Zrec.org – psychosocial phenomena studies in cyberspace

Martin Pavlíček, Tomáš Filip, and Petr Sosík

Silesian University in Opava - Institute of Computer Science
martin.pavlicek@protonmail.com

Abstract: We present a research project based on the use of bio-inspired computing methods which are applied to analyze psychosocial phenomena (group polarization, belief echo chamber and confirmation bias) and patterns in occurrences of world events and information dissemination in cyberspace. The aim of the project is to integrate infrastructure, tools, methods, data, AI researchers and end users to create a platform that can be used to understand these processes in human society based on social interaction on the surface Internet triggered by exposure to information about a world event. These processes are investigated and understood at the level of social super-systems as well as selected smaller units at the level of determining psycho-social phenomena as an individual's reaction to the world around in a form of received information and exposure dynamics. On the other hand, we focus on world events, their analysis in global scope and ability to predict them and find patterns in occurrence based on information dissemination. The text discusses selected methods of soft computing which can be effectively and prospectively used for data collection, information extraction, aspect-based sentiment analysis, monitoring phenomena of social interactions and world events occurrence and generating parametrized content for simulated and a real-life infrastructure. As a practical example, we include an analysis of data obtained from 689 days collection from sources targeting Czech population.

1 Introduction

1.1 Interactions

Social interactions between individuals and groups are increasingly moving into cyberspace [7]. Social networks, discussion forums and chats are a virtual space where interpersonal (and institutional) communication takes place and where these conversations affect individuals with each other. In addition, the effects of digital communication are multiplied by the hormonal responses of the human body and information overload [22] [30] . Opinions expressed in groups and repeated represent particular narratives and beliefs [2], which are passed on and significantly contribute to the acceptance or rejection of the worldview. News are also moving into cyberspace [17], bringing information about events in real time - social media with their content (whether true or fictional) is often the primary source of information for news agencies.

The shift of social interactions and intelligence to cyberspace highlights the effect that information itself has on social groups and individuals themselves, and the dissemination of information itself becomes subject to regulation and control. We see this effort on the part of providers of platforms [13] [3], political and power groups and governments, which approach information as an information weapon and a tool for psychological operations [5] [6] [4] [10]. At the level of national security, information and its dissemination has the ability to reach the target population outside the borders of the state and continents and thus influence perceptions at the level of economy, politics, religion, security.

1.2 A simple premise

We build our work on a simple premise.

An information about event appears in cyberspace (this information can be true, false or mixed together).

The individual (human, AI bot, intelligence agent, marketing agency, etc.) creator responds to these events in a dedicated way according to their "configuration". Accepts or rejects the event – in its configuration it expresses sentiment, towards set entities.

The behavior of individuals then grows into the behavior of the whole group. An individual who comes from outside then perceives the behavior of the whole group as a unified acceptance or a rejection of the event.

As a result, the group's behavior is used (media, political, security) as an approval (consent), narrative, understanding of the event and unfolding other events (Figure 1).

1.3 Group phenomena

We focus on three phenomena – group polarization, belief echo chamber, and confirmatory bias. We analyze this situation from a perspective of mining certain information from a text interaction (Figure 2). That information can be – a definite source, a concrete word or phrase, sentiment towards entity, or a complex belief.

One of the most striking phenomena is group polarization [1]. A typical example is when one group uses only one information source and the other uses another and the opinions in these sources are opposite and the information and biases towards specified entities can be described as opposite extremes. The sentiment towards particular entities is opposite within the social interaction.

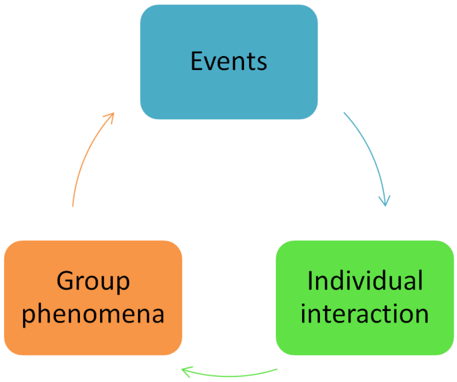

Figure 1: A simple premise of connection between an information about event, individual's interaction and a group behavior

Belief echo chamber [23] [33] is a part of imaginary space where repetitions, or worldview, or information sources are repeated (depending on what we are watching). In general we can understand thoughts, or narrative, sentiment to particular entities. Repetition of the opinion multiplies the effect of this information on the members of the group. And a person in an echo chamber encounters only that information that corresponds with their own.

Confirmatory bias [32] is a phenomenon when an individual responds to information in the opposite direction to the sentiment given in presented information, distorting incoming factual reality to fit preexisting worldview or to selectively cherry pick determined sources and information with corresponding belief.

1.4 Events

In the case of the analysis of world events, we monitor the development of events (conflict or support in a material or symbolic level) based on the past or similar developments at the global or local scale. We also monitor the dynamics of the spread, or conversely, the concealment and promotion of information of a specific event in cyberspace from a selected source. Therefore, we monitor both actual events and information about events. In addition to the analysis of events in the level of symbolic expression and computability, we observe the possibilities of suitable visualization of events and their dynamics and their further processing.

1.5 Motivation

Monitoring and quantitative understanding and interconnection of macro (events) and micro environments (individual and group interaction) within the monitored sources (social networks, surface internet) thus means a realistic understanding of the observed phenomenon in a certain ecosystem.

At the level of knowledge, as a society we are interested in the macro environment – events that have taken place, or we are just interested in mentioning them in cyberspace (within selected information sources) and micro environment – interactions between individuals and their positive expression, or a negative attitude towards individual situations, entities, opinions.

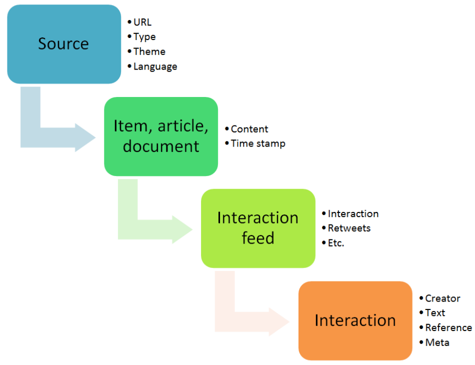

Figure 2: A common scheme of sources, items and interactions

As a result, we want to know the approval or opposition to a particular narrative, a belief in the form of a worldview, and thus a forward reaction to other (similar in pattern) world events, where there is already a positive or negative sentiment towards the actor.

The knowledge gained in this way has a societal character and we can approach this issue at the level of a long-term aim, namely the creation of tools and methods for collection, rather than just a one-time processing of selected data.

The goal of our efforts is to create a tool and platform that participates and can be used for long-term research on these distinct social phenomena at the global level. Furthermore, to connect the research in the field of computer science (artificial intelligence, scalability, parallel computation and visualization) and specific methods and individual researchers and institutions who will focus on phenomena understanding.

2 Main work

2.1 Pipeline

The first step is to collect the data. We are talking about different sources (surface internet discussion comments and social networks) different methods of getting this data (API, web scrapping, headless browsers) and storing it in a suitable temporary storage, that can be used to populate existing datasets and of course updating and collecting new data based on previous collections.

A common scheme (Figure 2), which can be applied both on social networks and surface internet describes a hierarchy of an information source (certain website, users Facebook page, Twitter feed, Reedit page, etc.), which is examined, information item which can be a document, article, tweet, posts. Finally, an interaction feed which is

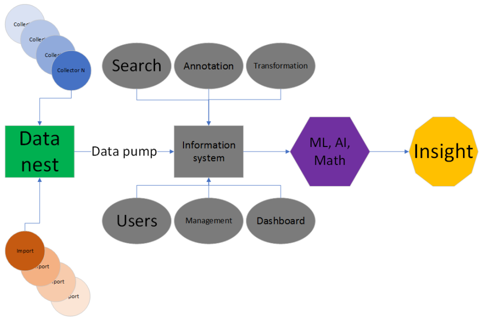

Figure 3: Architecture of the platform

constructed from single interactions presenting as a comment, or determined activity (reaction).

This is followed by data cleanup - such as corrupted, unusable, incomplete, duplicate pieces of data are discarded.

Furthermore, merging or separation from different sources, depending on which sources we want to process the data is present and storing the data in the mail computational grid.

Annotation is the most important step in the entire cycle because we assign some meaning to the text - identifying used references to referenced web resources, entity detection, sentiment analysis, or even detect ideas with handcrafted ontology or text summarization. We understand the annotation as a dynamic parametrized process that can be run with multiple parameters on specified data and thus it can be used as a benchmark within selected algorithms, NLP models. Over time, we can, for example, annotate image, video, sound and thus further the information collection to a whole new spectrum of media formats.

The following step is the data transformation for the model's calculation options – encoding semantic to symbols. We can also model a behavior of a super system as a simulation of profiled multi agent group.

The last step is to get the results and then apply the feedback to step up the accuracy of our models, store it and share it within the research group.

The whole pipeline architecture is very simplified and represents only one computational node. In the case of multi-institutional collection we add more complexity and managed layer of distribution of technology and data sharing.

2.2 Architecture

From the point of view of technical architecture (Figure 3) of a self-standing system we can talk about three units.

The first unit is a part of the system that is used for collecting and storing data from various sources. We collect data both through collectors that interactively retrieve data from web sources, and through imports. Each definite source has a group of imports and collectors which can be used to access it. Data about events can be accessed through annotation of information items or through event databases such as GDELT[1].

The second unit introduces a system for annotating and working with data both in the form of running tasks and manual annotation. It is about implementing a web interface where we can work with data that is pumped into the system. In this part of the system we can also create ontology describing complex ideas manually.

The third unit introduces the models of bio-inspired AI and their application in data processing. So we are talking about the model with the aim of classifying, detecting, predicting and generating content.

From the point of view of using technologies we choose several physical servers, GPU calculations, we consider the possibilities of deploying virtualization because of limitations and ensuring the operation for particular units. Our work is mainly done on a commodity based hardware with no use to special computation resources.

We choose relational and document databases, and tools for big data processing for the raw data storage. Artificial intelligence is solved using popular frameworks such as Tensorlow, Keras, PyThorch, as well as Python libraries NumPy, SciPy, Pandas, SciKit. After a model is fully tested we proceed to low level language implementation.

2.3 Bio-inspired methods of computing

Information extraction (IE) The goal of IE is to transform the input unstructured text into an output structured form. This process could be divided into two parts. First, extract named entities. Subsequently, classify existing relations between entities. The most common methods for IE solve both tasks separately.

Named entity recognition (NER) Linear-chain Conditional Random Fields (CRF) [18] is one of the classic methods for NER. Today, neural network is one of the most effective methods for NER. Various models based on Long Short-Term Memory (LSTM) have been tested [12]. Others combined LSTM with CNN [8] or CRF [16]. Many of the most advanced models use pre-trained language models as the BERT (Bidirectional Encoder Representations from Transformers) [9]. Further improvement was achieved using context representation [24].

Nested named entity recognition(NNER) One of the biggest challenges in NER is the extraction of nested named entities, for example the classification of the "Federal Bureau of Investigation" as single entity denoting an organization. Straková et al. [34] introduced method to solve NNER as sequence-to-sequence problem, where the input sequence contains tokens and output sequence labels. Another approach was used by Li et al. [21], which is based on creating queries and subsequent classification

[1] https://www.gdeltproject.org/

of relations between entities. The state-of-the-art model BERT-MRC [20] can extract both flat and nested entities. This model achieved the best results on ACE04, ACE5 and Genia datasets.

Relation extraction For a long time, models classifying relations between entities had problems with classifying long-distance relations, single sentences with multiple relations and overlapped relations on entity-pairs. The improvement was achieved by using models based on pre-trained Language Model. These complex models have been fine-tuned for Relation Extraction. Cheng Li [19] introduced a method that creates a model with pre-trained PLM parameters and achieve great success on SemEval, NYT, WebNLG datasets.

To meet our goals, it is essential to create methods capable of extracting entities and classifying relationships between entities. Our data contain searched nested entities. For this reason, our target model must handle these entities. We want to test a new approach combining NER and RE into one single model proposed by Bowel Yu et al. [38] and compare this model with the classical approach of two separate models for NER and RE. For these models we will use the parameters from pre-trained BERT. The resulting best approach will be applied to the dataset collected from social networks. We want to analyze which entities appear in the posts and what relations are between them. This is essential for further global social analysis. IE provides a tool for searching and sorting posts by contained entities.

Aspect based sentiment analysis (ABSA) ABSA is an essential method for understanding the content of a text. The aim of sentimental analysis is to determine opinion and emotions from the text. Compared to Sentiment Analysis, ABSA allows finer-grained determination of polarities with respect to individual aspects. Main task can be divided into two basic subtasks: (1) the detection of aspects in the source text; (2) the classification of polarity [25] [29].

Various methods have been used to solve the ABSA task. Portia et al. (2006) used the CNN architecture [26]. Jebbara et al. (2017) created stacked RNN and CNN [15]. Several high-success models have been developed for ABSA in recent years. Models based on transformer architectures such as OpenAI GPT2 were trained on various language tasks [27]. Significant success has been achieved with the BERT model using self supervised pre-training [9]. Raffel et al. (2019) demonstrates effectiveness of transformer architecture on various language tasks and achieved the highest accuracy on binary classification Sentiment Analysis dataset SST-2 [28]. For ABSA task, the state-of-the-art model LCF-ATEPC achieved high efficiency on SemEval-2014 Task 4 dataset [37]. A different approach to solve ABSA, that we would like to try, proposed Chi Sun et al. [35]. They transform ABSA task to a sentence-pair classification task.

For a deeper understanding of the input sentences, it is necessary to apply a contextualized word embeddings representation. ELMo model presented by Peters et al. [24] used LSTM layers to create deep contextualized word representation. A deeper understanding of the sentence context has been achieved with BERT model, which uses both left and right contexts. In our research, we create a model based on multi-head self-attention along with BERT architecture. The input strings will be processed by a sub-word tokenization. Transfer learning will be used for speed up training. We will optimize the architecture and hyperparameters to achieve the highest accuracy. For hyperparameters tuning we would like to test new optimization algorithms based on evolutionary algorithms such as PBT [14]. Once we find optimal architecture and parameters, we will compare it with other models on SemEval-2014 Task 4 dataset. Finally, the trained model will be tested on Twitter posts. With the final model, we will be able to classify whether the post deals with interest topic or entities and how it relates to it. Advanced methods of sentiment analysis will allow us to get an idea of the opinions of various social groups on individual aspects.

Polarization We can identify different opinion groups using cluster analysis and outputs from IE and ABSA. Society and its views evolve over time, so our models have to train online. These groups will be monitored over time depending on world events. Some of the analyzed data will be high-dimensional, thus we will use dimensionality reduction algorithms before further processing. We cannot expect that collected data have Gaussian shape. Therefore, we will use various algorithms applicable for data clustering, including BIRCH and K-means. The amount of collected data will be large so it would be inefficient to use a hierarchical clustering. Different unique opinion groups can be detected with anomaly detection techniques. From the collected data we will be able to determine what characterizes each group and how they interact with other groups. Once we analyze the polarity of groups to certain events and entities, we can analyze how groups differ and what they have in common. If new opinion groups appear during the monitoring, we will be able to detect them using novelty detection techniques. We can use Local Outlier Factor or Isolation Forest algorithms for this purpose. The collection of these processes shows us how world events influenced society views and what social groups think about the given topics.

Content generation Text generation is a method of artificial intelligence that generates natural text. GAN is a popular type of network for generating natural text. One of the models capable of generating coherent and semantically meaningful text is LeakGAN [11]. Shoeybi et al. [31] used model parallelism to create a state-of-the-art model on WikiText-103 dataset. A complex language model as

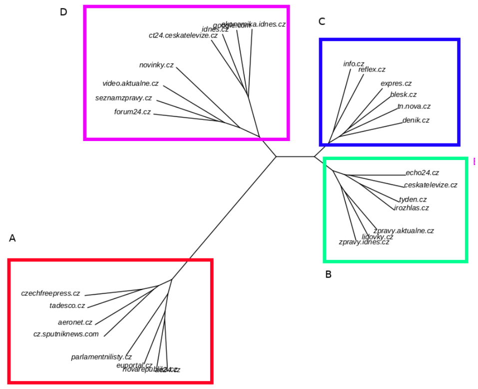

Figure 4: Polarization in use of sources, 689 days, 3,2M interactions

the GPT proved that it is able to generate convincing text. This model was trained to predict new word in sentence on large corpus of text. Once we have categorized posts by opinion groups, we will create a model based on GPT2. Recently released GPT2 pre-trained weights will be used as initial model parameters. We fine-tune model on the collected posts so that it can generate new posts syntactically similar to the posts of the selected opinion group. We will optimize the temperature parameter to get various different posts. For training we will use the same techniques as OpenAI used for GPT training. Twitter posts of specific group will be our training data and we use self-supervised task to predict next word in these posts. Finally, we create a binary classifier model that will classify which posts are real and which are artificially generated. We compare accuracy of this model with human. GAN architecture will also be tested to generate posts, but due to GAN training difficulties, this is not our preferred choice for the final model.

3 A practical example

As a real-life case study, we created a collector based on HTML content extraction and collected interaction from one of the most visited news portal in Czech republic novinky.cz. The collection was specific, because the website does not hold discussions to their news article open endlessly, but access to the discussion is removed in an undisclosed time.

Our collector was working in 24/7 mode and periodically checking for new interactions as well as new published articles with open discussion to appear in a cycle and collect them before they were disabled. From this point of view, we made a prospective study.

The data from novinky.cz (selected source) was collected since 18.9.2017 and the collection was active for 689 days. During the collection we obtained the 3 282 429 interactions from 24 787 anonymized creators who participated in the 54 073 of the observed discussions within news articles (information items). During the 689 days collection we obtained 3745 unique referenced sources. But in the group analysis we dealt only with the most used one.

From collected text (HTML) interactions we removed interactions which were censored, we removed duplicated texts (citation of other creators) and we focused on the use of concrete references to sources in interaction together with sentiment-based analysis towards definite entities in a content analysis. Our aim was to evaluate which sources and how are they used in groups from the point of view of polarization – a theme which was resonating in mainstream media and political discussion due to election season.

We used a hierarchical clustering analysis on top of modified normalized web distance computation [36] based on the premise that user used s specific source (Figure 4).

In this example we removed all sources that could be described as plural media (Youtube, Facebook, Twitter) – with content assumed both anti-systemic and systemic from a view of socio-political consensus in Czech republic and analyzed only sources that could be view as systemic (mainstream media, and corporate media) and anti-systemic (Russian outlets and independent media).

This simple analysis shows a very strong polarization between groups A versus groups B,C,D. We can see the polarization between the group that uses the so-called anti-systemic sources (Figure 4 group A) and the remaining resources (Figure 4 group B, C, D). In the group A we see aeronet.cz, sputniknews.com and others which were described by the mainstream media as media with disinformation and fake news content.

We used a various approaches for a content analysis (with different results) and aspect-based sentiment analysis based on the detection of entities. The approaches are mentioned in this text. In general we can describe the results as following:

In the group A there was a positive sentiment about China, Russia, Czech nationalism, some political figures namely Vladimir Putin and Donald Trump and a negative sentiment towards NATO, EU, Israel and the USA. USA intelligence agencies like CIA and FBI were mentioned with negative sentiment too.

Groups B, C, D had this sentiment expressed in reverse – negative sentiment was mainly towards Russia, China, North Korea. Politicians with negative sentiments were Vladimir Putin and Donald Trump. Positive sentiment was towards the USA, EU.

In this example we clearly see the polarization that is reflected in the choice of source, as well as in the expression of sentiment towards certain entities, corresponding to the narrative that the referenced sources inform about. This example illustrates the reality of the phenomenon. We have accessed specified sources in the analysis and proceed with sample content analysis of information items (articles) and interaction and found out that the sentiment towards entities corresponds with the analysis of the base dataset, so we can see those sources as a real-life example of belief echo chambers disregarding factual reality. The bias was with the same positive and negative tendencies as in analyzed interactions.

Our aim is to optimize used methods and present more robust and thorough analysis of the data-set we created. And to quantitatively evaluate results both in the form of output content and effectiveness and to publish those results both from phenomena level and level of bio-inspired methods and their effectiveness.

4 Discussion

4.1 Importance for society

Creation of an open, accessible and international global tool, which is intended for researchers of real social phenomena, as well as for informatics as a tool of applied research in the areas of AI, data processing, visualization is seen as a synergic effort. Our work combines research, technology, phenomena understanding and individuals who participate in their research.

Real quantitative mapping of phenomena that take place in a dedicated ecosystem and which are to some extent very generalized presents answers to today's world problems in the field of information warfare. Understanding the dynamics of information dissemination between individuals, groups and information sources. Monitoring and analysis of the development of events with regard to the distribution of information.

The ability to create simple profiles of both individuals and entire groups and place them in the context of events that take place in society. Ability to monitor and detect the performance of psychological operations or malignant work with information, both preventively and retrospectively.

The analysis of phenomena itself is only the first step. The second step is the prediction of the behavior of the individual, groups, and the world itself at the level of information about the possible upcoming typified event and the placement of even small phenomena in a massive context. With the ability to proceed with real time infrastructure content generation we can achieve a higher level of social dynamics and their response to particular information.

The storage of events and distinct interactions represents the preservation of the image of a part of the social groups which are present on internet and their reactions in a specific time and regional space. Thanks to the availability of this data, we can better understand current events and their development, both in the context of the past and present, even at the level of monitoring and understanding the deployment of information weapons and the implementation of psychological operations.

4.2 Importance for IT

The main idea is to create an organic global scalable platform, that could adapt to large datasets, easy plug-in model and AI integration and streamline pipeline.

Optimization, implementation and creation of suitable bio inspired computing methods, including already trained models that can be used on specific data sets within the project is a main goal. Modification of models to be suitable for general text corpuses and wider usage is step to a more general way of usage of the optimized outputs. These methods deal with - extraction of information, detection of entities, determination of sentiment, ability to predict, classify and generate parametrized content.

Creating a library of suitable (and interchangeable) bio inspired methods that can be used both for processing group phenomena and for working with global and local events, processing the dynamics of complex networks. Focus on optimization, implementation of methods and their comparability on living infrastructure.

Effective visualization methods that are applicable both for analytical human understanding of selected phenomena and for further machine processing at the input level to other computing subsystems are needed to be methodologically created. Use of unorthodox data representation as a form of transcription of symbols and numbers for further processing in bio inspired systems is also a priority.

The system should be presented both as a research tool and as a benchmarking tool for a plethora of bio inspired computing methods and their combination and particular optimization and fine tuning. Thus the usage of distinct methods, models and their optimization together with computed outputs should be used as an internal benchmark for finding specific insight for developers. Therefore a simple share and platform mechanism is needed to be implemented.

We want to focus on computation, visualization and understanding of complex network behavior and their time dynamics not just static properties of time defined system. Time dynamic and modeling in networks is important for a deeper understanding of selected phenomena and to understanding basic pattern similarities. Thus working on universal complex network computation is promising.

The concept of the system is built on a simplistic reduction of beliefs and thought to a form of ontology based on positive or negative sentiment towards entity. We can describe it as a very simple NLP reduction of text. We see a potential in the system 2 AI to focus more on abstract concepts and reasoning and gaining a higher level of machine understanding of higher and more complex ontology models.

4.3 Shortcomings

We realize that it is not possible to collect data from the entire surface Internet and all of the favorite social networks. The aim is to have the ability to analyze selected subsets from social networks and surface internet to a certain extent (theme, language, sections), and then train the computational models on the collected data as a whole.

The collection is primarily based on monitoring defined sources from the surface Internet. In future we assume a managed way to distribute sources which should be collected within partnered institutions based on location and language. As a result, data sharing on big data platform is a goal but we always assume a subset of potential sources.

As part of academic research, we do not violate the operating conditions of individual terms of services of sources we collect. Users are anonymized within the system and processing, so only textual interaction and available metadata are collected.

A distinct area of work is the creation of an autonomous collector, which receives only the URL of a target source and proceeds with all activities (detection of items, feeds, processing and detection of interactions, detection of users and scheduled updates) related to data collection. In the first phases, we rely on the creation of manual collectors always corresponding to the monitored source and having their own algorithm for data extraction and collection.

Our effort is to share data, infrastructure, training and optimization of models between institutions and creation of a unified multilingual environment for the work of the end users. At the infrastructure level, we see a priority in a simple addition of a computation node for collecting, updating, distributing models and scalable data sharing.

5 Conclusion

In this text we presented the Zrec project (www.zrec.org). The aim of the project is the analysis of psychosocial phenomena (group polarization, belief echo chamber and confirmatory distortion) on the surface internet. These phenomena are analyzed in the context of reactions (positive, negative) to information about local and world events. Our primary sources are social networks, and discussions and comment boards within webpages. Part of the project focuses on analysis and visualization of the dissemination of information about events on the surface Internet.

Žrec was an ancient highest Slavic priest and prophet who could influence even significant political decisions. Observing and understanding beyond the boundary of common communication, global processes and consciousness, that is the goal of our project

The topic of group behavior and individual's reaction to information resonates in the areas of politics, national security, religion, education and economics. In this case, we fulfill the need after quantitatively processing selected phenomena on the open Internet in the form of tool. The motivation is therefore to create a scientific research platform that contains methods, data, researchers in the field of Ai and end users who study the observed phenomena.

The core of the platform is built on a suitable combination of biologicallly inspired computing methods, which take care of the detection of entities, relationships between entities, extraction of information and determination of sentiment. Furthermore, methods that examine their own group behavior at the level of dynamic heterogeneous networks. We see the platform as a benchmarking tool for selected methods, which focus on the same result.

We include a 689 days prospective study which underlines the existence of select phenomena and perspective ability of select methods to proceed with deeper content analysis. We see a plug-in implementation of concurrent methods and their combination and optimization as a best way to achieve reliable quantitative results.

At the IT level, we focus on optimizing and finding new models and methods for working with text, effective computational models in the field of text annotation, complex network dynamics, data architecture scaling, unorthodox data representation and building a dynamic ontology with the ability to add remote computing nodes in the form of newly involved institutions.

Project's development represents involvement of other institutions to proceed with regional data collection. A creation of shared big data space to implement select AI models for content analysis and providing the platform together with collected data as a research and archive tool.

Acknowledgements

This work was supported by the Ministry of Education, Youth and Sports Of the Czech Republic from the National Programme of Sustainability (NPU II) project IT4Innovations Excellence in Science - LQ1602, and by the Silesian University in Opava under the Student Funding Scheme, project SGS/9/2019.

References

[1] C. A. Bail, L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. B. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky, *Exposure to opposing views on social media can increase political polarization*, Proceedings of the National Academy of Sciences, 115 (2018), pp. 9216–9221.

[2] D. Bar-Tal, *Group beliefs: A conception for analyzing group structure, processes, and behavior*, Springer Science & Business Media, 2012.

[3] BBC, *Twitter hides Trump tweet for 'glorifying violence'*, 2020. `https://www.bbc.com/news/technology-52846679`.

[4] *Annual report of the security information service for 2016*, 2017. `https://www.bis.cz/public/site/bis.cz/content/vyrocni-zpravy/en/ar2016en.pdf`.

[5] *Annual report of the security information service for 2017*, 2018. `https://www.bis.cz/public/site/bis.cz/content/vyrocni-zpravy/en/ar2017en.pdf`.

[6] *Annual report of the security information service for 2018*, 2019. `https://www.bis.cz/public/site/bis.cz/content/vyrocni-zpravy/en/ar2018en.pdf`.

[7] V. Blazevic, C. Wiertz, J. Cotte, K. de Ruyter, and D. I. Keeling, *Gosip in cyberspace: Conceptualization and scale development for general online social in-*

teraction propensity, Journal of Interactive Marketing, 28 (2014), pp. 87 – 100.

[8] J. P. CHIU AND E. NICHOLS, *Named entity recognition with bidirectional LSTM-CNNs*, Transactions of the Association for Computational Linguistics, 4 (2016), pp. 357–370.

[9] J. DEVLIN, M.-W. CHANG, K. LEE, AND K. TOUTANOVA, *Bert: Pre-training of deep bidirectional transformers for language understanding*, arXiv preprint 1810.04805, (2018).

[10] *A Europe that protects: The EU steps up action against disinformation*, 2018. `http://europa.eu/rapid/press-release_IP-18-6647_en.htm`.

[11] J. GUO, S. LU, H. CAI, W. ZHANG, Y. YU, AND J. WANG, *Long text generation via adversarial training with leaked information*, arXiv preprint 1709.08624, (2017).

[12] J. HAMMERTON, *Named entity recognition with long short-term memory*, in Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003, 2003, pp. 172–175.

[13] T. HATMAKER, *Youtube bans david duke, richard spencer and other white nationalist accounts*, 2020. https://techcrunch.com/2020/06/29/youtube-ban-stefan-molyneux-david-duke-white-nationalism/.

[14] M. JADERBERG, V. DALIBARD, S. OSINDERO, W. M. CZARNECKI, J. DONAHUE, A. RAZAVI, O. VINYALS, T. GREEN, I. DUNNING, K. SIMONYAN, C. FERNANDO, AND K. KAVUKCUOGLU, *Population based training of neural networks*, arXiv preprint 1711.09846, (2017).

[15] S. JEBBARA AND P. CIMIANO, *Aspect-based relational sentiment analysis using a stacked neural network architecture*, 2017.

[16] Z. JIE AND W. LU, *Dependency-guided lstm-crf for named entity recognition*, 2019.

[17] J. KAPUSTA, P. HÁJEK, M. MUNK, AND L'UBOMÍR BENKO, *Comparison of fake and real news based on morphological analysis*, Procedia Computer Science, 171 (2020), pp. 2285 – 2293. Third International Conference on Computing and Network Communications (CoCoNet'19).

[18] J. LAFFERTY, A. MCCALLUM, AND F. C. PEREIRA, *Conditional random fields: Probabilistic models for segmenting and labeling sequence data*, 2001.

[19] C. LI AND Y. TIAN, *Downstream model design of pre-trained language model for relation extraction task*, 2020.

[20] X. LI, J. FENG, Y. MENG, Q. HAN, F. WU, AND J. LI, *A unified mrc framework for named entity recognition*, 2019.

[21] X. LI, F. YIN, Z. SUN, X. LI, A. YUAN, D. CHAI, M. ZHOU, AND J. LI, *Entity-relation extraction as multi-turn question answering*, 2019.

[22] J. MILLER, *Information input overload and psychopathology*, The American Journal of Psychiatry, (1960), pp. 695–704.

[23] C. T. NGUYEN, *Echo chambers and epistemic bubbles*, Episteme, 17 (2020), p. 141–161.

[24] M. E. PETERS, M. NEUMANN, M. IYYER, M. GARDNER, C. CLARK, K. LEE, AND L. ZETTLEMOYER, *Deep contextualized word representations*, 2018.

[25] M. PONTIKI, D. GALANIS, J. PAVLOPOULOS, H. PAPAGEORGIOU, I. ANDROUTSOPOULOS, AND S. MANANDHAR, *SemEval-2014 task 4: Aspect based sentiment analysis*, in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, Aug. 2014, Association for Computational Linguistics, pp. 27–35.

[26] S. PORIA, E. CAMBRIA, AND A. GELBUKH, *Aspect extraction for opinion miningwith a deep convolutional neural network*, Knowledge-Based Systems, 108 (2016).

[27] A. RADFORD, J. WU, R. CHILD, D. LUAN, D. AMODEI, AND I. SUTSKEVER, *Language models are unsupervised multitask learners*, OpenAI Blog, 1 (2019), p. 9.

[28] C. RAFFEL, N. SHAZEER, A. ROBERTS, K. LEE, S. NARANG, M. MATENA, Y. ZHOU, W. LI, AND P. J. LIU, *Exploring the limits of transfer learning with a unified text-to-text transformer*, arXiv preprint 1910.10683, (2019).

[29] M. SAEIDI, G. BOUCHARD, M. LIAKATA, AND S. RIEDEL, *SentiHood: Targeted aspect based sentiment analysis dataset for urban neighbourhoods*, in Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, Osaka, Japan, Dec. 2016, The COLING 2016 Organizing Committee, pp. 1546–1556.

[30] A. SEUDKER, A. STEPUTAT, AND S. DRÖSSLER, *297 determinants and consequences of information overload – a systematic review*, Occup Environ Med, 75 (2018).

[31] M. SHOEYBI, M. PATWARY, R. PURI, P. LEGRESLEY, J. CASPER, AND B. CATANZARO, *Megatron-lm: Training multi-billion parameter language models using model parallelism*, arXiv preprint 1909.08053, (2019).

[32] K. SHU, A. SLIVA, S. WANG, J. TANG, AND H. LIU, *Fake news detection on social media: A data mining perspective*, SIGKDD Explor. Newsl., 19 (2017), p. 22–36.

[33] C. SINDERMANN, J. D. ELHAI, M. MOSHAGEN, AND C. MONTAG, *Age, gender, personality, ideological attitudes and individual differences in a person's news spectrum: how many and who might be prone to "filter bubbles" and "echo chambers" online?*, Heliyon, 6 (2020), p. e03214.

[34] J. STRAKOVÁ, M. STRAKA, AND J. HAJIC, *Neural architectures for nested NER through linearization*, in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, July 2019, Association for Computational Linguistics, pp. 5326–5331.

[35] C. SUN, L. HUANG, AND X. QIU, *Utilizing bert for aspect-based sentiment analysis via constructing auxiliary sentence*, arXiv preprint 1903.09588, (2019).

[36] P. VITÁNYI AND R. CILIBRASI, *Normalized web distance and word similarity*, CRC Press, 01 2010, pp. 293–314.

[37] H. YANG, B. ZENG, J. YANG, Y. SONG, AND R. XU, *A multi-task learning model for chinese-oriented aspect polarity classification and aspect term extraction*, arXiv preprint 1912.07976, (2019).

[38] B. YU, Z. ZHANG, X. SHU, Y. WANG, T. LIU, B. WANG, AND S. LI, *Joint extraction of entities and relations based on a novel decomposition strategy*, 2019.