The Algorithm of Search and Identification of Internet Sources on Historical Events

Andriy Peleshchyshyn [1], Tetiana Bilushchak [2]

[1]Social Communication and Information Science Department,
Lviv Polytechnic National University, UKRAINE, Lviv,
S. Bandery street 12, E-mail: apele@ridne.net

[2]Social Communication and Information Science Department,
Lviv Polytechnic National University, UKRAINE, Lviv,
S. Bandery street 12, E-mail: Tetiana.M.Bilushchak@lpnu.ua

Abstract – The paper presents the results of creating the algorithm of search and identification of Internet sources on historical events for efficient work of researchers of history.

Key words – algorithm of search, Internet sources, historical events, social networks, Facebook, consolidated information resource, search parameters, verification, establishing authorship, algorithm detalisation.

I. Topicality and Setting the Goal

The research of historical facts concerning Ukraine of the 21st century is also possible through the informational potential of virtual social networks, which makes urgent the issue of creating the algorithm of search and identification of Internet sources on historical events.

The goal of this research is creating the algorithm of information search in social media for the effective search index for historical information.

II. The Algorithm of Search and Identification of Internet Sources on Historical Events

For carrying out the analytical critic analysis of the found in the social media source it is suggested to see the following action algorithm (Fig.1). The algorithm of search and identification of Internet sources on historical events is repeated many times for various input search parameters. The formation of parameters groups is a separate task, which is solved by the ontology forms of the historical event. From Fig.1 we can see that the scheme of algorithm of search and identification of a source in the Internet is formed in the following way: first of all the researcher enters the parameters of search of historical information and the search according to the given parameters is performed, the next step is defining the set of the possible sources, after which a certain cycle of algorithm is entered – the process of finding the technical identification of the source, which appears before the researcher in the form of URL-address. That is, in case of finding a link in a post in social networks like Facebook, Twitter, VKontakte or an article in a forum, the researcher mostly finds not a complete source, but only its part. In that case the researcher needs to find the reference to the author or a community, where the post first appeared. The next step of the algorithm is the verification for the signs of bias or misrepresentation, if
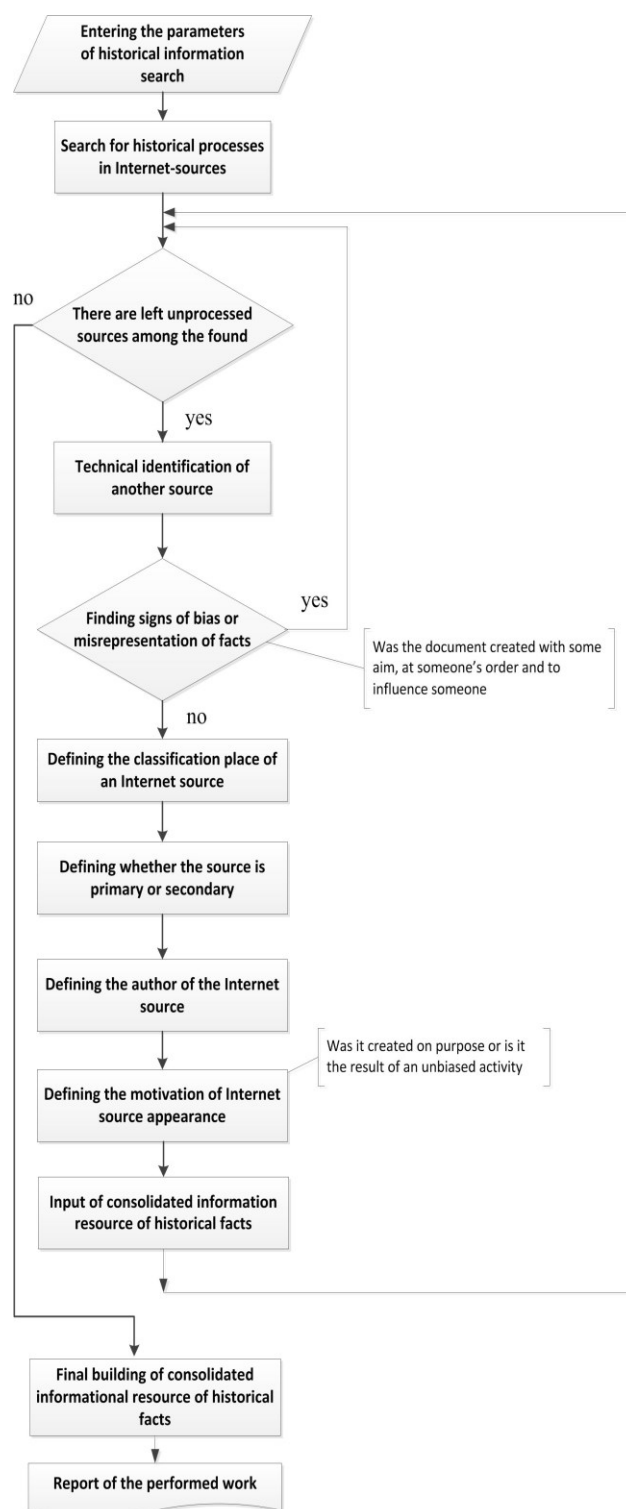

Fig. 1. The scheme of the algorithm of search and identification of Internet-sources on historical events

there is a certain misrepresentation, then it is necessary to repeat the previous step, that is to perform the search for historical information in another source. If the result is successful and there is no information distortion, the following steps should be carried out: the identification of the group the Internet source belongs to; defining whether the Internet source is primary or secondary and its author; finding out the motivation of the Internet source

appearance and forming a consolidated informational resource of historical facts. If the information resource already has a full representation of this historical event, the search according to the given parameters may be finished, the final consolidated information resource of historical facts is formed and the report of the performed task is created. However if the step of finding the mismatch of facts with other sources for the same question and the criteria of information resource content are not satisfactory, than it is necessary to go back to the step of the unprocessed sources from the number of found. For specification and better understanding of the main algorithm it is necessary to make each process more detailed, which will help to carry out each task more efficiently. In specification of process of the algorithm "Entering the parameters of historical information search" an important stage for entering the parameters is forming a terminology and key words in the researched sphere. Use of ontologies in information search enables to extend the possibilities on the basis of adding to the information requests. One of the important stages of finding the necessary historical information in the Internet by the researcher is performing successful information search, which requires certain experience and knowledge. In addition to this the typology of search is complicated and multiple, in particular after forming the search terms and requests connected with the historical event, the search is performed by several search engines as Google, Мета, Yahoo and others, and also through the search tools of social networks: Facebook, VKontakte, Twitter.

Depending on the received results of the requests and their specification there is a need of use of synonyms, which allows to include into results not only the terms, which are given in the request, but also the words with similar meaning. For even higher results of search for a historical event there is a possibility of finding a historical fact in several languages, which is called for by the efficiency of the search. For example, the process of search for the necessary information sources about the events of the "Revolution of Roses" in Georgia was performed in three languages, and a large portion of information was found in English, less in Russian and Ukrainian [1, 2, 3]. However, it should be kept in mind that the same event, phenomena or fact may be used and interpreted by various social media in a distorted way for influence of evaluations, intentions and orientation of population or separate persons in an informational-psychological war. The last process of specification of the process of algorithm "Search of historical processes in Internet sources" is establishing the suitability of the found sources, i.e. their relevance and pertinence. Relevance is the formal (objective) correspondence of the found document to the request or the level of adequacy to the request. Pertinence is the subjective evaluation by the user of the suitability level of the found information according to his/her needs or usefulness of the result [4]. Let us consider one of the most important stages of search and identification of Internet sources on historical events – the specification of the process of algorithm "Finding the features of bias or misrepresentation of facts". For futher

correct receiving of the result we identify three features: chronological order, consistency and relevance. Then it is possible to present the process of analysis of features of bias and misrepresentation of facts in the form of three blocks, which represent the verification of meaning of the logical expression of condition as: "Does the source correspond chronologically and gradually to the events described in it?", "Are there logical judgements and evidence?", "Is there the necessary level of statements relevance?". Then takes place the first check whether there is the feature (F) of chronological order and graduality (COG), and the meaning of TRUE or FALSE is assigned. The next block of verification is for the feature of logical judgements and evidence (LDE) and the verification for the feature of the level of statement relevance (LSR). That is, the algorithm consists of three consecutive blocks, which are not inserted into each other. Each block looks like an entrance into branching, if not than this change is assigned the meaning FALSE, if yes, than the change is TRUE, the blocks meet and go to the next such block. In the end there is one operator, where the result (R) is realized, i.e. R is TRUE, if the features are the elements of conjunction.

Conclusion

Thus, the use of the information potential of social networks is extremely important and significant factor of historical research, especially in study and analysis of the key events in the history of new era. Therefore the created algotrithm of search and identification of Internet sources on historical events and the specification of the main algorithm is a certain means of optimizing the search and receiving quality results of the search.

References

[1] A. Peleschyshyn, T. Bilushchak, "Analysis and classification of historical-oriented Internet sources," [Electronic resource] Proceedings 5th International Academic Conference of Young Scientists "Humanities and Social Sciences 2015" (HSS-2015) – Lviv: Lviv Polytechnic Publishing House, 2015. – Electronic edition on CD-ROM.– pp. 200-202.

[2] A. Peleschyshyn, T. Bilushchak, "Internet-dzherela yak zasib vykorystannia informatsii dlia dzhereloznavtsiv" ["Internet-Sources as a Means of Information Use for Source Researchers"], Sumy historical archival journal, vol. XXVI, pp. 16 – 23, 2016.

[3] Ia. S. Kalakura, I. N. Voitsekhivska, B. I. Korolov, "Istorychne dzhereloznavstvo: Pidruchnyk dlia studentiv istor. Spets. VNZ" [Historical Source Studies: a Textbook for the Students of Historical Specializations at Higher Educational Institutions] Kyiv: Lybid, 488p, 2002.

[4] V. O. Kulikov. (2007). Poshuk istorychnoi informatsii v interneti [Search for historical information on the Internet] [Online]. Available: http://dspace.univer.-kharkov.ua/bitstream/123456789/104/1/Kulikov_Poshyk_informacii.pdf.