

Establishing the Complexity of the Islamic State's Visual Propaganda

Vít Střítecký and Petr Špelda

Security analysts have not systematically studied visual discourses, even if they apparently play a prominent role in current propaganda efforts. The article intends to address this disciplinary insufficiency by introducing an inter-scientific approach to analysing large visual data samples. The article illustrates the method by applying it to the dataset comprised of the IS's visual stills. To achieve this goal, the article will first introduce an archive compiled by utilising the knowledge of IS's content dissemination strategies. Second, the article addresses narrative techniques used to effectively convey the message of an alternative worldview. Finally, the text introduces a computational method based on probabilistic topic modelling. Reflecting the results of its application, the article argues that with the growing complexity of the resulting topic sets there is an increase in the probability that the propagandist effort in question represents an attempt at a systematic articulation of a holistic socio-political order alternative to the status quo.

Keywords: Islamic State, probabilistic topic modelling, visual propaganda, revisionism, social media

This article argues that the Islamic State's (IS) visual propaganda should be perceived as an attempt to introduce a holistic socio-political order serving as an alternative to liberal worldviews. The audiences of Western media periodically seized by the virulent dreadfulness of decapitation scenes have a natural tendency to perceive the jihadist propaganda as pointless violence potentially arousing fears. This distorted



Vít Střítecký, Petr Špelda, Establishing the Complexity of the Islamic State's Visual Propaganda, *Central European Journal of International and Security Studies* 11, no. 4: 66-84.

© 2017 CEJISS. Article is distributed under Open Access licence: Attribution - NonCommercial 3.0 Unported (cc by-nc 3.0).

perspective readily summons a cycle not unlike the one discovered by August Ferdinand Möbius¹. The trajectory of its perpetual circles passing first through a repugnant provocation reaches the point of its departure by retaliatory strikes, only to begin a new round of a possibly endless iteration. Devised by its perpetrators as a ceaseless cause of socio-political ripples, it strives to embroil liberal democratic societies in an apocalyptic standoff resolved only through an annihilation of the other.

*Vít Stržítecký
Petr Špelda*

However, there is a coherent deeper level underpinning the grim realities that portrays the diverse schemes of vile atrocities perpetrated by the self-styled *is* or various al-Qaeda affiliates. The following analysis will show that the *is* complex propagandistic endeavour has employed various strategies to build alternative socio-political orders. Using social media imagery, *is* attempts to depict and consequently establish a utopic vision that treats the reality as an obstacle that needs to be levelled in order to yield a space for a new socio-political order. The article further claims that to fully grasp and understand these complex phenomena, machine-generated analysis is required to overcome the limits of human perception.

Empirically, the issue of *is* visual production has been tackled by several authors providing human interpretation of limited sets of visuals² or large sets of visuals that seem to surmount individual comprehension³. Additionally, there have been attempts to produce blueprints describing computational architectures assisting in the analysis of extremist content online.⁴ While recognizing this scholarship, we intend to overcome the limited epistemological capacity of these approaches producing idiographic studies with a constrained empirical reach. In doing so, we seek to contribute to the rising debates introducing computer-assisted qualitative analysis methodologies applied to processing big data within social science⁵. Thus, the aim of our endeavour is to develop a computational approach that would attempt to address the current methodological deficiencies of the analytical apparatus employed to process visual phenomena.

The article will be divided into the following parts. First, we will briefly address the foundations of the *is* strategy introducing enticing utopian visions. After this, we will elaborate on the *is* dissemination strategies employed to spread the visual content followed by the explanation of our approach to capturing the stills and forming the unique archive of *is* visuals. The next section will outline our methodology

CEJISS
4/2017

Source:
Author's
archive

combining a content-oriented computer assisted method based on probabilistic topic modelling with a reflexive analysis used to interpret the machine generated results. An intuitive understanding of the dataset would imply that at the holistic level, it represents an attempt to articulate a virtual alternative of the social, political, and economic order. However, given the volume of the dataset this hypothesis seems highly conjectural. Any of the anthropocentric epistemological approaches cannot settle this issue in a satisfying manner, for the puzzle lies beyond the reach of anthropic cognitive capabilities. Therefore, the phenomenon itself is directly unobservable and requires a methodological treatment which is not founded upon reflexive interpretation of the empirical input. The method proposed in the present text attempts to address this issue as follows. If the input dataset exhibits the level of complexity pertaining to a systematic attempt at articulating an alternative order, the method will output topics diverging with respect to



the low-level visual motives, however, converging on the conceptual level, that is regarding the patterns of socio-political activities. Empirically, we believe that one can judge the character of the propagandist effort according to the level of complexity of the product it generates. In other words, an increasing level of complexity increases the probability that the propaganda conveys an alternative conception of socio-political order. To better learn the difference, refer to the results of the analysis provided in Appendix I.

*Islamic
State's
Visual
Propaganda*

Propaganda as Socio-Political Alternative

When prescribing an order of socio-political conduct, most of the pre-15 jihadi groups could display only a bleak historical 'wreck record'. Despite the existence of a methodological groundwork established in Abu Bakr Naji's 2004 treatise *The Management of Savagery: The Most Critical Stage Through Which the Umma Will Pass*, inability to carve out an order from self-induced chaos prevailed. Perhaps the first indication of the upcoming change in the strategic conduct and propaganda undertakings was the inception of Ansar al-Sharia groups in Tunisia and more importantly Libya in 2011⁶. These groups, albeit inherently violent as their ideological forbearers, exhibited instrumental interest in socio-political conduct anchored in a purist interpretation of Sharia that they would impose in the areas suffering from a power vacuum. The mature version of this disposition was firmly embedded into the mechanisms guiding the daily routine of the self-styled 15. It would be mistaken to portray the group as endowed with a considerably higher potential to foster and maintain a socio-political order since its foothold is greatly self-exaggerated. The visual propaganda inflating the capacity to build a semblance of an ideal society is, however, painting precisely such deceptive picture. That is a picture of seamless utopia brimming with abundance of everything for anybody who answers the call and comes to the lands of the Caliphate. The virtual precursor is continuously filled with scenes of market stalls overflowing with food and shelves bending under the weight of the Hisbah-sanctioned consumer goods intertwined with a selective imagery of frenetic improvement and construction activities elevating the public spaces closer to the utterly surreal. The diorama foreground is a stage of repetitive charitable events scripted in the rigid fashion of a recurrent totalitar-

ian bliss. All of this unfolds against a poetic backdrop of rustic landscape imagery subtly focused on its innate pastoral traits.

If we could abstract away all the cases of recruits drawn to the ranks of the self-styled is due to their pathological fascination with violence, a considerable number of those that remain could be seen succumbing to the alluring effects of the group's visual propaganda. If one transmits an incessant social-media staccato of grandiose utopia imagery, some will come and help build the next more advanced virtual and (more importantly) physical layers and the cycle of exponential materialization begins. These strains of visual propaganda, although in a disparate phase of rudimentary, are now converging across the diverse terrain of Sunni jihadism clearly in a bid to replicate the success of the self-styled is. The case of the failed is Caliphate well demonstrated the effects of reality-suppressing visual propaganda flooding social media. This in an attempt to paint a deceptive picture of sustainable utopias capable of providing the potential recruits and their families with the righteous life-style the liberal democratic societies allegedly cannot offer.

Tracing the IS Visual Content

As outlined above, the ascent of the is prompted advancement in narrative efforts designed to usher in attempts at establishing alternative forms of social, political and economic organization. The is methodical approach to the daily portrayal of societal routine disseminated via the crowdsourced conduits of social networks exhibits a potential to create a deeply engaging experience. An ideologically receptive individual was met by an incessant staccato of world-building propaganda that sustains the feedback loop of cognitive estrangement. Indeed, this phenomenon is particularly noticeable in the case of the is's visual propaganda. For the present work, the main analytical interest lies in its subset comprised of images purporting to depict the daily life in sustainable and ever-expanding islands of the just societal order predicated upon a rigid interpretation of Islamic scripture.

The setting presented a unique analytical opportunity because it exhibited promising potential to yield a dataset of substantial size. The dataset would then be an ideal testbed for developing an empirical argument supporting the claim that the visual production represents a truly comprehensive rendering of an alternative socio-political order.



Source:
Author's
archive

The prerequisite for an analytical undertaking of this kind is continuous compilation of an offline archive of sufficient volume that would allow performing computer-assisted analysis. The archive itself is also

crucial for the specificities of the lifecycle of extremist content online. Even though it is posted and further disseminated in a way designed to attain high degrees of redundancy, long-term content retention does not reach the sufficient levels. When piggybacking on 3rd party publication platforms, the data are actively removed, and attempts at deploying self or friendly operated platforms proved to be similarly unsustainable.

The necessary software toolset needed for automatic archive compilation comprised of the image sets described above consists of a link detector and robot mirroring the discovered content for the subsequent analysis. The rate at which the data are downloaded successfully is partially determined by the delay between the content inception and the detection/mirroring countermeasures that 3rd party publication platforms frequently deploy to roll back the volume of uploaded extremist content. The speed was of the essence even in the cases of now largely defunct self-operated platforms, since they flickered under the pressure of hacktivist attacks or shutdowns by service providers. Any kind of deferred batch downloads thus yields an incomplete data archive.

Being aware of these constraints, shortly before the proclamation of the IS Caliphate in June 2014 we implemented and deployed a system based on the design described in the previous paragraph. The resulting computer assisted analysis described further below is based on the image dataset that was automatically compiled using this system. We could distinguish three consecutive shifts in the IS image distribution patterns, each signifying different stages in the attempted level of the group's virtual presence.

After the proclamation of the IS's Caliphate, the distribution of propaganda images stabilized at the first of the mentioned phases. The base unit of image propaganda assumed the format of an image set originating in one of the newly established *wilayats* (provinces) depicting a particular strain of civil or military activity. These sets, accompanied by the group's branding graphics, were then uploaded to a content sharing site (publication platform) ready for dissemination. Throughout the first phase, the group's platform of choice was the marginal justpaste.it. The site, although technologically trivial in providing a simple WYSIWYG editor ability to directly upload multimedia and generate sharable URLs, became to a considerable extent a standard, later repeatedly copied to attain similar functionality. The method

of sharing small sets of thematically convergent images was itself later imitated by a wide variety of other Jihadi groups.

The generated URLs leading to the prepared content were then disseminated via the group's Twitter social graph. In this case, what was needed for the successful compilation of the data archive was a scan of Twitter (via the Search/Streaming API) for the links to the content that needed to be mirrored. This, for the reason described above, was best accomplished in a fully automatic manner. However, it became apparent that the tenuousness of virtual presence was not limited to the group's Twitter social graph, due to the fact that the shared content was indeed affected in a similar way. The difference, however, is that unlike the social graph (which was repeatedly regenerated, although never to the original magnitude), the content when removed was not generally uploaded again, especially in the cases of image sets that are the subject of the present analysis. This raised a significant issue because the persuasiveness of sustainable alternative order is largely predicated upon the depictions of daily routine contained in now actively suppressed image sets. The solution that emerged involved introduction of greater redundancy levels which marked the beginning of the second phase.

At this stage, two different approaches were employed to attain sufficient levels of content redundancy. First, self and friendly operated platforms were deployed shadowing the core functionality of justpaste.it to bypass its measures against extremist content. The image sets were usually cross-posted to assure their survival at least in one copy. The second approach comprised of scaling of the piggybacking enterprise to include more 3rd party platforms to the same ends as in the case of imitators. Both approaches turned out to be flawed and thus unsustainable for the following reasons.

The deployment of the first justpaste.it epigone manbar.me occurred in November 2014 and was followed by several others throughout the rest of 2015, with the pace of epigone deployment slowing towards the last third of the year. The complete list of justpaste.it epigones included in the resulting dataset is located in Appendix II. However optimistic such a strategy of achieving content redundancy may seem, the lifespan of the epigones proved to be limited. Usually based on the popular WordPress content management system, a combination of lack of skill on the administering side and of various attacks or service providers' cancellations tended to make the career of an epigone,

*Islamic
State's
Visual
Propaganda*

CEJISS
4/2017

and of the data it carried, a short one. Despite the fact that number of epigones made a successful recovery, their existence remained mostly tenuous. Several months of operational history of the first generation became weeks in the second and days in the cases of the last, most amateurish, attempts. Some epigones were even proxied by Cloud-Flare (a company providing services helping to mitigate wide variety of cyber-attacks), which prolonged their survival but ultimately did not help in escaping their inevitable end.



Source:
Author's
archive

The strategy to scale up the dissemination using a broader set of 3rd party publication platforms analogous to justpaste.it ended in a similar fashion. The trend became observable in the first third of 2015 with the use of dump.to and lasted throughout 2016. The complete list of used/exploited 3rd party publication platforms included in the resulting dataset is also located in Appendix II. The failure of the piggybacking scheme had two possible causes. The first one is tied to the shutdown of a platform by its proprietor who realized the site plays an active role in the is content distribution enterprise. The second cause

of failure was the pressure of attacks that various hacktivist groups inflict upon the sites identified, sometimes incorrectly, as hosters of the is's content. The last standing website was shortwiki.org, which played a role in the dissemination of daily digests (see below). The failure of both strategies caused a fallback to the first phase methodology, which was then altered in several aspects. The inability to secure uninterrupted operation of friendly platforms and bleak prospects of exploited sites drove the is image content almost exclusively back to justpaste.it. The effort to attain higher levels of content redundancy was however not abandoned. Since actively chased and removed, the image sets were now uploaded to justpaste.it somewhat desperately in a multitude of copies. Putting the likelihood of such scheme to attain higher levels of redundancy aside, it exposed the fact that the current is Twitter social graph was unable to disseminate all the links to the uploaded copies. The only innovation of this iteration introduced in April of 2016 was the explicit trigger of content capture by Internet Archive (archive.org). Although the site is generally hostile towards extremist content, the mirrored image sets are intact.

Finally, a last remark about the image sets needs to be made regarding the phenomenon of the daily digest. Appearing in the second phase, it first introduced, among other information, a pictorial summary of



is activities published every twenty-four hours. Interestingly, its daily release time, despite some shifts, followed a consistent schedule. Originally, it contained only banners and links to the image sets, and with the ensuing crisis of the third phase whole image sets were included to further the quest for redundancy and the desired content endurance. Over time it became an information hub, with its pictorial core surrounded by links to videos, proselytization pamphlets, nasheeds, Islamic calendars, press releases, magazines, the is operated radios, links to the official is accounts on various communication platforms (such as Telegram), to name just a few examples. Overall “The Daily Prophet” represented perhaps the most advanced method of content syndication the is was able to devise. It successfully raised the redundancy levels and at the same simplified the access to the content itself. Nevertheless, it still depended on 3rd party publication platforms and content sharing sites that lay outside of the group’s control, thus failing to mitigate the original risk of content removal. Presumably realizing the unsatisfactory condition and intrinsic value of digests, in May of 2016 the group begun to package monthly archives of “The Daily Prophets” and share them by various means.

Reflecting upon the knowledge acquired throughout the process of the dataset compilation, although limited to the image sets, an assessment of the technological side of the group’s often recognized media operation does not render a competent picture. After nearly two years of constant attempts to develop a reliable sharing method that would maintain an easily accessible and comprehensive image archive, the group reached the point of departure, losing most of the publicly shared content. Despite the currently available technological means and the role of visual content in inducing the state of cognitive estrangement, is failed to solve the foundational issue of durable content distribution.

Methodology: Bayesian Mixture Models Applied towards Topic Modelling

The methodological apparatus of the traditional social science struggles with the vastness and dynamics of the empirical domain. Computer-assisted approaches generally help to process humanly ungraspable empirical domains, thus effectively scaling up individual research capability. Our intervention introduces stochastic-based (non-determin-

istic) methodology that has been largely dismissed by traditional social science. We tend to believe that this direction constitutes a potentially rich space for convergence of two methods – probabilistic modelling and reflexive analysis. In a nutshell, the contribution lies in the machine identification of the topics that do not necessarily correspond to the topics a human would identify, therefore better representing the Big Data virtual alternatives of the actual ones.

Source:
Author's
archive

The input dataset, whose formation was described above, amounts to a complete archive of the is's propaganda imagery we collected





Source:
Author's
archive

over the period between 16th August 2014 and 12th July 2015. The total number of pictures before deduplication reached 134 845. After we performed a simple checksum-based deduplication the total amount of images decreased to the final figure of 78 977. Despite this measure, a certain level of duplication remained as expected since not all the duplicates were bit-exact copies.

To address the complexity of the 1s’s visual propaganda we designed and implemented an exploratory tool able to deliver rapid insights into the vast, humanly ungraspable visual data domains. The following lines will describe the architecture of the successfully functioning model. The first pre-processing stage loosely adheres to the ‘bag of visual words’ model originally conceived as an ensemble of successive methods for image classification tasks. Image key points are extracted and descriptors subsequently computed using the SIFT (Scale-Invari-

ant Feature Transform) algorithm⁷. A codebook (vocabulary of visual words) is formed via a standard clustering method k -means⁸ executed over descriptors acquired from a suitable data subset. In the aforementioned experiment, the codebook (vocabulary) dataset covered a 14-day period around the first anniversary of the 1st Caliphate proclamation (25 June 2015 – 8 July 2015) supplemented with a small number of what I term “seasonal extras” – even to the lands of the Caliphate, spring comes only once a year. The composition of the codebook (vocabulary) dataset reflects the result of year-long sedimentation processes that have accrued the shape of the current 1st socio-political practice. In total numbers, the 5 000 visual-word vocabulary (codebook) was based on the dataset of 348 images, of which 10 % was allocated to the aforementioned seasonal extras. With the vocabulary prepared, SIFT descriptors were computed for the images of the main dataset; these were then matched, using L^2 norm, with the closest vocabulary word, thereby iteratively transforming the main dataset to its vocabulary representation.

*Vít Stržítecký
Petr Špelda*

After the pre-processing, the method proceeded with utilisation of a Bayesian mixed membership model. The goal of the mixed membership models is to describe high-dimensional multivariate datasets in terms of concise sets of latent (directly unobservable) variables. This means that there is an underlying model whose parameters describe the posterior (conditional) probability distribution over the latent variables in terms of the observed data. In other words, it expresses a probabilistic relation between fragmentary empirical observations and the unobservable structure that generates them. The process can be described as an advanced form of clustering assigning observational units to the latent structure. The difference between the traditional clustering methods and the mixed membership models lies in the assumption that controls the assignment process. The mixed membership models operate on the assumption that instead of a single affiliation, an observational unit belongs to, or in other words is generated by, a combination of several latent variables⁹. This one-to-many relation between an observational unit and the latent structure is given as a vector of continuous non-negative latent variables that add up to 1¹⁰.

The values measure a proportion which an individual latent variable assumes in an observational unit. This leads to the conclusion that each observational unit can be represented as a Bayesian mixture of the latent variables whose values are approximated by the model.

Then, even if there are two observational units consisting of identical latent variables, they still might be, and probably are, different for the proportions given by the values of the latent variables. This amounts to perhaps the most non-reductive approach towards knowledge representation. The latent components describing or generating the high-dimensional observational units to a certain degree repeat; their proportions structuring the individual mixtures differ, however.

Recently, the mixed membership models have been successfully implemented in several areas of applied statistics including the analysis of document collections, network analysis, analytical applications in social and health sciences or population genetics; the most innovative result are currently being generated in the confines of the mixed membership models of text¹¹. This area has been dominated by derivatives, optimizations and improvements of Blei's et al.¹² highly influential work on probabilistic topic modelling, more precisely by latent Dirichlet allocation (LDA).

LDA presumes that the observable phenomena, in this case textual documents, exhibit varying mixtures or proportions of unobservable topics¹³. The topics pertaining to the analysed document set are conceived of as distributions over a fixed vocabulary (unique words found in the set) and as such represent the latent structure¹⁴. LDA itself describes the probabilistic generative process responsible for producing both the latent topic structure and the content (words) of the observable phenomena (documents). The process is given as follows: '[...] for each document in the collection, I generate the words in a two-stage process, [i] randomly choose a distribution over topics, [ii] for each word in the document (a) randomly choose a topic from the distribution over topics in step [i], (b) randomly choose a word from the corresponding distribution over the vocabulary' ¹⁵. If reversed, the generative process is a case of Bayesian inference which, holding the document set constant, attempts to answer the question of 'What is the likely hidden topical structure that generated my observed documents?'¹⁶.

Conclusion: Reflexive Analysis

Building on the argument outlined in the introductory part of the present article the following section will illustrate the resolution of the puzzle. With respect to the empirical input comprised of the is propa-

ganda picture sets we would argue that it indeed represents a systematic attempt at articulating a holistic social, political, and economic order. Moreover, this statement stems directly from the result of applying a Bayesian mixture model to discern an unobservable generative structure represented as the revealed topics. A conventional way of utilising the results of probabilistic topic modelling, that is of the symptomatic documents pertaining to individual topics, is to interpret simply observable themes. This could be achieved by means of thematic discourse analysis which would code the pictures into predefined discrete categories. As a result, the only effect of this methodology is quantitative augmentation of interpretative capacity.

*Islamic
State's
Visual
Propaganda*

Even though we understand such an analytical move for it is an intuitive and pragmatic one, it however reduces the range of empirical puzzles which can be solved by it. As can be gathered from the introductory section, we have chosen the issue of complexity since it represents a natural complement to thematic-oriented studies of any kind. This would settle one of the most general characteristics regarding the propagandist efforts aimed at status quo revision. We would argue that according to the observed level of complexity, one can judge whether the propagandist efforts describe a systematic alternative to the status quo or merely attempt to erode it without an ambition to redefine it.

To empirically demonstrate what we have outlined above, the following can be stated regarding the picture sets contained in the Appendix. The picture sets below are comprised of the symptomatic images exemplifying individual topics. Picture sets 1-6 display the level of complexity associated with a systematic attempt at articulating an alternative order. This is because among the individual symptomatic images, which constitute the unobservable topics, there is a considerable difference regarding the depicted activity patterns. At the same time, these are, however, convergent with respect to the conventionally conceptualized attempts of arranging human societies according to a meaningful principle. The composition of image sets 1-6 stands in stark contrast to that of image sets 7 and 8. These sets exemplify the topics from the low complexity segment of the input dataset and as such do not exhibit any unifying principle that could be considered generative of a coherent alternative order. If the application of the above described methodology yields a topic set comprised mostly of simple patterns based on the low-level visual correspondence, the input dataset falls into the category of erosive efforts without redefining

ambitions. However, if for the most part the resulting topic set is comprised of the image sets which are similar in their composition to sets 1-6, we would argue that the underlying propagandist effort is indeed a constitutive one. This is because the depicted activity patterns exceed in their complexity the discrete categorisation resulting from visual-based thematic discourse analysis. This result is possible due the applied Bayesian mixture model able to represent the exemplifying sets in terms of proportions of the unobservable structure, i.e. topics. More importantly, underlying of these mixtures are probability distributions which do not reduce the observable phenomena to discrete categories. Hence, they do not commit to unnecessarily distorted accounts of the phenomena predicated upon the limits of the anthropocentric cognition. In practical terms, this is most usually manifested as artificially parsimonious typologies to fit the coding capacity of the researchers. The proportion between simple and complex reported in the Appendix pertains to the whole of the resulting topic set as well. As a result, we would argue that the input dataset, i.e. the IS's visual propaganda, indeed represents a systematic attempt to establish an alternative socio-political order. This result does not stem from a reflexive analysis, but from a machine learning model thus complementing the state-of-the-art disciplinary knowledge of the phenomenon.



VÍT STRÍTECKÝ is an Assistant Professor at the Departments of Security Studies, Charles University in Prague and Metropolitan University Prague. Can be reached at vit.stritecky@fsv.cuni.cz

PETR ŠPELDA is an Assistant Professor at the Department of Security Studies, Metropolitan University Prague. Can be reached at petr.spelda@mup.cz

This article was written within a Specific Academic Research Project of the Institute of Political Studies, Faculty of Social Sciences, Charles University, Prague, no. 260 342/2016, Contemporary Threats to Political Order.

Notes

- 1 This meant as a metaphor stemming from the topology of Möbius strip.
- 2 Aaron Zelin (2015), Picture Or It Didn't Happen: A Snapshot of the Islamic State's Official Media Output, *Perspectives on Terrorism* 9(4), pp. 85-97.
- 3 Charlie Winter (2015), *Documenting the Virtual 'Caliphate'* Quilliam Foundation, Available at: <http://www.quilliamfoundation.org/wp/wp-content/uploads/2015/10/FINAL-documenting-the-virtual-caliphate.pdf>;
Daniel Milton (2016), *Communication Breakdown: Unravelling the Islamic State's Media Efforts*, CTC Sentinel Report. Available at: https://ctc.usma.edu/v2/wp-content/uploads/2016/10/ISmedia_Online.pdf;
Carol K. Winkler and Cori E. Dauber (eds.) *Visual Propaganda and Extremism in The Online Environment*, Carlisle Barracks, PA: United States Army War College, 243 pp.
- 4 Saeid Balkesim (2014), Semantic Processing of Visual Propaganda in The Online Environment, in Carol K. Winkler and Cori E. Dauber (eds.) *Visual Propaganda and Extremism in The Online Environment*, Carlisle Barracks, PA: United States Army War College. pp. 193-214.
- 5 Paul DiMaggio (2015), Adapting computational text analysis to social science (and vice versa), *Big Data & Society* 2(2);
Ted Underwood (2015), The literary uses of high-dimensional space, *Big Data & Society* 2(2).
- 6 Henrik Gråtrud and Vidar Benjamin Skretting (2017), Ansar al-Sharia in Libya: An Enduring Threat, *Perspectives on Terrorism* 11(1), pp. 40-53.
- 7 David G. Lowe (1999), Object Recognition from Local Scale-Invariant Features, *The Proceedings of the Seventh IEEE International Conference on Computer Vision Volume 2*, pp. 1150-1157;
David G. Lowe (2004), Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision* 60(2), pp. 91-110.
- 8 Stuart P. Lloyd (1982). Least squares quantization in PCM, *IEEE Transactions on Information Theory* 28(2), pp. 129-137.
- 9 Edoardo M. Airoldi and David M. Blei and Elena A. Erosheva and Stephen E. Fienberg (2015), Introduction to Mixed Membership Models and Methods, in Edoardo M. Airoldi and David M. Blei and Elena A. Erosheva and Stephen E. Fienberg (eds.) *Handbook of Mixed Membership Models and Their Applications*, Boca Raton, FL: CRC Press, Taylor & Francis Group, pp. 618; Airoldi et al. (2015), pp. 3-4.
- 10 Ibid.
- 11 Airoldi et al. (2015), pp. 6-8.
- 12 David M. Blei and Andrew Y. Ng and Michael I. Jordan (2003) Latent Dirichlet Allocation., *Journal of Machine Learning Research* 3, pp. 993-1022.
- 13 David M. Blei (2012a), Probabilistic Topic Models, *Communications of The*

Vít Strátecký
Petr Špelda

ACM 55(4), pp. 77-84;

Blei (2012a), p. 78.

14 Blei (2012a), pp. 78-79.

15 Blei (2012a), p. 78.

16 David M. Blei (2012b), Topic Modelling and Digital Humanities, *Journal of Digital Humanities* 2(1).

CEJISS

4/2017