Zbigniew ŁĘSKI
https://orcid.org/0000-0003-4145-6955
Jan Długosz University in Częstochowa
e-mail: zleski@ujd.edu.pl

The Perspective of Using Big Data Technology for the Purposes of Educational Transactional Analysis

Summary

The 21st century is the time of dynamic technological progress. This phenomenon has a direct impact on educational processes – both at the level of needs and possibilities. And in this context, it is worth taking a closer look at what Transactional Analysis and Big Data technology possibilities can offer in this area. As D. Pankowska writes, "Educational Transactional Analysis applies the assumptions of Transactional Analysis with reference to educational goals, conditions of its efficacy, methods of didactic and educational work, attitude to problems" (Pankowska, 212, p. 23). Additionally, it is enjoying more and more interest. However, to make certain actions really efficient, it is necessary to diagnose educational reality in a detailed and multilayered way. Big Data analytics might prove useful in this task as "it allows for taking decisions on the basis of the data that have not beena vailable or useless before. Thanks to advanced analytical techniques, […] one can analyse data sources not used before, independently or together with existing and traditionally available information" (Tanaś, 2019, p. 13).

Keywords: Educational Transactional Analysis, Big Data, new media.

This journal has already published many articles on possibilities and benefits of using Educational Transactional Analysis. It allows both for diagnosing the whole rage of phenomena occurring in educational reality and taking effective action in this field. In this context, it is worth mentioning the following research: D. Pankowska (2010) and her studies pertaining to teachers from the point of view of Transactional Analysis, A. Pierzchała (2013) and her studies of school passivity, A. Sarnat-Ciastko (2015) who studied the effectiveness of school tu-

toring from the point of view of Transactional Analysis, and numerous publications by J. Jagieła, who shows various possibilities of using Transactional Analysis in pedagogical work. Yet, the first attempt to square the circle, i.e. to match Big Data analytics and TA terminology was made by Z. Wieczorek (2018), using an example of social media research.

What is Big Data analytics? It can be defined in the simplest way as an analysis of human activity in cyberspace. Nowadays, no one doubts that cyberspace constitutes a real and crucial part of social space. It is difficult to defend a thesis that is it a virtual and unreal world. Man works, earns, trades, studies, finds entertainment, etc. in cyberspace. Since 2011, this term has been functioning in Polish law (the Act of 30th August 2011 on the amendment of the act on martial law and competencies of the Chief of Staff of the Armed Forces and rules of their submission to constitutional bodies of the Republic of Poland and other acts – [Dz. U.] Journal of Polish Law No 222, item 1323). In February 2018, the CYBER.MIL.PL programme was launched. Its aim is to increase the state's and citizens' safety in cyberspace (https://www.gov.pl/web/obrona-narodowa/bronimy-polska-cyberprzestrzen). These facts clearly show that it is both a real and crucial issue.

Similarly, educational reality contributes significantly to cyberspace. On the one hand, it takes place at an implicit, formal level – in the shape of curricula, e-learning activities, independent or guided search for knowledge and information or a communication network between particular subjects of the educational process. On the other hand, there exists an informal, hidden level – in the shape of new forms of cheating, copying ready solutions, undertaking activities connected with entertainment at the time formally devoted to studying, informal contact among students, etc. (see also: Łęski, 2018, pp. 23–29). These two areas can be analysed with the help of Big Data technology. It gives us an opportunity to analyse all subjects engaged in the educational process, e.g. within opinions expressed by them, judgements, declared emotional attitude, etc. on social sites, discussion forums or in other places of their online activity. What is more, spontaneous comments are also subject to analysis. Such comments are not burdened with the awareness of taking part in a survey, thus it can be assumed that the probability of their sincerity and truthfulness is much higher than while using traditional research procedures and methods (especially those based on quantitative analysis).

One of the first scientists who started to use Big Data technology in social research is W. Gogołek. For instance, he writes that "Digital resources of Big Data are huge, non structured blocks of information whose volume has exceeded a critical size of conventional information sets – structured computer data bases. Already in 2012, the volume of BD exceeded two zettabytes of information, which in shape of printed books would form 8000 columns reaching from the Earth to the moon. To a large extent, these sets are formed by active individuals

using social media […], professional publications […], phone conversation registers and other sources of information in a digital form" (Gogołek, 2015, p. 98). Modern technology lets us search such data sets taking into account particular expressions, key words, etc. As a result, we can obtain data whose analysis makes it possible to learn about social moods, opinions, interests, etc.

A crucial issue that seems to occur at this moment is a question about an ethical side of this kind of research initiatives. We analyse online activity of an enormous number of Internet users, without their knowledge and consent. One of researchers that focused on this question was M. Androjevic: "If […] privacy is contextual (because of established expectations associated with particular information-collection contexts), then the big data era challenges people to develop "contextual" norms for the use of data whose uses can be radically, unpredictably decontextualized. Thanks to the proliferation of monitoring technologies (license plate readers, smart cameras, drones, RFID scanners, audio sensors, etc.), data scraping continues to extend its reach both online and off, so fewer places and activities are likely to be exempt from the logic of the big data divide, whereby people are separated from their data and excluded from the process of putting it to use. Overcoming the digital divide means exacerbating the big data divide. Greater access to and facility in the use of smartphones and networked laptops, tablets, and computers of one kind or another means more data to store, sort, and mine. More comprehensive forms of data mining promise to serve a growing variety of decision-making, forecasting, and sorting operations (2012, p. 1685)". The research conducted by him in the form of focused interviews showed that interned users who are aware of the fact that their data are collected and can be subject to any analysis at any moment react with frustration and powerlessness: "Overall, concern about actual harms came across less vociferously than did frustration over a sense of powerlessness in the face of increasingly sophisticated and comprehensive forms of data collection and mining. Focus group participants generally agreed with responses emphasizing that this sense of powerlessness extended to their lack of knowledge over how personal data might be used. As one respondent admitted, "We really don't know where things collected about us go – we don't understand how they interact in such a complex environment" [female, 22]. Interview respondents and focus group participants alike noted the seemingly endless appetite for personal data: "It's not just what you want – it's where you are, what you do. It's everything. You're not free any more. You're just a slave of these companies" [male, 22]. This may come across as hyperbolic, but nonetheless noteworthy is the stark contrast between this response and the rhetoric of freedom, empowerment and convenience that has long underpinned the promotion of the online economy. The contrast highlights the challenge posed by the power asymmetries ushered in by big data mining" (Andrejevic, 2014, p. 1685)

Unfortunately, it should be noticed that the specificity of Big Data analytics does not offer such a solution to the aforesaid problems that all subjects engaged

in the research process (especially at the side of researched people) can be fully informed about the time, place and aim of their data analysis. A certain solution deriving from the GDPR (General Data Protection Regulation) provisions is consent that a user of a given internet service must give before they start using it. However, it does not, to any extent, offer that individual any possibility to check who, when and why analyses these data. It is thus a given researcher who should be responsible for respecting personal data they have access to and making utterances anonymous enough in case of quotes. The researched individual should not know that their posts, together with thousands of others, are subject to a given analysis at that time.

At this moment, I would like to devote more attention to the process of data collection and analysis. The biggest challenge that a given scholar who wants to use Big Data analytics in their pedagogical research faces is, first of all, appropriately formed questions for the information system, and then, a reliable, qualitative analysis of the material obtained. This task is difficult as we talk about a new technology – especially in the context of its use in the pedagogical field. Thus, there are no ready examples of solutions or good practice that could constitute a reliable foundation. In a certain sense, a person deciding on the research based on Big Data technology contributes to the process of developing a new methodology, both in the area of collecting data in this way and their selection and analysis. It might seem that it is a relatively easy procedure. The fact that data are collected automatically by information systems might create an impression that the whole research process shall be relatively easy and fast to conduct. Unfortunately, this impression is wrong. Problems occur already at the stage of question formulation. Next, the researcher has to make a preliminary selection of the material obtained. Its analysis often makes them come back to the previous stage and correct their questions (make them more precise). The Internet is a huge set of random information and it is not always possible to predict in which contexts users' posts shall contain key words we are interested in.

What does it look like from the practical side? As it was mentioned earlier, the Educational Transactional Analysis Research Team made its first steps towards Big Data analytics using the software tools made available by Unamo company (https://unamo.com/), and their effects are already noticeable, e.g. in publications by Z. Wieczorek (2018). This tool facilitates searching social media regarding the countries such as Poland, Great Britain and the United States (gradual expansion of the offer is planned). While introducing key words that are to be searched, we also define words that shall accompany them and those that should be rejected from the results of our search. Additionally, it can be defined how far from the main entry key words should be found (e.g. not further than 10 words). Results are presented, among all, in the form of word clouds which accompany the defined search. For example, a search for the entry *Transactional Analysis* in Poland (the analysis carried out at the turn of November and December 2019)

renders a word cloud with such entries as: *disorders, interpretations, work, therapy, patients, health, scales*, etc. The researcher can verify the context in which a given word occurs, check where a given word was found by chance (being, for example, part of an advertisement and not contributing substantially to the analysis in question) and exclude it from search results. At this stage we face an arduous process of looking through the results obtained, determining accompanying words, removing particular speakers from the search results, etc.

What is more, the system can automatically estimate the so-called sentiment – i.e. determine emotional background of expressions linked with a searched entry. In case of the aforementioned analysis concerning *Transactional Analysis*, it turned out that 43.4% of posts are neutral, only 1.89% are negative, and 54.72% are positive. The place where mentions connected to that key word occur most frequently is Facebook (89.94%). Unamo makes it possible to analyse the results in the light of gender, to make comparisons between different analyses, etc.

Currently, Unamo is one of constantly growing ranges of possibilities of using Big Data analytics in pedagogical research. Whereas the system was established primarily to cater for business' needs, it can be successfully applied in scientific research in the field of social studies, offering ready tools for the analysis of collected material. However, it should be emphasised that it is not the only solution. Thanks to the cooperation of NASK (Research and Academic Computer Network) with representatives of the world of science, especially with M. Tanaś from the Academy of Special Education, a prototype version of an application CONTENT 1.0 was developed. This application was specially designed to support pedagogical research. Certainly, it will be developed soon, it shall get access to a bigger number of sources and become available to interested researchers. Functionalities offered by Google Analytics might also prove sufficient for many applications.

Z. Wieczorek drew the reader's attention to the fact that thanks to appropriately formed questions, Big Data analytics can be useful in research based on the concept of Transactional Analysis. He mainly focused on structural analysis and life positions. As he duly remarks, "If we write that something 'has to be done', 'should be done', most probably we deal with the Parent ego state. If we write that we 'want', 'desire', 'are afraid of', or 'have to', most probably we deal with the Child ego state. Descriptive language, relating events or asking matter-of-fact questions usually accompany actions from the level of the Adult ego state. […] using expressions 'they're always doing that to me', suggests that we have a feeling that others are not OK in their attitude towards us" (2018, p. 139). However, it seems that we can go much further there. Analysing the profiles of ego states, it is worth concentrating mainly on functional analysis as we deal here with well-defined, behavioural symptoms of functioning in a given ego state. The analysis of available egograms (i.e. questionnaires to research ego state profiles) can provide us with many inspirations for the right selection of key words. Apart from

the areas mentioned above, it is worth investigating scripts and drivers. Drivers consist of five phrases that are connected with specific behaviours: *Be Perfect, Be Strong, Try Hard, Please Others, Hurry Up.* Thus, they are linked with particular words, expressions or declarations, which can be analysed. Finally, an attempt at analysing the phenomenon of school passivity, which in Transactional Analysis is understood not only as lack of activity in the face of a problematic situation, but also as undertaking inefficient actions, seems very interesting. A. Pierzchała (2013) explored that issue more thoroughly writing that "One has to pay attention to the fact that passivity does not occur only in an individual's thinking, but it has external manifestations, which means it is observable. Additionally, such behaviour eliminates autonomy in human functioning, blocking in a certain way the objective perception of reality from the level of the Adult ego state. Taking into account what was said before, namely that passive activity is scriptural to some extent, it can be stated that passivity has its origins mainly in the Child ego state structure (and more precisely in the Adapted Child), but also in the Parent ego state, which derives from its symbiotic nature" (2013, pp. 86–87). Thus, taking into account the fact that a man expresses their behaviours and attitudes also while communicating with others, an analysis of such messages can render very interesting results in this field. Big Data allows for an analysis of messages expressed spontaneously and naturally, and not forcibly an intentionally. To a certain extent, one can assume that it resembles research based on participant observation in natural environment.

Due to its clear terminology and a series of available tools based on a word (egograms, questionnaires, semantic differentials, etc.), Transactional Analysis seems to go well with the specificity of research possible to conduct with the help of Big Data analytics. Thus, it may turn out to be the concept opening this new, unexplored research area to educators, considerably facilitating the development of methodologies and good practice, which is still lacking in this area.

Concluding, the following words of M. Tanaś are worth quoting: "[…] Big Data technologies form […] new, attractive cognitive processes. It is not only about the volume of data, but about their credibility, uniqueness and a possibility to undertake pioneer scientific research in the fields unexplored so far. […] the development of information sciences and technological progress, which helped to deliver this data, also formed new, powerful tools, which already prove very useful today, while collecting and sending but also while analysing and comprehending. It is time to conduct research with the use of Big Data also in pedagogy" (Tanaś, 2015, pp. 23–24). As far as the pace of new technologies and new media development is concerned, Big Data analytics itself is a phenomenon that can hardly be called a novelty. This article quotes publications devoted to that topic, dating from 2014 or 2015. Initially though, access to information on that issue and to necessary tools was difficult and reserved to a rather small group of specialists. In December 2015, W Gogołek wrote: "The problem does not lie in the

volume of Big Data but in the fact that most users do not have an appropriate tool platform for directed analysis at their disposal"(2015, p. 103). Currently, there are more and more solutions and services that make it possible to use that technology almost to everyone. Therefore, there is a new perspective of extending research conducted in pedagogy, also by that topic that so far has not been available to researchers publishing in that discipline.

Bibliography

Andrejevic, M. (2014). The Big Data Divide. *International Journal of Communication*, *8*, 1683–1689.

Gogołek, W. (2015). Big Data. Sieciowe źródło informacji dla edukacji. In: M. Tanaś, S. Galanciak (ed.), *Cyberprzestrzeń człowiek edukacja. Cyfrowa przestrzeń kształcenia* (pp. 97–103). Kraków: Impuls.

Łęski, Z. (2018). Nowa rzeczywistość – stara szkoła. Nieformalne aspekty szkolnej cyberprzestrzeni. *Problemy Opiekuńczo-Wychowawcze*, *574*(*9*), *23–29*. http://dx.doi.org/10.5604/01.3001.0012.8072.

Pankowska, D. (2010). *Nauczyciel w perspektywie analizy transakcyjnej*. Lublin: Wydawnictwo UMCS.

Pankowska, D. (2012). Analiza transakcyjna w edukacji czy edukacyjna analiza transakcyjna? *Edukacyjna Analiza Transakcyjna*, *1*, 13–30.

Pierzchała, A. (2013). *Pasywność w szkole*. Częstochowa: Wydawnictwo AJD w Częstochowie.

Sarnat-Ciastko, A. (2015). *Tutoring w polskiej szkole*. Warszawa: Difin.

Tanaś, M. (2019). Big Data – Informatyka w metodologii nauk pedagogicznych. In: M. Tanaś, M. Kamola, R. Lange, M. Fila (eds.), *Big Data w edukacji. CONTENT 1.0 – prototyp aplikacji do analizy treści Internetu* (pp. 9–24). Warszawa: Wydawnictwo Akademii Pedagogiki Specjalnej; NASK – Państwowy Instytut Badawczy.

Wieczorek, Z. (2018). Wykorzystanie analizy transakcyjnej w badaniach mediów społecznościowych. *Edukacyjna Analiza Transakcyjna*, *7*, 133–159. http://dx.doi.org/https://doi.org/10.16926/eat.2018.07.09.

Perspektywa wykorzystania technologii Big Data na potrzeby edukacyjnej analizy transakcyjnej

Streszczenie

XXI wiek to czas dynamicznego postępu technologicznego. Zjawisko to ma bezpośredni wpływ na procesy edukacyjne – zarówno na płaszczyźnie potrzeb, jak i możliwości. I właśnie w tym kontekście warto przyjrzeć się, co na tym polu może zaoferować analiza transakcyjna oraz możliwości, jakie niesie technologia Big Data. Jak pisze D. Pankowska „Edukacyjna analiza transakcyjna aplikuje założenia analizy transakcyjnej w odniesieniu do celów edukacji, uwarunkowań jej skuteczności, metod pracy dydaktyczno-wychowawczej, podejścia do problemów" (Pankowska, 2012, s. 23). Cieszy się przy tym coraz to szerszym zainteresowaniem. Aby jednak podejmowane działania mogły być naprawdę skuteczne, konieczna jest dokładna i wielopłaszczyznowa diagnoza rzeczywistości edukacyjnej. W tym zadaniu pomocne może okazać się wykorzystanie analizy Big Data, która „pozwala podejmować decyzje na podstawie danych, które wcześniej były niedostępne lub nieużyteczne. Dzięki zaawansowanym technikom analitycznym, […] można analizować wcześniej niewykorzystywane źródła danych niezależnie lub razem z istniejącymi i tradycyjnie dostępnymi" (Tanaś, 2019, s. 13).

Słowa kluczowe: edukacyjna analiza transakcyjna, Big Data, nowe media.