### Sentiment analysis as stress detector



Digital transformation is a process that touches every area of modern human life. This statement of Rob Kling<sup>1</sup> from 2000 is still up-to-date.

## Can you imagine living and working during the pandemic times as these of COVID -19 without digital transformation benefits?

Imagine yourself in 2000 or even in 2010. You start to work from your home. Can you imagine how we could switch to the remote work and meetings without the wide range of information and communication technology solutions accessible today?

Perhaps until today digital transformation benefits in your professional and private live have been so obvious and in the background that you do not pay attention to them anymore.

However, despite all these unquestionable digital transformation benefits the own process of digital transformation may be a source of employees' stress as a response to digital and social changes in the workplace. It is because of the redefinition of the work scope and responsibilities, a growing number of employees, requirements, new tasks, competences, and work mode, as well as changes in human team management. They are all examples of changes as consequences of digital transformation in workplace.

Organizational changes could give rise to resistance due to their unpredictability, as well as to the interference with the existing order and structures of the company. This resistance among employees can be expressed in terms of passive fears, severe stress, in some cases aggression, as well as professional burnout. The transformation of rooted patterns of behavior and value systems requires targeted and lengthy training measures to be carried out by managers, psychologists and educators. The main focus of the DT works is on project implementation and that is why, very often, change in training programs is limited to communication of changes, procedures and instructions.

<sup>&</sup>lt;sup>1</sup> R.Kling was a North American professor of Information Systems and Information. Father of social informatics.

Professional development of employees is by consequence neglected in this area when changes are introduced, which leaves them without the tools to manage these changes and challenges.



Thus, there appears the phenomenon named digital transformation stress, that is defined as an emotional response of an employee to a specific situation, which is the digital transformation process. To be more specific, we define it as employee's perception of the situation of digital changes or the IT implementation process, perceived as a potential threat to the current, familiar work style or/and to the current position. The digital transformation stress may even apply to employees who initially presented openness and positive attitude towards the digital transformation projects and implementations of new solutions.

# How do we know that this phenomenon (digital transformation stress) occurs? How do we know how to measure this? And what can we do with this knowledge?

One of the traditional ways to measure stress among employees has been the psychometric survey, which employees should complete. Because of digital transformation, nowadays ICT solutions in machine learning (ML) and data mining are more and more often used to examine possibilities to identify human emotions, cognitive functions, or disorders. Hence the idea to use IT solutions to have this possibility to automatically detect and monitor digital transformation stress among employees and then to give the HR department a possibility to counteract it.

#### OK, but what is machine learning?

Machine-learning is an application of artificial intelligence (AI) that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed. The process of learning begins with observations or data, such as examples, direct experience, or instructions, to look for patterns. In data models, machine learning has demonstrated great success in learning complex patterns that enable them to make predictions about unobserved data. Therefore, machine learning methods are being increasingly tested to identify specific features of stress. Such studies include the examinations of the effects of cognitive or physical stress, e.g., on specific context of writing, patterns used in smartphones texting or patterns of activity of the brain.

The second alternative method of stress measurement as sentiment analysis is the use of natural language processing, text analysis, computational linguistics to systematically identify, extract, quantify, and study affective states and subjective information. Sentiment analysis could analyze people's opinions, appraisals, attitudes, and emotions towards many entities such as products, services, organizations, phenomena, issues, subjects, and their attributes.

Together with sentiment analysis and machine learning there was a possibility to implement a tool to detect stress among employees by analyzing negative emotions occurrence by sentiment markers in employees' written communications, e.g., official emails and helpdesk tickets. Subjective expressions may contain explicit sentiment markers which can be identified by sentiment analysis of on-line texts in applications, opinion forums and service help desk for customers. Using these platforms, people express their subjective opinions and ratings, often with strong emotional load (especially when they are dissatisfied).

The study where the sentiment analysis method of stress detecting was used was based on texts from help desk (HD) ticketing system, used in the organization selected for research.

#### What is the help desk ticketing system?

The help desk applications or systems are very common in companies to help and accelerate solving different problems or issues in the companies. The main goal of IT help desk applications is to provide employees with support related to IT projects, software, computers, electronic equipment etc. Help desks allow to track and sort employee tickets with the help of a unique number, and can frequently classify problems by user, computer program, or into similar categories. Help desk (HD) communication, as it is aimed at solving issues, should be brief and specific. In each ticket registration the user should choose the category of issue, e.g., the name of an application affected. The HD ticket description should be short but contain all necessary details like screenshots or system alerts, that is all of the information which could be helpful in solving the issue. The employees register tickets mainly in case of technical issues, usually related to new systems implemented. This can happen as employees want to prioritize their issue, get it solved faster, identify the person responsible for solving their problem or simply to avoid starting lengthy email exchanges.



#### OK, but what is the relation between HD system and stress detecting?

According to Kamil Imbir's<sup>2</sup> research on emotion taxonomy, the identification of emotional markers was started by the initial analysis with limited, basic word phrases and syntax collection containing emotional markers like:

- imperative forms,
- •exclamation marks mixed with question marks,
- •generalizations ("always", "never", "nobody", "again"),
- •irritations ("it annoys me", "I'm sick of", "why")
- •and even curses or swearword expressions

Next, the frequency of occurrence of basic keywords in HD tickets was analyzed as well as the average number of characters in the ticket per user and the number of HD requests per user over two periods: the first one between January 2019 and June 2019 (6 months - 2036 requests and 208 users) and the

 <sup>&</sup>lt;sup>2</sup> "Toward a Human Emotions Taxonomy (Based on Their Automatic vs. Reflective Origin)"
Maria T. Jarymowicz, Kamil K. Imbir, <u>https://doi.org/10.1177/1754073914555923</u>.

second one between January 2020 and June 2020 (6400 HD requests and 223 users). Negative emotions' occurrence was detected in help desk tickets of 79% help desk users, i.e., employees of the selected organization. On charts below, there are presented: on the first chart differences between the number of registered HD tickets and negative emotional markers occurrences in HD ticket texts, in first six months of 2019 and 2020. On the second chart there are differences between gender in the number of HD tickets and the negative emotional markers occurrences in HD ticket texts, in first six months of 2019 and 2020.



number of registred HD tickets
Number of emotional markers in HD tickets

The main issue for sentiment analysis is to establish adequate lexicon and lexical clues, which are characterized for the specific domain. It is also very important for machine learning and sentiment analysis approach to identify correct classification methods. Therefore, the crosscheck was conducted in two ways: 1) with use of the Polish CLARIN-PL platform, and by comparing the result of sentiment analysis with psychometric survey results.

#### What is the CLARIN-PL platform?

CLARIN is an abbreviation for the Common Language Resources and Technology Infrastructure, <u>https://ws.clarin-pl.eu/sentyment.shtml</u>, which aims at providing easy and sustainable access for scholars in the humanities and social sciences to digital language data and advanced tools to discover, explore, exploit, annotate, analyze or combine them, independently of where they are located. The CLARIN-PL sentiment analysis tool has been developed on basis of short texts expressing opinions of users (i.e., *TripAdvisor, know medcin, ceneo*). An emotional marker has been assigned to separate words. However, there are some structures not possible to analyze, such as negations, two-words structures etc. Therefore, CLARIN was used as a supporting tool for exploration of Digital transformation stress sentiment analysis in the chosen organization's written materials – e.g., help desk tickets used in the research.

In the screenshot below, there is an example of CLARIN-PL functionality. You can put a sample of text and CLARIN-PL automatically counts words, identifies emotional markers of words and classifies words into categories. We use this verification of a text sample regarding the presence of negative emotions set in a help desk tickets text. In the CLARIN-PL tool, the first examination of the sentiment was performed, together with statistics of the most common words and their emotional markers. This first (CLARIN) verification of the occurrence of negative emotions markers in helpdesk tickets was successful, therefore research on an automatic tool for measuring stress was justified. Despite the CLARIN-PL limitations (which were mentioned above), the collection of HD ticket samples has been confirmed in context of sentiment occurrence.



Wynik



The next goal of the study on detecting digital transformation stress among employees was to find an adjusted method to verify if the prepared algorithm could be used in digital stress detecting only by sentiment analysis of written communications. Therefore, the dedicated algorithm was prepared, with use of negation phrases, helplessness syntax, emotional expressions etc. Then, a sentiment analysis of help desk request text was conducted to estimate how employees' stress could manifest in official written communication. Upon identifying negative emotions markers in help desk ticket text, the relationship between the frequency of ticket registration and negative emotion markers in help desk ticket text was analyzed.

#### How do we know that this algorithm is correct? Does the algorithm support measurement stress?

One of the ways to verify the correctness of algorithms is to compare them with psychometric tools, such as established surveys. This survey was conducted at the same time when the Help desk ticket texts were collected. Next, we matched logins of help desk system with the ones of survey<sup>3</sup> and examined the correlation. We confirmed the relationship between negative emotions markers in helpdesk tickets and results of digital transformation stress survey. Employees who reported high stress (both general stress at work and digital transformation stress) have presented high occurrence of negative emotion markers in registered helpdesk tickets texts. This relationship was very strong and confirmed by the correlation calculation.

The interdisciplinary research (using both psychological and informative tools and approach) confirmed that there is high and positive correlation between the psychometric stress measurement results, based on an established survey and sentiment analysis results of help desk ticket data set.

#### What are the implications? How could it help employees?

The novel proposed tool and approach will allow for the continuous monitoring of digital transformation stress among employees in any organization, without using psychometric surveys. This means it can be done without engaging employees to complete surveys, answer questions and, what is important for employers and employees' point of view, without disturbing and wasting people's time. Moreover, this tool will allow companies to make better use of their employees' time and to react quicker when an

<sup>&</sup>lt;sup>3</sup> All data was anonymized and all participants confirmed their contest to take participation in the research. Moreover, the concept of study was approved by the university Ethical Committee and each participant, before starting survey agreed to take part in the study. The board of directors of the organization approved using HD system data set to the study.

intervention, such as training, a tool upgrade or any other support is needed to safeguard employee's job satisfaction and their well-being.