Crisis"

Keyword/Object	AK-47	Financial Crisis
Categories: Level 1	AK-47	Financial Crisis
Categories: Level 2	1947 introductions; 7.62 mm rifles; Assault rifles; Infantry weapons of the Cold War; Rifles of the Cold War; Cold War weapons of the Soviet Union; Kalashnikov derivatives; Soviet inventions; Weapons of Russia; Military equipment 1945–1949	Crisis; Financial crises; Systemic risk; Economic problems; Economic bubbles
Categories: Level 3	20th-century introductions; 7.62 mm firearms; Rifles; Assault rifles; Personal weapons; Weapons of the Cold War; Infantry weapons of the Cold War; Weapons of the Soviet Union; Weapons by country; Cold War weapons by country; Inventions by country; Science and technology in the Soviet Union; Military equipment of Russia; 20th-century military equipment; 1940s	Risk; Business cycle; Economic crises; Economic problems; Financial economics; Financial system; Monetary economics; Money; Financial risk; Financial system; Economics; Hazards; Social problems; Behavioral and social facets of systemic risk; Economic problems; Economic booms

Figure 3: The Guardian January 28, 2016

For example, the article "ISIS hit by financial crisis" (Figure 3) from The Guardian newspaper contains the *Dabiq* image from Figure 1. Using Wikipedia categories (see Table 2), keywords in the linguistic text (e.g. "financial crisis", "salary cuts for fighters" and "ISIS government") are related to the concepts of 'crisis', 'systemic risk' and 'economic problems'. In this case, the linguistic text functions to co-contextualise and re-contextualise different aspects of the experiential meanings of the image that will have been machine learnt via the manual multimodal annotation.

That is, the text and image both function to legitimise ISIS as an officially recognisable government (e.g. "the ISIS government" and the standard army outfits in the image), and the positive emotional appeal of the image supports that legitimisation even further, in this case for Guardian readers. However, the linguistic text positions ISIS as undergoing an economic crisis, resulting in salary cuts for the mujahideen, making the celebratory activities of the army in the image incongruous. In addition, the sentiment analysis and use of non-core words (e.g. "hit', "slashed" and "rampaging soldiers of the Islamic State") together with the objective stance ("it has been decided to reduce the salaries" and "it is not allowed for anyone to be exempted") functions to negatively evaluate the situation and ISIS itself, again re-contextualising the celebratory activities in the image, although it would appear difficult for these linguistic meanings to override the strong emotional appeal of the image. As such, it becomes possible to analyse the co-contextualising and re-contextualising text and image relations from experiential and interpersonal viewpoints: in this case to legitimise the status and appeal of ISIS, while attempting to subvert its views in relation to jihadist recruitment. This example illustrates the significance of the proposed approach for understanding the impact of image selections in Western media and other online sources. Rather than being the result of logical and measured thought, image selections can be based on any number of factors (e.g. expediency, content, the pressure of reporting deadlines, and so forth). The proposed approach has even more significance in this context because it offers an interpretive lens for increased awareness of the meanings which images make when they are combined with text.

The use of data mining algorithms and Wikipedia classifications in a semantically focused, multiscale model addresses the current limitations of existing approaches such as bag-of-words modeling (where the semantic relations between elements are not addressed) and its basic probabilistic extensions[83] and image processing techniques which also do not address semantic relations between objects.[84] In this case, multiple semantic layers are introduced over the raw data, resulting in descriptions of the text, the image and their relations in relation to experiential and interpersonal meaning. In this way, it is possible to demonstrate how ISIS's images are re-used in relation to the terrorist group's agenda. The specification of text and image relations that conform to different profiles (e.g. legitimising, supporting, rejecting and resisting terrorism) are

verified manually by selecting instances of each type of relation and then checking to confirm that these cases are correctly identified.

Stage 4 – Information Visualisation: Visualisation techniques assist to determine features from complex multidimensional analyses and the interpretation of large-scale automated recognition analyses.[85] In this case, an interactive visualisation needs to be used to display the relations between the multimodal texts according to various dimensions: i.e. multimodal text type, interpersonal meaning, experiential meaning, time, source, location and other metadata. By comparing trends and patterns along several data dimensions, the visualisation (e.g. with clustering algorithms) brings forth insights and permits testing of hypotheses in an iterative and interactive way. The data visualisation will be defined by carefully defined design and style choices,[86] permitting exploration of patterns for profiling the nature of the text and image relations in the dataset.

Concluding Comments and Future Research Directions

Each stage of the proposed mixed methods approach is possible using existing resources. That is, the manual annotation of *Dabiq* using established text and image frameworks and multimodal analysis software will produce empirical-based patterns in the text and image relations in *Dabiq*. These semantic patterns are related to features of the text and images so they can be used for machine learning for automated analysis of the large database of multimodal texts with images from *Dabiq*. The use of Wikipedia categories for key words and objects detected by natural language and image processing algorithms in the large multimodal database will produce higher order semantic information to automatically determine the nature of the co-contextualising and re-contextualising relations for experiential and interpersonal meaning for the text and image relations. The resulting discourse patterns will be discernible using an interactive visualisation with clustering algorithms which function across various dimensions of the data.

The proposed approach is cutting-edge research that builds upon existing work to develop new techniques with evidence-based findings of discourse strategies employed by ISIS to perpetuate their views and how these are taken up and resisted by the global community in public online spaces, paving the way for next-generation data analytics that has applications for fields beyond

terrorism studies. As such, the approach is the first step towards the development of a digital mixed methods research design which makes use of the potential of data mining and information visualisation for exploring discourse patterns in large datasets.[87]

In summary, the integration of multimodal analysis, data mining and information visualisation provides the foundation for articulating theoretical contributions to discourse analysis and technological tools for visualising, analysing and extracting relevant information from social discourse. Moreover, the use of mixed methods approach and advanced computational techniques for investigating human communication is aligned with the goals of computational social science, which aims to develop quantitative approaches to modeling and understanding social systems characterised by multiple, interrelated layers of complexity.[88] The challenges of this endeavour cannot be underestimated, however. Representative datasets must be collected, and this is a complex undertaking when search engines such as Google have implemented restrictions with regards to automated data collection. The manual annotation of multimodal texts is timeconsuming and inaccuracies are inadvertedly introduced by the human analyst. Also, the task of associating higher order semantic categories to texts, images and their relations is difficult because meaning is context dependent and constantly evolving. We are seeking to overcome this problem by introducing contextual information to aid the intepretative process but this is fraught with difficulties, given the constant re-configurations of text and images in the wide range of media texts which are now available online. Lastly, we aim to visualise the resulting discourse patterns, building upon visual information research. This is equally a major undertaking given its current state of the art which only recently has expanded its scope beyond capturing the influence of human perception, towards identifying the social and culturally factors that construe meaning in data visualisations and infographics, such as overt figurative forms or more subtle visual conventions[89]. Regardless, we aim to make as much progress as possible in order to develop new digital tools and techniques for interpreting multimodal digital data as described in this paper.

Much research in the humanities has focused on the potential negative aspects of the 'Big Brother' scenario of digital data (e.g. surveillance, ethical issues of privacy and confidentiality, openness and transparency versus security issues)[90]. However, the objective of the approach

developed here is to address the "Big Problems of society, [thereby] avoiding crises and threats to its stability and healthy development".[91] Indeed, the aim is to develop "new tools and methods that would be applicable in any instances where Big Data are a key ingredient".[92] In this scenario, distinct discipline areas would be combined into a "truly disciplinary, non-compartmental science",[93] overcoming the traditional separation between humanities and the sciences which, as Snow[94] claimed over half a decade ago, has resulted in a division which has created a major barrier to solving the world's problems. This barrier can be dismantled, as suggested by the approach described above. As Raghavan[95] claims, "it is time to scale the science in the social sciences" to develop new methodologies which transform difficult problems "into a series of robust analyses that are replicable and composable"[96] to create "enormous progress for public good".[97] In this regard, multimodal analysis is poised to become an integral component of big data analytics, with the promise of a new research paradigm for studying violent extremist discourse in the world today.

Key Terms and Definitions[98]

Notes

[1] C. Jewitt (ed.), *The Routledge Handbook of Multimodal Analysis*, 2nd ed. (London: Routledge, 2014).

[2] J. Bateman, *Text and Image: A Critical Introduction to the Visual/Verbal Divide* (London & New York: Rouledge, 2014).

[3] K. L. O'Halloran, "Multimodal Digital Humanities" in *International Handbook of Semiotics*, ed. P. Trifonas (Dordrecht: Springer, 2015), 383-409; K. L. O'Halloran, S. Tan, D.-S. Pham, J. Bateman & A. Vande Moere, "A Digital Mixed Methods Research Design: Integrating Multimodal Analysis with Data Mining and Information Visualization for Big Data Analytics", *Journal of Mixed Methods Research* (2016), DOI: 10.1177/1558689816651015.

[4] A. Kovács, "The "New Jihadists" and the Visual Turn from Al-Qa'ida to ISIL / ISIS / Da'ish", *BiztPol Affairs* 2, no. 3 (2015): 67.

[5] J. Klausen, "Tweeting the Jihad: Social Media Networks of Western Foreign Fighters in Syria and Iraq", *Studies in Conflict and Terrorism* 38.1 (2015): 3.

[6] J. M. Berger & B. Strathearn, *Who Matters Online: Measuring Influence, Evaluating Content and Countering Violent extremism in Online Social Networks* (London: International Centre for the Study of Radicalisation and Political Violence, 2013), http://icsr.info/wp-content/uploads/2013/03/ICSR_Berger-and-Strathearn.pdf; , J. A. Carter, S. Maher & P. R. Neumann. #*Greenbirds: Measuring Importance and Influence in Syrian Foreign Fighter Networks* (International Centre for the Study of Radicalisation and Political Violence, 2014), http://icsr.info/wp-content/uploads/2014/04/ICSR-Report-Greenbirds-Measuring-Importance-and-Infleunce-in-Syrian-Foreign-Fighter-Networks.pdf; M. Sageman, *Leaderless Jihad: Terror*

Networks in the 21st Century (Philadelphia: University of Pennsylvania Press, 2014); I. von Behr, A. Reding, C. Edwards & L. Gribbon, *Radicalisation in the Digital Era: The Use of the Internet in 15 cases of Terrorism and Extremism* (RAND Europe, 2013), http://www.rand.org/content/dam/rand/pubs/research_reports/RR400/RR453/RAND_RR453.pdf ; G. Weimann *Terror on the Internet: The New Arena, the New Challenges* (Washington, DC: United States Institute of Peace Press, 2006).

[7] D. Murthy, "Towards a Sociological Understanding of Social Media: Theorizing Twitter" *Sociology, 46*(6), (2013): 1059-1073; K. Z. Sands, "Muslims, Identity and Multimodal Communication on the Internet", *Contemporary Islam, 4*, (2010): 139-155.

[8] Berger & Strathearn *"Who Matters Online: Measuring Influence, Evaluating Content and Countering Violent Extremism in Online Social Networks"*(see note 6); Carter, Maher & Neumann *#Greenbirds: Measuring Importance and Influence in Syrian Foreign Fighter Networks* (see note 6); Klausen "Tweeting the Jihad: Social Media Networks of Western Foreign Fighters in Syria and Iraq" (see note 5).

[9] Z. Papacharissi (ed.), *A Networked Self: Identity, Community and Culture on Social Network Sites* (New York & London: Routledge, 2011).

[10] M. Castells, *Communication Power* (Oxford: Oxford University Press, 2009).

[11] C. Jewitt, "An Introduction to Multimodality" in C. Jewitt (ed.), *The Routledge Handbook of Multimodal Analysis*, 2nd ed (London & New York: Routledge, 2014), 14-27: p. 1.

[12] J. Blascovich & C. McCall, "Social Influence in Virtual Environments" in K. E. Dill (ed.), *The Oxford Handbook of Media Psychology* (Oxford: Oxford University Press, 2013), 305-315.

[13] Ibid., p. 312.

[14] Jewitt "An Introduction to Multimodality": 24-25 (see note 11); G. Kress, "What is Mode?" in C. Jewitt (ed.), *The Routledge Handbook of Multimodal Analysis,* 2nd ed. (London & New York: Routledge, 2014), 54-67: 55-58.

[15] M. A. K. Halliday, *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. (London: Edward Arnold, 1978); T. van Leeuwen, *Introducing Social Semiotics*. (London: Routledge, 2015).

[16] G. Kress & T. van Leeuwen, *Multimodal Discourse: The Modes and Media of Contemporary Communication Discourse* (London: Arnold, 2001).

[17] J. L. Lemke, "Multimedia and Discourse Analysis" in J. P. Gee & M. Handford (eds.), *The Routledge Handbook of Discourse Analysis*. (London & New York: Routledge, 2013.

[18] M. Hynes & S. Sharpe, "Affected with Joy: Evaluating the Mass Actions of the Anti-Globalisation Movement", *Borderlands, 8*(3) (2009): 1-21.

[19] Jewitt, *The Routledge Handbook of Multimodal Analysis* (see note 1).

[20] J. L. Lemke, "Place, Pace, and Meaning: Multimedia Chronotopes" in S. Norris & R. H. Jones (Eds.), *Discourse in Action: Introducing Mediated Discourse Analysis* (London: Routledge, 2005), 110-122.

[21] http://www.clarionproject.org/news/islamic-state-isis-isil-propaganda-magazine-dabiq

[22] M. Lombardi, "Islamic State Communication Project", *Sicurezza, Terrorismo e Società* 1, no. 1 (2015): 99-133; A. Y. Zelin, "Picture or It Didn't Happen: A Snapshot of the Islamic State's Official Media Output", *Perspectives on Terrorism* 9, no. 4 (2015).

[23] M. Vergani and A.-M. Bliuc, "The Evolution of the ISIS' Language: A Quantitative Analysis of the Language of the First Year of Dabiq Magazine", *Sicurezza, Terrorismo e Società* 1, no. 2 (2015): 7-20, p. 8.

[24] Jewitt, *The Routledge Handbook of Multimodal Analysis* (see note 1).

[25] Bateman, *Text and Image: A Critical Introduction to the Visual/Verbal Divide* (see note 2).

[26] O'Halloran, "Multimodal Digital Humanities" (see note 3).

[27] J. Bateman, "Methodological and Theoretical Issues for the Empirical Investigation of Multimodality" in N.-M Klug and H. Stöckl (eds.) *Sprache Im Multimodalen Kontext / Language and Multimodality* (Berlin: de Gruyter Mouton, 2016).

[28] K. L. O'Halloran, A. Chua, and A. Podlasov, "The Role of Images in Social Media Analytics: A Multimodal Digital Humanities Approach", in D. Machin (ed.) *Visual Communication*, (Berlin: Gruyter, 2014), 565-588.

[29] A. Podlasov and K. L. O'Halloran, "Japanese Street Fashion for Young People: A Multimodal Digital Humanities Approach for Identifying Socio-Cultural Patterns and Trends", in E. Djonov & S. Zhao (eds.) *Critical Multimodal Studies of Popular Culture* (New York: Routledge, 2014), 71-90.

[30] J. Bartlett and L. Reynolds, *The State of the Art 2015: A Literature Review of Social Media Intelligence Capabilities for Counterterrorism* (London: Demos, 2015), 12.

[31] Ibid.

[32] Zelin, "Picture or It Didn't Happen: A Snapshot of the Islamic State's Official Media Output" (see note 22).

[33] B. Al-Hejin, "Covering Muslim Women: Semantic Macrostructures in BBC News", *Discourse & Communication* 9, no. 1 (2015): 19-46.

[34] K. Bhui and Y. Ibrahim, "Marketing the ''Radical'': Symbolic Communication and Persuasive Technologies in Jihadist Websites", *Transcultural Psychiatry* 50, no. 2 (2013): 216–234.

[35] F. Jan, *The Muslim Extremist Discourse: Constructing Us Versus Them* (London: Lexington Books, 2015).

[36] L. E. Wyszomierski, "Boko Haram and the Discourse of Mimicry: A Critical Discourse Analysis of Media Explanations for Boko Haram's Improved Video Propaganda Quality", *Critical Studies on Terrorism* 8, no. 3 (2015): 503-515.

[37] W. H. Allendorfer and S. C. Herring, "ISIS Vs the U.S. Government: A War of Online Video Propaganda", *First Monday* 20, no. 12 (2015): DOI: 10.5210/fm.v20i12.6336.

[38] J. E. Wright and M. Bachmann, "Al Qaida's Persuasive Devices in the Digital World", *Journal of Terrorism Research* 6, no. 2 (2015): 70-82.

[39] D. B. Skillicorn, "Empirical Assessment of Al Qaeda, ISIS, and Taliban Propaganda", *Intelligence and Security Informatics (ISI), 2015 IEEE International Conference* (2015): 61-66.

[40] Vergani and Bliuc, "The Evolution of the ISIS' Language: A Quantitative Analysis of the Language of the First Year of Dabiq Magazine" (see note 23).

[41] Kovács, "The "New Jihadists" and the Visual Turn from Al-Qa'ida to ISIL / ISIS / Da'ish." (see note 4).

[42] A. Aly, D. Weimann-Saks, and G. Weimann, "Making 'Noise' Online: An Analysis of the Say No to Terror Online Campaign", *Perspectives on Terrorism* 8, no. 5 (2014): 33-47.

[43] C. Archetti, "Terrorism, Communication and New Media: Explaining Radicalization in the Digital Age", *Perspectives on Terrorism* 9(1) (2015): 49-59.

[44] M. Grossman and H. Tahiri, "Community Perceptions of Radicalisation and Violent Extremism: An Australian Perspective", *Journal of Policing, Intelligence and Counter Terrorism* 10, no. 1 (2015): 14-24.

[45] M. Grossman, "Disenchantments: Counter-Terror Narratives and Conviviality", *Critical Studies on Terrorism* 7, no. 3 (2014): 319-335.

[46] K. O'Hara and D. Stevens, "Echo Chambers and Online Radicalism: Assessing the Internet's Complicity in Violent Extremism", *Policy & Internet* 7, no. 4 (2015): 401–422.

[47] Bartlett and Reynolds, *The State of the Art 2015: A Literature Review of Social Media Intelligence Capabilities for Counterterrorism* (see note 30).

[48] Ibid.

[49] Ibid., 22-24.

[50] For example, van Leeuwen, *Introducing Social Semiotics* (see note 15).

[51] Bateman, *"Methodological and Theoretical Issues for the Empirical Investigation of Multimodality"* (see note 27)*;* Jewitt, *The Routledge Handbook of Multimodal Analysis* (see note 1); K. L. O'Halloran et al., "Multimodal Digital Semiotics: The Interaction of Language with Other Resources", *Text and Talk: Special Edition for Michael Halliday (edited by Geoff Thompson)* 33, no. 4–5 (2013): 665-690.

[52] Halliday, *Language as Social Semiotic: The Social Interpretation of Language and Meaning* (see note 15).

[53] K. L. O'Halloran, S. Tan, and P. Wignell, "SFL and Multimodal Discourse Analysis", in *The Cambridge Handbook of Systemic Functional Linguistics*, ed. G. Thompson, et al. (Cambridge UK: Cambridge University Press, in press).

[54] M. A. K. Halliday and C. M. I. M. Matthiessen, *Halliday's Introduction to Functional Grammar*, 4th ed, revised by C. M. I. M. Matthiessen (London & New York: Routledge, 2014).

[55] J. R. Martin and D. Rose, *Working with Discourse: Meaning Beyond the Clause*, 2nd ed. (London: Continuum, 2007).

[56] M. O'Toole, *The Language of Displayed Art*, 2nd ed. (London & New York: Routledge, 2011).

[57] For example, see E. Margolis and L. Pauwels, eds., *Sage Handbook of Visual Research Methods* (London: SAGE, 2011).

[58] Zelin, "Picture or It Didn't Happen: A Snapshot of the Islamic State's Official Media Output" (see note 22).

[59] Bateman, *Text and Image: A Critical Introduction to the Visual/Verbal Divide* (see note 2).

[60] Ibid., 206.

[61] Martin and Rose, *Working with Discourse: Meaning Beyond the Clause* (see note 55*)*.

[62] J. R. Martin and P. R. R. White, *The Language of Evaluation: Appraisal in English* (London & New York: Palgrave Macmillan, 2005).

[63] http://www.clarionproject.org/news/islamic-state-isis-isil-propaganda-magazine-dabiq

[64] https://developers.google.com/custom-search/

[65] https://dev.twitter.com/overview/api

[66] https://www.instagram.com/developer/

[67] http://scrapy.org

[68] https://www.crummy.com/software/BeautifulSoup/

[69] http://www.clarionproject.org/news/islamic-state-isis-isil-propaganda-magazine-dabiq

[70] http://www.unixuser.org/~euske/python/pdfminer/

[71] http://www.tweepy.org

[72] https://github.com/NikolaiT/GoogleScraper

[73] Bartlett and Reynolds, *The State of the Art 2015: A Literature Review of Social Media Intelligence Capabilities for Counterterrorism* (see note 30).

[74] http://multimodal-analysis.com/products/multimodal-analysis-image/software/

[75] K. L. O'Halloran, M. K. L. E, and S. Tan, "Multimodal Semiosis and Semiotics" in J. Webster (ed.) *The Bloomsbury Companion to M.A.K. Halliday* (London & New York: Bloomsbury, 2015); O'Halloran, Tan, and Wignell, "SFL and Multimodal Discourse Analysis" (see note 53).

[76] O'Halloran, Chua, and Podlasov, "The Role of Images in Social Media Analytics: A Multimodal Digital Humanities Approach" (see note 28).

[77] See also O'Halloran et al., "A Digital Mixed Methods Research Design: Integrating Multimodal Analysis with Data Mining and Information Visualization for Big Data Analytics" (see note 3).

[78] http://wiki.dbpedia.org/

[79] J. Lehman, R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P. N. Mendes, S. Hellmann, M. Morsey, P. van Kleef, S. Auer, & C. Bizer, C. "DBpedia – A Large-Scale, Multilingual Knowledge Base Extracted from Wikipedia", *Semantic Web, 6*(2) (2015):167-195.

[80] M. Szczuka, M., A. Janusz, & K. Herba, "Semantic Clustering of Scientific Articles with Use of DBpedia Knowledge Base" in R. Bembenik, L. Skonieczny, H. Rybinski & M. Niezgodka (eds.), *Intelligent Tools for Building a Scientific Information Platform* (Berlin Heidelberg: Springer-Verlag, 2012): 61-76.

[81] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis", *Foundations and Trends in Information Retrieval* 2, no. 1-2 (2008): 1-135.

[82] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation " *Journal of Machine Learning Research* 3(2003): 993-1022.

[83] H. M. Wallach, "Topic Modeling: Beyond Bag-of-Words" in *Proceedings of the 23rd International Conference on Machine Learning*, (New York: ACM, 2006), 977–984.

[84] F. Feng, X. Wang, and R. Li, "Cross-Modal Retrieval with Correspondence Autoencoder" in *Proceedings of the 22nd ACM International Conference on Multimedia* (New York: ACM, 2014): 7-16.

[85] T. Munzner, *Visualization Analysis & Design* (Boca Raton, London and New York: CRC Press, 2014).

[86] A. Vande Moere, M. Tomitsch, C. Wimmer, C. Boesch and T. Grechenig, "Evaluating the Effect of Style in Information Visualization", *IEEE Transactions on Visualization and Computer Graphics December 2012 (Proceedings IEEE InfoVis)* 18, no. 12 (2012): 2739-2748.

[87] O'Halloran et al., "A Digital Mixed Methods Research Design: Integrating Multimodal Analysis with Data Mining and Information Visualization for Big Data Analytics" (see note 3).

[88] R. Conte, N. Gilbert, G. Bonelli, C. Cioffi-Revilla, G. Deffuant, J. Kertesz, V. Loreto et al., "Manifesto of Computational Social Science", *European Physical Journal Special Topics* 214 (2012): 325-346; D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis et al., "Computational Social Science", *Science* 323 (2009): 721-723; T. Preis, H. S. Moat, S. R. Bishop, P. Treleaven & H. E. Stanley, "Quantifying the Digital Traces of Hurricane Sandy on Flickr", *Scientific Reports* 3 (2013): Article no. 3141, DOI: 10.1038/srep03141.

[89] L. Byrne, D. Angus, D & J. Wiles, "Acquired Codes of Meaning in Data Visualization and Infographics: Beyond Perceptual Primitives", *IEEE Transactions On Visualization and Computer Graphics, 22*(1), (2016): 509-518.

[90] M. Craglia, K. de Bie, D. Jackson, M. Pesaresi, G. Remetey-Fulopp, C. Wang, A. Annoni, et al., "Digital Earth 2020: Towards the Vision for the Next Decade", *International Journal of Digital Earth* 5, no. 1 (2012): 4-21.

[91] Conte et al., "Manifesto of Computational Social Science", 342 (see note 88).

[92] Ibid., 341.

[93] Ibid., 341.

[94] C. P. Snow, *The Two Cultures and the Scientific Revolution* (Cambridge: Cambridge University Press, 1961).

[95] P. Raghavan, "It's Time to Scale the Science in the Social Sciences", *Big Data & Society* 1, no. 4 (2014): 1. DOI: 10.1177/2053951714532240.

[96] Ibid., 4.

[97] Lazer et al., "Computational Social Science", 723 (see note 88).

[98] Key Terms – *Big data analytics*: the study of large amounts of data to uncover hidden patterns, correlations and other insights; *Data mining*: Computational processes involving methods of artificial intelligence, machine learning, statistics and database systems to discern patterns in large data sets; *Information visualisation:* The study of interactive visual representations of abstract data to enhance human understanding and cognition; *Multimodal discourse analysis*: Study of the meaning arising from the integration of multiple semiotic resources (e.g. language, image, gesture and sound resources) in texts, interactions and events; *Semiotic resource*: Sign systems (e.g. language, images, gesture, sound resources) which are used to create meaning contextually in texts, interactions and events; *Social semiotics:* The branch of semiotics which focuses on the study of human sign systems and processes, conceptualised as interacting systems of meaning which constitute society and culture; *Systemic Functional Theory*: the approach developed by Michael A. K. Halliday that conceptualises language, images and other resources as systems of meaning which have evolved to fulfill different functions in society and culture; *Text and image relations*: the meanings arising from the interaction of linguistic text and images.