Text Big Data Analytics Case Study «Third Wave»: Internet of Words

Olga Kolesnichenko
Security Analysis Bulletin
Moscow, Russia
oykolesnichenko@list.ru

Gennady Smorodin
Dell EMC External Research and Academic Alliances
St. Petersburg, Russia
gennady.smorodin@dell.com

Dariya Yakovleva, Lev Mazelis
Vladivostok State University of Economics and Service
Vladivostok, Russia
darya.yakovleva15@vvsu.ru
lev.mazelis@vvsu.ru

Sergey Balandin
Open Innovations Association FRUCT,
University of Information Technologies, Mechanics
and Optics (ITMO University)
St. Petersburg, Russia
Sergey.Balandin@fruct.org

Yuriy Kolesnichenko
Cybersecurity Technology
Moscow, Russia
green-apple_2000@mtu-net.ru

Abstract—The results of Text Big Data Analytics Study «Third Wave» is presented. The Study is provided by Dell EMC External Research and Academic Alliances in Russia. Counted keywords are investigated in terms of Glocalization (Globalization plus Localization) issues. The findings include: Glocalization picture of countries' interest in Euro and Dollar; also Glocalization picture of countries' reaction to economic crisis and health impact; geographic picture of Sixth Kondratiev's Wave capture. The very significant role of US Dollar in the world social and economic stability was recognized. United States was identified as world leader of Sixth Kondratiev's Wave capture. Text Big Data Analytics using API access to Google and Yandex shows new knowledge about Global society, the investigation can be called «API-sociology».

Google and Yandex can be applied as «tools» to investigate Glocalization. The term «Glocalization» (Globalization plus Localization) has been proposed by sociologist, professor at University of Pittsburgh Roland Robertson [1]. It means reflection of Globalization process on local regional scale. Modern information technologies allow to collect data and investigate collected data in terms of reflection of Globalization process on local regional scale. The example of elaborated IT Glocalization's tool is the concept of Federation Business Data Lake, that has global scale due to Cloud computing and local scale due to peripheral data-centers at edge [2]. The similar features have Google and Yandex that are the global systems due to Internet and Cloud computing, and at the same time they collect data at local scale due to peripheral data-centers at edge.

The Text Big Data Analytics Study «Third Wave» is provided by Dell EMC External Research and Academic Alliances in Russia. In the Study some political and economic processes were investigated using API access from the point of view of Glocalization: mention currencies and topicality of crisis issue in texts in conjunction with the countries names.

Another approach to investigate Glocalization is Kondratiev's Waves geographic capture assessment.

Using API access to Google and Yandex data Morphological Matrix of several keywords was collected. Google and Yandex are considered as non-classical hybrid supercomputers with API access «as-a-Service» to data-storages (Cloud model). This Study focuses on an important task to investigate Text Big Data. Open resources analytics from Internet leads to better understanding of supercomputer's analysis of text arrays. Moreover Big Data Analytics of Internet's open data allows to investigate not just text data but the evolutive global system «People-to-IT» with understanding of Globalization and Glocalization processes.

I. GLOBAL + LOCAL PICTURE OF CURRENCIES PREFERENCE

In this Study the local picture of countries dependence on two global leading currencies Euro and Dollar is shown. The number of counted keywords «euro exchange rate» and «dollar exchange rate» in conjunction with countries names and 2015 year of texts publication, reflect the countries interest in Euro and Dollar. But also should be noticed that the word «dollar» is included in the names of several national currencies: Australian dollar, Barbadian Dollar, East Caribbean Dollar, Hong Kong Dollar, Canadian Dollar, New Zealand Dollar, Singapore Dollar, United States Dollar, New Taiwan dollar. In all 49 countries' names from different regions of the world were engaged into Data Mining. Dariya Yakovleva is the author of Data Mining technique using API-Google and API-Yandex. The common picture of all 49 countries is presented at Fig. 1; some countries of post-Soviet area with results of Data Mining are presented at Fig. 2; countries with big economics – at Fig. 3; countries from Top

Five of Index of Economic Freedom rating 2015 – at Fig. 4; some countries from Middle East, North Africa, Central Asia – at Fig. 5; from Asia-Pacific Region – at Fig. 6; from Europe – at Fig. 7; from South America and Caribbean – at Fig. 8.
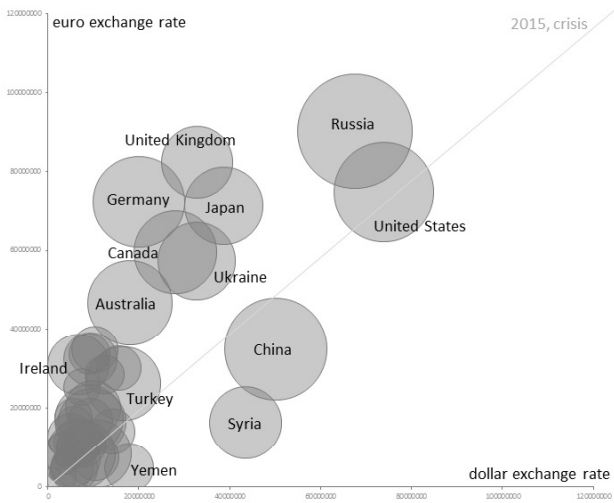

Fig. 1. Countries interest in Euro and Dollar. Counted keywords in million («euro exchange rate» and «dollar exchange rate») in conjunction with countries names and 2015 year of texts publication. The size of the circles is number of keywords «crisis» in million.
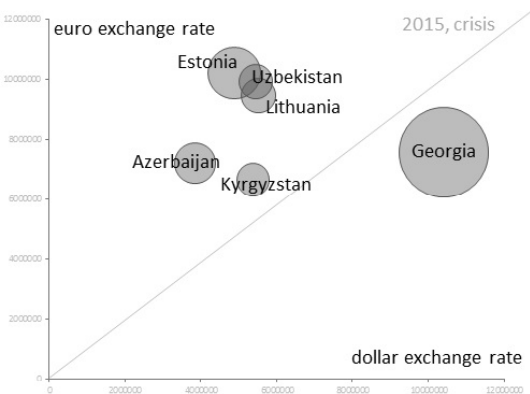

Fig. 2. Post-Soviet countries interest in Euro and Dollar. Counted keywords in million («euro exchange rate» and «dollar exchange rate») in conjunction with countries names and 2015 year of texts publication. The size of the circles is number of keywords «crisis» in million.
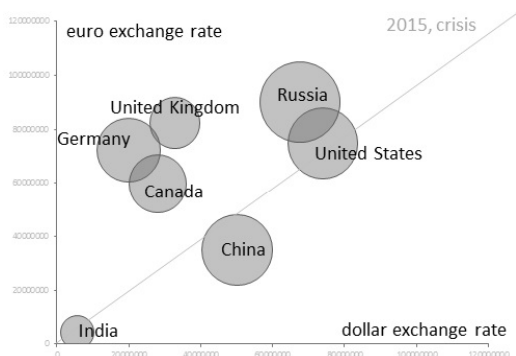

Fig. 3. Countries (with big economics) interest in Euro and Dollar. Counted keywords in million («euro exchange rate» and «dollar exchange rate») in conjunction with countries names and 2015 year of texts publication. The size of the circles is number of keywords «crisis» in million.
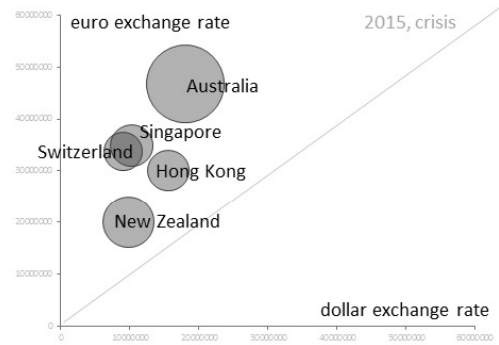

Fig. 4. Countries (from Top Five of Index of Economic Freedom rating 2015) interest in Euro and Dollar. Counted keywords in million («euro exchange rate» and «dollar exchange rate») in conjunction with countries names and 2015 year of texts publication. The size of the circles is number of keywords «crisis» in million.
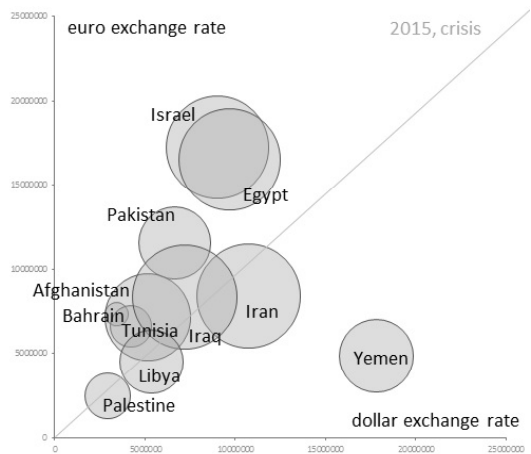

Fig. 5. Countries (Middle East, North Africa, Central Asia) interest in Euro and Dollar. Counted keywords in million («euro exchange rate» and «dollar exchange rate») in conjunction with countries names and 2015 year of texts publication. The size of the circles is number of keywords «crisis» in million.
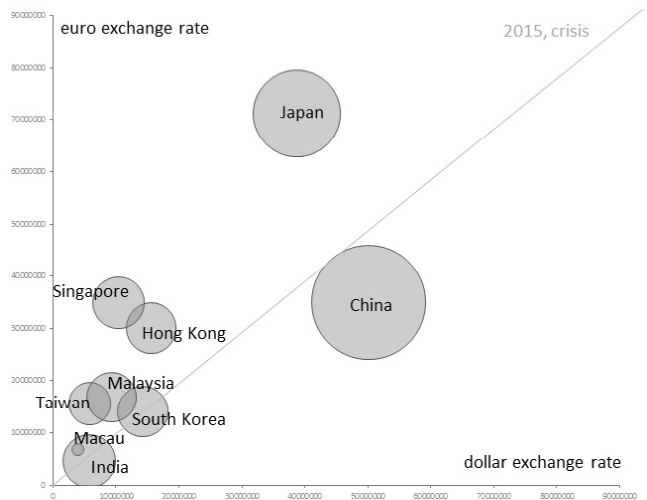

Fig. 6. Countries (Asia-Pacific Region) interest in Euro and Dollar. Counted keywords in million («euro exchange rate» and «dollar exchange rate») in conjunction with countries names and 2015 year of texts publication. The size of the circles is number of keywords «crisis» in million.
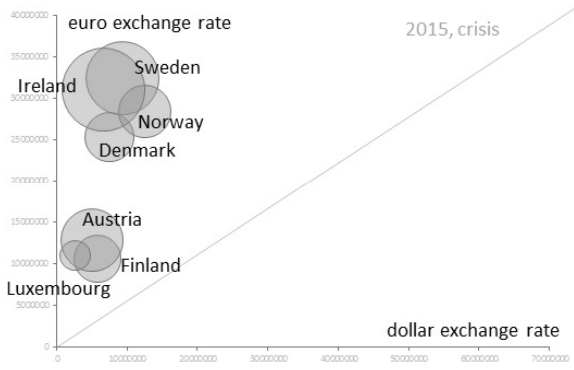
Fig. 7. European countries interest in Euro and Dollar. Counted keywords in million («euro exchange rate» and «dollar exchange rate») in conjunction with countries names and 2015 year of texts publication. The size of the circles is number of keywords «crisis» in million.
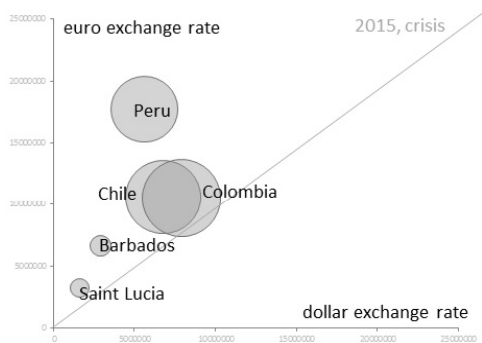

Fig. 8. Countries (South America and Caribbean) interest in Euro and Dollar. Counted keywords in million («euro exchange rate» and «dollar exchange rate») in conjunction with countries names and 2015 year of texts publication. The size of the circles is number of keywords «crisis» in million.

Looking to diagrams at Fig. 1-8 can be described Glocalization picture regarding two global currencies. Can be mentioned that Dollar has different features than Euro. There are some countries which showed more interest in Dollar than Euro – China, Syria, Yemen, Iran, Georgia. Another countries showed more interest in Euro than Dollar – Europe countries, Israel, Egypt, Japan, Peru, Canada, Ukraine, Azerbaijan.

The next set of diagrams (Fig. 9-14) shows the dependence between keywords that reflect interest in currencies and keywords that reflect reaction to crisis. Data includes results about all engaged into Data Mining 49 countries. The very strong Pearson correlation coefficient was found between keywords «dollar exchange rate» & «crisis» and the strong correlation coefficient was found between keywords «euro exchange rate» & «crisis». Also the very strong Pearson correlation coefficient was found between keywords «dollar exchange rate» & «inflation» and the strong correlation coefficient was found between keywords «euro exchange rate» & «inflation».

Correlation less than 0,7 was found for counted keywords «price increase». The strong Pearson correlation coefficient was found between keywords «dollar exchange rate» & «price increase» and also the strong correlation coefficient was found between keywords «euro exchange rate» & «price increase».

Based on this finding we can notice that Text Big Data Analytics shows the important role of US Dollar in the world economics. Such processes as crisis and inflation, reflected through keywords, have the very strong correlations with keywords «dollar exchange rate».
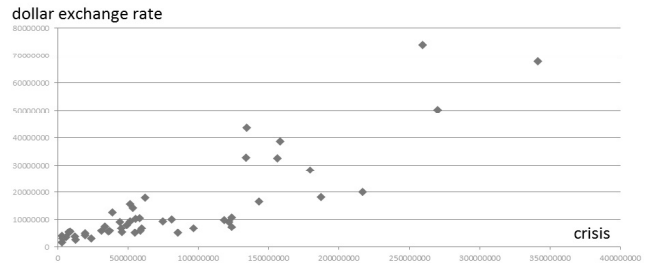

Fig. 9. Dependence between counted keywords in million, 2015, «dollar exchange rate» & «crisis», all 49 countries.
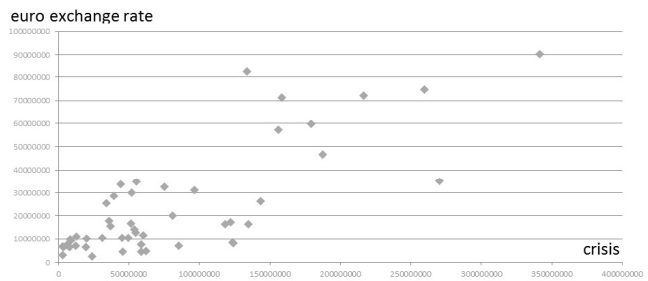Pearson correlation coefficient r = 0,85.

Fig. 10. Dependence between counted keywords in million, 2015, «euro exchange rate» & «crisis», all 49 countries.
Pearson correlation coefficient r = 0,78.

Fig. 11. Dependence between counted keywords in million, 2015, «dollar exchange rate» & «inflation», all 49 countries.
Pearson correlation coefficient r = 0,81.

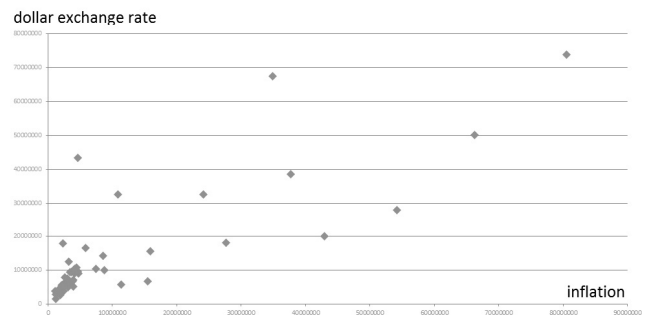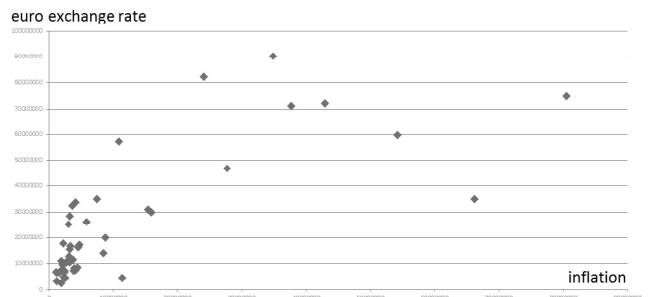Fig. 12. Dependence between counted keywords in million, 2015, «euro exchange rate» & «inflation», all 49 countries.
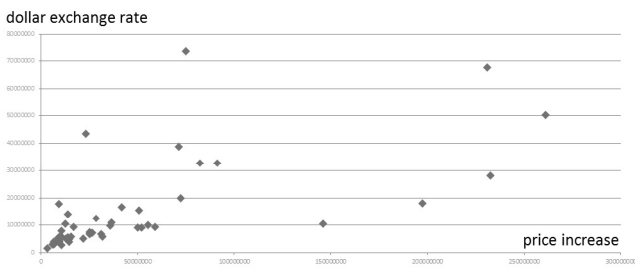Pearson correlation coefficient r = 0,75.

Fig. 13. Dependence between counted keywords in million, 2015, «dollar exchange rate» & «price increase», all 49 countries.
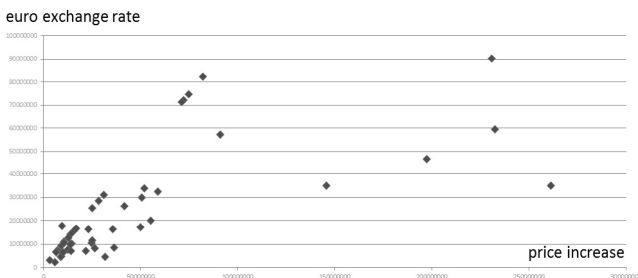Pearson correlation coefficient r = 0,65.

Fig. 14. Dependence between counted keywords in million, 2015, «euro exchange rate» & «price increase», all 49 countries.
Pearson correlation coefficient r = 0,68.

President of the World Economic Forum Klaus Schwab in the article «The Age of Adaptation» wrote that the «current technological revolution hits economies like a tsunami», causing crisis, high level of unemployment and social inequality [3]. On the other hand human capital is the main engine of the 21st century economics. Klaus Schwab proposed the term «Ecosystem of Change», this ecosystem consolidates government, society and business. According to the report of the World Economic Forum «The Global Risks 2016» [4], the significant risks for global society are the growing social inequality, social instability, financial crisis and unemployment. Text Big Data Analytics allows to investigate Glocalization picture of crisis in countries.

Group of countries named «Social Impact» was found (Fig. 15). These countries are characterized by significantly higher frequency of counted keywords «crisis», «inflation», «price increase», «unemployment», «dismissal». The group «Social Impact» was divided into three subgroups depending on list of keywords that have higher frequency: only one keyword – Social Impact-1, two or three keywords – Social Impact-2-3, four or five keywords – Social Impact-4-5.
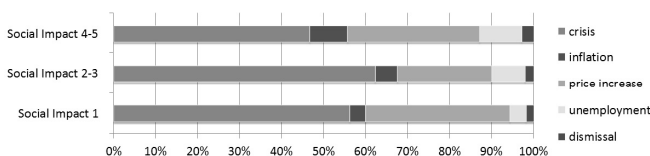

Fig. 15. Countries of group «Social Impact», 2015. The percentage of number of keywords in million for each country.
Social Impact 1 – Egypt, India, Iran, Iraq, New Zealand, Singapore, Sweden, Switzerland; Yemen;
Social Impact 2-3 – Hong Kong, Ireland, Israel, Syria, Turkey;
Social Impact 4-5 – Australia, Canada, China, Germany, Japan, Russia, Ukraine, United Kingdom, USA, Afghanistan.

Subgroup of countries with worse Social Impact (Social Impact-4-5) has bigger proportion of keywords «inflation», «unemployment», «dismissal». It means that countries of subgroup Social Impact-4-5 really undergo the new technological wave's negative impact to society.

Countries of subgroups Social Impact-2-3 and Social Impact-4-5 were analyzed in terms of relationship between keywords «crisis» & «dollar exchange rate» and «crisis» & «euro exchange rate». At Fig. 16 number of keywords «crisis» has the strong correlation with number of keywords «dollar exchange rate». But there isn't significant correlation between counted keywords «crisis» and «euro exchange rate» (moderate Pearson correlation coefficient, see Fig. 17) . This is the sign that countries of group «Social Impact» depend on US Dollar.

Fig. 16. Dependence between counted keywords in million, 2015, «crisis» & «dollar exchange rate». Countries of group «Social Impact».
Pearson correlation coefficient r = 0,76.

Fig. 17. Dependence between counted keywords in million, 2015, «crisis» & «euro exchange rate». Countries of group «Social Impact».
Pearson correlation coefficient r = 0,58.

At Fig. 18 and 19 countries from rating of Bloomberg Innovation Index 2016 (World's Most Innovative Economies) are compared [5]. Bloomberg Innovation Index includes countries with high activity in scientific researches and patents and with high percentage of innovation component in economics. There are 9 countries of group «Social Impact» among countries from Top of Bloomberg Innovation Index – Germany, Sweden, Japan, Switzerland, Singapore, United States, Israel, Russia, Ireland.

Countries of Top of Bloomberg Innovation Index showed less number of keywords «dollar exchange rate» in comparison with number of keywords «euro exchange rate». Very strong

correlation was found between number of keywords «crisis» & «dollar exchange rate» and «crisis» & «euro exchange rate». Significant role of Euro in innovation development of countries' economies can be noticed.
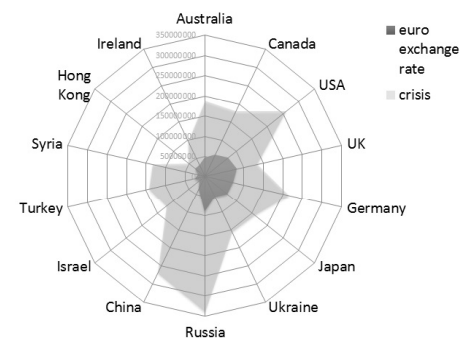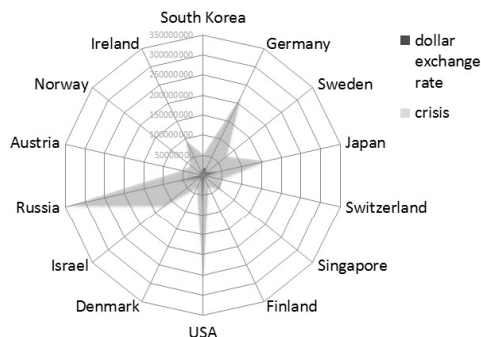

Fig. 18. Dependence between counted keywords in million, 2015, «crisis» & «dollar exchange rate». Countries of group Top of Bloomberg Innovation Index 2016 (World's Most Innovative Economies).
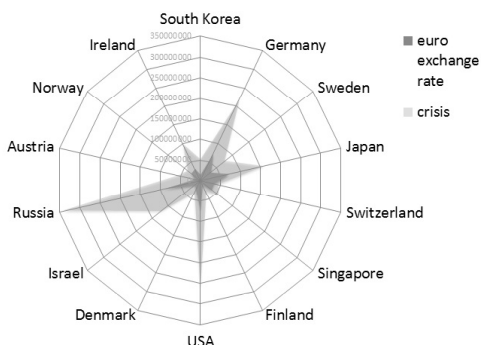Pearson correlation coefficient r = 0,88.

Fig. 19. Dependence between counted keywords in million, 2015, «crisis» & «euro exchange rate». Countries of group Top of Bloomberg Innovation Index 2016 (World's Most Innovative Economies).
Pearson correlation coefficient r = 0,90.

This investigation we call «API-sociology», that allows to measure the «Ecosystem of Change» of global society. The very important role of US Dollar in the global processes of social and economic stability or instability was showed. Innovative development of economics more connects to Euro.

II. GLOBAL + LOCAL PICTURE OF HEALTH IMPACT

Presence of different countries on the Internet is very varied. For investigation of health impact cluster analysis k-means was implemented by Dariya Yakovleva. Cluster analysis helped to divide countries into groups with different activity and different levels of citing on the Internet. Five clusters are found (Fig. 20).

High level of Internet activity and citing: Cluster № 1 includes United States, Canada, Germany, China – the most active countries on the Internet. Cluster № 3 includes United Kingdom, Australia, Japan, Russia. These countries have high level of activity and citing on the Internet.

Middle level of Internet activity and citing: Cluster № 2 includes Hong Kong, Singapore, New Zealand, Ireland, Malaysia, Ukraine, Georgia, Turkey, India. Cluster № 4

includes Switzerland, Chile, Sweden, Denmark, Norway, Colombia, Taiwan, South Korea, Austria, Israel, Afghanistan, Pakistan, Egypt, Iraq, Iran, Syria. Both clusters № 2 and № 4 contain countries with middle level of activity and citing on the Internet.

Low level of Internet activity and citing: Cluster № 5 includes Luxembourg, Estonia, Lithuania, Finland, Macau, Saint Lucia, Barbados, Peru, Libya, Bahrain, Tunisia, Yemen, Palestine, Azerbaijan, Kyrgyzstan, Uzbekistan. This cluster contains countries with low activity and citing on the Internet.
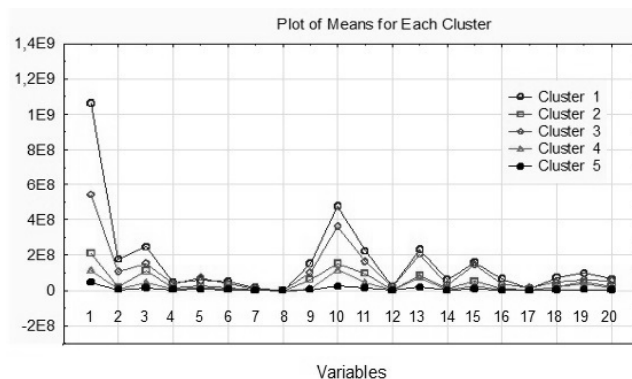

Fig. 20. Cluster analysis, k-means algorithm. 49 countries are divided into five clusters. Variables – matrix of 20 Keywords Phrases, 2015: 1 – Computer; 2 – Cloud computing; 3 – Mobile phone; 4 – Dollar exchange rate; 5 – Euro exchange rate; 6 – Electric cars; 7 – Solar panel; 8 – Drip irrigation; 9 – Gas supplies; 10 – Oil; 11 – Oil price; 12 – Nuclear power plant; 13 – Crisis; 14 – Inflation; 15 – Price increase; 16 – Unemployment; 17 – Dismissal; 18 – Migrants; 19 – Refugees; 20 – Muslims.

It's hard to take into account all the factors affecting the activity on the Internet, that is different for different countries and depends on number of computers in the country; level of literacy and IT education; number of IP connections and Internet accessibility; level of economic development and number of Internet services; amount of country's population; processes in politics and society, which are actively discussed; citing of country. Clusterization was chosen to arrange countries with different number of keywords on the Internet. These clusters allocation should be taken into account. But at the same time more or less activity and citing of countries on the Internet point out the economic development in countries.

After clusterization divided countries were analyzed in terms of mortality from cardiovascular diseases, cancer, diabetes mellitus, communicable diseases (data from WHO Global Health Observatory). The several very strong Pearson correlation coefficients were found.

Countries of High level of Internet activity and citing:
counted keywords in million, 2015, «crisis» &

- adult mortality rate (between 15 and 60 years per 1000 population), r = 0,84;
- cardiovascular diseases (age-standardized mortality rate per 100 000 population), r = 0,88;
- noncommunicable diseases (age-standardized mortality rate per 100 000 population), r = 0,89;

- malignant neoplasms (age-standardized mortality rate per 100 000 population), r = 0,70.

counted keywords in million, 2015, «dismissal» &

- adult mortality rate (between 15 and 60 years per 1000 population), r = 0,95;
- cardiovascular diseases (age-standardized mortality rate per 100 000 population), r = 0,81.

counted keywords in million, 2015, «unemployment» &

- cardiovascular diseases (age-standardized mortality rate per 100 000 population), r = 0,80.

counted keywords in million, 2015, «inflation» &

- diabetes mellitus (age-standardized mortality rate per 100 000 population), r = 0,80.

Countries of Middle level of Internet activity and citing:

counted keywords in million, 2015, «crisis» &

- adult mortality rate (between 15 and 60 years per 1000 population), r = 0,85;

Countries of Low level of Internet activity and citing:

counted keywords in million, 2015, «crisis» &

- communicable diseases (age-standardized mortality rate per 100 000 population), r = 0,78.
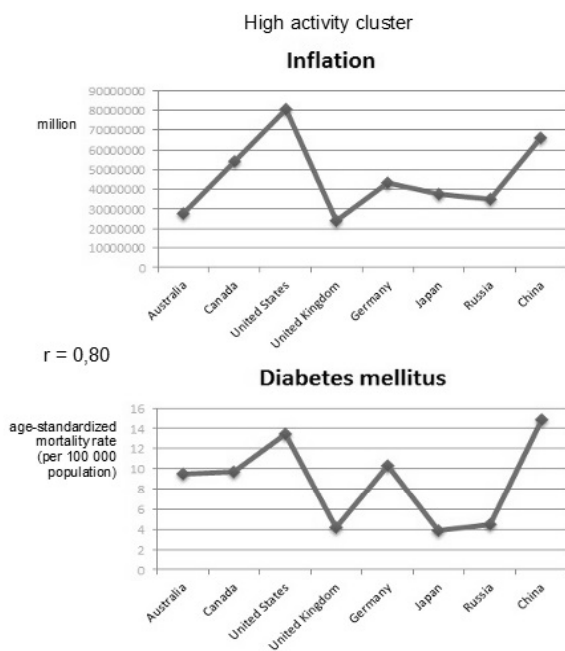

Fig. 21. The comparison of data of Text Big Data Analytics (keywords «inflation») and data of WHO Global Health Observatory (diabetes mellitus, age-standardized mortality rate per 100 000 population). Pearson correlation coefficient r = 0,80.

Text Big Data Analytics shows the important results, global economic crisis causes the population mortality from diseases. In countries of clusters with High level of Internet activity and citing (that are big and economically developed) keywords «crisis», «dismissal», «unemployment», «inflation» have the correlation with population mortality from noncommunicable diseases, including cardiovascular diseases and cancer. It is interesting that keywords «inflation» have the

correlation with population mortality from diabetes mellitus – financial crisis and rising prices constrain people to abandon the expensive diet products and to buy cheap harmful products (Fig. 21-23).
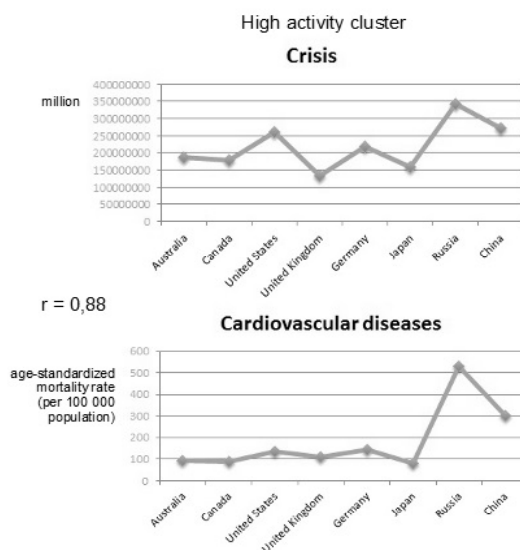

Fig. 22. The comparison of data of Text Big Data Analytics (keywords «crisis») and data of WHO Global Health Observatory (cardiovascular diseases, age-standardized mortality rate per 100 000 population). Pearson correlation coefficient r = 0,88.
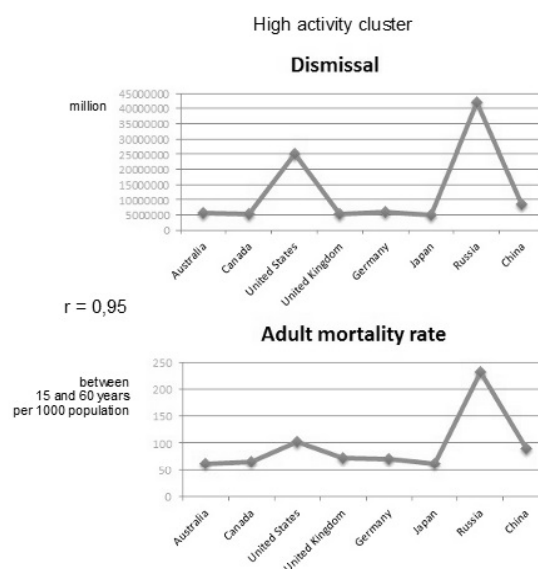
Fig. 23. The comparison of data of Text Big Data Analytics (keywords «dismissal») and data of WHO Global Health Observatory (adult mortality rate, between 15 and 60 years per 1000 population). Pearson correlation coefficient r = 0,95.

In countries of cluster with Low level of Internet activity and citing (that include not only small developed countries, but many countries with poor economics from Middle East, North Africa, Central Asia) was found the correlation between keywords «crisis» and mortality from communicable diseases (infectious diseases). Poor countries being under economic crisis incur epidemic of infectious diseases (Fig. 24).

Fig. 24. The comparison of data of Text Big Data Analytics (keywords «crisis») and data of WHO Global Health Observatory (communicable diseases, age-standardized mortality rate per 100 000 population). Pearson correlation coefficient r = 0,78.
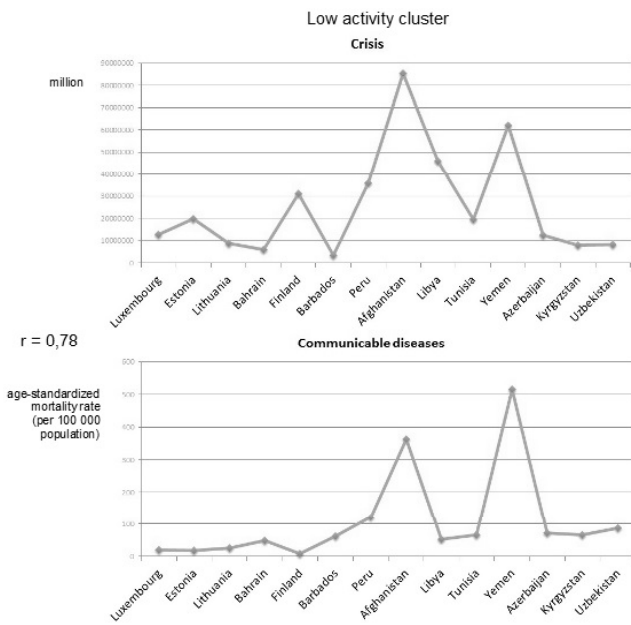
Thus Text Big Data Analytics allows to understand Glocalization picture of economic crisis impact to population health. World Health Organization and domestic national health care systems cannot achieve the objectives of health care alone without measures to overcome the economic crisis on the global scale.

IV. GLOBAL + LOCAL PICTURE OF ADVANCED ECONOMICS

The next part of the Study was dedicated to turnround of global Technological Order and Kondratiev's Waves. Collected Morphological Matrix of several keywords was divided to four units:

- Unit «Information»: computer; cloud computing; mobile phone; dollar exchange rate; euro exchange rate;

- Unit «Energy & Material»: electric cars; solar panel; drip irrigation; gas supplies; oil; oil price; nuclear power plant;

- Unit «Social inequality»: crisis; inflation; price increase; unemployment; dismissal;

- Unit «Stress load»: migrants; refugees; Muslims.

All 49 countries, that were included into Study, were grouped to three groups depending on keywords predominance: Unit «Information», Unit «Energy & Material» and Unit «Social inequality / Stress load» (Fig. 25-28).

The Big Data Analytics Multicenter Study «Third Wave» has been named after Alvin Toffler's «Third Wave» concept about information era [6]. The First Toffler's Wave had been agricultural, the economy had been based on land. The Second Toffler's Wave has been industrial, economy has been dependent on industry and hydrocarbon energy. Today weak

countries are still living in conditions of the industrial Second Wave. Third Wave is modern time and the future, the economy is linked to information technology and Globalization.
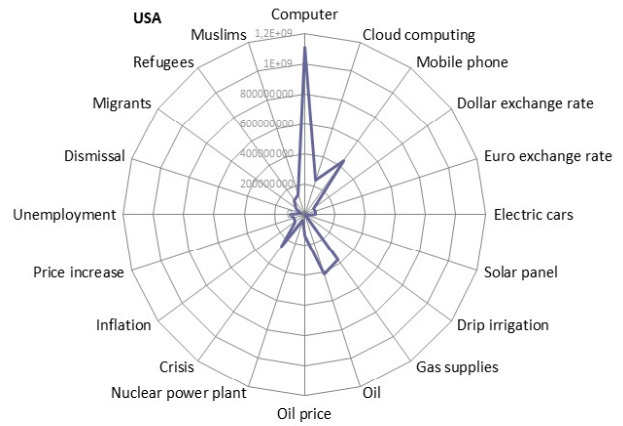
Fig. 25. Morphological Matrix of keywords for United States, 2015. World leader, best Sixth Kondratiev's Waves capture. Prepotency of keywords of Unit «Information».
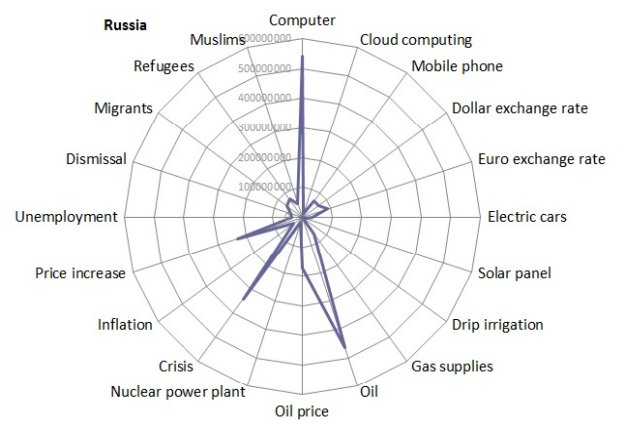

Fig. 26. Morphological Matrix of keywords for Russia, 2015. Prepotency of keywords of Unit «Energy & Material». Russia is still under Fifth Kondratiev's Wave.
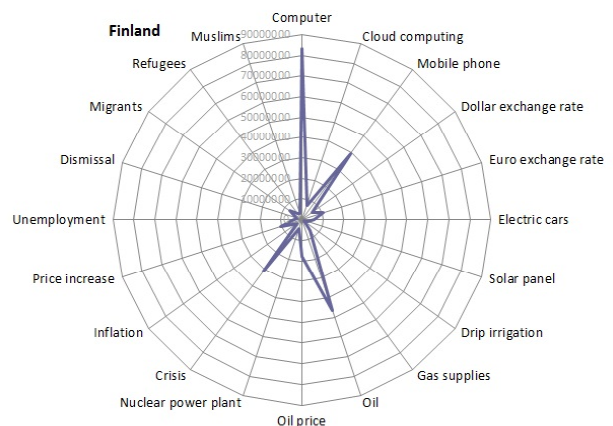

Fig. 27. Morphological Matrix of keywords for Finland, 2015. Prepotency of keywords of Unit «Information», Sixth Kondratiev's Waves capture.

Kondratiev's Waves are another measure of the evolution of Technological Order [7; 8; 9; 10]. These Waves aren't so extended during history as Toffler proposed his three-waves concept. Kondratiev's Waves endure 30-50 years. Start of Cybernetic revolution had coincided with the Fourth Kondratiev's Wave (1947-1991). Ten keywords of Morphological Matrix reflect the main modern economic trends and indicate different Kondratiev's Waves.
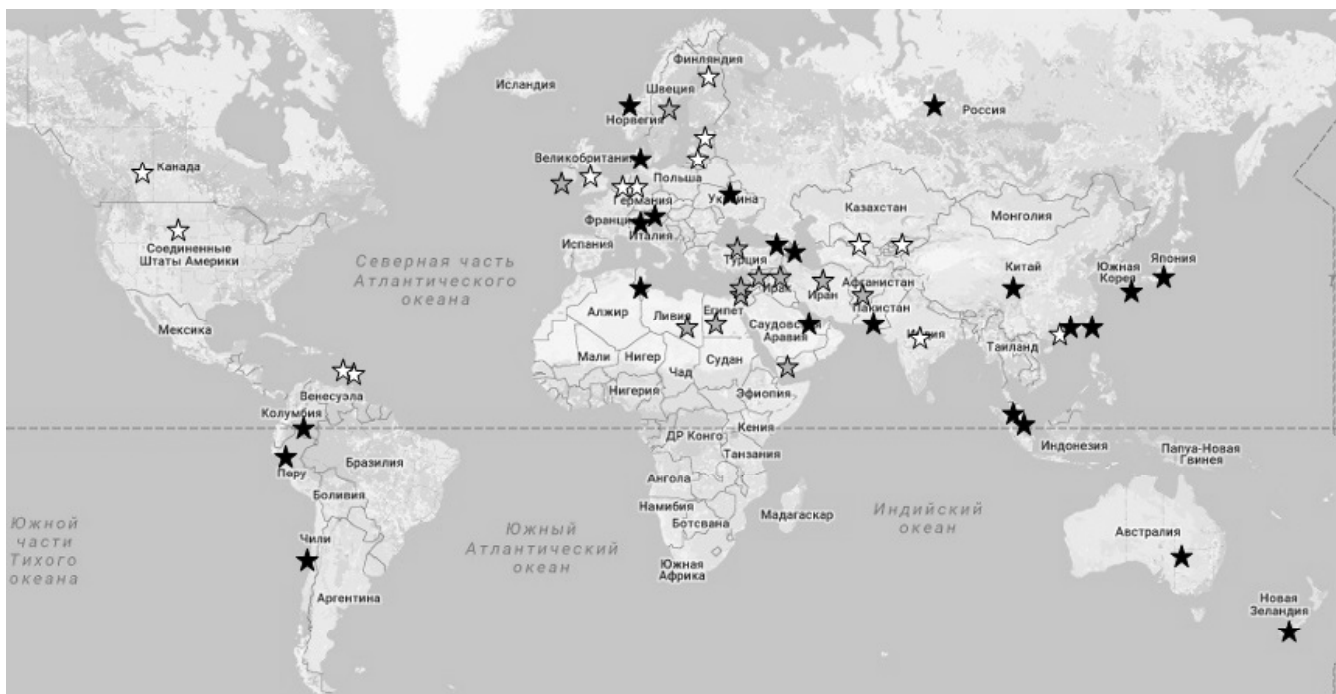

Fig. 28. The map with marks of countries that belong to group Unit «Information» (white marks), group Unit «Energy & Material» (black marks) and group Unit «Social inequality / Stress load» (grey marks).
Unit «Information» (white marks): Barbados, Canada, Estonia, Finland, Germany, India, Kyrgyzstan, Lithuania, Luxembourg, Macau, Saint Lucia, United Kingdom, United States, Uzbekistan.
Unit «Energy & Material» (black marks): Australia, Austria, Azerbaijan, Bahrain, Chile, China, Colombia, Denmark, Georgia, Hong Kong, Japan, Malaysia, New Zealand, Norway, Pakistan, Peru, Russia, Singapore, South Korea, Switzerland, Taiwan, Tunisia, Ukraine.
Unit «Social inequality / Stress load» (grey marks): Afghanistan, Egypt, Iran, Iraq, Ireland, Israel, Libya, Palestine, Sweden, Syria, Turkey, Yemen.

Keywords «Computer» have the highest frequency among all chosen keywords. This fact indicates the change of Technological Order to IT. Keyword «Computer» is the sign of Fourth Kondratiev's Wave, that today is completely finished.

Keywords «Mobile phone» also indicate the change of Technological Order to IT. Mobile phone can be attributed to the period of Fifth Kondratiev's Wave (1991-2020).

Keywords «Cloud computing» reflect the beginning of upcoming Sixth Kondratiev's Wave (2020-2070), during this Wave will be completing of Cybernetic revolution. The future ecosystem of IoT will be based on Cloud computing network of devices, vehicles, buildings and other items which are embedded with electronics, software, sensors. The most of these devices due to Cloud computing will be easy created by 3D-printing from plastic materials and also will include nanotechnology [11-16].

Keywords «Electric cars» reflect the interest in crucial technology that has already brought down the oil market. In other words, the electric cars are the real turning point in Technological Order. Bloomberg predicted a further downturn of oil price due to the electric cars growing market (Tesla Motors, Chevrolet, Ford, Volkswagen, Nissan, BMW, Apple,

Google, АвтоВАЗ, Local Motors). The most interesting trend is the emergence of 3D-printing technology for production of electric cars from carbon fibre composite tubes [17]. This means that oil and gas are no longer needed as fuel for new cars, but they are very needed as oil and gas chemistry products for 3D-printing of cars. Electric cars need electricity grid (Supercharger network), that more and more involve IoT, including household electricity. Common electricity grid is the inducement of transition to electric cars, because it allows to Big Data Analytics for cost savings. Keywords «Electric cars» apply to upcoming Sixth Kondratiev's Wave.

Keywords «Gas supplies»; «Oil»; «Oil price» indicate the use of hydrocarbons. During upcoming Sixth Kondratiev's Wave (2020-2070) the hydrocarbons will change the function from fuel to something as can be called «printing ink» for 3D-printing mass consumption products, including cars. People use of coal, oil and gas for industrial development in history can be considered as some episode of limited use of photosynthesis. The Kondratiev's Waves have been described since industrial obtaining of carbon has started.

The First Kondratiev's Wave (1803-1843) had been associated with industrial processing of coal; the Second

Kondratiev's Wave (1844-1896) had been related with ferrous metallurgy and railways; the Third Kondratiev's Wave (1896-1947) had been the raise of heavy machinery, electricity, steel and electric motors for industries (mechanical engineering, woodworking and other). The future progress of carbon industry related to artificial biogeochemical carbon cycle to obtain carbon from atmosphere.

Keywords «Drip irrigation» reflect agricultural IoT. Sensors of the drip irrigation systems at the fields create Big Data (electronic fields), that will lead to new economic value and efficiency. Drip irrigation as technology of IoT apply to Sixth Kondratiev's Wave.

Keywords «Solar panel» indicate the most important future oriented trend of transition to renewable energy sources. Solar panels associate with future of Global society.

Keywords «Nuclear power plant» reflect the traditional for Russia energy type. Russia also has such technology as floating (shipping) nuclear power stations.

United States is identified as world leader in terms of Sixth Kondratiev's Wave capture and new Technological Order accepting. United States is the root of Sixth Kondratiev's Wave expansion. Based on Text Big Data Analytics several countries can be included into zone of Sixth Kondratiev's Wave capture: Canada, Germany, United Kingdom, Finland, Estonia, Lithuania, Luxembourg, India. The most of another countries are still under Fifth Kondratiev's Wave.

For further analysis of Glocalization picture of new Technological Order accepting, the Value Chain Map was elaborated. Simon Wardley's approach was used for assess the economic situation in terms of competitiveness and shaping

the future consumer market [18], [19]. Wardley's map was modified, and counted keywords were added (Fig. 29, 30).
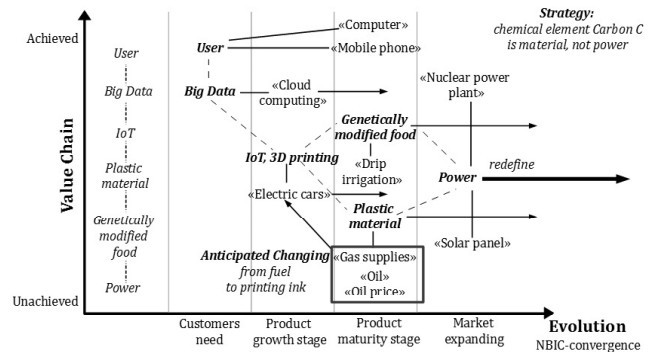

Fig. 29. The modified Value Chain Map for results of Text Big Data Analytics in Study «Third Wave». All the components of the Value Chain Map are evolving from left to right due to demand competition.

Making Value Chain Map we can look to patterns of keywords in dynamics. All components of the Value Chain Map are evolving from left to right due to demand competition. Value Chain Map allows to put keywords in needed sequence to determine what is already cover previous stage of development, and what is the innovative stage which determines future success during upcoming Sixth Kondratiev's Wave. The aim of Value Chain Map is to show a moment of advantage, a moment of competition. In presented chain the moment of competition, which will guarantee leadership, is energy power. Electricity generation will determine the acceleration of development of new technology and ecosystem of IoT.
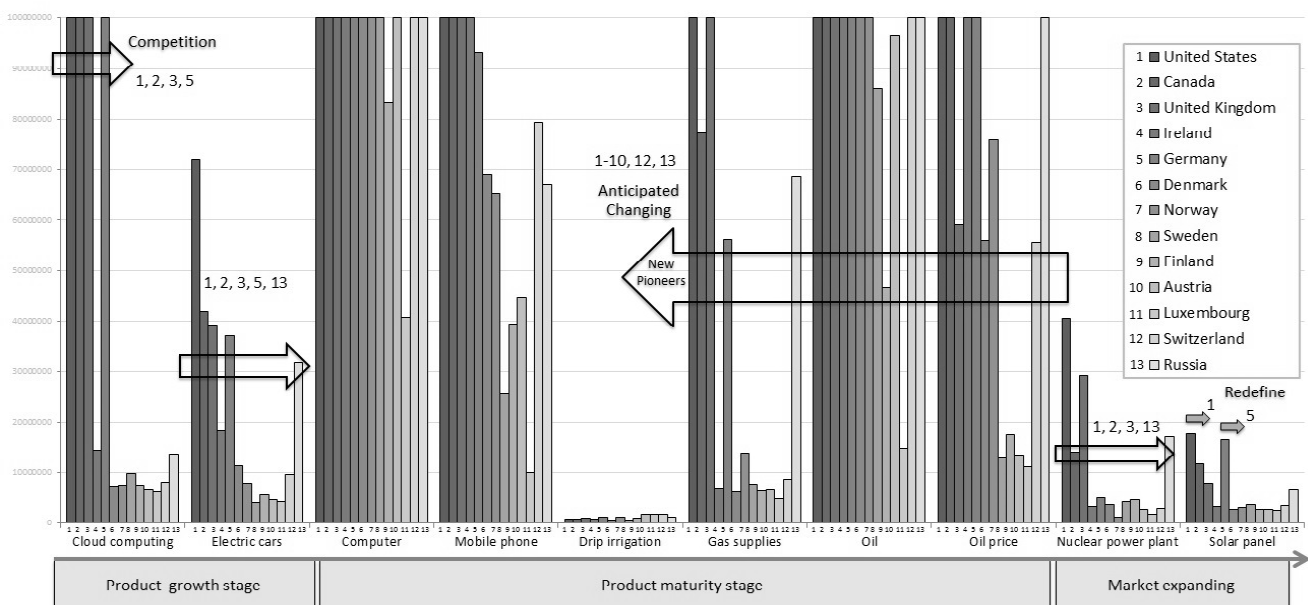

Fig. 30. The number of keywords in million during 2015 in conjunction with the countries names. Level above 100 million isn't shown. Patterns of keywords are presented in dynamics from left to right and are arranged relative to modified Value Chain Map. The arrows show the advantages for competition at the world marked during upcoming Sixth Kondratiev's Wave.

At the Fig. 30 the Value Chain Map for countries of Euro-Atlantic region is presented. The level of counted keywords «Cloud computing» increases for United States, Canada, United Kingdom and Germany. These countries achieved Sixth Kondratiev's Wave capture. United States, Canada, United Kingdom and Germany will be ahead among other countries achieving the widely distributed market of Cloud computing, that will ensure economic leadership.

United States, Canada, United Kingdom, Germany and Russia showed the increased level of counted keywords «Electric cars». This is also very crucial trend of development and the sign of future presence five countries at this new market on leading position. For example, today Russia has electric vehicles LADA Ellada, LADA Vesta, LIAS-6274.

United States and Germany show the growing interest in solar energy – the predominant number of counted keywords «Solar panel». The second row filled Canada, United Kingdom and Russia. During Sixth Kondratiev's Wave unlimited energy supply will ensure the advantages of economic development. For example, Russia has Orsk and Abakan Solar power plants with capacity 25 MW and 5,2 MW.

Russia traditionally focuses on nuclear energy that along with solar energy assure the advantages during increasing energy demand. On Fig. 30 four countries show the predominant number of counted keywords «Nuclear power plant» – United States, Canada, United Kingdom and Russia. It can be predicted that United States, Canada, United Kingdom, Germany and Russia will be the major electricity suppliers for world ecosystem of IoT during Sixth Kondratiev's Wave.

The most of the countries among presented on Fig. 30 show the growing interest in hydrocarbons. It can be predicted that Euro-Atlantic region will be pioneer at the new market of carbon materials, «printing ink» for 3D-printing, that will replace at hydrocarbons market fuel to materials.

V. CONCLUSION

Text Big Data Analytics is one of the High Performance Big Data Analytics (HPDA) priorities because the main goals of HPDA are creation of algorithms for texts «understanding» and consequently creation of Artificial Intelligence [20]. Big Data Analytics opposed to statistical methods allows to flexible and wide analysis [21-25]. Also Big Data Analytics allows to estimate countries in accordance with global system «People-to-IT».

API access to data storages is the modern and growing industry. In the Study «Third Wave» opportunity to extract new knowledge from Dark Data (data which wasn't recognized before as data) is investigated. It's impossible to apply classical statistical thinking to Text Big Data Analytics that needs to elaborate varied analytical methods. The methods of analysis should take into account the big common trends and specific differences at the same time.

Study allows to understand influence of Globalization on local scale that is called Glocalization process. Some findings were shown:

- the US Dollar has significant role in the global processes of social and economic stability or instability; innovative development of economics more connects to Euro;

- economic crisis has strong impact to population health; World Health Organization and domestic national health care systems cannot achieve the objectives of health care alone without measures to overcome the economic crisis at the global scale;

- United States is identified as world leader of Sixth Kondratiev's Wave capture and new Technological Order accepting; United States is the root of the Sixth Kondratiev's Wave expansion;

- based on Text Big Data Analytics several countries can be included into zone of Sixth Kondratiev's Wave capture: Canada, Germany, United Kingdom, Finland, Estonia, Lithuania, Luxembourg, India;

- Russia shows the advantage of energy supply due to nuclear energy; electricity generation will determine the acceleration of IoT ecosystem development during Sixth Kondratiev's Wave.

In this Study the capacity of morphological analysis using API is shown. Extraction the new knowledge from Dark Data illustrates how API access to Big Data storages can be used for producing different types of data analysis that help through the decision making process in many spheres – politics, economics, health care, industry, education, etc.

This Study can be called «API-sociology. It's capabilities are limited only by researchers creativity because many keywords about many issues can be investigated. The filling Internet by words is not simple process, it is the reflection of different complex processes from level of every human to local countries' level and global world level. The aim for researchers is the ability to find the hidden patterns in counted keywords.

REFERENCES

[1] R. Robertson, *Interpreting Globality. World Realities and International Studies Today*. USA, 1983. pp. 7-20.

[2] Dell EMC official website, Data Lake, Web: https://www.emc.com/en-us/big-data/data-lake/index.htm#accordion 4=0.

[3] K. Schwab, "The Age of Adaptation". *Project Syndicate*. Feb 19, 2015.

[4] "The Global Risks Report", 11th Edition. *World Economic Forum*, Geneva, 2016.

[5] "These Are the World's Most Innovative Economies". *Bloomberg Business*. USA, January 20, 2016.

[6] A. Toffler, *The Third Wave*, Transl., 1980. Moscow: AST Publishing, 2004.

[7] N.D. Kondratiev, Yu.V. Yakovets, L.I. Abalkin, *Big cycles of conjuncture and theory of foresight*. Moscow: Economics, 2002.

[8] A.V. Korotaev, S.V. Tsirel, "Kondratiev's waves in the world

economic dynamics", in *System monitoring. Global and regional development.* Moscow: Librokom / URSS, 2010. pp. 189-229.

[9] L.E. Grinin, "Kondratiev's waves, technological orders, and theory of industrial revolutions", in *Kondratiev's waves. Aspects and Perspectives.* Volgograd: Uchitel, 2012. pp. 222-262.

[10] L.E. Grinin, A.L. Grinin, "Cybernetic Revolution and the Sixth technological order", in *Kondratiev's waves. Heritage and Modernity.* Volgograd: Uchitel, 2015. pp. 83-106.

[11] Dell EMC official website, Hybrid Cloud computing. The Great Enabler of Digital Business, Web: http://www.emc.com/en-us/cloud/hybrid-cloud-computing/index.htm#accordion3=0&scroll Top=off.

[12] "3-D Printing, Cloud computing collide changing the face of small-scale manufacturing", *Global Manufacturing.* May 26, 2015. Web: http://www.manufacturingglobal.com/technology/449/3d-printing-cloudcomputing-collide-changing-the-face-of-smallscale-manufacturing.

[13] G. Davidson, "3-D Printing and the Cloud: a natural fit for manufacturers", *Industry Week.* 2015. Web: http://www.industryweek.com/emerging-technologies/3-d-printing-and-cloud-natural-fit-manufacturers.

[14] A. Cuthbertson, "How to turn your smartphone into a 3-D printer", *Newsweek,* 3.24.2016. Web: http://europe.newsweek.com/3d-printingolo-smartphone-printer-440033?rm=eu.

[15] R. Pepper, P. Biggs, J. Garrity, C. LaSalle, A. Polomska. "Harnessing the Internet of Things for Global Development", in *IoT Technologies (ITU).* Contribution to the ITU/UNESCO Broadband Commission for Sustainable Development, 2015.

[16] J. Manyika, S. Ramaswamy, S. Khanna, H. Sarrazin, G. Pinkus, G.

Sethupathy, A. Yaffe. *Digital America: A tale of the haves and havemores.* Report, McKinsey Global Institute. New York, 2015. Web: http://www.mckinsey.com/industries/high-tech/our-insights/digital-america-a-tale-of-the-haves-and-have-mores.

[17] L. Mearian, "Start-up plans to release 3D-printed, street-ready cars next year", *Computer World,* Jul 9, 2015. Web: http://www.computerworld.com/article/2946022/3d-printing/start-upplans-to-release-3d-printed-street-ready-cars-next-year.html.

[18] S. Wardley, "Anticipating the Future: An Introduction to Value Chain Mapping", in *Open Source Convention (OSCON),* 2014. Web: http://www.youtube.com/watch?v=NnFelt-uaEc.

[19] S. Wardley. "Situation Normal, Everything Must Change", in *Open Source Convention (OSCON),* 2015. Web: http://www.youtube.com/watch?v=Ty6pOVEc3bA.

[20] D. Dietrich, B. Heller, B. Yang, *Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data.* USA: Wiley, 2015.

[21] K. Cukier, V. Mayer-Schoenberger. "The Rise of Big Data", *Foreign Affairs,* vol. 92, № 3, 2013. pp. 28–40.

[22] V. Mayer-Schoenberger, K.N. Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think,* Transl. Moscow: Mann, Ivanov and Ferber, 2014.

[23] B. Schmarzo. *Big Data: Understanding How Data Powers Big Business.* USA: John Wiley & Sons. Inc., 2013.

[24] B. Schmarzo. "Thinking Like a Data Scientist" (Parts I-III), in *EMC Blog In Focus.* April-June 2015. Web: https://infocus.emc.com/author/william_schmarzo/.

[25] B. Schmarzo. *Big Data MBA: Driving Business Strategies with Data Science.* USA: John Wiley & Sons. Inc., 2016.