

Journal of Digital Science



ISSN 2686-8296

Volume 4 Issue 1

June 2022

© Institute of Certified Specialists

CONTENTS

An Empirical Examination of the Factors of Data Literacy	3
Ravi Nath, Joseph Kirby	
A conceptual framework for assessing information security management practices in selected universities in Uganda	21
Benjamin Ahimbisibwe, Peter Nabende	
Some Features of Social Structures and Institutions Transformation in the Digital Age	30
Artem Balyakin, Marina Nurbina, Sergey Taranenkov	
Geomatics and smart tools in Digital Land Resources Mapping and Sustainability of Coastal Agriculture, Egypt	43
Mohamed Zahran, Abd-Alla Gad	
Improving Business Processes by Applying the Kaizen Philosophy in a Macedonian Textile Company	56
Elizabeta Mitreva, Aneta Janeva	
On the fractal self-organization of the financial time series	71
Vladimir Hilarov	
Detectability of oncological diseases in the process of clinical examination of the adult population of Russia in 2013-2020	78
Olga Zakharchenko, Dina Terenteva, Irina Shikina	
Briefs in Assessing the Adequacy of Health Care Facilities' Fixed Assets ..	85
Tatiana Antipova, Alexander Zhelnin, Iuliia Zhelnina	

An Empirical Examination of the Factors of Data Literacy

Ravi Nath¹[0000-0002-3574-6262], Joseph Kirby²[0000-0002-8767-8409]

¹ Creighton University, Omaha, NE - USA, 68178

² Bellevue University, Omaha, NE - USA, 68005

https://doi.org/10.33847/2686-8296.4.1_1

Received 24.04.2022/Revised 30.04.2022/Accepted 22.05.2022/Published 12.06.2022

Abstract. To fully leverage the abundance of data and how data enhances decision-making, people must be data literate. Data literacy (DL) encompasses a set of interrelated skills in data management, data analysis, and the ability to interpret and communicate the results. Measuring an individual's DL level is an important first step toward designing and developing educational programs to improve one's DL skills. This paper considers a DL measurement scale referred to as the Global Data Literacy Benchmark survey and then explores the underlying constructs of this instrument. Data gathered from 311 university students across five universities in the United States is analyzed to identify and interpret the underlying factors of this DL scale. Also, the differences in DL scores among various subgroups of the students are investigated. The results show the existence of three DL factors. Also, the DL scores vary considerably among students depending upon the study areas and the comfort levels with data and analytics.

Keywords: Data Literacy, Factor Analysis, Global Data Literacy Benchmark survey.

1. Introduction

There is growing interest in measuring data literacy (DL). In 2020, according to some estimates, on average, every person on earth generated at least 1.7 Megabytes of data per second Bulao [1]. This turns out to be approximately 147 Gigabytes of data per person per day. The availability of data, reasonably low costs of data storage and advances in data analytics and Artificial Intelligence (AI) methodologies have enabled both public and private organizations to leverage this data and the resultant analytics and algorithms for the benefit of individuals, businesses, and communities [2].

Today, the term "big data" is commonly used to recognize the scale and speed at which data is generated and captured. Data and analytics present organizations with unique opportunities to use the insights provided by data to improve business processes and make better decisions. This is what many refer to as a "data-driven culture" [3-5]. Gauging the data-driven culture of an organization involves an assessment of the degree to which members of the organization (upper-level executives, middle-level managers, lower-level employees) make decisions based on insights gleaned from data [3]. And this is possible only if the workforce understands the basic concepts of data, data organization, analysis and interpretation, and data visualization. In other words, the workforce is data literate. In fact, Gartner [6] defines DL as "... the ability to read, write and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied, and the ability to describe the use-case application and resulting value." Thus, data literacy includes skills in not only data and analysis but also the ability to communicate the resulting insights (i. e, "the ability to tell the story").

DL, in many ways, is similar to financial literacy or health literacy. It is well established that financially literate people make better financial decisions [7]. And they understand the consequences of their actions on their financial outcomes.

Similarly, health literacy correlates positively with healthy living and results in better management of one's health [8, 9]. Therefore, to successfully navigate our data-rich, technology-driven, and ever-changing information society, it is essential to work towards a data-driven culture, i.e., for individuals to become "data literate." An initial step toward forming a data literate workforce is to assess the current status of their DL. Based on this information, then, appropriate training and professional development programs may be suggested to boost DL skills.

For this purpose, this paper examines the underlying constructs of the *Global Data Literacy Benchmark* survey instrument [10]. This is accomplished by factor analyzing the data collecting from a sample of university students in the United States. The resultant underlying factors of DL are then identified and interpreted. Further, the paper examines the DL scores for various subgroups of students and how they relate to their areas of study and other demographic variables.

In the following section, we review existing literature that has led to the importance of DL. Next, we present our findings from the examination of a commonly used DL instrument using factor analysis. Lastly, we provide conclusions and recommendations for further research.

2. Literature Review

Sitting at the base of the knowledge pyramid, data has always played a critical role in developing insight. Yet, the recent interest in data is peculiar and seems related to technology-enablers allowing for more advanced uses of data. The convergence of advances in data storage capabilities, low-cost computing power, and broad availability of advanced analytics algorithms are enabling organizations to derive value from data by leveraging techniques such as Machine Learning (ML) and Artificial Intelligence (AI) [11]. While the theoretical underpinnings of intelligent machines were outlined in the 1950s [12], it wasn't until just after the turn of the 21st century that technology had advanced to the point where these ideas could be realized through big data and analytics [13, 14]. These developments are fundamentally changing the roles of individuals, data, information, and computers and unlocking the potential of data [15]. Moreover, data provides the foundation across all phases of the evolution of analytics, from early decision support systems to, more recently, big data initiatives [16]. Organizations have taken note.

Recognizing data as a strategic asset, organizations have begun to create executive-level positions dealing with data and its role in creating tangible business value. Capital One – a financial firm based in the USA - signaled the importance of data as a critical organizational capability when it introduced the first Chief Data Officer (CDO) in 2003. Since then, additional organizations have followed Capital One's lead. Recognizing the value of data, in 2019, the US federal government passed a law mandating the role of a CDO in federal agencies [17]. The persistent growth in the CDO positions highlights the importance that businesses are attaching to data and, consequently, the role of DL as a critical skill in harnessing the full benefits of data and analytics. Growing interest exists among researchers as well.

Research papers on data analytics, machine learning, and artificial intelligence exhibit an exponential growth pattern. A search of Google Scholar shows 126 analytics-related publications in the year 2000 and increasing to 957 in 2010 and then jumping to 47,900 in 2020. This growth pattern serves as a proxy for the importance and currency that is being given to the study and applications of data analytics in creating value.

In light of the importance of data, the ability to properly analyze data and make sense of it, DL becomes an essential skill to possess [18]. A basic level of DL gives people the ability to analyze data intelligently and enhance their marketability in the workforce [19]. Also, it makes individuals more aware of deceptive statements and

information [20, 21], particularly when the information is data-derived. This, in turn, enables one to hold organizations and governments accountable [22] for false and deceptive information. Consequently, DL skills allow individuals to become more informed and thus fully participate in society [23-25]

Organizations also benefit from data-literate workers as data analytics, and their use is serving as a "disruptor" across many industries [26]. While many recognize the need to invest in talent development in analytics, the lack of basic DL skills among workers hinders the development of higher-level talent in data analytics [13]. Also, the lack of DL skills among college students or recent college graduates presents ongoing challenges for firms [27].

While there is consensus on the importance of DL as a necessary skill, there are a multitude of opinions and viewpoints on what it means to be data literate. This is very similar to the challenges that exist in defining business analytics, which continues to evolve as new technologies and capabilities emerge [28]. One factor driving many diverse definitions of DL is the fact that the need for DL spans different industries and academic disciplines, and each one has its own set of perspectives as to the skills a data literate person should possess. Thus, the multiple definitions of DL. Also, Corral [15] notes that DL rests at the center of other literacies, including computational, statistical, scientific, information, and media. In addition, DL also includes a focus on critical thinking and problem-solving [29, 30], the ability to argue with data or evaluate arguments [31, 32], to analyze data for meaning [33], to engage constructively in society with data [34], to leverage data as a part of a research or inquiry process [35], and to use data in an ethical manner [36]. All this makes the task of coming up with a single definition of DL that is acceptable to people in all disciplines and industries complicated.

D'Ignazio and Bhargava [35] propose a capability-based approach to DL focusing on "the ability to read, work with, analyze and argue with data." The motivation for this definition is guided by a practical approach to learning based on the complexity and flexibility provided by the tools associated with DL. In the same vein, [6] defines DL as "the ability to read, write and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied, and the ability to describe the use-case application and resulting value." This definition highlights the need to have a strong understanding of the provenance of the data and the ability to appropriately assess the accuracy of the data within the context of the technical environments in which the data was captured, cleaned, stored, and updated.

Yet another definition of DL [29] notes that becoming data-literate involves "the ability to consume for knowledge, produce coherently and think critically about data." They go on to state that being data literate further requires one to be competent in statistics and possess an ability to understand and work with large datasets, collect data from a variety of sources, and be able to interpret the resulting data. Further, at the more conceptual and societal outcome level, Bhargava, Deahl [34] describe DL as "the desire and ability to constructively engage in society through and about data." The authors decompose this definition and link the meaning of each word to either specific societal or technological capabilities inherently necessary to be considered Data Literate. Following the same logic, Bowen and Bartley [31] state that DL involves "understanding what good data and data analysis is so that you can make stronger arguments and better evaluate the arguments of others."

In addition to these perspectives regarding DL, other researchers have argued that discussion around DL is incomplete without the inclusion of concepts such as data quality [37], digital data infrastructures [38], data privacy, security, and safety [39], curiosity [40], algorithms and tools [11], ethical considerations [23, 25], domain knowledge and problem-solving disciplines [30].

In light of these broad and varying definitions of DL, the challenge of agreeing on a single definition that fits all contexts is understandable. In one well-referenced review of the literature published by Dalhousie University, Ridsdale, Rothwell [41] examined DL by collecting a set of best practices resulting in twenty-three (23) core competencies which are grouped into five broad categories: a conceptual framework, data collection, data management, data evaluation, and data application. This framework encompasses different DL perspectives, yet it is practical in providing guidance in developing relevant DL curricula for instructional and training purposes. Bonikowska, Sanmartin [42] summarize and discusses five different frameworks for DL as proposed by Datatothepeople.org [10], [41, 43-45]. While some of these frameworks are germane to measuring DL at the organization level, individual DL self-assessment instruments gauge skills such as data collection, manipulation, management, evaluation, decision-making, critical thinking, problem-solving capabilities, and data ethics. The organizational level models are designed to measure the degree to which data is used within ongoing operations and the extent to which the organizational culture is conducive to supporting DL.

For this study, we opted to use the Global Data Literacy Benchmark (GDLB) – a 25-item questionnaire for assessing an individual's DL levels [46]. This instrument is fairly comprehensive and it is inspired by the core DL competencies proposed by Ridsdale, Rothwell [41]. Also, it has been widely used in both academia and practice across the world. Details about this questionnaire are provided in the next section.

3. Data and Methodology

A questionnaire consisting of two sections was designed. Section I is based on the Global Data Literacy Benchmark (GDLB), which contains 25-items intended to establish a baseline of core DL skills [46]. Each item is designed to assess a capability related to a certain aspect of DL, and each item is posed as a question, "Which of these statements best describe you?" The respondent is then asked to choose one response from a set of options that range, in ascending order of the mastery of the capability, from "None of these describe me" to "I can teach and assist others (about a given capability)."

Out of the 25 items, there are 18 items each with seven (7) possible response options; two items with six (6) response options each; one item carries five (5) response options; two items each have four (4) options and another two items with two (2) possible responses. Responses are coded by assigning values ranging from a low of 0 ("none of these describe me...") to a maximum of 6 ("I can teach and assist.") for items with seven response options. For items with fewer response options, the maximum value remains at six, and the minimum assigned value is adjusted accordingly. Section II of the questionnaire collected demographic information about the participants (age, gender, years in school, field of study, extent of statistics and data analytics coursework, and comfort levels with data). A copy of the questionnaire appears in the Appendix.

Data were collected from students from five universities in the United States during the period from September through November 2020. Participants included both graduate and undergraduate students across different disciplines. Students were requested to complete the questionnaires either online or on paper. They were informed that the purpose of the survey was to collect information about DL and that their responses would be confidential, and only the aggregated results would be reported. A total of 315 surveys were collected, of which 311 were usable. The authors collected the survey information and analyzed the data using R-Script, the results of which are presented in the following tables. And all tables are generated by the authors.

Table 1 lists information regarding the respondents and their self-reported mastery of the basic DL skills. Nearly 70% of the respondents are males; most (85%) are under the age of 24 years; about 60% of them are in their third and fourth year of university studies, and only 9% are Masters' degree students. A majority seem somewhat or very comfortable with formal and informal data analytics and statistics (formal analytics: 64.3%; informal analytics: 62.7%). Further, about 83% of the respondents find their work, course work, or life moderately, highly, or extremely quantitative. Next, about 70% have taken at least one course in statistics and data analytics during their undergraduate studies. Finally, when asked to provide the level of their DL skills, nine out of ten rated themselves as average or above.

Table 1. Respondent Profiles (n=311)

Gender	Male	69.45%
	Female	26.05%
	Other	4.50%
Age	19 - 24	84.57%
	25 - 44	12.54%
	45 and older	2.57%
Year in School	First year	23.79%
	2nd year	8.04%
	3rd year	39.23%
	4th year	19.94%
	Masters	9.00%

Question	Rating	Percent*
Comfort level with formal Data Analytics and Statistics_(e.g., data, chance, probability, graphs, charts, etc.)	Very comfortable	21.9%
	Somewhat comfortable	63.3%
	Somewhat uncomfortable	11.3%
	Very uncomfortable	3.5%
Comfort level with informal Analytics and Statistics (e.g. reading and interpreting tables and graphs that use rates and percentages)	Very comfortable	23.8%
	Somewhat comfortable	59.8%
	Somewhat uncomfortable	13.5%
	Very uncomfortable	2.9%
Quantitative nature of your work, area of study or daily life	Extremely quantitative	7.1%
	Highly quantitative	36.3%
	Moderately quantitative	32.2%
	Minimally quantitative	19.9%
	Generally non-quantitative	4.5%
Number of undergraduate data analytics and statistics courses completed	None	29.2%
	One	33.7%
	Two	21.2%
	Three	8.7%
	Four or more	7.4%
Self-rated data literacy skill level	Extremely low	2.3%
	Low	7.8%
	Average	41.4%
	High	34.9%
	Extremely high	13.4%

Next, to assess the intensity of the responses for the twenty-five DL items, means, and standard deviations are calculated and are displayed in Table 2. Note that of all the items, items I3, I11, and I25 possess the highest mean values (4.10, 4.04, and 4.08, respectively). This shows that the respondents are fairly comfortable with identifying data sources, sharing data with others, and using data from myriad sources in decision-making. On the other hand, the three lowest-ranking items are I9, I10, and I12, with means of 2.24, 2.17, and 2.29, respectively. These items deal with deciding whether the data are clean and the best approaches to cleansing it and creating metadata.

Table 2. Item Mean and Standard Deviation (SD)

Item	Survey Question	Mean	SD
I1	Open and use datasets	3.22	1.59
I2	Browse and access various datasets	3.52	1.38
I3	Identify useful data from a variety of sources	4.1	1.19
I4	Identify whether data provided is trustworthy	3.18	1.61
I5	Data provided contains errors or problems	2.79	1.62
I6	Collect simple data provided to me	3.12	1.47
I7	Navigate folders, directories to find data	3.22	1.49
I8	Store and organize data in existing organizational structures	3.05	1.47
I9	Determine if data is clean using frameworks provided	2.24	1.58
I10	Determine the best method to clean data	2.17	1.51
I11	Make data available for others to use	4.04	1.73
I12	Create metadata associated with data sources	2.29	2.59
I13	Convert data sources to different formats as required	3.87	2.36
I14	Describe the analysis that needs performing	2.78	1.59
I15	Use simple tools provided to undertake analysis	2.78	1.55
I16	Ask and answer basic questions with data	3.09	1.42
I17	Read and understand simple tables, charts and graphs	3.89	1.54
I18	Find key points of interest in data	3.35	1.56
I19	Use visual methods and tools to understand and explore data	3.36	1.55
I20	Identify issues and problems in practical situations using data	3.15	1.56
I21	Create simple tables and charts to visually present data	3.31	1.64
I22	Verbally describe key points of interest in data	3.18	1.52
I23	Use data provided to support decision-making process	3.48	1.40
I24	Use data provided to inform decision-making process	3.67	1.30
I25	Use a range of data sources to evaluate a variety of decisions actions	4.08	1.51

4. Results

To identify the underlying factors of the DL, exploratory factor analysis is performed on the 25 items. Factor analysis is a statistical technique to identify constructs or factors using a set of data [47]. Based on a threshold of eigenvalue greater than 1 [48], this analysis identified three (3) factors and these factors accounted for 63.7% of the variation, which is a reasonable amount. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy of .948 is sufficiently close to 1, indicating the factor analysis is expected to produce distinct and reliable factors [49].

To have meaningful and interpretable factors, the initial factors were subjected to the Varimax rotation. Table 3 displays only those factor loadings that exceed 0.40. Note that the items bonding with Factor I, in general, deal with data analysis,

visualization, evaluation, and decision making. Thus, this factor is named *Data Analysis, Storytelling, and Decision-Making*. Factor II involves items related to working with a variety of data sources, storage, trustworthiness, and data organization. This factor is labeled *Data Acquisition, Navigation, and Quality Assessment*. Finally, items loading on Factor III involve data cleansing, enrichment, and conversion for use; this factor is termed *Data Wrangling*. Note that several items (I6, I7, I8, I14, and I15) bond with more than one factor. In such cases, either the item is assigned to the factor with the highest loading or bonded with that factor where it makes more conceptual sense. For example, as items I7 (Navigate folders, directories to find data) and I8 (Store and organize data in existing organizational structures) both involve data and its organization, they are assigned to Factor II (*Data Acquisition, Navigation, and Quality Assessment*) instead of Factor I even though the loadings are higher for the latter. Similarly, I15 (Use simple tools provided to undertake analysis) is bonded with Factor I as opposed to Factor III since it makes more sense to assign I15 to factor I. As a consequence of this process, the three factors and the associated items are as follows:

- Factor I: *Data Analysis, Storytelling, and Decision-Making* – (I14 – I25)
- Factor II: *Data Acquisition, Navigation, and Quality Assessment* – (I1 – I8)
- Factor III: *Data Wrangling* – (I9 – I13)

Table 3. Factor Loadings

Items	Factor Loadings		
	1	2	3
I1		0.680	
I2		0.671	
I3		0.711	
I4		0.712	
I5		0.682	
I6	0.492	0.585	
I7	0.576	0.443	
I8	0.434	0.415	
I9			0.700
I10			0.725
I11			0.587
I12			0.737
I13			0.639
I14	0.526		0.508
I15	0.477		0.551
I16	0.617		
I17	0.830		
I18	0.791		
I19	0.786		
I20	0.681		
I21	0.825		
I22	0.751		
I23	0.780		
I24	0.637		
I25	0.494		

* Only loadings greater than .40 are shown

An aggregate DL score was computed by averaging the 25-item scores. Factor scores were calculated by averaging the item scores that comprise the factor for each respondent. These three factors encompass all competencies needed to progress from raw data to developing insight for decision-making. To further explore the factors and their associations with other variables, the survey asked each respondent to self-rate their DL skill level by using a 5-point scale (1 = Extremely Low; 5= Extremely High).

Next, to determine the levels of associations between the self-rated DL and the 25-item instrument-based DL and the three derived DL factors, correlation coefficients were calculated. Table 4 displays the correlation coefficients. All correlations are found to be statistically significant at the $p < .01$ level confirming the internal consistency of the 25-item instrument as a measure of DL.

Table 4. Correlation between Self-rated DL, Factors, and 25-item DL

	F1	F2	F3	25-item DL
Self-rated DL	.47 ***	.39 ***	.38 ***	.46 ***

*** significant at the .01 level

Table 5 shows the score means and standard deviations (SD) for the three factors of DL, 25-item DL, and the self-rated DL when broken down by the number of analytics-focused courses taken and the academic level of the student, and whether they are undergraduates or graduate students. As one would expect, students completing higher number of analytics-focused courses have higher average DL scores across all dimensions. This indicates that additional coursework in analytics-oriented topics leads to improved DL skills.

Table 5. Mean (SD) of Factors, 25-item DL, and Self-rated DL Assessments by # of Analytics-focused Courses, Academic Level, and Graduate vs. Undergraduate Students

# of Courses	F1	F2	F3	25-item DL	Self-rated DL
None (n = 91)	3.03(1.15)	3.01(1.09)	2.57(1.36)	2.93(1.04)	3.36(0.96)
1-2 (n = 170)	3.31(1.27)	3.23(1.12)	2.82(1.53)	3.19(1.16)	3.43(0.82)
3 or more (n = 50)	3.89(1.20)	3.90(1.18)	3.70(1.42)	3.85(1.16)	3.97(0.92)
Academic Level					
First Year (n = 74)	3.16(1.19)	3.12(1.08)	2.64(1.16)	3.04(1.09)	3.40(0.84)
Second Year (n = 25)	3.34(1.06)	3.30(0.96)	2.91(1.42)	3.24(1.00)	3.76(0.91)
Third Year (n = 122)	3.23(1.24)	3.21(1.13)	3.21(1.13)	3.17(1.14)	3.50(0.91)
Fourth Year (n = 62)	3.41(1.24)	3.24(1.17)	2.85(1.56)	3.20(1.15)	3.40(0.89)
Masters (n = 28)	3.96(1.51)	3.99(1.37)	3.74(1.57)	3.92(1.40)	3.75(0.87)
Undergraduate vs Graduate					
Undergraduate (n = 283)	3.27(1.21)	3.21(1.11)	2.82(1.56)	3.15(1.11)	3.47(0.90)
Masters (n = 28)	3.96(1.51)	3.99(1.37)	3.74(1.57)	3.92(1.40)	3.75(0.87)
All Respondents					
All (n = 311)	3.33(1.25)	3.28(1.15)	2.90(1.50)	3.22(1.16)	3.50(0.90)

Next, when the data are examined by the student's academic level, generally, there are improvements in the mean DL scores with some exceptions. For instance, for F3, the mean for the fourth-year students (2.85) is significantly lower than that of the third-year students (3.21). This may be due to the chosen fields of study as well as the nature of the courses taken by the fourth-year students. Further, graduate students have higher mean DL scores than undergraduate students across all dimensions.

When the data are aggregated for all students, factor F3 has the lowest mean score (2.90) when compared to F1 (3.33) and F2 (3.28). Pairwise comparison tests reveal no significant difference between factors F1 and F2, but F3 is found to have a significantly lower mean than those of F1 and F2 ($p < .001$). Comparisons Mean performed with parametric t-test and non-parametric Wilcoxon Sign Rank tests provided similar results across all pairs. It is also worth noting that across all groupings, F3 produces the lowest scores of the three factors, indicating that data wrangling skills are lacking among the respondents compared to other DL skills. Also, the mean value of the self-rated DL is higher than the 25-item DL mean (3.49 versus 3.22), suggesting individuals tend to overrate their capabilities.

Students were asked to rate their comfort levels in dealing with formal analytics (data, probability, graphs, charts, etc.) and informal analytics (reading and interpreting tables and graphs that use rates and percentages) using a 4-point scale (1 = very uncomfortable; 4 = very comfortable). Also, they were prompted to describe how quantitative their area of study or work or daily life was using a 5-descriptors coded as follows: 5: Extremely quantitative; 4: Highly quantitative; 3: Moderately quantitative; 2: Minimally quantitative; 1: Generally non-quantitative.

Table 6 summarizes the results for comfort levels and the nature of the work in their field of study. For comfort levels, two groups are formed – Uncomfortable and Comfortable, and for the quantitative nature of work, three groups are formed- Generally Non-Quantitative, Moderately Quantitative, and Highly Quantitative. The mean scores for those who were comfortable with formal and informal analytics are found to be significantly higher across all dimensions (F1, F2, F3, 25-item DL, and self-rated DL) when compared to the mean scores of uncomfortable respondents ($p < .01$). Also, not surprisingly, when the nature of work or study of respondents was non-quantitative, the mean scores for F1, F2, F3, and the 25-item DL are statistically significantly lower in comparison to those involved in highly quantitative areas ($p < .05$), and the mean scores for F3 were also significantly higher for those involved in highly quantitative work in comparison to those in moderately quantitative work ($p < .05$). Interestingly, a pairwise comparison of the self-rated DL score revealed no differences between all levels of quantitative nature of work, consistent with other results suggesting that individuals overstate their self-rated DL abilities.

Table 6. Mean (SD) of Factors, 25-item DL and Self-rated DL Assessments by Comfort with Formal and Informal Analytics, and Quantitative Nature of Work or Study

Comfort with Formal Analytics	F1	F2	F3	25-item DL	Self-rated DL
Uncomfortable (n = 45)	2.25(1.38)	2.36(1.26)	2.02(1.62)	2.22(1.25)	3.02(1.13)
Comfortable (n = 265)	3.52(1.13)	3.44(1.05)	3.06(1.42)	3.39(1.04)	3.58(.83)
Comfort with Informal Analytics					
Uncomfortable (n = 50)	2.46(1.37)	2.52(1.22)	2.24(1.73)	2.42(1.27)	2.98(1.10)
Comfortable (n = 260)	3.50(1.15)	3.43(1.07)	3.04(1.42)	3.38(1.06)	3.60(.82)
Nature of Work or Study					
Generally Non-Quantitative (n = 75)	3.04(1.36)	3.00(1.25)	2.67(1.56)	2.95(1.24)	3.42(.98)
Moderately Quantitative (n = 100)	3.27(1.20)	3.26(1.07)	2.66(1.55)	3.13(1.14)	3.41(.80)
Highly Quantitative (n = 135)	3.53(1.20)	3.44(1.14)	3.20(1.40)	3.42(1.10)	3.62(.91)

Table 7 shows the mean DL scores broken down by comfort levels with formal and informal analytics and the completed number of analytics courses. Analysis of the scores for comfort with formal analytics shows students who have taken three or more analytics courses report significantly higher values than those who have taken 1-2 analytics courses ($p < .05$) and those who have taken no courses ($p < .01$); no differences were noted between students who took no analytics courses and those who took 1-2 analytics courses. Further, the scores for comfort with informal analytics show students who have taken three or more analytics courses report significantly higher scores than those who have taken 1-2 analytics courses ($p < .05$) and those who have taken no courses ($p < .001$), and students who took 1-2 analytics courses report significantly higher scores than those who have taken no analytics courses ($p < .05$). These findings suggest that greater exposure to analytics materials results in greater confidence in formal and informal analytics.

Table 7. Comfort with Formal and Informal Analytics by # of Analytics Courses

# of Courses	Comfort with Formal Analytics	Comfort with Informal Analytics
None (n = 91)	2.91(.77)	2.86(.71)
1-2 (n = 170)	3.04(.61)	3.08(.65)
3 or more (n = 50)	3.28(.67)	3.30(.71)

5. Summary and Conclusions

With the ascendancy of data-enabled approaches for decision making and the trend toward data democratization – where workers are empowered to perform their own data analysis and develop algorithms germane to their work – gauging the levels of workers' DL becomes important. To this end, this paper examines and then identifies the underlying factors for a DL instrument developed by DataToThePeople.org. By analyzing data collected from 311 university students, three factors are identified. The first factor involves skills related to data analysis and subsequent decisions based upon the analysis of the data. The second factor entails the basic abilities necessary to acquire data from multiple sources and then organize the data for analysis. The third factor deals with data preparation skills such as data cleaning, handling missing values, enriching data, and identifying data inconsistencies. It is worth noting here that all basic skills in the data value chain – from data acquisition to data cleansing to enriching data to data analysis to final decision making – are included in the inventory of DL skills.

An analysis of the factor scores revealed that students scored low on the factor dealing with data preparation and cleansing skills compared to their scores on the other two factors – data acquisition abilities and data analysis acumen. Considering that most organizations deal with the reality of less than pristine data sources, this finding provides critical insight for universities seeking to prepare students for the reality of their real-world work environments. Particular attention should focus on developing skills and training for identifying whether data needs cleaning and then the best approaches to clean it. One way to accomplish this would be for instructors to start with messy data and then show students the process of sanitizing the data sets before they are ready for analysis. Typically, in most analytics and statistics courses, students are given a clean and pristine dataset and then shown how to perform different analyses. And with this approach, students do not learn the essential skills of data cleansing, and they may even not be aware of this important aspect of data analytics.

Further analysis showed that students with three or more data-intensive courses had significantly higher average scores for a comfort level with analytics and the 25-item

DL than those with only one or two courses. Thus, it is plausible that there is a minimum threshold of exposure to data and analytics that is necessary to feel comfortable with data and analytics. Similarly, students involved in highly quantitative work had significantly higher average scores for a comfort level with analytics and the 25-item DL than those involved in generally non-quantitative work.

There are several limitations to this study. First, data for this research were collected from university students, and thus conclusions may not be fully applicable to other groups such as business workers and the general population. Second, some respondents found the DL instrument onerous to complete as it consists of twenty-five items, with each item being very wordy. Therefore, a revised shorter version of the instrument that simplifies the structure of the questions is needed as this will make the completion of the instrument less time consuming. Also, this can lead to a higher response rate.

For future research, we suggest replicating this study in corporate settings to examine similarities and differences between the students' and workers' DL skills. Also, measuring the baseline DL skills would be instrumental in designing training, professional development programs, and coursework for workers and students to boost their DL skills.

Acknowledgments

The authors thank Ms. Jane Crofts, founder of Data To The People (datatothepeople.org), for allowing us to use this instrument for this research. We also want to thank the students for participating in this study.

References

1. Bulao J. How much data is created every day in 2021? 2021 [Available from: <https://techjury.net/blog/how-much-data-is-created-every-day/#gref>].
2. Gupta A, Deokar A, Iyer L, Sharda R, Schrader D. Big data & analytics for societal impact: Recent research and trends. *Information Systems Frontiers*. 2018;20(2):185-94. <https://link.springer.com/article/10.1007/s10796-018-9846-7>
3. Gupta M, George JF. Toward the development of a big data analytics capability. *Information & Management*. 2016;53(8):1049-64. <https://www.sciencedirect.com/science/article/abs/pii/S0378720616300787>
4. McAfee A, Brynjolfsson E, Davenport TH, Patil D, Barton D. Big data: the management revolution. *Harvard Business Review*. 2012;90(10):60-8. <http://tarjomefa.com/wp-content/uploads/2017/04/6539-English-TarjomeFa-1.pdf>
5. Ross JW, Beath CM, Quaadgras A. You may not need big data after all. *Harvard business review*. 2013;91(12):90-+. <https://hbr.org/2013/12/you-may-not-need-big-data-after-all>
6. Gartner. Gartner Glossary - Data Literacy. 2018. <https://www.gartner.com/en/information-technology/glossary/data-literacy>
7. Costa VM, Teixeira NADS, Santos AC, Santos E. When more is less in financial decision-making: financial literacy magnifies framing effects. *Psychological Research*. 2020:1-11. <https://link.springer.com/article/10.1007/s00426-020-01372-7>
8. Lange M, Löwe A, Stassen G, Schaller A. Health literacy, health status and health behaviors of German students—study protocol for the “Healthy Habits” cohort study. *BMC Public Health*. 2021;21(1):1-10. <https://bmcpublihealth.biomedcentral.com/articles/10.1186/s12889-021-11542-w>
9. Netemeyer RG, Dobolyi DG, Abbasi A, Clifford G, Taylor H. Health literacy, health numeracy, and Trust in Doctor: effects on key patient health outcomes. *Journal of Consumer Affairs*. 2020;54(1):3-42. <https://onlinelibrary.wiley.com/doi/abs/10.1111/joca.12267#:~:text=We%20found%20that%20functional%20and,physician%20consultations%20and%20lower%20SWB.>
10. Datatothepeople.org. We're Data To The People 2018 <https://www.datatothepeople.org/>
11. Alpaydin E. *Machine Learning. The New AI*. Cambridge, MA: MIT Press; 2016. <https://mitpress.mit.edu/books/machine-learning>
12. Buchanan BG. A (very) brief history of artificial intelligence. *Ai Magazine*. 2005;26(4):53 <https://ojs.aaai.org/index.php/aimagazine/article/view/1848>

13. Davenport TH. From analytics to artificial intelligence. *Journal of Business Analytics*. 2018;1(2):73-80. <https://www.tandfonline.com/doi/full/10.1080/2573234X.2018.1543535>
14. Chen H, Chiang RHL, Storey VC. Business intelligence and analytics: from big data to big impact. (Special Issue: Business Intelligence Research)(Essay). *MIS Quarterly*. 2012;36(4):1165-88. <https://www.jstor.org/stable/41703503>
15. Corral S. Repositioning Data Literacy as a Mission-Critical Competence. 2019. <http://d-scholarship.pitt.edu/36975/>
16. Delen D, Ram S. Research challenges and opportunities in business analytics. *Journal of Business Analytics*. 2018;1(1):2-12. <https://www.tandfonline.com/doi/full/10.1080/2573234X.2018.1507324>
17. Shibu S. New law mandates chief data officers at federal agencies. *govloop*. 2019. <https://www.govloop.com/new-law-mandates-chief-data-officers-at-federal-agencies/>
18. McKinsey. Hal Varian on how the Web challenges managers. McKinsey Technology, Media & Telecommunications. 2009. <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/hal-varian-on-how-the-web-challenges-managers>
19. Brynjolfsson E, Mitchell T. What can machine learning do? Workforce implications. *Science*. 2017;358(6370):1530-4. <https://www.science.org/doi/10.1126/science.aap8062>
20. Bond Jr CF, DePaulo BM. Accuracy of deception judgments. *Personality and social psychology Review*. 2006;10(3):214-34. https://journals.sagepub.com/doi/10.1207/s15327957pspr1003_2
21. George JF, Carlson JR, Valacich JS. Media selection as a strategic component of communication. *MIS Quarterly*. 2013:1233-51. <https://www.jstor.org/stable/43825789>
22. Data-Revolution-Group. Accessible Data: open data, accountability and data literacy 2014 [Available from: <https://www.undatarevolution.org/data-revolution/>]
23. Letouze E. SHOULD 'DATA LITERACY' BE PROMOTED? 2016 [Available from: <https://unstats.un.org/unsd/undataforum/should-data-literacy-be-promoted/index.html>].
24. Montoya S. Defining Literacy. United Nations Educational, Scientific and Cultural Organization (UNESCO) October. 2018;17. https://qaml.uis.unesco.org/wp-content/uploads/sites/2/2018/12/4.6.1_07_4.6-defining-literacy.pdf
25. UNESCO. UNESCO - Literacy Website 2019 [Available from: <https://en.unesco.org/themes/literacy>].
26. Columbus L. 84% of enterprises see Big Data Analytics changing their industries' competitive landscapes In the next year 2014; (Journal, Electronic):[August 2019 p.]. Available from: <https://www.forbes.com/sites/louiscolombus/2014/10/19/84-of-enterprises-see-big-data-analytics-changing-their-industries-competitive-landscapes-in-the-next-year/#6a5aa2317de1>.
27. Pothier WG, Condon PB. Towards data literacy competencies: Business students, workforce needs, and the role of the librarian. *Journal of Business & Finance Librarianship*. 2019:1-24. <https://www.tandfonline.com/doi/abs/10.1080/08963568.2019.1680189?journalCode=wbfl20>
28. Power D, Heavin C, McDermott J, Daly M. Defining business analytics: an empirical approach. *Journal of Business Analytics*. 2018;1(1):40-53. <https://www.tandfonline.com/doi/full/10.1080/2573234X.2018.1507605>
29. Gray J, Chambers L, Bounegru L. The data journalism handbook: how journalists can use data to improve the news: " O'Reilly Media, Inc."; 2012. https://books.google.com/books/about/The_Data_Journalism_Handbook.html?id=W-sHp0ECqPUC
30. Mandinach EB, Gummer ES. Building educators' data literacy: Differing perspectives. *The Journal of Educational Research & Policy Studies*. 2013;13(2):1-5. http://www.waymandatause.com/wp-content/uploads/2014/04/JERPS_april-2013.pdf
31. Bowen M, Bartley A. The basics of data literacy: Helping your students (and you!) make sense of data: National Science Teachers Association; 2013. <https://my.nsta.org/resource/103665>
32. D'Ignazio C, Bhargava R, editors. Approaches to building big data literacy. Proceedings of the Bloomberg data for good exchange conference; 2015. <https://www.media.mit.edu/publications/approaches-to-building-big-data-literacy/>
33. Gemignani Z, Gemignani C, Galentino R, Schuermann P. Data fluency: Empowering your organization with effective data communication: John Wiley & Sons; 2014. <https://www.wiley.com/en-us/Data+Fluency%3A+Empowering+Your+Organization+with+Effective+Data+Communication-p-9781118851012>
34. Bhargava R, Deahl E, Letouzé E, Noonan A, Sangokoya D, Shoup N. Beyond data literacy: reinventing community engagement and empowerment in the age of data. 2015.

- <https://datapopalliance.org/item/beyond-data-literacy-reinventing-community-engagement-and-empowerment-in-the-age-of-data/>
35. D'Ignazio C, Bhargava R. DataBasic: Design principles, tools and activities for data literacy learners. The Journal of Community Informatics. 2016;12(3). <https://datapopalliance.org/item/beyond-data-literacy-reinventing-community-engagement-and-empowerment-in-the-age-of-data/>
36. Okamoto K. Introducing open government data. The Reference Librarian. 2017;58(2):111-23. <https://www.tandfonline.com/doi/abs/10.1080/02763877.2016.1199005>
37. Lawson RH, T.;Desroches, D. How to Embrace Data Analytics to Be Successful. Montvale, NJ: Institute of Management Accountants; 2019 February 2019. <https://www.imanet.org/insights-and-trends/technology-enablement/how-to-embrace-data-analytics-to-be-successful?ssopc=1>
38. Gray J, Gerlitz C, Bounegru L. Data infrastructure literacy. Big Data & Society. 2018;5(2):2053951718786316. <https://journals.sagepub.com/doi/10.1177/2053951718786316>
39. Markham AN. Taking data literacy to the streets: critical pedagogy in the public sphere. Qualitative Inquiry. 2020;26(2):227-37. <https://journals.sagepub.com/doi/abs/10.1177/1077800419859024?journalCode=qixa>
40. Dykes B. Data Curiosity: How To Cultivate An Inquisitive Workforce: Forbes; 2019 [Available from: <https://www.forbes.com/sites/brentdykes/2019/10/10/data-curiosity-how-to-cultivate-an-inquisitive-workforce/#7d72b1735471>.
41. Ridsdale C, Rothwell J, Smit M, Ali-Hassan H, Bliemel M, Irvine D, et al. Strategies and best practices for data literacy education: Knowledge synthesis report. 2015. <https://dalspace.library.dal.ca/handle/10222/64578>
42. Bonikowska A, Sanmartin C, Frenette M. Data Literacy: What It Is and How to Measure It in the Public Service. Statistics Canada; 2019. Contract No.: Report. <https://www150.statcan.gc.ca/n1/pub/11-633-x/11-633-x2019003-eng.htm>
43. Grillenberger A, Romeike R, editors. Developing a theoretically founded data literacy competency model. Proceedings of the 13th Workshop in Primary and Secondary Computing Education; 2018. <https://dl.acm.org/doi/10.1145/3265757.3265766>
44. Sternkopf H, Mueller RM, editors. Doing good with data: Development of a maturity model for data literacy in non-governmental organizations. Proceedings of the 51st Hawaii International Conference on System Sciences; 2018. <https://scholarspace.manoa.hawaii.edu/items/9323df5b-20e9-4478-99e5-ff24dea46f9d>
45. Wolff A, Gooch D, Montaner JJC, Rashid U, Kortuem G. Creating an understanding of data literacy for a data-driven society. The Journal of Community Informatics. 2016;12(3). <https://openjournals.uwaterloo.ca/index.php/JoCI/article/view/3275>
46. Crofts J. Global Data Literacy Benchmark. 2018. https://www.datatothepeople.org/gdlib?utm_term=&utm_campaign=Simple-dttp-campaign&utm_source=adwords&utm_medium=ppc&hsa_acc=4127976028&hsa_cam=11993952897&hsa_grp=113226491742&hsa_ad=489671633508&hsa_src=g&hsa_tgt=dsa-1021355669020&hsa_kw=&hsa_mt=&hsa_net=adwords&hsa_ver=3&qclid=CjwKCAjw4ayUBhA4EiwATWyBrTnsE4XIdnTqHBmH_F3BPyyTV3MOXszcvHqLBPGnJoFf32-hrRQPwxoC0hYQAvD_BwE
47. Kline P. An easy guide to factor analysis: Routledge; 2014. <https://www.routledge.com/An-Easy-Guide-to-Factor-Analysis/Kline/p/book/9780415094900>
48. Kaiser HF. An index of factorial simplicity. Psychometrika. 1974;39(1):31-6. <https://link.springer.com/article/10.1007/BF02291575>
49. Kaiser HF, Rice J. Little jiffy, mark IV. Educational and psychological measurement. 1974;34(1):111-7. <https://journals.sagepub.com/doi/10.1177/001316447403400115>

Appendix

Data Literacy Questionnaire [46]

Data literacy consists of skills to read, work with, analyze, and make decisions using data.

In the following, please check one and only one response that reflects your level of expertise. All your responses are confidential and for this research, they will be aggregated with other respondents' results.

1. Which of these statements best describe you?
 - o With guidance, I can open and use datasets provided to me.
 - o I can open and use datasets provided to me.

- I can identify and access data I need from a range of sources provided to me.
 - I can identify, locate and access data I need from a variety of sources.
 - I can assist others to identify and access data they need from a range of provided sources.
 - I can teach and assist others to identify, locate and access data they need from a variety of sources.
 - None of these describe me.
- 2.** Which of these statements best describe you?
- I can browse and access various datasets provided to me.
 - I can browse and access data from a range of locations.
 - I can search multiple sources to locate data related to my needs.
 - I can assist others to search a range of provided sources to locate data related to their needs.
 - I can teach and assist others to search multiple sources to locate data related to their needs.
 - None of these describe me.
- 3.** Which of these statements best describe you?
- I can identify useful data from a variety of sources provided to me.
 - I can locate and identify data from a variety of sources to meet my needs.
 - I can assist others to locate and identify data from a variety of provided sources to meet their needs.
 - I can teach and assist others to locate and identify data from a variety of sources to meet their needs.
 - None of these describe me.
- 4.** Which of these statements best describe you?
- With guidance, I can identify whether data provided to me is trustworthy.
 - I can identify whether data provided to me is trustworthy.
 - I can identify whether data I use is trustworthy.
 - I can identify whether data I use is trustworthy and locate alternate sources if required.
 - I can assist others to identify whether data they use is trustworthy.
 - I can teach and assist others to identify whether data is trustworthy and locate alternate sources if required.
 - None of these describe me.
- 5.** Which of these statements best describe you?
- With guidance, I can determine whether data provided to me contains any errors or problems.
 - I can determine whether data contains any errors or problems.
 - I can identify and take actions to correct any errors or problems in a range of data sources.
 - I can assist others to identify any errors or problems in a range of data sources.
 - I can teach and assist others to identify and take actions to correct any errors or problems in a range of data sources.
 - None of these describe me.
- 6.** Which of these statements best describe you?
- With guidance, I can collect simple data in a format provided to me.
 - I can collect simple data in a format provided to me.
 - I can collect data in simple and more complex forms.
 - I can collect data in a variety of forms to support my needs.
 - I can assist others to collect data in simple form to support their needs.
 - I can teach and assist others to collect data in a variety of forms to support their needs.
 - None of these describe me.
- 7.** Which of these statements best describe you?
- With guidance, I can navigate structured organization systems (including folders, directories and file naming) to find data I need.
 - I can navigate structured organization systems to find data I need.
 - I can navigate organization systems to assess and find data I need.
 - I can navigate internal and external systems and collections to locate and access data related to my needs.
 - I can assist others to navigate organizational systems and collections to locate and access data related to their needs.

○ I can teach and assist others to navigate internal and external systems and collections to locate and access data related to their needs.

○ None of these describe me.

8. Which of these statements best describe you?

○ With guidance, I can organize and store my own data in existing organizational structures.

○ I can organize and store my own data in existing organizational structures.

○ I can organize and store a range of data in existing organization systems.

○ I can organize and store data in internal and external systems.

○ I can assist others to organize and store data in internal systems.

○ I can teach and assist others to organize and store data in internal and external systems.

○ None of these describe me.

9. Which of these statements best describe you?

○ With guidance, I can determine if my data is 'clean' using frameworks provided to me.

○ I can determine if my data is 'clean' using frameworks provided to me.

○ I can assess and determine if my data is 'clean'.

○ I can determine if my data is 'clean' and take necessary actions to resolve any problems identified.

○ I can assist others to determine if their data is 'clean'.

○ I can teach and assist others to determine if their data is 'clean' and take necessary actions to resolve any problems identified.

○ None of these describe me.

10. Which of these statements best describe you?

○ With guidance, I can determine the best method to clean my data.

○ I can clean my data using a method provided to me.

○ I can determine the best method and use relevant tools to clean my data.

○ I can use a range of methods and tools to clean my data.

○ I can assist others to determine the best method and use relevant tools to clean their data.

○ I can teach and assist others to use a range of methods and tools to clean their data.

○ None of these describe me.

11. Which of these statements best describe you?

○ I can prepare and make my data available for others to use.

○ I can assist others to use data sources I have prepared and shared.

○ I can teach and assist others to use data sources I have prepared and shared.

○ None of these describe me.

12. Which of these statements best describe you?

○ I can create metadata to enrich data sources I prepare and share.

○ I can teach and assist others to understand metadata associated with data sources I have prepared.

○ None of these describe me.

13. Which of these statements best describe you?

○ I can convert data sources to a different format as required.

○ I can teach and assist others to use data in a variety of formats I have prepared.

○ None of these describe me.

14. Which of these statements best describe you?

○ With guidance, I can describe the analysis I need to perform.

○ I can develop a simple plan for analysis using frameworks provided to me.

○ I can develop a simple plan for my analysis.

○ I can develop a plan for analysis to better understand a range of problems.

○ I can assist others develop a plan for analysis.

○ I can teach and assist others develop a plan for analysis to better understand a range of problems.

○ None of these describe me.

15. Which of these statements best describe you?

○ With guidance, I can use simple tools provided to me to undertake analysis.

○ I can use simple tools provided to me to undertake my analysis.

○ I can select and use most appropriate tool for my analysis needs.

○ I can research and select the most appropriate tool for my analysis needs.

○ I can assist others to select and use the most appropriate tool to undertake their

analysis.

- I can teach and assist others to use a range of tools to undertake their analysis.
- None of these describe me.

16. Which of these statements best describe you?

- With guidance, I can ask and answer basic questions with data provided to me.
- With prompts, I can ask and answer a range of questions with provided datasets.
- I can ask and answer basic questions with data provided to me.
- I can ask and answer a range of questions using a variety of data sources.
- I can assist others to ask and answer basic questions using a variety of data sources.
- I can teach and assist others to ask and answer a range of questions using a variety of

data sources.

- None of these describe me.

17. Which of these statements best describe you?

With guidance, I can read and understand simple tables, charts and graphs provided to me.

- I can read and understand simple tables, charts and graphs provided to me.
- I can read and understand simple tables, charts and graphs.
- I can read and understand a range of tables, charts and graphs.
- I can assist others to read and understand a range of provided tables, charts and

graphs.

I can teach and assist others to read and understand a range of tables, charts and graphs.

- None of these describe me.

18. Which of these statements best describe you?

- With guidance, I can find key points of interest in data provided to me.
- I can find key points of interest in data provided to me.
- I can find key points of interest in a variety of data.
- I can identify points of interest and connections between a range of data sources.
- I can assist others to identify points of interest in provided data sources.
- I can teach and assist others to identify points of interest and connections between a

range of data sources.

- None of these describe me.

19. Which of these statements best describe you?

With guidance, I can use visual methods and tools to understand and explore data provided to me.

- I can use visual methods and tools to understand and explore data provided to me.
- I can use visual methods and tools to understand and explore data.
- I can use visual methods and tools to understand and explore a range of data sources.
- I can assist others to use visual methods and tools to understand and explore data.
- I can teach and assist others to use visual methods and tools to explore a range of data

sources.

- None of these describe me.

20. Which of these statements best describe you?

With guidance, I can identify issues and problems in practical situations using data provided to me.

- I can identify issues and problems in practical situations using data provided to me.
- I can identify and describe problems in practical situations using data provided to me.
- I can identify and describe problems in practical situations using a range of data

sources.

- I can assist others to identify problems in practical situations using provided data sources.
- I can teach and assist others to identify and describe problems in practical situations

using a range of data sources.

- None of these describe me.

21. Which of these statements best describe you?

With guidance, I can create simple tables and charts to visually present data provided to me.

- I can use templates to create meaningful tables and charts to visually present data provided to me.
- I can create meaningful tables and charts to visually present simple datasets.
- I can create meaningful tables and charts to visually present a range of data sources.

- I can assist others to create meaningful tables and charts to visually present simple data sources.
 - I can teach and assist others to create meaningful tables and charts to visually present a range of data sources.
 - None of these describe me.
- 22.** Which of these statements best describe you?
- With guidance, I can verbally describe key points of interest in data provided to me.
 - I can verbally describe key points of interest in data provided to me.
 - I can verbally describe simple datasets and data visualizations.
 - I can verbally describe a range of data sources and data visualizations.
 - I can assist others to verbally describe provided data sources and data visualizations.
 - I can teach and assist others to verbally describe a range of data sources and data visualizations.
 - None of these describe me.
- 23.** Which of these statements best describe you?
- With guidance, I can use data provided to me to support my decision-making process.
 - I can use data provided to me to inform my decision-making process.
 - I can use data to inform my decision-making process.
 - I can use data to analyze and inform my decision-making process.
 - I can assist others to use provided data sources to analyze and inform their decision-making process.
 - I can teach and assist others to use data to analyze and inform their decision-making process.
 - None of these describe me.
- 24.** Which of these statements best describe you?
- I can use data provided to me to inform my decision-making process.
 - I can identify key actions to take based on various data sources.
 - I can use a range of data sources to prepare and communicate actionable information.
 - I can assist others to use provided data sources to prepare and communicate actionable information.
 - I can teach and assist others to use a range of data sources to prepare and communicate actionable information.
 - None of these describe me.
- 25.** Which of these statements best describe you?
- I can use a range of data sources to evaluate a variety of decisions actions taken.
 - I can assist others to use provided data sources to evaluate a variety of decisions actions taken.
 - I can teach and assist others to use a range of data sources to evaluate a variety of decisions actions taken.
 - None of these describe me.

DEMOGRAPHICS

- 1.** What best describes how comfortable you are in dealing with formal Data Analytics and Statistics (e.g., data, chance, probability, graphs, charts, etc.)?
 - Very comfortable
 - Somewhat comfortable
 - Somewhat uncomfortable
 - Very uncomfortable
- 2.** What best describes how comfortable you are in dealing with informal Analytics and Statistics (e.g., reading and interpreting tables and graphs that use rates and percentages)?
 - Very comfortable
 - Somewhat comfortable
 - Somewhat uncomfortable
 - Very uncomfortable
- 3.** What best describes how quantitative your work, area of study or daily life is?
 - Extremely quantitative (e.g. mathematics, statistics, etc.)

- Highly quantitative (e.g., finance, econometrics, accounting, science, engineering, epidemiology, etc.)
- Moderately quantitative (e.g., psychology, sociology, MIS, market research, forecasting, etc.)
- Minimally quantitative (e.g., business management, education, journalism, health care, etc.)
- Generally non-quantitative (e.g., child-care, music, art, English, philosophy, etc.)

4. What best describes your fluency in English?

- English was a native language by primary school
- Became fluent in speaking and reading English after primary school
- Not yet fluent in speaking and reading English.

5. How many undergraduate data analytics and statistics courses have you completed?

- None
- One
- Two
- Three
- Four or more

6. Please rate your data literacy skill level using the following 5-point scale:

Extremely Low	Extremely High
1	5
2	4
3	3

7. What is your major (if declared)? If multiple majors, please list them.

8. Year in School: First year 2nd year 3rd year 4th year Masters

9. What is your age?

- 19 – 24
- 25 – 44
- 45 and older

10. What is your gender?

- Male
- Female
- Other _____(specify)

A conceptual framework for assessing information security management practices in selected universities in Uganda

Benjamin Ahimbisibwe ¹[0000-0002-0570-1274],
Peter Nabende ¹[0000-0003-2141-7940]

¹ Makerere University, Kampala, Uganda

https://doi.org/10.33847/2686-8296.4.1_2

Received 21.04.2022/Revised 03.05.2022/Accepted 22.05.2022/Published 12.06.2022

Abstract. The purpose of this paper is to present a conceptual framework for assessing managerial level information security practices, governance, and activities in selected university institutions in Uganda. Extant literature was drawn from existing information security management practices in different organizations. The proposed conceptual framework consisted of four manageable areas, namely, information security governance practices, information security practices, personnel management practices, and physical security practices. These areas are further subdivided into 25 categories that provide a formal checklist for assessing existing information security management practices in university institutions in Uganda.

Keywords: Conceptual framework, information security management practices, university institutions in Uganda.

1. Introduction

A conceptual framework describes the state of knowledge, identifies gaps in the phenomena under study, and outlines methodological underpinnings. Bilsky et al. [7] sum it up with two questions: why is the research important? and what contributions might a conceptual framework add to the body of knowledge? A conceptual framework is important because it helps to identify and clarify the central aspects of a study, connecting these aspects and their influence on the research being carried out [17].

Although several studies have been conducted in the information security management (ISM) domain, these studies did not provide a comprehensive guidance on specific practices to be followed by universities in the Ugandan context. None of them provides an adequate checklist of standards for assessing existing information security management practices (ISMPs) adopted by universities in Uganda. This paper proposes a conceptual framework that specifies such a checklist of standards. The conceptual framework was developed based on a generic inductive approach. The conceptual framework evolved during the study as new ideas, insights and knowledge concerned with the assessment of information security management practices got established [28]. At the beginning of the study, a tentative framework was constructed based on initially identified concepts and their relationships. Adjustments to the framework were then made as we gained a deeper understanding of the subject on information security management practices in organizations until the proposed conceptual framework that specifies 25 items for use as a checklist.

2. Literature Review

Several studies have been conducted to guide organizations to develop effective information security programs. The studies conducted in the Information Security Management (ISM) domain have focused mainly on individual aspects and the

application of different conceptual frameworks in various organizations. Current literature is limited in the explicitness to managerial level aspects of information security in organizations. This paper aimed at developing a conceptual framework that explicitly covers managerial level aspects of information security in selected universities in Uganda. In this section, we review some considerably related studies and discuss their relevance to the Ugandan context. We focus on the wider perspective of the ISM domain, information security management practices (ISMPs), and ISMPs at Kabale University and Bishop Barham University College both located in the western region of Uganda.

Information security is a multidimensional problem that is continually evolving and changing. Studies have evolved from approaching information security in a one-sided technical perspective to a managerial perspective with the view to providing a holistic and comprehensive approach [31]. A shared characteristic of existing studies is the aim to provide guidance to organizations by identifying activities that constitute a security program. The different studies use different terminologies such as processes, principles, practices, activity or theory to present their advice. Furthermore, existing studies use different levels of detail and analysis in describing the activities that organizations should implement. It is from the review of these studies that we identify gaps in terms of completeness, specificity, and implementation of ISMPs.

2.1. Information Security Management Practices in Organizations

In a study conducted by Keller et al. [9] about ISM in small organizations, nine best practices were recommended. These include: installation and proper configuration of firewalls; updated software; protection against viruses, worms, and trojans; implementation of strong password policy; implementation of physical security measures to protect computer assets; implementation of company policy and training; connecting remote users securely; locking down servers; and implementation of intrusion detection services. These practices are technological practices with the exception of implementation of policy and training.

Qingxiong and Pearson [16] proposed an ISM framework with the following five steps: assess organizational environment, establish information security objectives, analyze information security requirements, develop information security controls, and train or evaluate information security controls. However, due to ever changing information security patterns, new challenges associated with rising use of mobile and wireless systems have sprung up. Thus, the framework by Qingxiong and Pearson [16] does not adequately cover for new developments and new vulnerabilities.

According to Tryfonas [26], information security encompasses both managerial and technical aspects with issues emanating from internal and external sources. Tryfonas identified managerial practices such as establishment and use of policy, compliance with security standards, copyright protection, risk analysis and information security audits; and recommended technological practices such as requisite skills, cryptographic solutions, network security and use of firewalls, access control mechanisms, software security practices and intrusion detection techniques. Although Tryfonas identifies both managerial and technical issues, his recommendations are virtuously technological. Tryfonas further stated that simply deploying frameworks alone cannot solve all the information security and IT governance requirements. Therefore, organizations need to develop frameworks with elements that also include support from senior management, that include staff awareness and that include training in order to achieve all the intended benefits.

Other scholars like Oyelami and Ithnin [15] conducted a study on establishing a sustainable information security management policy in organizations. They proposed a guideline to ISMPs describing main processes and the activities of each process as

security management practices to be implemented in organizations to secure information. Accordingly, Oyelami and Ithnin [15] argued that an effective ISM process should consist of six sub-processes: policy, awareness, access, monitoring, compliance, and strategy; Oyelami and Ithnin [15] went ahead to describe each process and the activities undertaken thereof. Their study suggested that a policy process is a repetitive process continuously updated that starts by identifying what to secure, and document a draft policy that is to be published upon approval. However, a policy is exclusively managerial and therefore cannot handle all information security related issues of an organization.

In a study conducted to investigate ISM and its influence in the Nigerian banking sector by Babatunde and Selamat [5], a conceptual framework consisting of technological, organizational, and environmental factors was proposed. This framework identified factors that influence ISM practices among bankers from the perspective of reducing frauds and errors. If properly implemented, it could lead to improvement of information systems security and create a better investment climate in the banking industry.

A recent study was conducted by Zaini et al. [32] and was aimed at determining the extent to which ISM practices impact on the organizational agility. Zaini et al.'s [32] findings indicated that operational agility is significantly related to ISM practices in the sampled Malaysian organizations. Their study conceptualized three factors based on practices including: administrative security, technical security, and physical and environmental security controls. However, the suggested guidelines seemed to be complex to implement for some organizations especially those with difficulties in distinguishing security practices that influence agility [32].

2.2. Information Security Management Practices in Uganda

Uganda like most countries in the world has embraced the use of new Information Communication Technologies (ICTs) in all institutions including universities. This has been attained through introduction of information systems that ease handling of information in university institutions [13]. The adoption of information systems calls for a shift in the information handling mechanisms to match the new trends. There are some studies that have been carried out about information security related issues in Uganda. For instance, an investigation was conducted by Kisakye [10] into information security practices implemented by the Research and Education Networks of Uganda (RENU). Another study by Bogere et al. [8] was on the influence of ICT security to academic environment in universities in Uganda. Mbabazi et al. [12] assessed the implementation of information security policy in Ugandan universities. However, most of the Ugandan studies do not cover the attributes associated with the conceptual framework in this paper.

With the increased integration of ICTs in daily life, universities have also increased utilization of technology induced operations in their systems. This has encouraged the implementation of information systems security with corresponding practices to secure them. A number of university institutions have adopted reforms aimed at improving the implementation of information security measures. These include laws governing information security like the Computer Misuse Act 2011, Electronic Transactions Act 2011, Electronic Signatures Act 2011, Data protection and privacy Act 2019, etc. However, these laws have not been fully operationalized and incorporated in the information security management framework. To supplement these laws, the government of Uganda has developed an information security governance policy with elements of all activities required to manage information, personnel, equipment and physical security [27].

However, information management is still a big challenge to university institutions in Uganda. This is evidenced by the existing under developed information

management practices and strategies to guide information management processes and operations. These include but are not limited to information security governance practices, information security practices, personnel management practices, and physical security practices for information systems security as covered in the proposed conceptual framework in this paper.

3. Methodology

The purpose of this paper was to develop a conceptual framework with enhanced information security management practices in corporate organizations. It was motivated by lack of a comprehensive managerial-based information security framework for enhancing information security in organizations within the Ugandan context. In order to attain the intended objective, it was imperative that a conceptual framework suitable for the proposed solution be developed along the way.

To develop the conceptual framework, a systematic review of related literature in the domain area is recommended as the best approach ([4]; [24]) and it is the approach that was adopted for this paper. The systematic review involved identification of relevant studies, appraisal of their quality, and summarization of the evidence such that reasonable conclusions are reached [24]. This approach was followed to identify key concepts and develop the framework that identifies the different variables and their relationships as shown in Table 1.

4. Proposed conceptual framework for assessing ISMPs

The proposed framework identifies four manageable areas, namely: information security governance practices, information security practices, personnel management practices, and physical security practices. These four areas are further subdivided into 25 closely related subsections (activities) that form the checklist for collecting data and describing the status of ISMPs in an organization and their contribution to an organization’s information security.

The proposed conceptual framework shows the relationship between dependent and independent variables (Table 1). The framework describes the connection between the four main categories of ISMPs as well as the activities performed in each (as independent variables), against the organization’s information security (as a dependent variable).

Table 1. A conceptual framework for assessing information security management practices in selected universities in Uganda

Main practices	Sub practices (activities)	
	Independent variables	Dependent variable
Information security governance practices	<ul style="list-style-type: none"> • Policy statement on information security • Information security organization • Information systems risk management • Information security awareness, education and training among stakeholders • Have Business Continuity & Disaster Recovery mechanisms • Information systems incident risk management • Compliance with information system guidelines 	Organization’s information security
Information security practices	<ul style="list-style-type: none"> • Information system asset management • Secure information sharing • Information systems supply chain security • Access management to information systems • Install network security controls 	

	<ul style="list-style-type: none"> • Ensure portable and removable media security • Enhance remote access security • Ensure protective monitoring of information systems • Implement information back-ups • Secure accreditation for IS 	
Personnel management practices for information systems security	<ul style="list-style-type: none"> • Describe staff responsibilities and clear security roles • Ensure baseline security clearance for all users of information systems • Ensure Top secret clearance for all the users 	
Physical security practices for information systems security	<ul style="list-style-type: none"> • Install physical security measures • Install physical entry controls • Design internal data centre & physical access controls • Enhance information systems equipment security • Ensure secure equipment disposal & re-use 	

Source: Primary data

In the following subsections, we provide general descriptions of the different independent variables in Table 1 which are critical towards achieving an organization’s information security.

4.1. Information Security governance practices

One of the factors that influence how organizations manage an information security program is information security governance ([29]; [31]). Information security governance is a subset of the overall cooperate governance that include the adoption of a comprehensive approach to organize efforts and integrate various information security practices [30]. Adopting it would provide the right strategic direction, achieve objectives set, manage threats appropriately and use organizational resources correctly. However, this would require involvement of senior management and board of directors in information security governance issues. Therefore, implementing information security governance practices with support from higher organizational echelons would help to address the organization’s information security issues. This is an implication that involving senior management and board of directors to be held accountable for information security governance practices would provide necessary leadership. According to Trim et al. [25], good information security governance practices can be ensured if it is viewed as an integral part of corporate governance, processes and structures. Therefore, information security governance practices can serve as a major element to secure organization’s information. Such a scenario would require a clear policy statement on information security indicating information security aspects, systems risk management approaches, systems incident risk management strategies, education and training among stockholders, well stipulated Business Continuity and Disaster Recovery plans and measures for non-compliance with information systems security set guidelines.

4.2. Information Security Management

Information security practices are structured processes implemented to manage information security in organizations ([2];[3]). The aim is to maintain adequate levels of confidentiality, integrity and availability of information. This process involves planning, organizing, coordinating and controlling in order to establish acceptable level of security. The information security practices covered in the proposed

framework would entail activities performed to provide protection of information assets in organizations. These activities included asset management, secure information sharing, supply chain security, access management, network security controls, security of portable and removable media, remote access security, protective monitoring of information systems, implementation of information security back-ups as well information system accreditation by professional bodies. These activities if well implemented, would help organizations to maintain acceptable levels of integrity, confidentiality and availability of organization's information security [22]. Therefore, information security practices are regarded as an important element in the proposed information security management conceptual framework for selected universities in Uganda.

4.3. Personnel security practices

Information technology professionals are facing challenging tasks, analyzing, designing, and deploying solutions to protect information resources. This notwithstanding, previous studies concede that human beings are the major sources of many security failures ([11]; [1]). Human beings are vulnerable to a wide range of security attacks ranging from deliberate violation of security policy to circumvention of physical and technical security controls [23]. In addition, practitioners tend to underestimate the likelihood of occurrence of security breaches caused by human beings. A key area in information security research is understanding ways of motivating employees to participate in more secure behaviors [21].

Personnel security practices can address the problems associated with human oriented behaviors. The human resource practices of employee recruitment and selection, training and development, performance monitoring and appraisals are very important to improving organizational security routine [14]. Investing in staff training, awareness and development can motivate them to support attainment organizational goals. Staff training in information security management can serve as critical measure to safe guard an organization's information resources [6].

However, to achieve the best results, security training and awareness programs should be regularly evaluated to match the security threat levels [19]. Similarly, involving employees in overall security programs through commitment and engagement can be seen as a critical factor in improving job performance [18]. This is coupled with employee monitoring as an effort to maintain the high levels of productivity.

In order to ensure baseline security clearance for all users of information systems, employee background checks are important to ascertaining possible criminal records, character, and fitness for the position [20]. For staff to be held accountable and liable for their omissions or commissions with regard to their actions related to information security, their responsibilities and roles should be clearly described and clearance obtained to access top secret information in organizations.

4.4. Physical security

With regard to this conceptual framework, physical security meant activities aimed at installing physical security measures, installing physical security entry controls, designed internal data centre and physical access, all aimed at enhancing information systems equipment security. Physical security practices would stop unauthorized physical access, damage, and interference to information, premises and resources. This would curtail a range of physical security threats like crime, espionage, natural disasters, acts of terrorism and protecting personnel against violence and other harmful acts.

According to ISO/IEC 27001, organizations must protect equipment and personnel against physical and environmental threats. Such security measures help reduce the risk of unauthorized access to information and loss or damage to equipment. Physical security measures have strong importance to equipment, personnel and information. Organizations should also protect supporting facilities such as electricity supply and cabling infrastructure in order to reduce associated security threats or outcomes. Based on the proposed conceptual framework, a comprehensive managerial level framework would help to enhance an organization's information security.

5. Conclusion

The objective of this study was to develop a framework with a checklist of items for assessing information security practices and their relationship to an organization's information security. An in-depth examination of information security management practices in literature revealed lack of comprehensive guidance on assessing information security management practices in organizations. In this paper, we identify variables that constitute the proposed conceptual framework for information security management practices in organizations. We provide a summary of the contributions of the paper as follows:

- Information security has been known to be inadequate in most institutions including universities. This has been blamed mainly on the inadequacy of information security management practices. In this paper, we have presented a conceptual framework that would provide guidance for appropriate information security management practices for universities in Uganda.
- We also suggest that if the enacted laws to govern information security in Uganda like the Computer Misuse Act 2011, Electronic Transactions Act 2011, Electronic Signatures Act 2011, Data protection and privacy Act 2109, were fully operationalized and incorporated in the information security management framework, then system threats would be minimized.
- The proposed conceptual framework specifies an adequate checklist of standards that can be used to assess existing information security management practices adopted by universities in Uganda.

References

1. Abawajy, J. (2014). User preference of cyber security awareness delivery methods. *Behaviour & Information Technology*, 33(3), 237-248. URL: <https://doi.org/10.1080/0144929X.2012.708787>
2. Alshaikh, M., Ahmad, A., Maynard, S. B., & Chang, S. (2014). Towards a taxonomy of information security management practices in organisations. ACIS. URL: <https://openrepository.aut.ac.nz/handle/10292/8174>
3. Alshaikh, M. (2018). *Information security management practices in organisations*. PhD thesis, The University of Melbourne. URI: <http://hdl.handle.net/11343/208934>
4. Aromataris, E. (2014). The systematic review: an overview. *American Journal of Nursing*, March 2014, Vol. 114(3). URL: <https://doi.org/10.1097/01.naj.0000444496.24228.2c>
5. Babatunde, D. A., & Selamat, M. H. (2012). Investigating information security management and its influencing factors in the Nigerian banking industry: a conceptual model. *International Journal on Social Science & Art*, 2(2), 55-59. URL: <https://www.researchgate.net/profile/Dorcas-Adebola-Babatunde-2/publication/264884940>
6. Baxter, R. J., Holderness Jr, D. K., & Wood, D. A. (2016). Applying basic gamification techniques to IT compliance training: Evidence from the lab and field. *Journal of information systems*, 30(3), 119-133. URL: <https://doi.org/10.2308/isys-51341>
7. Bilsky, S. A., Cole, D. A., Dukewich, T. L., Martin, N. C., Sinclair, K. R., Tran, C. V. & Maxwell, M. A. (2013). Does supportive parenting mitigate the longitudinal effects of peer victimization

- on depressive thoughts and symptoms in children? *Journal of abnormal psychology*, 122(2), 406-419. URL: <https://doi.org/10.1037%2Fa0032501>
8. Bogere A., Haolader, F. A., & Mahbur, R. A. (2013). The influence of ICT security to academic environment at universities, case study Uganda: *International Journal of Innovative Research in Science, Engineering and Technology*, Vol 2, 4866-4873. ISSN: 2319-8753. URL: <http://www.rroij.com/open-access/the-influence-of-ict-security-to-academicenvironment-at-universities-case-study-uganda.pdf>
9. Coventry, W.L. & Keller, M. C. (2005). Estimating the extent of parameter bias in the classical twin design: A comparison of parameter estimates from extended twin-family and classical twin designs. *Twin Research and Human Genetics*, 8(3), 214-223. URL: <https://doi.org/10.1375/1832427054253121>
10. Kisakye, A. (2012). *An investigation into information security practices implemented by Research and Education Networks of Uganda (RENU)*. Masters thesis, Rhodes University. URL: <https://research.ict.ru.ac.za/snrg/Theses/Kisakye%202012%20MSc.pdf>
11. Komatsu, A., Takagi, D., & Takemura, T. (2013). Human aspects of information security: An empirical study of intentional versus actual behavior. *Information Management & Computer Security*, 21(1), 5-15. URL: <https://doi.org/10.1108/09685221311314383>
12. Mbabazi, B. P., Kareyo, M. and Muwanga-Zake, J.W.F. (2016). Assessing the implementation of information security policy in Ugandan Universities. *Global Journal of Engineering Science and Researches*, 3(11), 1-7. ISSN 2348-8034. URL: <http://www.gjesr.com/Issues%20PDF/Archive-2016/November-2016/1.pdf>
13. Mugenyi, R. (2017). Analysing information systems security in higher learning institutions of Uganda. *International Journal of Scientific & Technology Research*, 6(10), 385-392. ISSN: 2277-8616. URL: <https://www.ijstr.org/final-print/oct2017/Analysing-Information-Systems-Security-In-Higher-Learning-Institutions-Of-Uganda.pdf>
14. Naz, F., Aftab, J., & Awais, M. (2016). Impact of human resource management practices (HRM) on performance of SMEs in Multan, Pakistan. *International Journal of Management, Accounting and Economics*, 3(11), 699-708. URL: https://www.ijmae.com/article_116565.html
15. Oyelami, J. O., & Ithnin, N. B. (2015). Establishing a sustainable information security management policy in organization: A guide to information security management practice (ISMP). *International Journal of Computer and Information Technology*, 4(01), 44-49. URL: <https://www.ijcit.com/archives/volume4/issue1/Paper040107.pdf>
16. Qingxiong, M., Schmidt, M. B., & Pearson, J. M. (2009). An Integrated Framework for Information Security Management. *Review of Business*, 30(1). URL: link.gale.com/apps/doc/A220136074/AONE?u=googlescholar&sid=bookmark-AONE&xid=19347e2e
17. Ravitch, S. M., & Riggan, M. (2016). Reason & Rigor: How conceptual frameworks guide research. Sage Publications. URL: <https://doi.org/10.1177/105268461602600504>
18. Radhakrishna, A., & Raju, R. S. (2015). A Study on the effect of human resource development on employment relations. *IUP Journal of Management Research*, 14(3), 28. URL: https://www.iupindia.in/1507/Management%20Research/A_Study_on_the_Effect.html
19. Rantos, K., Fysarakis, K., & Manifavas, C. (2012). How effective is your security awareness program? An evaluation methodology. *Information Security Journal: A Global Perspective*, 21(6), 328-345. URL: <https://doi.org/10.1080/19393555.2012.747234>
20. Sarode, A. P., & Deore, S. S. (2017). Role of third-party employee verification and background checks in HR management: An overview. *Journal of Commerce and Management Thought*, 8(1), 86. URL: <https://indianjournals.com/ijor.aspx?target=ijor:jcmt&volume=8&issue=1&article=006>
21. Solaiman, B., Bosse, E., Pigeon, L., Gueriot, D., & Florea, M. C. (2015). A conceptual definition of a holonic processing framework to support the design of information fusion systems. *Information Fusion*, 21, 85-99. URL: <https://doi.org/10.1016/j.inffus.2013.08.004>
22. Soomro, Z. A., Shah, M. H., & Ahmed, J. (2016). Information security management needs more holistic approach: A literature review. *International Journal of Information Management*, 36(2), 215-225. URL: <https://doi.org/10.1016/j.ijinfomgt.2015.11.009>
23. Stewart, J. M., Chapple, M., & Gibson, D. (2015). *Certified information systems security professional study guide*. John Wiley & Sons. 7th Edition. URL: <https://www.wiley.com/en-us/CISSP+%28ISC%292+Certified+Information+Systems+Security+Professional+Official+Study+Guide%2C+7th+Edition-p-9781119042716>
24. Tawfik, G. M., Dila, K. A. S., Mohamed, M. Y. F., Tam, D. N. H., Kien, N. D., Ahmed, A. M., & Huy, N. T. (2019). A step-by-step guide for conducting a systematic review and meta-analysis

- with simulation data. *Tropical medicine and health*, 47(1), 1-9. URL: <https://tropmedhealth.biomedcentral.com/articles/10.1186/s41182-019-0165-6>
25. Trim, P. R. J., Lee, Y. I., & Weston, D. (2014). An interdisciplinary approach and framework for dealing with security breaches and organizational recovery. British Embassy Seoul. URL: <http://www.iaac.org.uk/media/1067/reporttrimyoumcybersecuritymarch14.pdf>
26. Tryfonas, T. H. E. O. (2010). Information security management and standards of best practice. *Handbook of Electronic Security and Digital Forensics*. World Scientific Publishing Co, 207-236. URL: <https://doi.org/10.1142/7110>
27. National Information Technology Authority (NITA) Uganda (2014). National Information Security Policy. National Information Security Framework (NISF) Publication, Uganda.
28. Varpio, L., Paradis, E., Uijtdehaage, S., & Young, M. (2020). The distinctions between theory, theoretical framework, and conceptual framework. *Academic Medicine*, 95(7), 989-994. URL: <https://doi.org/10.1097/ACM.0000000000003075>
29. Whitman, M., & Mattord, H. J. (2014). Information security governance for the non-security business executive. URL: <https://digitalcommons.kennesaw.edu/facpubs/3204/>
30. Williams, G. M., Kroes, R., & Munro, I. C. (2000). Safety evaluation and risk assessment of the herbicide Roundup and its active ingredient, glyphosate, for humans. *Regulatory toxicology and pharmacology*, 31(2), 117-165. URL: <https://doi.org/10.1006/rtph.1999.1371>
31. Yaokumah, W., & Brown, S. (2014). An empirical examination of the relationship between information security/business strategic alignment and information security governance domain areas. *Journal of Law and Governance*, 9(2). URL: <https://doi.org/10.15209/jbsge.v9i2.718>
32. Zaini, M. K., Masrek, M. N., & Sani, M. K. J. A. (2020). The impact of information security management practices on organisational agility. *Information & Computer Security*. URL: <https://doi.org/10.1108/ICS-02-2020-0020>

Some Features of Social Structures and Institutions Transformation in the Digital Age

Artem Balyakin¹[0000-0002-8655-7998],
Marina Nurbina¹[0000-0002-8063-9706], Sergey Taranenko¹

¹ NRC Kurchatov Institute, Moscow, Russia

https://doi.org/10.33847/2686-8296.4.1_3

Received 13.03.2022/Revised 28.04.2022/Accepted 22.05.2022/Published 12.06.2022

Abstract. The paper examines the peculiarities of digitalization processes influence on the architecture of emerging socio-economic relations. The legal regulation issues of digital technologies and the shifts they cause in public life are considered. The relations arising in connection with the regulation of big data are compared. The evolution of big data into smart content is described. The phenomenon of the "digital twin" is considered, as well as its impact on the social sphere. The tendency to move away from the policy of direct prohibitions in the field of digital technologies and the transition to the control of physical entities (data centers) and the regulation of methods and approaches to data processing (algorithms) is shown. It is noted that the existing expectations from digitalization are overstated. At the same time, the increasing influence of digital technologies significantly changes the existing socio-economic landscape, generating new risks. The answer to these challenges should be the joint work of authorities, business, society and the expert community on the formation of digital culture. It is shown that an important aspect should be the development of expert systems that translate qualitative characteristics into quantitative indicators.

Keywords: Big data, legal regulation, artificial intelligence, digital twin, digitalization, social systems, transformation.

1. Introduction and Literature Review

Digitalization intrudes in all aspects of life, and the fundamental importance of these processes is recognized both by business, society, and authorities [1,2,3]. Digitalization processes are based on scientific and technical developments in the field of data processing (big data technologies, artificial intelligence, decision-making systems, etc.). In the short term, we can expect the formation of a global infrastructure, an important element of which is correctly collected, structured and purposefully processed information that allows making optimal (or seemingly so) management decisions [4,5,6,7].

The most important elements of digitalization are, first of all, big data technologies and elements of artificial intelligence technology as a fundamental guarantee of the possibility of processing big data, as well as artificial intelligence in general as a decision-making tool, also based on "big data".

The downside of an excessive amount of information is the problem of sorting and choosing optimal solutions. One of the approaches being implemented is expert systems that translate qualitative characteristics into quantitative parameters. In particular, the methods of interviewing experts and forming a matrix of mutual influence [8,9] are increasingly being used in mathematical modeling of socio-economic processes [10].

Crucial point the authors stand for is the understanding that social and economic structures change constantly under the influence of digital technologies. That leads to the perception of non—neutrality of such technologies for the foresight purposes, both

theoretical-economic and societal. We stress that making a forecast without taking into account the processes and effects described below (following only the previously drawn trends) is erroneous.

Technically, the concept of "Big Data" aggregates a large number of different technologies that are actively used in all spheres of life, related to various fields of activity and, according to some viewpoint do not have signs of innovation [1]. Thus, the main criterion for referring to "big data" was the volume of processed information, "the size of which exceeds the capabilities of typical databases for entering, storing, managing and analyzing information" [3], and "big data" was defined by indicating the following main characteristics: volume, velocity and variety - the so-called 3 Vs [8].

Gradually, the definition of big data began to get rid of the restrictions focused on the current capabilities of machine information processing, and the focus moved to methods and approaches to processing primary and secondary data. The authors of [12] share the position that big data technologies are a complex phenomenon that involves both the development of technical solutions and understanding the humanitarian aspects of high technologies.

In our opinion, it is the humanitarian component of big data technology and the AI associated with it that is a perilous challenge generated by technological development. The risk of deceptive, false progress, when social changes caused by technological development must be considered as a regression, is a fact of our life. It is enough to mention the phenomenon of uberization, which consists in the fact that primitive social practices are revived through the most advanced technologies, returning the social fabric of society to primitive forms [13,14].

Thus, the importance of taking into account the social dimension of digital technologies is obvious: from the methods and approaches used, to the consequences of the practical implementation of digitalization achievements [14,15,16].

2. Obtained results

It is noteworthy that the term "big data" has not yet received an unambiguous legal definition in any of the world's jurisdictions. From a legal perspective, big data is currently not a legal term, but rather a description of a phenomenon that has many different implications for disciplines such as economics, engineering, legal sciences, and social sciences [1]. This circumstance is direct evidence of an incomplete understanding of the phenomenon of "big data" in changing societal and economic structures. In order to understand this role, in our opinion, it is necessary to trace a number of moments of the genesis and dynamics, including transformational dynamics, of this phenomenon. At the same time, it is immediately necessary to emphasize the symbiotic nature of the relationship between big data and artificial intelligence (AI) since the emergence of the "big data" category [16,17].

Currently, in relation to the symbiotic relationship of big data and individual AI technologies, the process of "data intellectualization", called the phenomenon of smart content, is intensifying. In our opinion, smart content is primary and secondary data of one or several subject areas processed and structured in a certain way based on big data technologies that are adequately interpreted with the help of intellectual procedures immanent to the subject area despite their immensity and irremediable heterogeneity [18]. Hereafter we note a number of characteristic features of this process, both allowing us to interpret data as smart, distinguishing them from mere data, and forcing us to consider smart content as a differently functioning object than "pure" data. Main characters of turning big data into smart content are illustrated in the Fig. 1.

Firstly, the impetus for the development of the technology of smart content was the technology of intelligent networks (i.e. smart grids), within which, by combining

heterogeneous data, building a complex of hierarchies, connections, potential completeness based on interpretation belonging to one or another notable area or several subject areas, a qualitative transition from "data" to "knowledge" was carried out.

At the same time, here the term "knowledge" must be understood in a different sense than in the term "human knowledge". Rather, it is a knowledge environment in which there is a part of the systemic connections that we interpret as "understanding". But this does not mean understanding as such, understanding immanent in the knowledge environment itself, the presence of a built-in "understanding" or any other equivalent of qualia. Understanding, as before, is an integral part of the human. Unfortunately, in recent times, in connection with the hopes for the expected (super)power of AI, such an approach is obscured. Thus, there is an expected, and due to expectations - imposed, change in the role of a person, including in the field of decision-making: production, economic, and others. The theoretical and economic "rational person" (despite all its controversy) is replaced by "a person with a rational decision-making system", and this rationality is external in relation to the person.

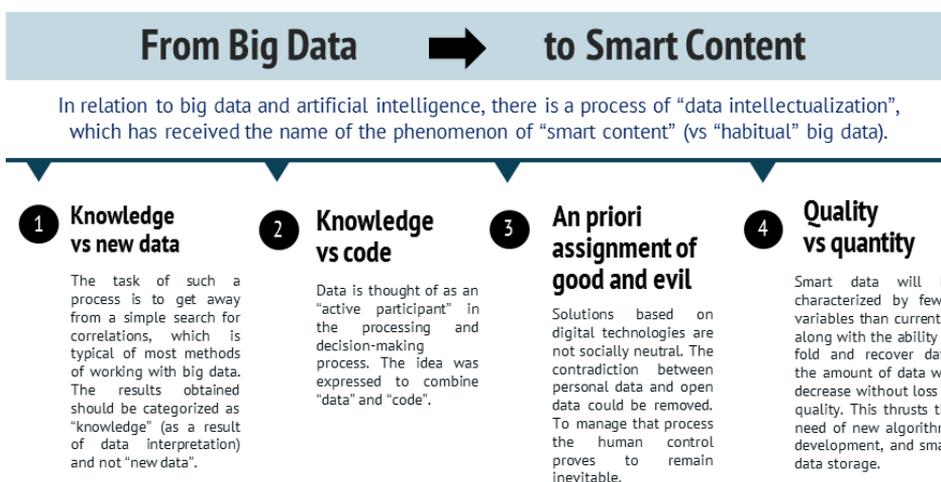


Fig. 1. The evolution of big data to smart content with main traits.
Source: original authors' design.

Secondly, the flow of big data and metadata from their previously inherent environments to others, in particular to technological environments and jurisdictions with the most convenient management protocols, has intensified. Data "behaves" in a reasonable way, giving "preferences" to get accumulated where it is "convenient" to get stored. In practical terms, there is an auto-accelerating concentration of big data in the form of "blue oceans" - isolated territories with their own rules. We note that such behavior is similar to memes [19].

Here, as part of the construction of predictive models, we must take into account two effects that lead to the differentiation of spatial, technological and cultural zones. The first one is significant for zones outside the "blue oceans", and it can be conditionally called the emasculation effect. The concentration of knowledge in some spaces is a source of possible changes in the qualities and role of other territorial and social spaces that are alienated from the "territory of knowledge".

The second characteristic of "oceans" is the potential presence of different routines and rules for handling data, including interpretative rules that separate "oceans" from each other. Knowledge potentially ceases to be an external effect, useful and applicable for everyone, if not equally, then to some significant extent, loses the universality of both access and meaning. Both can be an additional basis for the fragmentation of the World, or at least the presence of zones subject to independent analysis within a potentially unified forecasting system. In other words, the complexity of the predictive model increases inevitably.

Thirdly, not only the data itself, but also the algorithms that accompany this data and make it "smart", grow and begin to live their own lives and determine the evolution of the systems that they are destined to analyze. In science, the implementation of this approach is the FAIR principle (Findable, Accessible, Interoperable, Reusable) [20].

This is true both in terms of algorithms for interpreting data, when the evolution of meaning is determined not only by the evolution of data, but also by the evolution of algorithms, and in terms of algorithms for accessing data, their applicability or prohibition thereof [21].

In particular, in this regard, there is a problem of potential algorithmic unfairness, when approaches a priori embedded in algorithms evolve over time towards tightening and strengthening existing inequalities or begin to participate in decision-making [22].

The effect of algorithmic injustice, in our opinion, should be understood more broadly. In predictive models, we will be forced to abandon solutions focused on human behavior and the properties of human society that we understand (be it a model of rational behavior, the results of a social survey of the population, etc.) and derive the results (forming the picture of the future) from the dynamics of decision-making systems that we poorly comprehend [23,24].

Fourth, the process of "socialization" of data becomes obvious, meaning that with the growth of computing power and progress in analytics, there will happen a loss of meaning in the concept of "data protection", and access to information is going to be considered as one of "natural" rights [3].

Meanwhile, the potential general lack of data control, in our opinion, is a serious social challenge that requires its solution. For authorities, the process described above means a gradual transition from control over data to the grip over the means of storing, processing and transmitting information (for example, through a localization policy). At the same time, and in opposition to these potential limitations, a supranational big data circulation environment is emerging, one example of which is AIPO (OECD AI Policy Observatory), a comprehensive analytical platform for policy review and various national initiatives in the field of artificial intelligence.

The shift in the emphasis of state control over the key economic resource of our time, data and knowledge, its role in the construction of predictive systems requires its own understanding. The question whether the change in the flow of "data" will lead to a change in the dynamics of the flows of "knowledge" and, subsequently, to a change in the global structure of the quality of human capital remains open.

Continuing the list of features of the process of symbiotic interpenetration of big data and AI to be analyzed in the construction of predictive systems, one cannot but touch upon the issue of changing the quality of data in the technical, functional, as well as social areas.

Thus, fifthly, from the "technical" side, smart content will become "smaller", more compact, not only without loss of quality, but also with the acquisition of additional

characteristic properties. Issues of coherence, consistency and completeness of information will be addressed simultaneously with issues of compactness (density reduction). This involves the development of data convolution and recovery algorithms, primarily based on AI technologies. However, the problematic of convolution is at the same time the problematic of separating the essential from the non-essential, and therefore depends on the interpretation that we supposedly entrusted to AI. Taking into account this risk, it should be noted that the "convolution" also changes the dynamics and mobility of "knowledge" obtained through big data, the possibility of its operational application, which, presumably, can affect a variety of social and production processes and technologies, in including in terms of the width of the geographical and social coverage.

Confirming what has been said, sixth, big data, being AI-processed, becomes "useful", turning from "information" into "knowledge". This approach involves transferring the experience of working with big data available in the field of scientific research, where they are a priori selected according to specified algorithms within specific models, to other economic and social environments, including the business environment.

And in addition to changing social and technological practices, we should expect the emergence of fundamentally new ones. This is probably the most difficult moment to take into account in predictive models.

Seventh, the fundamental importance of the social and institutional role assumed by big data should be emphasized. There is no uniform understanding of what moral and ethical imperatives should underlie the evolution of big data. There is an agreement on the non-neutrality of digital technologies and the need for their "adjustment" in the direction necessary for society [22,25].

Such a possible adjustment is itself the subject of the forecast, the result of predictive models. Thus, we have a potential deadlock when building a predictive model and building a picture of the future. Avoiding this loop requires a prior understanding of the impact of digital technologies, at least at the scenario level.

For example, by now there is a scenario with probable danger of digital dictatorship and the system of new globalism [25]. A possible way to avoid this threat – and this is an alternative scenario – is to transfer key information to open data mode. The beginning of such a regime can be seen in such a concept as "open science". Within this approach, the idea of exemption from any restrictive legal regimes prevails, which some analysts describe as "nobody's data" [3,4].

At the same time, this decision can fairly be considered as a palliative, which does not solve the issue of moral and ethical imperatives underlying big data and AI, and which can become the basis for other scenario forks to be taken into account by predictive systems.

The foundation of the base scenario, within which it is possible and necessary to build predictive systems, according to the authors, should be based on the following assumptions. It can be taken for granted that in the nearest future, in practical terms, information ("data") will be produced from data lakes, understood as "nobody's data", using computing power and specialized methods and algorithms. According to the authors, the state policy in the field of digital technologies will be based on the logic of regulation of those entities over which physical control is possible. This raises the question of abandoning prohibitions that are not implemented in practice and focusing on the mandatory localization of processing centers with the possibility of providing their services to external customers.

A significant role will be played by the problems of ethical assessment of the big data and AI technologies usage, which will affect not only the dynamics of the relevant technologies, but also significantly impact their quality [22,26,27].

Research papers deal mostly with technical issues of data processing. The socio-economic aspects of high technologies are neglected or under-estimated, discussed fragmentarily and, for the most part, are considered as dependent, continuing the already developed trends (thus not changing the socio-economic landscape).

For instance, the author of [28] gives big data technology a "humanitarian connotation", proposing to adopt the concept of cultural and social capital developed by P. Bourdieu in relation to big data.

General philosophical issues of responsibility are considered, for example, in [29,30]. These authors proceed from the fact that algorithms silently structure our lives, removing responsibility from the actor, including the official, "impersonalizing" decisions. Instead of reliability and objectivity in relation to other uncertain procedures, the susceptibility of algorithmic models to human bias increases [29]. The way out seems to be both testing of algorithmic models used in practice [30], and the growth of "internal responsibility", increasing the involvement of employees developing algorithms and imposing ethical responsibility on them for the solutions being developed [30]. This approach looks controversial, if only because it does not mention the responsibility of the "customer" who sets the terms of reference. As a rule, this is the authorities [31], which, depending on its goals and objectives, implements different approaches (and, as a result, demonstrates the demand for different algorithms).

Despite the obvious fragmentation of socio-economic ideas in the ethical component, which is also poorly developed, in practical terms, we can talk about the formation of different ethical models.

Nowadays three main model approaches to ethical problems in general can be distinguished: an anthropocentric approach, characteristic of the EU; orientation to commercial interests (typical of the USA); the idea of the common good according to Confucius in the PRC (Fig. 2). So, in the EU, algorithms are divided into three categories: permitted, prohibited and partially permitted. There are no restrictions on algorithms in the USA, but the results obtained are regularly checked for compliance with socially approved norms, including gender aspects, issues of discrimination of social groups, etc. This position leads to an absurd situation when the results in a number of scientific fields turned out to be dependent on the prevailing socio-political discourse [22].

As a result of such a diversity of approaches, the created digital technology management systems are focused on and support different values and, in fact, create different digital worlds (divided into filter bubbles) with various related technologies, including AI [32]. So, in predictive models, it is necessary to take into account both points: firstly, the multi-variance of technological, including socially conditioned, development; secondly, different value orientations [27].

It should be noted that the novelty of emerging challenges is apparent [24]: this has already taken place in human history. Thus, it is enough to mention the value differences of different cultures, primarily marked by confessional differences. What is new at the present time is that the space, which we previously considered homogeneous, will not be same in the future. Moreover, the existing political tensions only contribute to the growing divergence between the main actors: filter bubbles receive legislative support. For example, according to the Chinese program for the

development of science and technology, the methods and approaches being introduced, the scientific ethics used should be “Chinese” ethics [33].

The legal regulation of evolving digital world will be changed

The source of law is shifting towards non-state institutions . A clearly visible effect of this kind of processes can be called “digital alienation”. Making decisions (based on the recommendations of the algorithm, decisions of a trained AI, etc.) alienates a person from law as an institution.



EU

Anthropocentric approach



USA

Emphasis on commercialization



PRC

Values of Confucianism

Competition of ethics is coming

The commencing race in the field of digitalization leads to a rivalry between both technical characteristics and moral and ethical standards that serve as the basis for the development of algorithms and methods (including data access policy).

Fig. 2. Approaches to legal regulation of big data in main World jurisdictions.
Source: original authors’ design.

This, as previously stated, is a serious argument in favor of the hypothesis of the digital defragmentation of the world, despite the persistent idea that the "digit" leads to the unification, globalization of all societal processes. Also social and technological changes generated by the phenomenon of big data can and should find their imprint in a fixing social institution, the institution of law.

These features of the development of digital technologies lead to the fact that within the framework of these processes a new law can and should be formed as a fundamental social institution of boundary conditions and potential development opportunities, which is of fundamental importance in forecasting constructions. It should be emphasized that this new law is based on the phenomena of digital reality, which in modern times did not have a decisive value or are considered as non-legal.

In terms of the analysis of the institution of law, the following can be said.

The large-scale use of the results of scientific research leads to the formation of a new complex branch of law - the law of science. This is a special area of knowledge, located in the system of the humanities, but comprehensively and inextricably linked with the legal regulation of scientific research [34].

Legal relations in the field of science, solving specific, but new tasks characteristic of science, contributed to the formation of fundamental law. Thus, in the field of scientific research, and characteristic of recent international cooperation and co-financing, the need to share risks and rights to the results of work, it is much easier for participants to agree on the conditions for participation and work at the beginning of the project [34].

The source of law is shifting towards non-governmental institutions. A well-marked effect of such processes can be called "digital alienation" as a new manifestation of alienation proposed by late Marxism [18,28].

In general, big data, as a mature phenomenon, is now a necessary component in making responsible decisions both at the large companies and at the government level [28]. But at the same time, decision-making based on the recommendations of an algorithm, decisions of a trained AI, etc. alienates a person from law as a social

institution, replacing the latter with the ideas of a generalized "programmer" ("coder"), including AI. Reliance on a "computer solution" as an explanatory factor of denial and/or acceptance is a common feature of our time [22]. Thus, there is a transition to a new, "impersonal" form of regulation, tied to algorithmic approaches.

The second aspect of law deformation is associated with a change in the nature of production processes. Degradation processes of different genesis are taking place within the digital economy [5,27]. Thus, there is a displacement effect when market players use digital technologies to circumvent existing laws, working conditions and fiscal systems. There is a real danger of uncontrolled development and unethical use of technologies, and it is still unclear how to avoid this [35].

Another genesis of possible deformations is a shift justified by the development of digital business, which requires the removal of previously existing institutional barriers, which today have become an obstacle to the introduction and use of digital technologies. Here we should expect the legalization of the institution of "corporate law", dictated by corporate interests and aimed at everyone who has relations with the corporation [36].

Concluding from above mentioned, we should also expect the effect of law fragmentation: the law will be different in different parts (by geographical, technological or sectorial characteristics). In general, there is an acceleration of the process of shifting the law from the model of "rules" to the model of "operational management", or, as emphasized, the seigniorate, which is a prerequisite for possible deep social transformations [13,36]. Moreover, the concept of "seigniorate" ceases to be exclusively geographical. In various sectors of human activity, primarily economic, the simultaneous existence of independent "seigniorates" is possible.

Thus, in addition to the "knowledge" fragmentation described above, we also assume a possible legal defragmentation of society occurring parallelly.

The economic order as part of the development of big data and AI technologies, taking into account what has already been described, will undergo a number of changes. Drivers of economic changes are, in particular, innovations. So, in addition to changing regulatory mechanisms, the development of digitalization is associated with a number of innovative solutions [5,24]. Thus, one of the ways to improve efficiency is the active introduction of methods and approaches of computer modeling. It is assumed that advances in the field of data processing, in the field of artificial intelligence will make it possible to transfer the identification of "real" problems to the virtual world, testing systems in it, analyzing them and working out decision-making strategies [2,37,38]. In the future, digital platforms are expected to come into use, while gradually adjusting into digital ecosystems (in order to consolidate all participants in the value creation process) [39].

Among such "modelling" approaches is the phenomenon of "digital twins", which is understood as "the real display of all components of the product life cycle using physical data, virtual data and data of interaction between them" [40]. The choice of digital twins as an object of research is fundamentally important for understanding the nature of the digital economy: we are not following a phenomenon, but are inclined to its copy or model.

According to 2019 estimates, three-quarters of the surveyed companies plan to use digital twins in the near future (e.g., a Gartner survey shows that digital twins are becoming mainstream [41]). The annual growth of the digital twin market from 2020 to 2026 is predicted to reach 58% [42].

The digital twin in science opens up the possibility of access to a unique scientific facility for a wide range of users and enhances the educational component [43,44]. In particular, in Russia, in order to support the operation of the world-class International Center for Neutron Research reactor, a new course "Methods based on synchrotron radiation in interdisciplinary research" is being developed using digital twin technologies of unique research equipment. The existing mega-installations, such

as the European X-ray Free Electron Laser (XFEL, Hamburg, Germany) and the Synchrotron Center of the National Research Center of the Kurchatov Institute (Moscow) will be used as a background [44].

Thus, the formation of the phenomenon of the "digital twin" can be considered to have taken place. According to the authors, it is premature to talk about its role from the point of view of the forecast, in addition to the above. This is a potential black swan. The digital twin, as well as the digital footprint (copy of an entity) is not the object under consideration, and, therefore, it leaves the possibility of misinterpretation. At the same time, the results (data) produced by them are attributed to a real object, which hypothetically leads to an incorrect interpretation. Solutions that look obvious when they are developed are likely to produce unplanned results and errors when launched in real life. As an example, we can cite various legal conflicts accompanying UAVs and automatic systems in medicine and transport [46]. For today, the possible effect of an error (substitution model and a real-life object to draw a conclusion) of a digital twin has not been described. Thus, the predictive value of this phenomenon is yet unknown.

Another significant innovative solution is the reuse of data in order to obtain new knowledge from them. This approach is based on the FAIR principles [44]. Similar processes are already being observed in the social sphere, when a person's "digital footprint" is being formed, and used for targeted advertising, risk assessment of reoffending or for the formation of an individual educational trajectory. Unlike physical systems, in this case, the question of the fullness of "data lakes" (initial unstructured arrays of information) inevitably arises. In order to ensure their uniform filling, the simplest method seems to be the introduction of forced formation of a digital footprint. That, remaining within the framework of moral and ethical assessments of modernity, should be considered as a serious challenge.

The implementation of this approach has become possible due to significant progress in the creation of storage systems and remote access to lakes and databases, as well as a significant reduction in the cost of information transfer and storage procedures. Note that in this case, the problem of interpreting the results, developing algorithms or models becomes secondary (postponed "for later"), and the most important tasks are recording, storing and providing remote access. This approach provides a link between the model and the algorithm with the real world (experimental data). It is essential that it is implemented only in the case of systems where redundancy of experimental data is possible. This situation is typical for high-energy physics and astronomy, but raises doubts about socio-economic systems. The reason for this is the variability of quantitative characteristics and the lack of an unambiguous transition of qualitative indicators into quantitative ones (including due to moral and ethical assessments external to the study).

4. Conclusion

Our study allows us to formulate a number of conclusions.

Firstly, the digital economy, AI technologies, big data, etc. have not yet become a fundamentally new area; they highlight pre-existing trends, challenges and risks. The perception of them as a revolutionary new tool is erroneous; they are only a technical way of solving current problems [6,24].

Secondly, there is currently a process of data accumulation, uncontrolled growth of information, methods, approaches, etc. Critical reflection on the impact of emerging data lakes is the next step in the development of the digital economy. This process should be accompanied by the formation of data processing standards, which will ultimately allow us to talk about the emergence of a new institution of digital technologies, an integral part of which will be "smart content".

Thirdly, the development of digitalization reflects another manifestation of globalization. In this regard, control over individuals (for example, the policy of localizing data centers in Russia) can serve as a deterrent to prevent the flow of "data lakes" to countries with different legal and political conditions for using such data.

Fourth, the policy of prohibitions and restrictions seems to be the simplest solution to the problems that arise, but this approach proves to be not effective. It is necessary to look for other methods and approaches to regulation, including a clear ethical framework for interaction with digital technologies. In practice, a new culture of the digital society must emerge. In the current conditions of science, technology and, above all, ethics, a universal artificial intelligence (equally strong in all areas and equally responsive to the same inputs) is impossible. In a pessimistic future scenario, decision-making systems will be based on implicit algorithms that exist in non-overlapping realities (filter bubbles) [25,32].

Fifth, such a phenomenon as a "digital twin" is becoming increasingly widespread. This means both a universal model of a real system and a set of technological solutions. In the social and educational sphere, a digital double is often understood as a digital footprint.

Sixth, there is an increase in personalization processes: from medicine to education. The combination of data (primarily digital footprint), algorithms and robotic production makes it possible to make a unique product for any need at a mass price [17]. This is expressed, among other things, in the personalization of artificial intelligence: in the EU, in May 2019, the AI-on-demand platform was launched, within which 6 artificial intelligence projects are being implemented [46].

Awareness of the described processes is not enough to overcome the challenges posed by digitalization. There should be an understanding at the imperative level that society, society (represented by the state) establishes rules and norms for business as a social institution, and not vice versa. That is why it becomes socially necessary to form a well-thought-out and effective regulatory system through the law of the relationship between man, science, the state, society and nature.

The next stage in understanding the processes of digitalization should be the study of possible socio-economic changes in society caused by the reformatting of ethical norms, and the formulation of proposals to curb negative trends in line with the primacy of social and humanitarian challenges [13,25,28]. In this regard, the authors actively support the idea of developing a set of international ethical principles for the use of advanced digital technologies. We believe that such a set of rules can be based on the Report on the Ethics of Robots, artificial intelligence and Algorithms of the UNESCO World Commission on the Ethics of Scientific Knowledge and Technology (signed in Paris on September 14, 2017).

The importance of studying the ethical aspects of AI is related to the limitations they produce [26]. The most alarming situation is in the case of codification of moral and ethical rules. A striking example is the GDPR (General Data Protection Regulation), which, according to experts, hinders scientific progress in Europe. As of 2019, more than 5,000 projects were affected by the restrictions prescribed in the GDPR, primarily in the field of personal information control [47,48]. Currently, discussions have been initiated on how to respond to this challenge. One of the options proposed during these discussions, shared by the authors, is the focus on improving the overall digital culture. The authorities actively interfere in this process: for example, in 2015-2020, 117 approaches in the field of ethics and morality were developed, of which 91 were in the EU and the USA [49]; China and Russia are following suit.

In general, the ongoing introduction of digital technologies carries a number of risks, dramatically changing social practices. It is required both to clearly articulate emerging challenges and develop responses to them, using both the arsenal of engineering sciences and applying the methods of the humanities and social sciences.

One of the possible solutions should be the active use of expert systems, laying the ethical component in the developed elements of the digital system and translating qualitative characteristics into quantitative indicators. Schematically the approach to digitalization the authors support and promote is presented in the Fig. 3. To our viewpoint, each component is important, and none can be neglected. Only joint efforts of society, business community, and authorities would lead to successful development and further implementation of digital technologies.



Fig. 3. The digitalization process (authors' design).

The figure reflects the importance of data (as a source), experts' system and scientific facilities (as a tool), and society in total (as a beneficiary of progress in digital technologies).

Acknowledgments

This work was supported by RFBR grant № 20-010-00576.
 Authors thank Zhulego V.G. for useful discussions.

References

1. Walker, M., Burton, B.: Hype Cycle for Emerging Technologies (2015), <https://www.gartner.com/en/documents/3100227>, last accessed 2022/03/31.
2. Gartner Top 10 Strategic Technology Trends For 2020, <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/>, last accessed 2022/02/21.
3. Manuka, J., Chui, M., Brown, B., Bughin, J., Hobbs, R., Roxburgh, C., Byers, A.: Big data: The next frontier for innovation, competition, and productivity / McKinsey Global Institute (2011).
4. Measuring the Digital Transformation: A Roadmap for the Future (2019), <https://www.oecd.org/publications/measuring-the-digital-transformation-9789264311992-en.htm>, last accessed 2022/03/31.
5. Melnik, M., Antipova, T. (2020) Organizational Aspects of Digital Economics Management. In: ICIS 2019. Lecture Notes in Networks and Systems, vol 78, pp. 148-162. Springer, Cham. https://doi.org/10.1007/978-3-030-22493-6_14.

6. Ross, G., Antipova, T., Konyavsky, V. (2022) Methodology for Innovative Projects' Financing in IT Business. Lecture Notes in Networks and Systems, 381, pp. 257–268. https://doi.org/10.1007/978-3-030-93677-8_22.
7. Nurakhov N.N. (2020) The Basic Processes of Creating a Megascience Project. Lecture Notes in Networks and Systems, 78, pp. 329–339. https://doi.org/10.1007/978-3-030-22493-6_29.
8. Saaty, T.L.: Decision making with dependence and feedback: The Analytic Network Process. Pittsburgh: RWS Publications (1996).
9. Saaty, T.L. (2008) The Analytic Hierarchy and Analytic Network measurement processes: Applications to decisions under risk / European Journal of Pure and Applied Mathematics, 1(1), pp. 122–196.
10. Balyakin A.A., Domnich A.S., Zhulego V.G., Taranenko S.B. Designing the Future: Nonlinear Dynamics in Economic Models. Integral Scientific and Practical Interindustry Journal. 2011. No. 1 (57), pp. 33–35.
11. Lynch, C.: Big data: How do your data grow? Nature, 455, 28–29 (2008).
12. Mayer-Schönberger, V., Cukier, K.: Big Data: A Revolution That Will Transform How We Live, Work, and Think. Moscow: Mann, Ivanov & Ferber (2014).
13. Balyakin, A.A., Nurbina, M.V., Taranenko, S.B. (2021) Ethics in Big Data: Myth or Reality. In: Á. Rocha et al. (eds.) Information Technology and Systems, AISC 1330, pp. 14–22.
14. A. Pentland Social Physics: How Social Networks Can Make Us Smarter. Penguin Books. 2015
15. Cathy O'Neil Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy Crown, 2016
16. Ivanov V.V., Malinetskii G.G. Contours of digital reality. Humanitarian technological revolution and the choice of the future. Lenand, Moscow, 2018, 344 p.
17. Kai-Fu Lee, AI Superpowers: China, Silicon Valley, and the New World Order, Boston, Mass, 2018
18. Balyakin, A.A., Nurbina, M.V., Taranenko, S.B. (2021) Some Current Aspects of Big Data Evolution. In T. Antipova (Ed.). ICADS 2021, AISC 1352, pp. 444–450. https://doi.org/10.1007/978-3-030-71782-7_39.
19. Dawkins, Richard: The Extended Phenotype, Oxford University Press, p. 109 (1982).
20. FAIR Principles, <https://www.go-fair.org/fair-principles/>, last accessed 2022/05/13.
21. Balyakin, A.A., Malyshev, A.S.: Big data management in research infrastructures. In: Open Systems. DBMS. 2020(3), pp. 40–42. Moscow (2020).
22. Birhane, A.: Algorithmic injustice: a relational ethics approach. Perspective, Patterns 2, vol. 2(2), 100205, February 12 (2021).
23. Griffy-Brown C., Chun M., Miller H., Lazarikos D. How Do We Optimize Risk in Enterprise Architecture when Deploying Emerging Technologies? J. Digit. Sci. **3**(1), 3 – 13 (2021). https://doi.org/10.33847/2686-8296.3.1_1.
24. Balyakin A.A., Malyshev A.S., Nurbina M.V., Titov M.A. (2020) Big Data: Nil Novo Sub Luna. In: ICIS 2019. Lecture Notes in Networks and Systems, vol 78, pp. 364–373. Springer, Cham. https://doi.org/10.1007/978-3-030-22493-6_32.
25. Harari, Y.N: 21 Lessons for the 21st Century. p. 416, Vintage Digital (2018).
26. Ustinova, Y. (2021) The true and fair view concept: the palette of controversial points (of «worth banning» to «worth keeping»). J. Digit. Art Humanit., **2**(1), 39–47. https://doi.org/10.33847/2712-8148.2.1_4.
27. Balyakin, A.A., 1, Taranenko, S.B., Nurbina, M.V., Titov, M.A. Social Aspects of Big Data Technology Implementation. J. Digit. Sci., **1**(1), 15–24 (2019). https://doi.org/10.33847/2686-8296.1.1_2.
28. Sadowski, J.: When data is capital: Datafication, accumulation, and extraction. Research Article, 6(1), 1–12 (2019).
29. Kemper, J., Kolkman, D.: Transparent to whom? No algorithmic accountability without a critical audience. Information, Communication & Society, 2081–2096 (2018).
30. Kirsten, M.: Ethical Implications and Accountability of Algorithms. Journal of Business Ethics, 160(4), 835–850 (2019).
31. Engin, Z., Treleaven, Ph. (2019). Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. The Computer Journal, 62(3), pp. 448–460.
32. Pariser, E.: The Filter Bubble: What the Internet Is Hiding from You. New York: Penguin Press. (2011).
33. Viewpoint: China risks damaging science ties by forging own path on research ethics. <https://sciencebusiness.net/news/viewpoint-china-risks-damaging-science-ties-forging-own-path-research-ethics>, last accessed 2021/09/23.

34. Chetverikov, A.O. (2018). Organizatsionno-pravovye formy bolshoy nauki (megasayens) v usloviyakh mezhdunarodnoy integratsii: sravnitelnoe issledovanie. Legal forms of big science (megascience) in the context of. Legal science, 1 (1), 13-27; 2 (2), 34-50.
35. Von Weizsaecker, E., Wijkman, A.: Come On! Capitalism, Short-termism, Population and the Destruction of the Planet. Springer Science+Business Media LLC. (2018).
36. Balyakin, A.A., Nurbina, M.V., Taranenko, S.B. (2020) Comparative legal features of the formation of a digital ecosystem. Int. Legal Courier 1-2(37-38), 42-45.
37. The decree of the President of the Russian Federation. About the strategy for the development of the information society in the Russian Federation for 2017-2030. No. 203, 09.05.2017
38. Order of the Government of the Russian Federation of July 28, 2017 No. 1632-r Program "Digital Economy of the Russian Federation", <http://static.government.ru/media/files/9gFM4FHj4PsB79I5v7yLVuPgu4bvR7M0.pdf>, last accessed 2022/02/02.
39. URL: <https://e-cis.info/news/568/82077/>, last accessed 2019/05/20.
40. Fei Tao et al. (2018). Digital twin-driven product design framework. International Journal of Production Research, 57(12), pp. 3935-3953.
41. Gartner Survey Reveals Digital Twins Are Entering Mainstream Use (2019) / Gartner. <https://www.gartner.com/en/newsroom/press-releases/2019-02-20-gartner-survey-reveals-digital-twins-areentering-maj>, last accessed 2019/03/31.
42. Digital Twin Market by Technology, Type (Product, Process, and System), Application (predictive maintenance, and others), Industry (Aerospace & Defense, Automotive & Transportation, Healthcare, and others), and Geography - Global Forecast to 2026. (2020), <https://www.marketsandmarkets.com/Market-Reports/digital-twin-market-225269522.html>, last accessed 2020/05/13.
43. Singh, M., Fuenmayor, E., Hinchy, E.P., Qiao, Y., Murray, N., Devine, D.: Digital Twin: Origin to Future. Appl. Syst. Innov, 4, 36 (2021).
44. Balyakin, A.A., Nurakhov, N.N., Nurbina M.V. (2021). Digital Twins vs Digital Trace in Megascience Projects. In Á. Rocha et al. (Eds.). ICITS 2021, AISC 1330, pp. 534-539.
45. Crease, R.P., Westfall, C.: The New Big Science Physics Today, 69 (5), pp. 30-36 (2016).
46. URL: <https://digital-strategy.ec.europa.eu/en/news/new-projects-enrich-ai-demand-platform>, last accessed 2022/03/31.
47. URL: https://www.feam.eu/wp-content/uploads/International-Health-Data-Transfer_2021_web.pdf, last accessed 2021/10/10.
48. URL: <https://sciencebusiness.net/news/data-protections-rules-harming-eu-leadership-health-research-says-report>, last accessed 2022/05/22.
49. URL: <https://sciencebusiness.net/news/time-harmonise-artificial-intelligence-principles-experts-say>, last accessed 2022/03/10.

Geomatics and smart tools in Digital Land Resources Mapping and Sustainability of Coastal Agriculture, Egypt

Mohamed Zahran¹ [0000-0002-4747-6191] and
Abd-Alla Gad¹ [0000-0002-2249-4246]

¹ National Authority for Remote Sensing and Space Science (NARSS), Cairo, Egypt

<https://doi.org/10.33847/2686-8296.4.1> 4

Received 10.04.2022/Revised 19.04.2022/Accepted 22.05.2022/Published 12.06.2022

Abstract. The northwestern coast of Egypt is characterized by an international interest due to its history and magnificent environment. The area was known as being the bread basket during the Greek and Roman periods. Recently, drastic changes in land use resulting in destructing many of water harvesting tools, thus diminution of the agriculture importance. Restoration of the area and planning self-sufficient communities needs to develop a sustainable land resources database for these regions. Multi concept of remote sensing and the Geographic Information System (GIS) permit to store, merge, and manipulate the huge amounts of thematic maps and attribute data. Sentinel satellite image 2018 scenes, covering the study area at the Egyptian northwestern coast, were acquired. ENVI software was used for image processing. A number of 53 topographic maps at scale 1:50000 were used to input GIS thematic layers relevant to land resources, using Arc_GIS 10.2 system. Field investigation was carried out to represent different soil units and collect ground control points. Chemical and physical soil properties were determined to assist soil classification. Soil map was produced including dominant geographic units and soil association. MicroLEIS system was employed to define soil suitability classes to olives, peach, wheat, beans, and sunflower crops. An intelligent module will be added to analyze the digital maps, interact the given data with learning tool (layer) to provide the decision makers with suggested solution not only information. The results showed that the soils are generally characterized by the presence of Calcic, Petrogypsic and Salic horizons. The limiting factors found in the piedmont and coastal plains are salinity, soil depth and texture. These factors decrease the suitability classes to be between S2 and S5. It can be concluded that the digital mapping of land resources using Geographic Information System (GIS) and satellite data preserve in the investment spent in soil and other thematic mapping.

Keywords: Soils, Space data, GIS, Digital soil mapping, Egypt, IoT.

1. Introduction

With the great explosion in computation and information technology come vast amounts of data and tools in all fields of endeavor. Soil science is no exception, with the ongoing creation of regional, national, continental and worldwide databases. The challenge of understanding these large stores of data has led to the development of new tools in the field of statistics and spawned new areas such as data mining and machine learning [9] In addition to this, in soil science, the increasing power of tools such as geographic information systems (GIS), GPS, remote and proximal sensors and data sources such as those provided by digital elevation models (DEMs) are suggesting new ways forward. Fortuitously, this comes at a time when there is a global clamor for soil data and information for environmental monitoring and modelling. Consequently, worldwide, organizations are investigating the possibility of

applying the new spanners and screwdrivers of information technology and science to the old engine of soil survey. The principal manifestation is soil resource assessment using geographic information systems (GIS), i.e., the production corresponding to national to global, catchment to landscape and local extents. In the language of digital soil maps, different from that of, scale is a difficult concept, and is better replaced by resolution and spacing [2].

The northwestern coast represents a promising region for extensive development both for local and expected new inhabitants. The concern of coastal resources has increasingly risen during the last two decades mainly because of the great pressure of human actions (urban expansion, industry, tourism, infrastructure, aquaculture, fisheries ports and marinas, energy production and transportation) but also due to the ineffective information, policies, planning and management tools for controlling or regulating human actions and natural processes (natural risks or hazards) in such sensitive environments as that of coastal zones.

The study area dominates the northwestern coast of Egypt between Burg El Arab and El Sallum (Fig. 1). It is bounded by latitudes 30° 30' N and 31° 45' N and longitudes 25° 00' E and 29° 30' E. The distance from Alexandria to the extreme east of the study area is about 390 km, while extending for some 600 km to the extreme west.

The cultivable soils in the northwestern coast are originated from transferred sedimentary rocky material. The sediments have been transported by water to alluvial fans and flood plains. Soils were formed also by Aeolian sediments in some locations. The subsoil layers are formed locally from the marine limestone. The soil depth varies according to its location, found shallow in the sloping and plateau landscape, and deep in the coastal plain and alluvial fans. The occurrence of calcium carbonates ranges from 30 to 70% and may reach 99% in the calcareous sands [8].

The area from the coast to the Libyan plateau includes calcareous formation belong to the Pliocene and Pleistocene covered with recent sediments. The existence of parallel ridges along the coast characterizes the area. The ridges are absent in some locations, and consist of calcareous sedimentary material differ in their coherence. The Libyan plateau occupies huge area and extends southwards [3].

The area includes a narrow coastal plain, followed at the south by a sand dune area. Southwards of the dunes, the plain rises gradually till the altitude of the plateau this reaches 50 to 150 meters above sea level. The coastal plain stretches in east-west direction, bounded by the sea to the north and the pediment plain to the south. Its width varies, controlled by the geologic formations from some meters to about 10 km. This plain mainly consists of alluvial fans, descending from the plateau, wad's extensions, rocky plains sabkhas, sand sheets and sand dunes. The pediment plain is clear between Ras El-Hekma to Ras Alam El-Room. It is a low lying plain where rain water, descending from the plateau is collected. This area has a considerable potentiality for agriculture expansion. The plateau is rocky, covered mostly by a thin depth of soil. It plays an important role in distributing winter rainfall [6].

Most of the cultivable soils in the northwestern coast are alluvium. The sediments have been transported by water to alluvial fans and flood plains. However, Aeolian sediments in some locations are being cultivated. The subsoil layers are formed locally from the marine limestone. The soil depth varies according to its location, found shallow in the sloping and plateau landscape, and deep in the coastal plain and alluvial

fans [11].

According to Egyptian Meteorological Authority [5], the average annual rainfall ranges between 156 – 180 mm. and the mean minimum and maximum annual temperatures are 16.4 and 23.0 C° respectively. The evaporation rates are coinciding with temperatures where the lowest evaporation rate (6.9 mm/day) was recorded in January while the highest value (8.8 mm/day) was recorded in September.

This study aims to use the remote sensing data and Geographic Information system to produce the land resources digital maps of the Northwestern Coast of Egypt, which can be used as a base for land use planning and sustainable development.

2. Geomatics and smart city

From the perspective of geomatics, a smart city is the full integration of a digital city, the Internet of Things and cloud computing technology. A digital city provides a 3D geospatial framework for cities, while the Internet of Things embedded in the ubiquitous sensor network realizes the real-time sensing, measuring, and data transmitting of still or moving objects. Cloud computing, performing like a human brain, is responsible for massive and complex calculations, data mining, and analysis; and it then helps in the automatic discovery of patterns, rules, and knowledge and provides remote monitoring, control, and feedback to the real world for intelligent city management and public services.

Most of the current smart city efforts are focused on how to build a complete Internet of Things, including men, machines, and city infrastructure through a variety of wireless sensor networks, as well as how to accomplish real-time analysis and control by super-computers on cloud computing platforms.

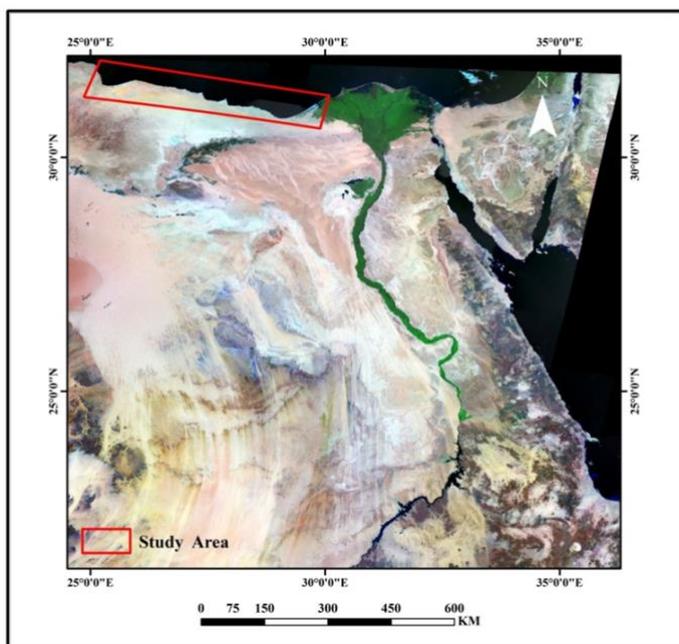


Fig. 1. Location map of study area

2.1. Concept of digital earth

From a geomatics point of view, the supporting techniques of a digital city consist of the following:

- 1) *Space, air, and land-based sensor webs for geospatial data acquisition and updating plus modern communication devices;*
- 2) *Building extraction and 3D/4D modeling;*
- 3) *Multiresolution, multiscale, and multidimensional visualizations of geospatial data;*
- 4) *Distributed spatial data archiving and management for federal databases with interoperability;*
- 5) *Spatial data analysis and mining techniques; and*
- 6) *Global navigation satellite system (GNSS) and location-based service techniques.*

Although these technologies are becoming mature and stable, the future generation of digital cities still requires the complete cooperation of various governments and institutions, the standardization of data and products, and the interoperability of nonprofit organizations and cooperation. The ultimate goal of a digital city, with the help of all the above technologies, is to provide the "right data" to the right person at a right time and a right place (10).

3. Materials and methods

This study is based on the multi concept of remote sensing data and techniques, thus, materials and methods of different sources are used as the following:

- Sentinel satellite image 2018 was used to obtain the space images mosaic of the studied area (Fig. 2).

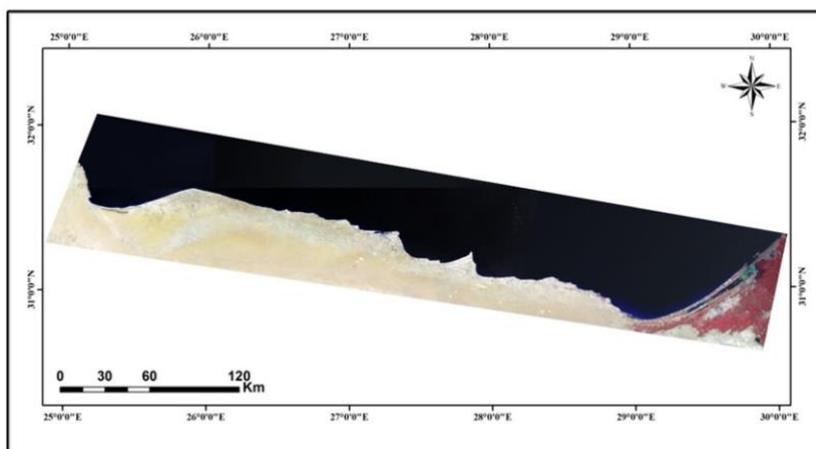


Fig. 2. Sentinel satellite image 2018 of the whole area.

- A number of thematic maps were obtained from different sources. (eg. 53 Topographic maps of Military survey authority (MSA) at scale 1: 50000). Other maps were extracted from different sources and were input as separate layers. These thematic layers includes water bodies, altitude points, contour lines, roads, railways, tracks, pipelines, telephone lines network and mine fields distribution.
- Field missions were carried out with the purpose of collecting ground truth information concerning landscape, soil and vegetation. A number of 149 observation

sites were comprehensively field studied, where different environmental parameters were described according to [7]. Representative soil and water samples were also collected manual from different horizons for laboratory analysis using the soil survey laboratory methods [12].

- Digital Elevation Model (DEM) of the study area has been obtained from the SRTM images (Fig. 3).
- Rectification of studied scenes, (ETM 2001) was performed using sufficient number of GCP's, which are distributed randomly all over the images. The root mean square (RMS) error was found to be 0.74, the process was applied first on the ETM of 1990 and hence, image to image registration was accomplished.
- Arc\Info 7.2 software was used to create GIS coverage's from the CAD file of thematic layers. The same system was used for map features coding, editing, building topology, creating feature attribute tables FAT, edge matching and map projection. Generating check plots, compared with source maps, was fulfilled for quality assurance. This helped in detecting and editing digitizing mistakes. Join item function was used to link the tabular attributes with the spatial features.
- Arc View 3.2 software was used in data analysis, the first step in analysis began with locating the field observation sites on the thematic layers with their attributes (i.e. soils and landscape properties). Using the 3D module of Arc View the interpolation of the spatial distribution of the land use classes was performed. Spatial analyst of Arc View was used to classify the soil parameter ranges on the map and deduct the relation between the soil conditions and the land features. Also, the 3D analyst was used for generating digital Elevation Model (DEM) from the contour lines and spot heights. The DEM creation depends on the nearest neighboring function. Statistical parameters and presentations were used to find out the relation between ground truth and image classification.
- Arc- GIS 9.0 software was used for this function. Sentinel 2018 images and Digital Elevation Model (DEM) were grouped and processed in ERDAS Imagine 8.7 software to define the different landforms of the studied area [4] and [13].

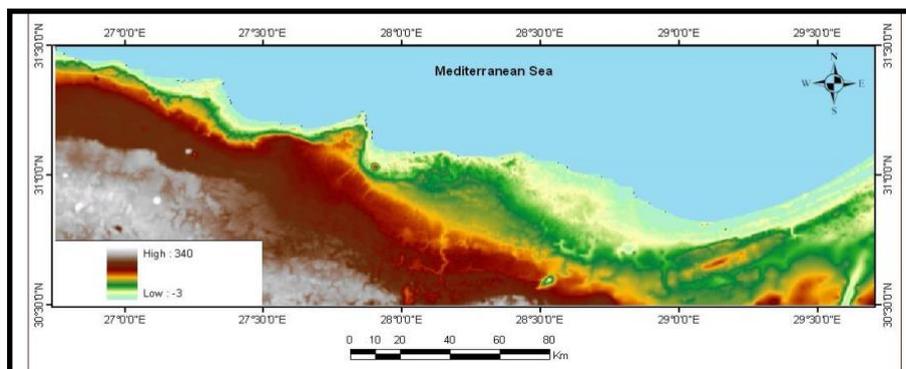


Fig. 3. Digital elevation model of the studied area.

4. Results and discussion

4.1. Producing base layers

The planned schedule was completely fulfilled for this task, as it includes digitizing of 53 topographic map sheets. The maps performed to be available in the digital format; however, their preparation as GIS ready maps has been completed. The mosaic of database layers is represented in Fig.4 – Fig.10.

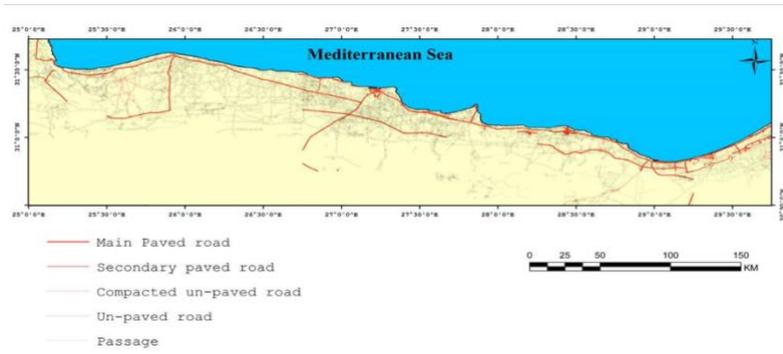


Fig. 4. Roads networks in the Northwestern Coast.

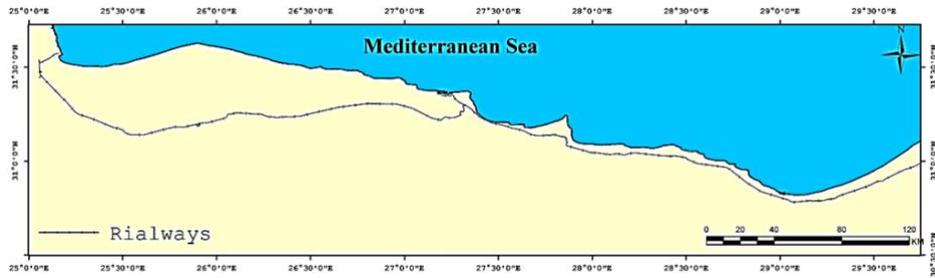


Fig. 5. Railways networks in the Northwestern Coast.

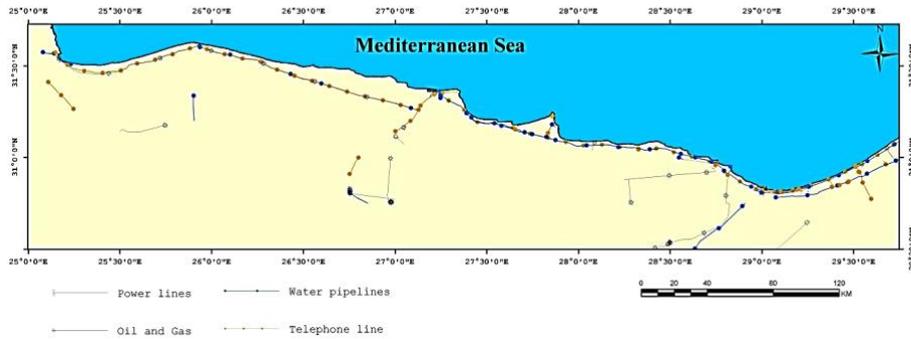


Fig. 6. Utilities layer in the Northwestern Coast.

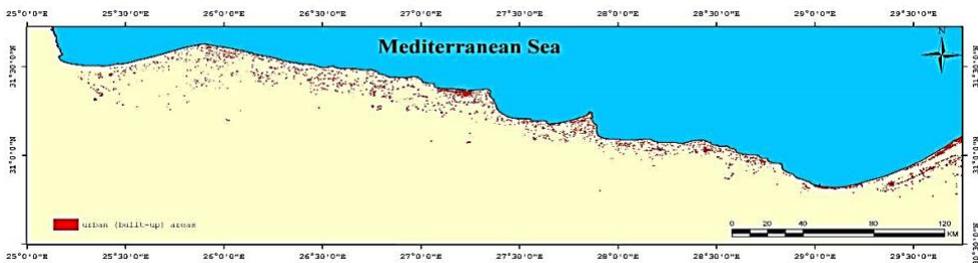


Fig. 7. Urban areas in the Northwestern Coast.

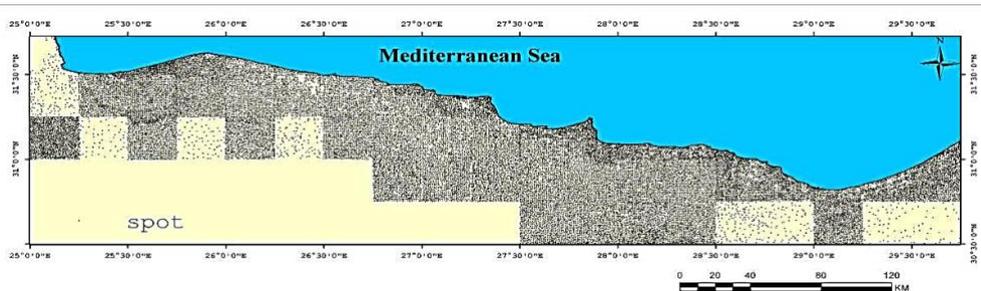


Fig. 8. Spot heights layer in the Northwestern Coast

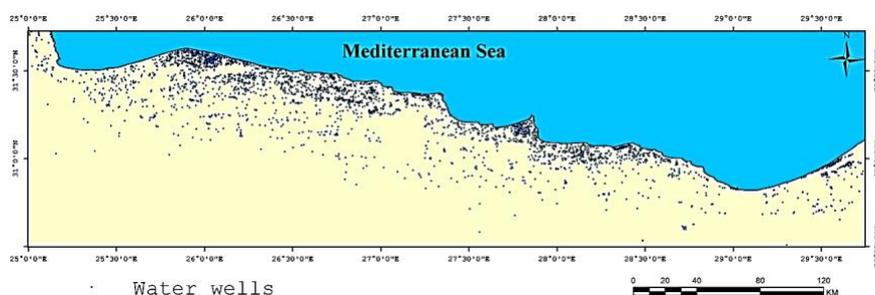


Fig. 9. Wells distribution layer in the Northwestern Coast

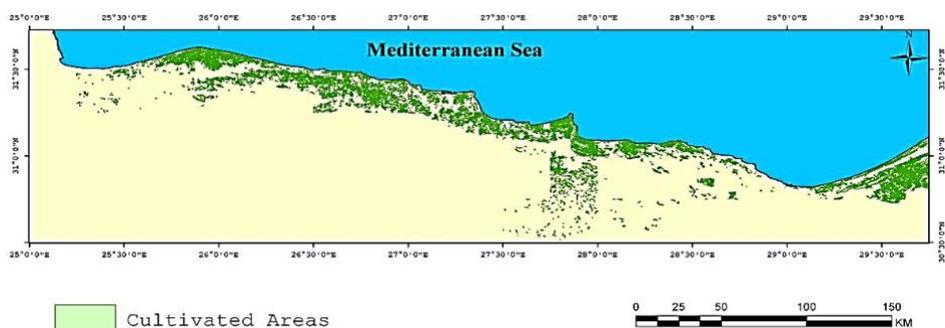


Fig. 10. Cultivated areas in the Northwestern Coast

4.1.1. Defined the physiographic units of the studied area

Physiography is assumed to be one of the driving soils forming factors and soil mapping criteria, concepts provided by this discipline can conveniently be used for soil data structuring. The combination of the geomorphic approach as a hierarchic classification system of geofoms using the existing body of knowledge in geomorphology, with the satellite data and field observations improved the results and allow us to use the computer-assistance procedures.

The delineation of the physiographic units from the satellite data needs a high spatial resolution image; therefore, the spatial resolution of the used Sentinel was enhanced through the data merge process. This process is commonly used to enhance the spatial resolution of multi-spectral datasets using higher spatial resolution band data or single band [10]. In this study merged data were performed using multi-

spectral bands (20 m) as a low spatial resolution with band 8 of Sentinel satellite image as a high spatial resolution (10 m) resulting in multi-spectral data with high spatial resolution. The enhanced image was draped over the Digital Elevation Model (DEM) of the area to delineate the physiographic map. The physical and chemical analyses of the studied soils were linked to the attribute table of the mapping units.

The physiographic description of the investigated observation sites shows that the relief in the study area ranges between almost flat to slope, while the lithology varies from marine deposits in the coastal plain to Aeolian deposits in plateau landscape and colluvium in the basins, terraces and slope ones. The dominant land cover is sandy sheets in the coastal plain. Scattered areas are cultivated with fig and olive trees in both coastal plain and plateau. Gravel surface and low dense shrubs exhibit the plateau, while highly dens shrubs exist in the basins. Boulders and stony fragments often exist in the gently sloping areas. The altitudes in the coastal plain ranges between 12 and 19 meters a.s.l., while in the plateau ranges between 92 and 185 meters a.s.l. Fig. 11 represents the different physiographic units in the studied area; the obtained data reveal that the area includes three main landscape units as the following.

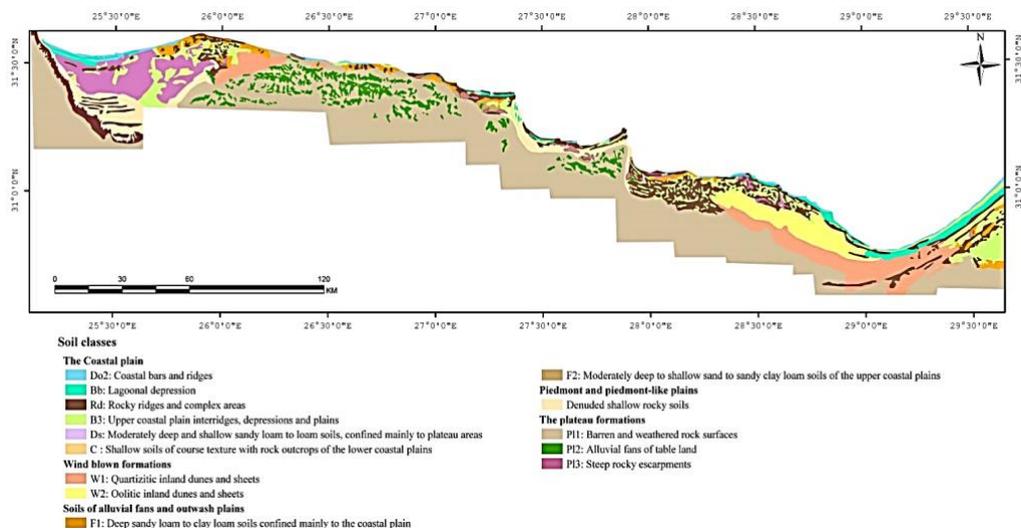


Fig. 11. The main physiographic units in the Northwestern Coast

4.1.2. Coastal plain

This type of landscape is found near to the coast of the Mediterranean Sea and includes different types of land forms such as sand sheets, terraces, vales and basins. It is characterized by the low elevation as the elevation differ from zero to 20 m a.s.l. the elevation increases generally in the southern parts of the coastal plain. This type of landscape contains sand and colluvial deposits with almost flat to gently undulating relief type.

4.1.3. Piedmont

This type is located between the plateau and the coastal plain and has an elevation range between 20 to 65 m a.s.l. The main land forms in this type of landscape are terraces, basins and sand sheets of gently sloping to undulating relief type.

4.1.4. Plateau

The plateaus are found in the south of the study area and have an elevation ranging from 65 up to 275 m a.s.l. and characterized by the limestone deposits as desert pavement and rock outcrops. It was possible to confirm that the rocky surfaces exhibit the plateau landscape. It was found that the surface relief of the plateau type differs from almost flat to undulating relief.

The detailed description of the landscape, relief, lithology, land forms and laboratory analyses are attached to the attribute table of the physiographic digital units (Fig. 12). The cultivations in the investigated area are wheat, barley, fig and olive trees as rain fed cultivation. The cultivated lands are found in the coastal plain and large parts of the piedmont because of the relatively high number of annual rains and the absence of rock outcrops. The cultivation activity is found in the vales, basin, terraces and sand sheet area. The grazing activity is found in different areas depending on the density of the natural vegetation.

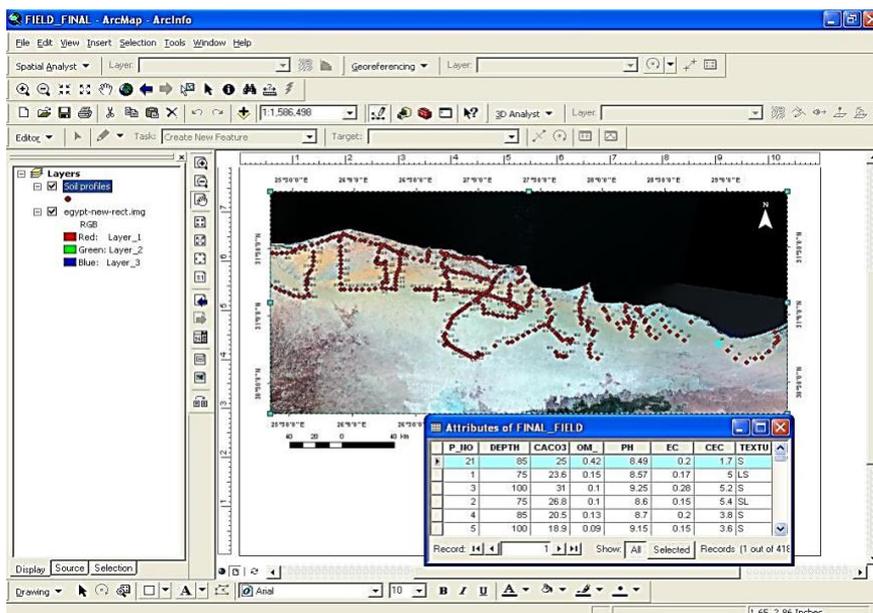


Fig. 12. The geographical distribution of the representative soils profiles with attributed data.

4.1.5. Compilation of digital soil map

The northwestern coastal region attracted the attention of several investigators and with the advantages of satellite images of the earth, reviewing of the previous work was necessary. In view of available resolution (30m) and on regional scale, the soil maps produced formerly were modified as shown in Fig. 13 [6, 7 and 8].

Since 1960, several colleagues in the Desert Institute studied separate parts. Most of their studies were accomplished by aerial photo interpretation. The modern technique at that time provided an adequate and excellent tool as base maps and for interpretation. The soil maps presented were highly predictable and credible.

The soils of the region in general are highly calcareous as the dominant rock is limestone. However; existence of sandstones and shales is reported. The following origins of these soils were identified.

1. Marine origin for the oolite sands of the ridges and dunes either consolidated or loose.
2. Alluvial and /or fluvio-marine origins, for the soils of the coastal plains, alluvial fans and depressions

3. Lacustrine origin for the sediments of lagoons and the deep lagoonal deposits
4. Aeolian origin for the sand dunes, hummocks and sheets of some tracts along the region.

The soils of the studied region are classified according to the Soil Taxonomy, Table 1 and Fig. 13 show the geographical distribution of soil units in the Northwestern coastal region. Both Aridisols and Entisols soil orders are found covering 42.1 and 57.9% of the mapped soils respectively. The Calcids sub-order is mostly clustered in areas of Burg El-Arab, Marsa Matrouh and Sidi Barani exhibiting 14.45% of the area. The Salids sub-order exist around both Matiout and Salum lagoons representing an 6.11% of soils The Gypsids sub-order soils is restricted in the area between El-Hammam and Sedi Heneish covering 21.54 % of the mapped soils. The Entisols soil order includes the sub-orders Orthents and Psamments representing 33.99 and 23.91% of the mapped soils respectively. It can be noticed that the area from Burg El-Arab to Matrouh is characterized by variability of soil units (e.g. Torripsamments, Torriorthents), while the area from Matrouh to El-Saloum is occupied by the Torriorthents great group soils.

Table 1. Areas of sub-great groups, in the Northwestern coast of Egypt

Order	Sub-Order	Type (Great group)	Area (Km2)	%
Aridisols	Calcids	Haplocalcids	542.17	14.45
	Salids	Haplosalids	229.4	6.11
	Gypsids	Petrogypsids	808.45	21.54
Entisols	Orthents	Torriorthents	1275.57	33.99
	Psamments	Torripsamments	897.14	23.91
Total area			3752.73	100.00

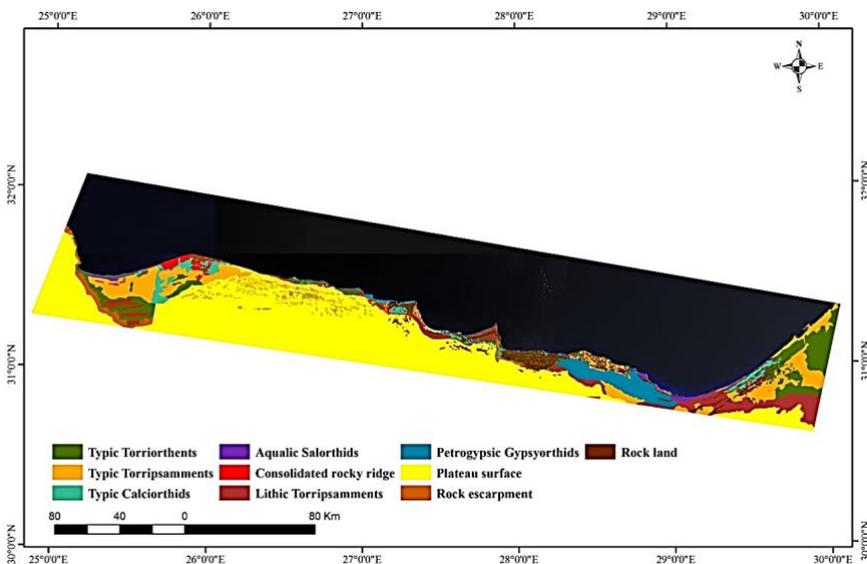


Fig. 13. Soil map of the North Western Coast region.

4.1.6. Application of the digital soil map in crop suitability classification:

Land capability and suitability for some crops (i.e. wheat, melon, sunflower, olive, peach and alfalfa) was determined using MicroLeis software [1]. The obtained data were linked to the attribute tables of the established database for defining the land suitability, Arc-GIS spatial analyst was used perform this task.

The soil characteristics such as soil depth, texture, calcium carbonate content, salinity, natural drainage and development of the soil profile were used in this system to determine the suitability class. Results of this determination are demonstrated in table (2). The suitability of 67 soil profiles representing the different soil types in the studied area was carried out; the results indicate that the olive, peach, wheat, melon and sunflower are the most suitable crops in the study area. Fig trees and barley are already found in the area with a good productivity. The soils in the area were classified to five groups (S2, S3, S4, S5 and N) according to their suitability classes and limiting factors. The map shows that the southern part of the study area is non suitable (N) for cultivation due to its very shallow depth and very high content of calcium carbonates as well as the domination of rock outcrops. The soils of piedmont and coastal plain have a good potentiality for cultivating by the selected crops, as they are classified S2 to S5. These classified soils are characterized by the following:

- Useful depth: 25 to 100 cm
- Soil texture: sandy to sandy clay loam
- Salinity: 0.25 to 31.5 dS/m
- CaCO₃ : 10 to 57.1%
- Natural drainage: poor to excessively
- Stoniness: 0.2 to 25%
- Development of the soil profiles: incipient

The limiting factors in the soils of the piedmont and coastal plain are salinity, soil depth and texture. These factors decrease the suitability class to S2, S3, and S4 and sometimes to S5. The classes of S2, S3 are found mainly in the coastal plain where the classes of S4, S5 are exhibited in the soils of the piedmont.

In general, the investigated area could be cultivated by wheat, olive and peach with a suitability class (S2), however melon, sunflower and citrus with suitability classes S3, S4 and S5.

The barren lands are covered with gravels, stones, boulders and few patches of natural vegetation (small shrubs). These patterns of land cover are found mainly in the plateaus and the high parts of the piedmont.

Table 2. Limitation factors and land suitability classes of the studied soil profiles

Profile no.	Limiting factors	Suitable crops	Class
1	Useful depth, Texture, Drainage	Wheat	3, 4
2	Useful depth, Texture, Salinity	Wheat	3, 4
3	Texture, Carbonate %	Olive, Peach,	2
4	Useful depth, Texture, Drainage	Non	5
5	Useful depth, Drainage	Wheat	3, 4
6	Useful depth, Texture	Olive, Wheat, Melon, Peach	2, 3
7	Texture	Olive, Wheat, Melon, Peach	2, 3
8	Useful depth, Texture	Non	5
9	Useful depth, Texture, Drainage	Non	5
10	Useful depth	Wheat	3, 4
11	Texture	Olive, Peach, Alfalfa, Melon, Wheat	2, 3
12	Useful depth	Wheat	3, 4
13	Texture, Carbonate %	Olive, Peach, Wheat, Melon	2, 3
14	Texture, Carbonate %	Olive, Peach, Wheat, Melon	2, 3
15	Texture	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2

16	Non	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2
17	Carbonate %	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
18	Useful depth, Texture, Carbonate%	Olive, Peach, Wheat, Melon	3
19	Texture	Olive, Peach	4
20	Salinity, Useful depth, Drainage	Non	5
21	Texture, Salinity	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	3
22	Non	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2
23	Texture	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
24	Texture	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
25	Non	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2
26	Texture	Olive, Peach	3
27	Carbonate %	Olive, Peach, Wheat, Melon,	2, 3
28	Non	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2
29	Non	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2
30	Texture, Carbonate %	Olive, Peach, Wheat, Melon	2, 3
31	Texture	Olive, Peach	3
32	Useful depth, Texture, Carbonate%	Olive, Peach, Wheat, Melon	3, 4
33	Carbonate %	Olive, Peach, Wheat, Melon,	2
34	Texture, Carbonate %	Olive, Peach, Wheat, Melon	3, 4
35	Useful depth, Texture, Salinity	Non	5
36	Useful depth	Olive, Peach, Wheat, Melon	2, 3
37	Texture	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
38	Texture, Carbonate %	Olive, Peach, Wheat, Melon	2, 3
39	Useful depth, Drainage, Carbonate%	Non	5
40	Useful depth, Carbonate%, Salinity	Wheat, Melon, Sunflower	2, 3
41	Texture	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
42	Useful depth, Carbonate%, Salinity	Wheat, Melon, Sunflower	2, 3
43	Texture	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
44	Useful depth, Texture, Drainage	Non	5
45	Useful depth, Carbonate%, Salinity	Wheat, Melon, Sunflower	2, 3
46	Useful depth, Texture, Drainage	Non	5
47	Useful depth, Texture, Drainage	Non	5
48	Useful depth	Wheat	3
49	Useful depth, Drainage	Non	5
50	Useful depth, Texture, Drainage	Non	5
51	Texture	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
52	Texture, Salinity	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
53	Texture	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
54	Useful depth, Texture, Carbonate%	Wheat, Melon, Sunflower	2, 3
55	Useful depth, Drainage	Non	5
56	Useful depth, Drainage	Non	5
57	Useful depth, Drainage	Non	5
58	Useful depth	Non	5
59	Carbonate%	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	2, 3
60	Carbonate%	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	3, 4
61	Carbonate%	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	3, 4
62	Useful depth	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	3, 4
63	Useful depth	Olive, Peach, Wheat, Melon, Alfalfa, Sunflower	3, 4
64	Useful depth, Drainage	Non	5
65	Useful depth, Drainage	Non	5
66	Useful depth, Drainage	Non	5
67	Carbonate%	Olive, Peach, Wheat and Melon	3, 4

5. Conclusion

It can be concluded that the digital mapping of land resources encouraged by the progress of Geographic Information System (GIS) and data provided by satellite

images. Such approach may preserve in the investment spent in soil and other thematic mapping, as the digital maps are more granted compared with analogue ones. Updating and manipulating the digital thematic maps are accessible and economically effective. Usage of the digital maps and their attribute tables assist the decision support systems and may result in obtaining maps required for controlling sustainable development projects. The digital format of the soil map facilitate the linkage with the different software, this allow the integration of data for defining the optimum land uses of the studied region. The obtained results from the established database recommend that the soils of alluvial fans and watershed basins are most suitable for olives, peach, wheat, beans, and sunflower cultivation.

References

1. Berger, K., Atzberger, C., Danner, M., D'Urso, G., Mauser, W., Vuolo, F., Hank, T., (2018). Evaluation of the PROSAIL model capabilities for future hyperspectral model environments: a review study. *Remote Sens. (Basel)* 10 (851).
2. Bishop, T.F.A., McBratney, A.B., Whelan, B.M., (2001). "Measuring the quality of digital soil maps using information criteria." *Geoderma*, 105, 93– 111.
3. CONOCO, (1989). "Stratigraphic Lexicon and Explanatory Notes to the Geographical Map of Egypt 1: 500,000" Editor: Maurice Harmina, Eberhard Klitzsch and Franz K. List, Conoco Inc., Cairo, Egypt.
4. Dobos, E. Norman, B. Bruee, W. Luca, M. Chris, J. and Erika, M. (2002). "The Use of DEM and Satellite Images for Regional Scale Soil Database" 17th World Congress of Soil Science (WCSS), 14-21 August 2002, Bangkok, Thailand
5. Egyptian metrological Authority (1996). "Climatic Atlas of Egypt" Published., Arab republic of Egypt. Ministry of transport.
6. FAO (1970). "United Arab Republic: Pre-investment survey of the northwestern coastal region" Comprehensive account of the project. Technical Report 1 ESE: SF/UAR 49: 109 P., 5 Maps.
7. FAO (2006). "Guidelines for soil description" Fourth edition, FAO, Rome, ISBN 92-5-105521-1.
8. Hammad, M.A., Haraga, A., El Shazly M., and Omar M. (1981). "Morphopedological studies of the soils of North Western Coast of Egypt" *Egypt. J. soil sci.* 21, No. 1 pp. 79-97, 1981.
9. Hastie, T., Tibshirani, R., Friedman, J., (2001). "The elements of statistical learning: data mining, inference and prediction." *Springer Series in Statistics*, Springer-Verlag, New York.
10. Lanaras, C., Bioucas-Dias, J., Baltsavias, E., Schindler, K., 2017. Super-Resolution of multispectral multiresolution images from a single sensor. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* 1505–1513.
11. NARSS (2005). "Environmental Evaluation of Land Resources in the Northwestern Coast of Egypt, Using Space Data and Land Information Systems" Final report, National Authority for Remote Sensing and Space science (NARSS), Cairo, Egypt.
12. USDA, (2004). "Soil Survey Laboratory Methods Manual" Soil Survey Investigation Report No. 42 Version 4.0 November 2004.
13. Zinck, J.A. and Valenzuela, C. R. (1990). "Soil geographic database: structure and application examples" *ITC j.* 3, 270.

Improving Business Processes by Applying the Kaizen Philosophy in a Macedonian Textile Company

Elizabeta Mitreva ¹[0000-0002-1816-2185], Aneta Janeva¹

¹ "Goce Delčev" University, Štip, Republic of Macedonia

https://doi.org/10.33847/2686-8296.4.1_4

Received 19.04.2022/Revised 02.05.2022/Accepted 22.05.2022/Published 12.06.2022

Abstract. This paper presents the research aimed at developing a solution for advancing of business processes in a Macedonian textile company, by implementing techniques and methods of the Kaizen philosophy. The main objective of the paper is to make a comprehensive analysis of the factors that enable the improvement of the overall production process, thus achieving greater effectiveness and efficiency in the production operations, as well as greater market competitiveness. The solutions were based on the concept of improving the efficiency of operations, especially in the area of organization of production operations, in a way that enabled complete elimination of errors, greater productivity and increased quality of the final product, within a precisely defined time frame. Additionally, the paper also shows the key elements from the implementation of the Kaizen philosophy in the process of competitiveness improvement of the company.

Keywords: Kaizen philosophy; management tools and techniques; business processes; textile company; business performance.

1. Introduction

The textile production is one of the best developed industries in RN Macedonia, but only with the application of the loan - system of work. The very process of organizing these processes begins with the signing of a cooperation agreement between the textile company and a partner company from one of the Western European countries. The partner company is committed to full and continuous supply of raw materials, textile, accessories and everything needed for a smooth flow of production, while the textile factory entirely organizes the production process – from the receipt of raw materials to the final control of the final product and its preparation for export and delivery to customers, Fig. 1.

The efficient operation of the textile company is a key factor in determining its competitiveness, which is a set of multiple external and internal factors [1]. The company cannot directly influence some of these factors, because they are determined by the environment, the business climate in the country and the global economic trends [2]. However, a good part of the factors that determine the competitiveness of the company is in the direct hands of the company itself, that is, its management [2]. One such factor is the efficiency and effectiveness of the day-to-day operations in the overall production [3]. Increasing the efficiency of this factor directly affects the efficiency of the company as a whole.

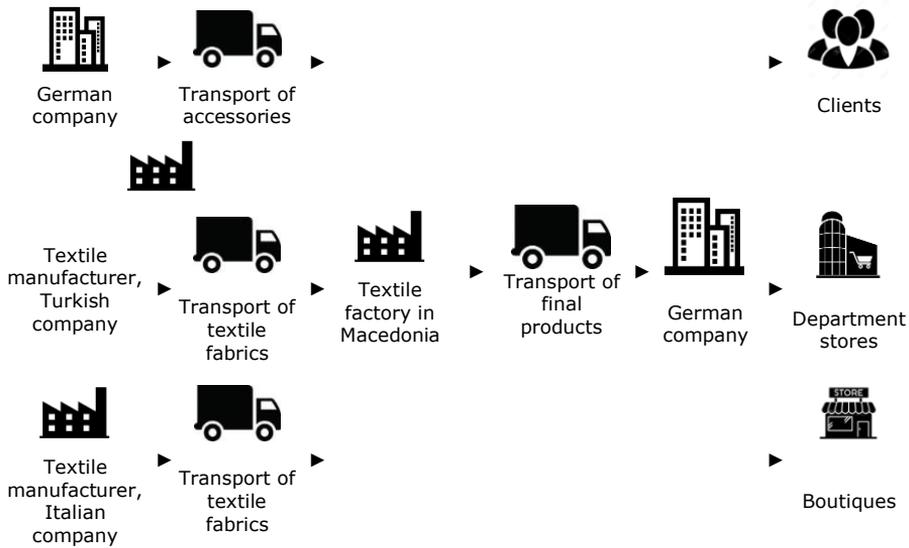


Fig. 1. Textile-loan production in RN Macedonia – starting with basic textile fabrics and accessories supply to delivery of the final products

There are several management tools and techniques to deal with this challenge, the best of which is the Japanese philosophy called Kaizen. It provides a systematic approach to workplace organization and is a sophisticated mix of organizational efficiency, competitiveness and organizational survival [2, 3]. Fig. 2 summarizes the organizational approach to the realization of business processes in the Macedonian textile organization.

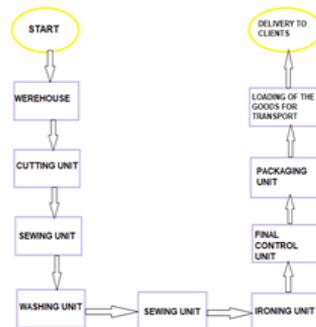


Fig. 2. Production flow diagram in a textile company
Source: Own sources.

The main topic of the research in this paper is to analyze the factors that determine the improvement of the operation of a Macedonian textile company and its competitiveness through the improvement of business processes in all aspects. The aim of the research is to determine the benefits and results of the adoption of the Kaizen philosophy and its methods and techniques, which are of strategic importance for continuous success in the implementation of business processes [3, 4].

2. Literature Review

Competitiveness of a company is the ability to produce a product or service that meets international standards accepted by the international market, while ensuring sustainable growth and well-being of the employees [5]. Competitiveness of a company is achieved by constantly increasing productivity which results in higher profits, wages and development, increased domestic income, constant quality and innovation in operations, acceptance of new technologies, information and connection to the world and economic operations [6] [7].

The organization of the work has a key role in the whole process of improving the business processes and the competitiveness of the company [8]. The management of the company must review and establish the organizational and business processes in a new environment and at a higher level [9]. It is extremely important to establish modern work management and separation of the managerial and ownership functions and to accept new methods and techniques of work which will include all employees [10].

Hence the need to create a concept for total quality management in all segments of operations, its planning, control, improvement – all in order to create a top-quality product or service that will meet and exceed the needs and expectations of the market [11].

Kaizen is a Japanese philosophy that is a set of tools and techniques to improve the quality of business processes, products and services [12]. The Japanese philosophy is based on small and continuous process improvements, which increase the efficiency of the organization and production and achieve greater results, by including all members of the organization, regardless of their hierarchical level, without making large capital investments [13, 14, 15].

3. Data and Methodology

This methodology starts by taking the initiative by the management or the management structure of the company for use of modern tools and techniques for quality improvement in order to achieve the planned objectives [16]. The main reasons of the management are the large percentage of defect pieces of the final product and the quantitatively insufficient number of finished pieces in a precisely determined time frame [17].

The implementation of the preparations for the activities begins with the establishment of teams for improvement, problem detection and selection of Kaizen techniques. This is followed by defining a plan of activities and setting an objective and projection for the outcome of the planned corrective measures [18]. Quality control tools and techniques are used in order to set measurements at the critical points, where results are monitored and ongoing control is set [19]. After receiving and analyzing the results, an internal analysis is made and the business process are standardized. Finally, the self-assessment method is used to make the final assessment of the results and outcomes of the designed and applied methodology [19] [20] [21], Fig. 3.

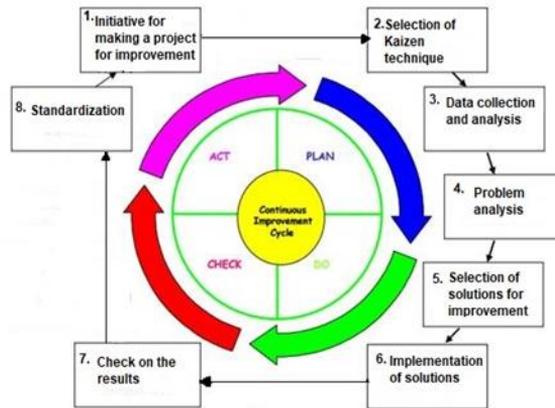


Fig. 3. Steps in implementing the Kaizen methodology [based on 17].

The distribution of activities when implementing the Kaizen strategy starts from the top management which is responsible for its introduction, allocation of resources and determination of policies and objectives for full realization of the concept [16]. Secondly, the middle management is in charge of implementing the objectives set by the top management. This management level designs the operating procedures and trains the employees to apply the Kaizen and development of appropriate techniques and skills. Thirdly, the workers are required to apply all of Kaizen methods and techniques to which they are trained [15, 17, 20, 21]. A great discipline is required in their work as well as continuous education and self-development in order to better resolve all the current problems [21]. This will ensure continuous improvement, increased productivity and better quality of the entire production [14]. Further in this paper we show the application of the Kaizen methodology and its tools and techniques for improving the business processes and the quality of work in a Macedonian textile company. The main objective of the research is to increase the productivity and the number of final products with minimal errors and defects in the production. The benefits from the application of this methodology are presented through the analysis performed for the individual operations of the textile production.

4. Results

4.1. Implementing the Kaizen philosophy in a Macedonian textile company

The Macedonian textile company based in the city of Shtip was founded 20 years ago and with its production and the established reputation is a real success story on the Macedonian market. With its modern production facility, modernly arranged plants and warehouse space as well as specialized machines for production of men's fashion pants, the company successfully does business and is already well known at the textile fairs in Europe. New fashion trends, international standards and customer requirements, as well as the shorter deadline for production and delivery of the finished product, inevitably impose the need to improve business processes and improve all segments in the textile production. The company's management made a strategic decision to adopt the Kaizen philosophy, as an effort of introducing new way of work and increase of the success performance.

The design and implementation of this philosophy in this company was realized through the following steps [17]:

➤ ***Step 1: Initiative for preparation of improvement project.***

The management of the company decided to take concrete steps by applying the technique of establishing teams for improvement. These teams are composed of members of the top management of the company, members of the middle management as well as representatives of the workers from each sector of the business processes. In that way, the company got the best picture of the internal state of affairs, in terms of what is good and what does not work, where are the impediments and which segment works properly. The main task of these teams was to develop an improvement project, to come up with a list of problems and to assess their impact and scope.

All business processes were identified and analyzed in which there is a need for improvement and betterment. After identifying the problems, the critical points and issues that need to be addressed, brainstorming was used as a management tool to generate ideas and possible solutions [12] [18]. Starting from the fact that every idea is a good idea, many ideas have been generated in order to solve the current critical points [12]. The ideas were coming from the management, but also from the people directly involved in the processes: brigadiers (persons responsible of the whole brigade/ sewing line, person responsible for cutting department, responsible person for warehouse, responsible person in the finalization department, technical persons, etc.).

After the end of the brainstorming meeting, the following key points were singled out which should be approached with an appropriate methodology, with aim to improve them:

- complete cleaning and arrangement of the warehouse space;
- arranging the machine part of the plant (sewing and cutting units);
- simplification of the work and organization of the unit for control and packaging of the finished/ final products;
- elimination of the reasons for the increased number of final products with defects/ defected garments in the production;
- production of a larger number of units/ pieces for a shorter period of time without defected pieces.

➤ ***Step 2: Selection of Kaizen techniques that have been applied in the improvement of the business process.***

The choice of the Kaizen techniques should be in line with the issues that require improvement. Given that the Kaizen philosophy is based on small but continuous changes in all processes and functions of the company that involve all employees in the organization, the success and improvement will definitely happen. This paper also elaborates several business processes which were pointed out by the teams to be improved with the application of appropriate methods and tools from the Kaizen philosophy [18].

a) Complete cleaning and arrangement of the warehouse space.

To solve this problem, the 5S Kaizen methodology was applied, which refers to the organization and maintenance of the workplace. The 5S method is a set of rules and concepts for organizing the workplace of employees, in order to be neat and clean and easier to operate. 5S Kaizen is an easily applicable technique that is realized through the following stages [15] [16] [17]:

- sort (seir), which means: pick up or remove the unnecessary items, but also arrange and dispose of them;

- set in order (seiten), which means: arrange the necessary items for easier and faster use;
- systematic cleaning, which means: clean your work equipment, workplace and environment;
- standardize, which means: maintain a high standard of keeping your domestic and your own order;
- constant observance of the rules (self-discipline), which means: spontaneously maintain the objects and the environment without someone ordering you to do it.

The 5S Kaizen methodology in the company was realized by throwing out everything that is not needed in the warehouse space, including remnants of fabric rolls and rolls of interlining or lining that are small in size and pertain to orders that have been already completed. If their reuse is impossible, they just take up space of new rolls coming from new orders. Every roll was properly tagged with the number of the order that was showing where it is to be used and with the S-number of the customs import declaration for easier and more appropriate dealing in the processes of further production. The rolls were properly arranged on the shelves according to the order number, size, color, etc., in order to find them more easily when needed. Cleaning and arranging the accessories, although time consuming, was still an extremely important process for simplification of the search and finding everything you need at any moment.

International research says that employees lose half of the time allotted for work in search of what is needed [13]. In order to minimize that process, the first thing that was done was selection of the types of items, the material from which they are made, the quantity of the packaging, etc. Each item was separately selected and divided into accessories boxes, which are tagged with an appropriate code, order number and quantity. The tagging of the items was in the following order: necessary, critical, most important, unusable, not for use now and so on. The boxes were then stored on shelves for visually easier retrieval. With these procedures, more than half of the work was completed. Then it is up to the staff, responsible for the warehouse section, to maintain and arrange it on a daily basis, to dispose of the unnecessary items on a monthly basis and to take care of all the new goods that arrive - to properly distribute, select and order according to the already established principles. The self-discipline of the staff in this aspect of the improvement was essential for the successful implementation of the Kaizen philosophy.

b) Red tag

Identifying the unnecessary parts, tools or materials found in plants of an organization is not always an easy task. The employees and the managers are accustomed to the chaos. The use of the so-called red tag in the company, especially in the warehouse section, was a great way to identify all the unnecessary items that needed to be dislocated from the workplace. With the help of this tag, everyone else was informed that the tagged item should be evaluated [10], Fig. 4.



Fig. 4. Example of a Red Tag.

Source: Own sources

The implementation of the 5S Kaizen methodology within the unit for receipt and distribution of goods was completed by applying the method of visual factory, which provided normal working conditions in the company, that is, everything necessary for realization of the business processes was placed in a visible place. In this way, standards were created in the workplace, but also quick detection if something is out of the plan.

c) Arranging the machine part in the plant (sewing and cutting department).

The purpose of this segment was more efficient and effective operation, without interruption in the production process in the sewing department and the cutting department, by saving time (a worker on one machine waits while another worker on another machine finishes her part of the work) and increasing of the number of final products in a shorter period of time. In order to solve this problem, the management applied the SMED and OTED Kaizen methods. These techniques are performed sequentially, one after the other, and allow the tool/ machine change time to be reduced.

The SMED (Single Minute Exchange of Die), aims to reduce the tool change time to less than 10 minutes, and OTED (One touch exchange of Die) allows the tool to be changed with just one touch, that is, by pushing a single button [14].

In the textile production, there was no change of the tool but a change of the activities in the production processes from one type of machine to another. For example: the legs of the pants are sewn at one machine, and the leg seam is made on another. A third type of machine makes the belt of the pants.

All these operations were fully coordinated according to the minutes required for making each separate part of the pants, in a way that will provide for zero downtime, that is, no worker will wait from another worker to do her part. Because only men fashion pants are sewn in this textile company, the principle and operations are almost identical in almost all models, with small differences in the number of pockets, decorative seams or some effects (which are additionally completed). Therefore, the layout of the sewing machines (by type of operation) could have been distributed and standardized, thus completely facilitating and accelerating the overall production. If we take into account that for the most important thing for the client is to meet the delivery deadline (between 2 and 3 weeks after receiving the raw materials and fabrics in the factory warehouse), then the reduction of the production time is extremely important.

The 5S Kaizen methodology is also applied in the production processes implemented in the cutting and sewing units. Cleaning the workspace in the facility and around each machine individually and on a daily basis is especially important. For this purpose, all waste that is generated every day was selected, separated and what could be reused and combined was tagged in separate boxes. It took only 15 minutes of less work (compared to the 8 hours of active work planned) for each worker to tidy up his/her workplace and waste. The management motivated its employees by rewarding them on a weekly basis. Most often, they were divided according to the brigades (sewing lines) in which they work, and the reward was appropriate to their needs: days off, extra working hours paid, cash bonus, etc. In this way, better and faster production was achieved, including satisfied and motivated workers, tidy and proper work space, and thus the time needed to make each model was reduced.

d) Simplification of the work and organization of the department for control and packaging of finished/ final products.

The department for final control and packaging of the final products is one of the key links in the entire production process. The further cooperation with the client after the products are delivered to the final destination depends on the quality of the work of this unit. Poka Yoke - Kaizen visual tool [17] – was implemented for implementation

and realization of as many final products as possible, without errors and defects and within shorter time of delivery. Its main purpose is to design a flawless or error-free process by removing all potential areas where an error may occur. The basic value of the Poka Yoke technique is seen when performing some tasks for the first time, making it impossible for an error to occur [12].

The application of this tool in the unit for final control and packaging of final products proved to be very effective. First, errors were identified, then appropriate norms and rules for control and manner of efficient packaging were established, which resulted in increased efficiency and time effectiveness of these processes.

e) Elimination of the reasons for increased number of final products with errors/ defect pieces in the production.

During the production, it can happen that some of the final products have some error or defect, which may be smaller or larger, technically permissible or not. However, the emergence of an increasing number of products with error inevitably leads to the need to identify the causes and resolve them. Following the detailed analysis, statistics and complaints by the client, the following reasons for this phenomenon were identified:

- "Heavy" or problematic material/ textile fabric to work with. Due to its composition, the fabric often causes a number of problems such as slipping through the machine, tearing, missed threads, factory stains, etc.
- All this causes delays or problems in the process of tailoring, and further in the process of sewing.
- malfunction of sewing machines or their parts due to irregular servicing.
- In order to achieve flawless operation without downtime and problems, the machines need regular maintenance and servicing.
- insufficient expertise and training of the working staff.

Due to lack of professional and qualified staff and increased workload (especially in the season when the clothes collections for the clients are made), people were employed but they were not sufficiently professional and capable of working without errors. The errors were inevitable, and the repair of the defect pieces cost much more time and money than the quality and continuous training of the staff.

To reduce the defect pieces of the finished product, it was necessary to prevent those pieces from happening in the first place. In the prevention, the management adopted the principle of Kaizen technique - Gembutsu, whose implementation means a detailed inspection and control of the following: textile (arriving directly from the supplier), inventory, machinery and tools before the start of the order [10]. Input control was realized in such a way that the textile is completely rewound and the possible defects, holes, missed threads or the composition of the fabric (which could cause further problems during production) were immediately detected. The machines and tools were regularly serviced, lubricated, leveled and the worn parts were replaced. Sometimes, for various reasons, significant time is lost for a specific sewing procedure, and the client is not aware of that at all. If this significant additional time is not included in the final price for the Lohn - production of that item, then it is a net loss for the textile company. Gembutsu is the perfect tool for detecting exactly such things [3] [9]. For each order, detailed norms are prepared separately with a description of all individual operations, which can help to accurately determine whether there is some specific procedure in place and how much it should be evaluated.

f) Production of a large number of items in a shorter period of time without errors and damaged pieces.

The basic and most important objective in the textile production is to produce as many final products as possible in the shortest possible time, by observing all the

conditions for quality of the products and deadlines for delivery to the final destination - the client. Achieving this requires a perfect organization of the steps before and after production, and they do not depend on the textile company but on external factors, companies and associates. The course of the whole process takes place in the following order: The client orders the basic and auxiliary materials for the production from his suppliers. The materials arrive at the client and are then sent to the textile company that does the production. This includes a transport company that imports and exports the final products. After the customs clearance of the materials, they are taken to the warehouse unit of the textile company and from there the production process follows. In order to achieve the best results, especially in meeting the deadlines, the Kaizen strategy JIT (Just in Time) is the one that is the most appropriate [17, 18, 21]. This management strategy is based on aligning the orders of the basic fabric or accessories with the dynamics and production plans. Companies adopt and implement this strategy to increase efficiency, meet deadlines and reduce the risk of delays of the materials, thus automatically reducing the overall costs and the wasted time. In order to be able to successfully implement this strategy, a quality and reliable supplier is needed, which will continuously supply the company with quality materials and will fully adhere to the deadlines. That way the client does not have to pay for warehousing or to waste valuable production time waiting for the necessary materials. On the other hand, the suppliers have a continuous delivery plan, planned at least a month in advance and thus do not remain without orders and gaps in their work.

In the Macedonian textile company, such a management strategy is of exceptional importance, because it is a matter of very short and fast production and delivery deadlines, and at the same time significant financial savings are made, because the order equals to the quantity planned for production. Applying this strategy saves warehouse space, waste management and disposal. The only drawback of this strategy is the potential delay, defect or disruption of the supply chain.

The Macedonian textile company that is the subject of research in this paper established business contacts directly with textile suppliers from Turkey, which enabled direct import of the required material and even further reduced the production cycle time. The application of Just In Time in this part of the production process provided huge savings of time and financial resources.

4.2. Benefits for the textile company from the implementation of the Kaizen modern philosophy

The implementation of the Kaizen measures and solutions was carried out in full coordination with the management of the company, the improvement teams, the responsible persons of the individual sectors and all the employees. The solution was a set of measures and activities which referred to implementation and execution of the following work activities:

- daily coordination of the activities planned by the management and the responsible persons in the separate sectors/units: warehouse, cutting unit, sewing unit, ironing unit, finishing and packing unit and the unit for loading of the goods;
- preparation of a daily plan for work with specific work tasks according to a previously determined agenda;
- complete cleaning, arrangement and adaptation of the warehouse space and of the plant with machines for cutting, sewing and ironing;
- raising awareness and self-discipline of employees;
- proposals for improvement of the overall work activity and motivation of

- the employees;
- continuous work in realization of the objectives aimed at achieving greater and better production with reduced costs and increased profits;
- continuous activity for finding new foreign strategic partners, clients and suppliers of materials;
- quarterly evaluations;
- modernization of the machine park;
- standardization of the work tasks.

Some of the activities of the working team in charge of implementing the Kaizen methodology were aimed at monitoring the consistent implementation of the measures and facilitating the implementation of the objectives. With the application of these measures, the responsible persons of the sectors started with greater preparations in order to carry out all the methods and tools determined according to the needs and expectations of each business process.

➤ **Step 3: Analysis and control of the results.**

The Kaizen philosophy is based on slow but efficient and long-term changes in all segments of operations, so the first visible results can be expected in the long run after its implementation. It takes time for adoption and acceptance of the whole strategy by the employees, the management as well as the teams initially created for its design, implementation and realization. The effects of the application of the Kaizen techniques and tools in the Macedonian textile company were visible several months after the implementation and showed excellent results on many different levels.

Using the 5S Kaizen technique, the warehouse space was completely cleaned and arranged in a way that put the following things first: functionality and visibility, easy access and complete organization of the basic material: textile rolls, interlining and lining, Fig. 5.



Fig. 5: Warehouse view before implementing the 5S Kaizen philosophy
Source: Own sources

Each roll was re-examined according to its size, the order number to which it belongs and the correctness of the material (in terms of holes, stains, missing threads, fading, etc.). The unnecessary rolls were removed in a special waste compartment, leaving a large space for everything to be neatly arranged on the shelves and tagged according to the required markings. It took a little longer to clean and arrange the auxiliary material, as it involved large quantities of buttons, metal buttons, zippers, labels, hangers, threads, and other plastic or metal utensils. It was arranged in boxes, drawers and packages with appropriate tags for the order number, the S-number of the customs declaration, quantity and item code. It was decided to arrange the rolls together with the auxiliary material (those that belong together) for each individual item and thus significant time was saved in searching for it through the warehouse space. If it used to take several hours and even one working day to make all the

selections, to arrange the basic and auxiliary material for each subsequent order that is put into production, now literally in 30 to 60 minutes everything was selected and the process of tailoring of the order was started. Cleaning and removal of leftovers from old and non-functional sewing machines, worn-out packaging material, loading pallets, and debris from long-completed and delivered orders was also done. The warehouse space has become functional at the highest possible level and the employees are investing daily care in order to keep it that way – the materials arriving from the imports for new orders on weekly basis are immediately selected, tagged and stored at the appropriate place, Fig. 6.



Fig. 6. View of the warehouse following the application of the 5S Kaizen philosophy
Source: Own sources

The arrangement of the machine unit was aimed at obtaining the highest possible productivity and effectiveness in the time required for preparation of each individual operation in the textile production. With the help of the SMED and OTED tools, a scheme of setting up, arranging the machines in a chain/ brigade for a more rational utilization time was carried out. In order to optimize the work of the brigade (and yet in accordance with the number of production operations required by the model of pants) it was determined that it would be best to have between 18 to 22 machines (plus machines for intermediate activities, ironing, etc.). It takes about 30 employees to complete all production operations on an average complex model of pants.

First, 12 ordinary single needle lock stitch sewing machines were allocated in order to produce a straight stitch. These machines are mostly used because 80% of the steps are made on an ordinary single needle lock stitch machine. Among these machines, 3 overlocking sewing machines were placed - machines that serve to assemble the inner stitches of the pants, that is, assembling of the individual parts of the pants, then 2 double-headed machines for making the decorative stitches.

Next to these machines, each of the following machines were placed: a bartack sewing machine, double needle lock stitch machine, button hole machine, flatlock cover stitch sewing machine, feed of the arm machine, fixer press machine, ironing machine etc.

After the complete sewing process of the model was completed, the following phases were also organized: thread cleaning, control, ironing, final control and packaging according to the client's order.

According to the norm given by the client, daily production of 400 pieces was expected for a simple model of pants, while for more complex models up to 300 pieces.

After applying the Kaizen techniques for optimal sorting of the machines in the brigade, the production on a daily basis increased by 12%, so about 450 pieces of the simple model of pants were made per day, about 350 pieces of the more complex model were made per day.

The 5S Kaizen methodology for complete cleaning of the workspace was applied in the unit with brigades and sewing machines. This methodology continued to be applied on daily basis as an integral part of the work responsibilities. The effects of the implementation are that the space around the machines is clean, each worker takes care of her micro space and every necessary utensil or material is easily and

immediately available.

The application of the Poka Yoke method in the company led to finding creative solutions to avoid or detect manufacturing errors with almost no investment.

The application of these activities in the process of final control and packaging of the finished product proved to be very effective. First, the errors detected in the control of the final products were identified: the control was carried out in the order, as the pieces were finished in the previous production operations. With already established and prescribed measures and norms of what the final piece should look like, no time was wasted in changing them if the piece was in a completely different size.

A similar activity was also carried out during the packaging process. If the packaging was in low order, first all the pieces of the same size were packed in bags, then they were tagged with a sticky label that showed all the data, such as order number, model, item, size, etc. According to the sizes and the previously determined space, the pieces were packed in cardboard boxes. If the packaging was to be on a hanger, the principle was similar – first, a bag of each product was packed, and then the pieces were lined up on hangers and stands according to the number of pieces in each size. In this way, time is saved and efficiency is achieved in every activity. The only deviation was made if the manner of packaging was changed by the client and a completely new way of stacking and packaging of the final products was required. The likelihood of errors in the use of this principle was indeed very small and the efficiency and, above all, the time efficiency were absolutely improved.

The Implementation of the Gembutsu methodology enabled the company to work much more efficiently in the field of prevention of defected final products and to provide detailed description of all work activities with accurate time required and complexity of operations applied in the production process of each individual item. This procedure provided a detailed norm which affects the determination of the final price for production of that item. Thus, neither side was damaged in the calculation of the final prices.

Using the Kaizen JIT strategy (Just in Time) outstanding results have been achieved in the field of ordering the basic fabrics, interlining and lining directly from manufacturers and suppliers of materials, most of which are based in Turkey. The company needed more time to establish all the correlations, conditions and rules for cooperation, to establish all customs rules for direct import from Turkey to the RN Macedonia (for which the two countries have concluded a special agreement for diagonal cumulation) and to make the necessary selections of the materials or the textile. Since the client has determined in advance the collection of models that will be made for that season, and already has the exact quantities per model that he will need, it was possible to make accurate calculations from which model, which color, quantity and type of fabric will be needed. Then, a list of textile was made in an order according to the production term of the respective model and a delivery was agreed once a week for quantities according to the needs of the current orders. This way of work simplified the production from the aspect of saving time and finances for transport of the ordered textile from the supplier in Turkey to the client in Germany and then from the client to the manufacturer in Macedonia. At the same time, there is room for the possibility of quick replacement or further delivery of the textile in case of errors or defects in the fabric itself or in case of need to sew urgent orders. In this way, timely and accurate delivery of the textile was achieved, which means absence of downtime in production, which also means that for each order it can be specified exactly when it will be made and delivered to the client, and the benefits are mutual for both the client and the producer.

5. Conclusions

The results of the implementation of the Kaizen methodology in the textile company for the most part contributed to the improvement of the business processes in production, but also to improvement in all segments of the business processes and change in the way the employees perform the work tasks and obligations.

At the same time, the textile company continued with the process of continuous improvements in other segments, which were aimed at greater investment in quality and changing the approach of the employees in performing their tasks. As a result of the positive effects of the implemented measures and activities, a standardization of the processes in the execution of the following work tasks was adopted:

- regular selection, tagging and proper storage of all imported shipments of basic fabrics and accessories in the warehouse of the production plant;
- regular updating of the location of the machine park in the cutting department as well as continuous servicing of the machines;
- daily cleaning of the micro space for each employee and greater self-discipline and pro activity of the employees;
- direct import of textile fabrics is established from manufacturers or suppliers of textile from Turkey;
- defining of dynamics and plan of final control and packaging of the final products;
- observation and timely removal of the causes of defect products;
- achievements of the monthly production plans and objectives;
- plan and realization of contacts with new and existing clients.

As a result of the changes implemented in all production processes, the responsible persons of each unit began to standardize their daily activities and to implement greater coordination in their teams. Continuous corrections of the weaknesses in the teams are made, by rejecting all the unnecessary activities, which are not aimed at making progress and do not contribute to the creation of new values.

The introduction of Kaizen modern tools and techniques in order to improve the business processes in the textile production enabled slow but continuous changes at all levels and in all areas of operation. The awareness of the management and all the employees was growing in a sense that something is being done every day that should be fruitful, that will improve the work, productivity, efficiency and effectiveness and will make the workplace a pleasant place to work. In the end, that would result in increased profits and higher wages.

The adoption of the Kaizen philosophy in the company has led to full customer orientation, which is a key factor in the textile production for successful production, gaining of reputation and following of all current fashion requirements and criteria. Hence the need for continuous progress and change, small but significant steps in every segment of daily work.

The implementation of the Kaizen philosophy is expected to yield even greater results in the long run because, over time, it turns into a way of work that is accepted and approved by all employees and management.

References

1. Carnerud D, Jaca C, Bäckström I. Kaizen and continuous improvement—trends and patterns over 30 years. *The TQM Journal*. 2018 Jun. <https://doi.org/10.1108/TQM-03-2018-0037>.
2. Iwao S. Revisiting the existing notion of continuous improvement (Kaizen): literature review and field research of Toyota from a perspective of innovation. *Evolutionary and Institutional Economics Review*. 2017 Jun;14(1):29-59. <https://doi.org/10.1007/s40844-017-0067-4>.
3. Álvarez-García J, Durán-Sánchez A, del Río MD. Systematic bibliometric analysis on Kaizen in scientific journals. *The TQM Journal*. 2018 Apr 10. <https://doi.org/10.1108/TQM-12-2017-0171>.
4. Alvarado-Ramírez KM, Pumisacho-Álvaro VH, Miguel-Davila JÁ, Barraza MF. Kaizen, a continuous improvement practice in organizations: a comparative study in companies from Mexico and Ecuador. *The TQM Journal*. 2018 Jun 11. <https://doi.org/10.1108/TQM-07-2017-0085>.
5. Chung CH. The Kaizen Wheel—an integrated philosophical foundation for total continuous improvement. *The TQM Journal*. 2018 Jun 11. <https://doi.org/10.1108/TQM-03-2018-0029>.
6. Gonzalez-Aleu F, Van Aken EM, Cross J, Glover WJ. Continuous improvement project within Kaizen: critical success factors in hospitals. *The TQM Journal*. 2018 Apr 20. <https://doi.org/10.1108/TQM-12-2017-0175>.
7. Jaca C, Ormazabal M, Viles E, Santos J. Environmental comfort based (ECB) methodology as a tool for preparing Kaizen application in a catering service company. *The TQM Journal*. 2018 Jun 11. <https://doi.org/10.1108/TQM-10-2017-0117>.
8. Al-Hyari KA, Zaid MK, Arabeyyat OS, Al-Qwasme L, Haffar M. The applications of Kaizen methods in project settings: applied study in Jordan. *The TQM Journal*. 2019 Oct 10. <https://doi.org/10.1108/TQM-03-2019-0078>.
9. Enshassi A, Saleh N, Mohamed S. Application level of lean construction techniques in reducing accidents in construction projects. *Journal of Financial Management of Property and Construction*. 2019 Nov 4. <https://doi.org/10.1108/JFMPC-08-2018-0047>.
10. Vinodh S, Antony J, Agrawal R, Douglas JA. Integration of continuous improvement strategies with Industry 4.0: a systematic review and agenda for further research. *The TQM Journal*. 2020 Aug 20. <https://doi.org/10.1108/TQM-07-2020-0157>.
11. Fonseca LM, Domingues JP. The best of both worlds? Use of Kaizen and other continuous improvement methodologies within Portuguese ISO 9001 certified organizations. *The TQM Journal*. 2018 Apr 10. <https://doi.org/10.1108/TQM-12-2017-0173>.
12. Debnath RM. Enhancing customer satisfaction using Kaizen: a case study of Imperial Tobacco Company (ITC). *Journal of Advances in Management Research*. 2019 Jul 15. <https://doi.org/10.1108/JAMR-01-2018-0009>.
13. Omotayo TS, Kulatunga U, Bjeirmi B. Critical success factors for Kaizen implementation in the Nigerian construction industry. *International Journal of Productivity and Performance Management*. 2018 Nov 19. <https://doi.org/10.1108/IJPPM-11-2017-0296>.
14. Vo B, Kongar E, Barraza MF. Kaizen event approach: a case study in the packaging industry. *International Journal of Productivity and Performance Management*. 2019 Sep 9. <https://doi.org/10.1108/IJPPM-07-2018-0282>.
15. Kumar S, Dhingra A, Singh B. Lean-Kaizen implementation: A roadmap for identifying continuous improvement opportunities in Indian small and medium sized enterprise. *Journal of Engineering, Design and Technology*. 2018 Feb 5. <https://doi.org/10.1108/JEDT-08-2017-0083>.
16. Singh J, Singh H. Enigma of KAIZEN approach in manufacturing industry of Northern India—a case study. *International Journal of Quality & Reliability Management*. 2018 Jan 2. <https://doi.org/10.1108/IJQRM-12-2016-0220>.
17. Chan CO, Tay HL. Combining lean tools application in kaizen: a field study on the printing industry. *International Journal of Productivity and Performance Management*. 2018 Jan 8. <https://doi.org/10.1108/IJPPM-09-2016-0197>.

18. Suárez-Barraza MF, Miguel-Dávila JÁ, Jaca-García C. Kaizen: An Ancient Operational Innovation Strategy for Organizations Of The XXI Century. <https://doi.org/10.1108/TQM-06-2018-180>.
19. Solaimani S, van der Veen J, Sobek II DK, Gulyaz E, Venugopal V. On the application of Lean principles and practices to innovation management: A systematic review. The TQM Journal. 2019 Nov 29. <https://doi.org/10.1108/TQM-12-2018-0208>.
20. Kumar R. Kaizen a tool for continuous quality improvement in Indian manufacturing organization. International Journal of Mathematical, Engineering and Management Sciences. 2019;4(2):452-9. <https://dx.doi.org/10.33889/IJMEMS.2019.4.2-037>.
21. Goyal A, Agrawal R, Chokhani RK, Saha C. Waste reduction through Kaizen approach: A case study of a company in India. Waste Management & Research. 2019 Jan;37(1):102-7. <https://doi.org/10.1177/0734242X18796205>.

On fractal self-organization of the financial time series

Hilarov V.L.¹[0000-0002-9211-6144]

¹Ioffe Institute, St. Petersburg, Russia

https://doi.org/10.33847/2686-8296.4.1_6

Received 20.04.2022/Revised 29.04.2022/Accepted 22.05.2022/Published 12.06.2022

Abstract. Time series of five financial indexes daily returns were analyzed by means of multifractal and recurrence quantification analysis (RQA) methods. It is shown that a financial crisis in 2008 year is accompanied with the increase in determinism and fractal self-organization. Such regularity is noted as analogous to other nonlinear systems behavior in catastrophic situations. At the same time, the global Hurst coefficient is minimal during the crises instead of maximum for physical systems.

Keywords: nonlinear dynamical systems, multifractals, recurrence quantification analysis, catastrophes.

1. Introduction

Recently, a certain amount of data have appeared that complex systems of different physical nature behave in the same way in catastrophic situations. The multifractal singularity spectra (SS) of dynamic variables are significantly narrowed, the global Hurst coefficients increase, and the determinism of the process calculated from the time dependences of the quantitative characteristics of recurrent analysis (RQA) increases when approaching a catastrophic event. So, the aim of the research was to find out whether these patterns are present in systems of quite a different nature: financial time series of stock indices. Presence of the same peculiarities in very different systems suggests that these peculiarities are not connected with the details of the system elements interaction but rather with the character of these elements collectivization.

2. Literature Review

The multifractal singularity spectra of the surface profiles of amorphous alloys, measured by probe microscopy, are transformed during mechanical loading and narrow before the destruction of the material [1]. Narrowing of the SS was also observed during the crystallization phase transition in the NiTi alloy [2]. It is known that strong earthquakes and rock bursts are preceded by a decrease in the fractal dimension (an increase in the Hurst coefficient) [3], as well as a decrease in the width of the SS seismic time series [4]. Same effects were noted in the time series of acoustic emission during the formation of a main crack in laboratory samples [5-6]. Multifractal spectra of heart rate for the healthy people are wider, and the Hurst coefficients are lower than for the people suffering from heart disease (chronic heart failure or decreased systolic function of the left ventricle) [7]. Epileptic seizures in children are accompanied by an increase in the determinism of the process, manifested by an increase in the parameters of recurrent plots associated with diagonal lines, as well as an increase in global Hurst coefficients (an increase in persistence) [8]. It was suggested ([8]) that such behavior is associated with the fractal self-organization of systems due to the collectivization of their structural elements at different scales.

3. Data and Methodology

In this paper, we study the time dependences of the SS and the RQA characteristics in the systems of a completely different type –the time series of the S&P500, DJI, Nasdaq-100, DAX and Nikkei 225 stock indices from 1990 to 2019. In particular, it is of interest to study how the multifractal characteristics of these indices and the numerical parameters of RQA behaved during the financial crisis of 2008 and whether they have the patterns described above. It should be noted that there are many papers devoted to the study of multifractal characteristics of financial markets, but these investigations, as a rule, pursued other goals (mainly the study of market efficiency and prediction of its volatility) (see review [9]).

We studied the time series of daily closing prices of stock indices. The daily returns r_i were chosen as a random variable:

$$r_i = \frac{C_i - C_{i-1}}{C_{i-1}} = \frac{\Delta_i}{C_{i-1}}, \quad (1)$$

where C_i and C_{i-1} are the close prices of i of that day and $i-1$ of the day, Δ_i is Bachelier's absolute return. Note that with a small change in daily prices for developed markets, this value practically coincides with the logarithmic return:

$$r \log_i = \ln \frac{C_i}{C_{i-1}} = \ln \left(1 + \frac{\Delta_i}{C_{i-1}} \right) \approx \frac{\Delta_i}{C_{i-1}} \quad (2)$$

We used the p-leaders method developed in [10,11] to determine the multifractal characteristics. The p-leaders method is a generalization of the wavelet leaders method developed earlier by the same authors [12]. P-leaders at a given scale level consist of locally weighted l^p norms of wavelet coefficients calculated over a narrow range of times at all scales less than a given one [12]. The difference from wavelet leaders is the use of the l^p norm instead of l^∞ one.

The Hurst coefficient H – the value of the Hölder coefficient, at which the magnitude of the singularity spectrum reaches its maximum, and the width of this spectrum were chosen as the characteristics to be studied. It should be noted that generalized Hölder coefficients (p exponents) [10,11] are used here. The entire time series was divided into partially overlapping sections of 2048 signals, on each of which the calculation was performed. Partial overlap provides smoother time dependences of the studied quantities.

For RQA analysis, the entire financial time series was divided into partially overlapping sequences of 1024 daily time counts. For these sequences, recurrent matrices of the form were constructed:

$$R_{i,j}(\varepsilon_i, m) = \theta(\varepsilon_i - \|\mathbf{X}_i(m) - \mathbf{X}_j(m)\|), \quad (3)$$

where $\mathbf{X}_i(m)$ is the phase trajectory of a dynamical system in m -dimensional phase space, θ is the Heaviside function, and $\|\cdot\|$ is the vector norm. The recurrent matrix constructed in the form of a plot is a two-dimensional set of points corresponding to relation (3) and consists of black dots for the close (recurrent) points with the values of ones and white dots for the distinct points with the value of zero. The proximity of point i neighbors is determined by the parameter ε_i . The choice of large values of m

– the dimension of the embedding space is necessary in the case of restoring the attractor based on the Takens theorem for scalar series. Since this procedure is not used in this paper, the value was chosen to be $m=1$. The choice of the threshold value \mathcal{E}_i significantly affects the results; therefore, it is discussed many times, see the review [13]. If one chooses \mathcal{E}_i too large, then the points will mostly close to each other (recurrent), and the graph will be a black field; while if \mathcal{E}_i is too small, then there will be almost no recurrent points. In this work, the value was empirically chosen to be 0.1 for all points. The patterns obtained were observed in a wide range of threshold values. The free CRP Toolbox software package [13] has been used.

The aim of RQA analysis was to calculate the parameter called determinism and to find its time dependence. In order to define this parameter one has first to introduce the recurrence rate parameter RR:

$$RR(\mathcal{E}_i) = \frac{1}{N^2} \sum_{i,j=1}^N R_{ij}(\mathcal{E}_i, m), \quad (4)$$

representing the recurrent points density on the plot. The determinism value is introduced as the ratio of the number of recurrent points forming diagonal lines parallel to the main diagonal to the total number of recurrent points:

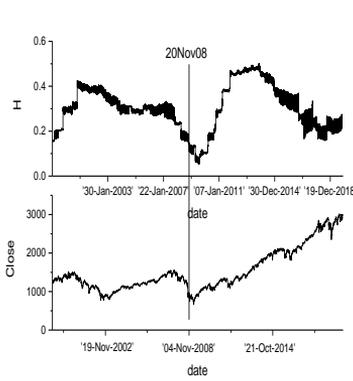
$$DET = \frac{\sum_{i,j=1}^N D_{ij}}{\sum_{i,j=1}^N R_{ij}} \quad (5)$$

Chaotic, weakly correlated processes either do not have diagonal lines at all, or these lines are very short, while deterministic processes have longer diagonals and fewer isolated recurrent points [13].

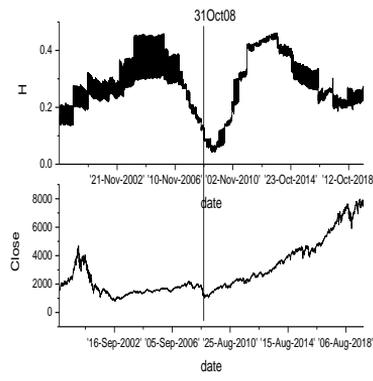
4. Results

Results on the Hurst coefficient calculation are shown in figure 1a. All dependencies show a clear minimum that accompanies the financial crisis of 2008. Figure 1b shows the calculated time dependences of the singularity spectrum widths w . This data was obtained for the case of $p=1$. It can be seen that the financial crisis is preceded by a decrease in w thus demonstrating the tendency to monofractalization. For the Nasdaq-100 index this trend is less pronounced. Perhaps this is because the Nasdaq-100 index does not include the financial companies most exposed to stock changes. The close prices, shown in figure 1 were obtained from [14].

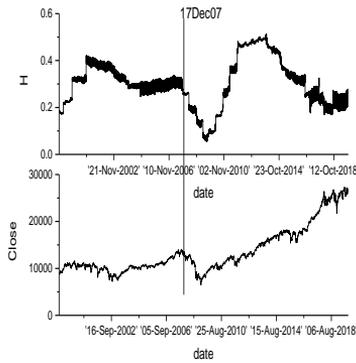
Note that a decrease in the spectrum width and Hurst coefficient is also observed for the latest data processed. This suggests that they reflect the beginning for the new ongoing financial crisis. Multifractal methods require time series of considerable length so that they can be analyzed over a wide range of scale levels. Significantly shorter series lengths are required for Recurrent Plots (RP) and Quantitative Recurrent Analysis (RQA).



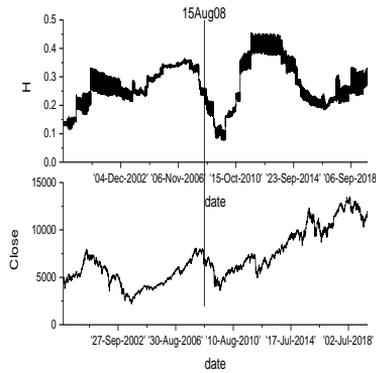
S&P500



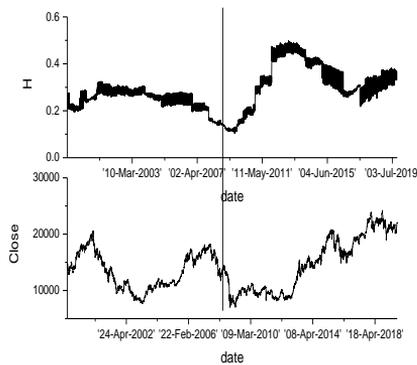
Nasdaq



DJI

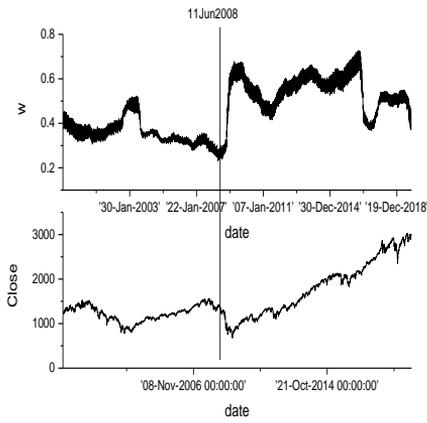


DAX

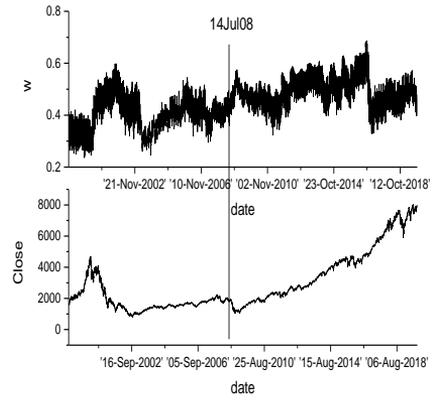


Nikkei225

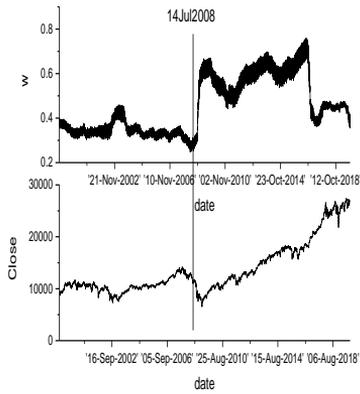
Fig. 1a. Calculated time dependencies of the Hurst coefficients of stock indices and close prices of these indices obtained from [14].



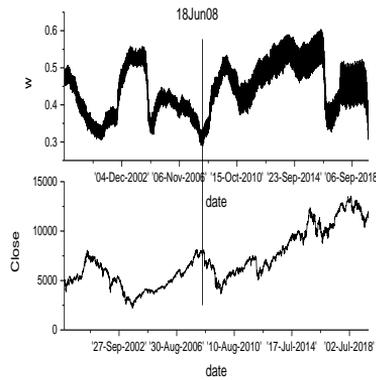
S&P500



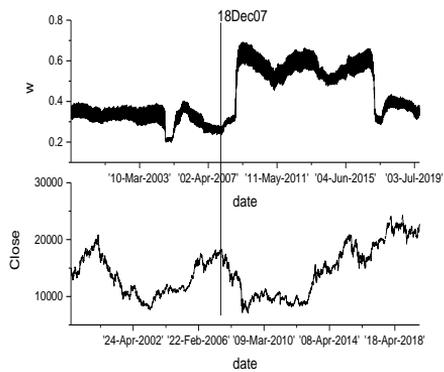
Nasdaq



DJI



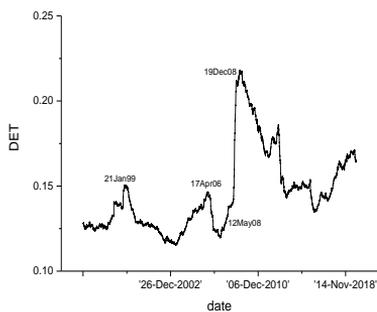
DAX



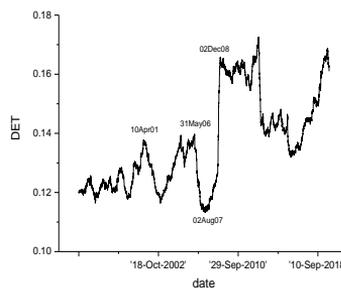
Nikkei225

Fig. 1b. Calculated time dependencies of the SS widths of stock indices and close prices of these indices obtained from [14].

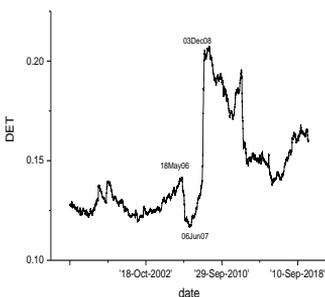
Fig. 2 shows the time dependencies of the determinism value (DET) for the considered exchange indices calculated using equation (5). It can be noted that the financial crisis of 2008 is accompanied by an increase in determinism. There is also an increase in determinism, starting from 2017 until the end of the studied time range, which, along with the results of multifractal analysis, indicates the possible preparation of a new financial crisis.



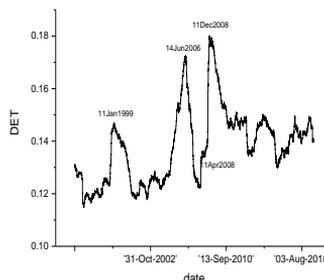
S&P500



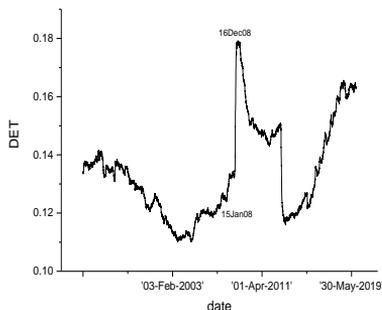
Nasdaq



DJI



DAX



Nikkei

Fig. 2. Calculated time dependencies of the DET parameter for stock indices.

5. Conclusion

We performed multifractal and recurrence quantification analysis of financial time series in order to investigate fractal self-organization. We can conclude that financial catastrophes, based on the changes in the parameters of daily returns over time, are similar in some respects to natural ones: they are accompanied by a decrease in the width of the singularity spectra and an increase in determinism, demonstrating the tendency to monofractalization. At the same time, in contrast to natural disasters, the Hurst coefficient in financial catastrophes has an antipersistent minimum and turns out to be significantly less than its characteristic value for an efficient market (0.5). This is a further confirmation of the fact that the observed fractal self-organization in the systems of quite different nature is not determined by the details of the system elements interaction but rather with the character of these elements collectivization. Besides, the obtained results can be used for the purpose of predicting financial crises

The presence of other extremes on the charts requires further economic justification.

References

1. Hilarov, V.L., Korsukov, V.E., Butenko, P.N., Svetlov, V.N.: Wavelet transform as a method for studying the fractal properties of the surface of amorphous metals under mechanical load. *Phys. Solid State*, **46**, 1868–1872 (2004)
2. Butenko P.N., Gilyarov V.L., Korsukov V.E., Korsukova M.M., Obidov B.A.: The Effect of Isochronous Annealing on the Surface Characteristics of Ni50Ti50 Metal Glass Ribbons. *Technical Physics*, **65**, 205–210, (2020).
3. Gibowicz S.J., Lasocki S.: Seismicity induced by mining: Ten years later In: Dmowska R., Saltzman B. *Advances in Geophysics*. vol. 44. Elsevier, pp. 39-181, (2001).
4. Kasimova V.A., Kopylova G.N., Lyubushin A.A.: Variations in the Parameters of Background Seismic Noise during the Preparation Stages of Strong Earthquakes in the Kamchatka Region. *Izvestiya. Physics of the Solid Earth*, **54**, 269–283, (2018).
5. Hilarov V.L.: Detection of the Deterministic Component in Acoustic Emission Signals from Mechanically Loaded Rock Samples. *Phys. Solid State*, **57**, 2204–2211, (2015).
6. Damaskinskaya E.E., Hilarov V.L., Pantelev I.A., Gafurova, D.R., Frolov, D.I.: Statistical Regularities of Formation of a Main Crack in a Structurally Inhomogeneous Material under Various Deformation Conditions. *Physics of the solid state*, **60**, 1821–1826, (2018).
7. Ivanov P. Ch., Amaral N L. A., Goldberger A. L., Havlin S., Rosenblum M. G., Struzikk Z R., Stanley H. E.: Multifractality in human heartbeat dynamics. *Nature*, **399**, 461–465, (1999)
8. Hilarov V.L.: Epileptic seizures regularities, revealed from encephalograms time series by nonlinear mechanics methods. *J. Phys.: Conf. Ser.*, 1400 033011, (2019).
9. Jiang Z.-Q., Xie W.-J., Zhou W.-X., Sornette D.: Multifractal analysis of financial markets: a review. *Rep. Progr. Phys.*, **82**, 125901, (2019).
10. Jaffard S., Melot C., Leonarduzzi R., Wendt H., Abry P., Roux S.G., Torres M.E.: P-exponent and p-leaders, Part I: Negative pointwise regularity. *Physica* **A448**, 300–318, (2016).
11. Leonarduzzi R., Wendt H., Abry P., Jaffard S., Melot C., Roux S.G., Torres M.E.: P-exponent and p-leaders, Part II: Multifractal analysis. Relations to detrended fluctuation analysis. *Physica*, **A448**, 319–339, (2016).
12. Wendt H., Roux S.G., Jaffard S., Abry P.: Wavelet leaders and bootstrap for multifractal analysis of images. *Signal Process.* **89**, 1100–1114, (2009).
13. Marwan N., Romano M.C., Thiel M., Kurths J. (2007). Recurrence plots for the analysis of complex systems. *//Phys. Reports*. V. 438. P. 237–329.
14. <https://finance.yahoo.com/world-indices/>

DETECTION OF NEOPLASMS IN THE PROCESS OF CLINICAL EXAMINATION OF THE ADULT POPULATION OF RUSSIA

Olga Zakharchenko ¹[0000-0002-6234-2992],
Dina Terenteva ¹[0000-0003-1669-939X], Irina Shikina ¹[0000-0003-1744-9528]

¹Federal Research Institute for Health Organization and Informatics of the
Ministry of Health of Russia, Moscow, Russia

https://doi.org/10.33847/2686-8296.4.1_7

Received 11.03.2022/Revised 25.04.2022/Accepted 22.05.2022/Published 12.06.2022

Abstract. The article analyzes the detection of neoplasms, including malignant ones, in the process of clinical examination of the adult population of Russia. It is shown that during the period of 2013-2020, 312206 cases of neoplasms, including 183436 malignant ones, were detected during the clinical examination. There is a tendency to an increase in the detection of neoplasms, including malignant ones, the average annual growth rate is 6.6%. Even taking into account the difficult epidemiological situation in 2020 caused by the new coronavirus infection COVID-19, which led to the termination for some time of the clinical examination, the detection of neoplasms, including malignant ones, continued to grow. Thus, detection of oncological diseases is increasing, which indicates the effectiveness of large-scale preventive measures, in particular, the clinical examination.

Keywords: active detection, medical examination of the adult population; screening; neoplasms; preventive measures.

1. Introduction

In 2013, the World Health Organization published the Plan of Action for the Prevention and Control of Noncommunicable Diseases for 2013-2020, one of the global goals of which was to reduce premature mortality from cardiovascular diseases, cancer, chronic respiratory diseases and diabetes by 25% [1]. In Russia in 2013 a large-scale preventive program was launched to detect chronic non-infectious diseases at an early stage of their development — the clinical examination of certain groups of the adult population, hereinafter referred to as a 'clinical examination' (Order of the Ministry of Health of the Russian Federation No. 1006n "On Approval of the Procedure for Clinical Examination of Certain Groups of the Adult Population" of 03.12.2012).

Special attention is paid to the detection of circulatory system diseases, diabetes mellitus, chronic lung diseases and malignant neoplasms. According to previous studies, malignant neoplasms occupy one of the leading positions among causes of morbidity, disability and mortality, which cause significant damage to labor, production and financial resources both in the world and in Russia [2, 3]. In this context, cancer control is an important part of all existing healthcare systems in the world: public, private and mixed. In Russia, where the public health system prevails, the fight against cancer is carried out within the framework of the national project 'Healthcare' of which the federal project 'Combating malignant neoplasms' is a part (Passport of the national project 'Healthcare' (approved by the Presidium of the Presidential Council for Strategic Development and National Projects, protocol of 24.12.2018 N 16; Federal project 'Combating malignant neoplasms'; Presidential Decree of 07.05.2018 № 204 'On national goals and strategic objectives of the Russian Federation for the period until 2024') [4]. One of the target indicators of the federal project 'Combating malignant neoplasms' is the share of malignant neoplasms

detected at the early stages (stages 1-2). The leading direction for detecting malignant neoplasms at early stages, along with preventive examinations and population screening programs, is the dispensary. Early diagnosis of malignant neoplasms gives an opportunity to start treatment in time and reduce both human losses and costs for health care system [5, 6].

The specific feature of the clinical examination developed in Russia is the two-stage principle of its conduct. At the first screening stage, signs of chronic non-infectious diseases and risk factors for their development are identified. Medical indications are determined for additional examinations by specialists in order to clarify the diagnosis. At the second stage of the clinical examination, additional examinations are carried out and the diagnosis of the disease (condition) is clarified. Currently, in the first part of the two-stage clinical examination, cancer screening is performed for the seven most common localizations of cancers. The second part includes additional examinations to clarify the diagnosis of the disease (condition) if there are medical indications and in accordance with the clinical recommendations of specialists (Order of the Russian Ministry of Health No. 404n of April 27, 2021 'On approval of the procedure for preventive medical examinations and clinical examination of certain groups of citizens'). Our study shows the state of active detection of oncological diseases in the early stages of development in the process of clinical examination of the adult population in the Russian Federation.

Our study contributes to the identification of organizational problems faced by a physician of a medical organization providing medical care in outpatient settings when providing preventive medical services. Let us note that clinical examination is a typical preventive measure in Russia, covering more than 50% of the adult population. The study of the dynamics of the detection of cancer diseases in the process of clinical examination for eight years will make it possible to determine the ways of further development of clinical examination of the adult population. The experience of Russia presented in this study may be extrapolated to other regions of the world, especially to those countries where there is a high degree of state participation in the health services provision.

2. Literature Review

In accordance with the normative documents regulating the conduct of clinical examinations, the main criterion of the effectiveness of clinical examinations is the coverage by health examinations of the population who are covered by medical care in a medical organization and are subject to a scheduled medical examination in the current year [7].

According to our earlier analysis, in 2013-2019 one fifth of the adult population of Russia underwent annual clinical examinations. The population older than working age (over 60) was more involved [8].

The analysis of the form №131/o has shown that the coverage of the population by screening increases every year, and its contribution to the detection of cancerous diseases also grows. A statistically significant correlation of strong and moderate strength between the coverage of the population by clinical examination and the primary incidence of malignant diseases has been revealed [9]. However, despite rather high coverage of the population by clinical examinations, it is necessary not only to continue health education work with the population, but also to find new ways to attract citizens to undergo medical examinations [10].

During studying and analyzing the data of form №131/o, it became clear that in the structure of diseases annually detected during the clinical examination, neoplasms occupy the 10th rank out of 11. The low detection rate of oncological diseases during

clinical examination, compared to other chronic non-infectious diseases, is caused by a number of factors, including:

- insufficient cancer awareness among doctors at medical institutions providing primary health care, and, as a result, untimely registration of citizens for follow-up care;
- under-fulfillment of the volume of examinations provided for by the procedures for conducting clinical examination, including its oncological component;
- inadequate organizational and methodological guidance in terms of controlling the routing of patients with suspected or already diagnosed cancer at the level of the subject of Russia [11,12];
- different level of material, technical, and staffing support of medical organizations in Russian regions, what determines the availability and quality of medical care in the 'oncology' profile [3,4,12,13].

In order to eliminate the negative factors outlined above, it is necessary to actively study and improve approaches to organizational technologies for conducting clinical examinations.

A good example of a quick response to the new requirements for the health care system caused by the COVID-19 outbreak is the introduction of changes in the procedure for conducting clinical examinations for citizens who have had a new coronavirus infection. From 01.07.2021, the above-mentioned citizens undergo an in-depth medical examination.

3. Data and Methodology

Based on the data from the official statistics of the sectoral statistical reporting form No.131/o 'Information about the clinical examination of certain groups of the adult population' and 'Information about the preventive medical examination and clinical examination of certain groups of the adult population' (form No.131/o) the dynamics of detection of neoplasms, including malignant ones, during clinical examination in the period from 2013 to 2020 in Russia were assessed. Clinical examinations in Russia are conducted according to the district (territorial) principle and serve the population assigned to the medical organization that provides primary health care in outpatient settings. In the present work we relied on information provided directly by primary health care medical organizations, the activity of which includes the identification and elimination of risk factors, including preventive measures for early detection of malignant tumors. Generalization, system and functional approaches as well as methods of descriptive statistics were used in the study. The indicators of dynamic rows, intensive and extensive indicators were calculated.

4. Results

At the first stage of the study, we studied the coverage of clinical examination of the adult population of the Russian Federation in 2013-2020. It should be noted that in 2020, due to the complicated epidemiological situation caused by the new coronavirus infection COVID-19 [14,15], the clinical examinations were suspended, and the report includes data only for the first four months.

The analysis of the data of form №131/o for 2013-2020 showed that, according to medical organizations of Russian regions, the share of persons who underwent stage I medical examinations varied from 77.1% (2015) to 97.4% (2019) of those

subject to the plan. The proportion of persons who underwent clinical examination in 2020 was 44.7% of those subject to the plan (Fig. 1).

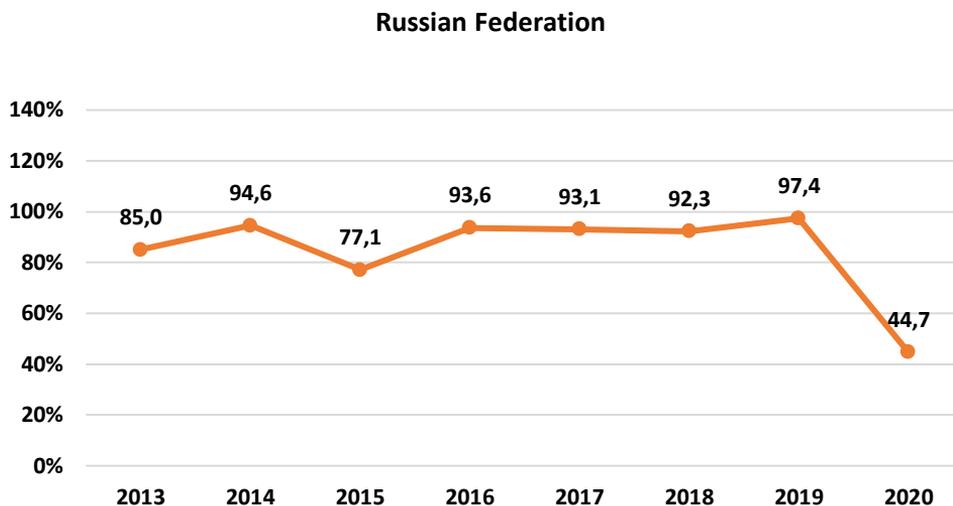


Fig. 1. Percentage of persons who underwent Stage I clinical examination of those subject to the plan for the period 2013-2020 in the Russian Federation (%).

Source: Own data analysis based on industry statistical reporting forms №131/o.

The decrease in the share of persons who underwent medical examinations in 2015 was due to changes in the procedures for undergoing medical examinations, including reporting and accounting documents.

Taking into account the fact that in 2020 the clinical examination was carried out for only four months, we calculated the average indicators of the series of dynamics of the coverage of the Russian adult population with and without 2020. Over the period 2013-2019, an average of 21.481 million people underwent annual clinical examination for the entire analyzed period, the growth of the coverage rate was 1.0464, that is, on average, the coverage each year increased by 4.6% (1.013 million people). Taking into account the difficult epidemiological situation of 2020, the average indicators of the dynamics of the series show not anticipated growth but decrease. Thus, over the entire analyzed period (2013-2019) the decrease in the index of coverage of the adult population by medical examinations was 0.9082, i.e. the average coverage decreased every year by 9.2% (1.360 million people).

In the next stage of the study, we've analyzed the detection of neoplasms, including malignant neoplasms, in the course of clinical examinations in 2013-2020.

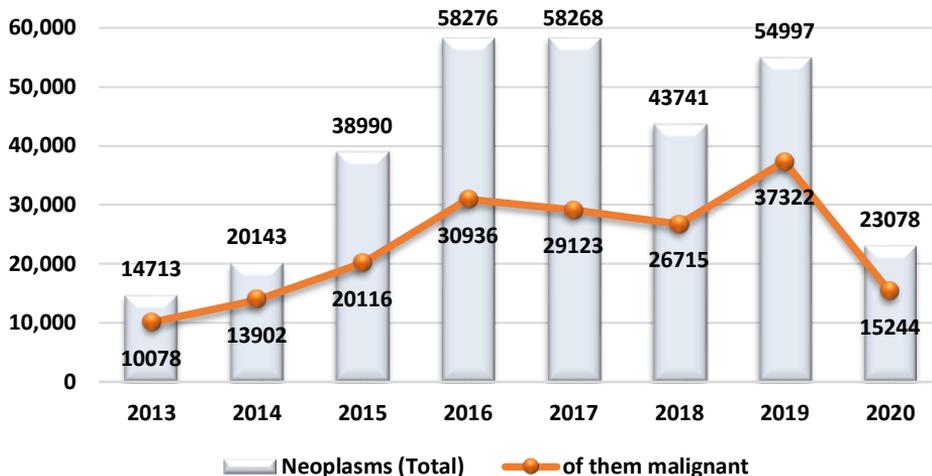


Fig. 2. The number of persons with newly detected neoplasms, including malignant neoplasms, during clinical examination of the adult population for the period 2013-2020 in the Russian Federation (abs. numb.).

Source: Own data analysis based on statistical reporting forms №131/o.

Fig. 2 shows the number of first-time detected neoplasms, including malignant ones. During the analyzed period, 312206 cases of oncological diseases were detected, of which 183436 (58.8%) were malignant. The largest number of cancer diseases was detected in 2016, 2017 and 2019, which may be due to changes in the frequency of medical examinations for certain categories of citizens, the volume and multiplicity of examinations and tests. The smallest number of oncological diseases was detected in 2013 and 2014, at the initial stage of implementation of the dispensary project, as well as in 2020, during the COVID-19 pandemic. Nevertheless, the share of registered malignant neoplasms from all oncological diseases fluctuates in the range of 50-69%.

The average values of the detected neoplasms, including malignant ones, during the clinical examination also depend on the consideration in the analysis of 2020, when the clinical examination was carried out for only four months. Despite the fact that in 2020 the number of oncological diseases detected for the first time during the clinical examination was expected to decrease, the dynamics of the index remained positive (Table 1).

Table 1. Average rates of detection of neoplasms, including malignant ones, during clinical examination of adult population in Russia with and without taking into account the data of 2020

	Average value	Average growth rate, %	Average growth rate, %	Average absolute gain, abs.
2013-2019	42378,83	1,25	24,6	6714
2013-2020	41901,5	1,07	6,6	1195

Source: Own data analysis based on statistical reporting forms №131/o.

5. Conclusion

Thus, the goal of our study, which was to study the dynamics of cancer detection in the process of clinical examination of the adult population, was achieved by using the methods of descriptive statistics, generalization and compilation of dynamic rows and mean values. Based on our study we have made the following conclusions:

1. There is a positive dynamic of coverage of the adult population by medical examinations during 2013-2019.

2. The growth of oncological diseases detection rate in the study period, including in 2020 on the background of a difficult epidemiological situation caused by a new coronavirus infection was recorded.

3. The share of detected malignant neoplasms out of all neoplasms is 58.8%.

4. The results of 2020 are comparable with the data of the first years of the clinical examination (2013-2014) both in terms of coverage of the adult population by dispensary and detection of oncological diseases.

5. The results of our study can be useful for both health organizers and clinicians, as well as for epidemiologists, because part of the study was conducted in a difficult epidemiological situation caused by a new coronavirus infection.

6. For a more encompassing analysis of cancer detection in the process of clinical examination, it is necessary to increase the level of continuity between medical organizations providing care in outpatient settings and specialized medical organizations.

Acknowledgments

The study has no sponsorship. The authors declare that there is no conflict of interest.

References

1. Action Plan against Non-Communicable Diseases for 2013-2020, https://apps.who.int/iris/bitstream/handle/10665/94384/9789244506233_rus.pdf?sequence=5 (Circulation Date: 10.02.2022)
2. Kaprin A.D., Starinsky V.V., Shakhzadova A.O. State of oncological assistance to the population of Russia in 2019. M.: MNIOI named after P.A. Herzen - a branch of the Federal State Budgetary Institution "NMITS of Radiology" of the Ministry of Health of Russia, 252 p., (2020). (In Russian).
3. Antipova T., Shikina I. Informatic indicators of efficacy cancer treatment. 12th Iberian Conference on Information Systems and Technologies (CISTI), Lisbon, Portugal, (2017) pp. 1–5, <https://doi.org/10.23919/CISTI.2017.7976049>
4. Antipova T.V., Melnik M.V., Nechaeva O.B., Shikina I.B., Vechorko V.I., Lutseva E.M. Social aspects of public health, 1(47), (2016). <https://doi.org/10.21045/2071-5021-2016-47-1-3> [Electronic Resource] URL: <http://vestnik.mednet.ru/content/view/730/30/lang, ru/>, last accessed 2021/09/10 (In Russian).
5. Efremov S.A., Petkau V.V., Gruzdeva E.A. Implementation of a system to support the work of the oncological service at the regional level. Health manager, 2, pp. 25-30, (2021). <https://doi.org/10.21045/1811-0185-2021-2-25-30> (In Russian).
6. Voskanyan Y., Shikina I., Andreeva O., Kidalov F., Davidov D. Multifactorial model of adverse events and medical safety management system. Journal of Digital Science, vol 2(1), pp. 29-39, (2020). https://doi.org/10.33847/2686-8296.2.1_3
7. Chernobrovkina A.E. Contribution of the medical examination of the adult population to the early detection of oncological diseases of the female genital organs. Preventive medicine, 25(1), pp. 7-13, (2022). <https://doi.org/10.17116/profmed2022250117> (In Russian).
8. Zakharchenko O.O., Terentyeva D.S. Sexual age characteristics of the coverage of the adult population of the Russian Federation in 2013-2019. Social aspects of population health, 3 (67), (2021). <https://doi.org/10.21045/2071-5021-2021-67-3-1> [Electronic Resource] URL: <http://vestnik.mednet.ru/content/view/1263/30/lang, ru> (In Russian).

9. Bryzgalova O.E., Armashevskaya O.V. Online resource as a tool for preventive measures in providing medical care to women during the Covid-19 pandemic. *Obstetrics and gynecology*, (1), pp. 113-121, (2022). <https://doi.org/10.18565/aig.2022.1.113-121>. (In Russian).
10. Zakharchenko O.O., Terentyeva D.S., Surayeva N.A., Komarov Yu.I. Detection of malignant neoplasms in the process of medical examination of the adult population in the Northwestern Federal District in 2016-2019. *Social aspects of public health* [online publication], 6 (67), 5 p., (2021). <https://doi.org/10.21045/2071-5021-2021-67-6-5> [Electronic Resource] URL: <http://vestnik.mednet.ru/content/view/1319/27/lang,ru/> (In Russian).
11. Kaprin A.D., Alexandrova L.M., Starinsky V.V., Kalinina A.M., Ipatov P.V., Fighters S.A., Drapkina O. M. Dispensation of certain groups of the adult population of Russia as an instrument for early detection of malignancies (results of 2015-2016). *Preventive medicine*, 21(4), pp. 13-19 (2018). <https://doi.org/10.17116/profmed201821413> (In Russian).
12. Khodakova O.V., Koshevaya N.V. Self-esteem of health as an element of self-preservation behavior and commitment to medical examination of adult population. *Current health and health statistics issues*,(4), pp.309-326, (2019). <https://doi.org/10.24411/2312-2935-2019-10099>. URL: <https://healthproblem.ru/ru/magazines?text=317> (In Russian).
13. Nechaeva O.B., Shikina I.B., Chukhrienko I.Yu. and others. Resource support for medical organizations providing assistance in the profile of "oncology." *Current health and health statistics issues*, (2), pp. 269-279, (2019). <https://doi.org/10.24411/2312-2935-2019-10042> [Electronic Resource] URL: <https://healthproblem.ru/ru/magazines?text=259> (In Russian).
14. Zakharycheva T., Makhovskaya T., Shirokova A., Shikina I. (2022) Autonomic Dysregulation Syndrome in Covid-19 Convalescents: Possible Causes and Approaches to Its Correction. In: Antipova T. (eds) *Comprehensible Science. ICCS 2021. Lecture Notes in Networks and Systems*, vol 315. Springer, Cham. https://doi.org/10.1007/978-3-030-85799-8_34
15. Zakharycheva T., Makhovskaya T., Shirokova A., Shikina I. (2021) In: Antipova T. (eds) *2021 International Conference on Advances in Digital Science (ICADS 2021)*, AISC 1352, pp.191-197, (2021). https://doi.org/10.1007/978-3-030-71782-7_17

Briefs in Assessing the Adequacy of Health Care Facilities' Fixed Assets

Tatiana Antipova¹[0000-0002-0872-4965], Alexander Zhelnin²[0000-0003-1436-6145], Iuliia Zhelnina²[0000-0002-9586-1465]

¹ Federal Research Institute for Health Organization and Informatics of Ministry of Health of the Russian Federation, Moscow, Russia

² Perm Krai Clinical Hospital, Perm, Russia

https://doi.org/10.33847/2686-8296.4.1_8

Received 01.03.2022/Revised 13.04.2022/Accepted 22.05.2022/Published 12.06.2022

Abstract. This work considers some procedures to assess the adequacy of fixed assets for Health Care facilities activity. In terms of effectiveness, equitable distribution examines whether limited resources are directed toward improving the health of the population in the delivery of health care services. Assessing the adequacy of health care resources examines the extent to which resources are used to provide health system outcomes and/or achieve health system goals. The results of such an assessment: on the one hand can prevent waste of limited health care resources, and another hand to increase effectiveness of health care services.

Keywords: health care, facilities, nonfinancial assets, assessing, adequacy, effectiveness, efficacy, evaluation.

1. Introduction

Health Care facilities are an important part of the healthcare system. They are critical infrastructures that provide essential services to the social functioning of a community [1]. Health Care facilities have to ensure the highest standards by delivering routine health service [2], and they must continue to operate during disasters [3]. Indeed, communities are at risk of disasters due to hazardous events including biological, environmental, geological or geophysical, hydrometeorological, societal and technological hazards [4]. Therefore, the properly functioning of health care, i.e., hospital facilities and the reduction of consequences due to disasters are vital for local, national, and global health security [5]. This reduces the cascading effects and helps build the resilience of communities, countries, and health systems [6]; this has been particularly evidenced by the COVID-19 pandemic [7] since the COVID-19 pandemic is a timely reminder of how hazards within the complex and changing global risk landscape can affect lives, livelihoods and health [4].

In this context, the World Health Organization (WHO) states that 'measures to ensure the safety, security and functionality of health infrastructure are needed at both national and community levels. Countries and communities need to prioritize the protection of new and existing hospitals and other health facilities from identified hazards and should ensure the physical integrity of buildings, equipment and critical hospital systems [3]. More and more attention has been paid to creating comfortable and functional hospital environments, where the patients can feel good and at ease maintaining the same efficiency of medical activities. In addition, flexibility has been particularly considered because hospital facilities must adapt to the continuous needs of change. [8]

In some cases, higher costs of equipment are justified by patients requiring more complex care. In other cases, differences in costs are caused by differences in

medical practice. A difference could be caused, for example, by the use of different equipment or other tools, but also by different surgical techniques. Such differences, once visible, form the basis of discussions among clinicians aimed at confirming a shared understanding of when variation is appropriate, and align diverse clinical practice where evidence shows a clinically determined balance of health outcomes and cost. Discussions around comparisons of costs between the service lines of different facilities can also reveal differences in resource/assets consumption for indirect costs. If activity-based costing is in place, decision-makers can then spot more easily the origin of such differences and potentially make services more efficient. While activity-based costing has been shown to reduce overall costs, its main advantage may actually lie in using the existing resources/assets more effectively. This is particularly important in a context of coping with rising demand without increasing the available resources.[15]

The Care Pathway Simulator (CPS) allows the design and comparison of different configurations of services to assess benefits and resource needs. Users can specify different care pathway models to determine parameters of interest including resource usage and patterns of care for each scenario. The simulation, which is based on discrete event modelling, allows predictions of capacity constraints in restructuring care pathways and mapping of performance to resource needs. The CPS has been used to redesign an outpatient clinic for vascular surgery in Good Hope Hospital, United Kingdom. In this application, the model used three inputs: patient lists to represent demand at the clinic; the care pathway or sequence; and the resources required to carry out the necessary care. The model was then applied to predict performance under different clinic process designs.[15]

But assessment methodologies of the adequacy of Health Care Facilities' Fixed Assets with specific focus on performance evaluation are limited. While primary care remains the principal point of patient contact with the delivery system for most citizens, it is the hospital sector that, because of the scale of the facilities and their resource requirements [14], has elicited the bulk of academic and official focus on the adequacy of these resource (fixed assets). This work considers the Adequacy of Health Care Facilities' Fixed Assets. For ease of perception of this study, fixed assets will refer to as resources.

2. Data and Methodology

The data was obtained from financial and statistical reports of health care facilities for time period 2005 – 2022 with once-a-year frequency of the data observations and calculation. From the organizational point of view, the authors have organized and participated multidisciplinary and thematic meetings (management, engineering, technical, logistic, health area, artificial intelligence). Experts and stakeholders were involved in the working groups with different purposes according to the development phase of the methodology. All key points of these measures as a part of future methodology (strategic goals, reference standards, algorithms and outcomes) were discussed in periodic meetings during nine International Conferences 2017-2021.

3. Results

The pursuit of efficiency is one of the central preoccupations of health policymakers and managers, and it is justifiably a cause for such concern. Most immediately, inefficient care can lead to unnecessarily poor outcomes for the patients directly affected, measured either in terms of their health improvement, or in their broader satisfaction with the health system. More generally, inefficiency somewhere in the health system is likely to deny treatments and health improvement to patients

who would otherwise have received treatment if resources had been better used, especially in systems operating with a fixed global budget. [15].

To assess the adequacy of a medical facilities' resources, we must understand what those resources are and how and what they are used for. Some of these resources are illustrated in Fig. 1.

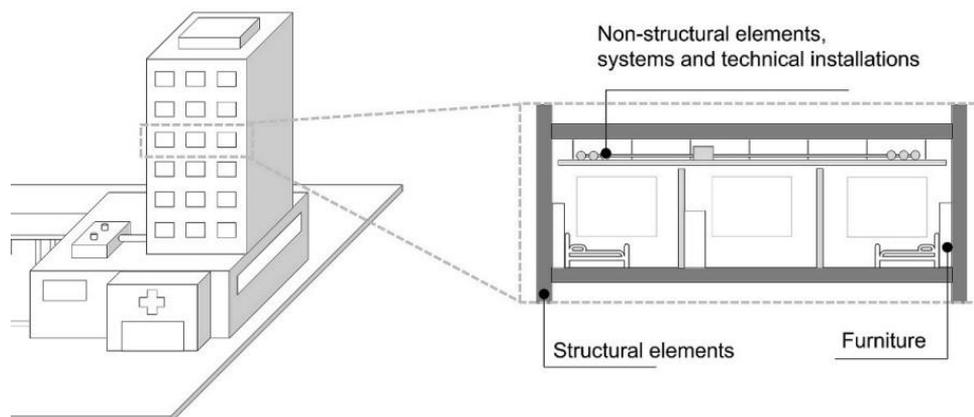


Fig.1. Example of health care facility's fixed assets.

Source: authors' interpretation of Fig.3 [8].

Fig.1 shows that all of highlighted elements (Non-structural elements, systems and technical installations; Structural elements (part of building), and Furniture) include into fixed assets in any Russian hospital according to Russian legislative. So, we can see that facilities' fixed assets include:

- Dwellings
- Buildings and Constructions
- Investment property
- Machinery and equipment
- Transport vehicles
- Inventories
- Biological resources
- Other fixed assets.

More details of fixed assets considered in previous work [10].

Before we can use resource performance measures, we need to make sure that resources are available, secure, and sufficient. The only way to verify the actual availability of the resources accounted for is through an inventory. If it is not possible to conduct an inventory at the current moment, the results of a recent inventory (e.g. at the end of the previous year) can be used.

When securing resources, a strong internal control system is necessary to ensure rational access to resources, accountability for non-compliance.

Access to resources and reports should be limited to a list of persons authorized and responsible for their preservation and use. It can avoid inefficient use of resources.

Inefficient use of health system resources poses serious concerns, for a number of reasons:

- it may deny health gain to patients who have received treatment because they do not receive the best possible care available within the health system's resource limits;

- by consuming excess resources, inefficient treatment may deny treatment to other patients who could have benefited from treatment if the resources had been better used;

- inefficient use of resources in the health sector may sacrifice loss of consumption opportunities elsewhere in the economy, such as education or nutrition;

- wasting resources on inefficient care may reduce society's willingness to contribute to the funding of health services, thereby harming social solidarity, health system performance and social welfare.[15]

To avoid above, mandatory periodic inventory of resources must be organized and conducted. Accountability must be provided to reduce risk due to possible errors, fraud, or misuse.

Let us consider basic assessing procedures for resource preservation.

1. Verification of the completeness and accuracy of accounting facts. At this stage it is necessary to make sure that the measurement (or calculation) of quantitative indicators is made with a sufficient degree of accuracy when accounting for resources. It is also necessary to make sure that all resources without exception are measured or counted. The procedure is performed using the methods of observation and counting. If necessary, you can use the method of interviewing personnel. The researcher should determine the methods, techniques and technical means used in the institution to measure resources, assess the condition of the weighing facilities and measuring instruments, and check how often the instruments should be checked and when they were last checked. Find out whether facilities' warehouse employees are satisfied with the state of the hand-held instruments and whether, in their opinion, these instruments ensure completeness and accuracy in measuring and counting operational facts. It is helpful to review the job descriptions of facilities' warehouse employees and the liability agreements they have in place. It is advisable to conduct a survey of facilities' warehouse employees and resource recipients.

The results of this survey will assist the researcher in evaluating the use of resources and the need for recommendations to optimize this process.

In addition, restricting access to the warehouse is important. Without access restrictions, an organization's warehouse is like a large store with no prices-just take whatever you want. This by no means that employees are taking goods for personal use, but they may be taking excessive quantities of them for medical purposes, leading to piling up in the warehouse area. It also puts managers in the difficult and almost impossible task of trying to determine what is in stock and what needs to be purchased for today's medical needs. Consequently, there must be mandatory control of inventory, as well as tightly restricted access to it.

Only a very narrow circle of well-trained employees should have access to documents containing raw material and inventory records. This will minimize the risk of accidental or intentional changes to these critical records. The security system should also retain the user passwords of everyone who has had access to these records, in case evidence is needed to prove that fraudulent activity has taken place.

If the researcher has doubts about the efficiency of the resources and there is a possibility of failures and irregularities, then an additional procedure should be carried out to identify the possibility of errors and irregularities.

2. Verify the promptness of recording resources on the media. It is possible to obtain such information (with an appropriate degree of reliability) using digital technology. As a rule, several users have access to records in the medical facilities.

3. Verification of the reliability of primary accounting records that formalize resources. The purpose of this procedure is to obtain sufficient evidence that the entire body of primary resource records is legally valid. The following sources of information are used during the procedure:

claims - bills of lading or limit - intake cards;

material accounting cards;

orders on approval of persons entitled to sign primary documents.

Particular attention should be paid to documents in which the following requisites are missing or corrected:

name of the structural subdivision - the recipient;

name of the type of activity (type of work) for which resources are requested;

the number of resources requested and especially the number of resources released;

personal signatures of the participants of the operation.

Summarizing the results of all analytical procedures described above, the researcher should group all detected errors and violations, give recommendations for their elimination, if possible, as well as indicate the main ways to optimize the internal control system.

When the availability, integrity and adequacy of resources are assured, it is possible to move on to the assessment of resource efficiency.

The concept of achieving resource efficiency and its application is carried out in various organizational contexts:

- * Introduction of managerial responsibility for action;

- * Implementation of a results-oriented strategy;

- * Development and approval of the standard of accountability (individually and departmentally) for top managers.

Assessment of compliance of the goals of providing budget funds with the goals of their use.

Since medical facilities are participants of the budgetary process and receive funds directly from the budget, a great role is played by compliance with the objectives for which these funds were allocated. The goals of allocating budgetary funds to medical facilities are determined by the state assignment for providing a certain type of services [9].

Therefore, first of all, it is necessary to analyze the compliance of the types of activities of a medical institution with the state assignment. The analysis can be conducted in two directions: by volume and by content.

Analysis by volume involves assessing whether the amount of expenses is equal to the number of subsidies or subventions provided.

Some additional negative events or transactions that might be anticipated or projected include the following:

- Health care facilities may misrepresent their financial or other circumstances to change the qualifications for grants or subsidies.

- health care facilities may miscalculate the number of required budgets in order to receive more budgeted funds than necessary.

Assessment by content includes an evaluation of the purpose of spending budget funds.

After the calculation and analytical activities, it is necessary to implement the results of the performance evaluation. In this case, there are two options depending on the nature and obligation of performance evaluation:

- 1) if the performance evaluation is carried out at the initiative of the state, not satisfied with the activities of the use and disposal of its property or funds, then its results will also be mandatory for subsequent application. In this regard, it is possible

to complete its conduct with a report, and the implementation of the recommendations will be traced by the relevant state authorities;

2) evaluation of activities is carried out at the initiative of the object itself. In this case, the use of the recommendations reflected in the report is left to the discretion of the management of the medical facilities.

4. Conclusion

The aspects to assess have been identified to provide decision-makers with an overview of the situation of health care facilities. The overview is finalized to modernization and to outline how to increase the physical environment contribution to the resilience of the entire health care system.

Implementing modernization strategies of health care facilities to ensure high standard care and effective using of given resources is one of the duties of the decision-makers. This process aimed to elicit expert reasoning by using the modern equipment, i.e., magnetic resonance scanner, computer scanner, etc. The customization has concerned objectives of the assessment, main aspects, algorithms, metrics, and outcomes representation. Aspects with reference to specific established goals and they are based on a hierarchy of main indicators which assign greater importance to safety and functionality. In addition, the algorithms take into account the relationships between the hospital facilities parts, considering them as complex systems of result-oriented health care.

The safety of utility employees, contractors, and the public is paramount during this time, while addressing the new power system operating challenges [11].

This will be achieved by adopting the Integration with municipal/federal disaster response structures, as appropriate.

References

1. United Nations Secretary-General, Report of the Open-Ended Intergovernmental Expert Working Group on Indicators and Terminology Relating to Disaster Risk Reduction, A/71/644), 2016, pp. 1-41, https://www.preventionweb.net/files/50683_oiewgreportenglish.pdf. (Accessed 06 June 2022).
2. World Health Organization (WHO), Organisation for Economic Co-operation and Development (OECD), International Bank for Reconstruction and Development, Delivering Quality Health Services: a Global Imperative for Universal Health Coverage, 2018, p. 93.
3. World Health Organization (WHO), Comprehensive Safe Hospital Framework, 2015, pp. 1-12.
4. United Nations Office for Disaster Risk Reduction (UNDRR), International Science Council (ISC), Hazard Definition & Classification Review: Technical Report, 2020, pp. 1-88. <https://www.undrr.org/publication/hazard-definition-and-classification-review>. (Accessed 10 June 2022).
5. World Health Organization (WHO), Health Emergency and Disaster Risk Management Framework, 2019, pp. 1-31. <https://www.who.int/publications/i/item/9789241516181>. (Accessed 10 June 2022).
6. United Nations Office for Disaster Risk Reduction (UNDRR), GAR, Global Assessment Report on Disaster Risk Reduction, 2019, pp. 35-78. <https://gar.unisdr.org>. (Accessed 2 September 2021).
7. Antipova T. (2021) Digital View on COVID-19 Impact. In: Antipova T. (eds) *Comprehensible Science. ICCS 2020. Lecture Notes in Networks and Systems*, vol 186, pp. 155-164. Springer, Cham. https://doi.org/10.1007/978-3-030-66093-2_15.
8. Grimaz S., Ruzzene E., Zorzini F. Situational assessment of hospital facilities for modernization purposes and resilience improvement. *International Journal of Disaster Risk Reduction*, **66** (2021), 102594. <https://doi.org/10.1016/j.ijdr.2021.102594>.

9. Antipova T. Insights from Some Governments' Budget Functional Expenditures for the Fifteen Years: 2005–2019. In: T. Antipova (Ed.): ICCS 2021, LNNS 315, 1–11, 2022. https://doi.org/10.1007/978-3-030-85799-8_6.
10. Antipova, T., Bourmistrov A. (2013) "Is Russian Public Sector Accounting in the Process of Modernization? An Analysis of Accounting Reforms in Russia". *Financial Accountability & Management*, 29(4), November 2013, pp. 442 – 478. <https://doi.org/10.1111/faam.12021>.
11. Skarvelis-Kazakos, S. et al. Resilience of electric utilities during the COVID-19 pandemic in the framework of the CIGRE definition of Power System Resilience. *International Journal of Electrical Power & Energy Systems*, Vol. 136 (2022) 107703, 1-17. <https://doi.org/10.1016/j.ijepes.2021.107703>.
12. C. Lankford Walker (1993) A Cross-Sectional Analysis of Hospital Profitability, *Journal of Hospital Marketing*, 7:2, 121-138, DOI: 10.1300/J043v07n02_11
13. Whinney Ernest S. Health Care Notes (1985) Considerations in Hospital Acquisitions, *Hospital Topics*, 63:1, 33, DOI: 10.1080/00185868.1985.9948394
14. Davis, Peter et al. Efficiency, effectiveness, equity (E3). Evaluating hospital performance in three dimensions. *Health Policy*, Vol. 112. <https://doi.org/10.1016/j.healthpol.2013.02.008>
15. Health System Efficiency - WHO European Region. <https://www.euro.who.int>.

Aims and Objectives

Published online by ICS two times a year, Journal of Digital Science (JDS) is an international peer-reviewed journal which aims at the latest ideas, innovations, trends, experiences and concerns in the field of digital science covering all areas of the scholarly literature of the sciences, social sciences and arts & humanities. The main topics currently covered include: Artificial Intelligence Research; Digital Economics, Education, Engineering, Finance, Health Care.

The main goal of the journal is the effective dissemination of original incites/results generated by the human brain and presented/reflected in articles using modern information/digital technology.

Editorial Board

Editor-in-Chief Tatiana Antipova, ICS,
<https://orcid.org/0000-0002-0872-4965>

Associate Editor Julia Belyasova, Catholic University of Louvain, Louvain-la-Neuve, Belgium;
<https://orcid.org/0000-0001-6983-2129>

Editors

Abdulsatar Sultan, Catholic University in Erbil, Erbil, Iraq;

<https://orcid.org/0000-0001-5090-5332>

Achmad Nurmandi, Universitas Muhammadiyah Yogyakarta, Indonesia

<https://orcid.org/0000-0002-6730-0273>

Jelena Jovanovic, University of Nis, Nis, Serbia;

<https://orcid.org/0000-0001-7238-6393>

Indra Bastian, Universitas Gadjah Mada, Yogyakarta, Indonesia;

<https://orcid.org/0000-0003-4658-8690>

Indrawati Yuhertiana, Universitas Pembangunan Nasional Veteran Jatim, Surabaya, Indonesia;

<https://orcid.org/0000-0002-1613-1692>

Lucas Tomczyk, Uniwersytet Jagielloński, Krakow, Poland

<https://orcid.org/0000-0002-5652-1433>

Narcisa Roxana Moşteanu, American University of Malta, Bormla, Malta

<https://orcid.org/0000-0001-5905-8600>

Olga Khlynova, Russian Academy of Science, Moscow, Russia

<https://orcid.org/0000-0003-4860-0112>

Omar Leonel Loaiza Jara, Universidad Peruana Unión, Lima, Peru

<https://orcid.org/0000-0002-3262-709X>

Roland Moraru, University of Petrosani, Romania

<https://orcid.org/0000-0001-8629-8394>

Tjerk Budding, Vrije Universiteit Amsterdam, Netherland

<https://orcid.org/0000-0002-5343-7535>

Zhanna Mingaleva, National Research Polytechnic University, Perm, Russia

<https://orcid.org/0000-0001-7674-7846>

Quang Vinh Dang, Industrial University, Ho Chi Minh City, Viet Nam

<https://orcid.org/0000-0002-3877-8024>

Contact information

Website: <https://ics.events>

Email: conf@ics.events