Applying Semantic Computing for Health Care Professionals: the Timing of Intervention is the Key for Successful Rehabilitation

Linda Karoliina Nieminen[1], Jari Juhani Vuori[2,3,4], Harri Juhani Ketamo[2], Markku Juhani Kankaanpää[1,4]
[1] Tampere University Hospital, Tampere, Finland
[2] Headai Ltd., Pori, Finland
[3] Arizona State University, USA
[4] Tampere University, Tampere, Finland
(linda.k.nieminen, markku.kankaanpaa)@pshp.fi, (jari.vuori, harri.ketamo)@headai.com

Abstract—**This study aims to apply Graph Machine Learning, a subset of artificial intelligence, in labeling electronic health records. The theoretical approach of the study stems from the studies of AI, machine learning, health policy, and physical medicine and rehabilitation.**

The data of chronic low back pain patients (n=93) were collected from electronic health records in form of a free text. The comparative analysis between the AI and medical expert was executed with the data of randomly selected patients (n=5).

The International Classification of Functioning, Disability, and Health was used as a scientifical frame to identify the factors affected by a patient's medical status. A medical expert identified the factors stated in the electronic health records. Data was analyzed and labeled with the graph (semantic networks) based machine learning engine, Headai Graphmind. Headai Graphmind automatically converted the findings to a readable map of factors, which are relevant concerning the timing of rehabilitation.

Headai Graphmind found 56% of identical factors in relation to the medical expert. In future studies, the analyses of mutuality between Headai Graphmind, the health care professional and the patient are crucial to set the right timing for rehabilitation.

I. Introduction

It has become tremendously obvious that the authority of artificial intelligence (AI) technologies will challenge the digital world of healthcare in the future. Advocates of critical scrutiny such as Bartlett [1] are taking a pioneering roles in Armageddon. They are convinced that the technologies of AI, big data, mobile and social media will even destroy democracy in the world of technology companies that scale fast but do not question anything. This raises a question in a global setting: is embraced "big tech" a real threat to effective, equitable, and personalized health service delivery [2]? Perhaps, but are we a bigger threat to ourselves than tech if we are not able to ensure that we control our machines, rather than the other way around [3]?

Admittedly, this trend is daunting to healthcare, but both AI data and evidence based current care guidelines are still in many diseases general and indefinite in terms of personalized medical decisions. Likewise, allegorically it makes sense to argue according to Karl Polanyi's [4] paradox: "me medical expert know more than I can tell and document". Specifically, medical experts know how to treat patients tacitly but cannot tell all of it to colleagues. Due to machine learning the investigation of the paradox proceeds rapidly [5]. On the other hand, AI and Machine Learning (ML) are umbrellas of thousands of algorithms, methods, and setups, all performing well in certain areas and poorly in other areas, thus making the selection of AI/ML algorithm difficult. For example, Deep Learning algorithms perform well in categorizing tasks when the task is well defined and the training material is big enough. Nevertheless, Deep Learning cannot perform well in cases where the task is ill-defined and requires humankind of reasoning to work with unknown factors. At present, as Panch et al [2] put it, the algorithms that feature prominently in the research literature are not very much, if at all, executable at the frontlines of clinical practice. On the other hand, Graph/Semantic Networks based Machine Learning seems to perform well [6], [7].

In this study, the semantic network based machine learning engine, Headai Graphmind (HGM), does, at its best, reasoning to supply best guess answers where formal procedural rules are unknown. The biggest difference between Headai Graphmind and common Graph Machine Learning is in the fundamentals of how Graph (detailed Semantic Network) is processed. HGM adds, modifies, and reasons according to conceptual learning theories [8] and Semantic Network is a storage structure for all the leaned data. I.e. HGM's Semantic Network include only processed data with explanations, not just nodes and edges. This is promising in a frame of patients with multiple morbidities and where the steps needed to achieve adequate health services are considered exceptionally highly complex.

The fact is that physicians and other health care professionals cannot be replaced with machines and robots as fast as the "optimistic hypers of AI" promise. That does not necessarily follow so far as machines can mainly be assistants in heavy lifting and logistics. Currently, approximately only 9 % of the worktime of an expert can be automated compared with 78 % of worktime in predictable physical work [9].

But cynical arguments against AI in healthcare are not very well articulated either. Conversely, according to the opponents, the real benefit will be realized in a continuous move in the managed health value chain from a labour-driven and technology-enabled model to a digital-driven and human-enabled one [10]. Most importantly, however, transforms toward personalized medicine do not happen only with simple decision support systems driven by AI and data. Likewise, in a rehabilitation process, successful personalization presupposes both right timing in the intervention and a specific profile produced most effortlessly by AI. Therefore, this paper highlights the importance of timing in the rehabilitation process in the frame of individual and societal, professional needs for rehabilitation (see Fig. 1).

A. Research design

In Fig. 1 is described the research design of the study. At the beginning of rehabilitation (T [1]) an individual's need for rehabilitation and its intensity is almost always higher than the need for rehabilitation by health care professional/ societal expertise (HP). Logically, human beings suffer first and society, in this case, health care professionals with medical experts start the treatment and rehabilitation much later along the clinical pathway.

Nevertheless, even if an active approach toward rehabilitation is taken at the beginning of the process, the intensity is quite low and close to non-existent, if professionals have adopted a "wait and see approach" [11](see Fig. 1, HP/ T[1]). This causes a dilemma in which individuals are not always taken care of at the right time, at the right place, and the right intensity. In some cases, time delay in rehabilitation leads to individuals' frustrations and other symptoms (e.g. psychosocial). The dilemma manifests itself usually in the phase of T [3]or later. In these cases, rehabilitation becomes more ineffective if individuals are not motivated to self-manage themselves anymore for many reasons (e.g. unemployment, isolation, depression, etc.) [12].

On this basis, it makes sense to believe that by applying Headai Graphmind (HGM) at T [1] in a very profiled way, health care professionals can obtain knowledge more quickly and thus, begin the rehabilitation planning (T [2] at the latest). This multidisciplinary knowledge is based on theoretical and scientific knowledge of machine learning, health policy, and physical medicine and rehabilitation.

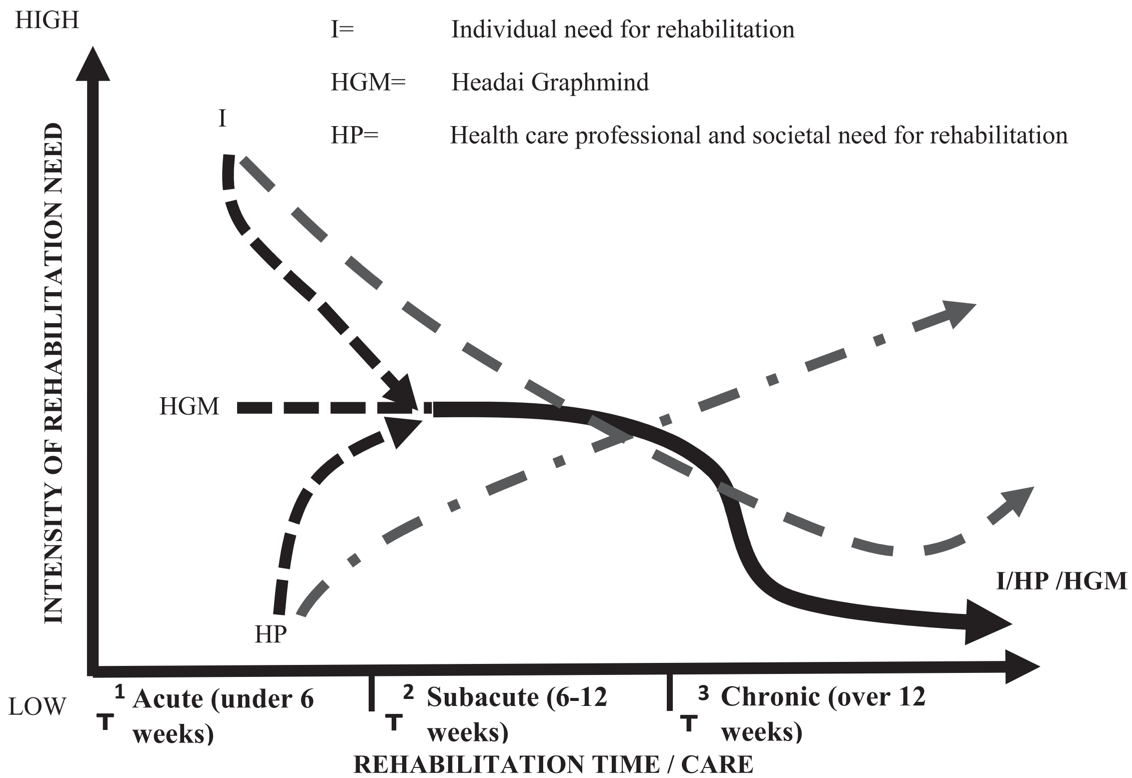

Fig. 1. The research design of rehabilitation processes embedded in patient involvement (I), Headai Graphmind (HGM) and medical expertise (HP)

B. Aims of the study

Low back pain (LBP) is the most burdensome medical condition worldwide in terms of disability [13] and when reaching the chronic stage (> 12 weeks), it comes with enormous individual and societal burden, let alone economic burden [14], [15]. There are numerous identified risk factors for developing LBP chronicity [16], but their early detection is often missed at the latest for the sake of inoperative clinical pathways. The timely recognition and targeted rehabilitation would help in the prevention of pain chronicity, but achieving such a plan can be a difficult task even for experienced health care professional. This study aims to apply a semantic network based machine learning engine, Headai Graphmind, in physical medicine and rehabilitation. The free text from electronic health records (EHR) is automatically converted into a readable map of factors, which are relevant to the timing of rehabilitation. To our knowledge, this form of method has not been used to date. The research questions are:

a) do the findings of a medical expert (ME) differ from Headai Graphmind's findings?

b) what is the potential impact of Headai Graphmind's findings on the timing of the rehabilitation process?

II. DATA AND METHODS

Overall, 1569 patient records were screened. Inclusion and exclusion criteria were used to form an eligible patient sample. The included patients were adults (18 to 65 years) suffering from chronic LBP (duration over 12 weeks). Specific reasons for LBP, such as nerve root disorders, or fractures of the spine were excluded. 93 patients fulfilled the criteria. These patients were suffering from non-specific LBP, where a specific biomechanical reason for the pain could not be identified. The data was collected in form of a free text from EHR between October 2019 and February 2021. The data was in the Finnish language. A longitudinal dataset of five patients was used for the result comparison (n=15 EHR notes) between medical expert and Headai Graphmind. The data was retrieved under a data transfer contract from Tampere University Hospital (Finland), where patients had been visiting the unit of Rehabilitation and Psychosocial support for their prolonged back pain. At the data collection time, there were 10 physicians (specialists and residents of physical medicine and rehabilitation) working in the unit, who were responsible for producing the EHR notes.

The International Classification of Functioning, Disability and Health (ICF) by World Health Organization (WHO) [17] was used as a scientifical frame to identify the factors affected by their medical status (Fig.2). A medical expert identified the factors stated in the patient's EHR produced by physicians. The ME was one of the physicians working in the unit of data retrieval, which minimized the misunderstanding of the data's medical content. The data was analyzed with HGM that imitates human reading and processing of the texts. HGM automatically converted the findings to a readable map of factors, which are relevant concerning the timing of rehabilitation.

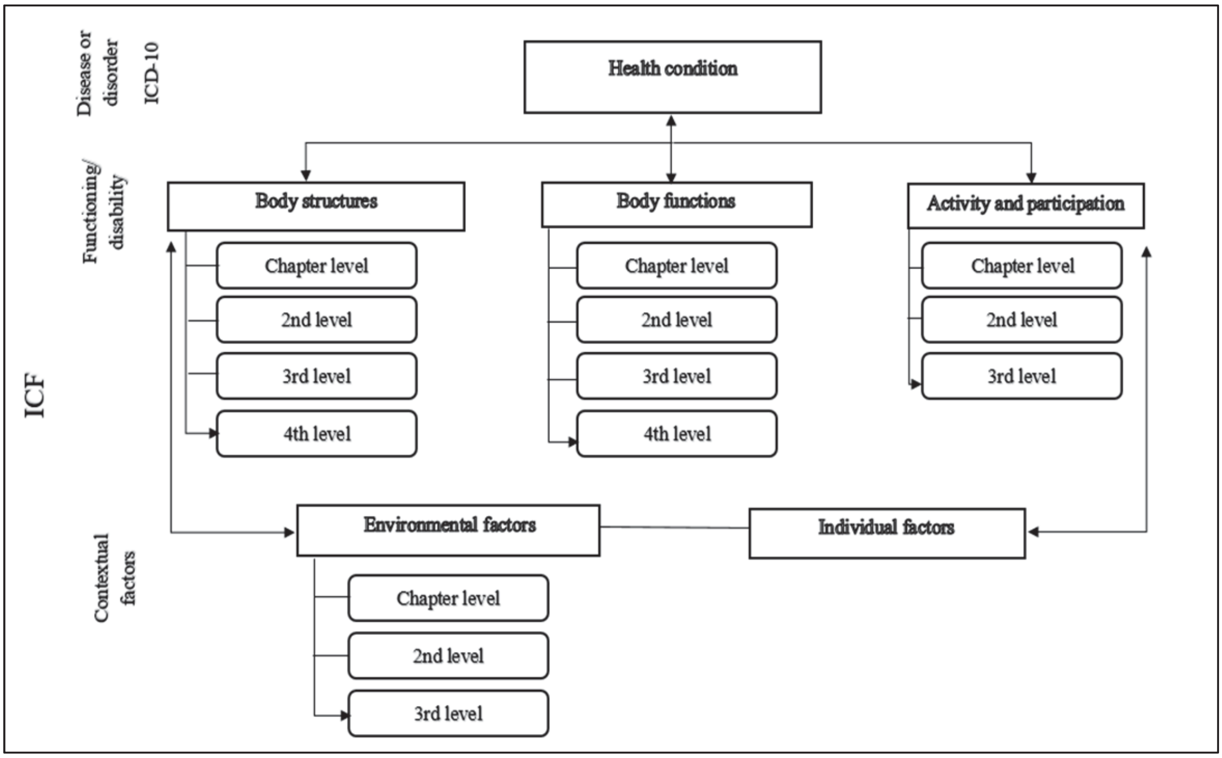

Fig. 2. The structure of International Classification of Functioning, Disability and Health (ICF). Different domains (e.g. body functions) are divided to three to four levels, which represent the ontology of the coding. For example in body functions, chapter level: b2 sensory functions and pain, 2nd level: b280 sensation of pain, 3rd level: b2801 pain in body part, and 4th level: b28013 pain in back. Individual factors are not coded because of the wide variability among cultures. Adopted from WHO Beginner's guide

The automated conversion of known LBP factors from EHR to ICF codes was executed with the cognitive text analyses, using natural language processing algorithms and semantic networks based machine learning.

Overall, 12 different setups were tested within the algorithm (Table I). The inputs included the ICF ontology, Medical Subject Headings (MeSH, a hierarchical vocabulary for life sciences), and medical experts' view of the language used by physicians in their texts.

The abilities of the algorithm to detect and convert the known factors were tested on a longitudinal dataset of five patients. A medical expert (ME) read the free texts of physicians (n=15 EHR notes) several times and searched for words, terms, and short sentences that could have a link to the ICF codes. The codes and the free text in question were listed as the reading went along so that the similar words and terms would be coded iteratively. The results of Headai Graphmind's different setups were compared to the codes converted by the ME. This comparative analysis was made by the same ME who had done the conversion. The matching results, the false conversions of the algorithm, and the results that were detected by the algorithm but not by the medical expert were listed.

III. Results

The dataset of each patient consisted of two to five notes in their EHR (n=15), that were entered during their period of treatment. The notes consist of referrals from primary health care, occupational health care, or private sector (n=4), physical appointments (n=6), contacts by phone call (n=4), or by letter (n=1).

The semantic networks based machine learning engine, HGM, and the ME found ICF domains (n=355) differently from five patients' EHRs (Table II). First, HGM and ME found partially the same codes in all the domains. In category 1 from body functions and structures domains, HGM found 68% of the ME's findings (20,3% of total findings), 76% from environmental domains (4,6% of total findings), but in the activity and participation domain the findings were in line in only 20% of the codes (2,5% of total findings). HGM found also codes that the ME did not. In category 3 there were 24 codes (6,8% of total findings) that HGM found better than the ME; these are explained by human error and different (but logical) interpretations of the text. For example, the ME coded "walking on different surfaces" and HGM "moving around in different locations", which both suited to the context. In addition, 61 code findings (17,2% of total findings) in all domains were not to be interpreted as correct findings after several appraisals of the data. Headai Graphmind missed 44% (n=76, 21,4% of total findings) of the findings of the ME's (category 1), and vice versa, ME missed 20% (n=24, 6,8% of total findings) of correct findings of HGM (category 3).

The most promising setups (Table I) were "ICF title" and "ICF real life fuzzy", where the relation of the correct findings to the false conversions was the highest. As far as the information and medical scientist's collaboration are concerned, it seems, that ICF codes combined with medical expert's view on the medical language will lead to intriguing results. The most important implication for AI scientists is that interdisciplinary research cooperation with medical experts should be encouraged. First, the prediction of risk scenarios in complex services might become easier in the future. Second, information scientists can innovate novel research designs with a better terminological understanding of medical counterparts. Finally, the research collaboration will lead to better applications also in other health care services.

TABLE I. The Explanation Of Different Setups. ICF= International Classification Of Functioning, Disability And Health; MESH= Medical Subject Headings (A Hierarchical Vocabulary For Life Sciences)

Setup abbreviation	**Setup name**	**Explanation of the input/algorithm-ontology configuration**
a	ICF title	The ontology of the ICF (title level)
b	ICF title fuzzy	The ontology of the ICF (title level) analyzed with fuzzy logic
c	ICF description	The ontology of the ICF (description level)
d	ICF description fuzzy	The ontology of the ICF (description level) analyzed with fuzzy logic
e	ICF real life	The ontology of the ICF was extended with the language used by physicians from ME point of view, e.g. b1342 onset of sleep= to fall asleep
f	ICF real life fuzzy	The ontology of ICF was extended with the language used by physicians from ME point of view, and analyzed with fuzzy logic
g	MESH-ICF	The ontology of ICF (title level) was extended with MeSH vocabulary
h	MESH-ICF fuzzy	The ontology of ICF (title level) was extended with MeSH vocabulary, and analyzed with fuzzy logic
i	MESH-ICF description	The ontology of ICF (description level) was extended with MeSH vocabulary
j	MESH-ICF description fuzzy	The ontology of ICF (description level) was extended with MeSH vocabulary, and analyzed with fuzzy logic
k	MESH-ICF real life	Setup e. was further extended with MeSH vocabulary
l	MESH-ICF real life fuzzy	Setup e. was further extended with MeSH vocabulary, and analyzed with fuzzy logic

TABLE II. HEADAI GRAPHMIND VS. THE MEDICAL EXPERT: THE COMPARISON OF THE CONVERSION FROM LOW BACK PAIN PATIENT'S ELECTRONIC HEALTH RECORDS TO ICF DOMAINS. *BODY STRUCTURE AND BODY FUNCTION CODES COMBINED

ICF DOMAINS (N=355)	FINDINGS OF HEADAI GRAPHMIND (HGM) AND THE MEDICAL EXPERT (ME)						
	(1) Graphmind found the same as ME		*(2) Graphmind found something*		*(3) Graphmind found better*		*Total*
	ME n (%)	**HGM n (%)**	**ME n(%)**	**HGM n (%)**	**ME n(%)**	**HGM n (%)**	**n (%)**
Body structures and functions*	106 (29.9)	72 (20.3)	N/A	34 (9.6)	N/A	10 (2.8)	222 (62.5)
Activity/ participation	46 (13.0)	9 (2.5)		14 (3.9)		9 (2.5)	78 (22.0)
Environmental factors	21 (5.9)	16 (4.5)		13 (3.7)		5 (1.4)	55 (15.5)
Total	173 (48.7)	97 (27.3)		61 (17.2)		24 (6.8)	**355 (100)**

IV. DISCUSSION

This paper introduces the principles of applying Semantic Network based Machine Learning engine Headai Graphmind to the framework of functioning, disability, and health (ICF), which can help our health care professionals in co-operation with the individuals plan the rehabilitation processes needed in a personalized healthcare fashion. In precise, the study highlights the correct timing in rehabilitation in minimizing the time of disability and waste of resources. The future key should be in the combination of knowledge of individuals, health care professionals, and advanced machine learning.

In Fig. 1. were described the potential shift from individual (I) or health professional (HP) driven planning and rehabilitation process to the more effective process constructed by the semantic fields of the texts habituated into reciprocal roles played by the individual, the health professional and HGM (I/HP/HGM) [18]. The described results of data analysis (Table II.) give arguments for the shift within the categories 2 and 3. The categories are highly promising for the inquiries of a new ontology of ICF domains. At its worst, health professionals only maintain, modify, and reconstruct the reality of unquestionable ICF domains. However, the category 2 findings go in line with the fuzzy logic and are not ontologically based on the ICF domains, therefore offering health professionals new ways to approach LBP patients in general. The findings of category 3 challenge the health professionals as well. Fictionally, HGM can define better the status of the patient following the ontology of ICF domains and could easily ask a physician: "you didn't change new sunglasses to see the whole picture and specific needs of individuals in the rehabilitation process, did you?" The applications of supervised machine learning (e.g. Deep Learning) in this case would have not produced the findings described in categories 2 and 3. Therefore, HGM promises a lot in a frame of risk analysis for preventive rehabilitation. In addition, these findings may pave the way for new interesting studies of fuzzy logic in risk analysis and scenarios, particularly in information sciences.

HGM reached identical conclusions in 56% of the ME's results. Different reasons for this can be that HGM was not yet learned well, or data was too limited and not rich enough for a more precise conversion. Further development of the most functional setups should lead to higher accuracy of HGM in relation to the medical experts and can even give discoveries on the individual's functioning and disability.

V. CONCLUSIONS

This paper underlines the importance to provide profiled knowledge of patients with managed machine learning to create a more effective rehabilitation process from the beginning. The first research question concerned the difference between a medical expert and the machine. Table II answered the question simply with the existence of categories 2 and 3. These categories remind readers of the classical learning curves that are more often non-linear than linear. The second research question was exploring the potentiality of HGM's effectiveness in the right timing of rehabilitation. Answering this question comprehensively is not possible with this pilot study. First, we must test HGM with a prospective and larger data, and conduct a study with a reasonable follow-up time. Second, the questions of timing and time need to be considered carefully in this kind of analysis. In the work of health professionals it is dynamic, and the issues of the right timing in treatments prior in many cases the question of duration of the treatments.

It seems promising that the Semantic Network based Machine Learning engine Headai Graphmind is capable to do conceptual reasoning in challenging domains. However, it must be highlighted that Headai Graphmind's performance was at its best in two cases: with training data based on ICF titles and with training data based on domain professional's short explanations of the ICF code written in professional language. When applying MeSH vocabulary or too generic definitions as training data, the results were not that good. Furthermore, this is nothing unexpected. In fact, this is aligned with earlier studies on semantic computing: the smaller the training data is, the more critical the quality of the data is, no matter what the algorithm is. Finally, semantic computing cannot solve ICF coding alone, but it can be exploited wisely by experienced health care professionals.

References

[1] J. Bartlett. *The people vs. tech. How the internet is killing democracy (and how we save it).* London: Ebury Press, 2018

[2] T. Panch, H. Mattie, L. Celi."The "inconvenient truth" about AI in health care". *Npj Digital Medicine,* 2019, 2:77, 1-2.

[3] G. Petriglieri. "Technology is not threatening our humanity — We are", *Harvard Business Review* October, 30, 2015.

[4] M. Polanyi. *The tacit dimension*. Page 4. Garden City, NY: Doubleday and Co., 1966.

[5] D.H. Autor. "Why are there sill so many jobs? The history and future of workplace automation", *Journal of Economic Perspectives.* 2015, Vol. 29: 3, pp. 3–30

[6] H. Ketamo, H. "Self-organizing content management with semantic neural networks", *Recent Advances in Neural Networks: Proceedings of the 10th WSEAS International Conference on Neural Networks (NN'09)*, Prague, Czech Republic, 23-25.3. 2009, pp.63-69.

[7] T. Gaudelet, B. Day, A.R. Jamasb, J. Soman, C. Regep, G. Liu, J.B.R. Hayter, R. Vickers, C. Roberts, J. Tang, D. Roblin, T.L. Blundell, M. Bronstein, J.P. Taylor-King. "Utilizing graph machine learning within drug discovery and development", *Briefings in Bioinformatics*, Volume 22, Issue 6*,* November 2021, bbab159, https://oi.org/10.1093/bib/bbab159

[8] S. Vosniadou. *Conceptual change approach and its re-framing*. In Vosniadou, S., Baltas, A. & Vamvakoussi, X., (Eds), *Re-framing the conceptual change approach in learning and instruction.* Oxford: Elsevier Press, 2007, pp. 1-15.

[9] A. Larson and R. Teigland,(2020). *The digital transformation of labor: automation, the gig economy and welfare.* New York: Routledge, 2020

[10] D. Kaul, H. Raju, B.K. Tripathy. (2022) "Deep Learning in Healthcare". In: Acharjya D.P., Mitra A., Zaman N. (eds) Deep Learning in Data Analytics. Studies in Big Data, vol 91. *Springer Nature Switzerland*. https://doi.org/10.1007/978-3-030-75855-4_6

[11] S.J. Linton, M. Nicholas, W. Shaw. "Why wait to address high-risk cases of acute low back pain? A comparison of stepped, stratified, and matched care". *Pain.* 2018;159(12):2437-2441. doi:10.1097/j.pain.0000000000001308

[12] S. Michie, M.M. van Stralen, R. West. "The behaviour change wheel: A new method for characterising and designing behaviour change interventions". *Implementation Science*. Vol 6:42, 2011. https://doi.org/10.1186/1748-5908-6-42

[13] GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. "Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016" [published correction appears in Lancet. 2017 Oct 28;390(10106):e38]. *Lancet*. 2017;390(10100):1211-1259. doi:10.1016/S0140-6736(17)32154-2

[14] J. Hartvigsen, M.J. Hancock, A. Kongsted, Q. Louw, M.L. Ferreira, S. Genevay, D. Hoy, J. Karppinen, G. Pransky, J. Sieper, R.J. Smeets, M. Underwood, & Lancet Low Back Pain Series Working Group. "What low back pain is and why we need to pay attention". *Lancet*. 2018;391(10137):2356-2367. doi:10.1016/S0140-6736(18)30480-X

[15] T.J. Marin, D. Van Eerd, E. Irvin, R. Couban, B.W. Koes, A. Malmivaara, M.V. van Tulder, & S.J. Kamper. "Multidisciplinary biopsychosocial rehabilitation for subacute low back pain". *Cochrane Database Syst Rev.* 2017;6(6):CD002193. Published 2017 Jun 28. doi:10.1002/14651858.CD002193.pub2

[16] L.K. Nieminen, L.M. Pyysalo, M.J. Kankaanpää. "Prognostic factors for pain chronicity in low back pain: a systematic review". *Pain Rep*. 2021;6(1):e919. Published 2021 Apr 1. doi:10.1097/PR9.0000000000000919

[17] World Health Organisation WHO website, ICF Beginner's Guide: Towards a common language for functioning, disability and health. 2002. Accessed: January 22, 2022. Webpage: https://www.who.int/publications/m/item/icf-beginner-s-guide-towards-a-common-language-for-functioning-disability-and-health

[18] P.L. Berger, and T. Luckmann. *The social construction of reality: a treatise in the sociology of knowledge.* Garden City, NY: Anchor Books, 1966