International Journal of Pure and Applied Mathematics
Volume 119 No. 2 2018, 321-330
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: 10.12732/ijpam.v119i2.5

ijpam.eu

THE DYNAMICS OF INFORMATION REQUEST IN THE FIELD OF BIG DATA

Victor Kolesov[1], Artur Petrov[2 §]
[1,2]Tyumen Industrial University
Tyumen, RUSSIAN FEDERATION

Abstract: The concepts of dynamic characteristics of the innovation process are introduced. The structural and parametric identification of the basic models of the dynamic characteristics of the innovation process is carried out, allowing to solve the tasks of forecasting and managing the efficiency of innovative technologies. Using the example of statistics of information queries of the phrase “Big Data” in the Google search engine, the results of calculations of the dynamic characteristics of the innovation process are shown.

AMS Subject Classification: 46T10, 93A10
Key Words: dynamic process characteristics, model identification, dynamics, management, forecast

1. Introduction

The wide introduction of innovative technologies into the production practice transforms the problem of analyzing their dynamics into the category of primary ones. One of the large-scale examples is the Big Data meta-technology [1, 2, 3, 4, 5, 6, 7, 8, 9], whose progress predetermines the country's sustainable innovative development.

Big Data meta-technology, like most social and technological processes, develops over time and has a finite life cycle. Unfortunately, in the question what to consider as the basic dynamic characteristics of such processes, there is no complete clarity or clear engineering algorithms for their quantitative evalua-

Received: April 11, 2018
Revised: June 20, 201852
Published: July 4, 2018

§Correspondence author

tion. An attempt is made to remove the existing gap, at least partially.

2. Solution of the Problem

2.1. Process and its Dynamic Characteristics

First of all, it is necessary to clarify what is generally meant by "sustainable innovative development", which economists talk so much about. From the engineering point of view, as our analysis [10] has shown, sustainable innovation development is the increase in efficiency due to the growth of the efficiency by innovation, i.e. by increasing the level of creativity of developers.

Further, proceeding to a general analysis, it is necessary to formalize the very concept of "process". We consider the normalized (by the value of the function and the argument) monotonic processes in which, as the argument x grows from 0 to 1, the value of the function y also increases from 0 to 1. Usually, if Q is an increasing parameter (when Q grows with the growth of u), for the normalization of y and x, we use the interpolation formula (1):

$$u = \frac{Q - Q_{\text{min}}}{Q_{\text{max}} - Q_{\text{min}}}, \tag{1}$$

Where Q, Q_{min}, Q_{max} are the current, minimum and maximum value of the parameter, respectively. Otherwise:

$$u = \frac{Q_{\text{max}} - Q}{Q_{\text{max}} - Q_{\text{min}}}. \tag{2}$$

Thus, the normalization of the function y, depending on the nature of the parameter, takes the form:

$$y = \frac{A - A_{\text{min}}}{A_{\text{max}} - A_{\text{min}}} \text{ or, } y = \frac{A_{\text{max}} - A}{A_{\text{max}} - A_{\text{min}}} \tag{3}$$

and for the argument x, the normalization is expressed as:

$$x = \frac{B - B_{\text{min}}}{B_{\text{max}} - B_{\text{min}}} \text{ or, } x = \frac{B_{masx} - B}{B_{\text{max}} - B_{\text{min}}} \tag{4}$$

Consider the dynamics of information requests in the area of Big Data, taking into account that, firstly, it is actually a quantitative indicator of the evolution of meta-technology, and, secondly, it is clearly monitored by the Google search system [11].

The obvious goal of Big Data technology is to deliver a positive effect to the user during the life cycle. If we approach the question from the standpoint of the theory of utility, then as a function, it is expedient to choose a normalized general utility TU (5):

$$y = \frac{TU - TU_{\min}}{TU_{\max} - TU_{\min}}, \tag{5}$$

and as an argument – the normalized benefit Q or the product (6):

$$x = \frac{Q - Q_{\min}}{Q_{\max} - Q_{\min}}. \tag{6}$$

An important characteristic of the process is the rate of increase in the normalized utility $MU = dy/dx$, which has the meaning of a normalized boundary utility.

Strictly speaking, all the characteristics of the process (x, y and dy/dx) change during the life cycle T, i.e. depend on the normalized time $t_v = t/T$, therefore the essence of the problem does not change if we consider it an argument.

In dynamic analysis, as the function , the selected normalized technological indicator is considered, and the argument is $x = t_v$. Thus, the function $y(t_v)$ is the analyzed dynamic characteristic of the process. An important issue is the nomenclature of basic characteristics.

As the basic dynamic characteristics by default, we have chosen:

– rate of change in the utility of the process $V_p(t_v) = dP(t_v)/dt_v$;

– utility of the process $P(t_v) = \int\limits_0^{t_v} V_p(t_v) \cdot dt_v$;

– life force of the process $A(t_v) = P(t_v) \cdot V_p(t_v)$;

– life force reserve of the process $ZA(t_v) = \int\limits_0^{t_v} A(t_v) \cdot dt_v$.

2.2. The Mechanism of Process Characteristics Formation

Let us consider the mechanism of formation of the dynamic characteristics of the process (Fig. 1).

2.2.1. Rate of Change in the Utility of the Process (V_p)

The catalyst of the process is the dynamics of the rate of change in utility. It is postulated that the rate of increase in utility decreases with the increase in

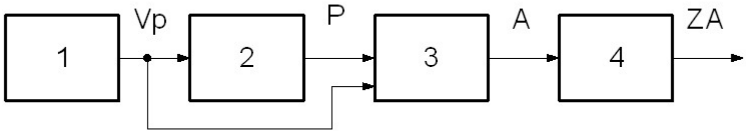

Figure 1: Functional diagram of the process

utility $P(t_v)$ in accordance with the differential equation (7):

$$\frac{dV_p}{dt_v} = -m \cdot V_p, \tag{7}$$

where m is the coefficient, generally depending on external conditions.

Separating the variables and integrating the right and left sides, we obtain $\int\limits_{V_{p\ \max}}^{V_p} \frac{dV_p}{V_p} \cdot dt = -m \cdot \int\limits_0^{t_v} dt_v$;

$\ln(V_p) - \ln(V_{p\ \max}) = -m \cdot t_v$,

i.e.

$$V_p = V_{p\ \max} \cdot \exp\left(-m \cdot t_v\right). \tag{8}$$

The normalized value of the speed is equal to

$$V_{p\ n}(t_v) = V_p / V_{p\ \max} = \exp\left(-m \cdot t_v\right) \tag{9}$$

2.2.2. Dynamics of Changes in the Utility of the Process ($\boldsymbol{P}$)

The nature of the change in the utility of process $P(t_v)$ is determined by its speed $V_p(t_v)$:

$$P(t_v) = \int\limits_0^{t_v} V_p(t_v) \cdot dt_v = \int\limits_0^{t_v} V_{p\ \max} \cdot \exp\left(-m \cdot t_v\right) \cdot dt_v =$$
$$= P_{\max}\left[1 - \exp\left(-m \cdot t_v\right)\right], \tag{10}$$

where $P_{\max} = V_{p\ \max}/m$.

We introduce the normalized utility value (11):

$$P_n(t_v) = P(t_v)/P_{\max} = 1 - \exp\left(-m \cdot t_v\right). \tag{11}$$

2.2.3. The Life Force of the Process (A)

The question of the effectiveness level of the process is crucial. But what is meant by this? On the one hand, it is desirable to have a high level of process (technology) perfection, determined by the level of overall utility $P(t_v)$, and on the other, of course, it would be appreciated to ensure high growth rates $V_p(t_v) = dP(t_v)/dt_v$.

As a criterion of the effectiveness (life force) of the process, it is suggested to take the value

$$A(t_v) = y \cdot dy/dt_v = P(t_v) \cdot V_p(t_v) = B \cdot \left[F(t_v) - F^2(t_v)\right], \qquad (12)$$

where $B = P_{\max} \cdot V_{p\ \max} \cdot m$; $F(t_v) = \exp(-m \cdot t_v)$. From an engineering point of view, the question of the extremum of the life force $A(t_v)$ is interesting. Its maximum is achieved when $dA/dt_v = 0$. Solving this equation, we take into account the relation (12), then we obtain $B \cdot [dF/dt_v - 2 \cdot F \cdot dF/dt_v] = B \cdot dF/dt_v \cdot (1 - 2 \cdot F) = 0$. When $B \cdot dF/dt_v \neq 0$, the equation is reduced to the form $1 - 2 \cdot F = 0$. Its solution $F = \exp(-m \cdot t_{v\_opt}) = 0.5$, which allows estimating the time to reach the extremum t_{v_opt}:

$$t_{v\_opt} = -\ln(0.5)/m. \qquad (13)$$

The maximum value of the life force of the process in this case is

$$A_{\max}(t_v) = B\left[F(t_{v\_opt}) - F^2(t_{v\_opt})\right], \qquad (14)$$

where $F(t_{v\_opt}) = \exp\left(-m \cdot t_{v\_opt}\right) = \exp\left(\ln(0.5)\right) = 0.5$.
Thus, $A_{\max}(t_v) = B\left[F(t_{v\_opt}) - F^2(t_{v\_opt})\right] = B \cdot \left(0.5 - 0.5^2\right) = B/4$.
Introduce the normalized value of the life force of the process

$$A_n(t_v) = \frac{A(t_v)}{A_{\max}} = \frac{F(t_v) - F^2(t_v)}{F(t_{v\_opt}) - F^2(t_{v\_opt})} = 4 \cdot \left[F(t_v) - F^2(t_v)\right]. \qquad (15)$$

2.2.4. Life Force Reserve of the Process (ZA)

Life force reserve of the process ZA is interpreted as the accumulated sum of life forces in the interval t_v

$$ZA(t_v) = \int\limits_0^1 A(t_v) \cdot dt_v = \int\limits_0^{t_v} y \cdot dy/dt_v \cdot dt_v = \int\limits_0^{y(t_v)} y \cdot dy = \frac{y(t_v)}{2}$$

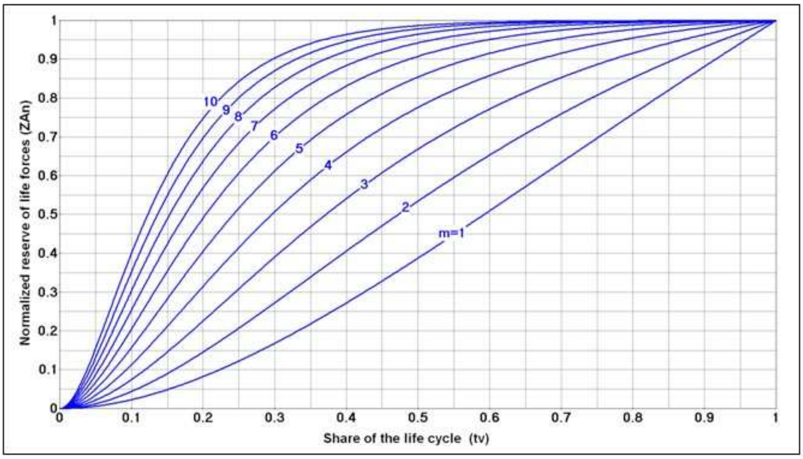

Figure 2: Normalized reserve of life forces of the process $ZA_n(t_v)$

$$= \frac{P_{\max}^2[1-\exp(-m\cdot t_v)]^2}{2} = \frac{V_{p\ \max}^2\cdot[1-\exp(-m\cdot t_v)]^2}{2m^2}. \quad (16)$$

The level of life force reserve at the end of the life cycle is

$$ZA_{\max} = \int_0^1 A(t_v)\cdot dt_v = \int_0^1 y\cdot dy/dt_v\cdot dt_v = \int_0^{y_{\max}} y\cdot dy = \frac{y_{\max}^2}{2}$$
$$= \frac{P_{\max}^2[1-\exp(-m)]^2}{2} = \frac{V_{p\ \max}^2\cdot[1-\exp(-m)]^2}{2m^2}. \quad (17)$$

The introduction of the normalization leads to the following result (Fig. 2).

$$ZA_n(t_v) = ZA(t_v)/ZA_{max} = [(1-exp(-m\cdot t_v))/(1-exp(-m))]^2 \quad (18)$$

The intensity of expenditure of life forces is predetermined, as we see, by the coefficient . All the introduced dynamic characteristics of the process are shown in Fig. 3.

2.3. Experimental Verification of the Working Hypothesis

As a proof, the results of monitoring by the Google search engine information requests in the field of Big Data were used [11]. The equivalent of a measure

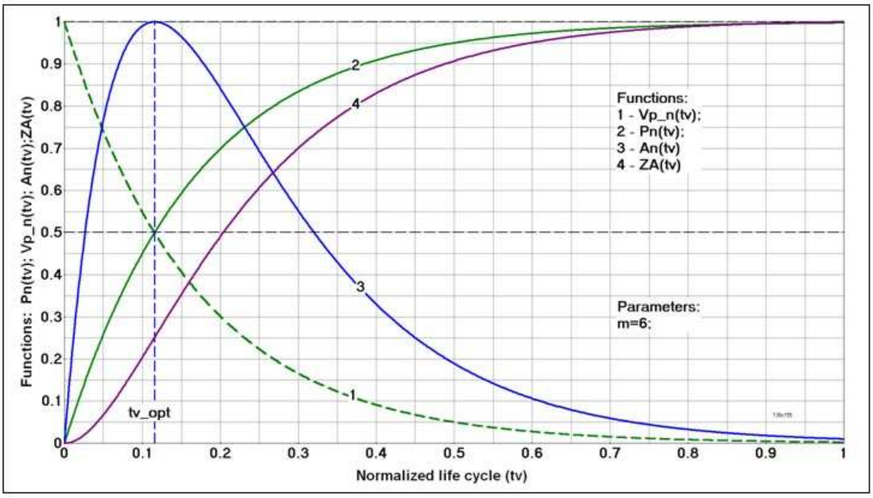

Figure 3: Dynamic characteristics of the process

of utility is the number of requests per month (the intensity of requests). The normalized monitoring data and the results of their processing are shown in Fig. 4. The reference point corresponds to 01/01/2004. The rate of data registration is 1 month. The result of the processing are two models: Y_1 and Y_2. The first of them corresponds to the accepted hypothesis (11), and the second - to the maximum adequacy of the analyzed process.

Both models almost coincide in the dynamic area of the process ($100 \leq t \leq 150$),but the second model has a clearly defined "background" at 0.05. The analysis of Fig. 4 indicates that the basic model (11) does not contradict the experiment and, therefore, reasonably interprets the mechanism of the evolutionary process.

As for the structural and parametric identification of the life-force model of process A(t), it was realized on the basis of an analysis of the Google system data for the USA and Russia [11] (Fig. 5)

Model A_1(t) corresponds to the USA, model A_2(t) – to Russia. Both models have a structure corresponding to the relation (12), but with different life cycles. Their analysis shows that the basic model (12) is consistent with the experimental data.

Practical interest is the trend of coefficient $K_{bd}(t) = A_1(t)/A_2\ (t)$, which characterizes the multiplicity of losses in the life force of the process in Russia relative to the United States.

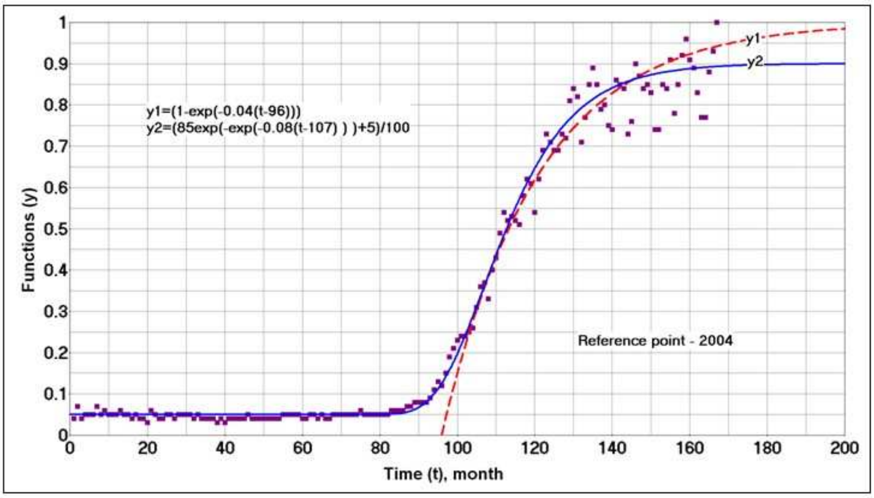

Figure 4: Time trend of the process

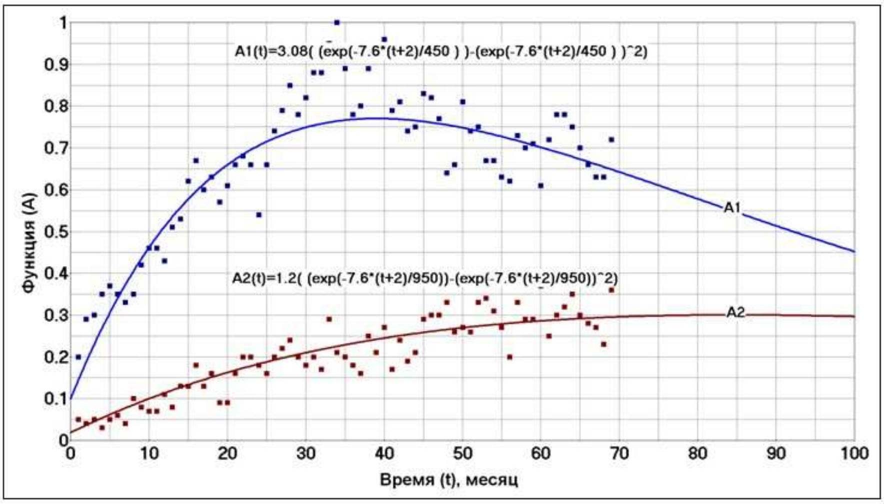

Figure 5: The life force of the process

As we can see, in this indicator, Russia should catch up with the United States in the near future.

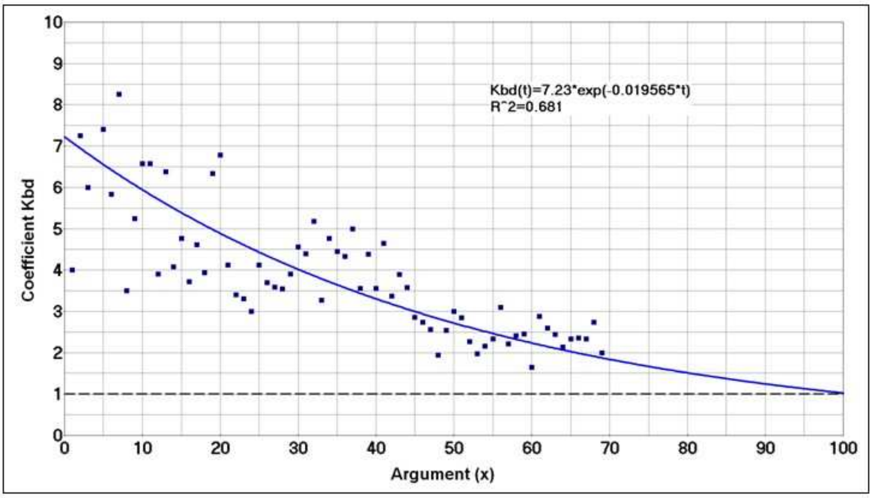

Figure 6: Coefficient trend Kbd

3. Conclusions

The results obtained have a number of promising engineering applications [12]. In particular, they are proposed to be used in solving the following tasks:

– structural and parametric identification of models of social and technological processes' effectiveness;

– forecast of the dynamics of the processes analyzed;

– development and examination of programs on increasing the effectiveness of scientific and technological projects;

– development of algorithmic support for tools for analyzing the system dynamics of innovation processes.

References

[1] I.L. Avdeeva, *Analysis of foreign experience using global technologies "Big Data"*, URL: http://naukovedenie.ru/PDF/13EVN616.pdf

[2] S. Artemov, *Big Data: new opportunities for growing business*, URL: http://www.pcweek.ru/upload/iblock/d05/jet-big-data.pdf

[3] A. Afanasiev, *Analytical review of the market Big Data*, URL: https://habrahabr.ru/company/moex/blog/256747/

[4] S.A. Izmalkova, T.A. Golovin, Using global technology "BIG DATA" in the management of economic systems, *Izvestiya TulGU. Economic and Legal Science*, 4-1 (2015), 151-158.

[5] V. Meyer-Shenberger, K. Kukier, *Big Data. A revolution that will change the way we live, work and think*, Mann, Ivanov and Ferber, Moscow (2014).

[6] V.V. Sokolyansky, B.S. Pashkov, Technology BIG DATA and their installation in the economic studies, LLC "Sputnik +", Moscow (2015), pp. 169-171.

[7] L. Chernyak, Big data as a new theory and practice, Open systems. *The Database Management System*, 10 (2011),18.

[8] K. Cukier, Data, data everywhere: A special report on managing information,
URL: http://www.economist.com/node/15557443

[9] M. Hilbert, How to measure the world's technological capacity to communicate, store and compute information? *International Journal of Communication*,
URL: http://www.ijoc.org/ojs/index.php/ijoc/article/viewFile/1562/742

[10] V.I. Kolesov, Dominant technology in terms of sustainable innovation development of the region, Energysaving and innovative technologies in the fuel and energy complex: Materials of the regional scientifically-practical conference. TSOGU, Tyumen (2012), pp. 125-128.

[11] Google Trends. Analysis,
URL: https://trends.google.ru/ trends/explore?date=all&q=Big%20Data

[12] V.I. Kolesov, Model of the dynamics of transport and social risks in the country, New information technologies in the oil and gas industry and educational research institute: Materials of the VI all-Russian scientific-technical conference with international participation, TSOGU, Tyumen (2015), pp. 68-74.