ITHEA

International Journal INFORMATION THEORIES & APPLICATIONS

2013 Volume 20 Number 1

International Journal
INFORMATION THEORIES & APPLICATIONS
Volume 20 / 2013, Number 1

International Journal "INFORMATION THEORIES & APPLICATIONS" (IJ ITA) is official publisher of the scientific papers of the members of the ITHEA International Scientific Society

IJ ITA welcomes scientific papers connected with any information theory or its application.
IJ ITA rules for preparing the manuscripts are compulsory.
The **rules for the papers** for IJ ITA are given on *www.ithea.org*.

Responsibility for papers published in IJ ITA belongs to authors.

General Sponsor of IJ ITA is the **Consortium FOI Bulgaria** (*www.foibg.com*).

International Journal "INFORMATION THEORIES & APPLICATIONS" Vol. 20, Number 1, 2013

Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with:
Institute of Mathematics and Informatics, BAS, Bulgaria,
V.M.Glushkov Institute of Cybernetics of NAS, Ukraine,
Universidad Politécnica de Madrid, Spain,
Hasselt University, Belgium,
St. Petersburg Institute of Informatics, RAS, Russia,
Institute for Informatics and Automation Problems, NAS of the Republic of Armenia

Printed in Bulgaria
Publisher ITHEA®
Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, e-mail: info@foibg.com
Technical editor: **Ina Markova**

ISSN 1310-0513 (printed) **ISSN 1313-0463 (online)** **ISSN 1313-0498 (CD/DVD)**

XX-th anniversary of IJ ITA !

PREFACE

Verba volant, scripta manent !

The "International Journal on Information Theories and Applications" (IJ ITA) has been established in 1993 as independent scientific printed and electronic media. IJ ITA is edited by the Institute of Information Theories and Applications FOI ITHEA in collaboration with the leading researchers from the Institute of Cybernetics "V.M.Glushkov", NASU (Ukraine), Universidad Politécnica de Madrid (Spain), Hasselt University (Belgium), St. Petersburg Institute of Informatics, RAS (Russia), Institute for Informatics and Automation Problems, NAS of the Republic of Armenia, and Institute of Mathematics and Informatics, BAS (Bulgaria).

During the 20 years, IJ ITA became as well-known international journal. Till now, excluding this volume, more than 820 papers have been published. IJ ITA authors are widespread in 43 countries all over the world: Armenia, Austria, Belarus, Brazil, Belgium, Bulgaria, Canada, Czech Republic, Denmark, Egypt, Estonia, Finland, France, Germany, Greece, Hungary, India, Iran, Ireland, Israel, Italy, Japan, Kirghizia, Latvia, Lithuania, Malta, Mexico, Moldova, Netherlands, Poland, Portugal, Romania, Russia, Scotland, Senegal, Serbia and Montenegro, Spain, Sultanate of Oman, Switzerland, Turkey, UK, Ukraine, and USA.

The great success of IJ ITA belongs to the whole of the ITHEA International Scientific Society. We express our thanks to all authors, editors and collaborators who had developed and supported the International Journal on Information Theories and Applications.

More information about the IJ ITA rules for preparing and submitting the papers as well as how to take out a subscription to the Journal may be obtained from www.ithea.org.

Krassimir Markov

IJ ITA Founder and Editor in chief

USEFULNESS OF SCIENTIFIC CONTRIBUTIONS

Krassimir Markov, Krassimira Ivanova, Vitalii Velychko

Abstract: *The prevailing role of counting citations over the added scientific value evaluating distorts the scientific society. As result, the scientific work becomes a kind of business, for instance, to obtain as more citations as possible. It is important to counterbalance the role of counting citations by using additional qualitative criteria. The aim of this survey is to discuss an approach based on measure of "usefulness of scientific contribution" called "usc-index" and published in [Markov et al, 2013]. It is grounded on theory of Knowledge Market. In accordance with this, we remember main elements of this theory. After that we recall some information about Bibliometrics, Scientometrics, Informetrics and Webometrics as well as some critical analyses of journals' metrics and quantity measures. Finally, we outline the approach for evaluation usefulness of the scientific contributions.*

Keywords*: Information Market, Knowledge Market, Usefulness of the Scientific Contributions*

ACM Classification Keywords*: A.1 Introductory and Survey*

Introduction

The main goal of this paper is to continue the investigation of Knowledge Markets started in [Ivanova et al, 2001; Markov et al, 2002; Markov et al., 2006; Ivanova et al, 2006].

Now, our attention will be paid to the *Usefulness of the Scientific Contributions (USC)*.

What is "scientific contribution"? May be the most popular understanding is:

(1) The *added scientific value* of the published researcher's results;

(2) Its impact on obtaining new scientific results registered by corresponded *citations*.

It is very difficult to measure the added scientific value.

Because of this, in recent years, it became very popular to measure the second part – the citations.

There are a number of ways to analyze the impact of publications of a particular researcher. A longtime favorite has been ISI's (Social) Science Citation Index, which has come to the web as Web of Science. The web has introduced a number of other tools for assessing the impact of a specific researcher or publication. Some of these are Google Scholar, Scopus, SciFinder Scholar, and MathSciNet among many others. In addition, Publish or Perish uses data from Google Scholar, but it automatically does analysis on the citation patterns for specific authors. After searching for an author one can select the papers to analyze and to get metrics such as total citations, cites per year, h-index, g-index, etc. [Peper, 2009]. In the same time, a negative tendency appears.

The prevailing role of counting citations over the added value evaluating distorts the scientific society.

As result, the scientific work becomes *a kind of business*, for instance, to obtain as more citations as possible.

For examples see [Harzing, 2012].

It is important to counterbalance the role of counting citations by using additional qualitative criteria [DORA, 2012; ISE, 2012].

In an early work (1964) Garfield suggested 15 different reasons for why authors cite other publications (reprinted in [Garfield, 1977]). Among these were: paying homage to pioneers; giving credit for related work; identifying methodology; providing background reading; correcting a work; criticizing previous work; substantiating claims; alerts to a forthcoming work; providing leads to poorly disseminated work; authenticating data and classes of fact – physical constants, etc.; identifying original publications in which an idea or concept was discussed; identifying original publication or other work describing an eponymic concept; disclaiming works of others and disputing priority claims.

Similarly, the textual function of citations may be very different. In a scientific article some of the references will represent works that are crucial or significant antecedents to the present work; others may represent more general background literature. For example, reviewing the literature published on this topic during 1965–1980, Henry Small identified *five distinctions*: a cited work may be

1) *Refuted;*
2) *Noted only;*
3) *Reviewed;*
4) *Applied;*
5) *Supported by the citing work*.

These categories were respectively characterized as [Small, 1982]:

1) ***Negative***;
2) ***Perfunctory***;
3) ***Compared***;
4) ***Used***;
5) ***Substantiated***.

Thus, the different functions that citations may have in a text are much more complex than merely providing documentation and support for particular statements [Aksnes, 2005].

The aim of this survey is to discuss an approach for evaluating the "usefulness of scientific contribution" called "***usc-methodology***" [Markov et al, 2013]. It is grounded on theory of Knowledge Market. In accordance with this, the next chapter remembers main elements of this theory. After that we recall some information about Bibliometrics, Scientometrics, Informetrics and Webometrics as well as some critical analyses of journals' metrics and quantity measures. Finally, we outline the approach for evaluation usefulness of scientific contributions. In more details, the chapters of the paper concern:

- Basic concepts of Knowledge Markets' Theory;
- Structure of the Knowledge Market;
- Science, Publishing, and Knowledge Market;
- National and International Knowledge Markets;
- Bibliometrics, Scientometrics, Informetrics and Webometrics;
- Citation tracking and Evaluation of Research;
- Journal metrics;
- Quantity measures;
- Disadvantages of journal metrics and quantitative measures;
- Evaluation of Scientific Contributions;

Basic concepts of Knowledge Markets’ Theory

Information society

At the stage of social growth, called “information society”, the information and information activities get decisive value for existence of the separate individuals or social teams. Certainly, at earlier stages of development of mankind, the information had important value too. But never, in all known history, other means for existence have been so dominated by the information means as it is in the information society [Markov et al., 2006].

From the origin, human society has been "information" one, but levels of information service differ in different periods of existence of societies. It is possible to allocate following levels of information society:

- *Primitive* (people having knowledge, letters on stones etc.);
- *Paper based* (books, libraries, post pigeons, usual mail etc.);
- *Technological* (telephone, telegraph, radio, TV, audio- and video-libraries etc.);
- *High-Technological* (computer systems of information service, local information networks etc.);
- *Global* (global systems for information service, opportunity for everybody to use the information service with help of some global network etc.).

The information society does not assume compulsory usage of the information services by a part or all inhabitants of given territory. One very important feature thus is emphasized: for everyone will be necessary diverse and qualitative (from his point of view) information, but also everyone cannot receive all necessary information. *The enterprising experts will accumulate certain kinds of information* and will provide existence through favorable to them information exchange with the members of the society. Thus, in one or other form, they will carry out **payable information service (carrying out information services for some income)** [Ivanova et al, 2001]. This is the background of ***Information Market***.

Knowledge Information Objects

The usual understanding of the verb "**to know**" is: "*to have in the mind as the result of experience or of being informed, or because one has learned*"; "to *have personal experience of something*” etc. The concept "**knowledge**" usually is connected to concepts "understanding" and "familiarity gained by experience; range of information" [Hornby et al, 1987] or "organized body of information" [Hawkins, 1982].

V.P. Gladun correctly remarks that the concept “*knowledge*” does not have common meaning, especially after beginning of it’s using in technical lexicon in 70-ies years of the last century. Usually, when we talk about the human knowledge we envisage all information one has in his mind.

Another understanding sets the “knowledge” against the “data”. We talk about data when we are solving any problem or are making logical inference. Usually the concrete values of given quantities are used both as data and descriptions of objects and interconnections between objects, situations, events, etc.

During decision making or logical inference we operate with data involving some other information like descriptions of the solving methods, rules for inference of corollaries, models of actions from which the decision plan is formed, strategies for creating decision plans, and general characteristics of objects, situations, and events.

In accordance with this understanding, the “knowledge” is information about processes of decision making, logical inference, regularities, etc., which, applied to the data, creates any new information [Gladun, 1994].

The knowledge is a structured or organized body of information models, i.e. the knowledge is information model, which concerns a set of information models and interconnections between them.

Let remember, in general, the information model is a set of reflections, which are structured by Subject and, from his point of view, represents any entity [Markov et al, 2001].

The information objects, which contain information models, are called ***"knowledge information objects"***.

Knowledge Market

The growth of societies shows that the knowledge information objects become important and necessary articles of trade. The open social environment and market attitudes of society lead to arising of *knowledge customers* and *knowledge sellers*, which step-by-step form "Knowledge Markets" [Markov et al, 2002].

As the other markets, the ***Knowledge Market*** *is organized aggregate of participants, who operate following common rules and principles*. The knowledge market structure is formed by a combination of mutually-connected elements with simultaneously shared joint resources.

Staple commodities of knowledge market are knowledge information objects.

The knowledge information bases and tools for processing the knowledge information objects, such as tools for collecting, storing, distributing, etc., form the knowledge market environment. The network information technologies enable to construct uniform *global knowledge market environment*. It is very important, it to be friendly for all knowledge market participants and open for all layers of the population without dependence from a nationality, social status, language of dialogue, place of residing. The decision of this task becomes a crucial step of humanization of all world commonwealths.

In the global information society, on the basis of modern electronics, the construction of the global knowledge market, adapted to the purposes, tasks and individual needs of the knowledge market participants is quite feasible, but the achievement of this purpose is connected to the decision of a number of scientific, organizational and financial problems. For instance, the usual talk is that *at the Knowledge Market one can buy knowledge*. But, from our point of view, *this is not so correct*.

In global information society, the e-commerce becomes fundamental for the Knowledge Market. The advantages of e-commerce are obvious. In the same time there exist many risks for beginners at this kind of market. From this point of view, the society needs to provide many tasks for training the citizens to use properly opportunities of the new environment [Markov, 1999]. Let consider an example.

When an architect develops any constructive plan for future building, he creates a concrete "*information object*". Of course, he will sell this plan. This is a transaction in area of the *Information Market*.

Another question is: from where does architect have received the skills to prepare such plans? It is easy to answer – he has studied hardly for many years and received knowledge is the base for his business. Textbooks and scientific articles are not concrete information for building concrete house, but they contain the knowledge needed for creating such plans.

The scientific books and papers written by the researchers (lecturers) in the architectural academy are special kind of "information objects" which contain special generalized information models. They are "*knowledge information objects*" which have been sold to students and architects.

Here we have a kind of transactions at the "*Knowledge Market*".

We have to take into consideration the *difference between responsibility* of architect and lecturer (researcher).

If the building collapses, the first who will be responsible is architect, *but never lecturer!*

In beginning of the XX-th century, the great Bulgarian poet Pencho Slaveikov wrote:

"The speaker doesn't deliver his thought to listener, but his sounds and performances provoke thought of the listener. Between them, a process performs like lighting the candle, where the flame of the first candle is not transmitted to another flame, but only cause it."

If one buys a candle what does he really buy – "wax" or "light" of candle? The light is not for sale in the store… But one really may see the example how the candle works and how it may be used. Based on this he/she may decide whether to buy the candle or not.

We came to the main problem we need to point – t*he authors and publishers are not responsible for what they sold to the customers*. Pros and Cons of (electronic) Publishing are discussed many times (see for instance [NLC, 2004]). From customers' point of view, it is difficult to discover what really we will receive if we will buy one (electronic) publication. The title and announcement of the publications are not their content. The customers could not claim damage if the content is not what it is needed. To regulate this process we need specialized rules and standards for knowledge markets as well as corresponded laws for *authors' and publishers' responsibility*.

The scientific work usually is reported as series of publications in scientific journals. The practice is to delegate social rights to editors and reviewers to evaluate the quality of reported results.

And here we see serious problem – *is their evaluation enough? Of course,* ***it isn't!***

Because of this, counting of citations became important. But, the citations may be of different types including negative ones. We need methodology for evaluating *Usefulness of the Scientific Contributions (USC).*

Structure of the Knowledge Market

The Structure of the Knowledge Market was presented in [Markov et al, 2002]. The updated scheme of the basic structure of Knowledge Market is outlined on Figure 1 below.

Let's remember basic elements of the knowledge market.

Employer (Er) is the initial component of the Knowledge Market whose investments support providing the scientific research. The concept of Employer means men or enterprise, which need to buy manpower for the purposes of the given business. A special case is the government of the state which may be assumed as representative of the society as Employer. In addition, different scientific or not scientific foundations, social organizations, etc., may invest in scientific activities and this way to become Employers.

The concept of the ***Employee (Ee)*** means a man who is already taken as a worker in the given business or is potentially to be taken in it. The main interest of the employee is to sell his received knowledge and skills. The main goal of the Employee is to receive maximal financial or other effects from already received knowledge and skills. This means that the Employee is not internally motivated to extend them if this knowledge and skills are enough for chosen work activity. From other point of view the Employee motivation closely depends to future expectations for his social status. The Employee became as converter of the learned knowledge and skills into real results of his workplace. Let remark, that scientific organizations, institutes, groups, etc. may be employed to fulfill some scientific projects and to be in the role of Employee at the KM.

In other words, *Employer* hires *Employees*. During the work processes, the knowledge and skills of Employees are transformed in real products or services. This process is served by the Manpower Market. Employees, even owning a high education level, need additional knowledge to solve new tasks of the Employers. Still, they are **customers of new knowledge**, who arouse necessity of the Knowledge Market, which should rapidly react to the customers' requests. In other words, the Manpower's Market causes activity of the Knowledge Market (KM). These two members of KM are main its components – *the knowledge customers*.

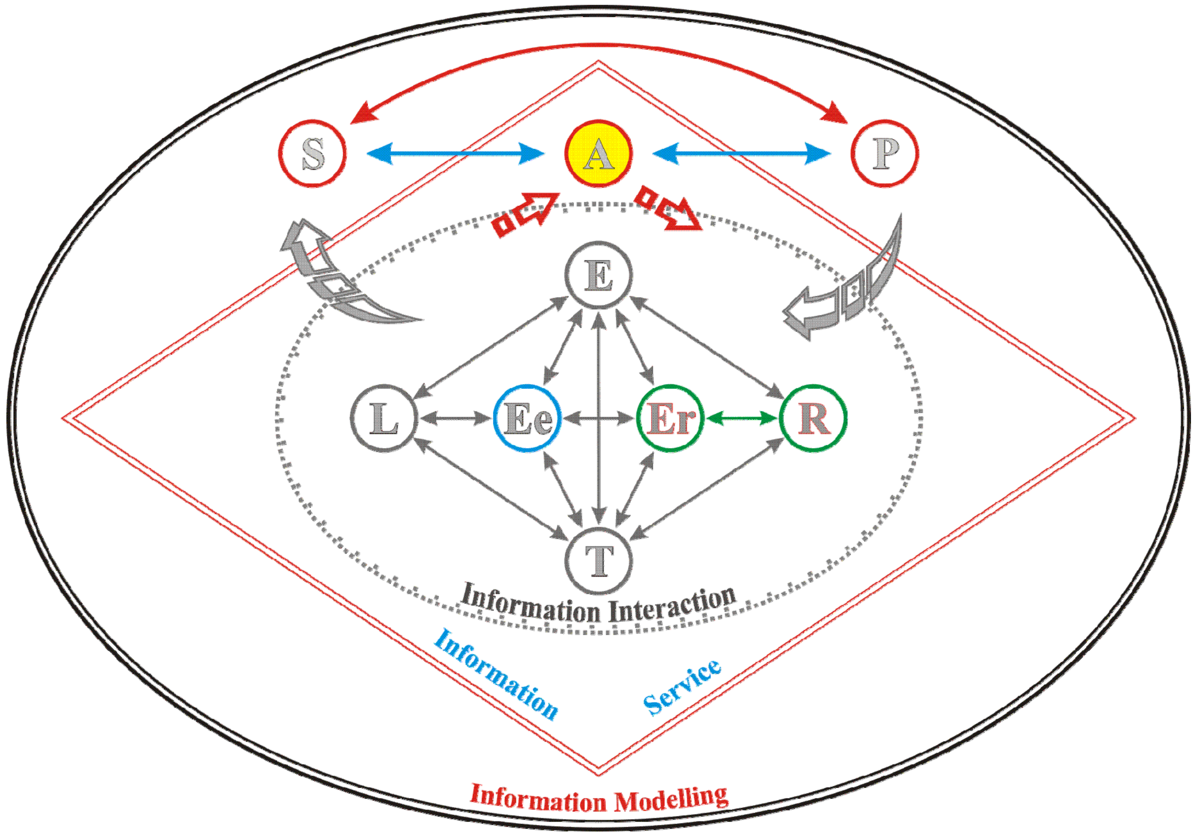


Figure 1. Structure of the Knowledge Market

It is clear that the business needs the high-skilled workers. The employer buys the final result of the cycle in the Knowledge Market - the educated and skilled workers. The continuous changing of technological and social status of the society leads to appearance of new category – industrial **Researchers (R)** – peoples/organizations, who have two main tasks:

- To invent and/or promote new technologies to Employers in convenient way to implement them in practice;
- To determine the educational methods for training the staff for using the new technologies.

The educational process is carried out by the **Lecturers (L)**, who transforms new scientific knowledge into pedagogical grounded lessons and exercises. During realizing concrete educational process, Lecturers are assisted by **Tutors (T)** who organize the educational process and supports the Employees to receive the new knowledge and to master theirs skills. At the end of the educational process, a new participant of KM appears – **Examiners (E)** – who test results of education and answer to the question "have the necessary knowledge and skills been received".

These six components of the Knowledge Market, which contact each other via global information network, form the first knowledge market level called **“*information interaction*”**. As far as these components are too much and distributed in the world space, the organization and co-ordination of theirs information interaction needs adequate **“*information service*”**. It is provided by a new component called **Administrators (A)**. Usually the Administrators are Internet and/or Intranet providers or organizations. They ***collect, advertize and sell knowledge objects, sometimes without understanding what really they content.***

The rising activity of knowledge market creates need of developing new general or specific knowledge as well as modern tools for the information service in frame of the global information network. This causes the appearance of high knowledge market level, which allows observing processes, as well as inventing, developing and implementing new knowledge and corresponded systems for information service. This is the **“*information modeling*”** level. It consists of two important components – the academic researchers called here **Scientists (S)** and the **Publishers (P)**. In this paper we will discuss more deeply characteristics and activities of both of them.

Of course, the Knowledge Market as a kind of Market follows rules and laws given by social environment. The interrelation between government, social structures, and Knowledge Market need to be studied in separate investigation. In several papers we have already investigate different problems of the Knowledge Market [Ivanova et al, 2001; Markov et al, 2002; Ivanova et al, 2003; Markov et al, 2003].

For years we have seen that the Knowledge Market is very important for growth of science and in the same time it is important scientific area and need to be investigated.

Science, Publishing, and Knowledge Market

Preparing this survey, we have collected more than hundred definitions of terms “*science*” and “*scientific methodology*”. Analyzing them we chose the one of the Britain's Science Council, which has spent a year working out a new definition of the word “science”. The Science Council is a membership organization that brings together learned societies and professional bodies across science and its applications. It was established under Royal Charter in October 2003 and was registered as a charity with the Charity Commission in September 2009. The principal activity of Science Council is to promote advancement and dissemination of knowledge and education in science, pure and applied, for public benefit [BSC, 2013].

The Science Council definition focuses on the pursuit of knowledge rather than established knowledge. It may be the first "official definition of science" ever published. Here's what they've come up with:

> ***"Science is the pursuit of knowledge and understanding of the natural and social world following a systematic methodology based on evidence"*** [BSC, 2013].

It defines science as a pursuit, an activity, related to the creation of new knowledge, rather than established knowledge itself. Science is seen as a species of research.

Scientific methodology includes the following [BSC, 2013]:

- Objective observation: measurement and data (possibly although not necessarily using mathematics as a tool);
- Evidence;
- Experiment and/or observation as benchmarks for testing hypotheses;
- Induction: reasoning to establish general rules or conclusions drawn from facts or examples;
- Repetition;
- Critical analysis;
- Verification and testing: critical exposure to scrutiny, peer review and assessment.

The last point is closely connected to publishing activities which are the main way to provide critical exposure to scrutiny, peer review and assessment. In addition, previous published research results have to be taken in account and current results have to be compared and evaluated in accordance to them.

Due to very great number of results to be published, *scientific publishing activities became an industrial branch*. Nowadays, the scientific publishing companies (Publishers "**P**" on Figure 1) compete with others at the knowledge markets in two main areas:

- Collecting original scientific results to be published;
- Market shares where the publications may be sold.

The basic difference between knowledge markets and other kinds of markets consists in the following.

To publish the results of their research is an obligation that professional scientists are compelled to fulfill [Merton, 1957b]. New knowledge, updated by researchers, has to be transformed into information made available to the scientific community. Not only do scientists have to make their work available to the public at large, but they in turn are supposed to have access to the work of their peers. Research is carried out in a context of "*exchange*". Even so, the fact that the system of scientific publication has survived in modern science is due, paradoxically, to scientists' desire to protect their intellectual property. New scientific knowledge is a researcher's personal creation, and claim to its discovery can be laid only through publication [Merton, 1957a].

The "reward system", based on the recognition of work, merely underscores the importance of publication: the only way to spread the results of research throughout the world is to have them published. Publication therefore has three objectives: *to spread scientific findings, protect intellectual property and gain fame* [Okubo, 1997].

The academic researchers (Scientists "**S**" on Figure 1) who produce the new knowledge (presented by knowledge objects to be published) are, in the same time, ***main clients***. In other words, *the source and target groups partially coincide* but they are distributed all over the world. Because of this, information about the published results is accumulated by knowledge market organizers (Administrators "**A**" on Figure 1) who, using special kinds of data bases, serve the interactions between scientists and publishers as well as between both of them and the rest participants of the knowledge markets.

Due to serious *competition between publishers*, the administrators play an extra role – to *range* those using different criteria and this way *to control the knowledge objects' flows*. This is a play for billions of Dollars, Euros, etc. Let see an example from our practice.

We were invited to write a chapter in a scientific monograph to be published by a leading scientific publishing company [Markov et al, 2013a]. The book was published and it became as a staple commodity at the knowledge market. Depending of the format, its price varies between $195 and $390 [Naidenova & Ignatov, 2013]. We were glad to understand that our chapter was evaluated as a good one to be included in an encyclopedic four volumes comprehensive collection of research on the latest advancements and developments [Markov et al, 2013b]. Again, depending of format, the price of the collection varies between $2050 and $4100 [AIRM, 2013].

Let see what income will be received if we assume that the editions have only 250 exemplars and if the editions have 1000 exemplars sold.

In the case with 250 exemplars sold, the income is:

- min: 195x250 + 2050x250 = 48750 + 512500 = 561250 USD;
- max: 390x250 + 4100x250 = 97500 + 1025000 = 1122500 USD.

In the case with 1000 exemplars sold, the income is:

- min: 195x1000 + 2050x1000 = 195000 + 2050000 = 2245000 USD;
- max: 390x1000 + 4100x1000 = 390000 + 4100000 = 4490000 USD.

Concluding this hypothetical accounting we may say that expected income may vary between **500 thousands** and **4.5 millions** of Dollars. Because of this, it is very important to be a "leading" publisher who publishes new and useful results which can be sold. *Unfortunately our income from these editions was **0 (zero) cents**.*

National and International Knowledge Markets

One may remark that for our scientific work we had received salaries, society spend resources for supporting our research via buildings, service workers, etc. Yes, it is truth. But let analyze the situation according the scheme on Figure 1. Two variants of knowledge markets are shown on Figure 2 and Figure 3. The first one is "*national*" KM and second – "*international*" KM. Let analyze them step by step.

The ***National knowledge market*** (Figure 2) is included in the clear boundaries and all processes are connected.

1. The society, via government subsidies and/or concrete national projects, provides financial and organizational support of the scientists and their work.
2. The received results are published and indexed again on the base of financial and organizational support of government subsidies and concrete national projects.
3. Selling the results as printed publications and implementations in practical realizations as well as via the tax mechanism, the society receives some income which in some degree covers the initial expenses.

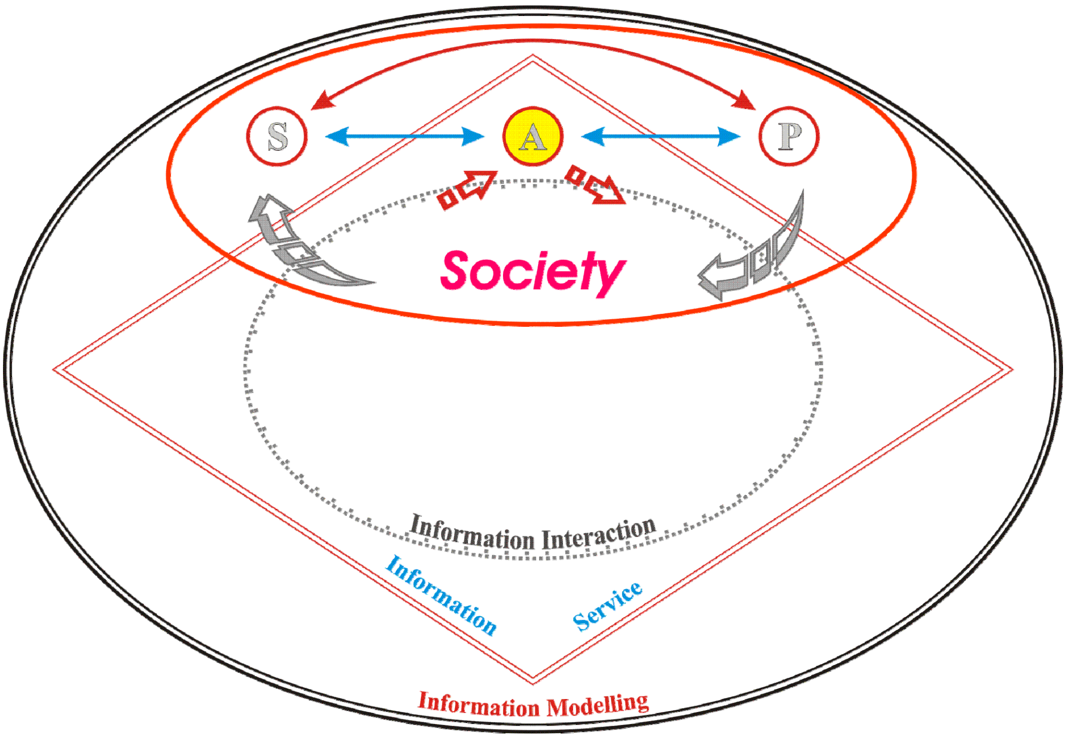


Figure 2. National Knowledge Market

The ***International knowledge market*** (Figure 3) is distributed in the boundaries of separated societies and all processes are financially disconnected.

1. The Society 1, via government subsidies and/or concrete national or international projects, provides financial and organizational support of the scientists and their work.
2. The received results are published in Society 3 and indexed in Society 2 on the base of financial and organizational support of government subsidies and concrete national or international projects.
3. Selling the results as printed publications and implementations in practical realizations as well as via the tax mechanism, the Society 3 receives some income which covers its initial expenses and realizes some profit.

4. Selling informational services based on indexed publications, Society 2 covers its initial expenses and realizes some profit.
5. Only Society 1 has *no profit* but some losses because it spends resources for supporting its scientists but the *surplus value* of their work is accumulated in Society 2 and Society 3.
6. Finally, Society 1 became poor and slowly perishes, but Society 2 and Society 3 became rich and grow.

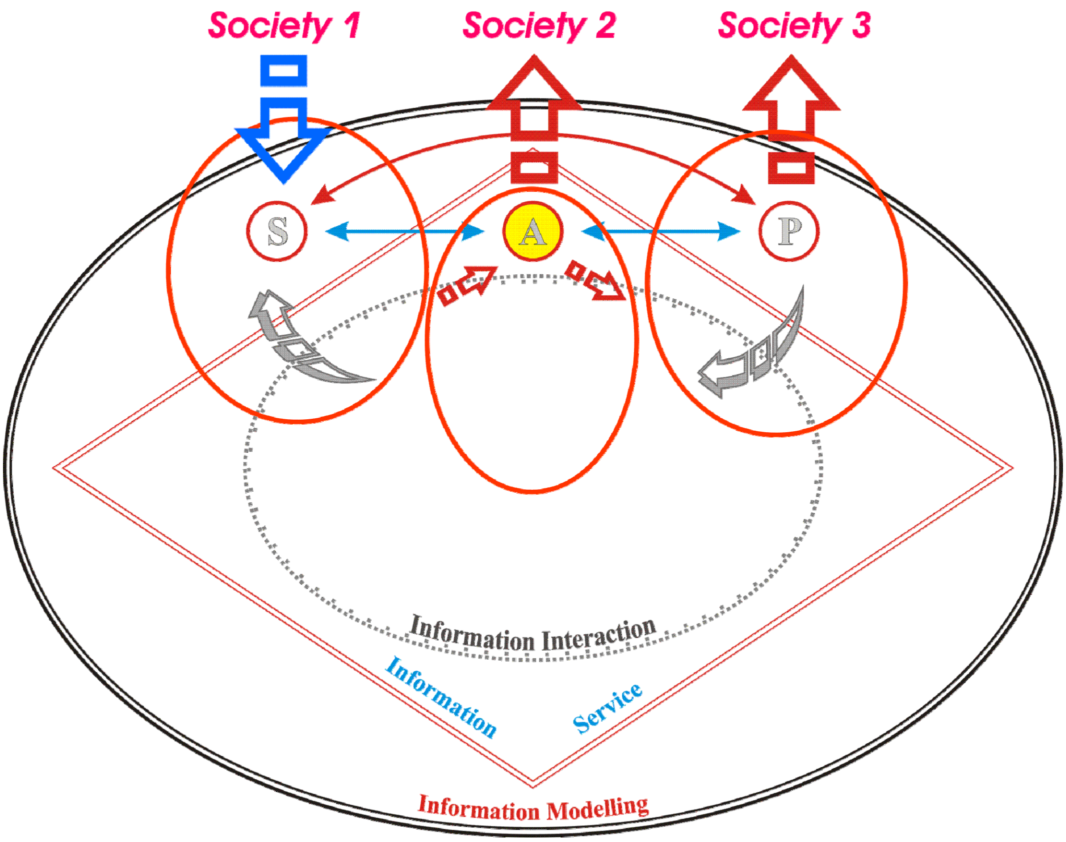


Figure 3. International Knowledge Market

It is important to comment the role of ***international scientific projects***. They give some financial support to the Society 1 but in the same time they orientate scientists towards interests of sponsoring society, usually it is Society 2 or Society 3, both two societies together or one and the same society which plays both roles. As result, the national knowledge market of Society 1 will be *destroyed* and its rebuilding becomes impossible. In opposite, the national knowledge markets of other societies will grow.

Now the main question is "*How to influence to the Society 1 to participate in such unequal battle?*"

The answer is: *By using the power of*

- *Developed national knowledge markets;*
- *Advertising, mainly indirect.*

The best influence is ***the developed national knowledge market*** with participants who are high level specialists in their area. This generates the willingness to join, to be part of them. As more people are involved so great is the influence to other societies. Opening the national knowledge market is very important step. Possibility to be published on such authoritative level is a possible dream. And the result is total influence. In addition, opening the manpower market for specialists from abroad make this dream reality and many scientists start working following the rules of this national knowledge market to ensure possibility for immigration. Finally, they influence on developing the own national knowledge markets to be organized in the same manner and rules as of the prototype one *without taking in account the national specifics and interests*.

The ***advertizing*** (mainly – indirect) of developed national knowledge markets increase their influence. *Advertising* was originated from a Latin term — "*advertire*", which means — "***to turn to***". The American Marketing Association (AMA) has defined *Advertising* as — the placement of announcements and persuasive messages in time or space purchased in any of the mass media by business firms, nonprofit organizations, government agencies, and individuals who seek to inform and/or persuade members of a particular target market or audience about their products, services, organizations, or ideas [AMA, 2013].

Indirect advertising is a form of marketing that does not use the formal everyday methods such as newspapers and magazines. This type of advertising uses: a product in a television show; giving a product away for free; sponsoring of events or activities (= paying for them); etc. [Jeeves, 2013; CBED, 2013].

*"**Audience reach measures**" have been used to determine how many people see the advertising and how often.* Measurement systems exist across the globe that determine how many people in total read certain magazines and newspapers, watch TV programs, listen to radio stations, etc.

For instance, in the US, Roy Morgan Single Source shows that, in year 2005, television is still the most widely used medium (see Figure 4). However, magazines, as a group, reach as many people as 'free to air' TV, and more people than newspapers or the Internet. Of course, specific magazines or genres of magazines often outperform specific television 'shows' [Levine et al, 2005].

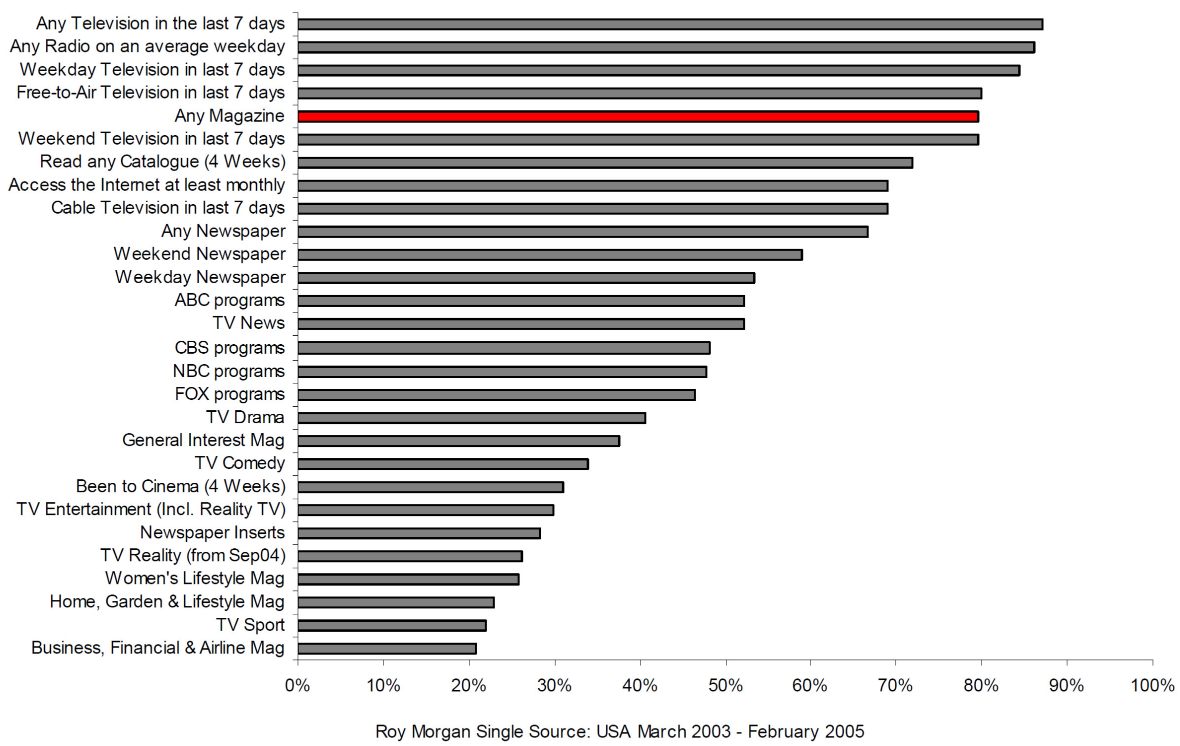


Figure 4. Media Usage in USA for year 2005

One of the movements happening on the internet is that of indirect marketing and advertising. Publishers and manufactures are catching on to what customers want, which is proof that they must invest having a business. Indirect advertising and marketing is often a technique to obtain this, as in most circumstances it supplies something of worth upfront *for totally free*. You are going to see this with no cost eBooks, blogs, and videos all dedicated to helping the visitor.

If the content delivers enough enable, the visitor may just check out the rest of the site and sign up for membership region or buy their premium book. Indirect marketing makes use of a funnel pointing toward the location where the business can make money. Another instance is often observed with no cost apps tied to movies. By downloading the app, you might just want to go see or obtain the movie [EzineMark, 2013].

In order to determine how to create an effective advertising campaign decision makers in the industry use a range of ***measures*** to try to predict the outcome of the campaign. Those who make decisions each year about where to place billions of dollars in advertising have focused in the past primarily on audience or “opportunity-to-see” measures – the task being to create chance that target audience will see advertisement with assumption that everything else will run its course.

Bibliometrics, Scientometrics, Informetrics and Webometrics

The advertisers need to know their audience and ***to measure results achieved – shifts in sales or shifts in attitude*** among the intended audience. Today all marketing and advertising people are judged by the overall performance of their company, each quarter of every year. Research and information is not a substitute for ingenuity. But ignoring intelligent and reliable research and information altogether is a luxury nobody can afford! [Levine et al, 2005]. At the knowledge markets there are two main kinds of indirect advertizing:

- *Ranging selected journals* and this way to raise the income of publishers of these journals and Society 3;
- *Counting citations and computing scientific indexes* based only on digital libraries of collected papers from selected journals and this way to raise income of administrators of these libraries and Society 2.

Measuring science has become an “industry”. Governments and their statistical offices have conducted regular surveys of resources devoted to research and development (R&D) since the 1950s. A new science had raised – *Scientometrics*.

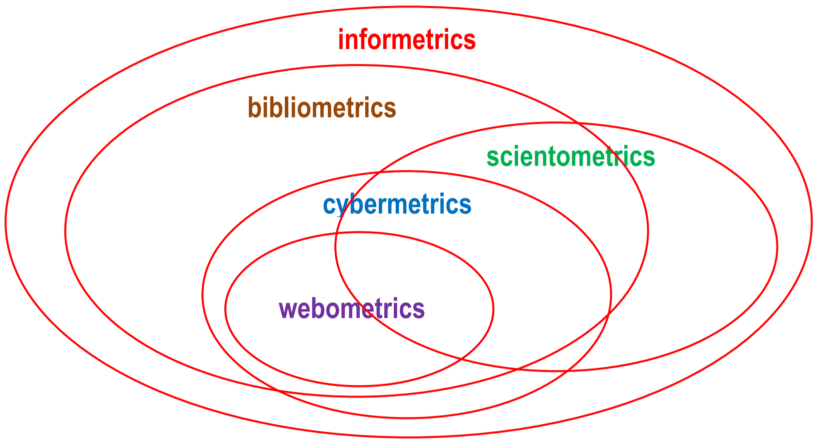
“Scientometrics” is the English translation of the title word of Nalimov’s classic monograph “**Naukometriya**” in 1969, which was relatively unknown to western scholars even after it was translated into English. Without access to the internet and limited distribution, it was rarely cited. However, the term became better known once the journal “Scientometrics” appeared in 1978 [Garfield, 2007] and term has grown in popularity and is used to describe the study of science: growth, structure, interrelationships and productivity [Mooghali et al, 2011].

Scientometrics is related to and has overlapping interests with Bibliometrics and Informetrics. The terms Bibliometrics, Scientometrics, and Informetrics refer to component fields related to the study of the dynamics of disciplines as reflected in the production of their literature [Hood & Wilson, 2001]. A whole community of researchers concerned with counting papers and citations called themselves bibliometricians [Godin, 2005].

Among the many statistical analyses of scientific publications, bibliometrics holds a privileged place for counting scientific papers. Bibliometrics is one of the sub-fields concerned with measuring the output of scientific publications. Bibliometrics owes its systematic development mainly to the works of its founders V.V. Naliv, D.J. D. Price and Eugene Garfield in the 1950s. Since 1958 Bibliometrics has evolved as a field, taught in library and information science schools and it emerged as a tool for scientific evaluation for a number research groups around the world. This process was made possible by the work of Eugene Garfield and his “Science Citation Index”. Castell, an American psychologist, was credited with the launching of Scientometrics, when he produced statistics on a number of scientists and their geographical distribution, and ranked the scientists according to their performance. He introduced two dimensions into the measurements of science, namely, *quantity* and *quality*. The term informetrics was introduced by Blackert, Siegel and Nacke in 1979, but gained popularity by the launch of the international informertics conferences in 1987. A recent development in informetrics called the webometrics/cybermetrics, has become a part of the main stream library and information science research area. The term webometrics refers to the quantitative studies of the nature of scientific communication over the internet and its impact on diffusion of ideas and information. The inter-relations between Infor-, biblio-, sciento-, cyber-, and webometrics are illustrated on Figure 5 [Thelwall, 2006].

Dirk Tunger gave the next definitions [Tunger, 2007]:

- ***Bibliometrics*** is a study or measurement of formal aspects of texts, documents, books and information;
- ***Scientometrics*** analyses the quantitative aspects of the production, dissemination and use of scientific information with the aim of achieving a better understanding of the mechanisms of scientific research as a social activity;
- ***Informetrics*** is a sub-discipline of information sciences and is defined as the application of mathematical methods to the content of information science;
- ***Webometrics*** is the application of informetrical methods to the World Wide Web (WWW).



Firure 5. Infor-, biblio-, sciento-, cyber-, and webometrics**.**

The sizes of the overlapping ellipses are made for sake of clarity only. [Thelwall, 2006]

Citation tracking and Evaluation of Research

Citation tracking is very important. It allows for tracking of authors own influence, and therefore the influence of organization. It allows tracking the development of a technology, which may be the basis for progress undreamt of when a paper is written. Citation tracking provides information on other organizations and authors who are doing similar work, potentially for collaboration, and identifies publications that cover similar topics. Finally, tracking back in time can find the seminal works in a field [Fingerman, 2006].

The use of scientometric indicators in ***research evaluation*** emerged in the 1960s and 1970s, first in the United States and then also in various European countries. Before that time, research evaluation had not been formalized other than through the peer review system, on the one hand, and through economic indicators which could only be used at the macro-level of a national system, on the other.

The ***economic indicators*** (e.g., percentage of GDP spent on R&D) have internationally been developed by the Organization of Economic Co-operation and Development (OECD) in Paris. For example, the Frascati Manual for the Measurement of Scientific and Technical Activities form 1963 (or its new edition [Frascati Manual, 2002]) can be considered as response to the increased economic importance of science and technology which had become visible in economic statistics during the 1950s.

The idea that scientific knowledge can be organized deliberately and controlled from a mission perspective (for example, for military purposes) was a result of World War II. Before that time the intellectual organization of knowledge had largely been left to the internal mechanisms of discipline formation and specialist communications. The military impact of science and technology through knowledge-based development and mission-oriented research during World War II (e.g., the Manhattan project) made it necessary in 1945 to formulate a new science and technology policy under peacetime conditions.

In 1945, Vannevar Bush's report to the U.S. President entitled The Endless Frontier contained a plea for a return to a liberal organization of science. ***Quality control should be left to the internal mechanisms of the scientific elite, for example, through the peer review system.*** The model of the U.S. National Science Foundation from 1947 was followed by other Western countries. For example, the Netherlands created its foundation for Fundamental Scientific Research (ZWO) in 1950. With hindsight, one can consider this period as the institutional phase of science policies: the main policy instrument was the support of science with institutions to control its funding [Okubo, 1997].

The attention for the measurement of scientific communication originated from *an interest other than research evaluation*. During the 1950s and 1960s, the scientific community itself had become increasingly aware of the seemingly uncontrolled expansion of scientific information and literature during the postwar period. In addition to its use in information retrieval, the Science Citation Index produced by Eugene Garfield's Institute of Scientific Information came soon to be recognized as a means to objectify standards [Price, 1963; Elkana et al, 1978]. The gradual introduction of output indicators (e.g., numbers of publications and citations) could be legitimated both at the level of society - because it enables policy makers and science administrators to use arguments of economic efficiency - and internally, because quality control across disciplinary frameworks becomes difficult to legitimate unless objectified standards can be made available in addition to the peer review process [Leydesdorff, 2005].

In 1976 Francis Narin's pioneering study "Evaluative Bibliometrics" [Narin, 1976] was published under the auspices (not incidentally) of the U.S. National Science Foundation. In 1973 Henry Small had proposed a method for mapping the sciences based on the co-citations of scientific articles. While Small's approach tried to agglomerate specialties into disciplinary structures, Narin focused on hierarchical structures that operate top-down [Carpenter & Narin, 1973; Pinski & Narin, 1976]. This program appealed to funding agencies like the N.S.F. and N.I.H. that faced difficult decisions in allocating budgets across disciplinary frameworks [Leydesdorff, 2005].

Recent years have seen quantitative bibliometric indicators being increasingly used as a central element in the assessment of the performance of scientists, either individually or as groups, and as an important factor in evaluating and scoring research proposals.

These indicators are varied (see [bibliometric, 2012]), and include e.g.:

- Citation counts of individual papers published by researchers;
- Journal metrics (the impact factors of the journals);
- Measures that quantify personal research contributions over an extended period.

Journal metrics

Journal metrics measure the performance and/or impact of ***scholarly journals***. Each metric has its own particular features, but in general, *they all follow the theories and practices of advertizing and aim to provide rankings* and insight into journal performance *based on citation analysis* (very similar to "audience reach measures" and rankings).

They start from the basic premise that a citation to a paper is a form of endorsement, and the most basic analysis can be done by simply counting the number of citations that a particular paper attracts: more citations to a specific paper means that more people consider that paper to be important.

Citations to journals (via the papers they publish) can also be counted, thus indicating how important a particular journal is to its community, and in comparison to other journals. Different journal metrics use different methodologies and data sources, thus offering different perspectives on the scholarly publishing landscape, and bibliometricians use different metrics depending on what features they wish to study [Elsevier, 2011].

For example, let remember four metrics:

- Journal Impact Factor (IF);
- SCImago Journal Rank (SJR);
- Eigenfactor;
- Source-Normalized Impact per Paper (SNIP).

Journal ***Impact Factor***; is a measure of a journal's average citations per article. The impact factor was computed by dividing the number of citations by the number of articles contained in the journal. This made it possible to eliminate any bias stemming from a journal's size, rendering citation proportional to the number of articles.

The Impact Factor (IF) is the brainchild of Dr. Eugene Garfield, who devised a system of quantifying the number of times a manuscript is referenced in the literature [Teixeira da Silva & Van, 2013]. As indicated by Thomson Reuters (http://thomsonreuters.com/products_services/science/free/essays/impact_factor/), the IF is calculated as an extremely simple equation:

Year impact factor ***IF = C/N***, where ***C*** = Cites to articles published in two previous years (Year-1) and (Year-2) (this is a subset of total cites in current Year); ***N*** = number (sum) of articles published in Year-1 and Year-2.

Developed by Professor Félix de Moya, ***SCImago Journal Rank*** (SJR) [SCI, 2013] is a prestige metric based on the idea that "*all citations are not created equal*". With SJR, the subject field, quality, and reputation of the journal have a direct impact on the value of a citation. This means that a citation from a source with a relatively high SJR is worth more than a citation from a source with a lower SJR.

The essential idea underlying the application of these arguments to the evaluation of scholarly journals is to assign weights to bibliographic citations based on the importance of the journals that issued them, so that citations issued by more important journals will be more valuable than those issued by less important ones. This "importance" will be computed recursively, i.e., the important journals will be those which in turn receive many citations from other important journals [González-Pereira et al, 2009].

SJR assigns relative scores to all of the sources in a citation network. Its methodology is inspired by the Google PageRank algorithm, in that not all citations are equal. A source transfers its own 'prestige', or status, to another source through the act of citing it. A citation from a source with a relatively high SJR is worth more than a citation from a source with a lower SJR. A source's prestige for a particular year is shared equally over all the citations that it makes in that year; this is important because it corrects for the fact that typical citation counts vary widely between subject fields. The SJR of a source in a field with a high likelihood of citing is shared over a lot of citations, so each citation is worth relatively little. The SJR of a source in a field with a low likelihood if citing is shared over few citations, so each citation is worth relatively much. The result is to even out the differences in citation practice between subject fields, and facilitate direct comparisons of sources. SJR emphasizes those sources that are used by prestigious titles [Elsevier, 2011].

The ***Eigenfactor® score*** of a journal is an estimate of the percentage of time that library users spend with that journal. The Eigenfactor algorithm corresponds to a simple model of research in which readers follow chains of citations as they move from journal to journal. Imagine that a researcher goes to the library and selects a journal article at random. After reading the article, the researcher selects at random one of the citations from the article. She then proceeds to the journal that was cited, reads a random article there, and selects a citation to direct her to her next journal volume. The researcher does this *ad infinitum*.

The amount of time that the researcher spends with each journal gives us a measure of that journal’s importance within network of academic citations. Moreover, if real researchers find a sizable fraction of the articles that they read by following citation chains, the amount of time that our random researcher spends with each journal gives us an estimate of the amount of time that real researchers spend with each journal. While we cannot carry out this experiment in practice, we can use mathematics to simulate this process [Bergstrom, 2007].

Source-Normalized Impact per Paper (SNIP) corrects for differences in the frequency of citation across research fields. SNIP measures a source’s contextual citation impact. It takes into account characteristics of the source’s subject field, especially the frequency at which authors cite other papers in their reference lists, the speed at which citation impact matures, and the extent to which the database used in the assessment covers the field’s literature. SNIP is the ratio of a source’s average citation count per paper, and the ‘citation potential’ of its subject field. It aims to allow direct comparison of sources in different subject fields.

A source’s subject field is the set of documents citing that source. The citation potential of a source’s subject field is the average number of references per document citing that source. It represents the likelihood of being cited for documents in a particular field. A source in a field with a high citation potential will tend to have a high impact per paper.

Citation potential is important because it accounts for the fact that typical citation counts vary widely between research disciplines – they tend to be higher in Life Sciences than in Mathematics or Social Sciences, for example. If papers in one subject field contain on average 40 cited references while those in another contain on average 10, then the former field has a citation potential that is four times higher than that of the latter. Citation potential also varies between subject fields within a discipline. For instance, basic journals tend to show higher citation potentials than applied or clinical journals, and journals covering emerging topics tend to have higher citation potentials than periodicals in well established areas.

For sources in subject fields in which the citation potential is equal to the average of the whole database, SNIP has the same value as the ‘standard’ impact per paper. But in fields with a higher citation potential – for instance, a topical field well covered in the database – SNIP is lower than the impact per paper. In fields in which the citation potential is lower – for instance, more classical fields, or those with moderate database coverage – SNIP tends to be higher than the impact per paper. In this way, SNIP allows you to rank your own customized set of sources, regardless of their subject fields [Elsevier, 2011].

Concluding this chapter we have to remember that a metric in business is a measure used to gauge some quantifiable component of an organization’s performance, such as return on investment (ROI), or revenues. Metrics are part of the broad area of business intelligence used to help business leaders make more informed decisions. Organizations often use metrics to develop a systematic approach to transform an organization’s mission statement and strategy into quantifiable goals, and to monitor the organization’s performance in terms of meeting those goals [GPM, 2010]. At the knowledge market, the journal metrics are aimed for quantitative evaluation the popularity and importance of the journals as well as their impact. These metrics have to be used carefully. They are useful for publishers, librarians and administrators, but are not applicable for evaluating of personal scientific contributions. At first, the quantity personal measures were introduced to achieve this goal.

Quantity measures

Quantity measures that quantify personal research contributions over an extended period are based mainly on the idea of [Hirsch, 2005]. Several papers related to research indices were proposed to assess the quality of the academic research publications. Each one of those indices has its own strengths and weaknesses. The idea of having research indices started when J. Hirsh proposed the H-index [Hirsch, 2005].

Although the H-index has many limitations and seems biased or unfair in many cases, the other proposed indices such as: G-, H(2)-, HG-, Q^2 -, AR-, M-quotient, M-, W-, H_w- ,E-, A-, R- , W-, J-index, etc. considered H-index as a suitable base to produce those other indices with some behavioral enhancements in order to overcome its limitations. In fact, all the other indices are calculated based on the number of citations (originally proposed in H-index) which the authors' papers received. The differences between those indices can be shown through how the index deals with the citations number, as in H-index, G-index, W-index, or in adding new attributes such as time, average…etc as in Contemporary H-index, M-quotient, and AR- index [Maabreh & Alsmadi, 2012]. A review focused in h-Index variants, computation and standardization for different scientific fields is given in [Alonso et al, 2009]. Following [Bornmann et al, 2008] in Table 1 below we remember some definitions of popular indexes.

Table 1. Definitions of the h index and its variants [Bornmann et al, 2008]

Index	Definition
N/yr	Total number of publications (N) divided by years of publishing (yr)
N_{pr}/yr	Number of peer-reviewed publications (N_{pr}) divided by years of publishing (yr)
Cit	Total number of citations (Cit) received by an author
Cit/N	Citations per publication
H index [Hirsch, 2005]	A scientist has index *h* if *h* of his or her N_p published papers have at least *h* citations each and the other (N_p - *h*) papers have fewer than ≤ *h* citations each"
M quotient [Hirsch, 2005]	$\frac{h}{y}$ where *h* = *h* index, *y* = number of years since publishing the first paper
G index [Egghe, 2006]	"The highest number *g* of papers that together received g^2 or more citations"
H(2) index [Kosmulski, 2006]	"A scientist's *h*(2) index is defined as the highest natural number such that his *h*(2) most-cited papers received each at least $[h(2)]^2$ citations"
A index [Jin, 2006]	$\frac{1}{h}\sum_{j=1}^{h} cit_j$ where *h* = *h* index, *cit* = citation counts
M index [Bornmann et al, 2008]	The median number of citations received by papers in the Hirsch core (this is the papers ranking smaller than or equal to *h*)
R index [Jin et al, 2007]	$\sqrt{\sum_{j=1}^{h} cit_j}$ where *h* = *h* index, *cit* = citation counts
AR index [Jin et al, 2007]	$\sqrt{\sum_{j=1}^{h} \frac{cit_j}{a_j}}$ where *h* = *h* index, *cit* = citation counts, *a* = number of years since publishing
H_w index [Egghe & Rousseau, 2008]	$\sqrt{\sum_{j=1}^{r_o} cit_j}$ where *cit* = citation counts, r_o = the largest row index *j* such that $r_w(j) \leq cit_j$
Creativity index (C_a) [Soler, 2007]	$\sum_{i=1}^{N_p} \frac{c(n_i, m_i)}{a_i}$ where: N_p=Number of published papers; n_i=Number of references for paper "i"; m_i=Number of citations for paper "i"; a_i=Number of authors for paper "i"; c=not clearly defined in reference

Disadvantages of journal metrics and quantitative measures

At the first glance, the variety of scientific measures seems to be very great and with great differences.

Really, they all are based on counting the citations and similar formulas based or not on additional criteria like prestige of the journals, time periods, number of authors, etc.

The indexes for quantifying personal research contributions are based on same idea of the Hirsh with modifications.

The subject of limitations in research indices is still evolving and with all proposed indices, there are still limitations and weaknesses. Moreover, the large number of available indices may lead to the dispersion of the evaluation, and therefore produce differences in values among research communities or even countries [Maabreh & Alsmadi, 2012].

References may also be negative. An author may be cited for research of a controversial nature or for an error of methodology. Here too, citation does not always measure the quality of research but rather the impact of a particular piece of work or of an individual scientist [Okubo, 1997].

At the end, if an academic shows good citation metrics, it is very likely that he or she has made a significant impact on the field. However, the reverse is not necessarily true. If an academic shows weak citation metrics, this may be caused a lack of impact on the field. However, it may also be caused by: working in a small field; publishing in a language other than English (LOTE); or publishing mainly (in) books [Harzing, 2008].

Sites and tools that are interested in the evaluation of researchers and research publications may have to calculate and display all the indices, and this may cause two issues [Maabreh & Alsmadi, 2012]:

- Large number of indices, if used, may clutter pages and make them unreadable;
- Since most likely values will be different among those indices, and in some cases they may even contradict with each other, such information will be misleading to the reader rather than being helpful or informative.

From the beginning, the quantitative measuring of scientific work has been criticized due to problems raised during evaluation of scientific results. Let point one of the earliest papers "Why the impact factor of journals should not be used for evaluating research" [Seglen, 1997]. Its arguments are still valid:

Problems associated with the use of journal impact factors [Seglen, 1997]

- Journal impact factors are not statistically representative of individual journal articles;
- Journal impact factors correlate poorly with actual citations of individual articles;
- Authors use many criteria other than impact when submitting to journals;
- Citations to "non-citable" items are erroneously included in the database;
- Self citations are not corrected for;
- Review articles are heavily cited and inflate the impact factor of journals;
- Long articles collect many citations and give high journal impact factors;
- Short publication lag allows many short term journal self citations and gives a high journal impact factor;
- Citations in the national language of the journal are preferred by the journal's authors;
- Selective journal self citation: articles tend to preferentially cite other articles in the same journal;
- Coverage of the database is not complete;
- Books are not included in the database as a source for citations;

- Database has an English language bias;
- Database is dominated by American publications;
- Journal set in database may vary from year to year;
- Impact factor is a function of the number of references per article in the research field;
- Research fields with literature that rapidly becomes obsolete are favored;
- Impact factor depends on dynamics (expansion or contraction) of the research field;
- Small research fields tend to lack journals with high impact;
- Relations between fields (clinical v basic research, for example) strongly determine the journal impact factor;
- Citation rate of article determines journal impact, but not vice versa;

Summary points [Seglen, 1997]:

- Use of journal impact factors conceals the difference in article citation rates (articles in the most cited half of articles in a journal are cited 10 times as often as the least cited half);
- Journals' impact factors are determined by technicalities unrelated to the scientific quality of their articles;
- Journal impact factors depend on the research field: high impact factors are likely in journals covering large areas of basic research with a rapidly expanding but short lived literature that use many references per article;
- Article citation rates determine the journal impact factor, not vice versa.

These problems still exist and are object for current discussions. For example, the major disadvantage of the Web of Science is that it may provide a substantial underestimation of an individual academic's actual citation impact. This is true equally for the two functions most generally used to perform citation analyses – for the "general search" and for the Web of Science "cited reference". However, the Web of Science "general search" function performs more poorly in this respect than the "cited reference" function. There are a number of reasons for the underestimation of citation impact by Thomson ISI Web of Science, for instance [Harzing, 2008]:

- Web of Science General Search is limited to ISI-listed journals - In the General Search function Web of Science only includes citations to journal articles published in ISI listed journals [Roediger, 2006]. Citations to books, book chapters, dissertations, theses, working papers, reports, conference papers, and journal articles published in non-ISI journals are not included;
- Web of Science Cited Reference is limited to citations from ISI-listed journals - In the Cited Reference function Web of Science does include citations to non-ISI publications. However, it only includes citations from journals that are ISI-listed.

Both Google Scholar and Thomson ISI Web of Science have problems with academics that have names including either diacritics (e.g. Özbilgin or Olivas-Luján) or apostrophes (e.g. O'Rourke) [Harzing, 2008]:

- In Thomson ISI Web of Science a search with diacritics provides an error message and no results;
- In Google Scholar a search for the name with diacritics will generally not provide any results either.
- For both databases doing a search without the diacritic will generally provide the best result.

The popularity and the wide use of the h-index have raised a lot of criticism.

The most notable and well-documented example of critical view on the h-index (and other "simple" measures of research performance) is the report by the joint Committee on Quantitative Assessment of Research [Adler et al, 2008]. In this report, the authors argue strongly against the use (or misuse) of citation metrics (e.g., the impact factor or the h-index) alone as a tool for assessing quality of research, and encourage the use of more complex methods for judging scientists, journals or disciplines, that combine both citation metrics as well as other criteria such as memberships on editorial boards, awards, invitations or peer reviews. With regard to the h-index (and associated modifications), specifically, [Adler et al, 2008] stress that its simplicity is a reason for failing to capture the complicated citation records of researchers, loosing thus crucial information essential for the assessment of a scientist's research. The lack of mathematical/statistical analysis on the properties and behavior of the h-index is also mentioned. This is in contrast to the rather remarkable focus of many articles to demonstrate correlations of h-index with other publication/citation metrics (i.e. published papers or citations received), a result which according to the authors is self-evident, since all these variables are essentially functions of the same basic phenomenon, i.e. publications [Panaretos & Malesios, 2009].

Besides the above-mentioned works, there are many more articles referring to disadvantages of the h-index. In what follows we list some of the most important disadvantages of the h-index [Panaretos & Malesios, 2009]:

- The h-index is bounded by the total number of publications. This means that scientists with a short career (or at the beginning of their career), are at an inherent disadvantage, regardless of the importance of their discoveries. In other words, it puts newcomers at a disadvantage since both publication output and citation rates will be relatively low for them;
- Some authors have also argued that the h-index is influenced by self-citations. Many self-citations would give a false impression that the scientists' work is widely accepted by the scientific community. Both self-citations and "real" (independent) citations are usually used in the calculation of the h-index. In this context, the emerging problem is that scientists with many co-operating partners may receive many self-citations, in contrast to scientists that publish alone;
- The h-index has slightly less predictive accuracy and precision than the simpler measure of mean citations per paper;
- Another problem is that the h-index puts small but highly-cited scientific outputs at a disadvantage. While the h-index de-emphasizes singular successful publications in favor of sustained productivity, it may do so too strongly. Two scientists may have the same h-index, say, h = 30, i.e., they both have 30 articles with at least 30 citations each. However, one may have 20 of these papers that have been cited more than 1000 times and the other may have all of his/hers h-core papers receiving just above 30 citations each. It is evident that the scientific work of the former scientist is more influential;
- Limitations/differences of the citation data bases may also affect the h-index. Some automated searching processes find citations to papers going back many years, while others find only recent papers or citations;
- Another database related problem often occurring with a significant effect on the correct calculation of the h-index, is that of name similarities between researchers. It is almost impossible to find a scientist with a unique combination of family name and initials while searching the most known citation databases. As a result, in many cases the h-index will be overestimated, since in its calculation the works of more than one researcher are added;
- It seems that the h-index cannot be utilized for comparing scientists working in different scientific fields. It has been observed that average citation numbers differ widely among different fields;

— General problems associated with any bibliometric index, namely the necessity to measure scientific impact by a single number, apply here as well. While the h-index is one 'measure' of scientific productivity, some object to the practice of taking a human activity as complex as the formal acquisition of knowledge and condense it to a single number. Two potential dangers of this have been noted:

(a) Career progression and other aspects of a human's life may be damaged by the use of a simple metric in a decision-making process by someone who has *neither the time nor the intelligence* to consider more appropriate decision metrics;

(b) Scientists may respond to this by maximizing their h-index to the detriment of doing more quality work.

This effect of using simple metrics for making management decisions has often been found to be an unintended consequence of metric-based decision taking; for instance, governments routinely operate policies designed to minimize crime figures and not crime itself.

The disadvantages of the h-index may be seen in the indices which inherit its properties. For instance, some advantages and disadvantages of quantity metrics were outlined by [Thompson, 2009] (see Table 2).

Table 2. Some advantages and disadvantages of quantity metrics [Thompson, 2009]

Metric	Advantages	Disadvantages
N/yr	Measures gross productivity	Definition of “publication” can be arbitrary; No insight into the importance or impact of published works
N_{pr}/yr	Measures gross productivity Eliminates marginal publications	No insight into the importance or impact of published work
Cit	Measures total impact of a body of work	Can be inflated by a small number of papers with high citation counts.
Cit/N	Measures total impact of a body of work normalized by the number of published papers.	Tends to reward low productivity Can penalize high productivity
h-index	Combines quantitative (publication numbers) and impact (citation counts) into a simple whole number. Identifies a set of core, high performance journal articles (“Hirsch core”)	Insensitive to highly cited work
M quotient	Allows h-index comparisons between faculty that differ in seniority	Insensitive to highly cited work
G index	Once a paper makes the Hirsh core, additional citations in this group are not counted further; the g index takes these further citations into account	Gives more weight to highly cited papers
H(2) index	Since h(2) index is always smaller then h-index, it is less open to problems of citations accuracy	Possibly overly sensitive to a few highly cited papers

Metric	Advantages	Disadvantages
A index	Calculates the average number of citations in the Hirsch core	Emphasizes more of the impact of the Hirsch core than quantity. Can be very sensitive to a few highly cited papers
M index	Median value may be a better measure of central tendency because of the skewed nature of citation counts	Emphasizes more of the impact of Hirsch core than the quantity.
R index	Involves the Hirsch core but does not "punish" an author for having a high h-index unlike the a-index	Emphasizes more of the impact of the Hirsch core than quantity. Can be very sensitive to a few highly cited papers
AR index	Normalizes the r index by the number of years publishing allowing comparison of younger and more seasoned faculty	Similar to r index
Creativity index (C_a)	Only scholarship metric that proposes to measure creativity	Insufficient data to validate this metric at present. The calculation of the creativity index is not simple, however the author of paper has a free download of a program that will calculate the index

Very important disadvantage of quantitative measures is that they are applicable only to cited papers.

In 1991, David A. Pendlebury of the Philadelphia-based Institute for Scientific Information had published the startling conclusion that

> ***55% of the papers published in journals covered by ISI's citation database did not receive a single citation in the 5 years after they were published*** [Hamilton, 1991].

In his further publication, Pendlebury gave more concrete data. He had written [Pendlebury, 1991]:

"The figures -- 47.4% un-cited for the sciences, 74.7% for the social sciences, and 98.0% for the arts and humanities -- are indeed correct.

These statistics represent every type of article that appears in journals indexed by the Institute for Scientific Information (ISI) in its Science Citation Index, Social Sciences Citation Index, and Arts & Humanities Citation Index. The journals' ISI indexes contain not only articles, reviews, and notes, but also meeting abstracts, editorials, obituaries, letters like this one, and other marginalia, which one might expect to be largely un-cited. In 1984, about 27% of the items indexed in the Science Citation Index were such marginalia. The comparable figures for the social sciences and arts and humanities were 48% and 69%, respectively.

If one analyzes the data more narrowly and examines the extent of un-cited articles alone, the figures shrink, some more than others: **22.4%** of 1984 science articles remained un-cited by the end of 1988, as did **48.0%** of social sciences articles and **93.1%** of articles in arts and humanities journals.

If one restricts the analysis even further and examines the extent of un-cited articles by U.S. authors alone, the numbers are even less "worrisome."

Only 14.7% of 1984 science articles by U.S. authors were left un-cited by the end of 1988.

We estimate the share of un-cited 1984 articles by non-U.S. scientists to be about 28%" [Pendlebury, 1991].

Authors from developing countries

Whatever performance metrics we may use, it appears that ***authors from developing countries*** do face certain constraints in terms of achieving higher performance indices and therefore recognition for themselves and their country. *It is quite possible that authors from advanced countries may tend to cite publications from organizations located in their own countries, leading to a disadvantage for authors working in difficult situations, with less funding opportunities* Since there is a limited page budget and increased competition in many "high-profile" journals, it is *not always possible to publish in these journals*.

One way to overcome this problem is to encourage and give value to papers published in national journals. There are many scientists from developing countries such as India working in highly developed countries with advanced scientific infrastructure and huge funding. These scientists should seriously consider publishing their work in journals originating from their native countries. This will bring an international flavor to the national journals, attracting more international authors and ultimately making them mainstream international journals. When these journals become more visible and easily accessible through their online versions, there is a chance that papers published in these journals are more often cited [Kumar, 2009].

In other words, *developing national knowledge markets became mission important and considerable*.

Mentoring abilities

In addition, we should measure the ***mentoring abilities*** of a scientist. Scientists do research and also mentor younger colleagues. Good mentoring should be a significant consideration of one's contribution to science. The h-index might measure research productivity, but currently there does not appear to be a "*mentoring index*" [Jeang, 2008]. If the coauthors of a scientist are his or her own trainees or students and if they continue to make a scientific impact after leaving their supervisor, it does point to the quality of the mentoring by the scientist and to the impact made by the scientist, as a result of his/her mentoring abilities, in a given area during a given period. This is a very important but totally neglected aspect of the contribution made by a scientist or an academic.

However, *we do not yet have a well–worked out formula to measure such mentoring abilities* [Kumar, 2009].

Evaluation of Scientific Contributions

The products of science are not objects but ***ideas***, means of communication and reactions to the ideas of others. While it is possible simultaneously to track scientists and money invested, it is far more difficult to measure *science as a body of ideas*, or to grasp its interface with the economic and social system. For now, indicators remain essentially a unit of measure based on observations of science and technology as a system of activities rather than as a body of specific knowledge [National Science Foundation, 1989].

Research papers and publications are important indicators for the ability of an author or an education community to conduct research projects in the different human science fields. In general, the number of publications and the increase in this number is a direct indicator of the size or the volume of research activities for a particular author or university. Nonetheless, the number of publications merely, is showed to be a limited indicator to show the impact of those publications. The number of citations for a particular paper is shown to be more relevant and important in comparison to the number of publications. This is why early citation indices such as H-index and G-index gave more weight and important to the number of citations in comparison to the number of publications [Maabreh & Alsmadi, 2012].

Each indicator has its advantages and its limitations, and care must be taken not to consider them as "absolute" indices [Atanassov & Detcheva, 2012; Atanassov & Detcheva, 2013]. The "convergence" of indicators has to be tested in order to put the information they convey into perspective [Martin & Irvine, 1985]

Usefulness of Scientific Contribution

The Main Phases of the Science are

(1) Creation of a Scientific Result;

(2) Registration of the Scientific Result;

(3) Implementation and Using of the Scientific Result.

The bibliometric indexes analyze the second phase – registration of scientific result as (primary) publications and as (secondary) citations. The first and third phases are out of bibliometric scope. This way the evaluating of scientific work became partial and not significant. Practically, the evaluation of scientific results is closed in the contours of the Knowledge Markets (KM) shown at Figure 2 and/or Figure 3, i.e. without taking in account the main knowledge customers of the KM.

A possible step, to counterbalance and to extend consideration to all KM elements shown at Figure 1, is to analyze the publications and citations from point of view of the third phase – ***implementation and using the scientific results*** *by the members of KM.*

A wide spread understanding is that only high qualified ***academic researchers*** (Scientists (S), Figure 1) can evaluate published ideas. They have knowledge and skills to continue research and developing of proposed ideas and via citations they recognize previous research done by other scientists or by themselves. In accordance to *usefulness of cited ideas*, we may separate academic citations on three main groups:

– ***Substantial citations***, which applied or supported the citing work indicating implementation and using the citied results, including “mentoring impact”;

– ***Casual citations***, which noted only or reviewed the citing work;

– ***Refuting citations,*** which indicate that the citing work (possibly) has no scientific added value.

Regarding ***industrial researchers*** (Researchers (R), Figure 1) we may make the similar consideration. They have knowledge and skills to implement the published ideas and to evaluate their usefulness for industrial applications. Here the citations are mainly in two groups:

– *Substantial citations*, which applied or supported the citing work indicating implementation and using the citied results, including “mentoring impact”;

– *Refuting citations,* which indicate that the citing work (possibly) has no scientific added value to be implemented.

Further analysis of the KM-scheme concerns the educational cycle done by ***Lecturers*** ((L), Figure 1), ***Tutors*** ((T), Figure 1) and ***Examiners*** ((E), Figure 1). Their main goal is to assist Employees in learning of the published ideas. In this cycle, the citations are in text-books, methodical or other supporting publications, and educational learning materials. All such citations we may classify as:

– *Casual citations*, which noted only or reviewed the citing work.

The ***Employees*** ((Ee), Figure 1) may use the received knowledge in their everyday activities. During educational process they may create some new knowledge information objects with or without new ideas. For instance, they may prepare different theses, surveys, guides, papers, etc. In such case, the types of citations may vary, i.e. it may be:

– *Substantial citations*, which applied or supported the citing work indicating implementation and using the citied results, including “mentoring impact”;

– *Casual citations*, which noted only or reviewed the citing work;

– *Refuting citations,* which indicate that the citing work (possibly) has no scientific added value.

The ***Employers*** ((Er), Figure 1) are the most important members of KM. They invest both in developing man power as well as in research activities. In both cases the evaluation of usefulness of scientific results is not by citations in papers but by amount of invested assets. This way their citations may be classified only as

– *Substantial citations*, which applied or supported the citing work indicating implementation and using the citied results, including "mentoring impact"

if the amount of investments is over some normalized limit. Usually the investments are provided by scientific or educational projects and because of this we may assume that one project corresponds to one substantial citation.

At the end we have to pay attention to two main distributors of knowledge ***Publishers*** ((P), Figure 1) and ***Administrators*** ((A), Figure 1). After first publishing of the knowledge information objects (papers, books, etc.), Publishers start selling and corresponded advertizing. Main part of advertizing activities is indexing of published materials by different scientific digital libraries and data bases which are inherent for Administrators. All their citations may be classified as:

– *Casual citations*, which noted only or reviewed the citing work.

Transitive citations

The useful scientific results may cause a chain of publications which further use and develop them. This way, transitive citations will exist. Citation chain has to start from a substantial citation and to continue by same type citations because casual citations could not generate such citation chain.

The influence of the scientific ideas is greatest when citation chains exist. Because of this, the *transitive substantial citations* have to be counted as native characteristic of the scientific publications. It is correct to assume that a transitive substantial citation is equal to direct one.

Temporal dimension

There is also a ***temporal dimension*** to the citation process. An article may first be cited for substantial reasons (e.g., its content has been used). Later when a paper is widely known and has obtained many citations the importance of the other mechanisms will increase (authors citing authoritative papers, the bandwagon effect, etc.). In other words, ***visibility dynamics*** become more important over with time because of the self-intensifying mechanisms that are involved. This explains why the relative differences in citation rates between poorly cited and highly cited papers increase over time. Another temporal effect is the phenomenon termed "obliteration by incorporation", meaning that basic theoretical knowledge is not cited anymore. As a consequence, the most basic and important findings may not be among the most highly cited papers because they have been rapidly incorporated into the common body of accepted knowledge [Aksnes, 2005].

Concluding this short survey we have to draw attention to one very important fact.

A great number of publications have no chance to be viewed and further studied because they are published in media with limited and/or payable access. In this case only well-known authors have chance to be recognized and possibly – cited.

Only what is needed is publications to be included in different digital libraries with open access and *as more such libraries exist* in the world so greatest chance these publications have to become useful. The variety of digital libraries and index data bases with open access to scientific publications and reviews is a crucial factor for further grow of the science. One may say that such practice will destroy the knowledge markets. This is partially true. The societies invest in science by direct or indirect financing and further business with scientific results is not admissible

USC-methodology

Following considerations discussed above, we assume that for evaluating of usefulness of scientific contributions more-less important are:

- ***p*** – Number of the papers;
- ***q*** – Number of monographs;
- ***s*** – Number of the substantial citations;
- ***c*** – Number of the casual citations;
- ***r*** – Number of the refuting citations;
- ***Y*** = y_e - y_b +1 – Length of the interval of publications;
- ***z*** = y_c - y_b – Length of the interval of citations,

where

- y_b – starting year (beginning) of the period of publications;
- y_e – last year (end) of the period of publications;
- y_c – last year (end) of the period of citations.

In this list we have three different types of values which we have to reduce to common measurement unit. We propose to use "paper" as such unit because it may be assumed that *one paper represents a single idea*.

In accordance with this, we propose to use four coefficients of correlation:

- ***m*** – coefficient of the monograph correlation:
  - ⇒ ***m** : **1 monograph** = **m papers**;* example: if 1 monograph = 5 papers than m = 5;
- ***a*** – coefficient of the substantial citation correlation:
  - ⇒ ***a : 1 substantial citation** = **1/a paper**;* example: if 5 substantial citations = 1 paper than a=5;
- ***b*** – coefficient of the casual citation correlation:
  - ⇒ ***b : 1 casual citation** = **1/b paper**;* example: if 10 casual citations = 1 paper than b = 10;
- ***v*** – coefficient of the refuting citation correlation:
  - ⇒ ***v : 1 refuting citation** = **1/v paper**;* example: if 10 refuting citation = 1 paper than v = 10.

This way we have the methodological formula for *Usefulness of Scientific Contributions* (***usc-index***):

$$usc = \frac{p + mq + z}{Y} + \frac{s}{aY} + \frac{c}{bY} - \frac{r}{vY}$$

This formula is ***only a formal representation*** of the understanding that the ***scientific contributions have to be evaluated completely*** taking in account as more parameters as possible. All types of publications have to be included in the evaluation process as well as mentoring activities, learning materials, and all types of citations including transitive citations, implementations, scientific projects, received funding, etc.

Special comment is needed for ***substantial self-citations***. They are indicator that the scientists provide longtime investigation and step by step publish new results. This is normal cycle of science. Ignoring this means that we expect receiving the results in one "genius" invention. In addition, mentoring students and young researchers lead to publishing of co-authored papers which cause ***substantial citations from co-authors*** in further their independent work and publications. As the received knowledge is more qualitative so more important are the further citations from co-authors. Ignoring this means that we do not acknowledge the high level skills and leading ideas of the advisors.

Example

Results from an experiment with real data taken from DBLP (http://dblp.uni-trier.de/) are presented in Table 3. In the real data there was no data for monographs and refuting citations. Because of this the corresponded columns contain zeroes.

Table 3. Experimental data for usc-index

scientist	usc	y_b	y_e	y_c	Y	z	m	a	b	v	p	q	s	c	r
S1	**26.07**	1991	2011	2009	**21**	18	5	5	10	10	405	0	15	1215	0
S2	**13.74**	1983	2011	2011	**29**	28	5	5	10	10	109	0	208	2200	0
S3	**13.52**	1995	2011	2011	**17**	16	5	5	10	10	110	0	32	975	0
S4	**11.66**	1981	2011	2011	**31**	30	5	5	10	10	181	0	50	1406	0
S5	**8.48**	1999	2011	2010	**13**	11	5	5	10	10	44	0	8	537	0
S6	**8.23**	1972	2011	2011	**40**	39	5	5	10	10	98	0	66	1789	0
S7	**6.68**	2000	2011	2007	**12**	7	5	5	10	10	53	0	10	182	0
S8	**5.57**	1985	2011	2011	**27**	26	5	5	10	10	68	0	22	520	0
S9	**4.36**	2007	2011	2010	**5**	3	5	5	10	10	16	0	1	26	0
S10	**3.87**	1991	2010	2010	**20**	19	5	5	10	10	44	0	1	142	0
S11	**3.71**	2003	2011	2008	**9**	5	5	5	10	10	26	0	0	24	0
S12	**3.62**	2004	2009	2011	**6**	7	5	5	10	10	8	0	0	67	0
S13	**3.62**	1983	2009	2011	**27**	28	5	5	10	10	47	0	2	223	0
S14	**3.54**	1973	1986	2008	**14**	35	5	5	10	10	11	0	2	32	0
S15	**3.33**	2009	2011	2010	**3**	1	5	5	10	10	8	0	1	8	0
S16	**3.16**	1995	2009	2011	**15**	16	5	5	10	10	18	0	2	130	0
S17	**2.42**	1986	2011	2006	**26**	20	5	5	10	10	34	0	2	85	0
S18	**2.35**	2008	2011	2011	**4**	3	5	5	10	10	6	0	1	2	0
S19	**1.63**	2001	2011	2008	**11**	7	5	5	10	10	10	0	1	7	0
S20	**0.96**	1991	2006	2001	**16**	10	5	5	10	10	5	0	1	1	0

USC-index reflects the dynamics of scientific development during the analyzed period. For instance, scientist S2 has more long scientific career and more citations than S1 but his usc-index is less than that of S1 due to less number of papers for longer period.

It is important to remark: *periods have different lengths (column Y) and for further analysis it has to be accounted.*

It is complicated to compute usc-index for all scientists of a given organization and many times more complicated to do this for all researchers from given scientific area. Because of this, the computer linguistic analysis of the scientific publications (to obtain values of the main parameters of usc-index) is serious scientific problem which has to be solved. Some preliminary considerations about possibility for solving it may be done. For instance, it is typical that the introduction of a scientific article is structured as a progression from the general to the particular. References have been found to be most frequent in the introductory section of paper. Thus, in the introduction, an article typically refers to more general or basic works within a field. The net effect of many articles referring to the same general works, therefore, is that such contributions get a very large number of citations. References to highly cited publications seemed to occur more in the introduction than anywhere else in the articles. Similarly, since most scientific articles contain a methodology section in which the methods applied in the study are documented, authors typically cite the basic papers describing these methods. This may explain why some papers containing commonly used methods sometimes receive a very large number of citations [Aksnes, 2005].

Conclusion

Starting point of our consideration was the introduction of the "Information Market" as a payable information exchange and based on it information interaction. In addition, special kind of Information Markets - the Knowledge Markets (KM) were outlined. Basic understanding of our work is that we have to evaluate the usefulness of scientific contributions from point of view of those for whom the results are created. This is not simple task because the KM customers are of many kinds.

The identifying of the staple commodities of the knowledge markets was a step of the process of investigation of contemporary situation in the global knowledge environment. Investigation of the staple commodities of the knowledge markets is very difficult but useful task. We have introduced them as kind of information objects, called "knowledge information objects". The main their distinctive characteristic is that they contain information models, which concerns sets of information models and interconnections between them.

We belong to the modern knowledge market and perhaps we shall agree that "à la marché comme à la marché" ("at the market as at the market"). In the world of science, there exist commercial interests that set the trends to redistribute the money given for science by the societies. Unfortunately, for instance, the "impact factor" is just such trend, borrowed from advertising industry, to force scientists to invest in selected retailer chains.

It is not permissible to replace the quality of a scientific publication, with qualities of the media in which it has been published.

In science, the incorrect management decisions lead to a decline in its development. *If a complete scientific "industry" is not developed, the "complete" administrative attitude to science grows, which inevitably will kill it.* Exuberant dependence on single numbers to quantify scientists' contribution and make administrative decisions can affect their career progression or may force people to somehow enhance their h-index instead of focusing on their more legitimate activity, i.e., doing good science. Considering the complex issues associated with the calculation of scientific performance metrics, it is clear that a comprehensive approach should be used to evaluate the research worth of a scientist. We should not rely excessively on a single metric [Kumar, 2009].

Although the use of such quantitative measures may be considered at first glance to introduce objectivity into assessment, the exclusive use of such indicators to measure science "quality" can cause severe bias in the assessment process when applied simplistically and without appropriate benchmarking to the research environment being considered. Funding agencies are aware of this, nevertheless experience shows that the reviewing of both individuals and projects on the national and European level is still relying excessively on the use of these numerical parameters in evaluation. This is a problem of much concern in the scientific community, and there has been extensive debate and discussion worldwide on this topic [bibliometric, 2012].

Since the very first applications of bibliometric indicators in this way, scientists and science organizations have taken strong positions against such purely numerical assessment. Various organizations in Europe have published studies on their potential adverse consequences on the quality of funded scientific research. A prime example is the publication of the Académie des Sciences of the Institut de France that has presented clear recommendations on the correct use of bibliometric indices [IDF, 2011]. Other publications have addressed the role of peer review in the assessment of scientists and research projects e.g. the European Science Foundation Peer Review Guide published in 2011 [ESF, 2011a] with recommendations for good practices in peer review following an extensive European survey on peer review practices [ESF, 2011b]. Other recent examples are a study of peer review in publications by the Scientific and Technology Committee of the House of Commons in the UK [STC, 2011], the peer review guide of the Research Information Network in the UK [RIN, 2010] and the recommendations formulated at a workshop dedicated to quality assessment in peer review of the Swedish Research Council [SRC, 2009].

A common conclusion of these studies is the recognition of the important role of peer review in the quality assessment of research, and the recommendation to apply bibliometric performance indicators with great caution, and only by peers from the particular discipline being reviewed [bibliometric, 2012].

A considerable step toward this goal is ***The San Francisco Declaration on Research Assessment*** (DORA), [DORA, 2012] initiated by the American Society for Cell Biology (ASCB) together with a group of editors and publishers of scholarly journals, who recognize the need to improve the ways in which the outputs of scientific research are evaluated. The group met in December 2012 during the ASCB Annual Meeting in San Francisco and subsequently circulated a draft declaration among various stakeholders. DORA as it now stands has benefited from input by many of the original signers. It is a worldwide initiative covering all scholarly disciplines.

A special press release of *Initiative for Science in Europe (ISE)* called ***"Initiative to put an end to the misuse of the journal impact factor (JIP)"*** has been published [ISE, 2012]. We have kind permission of ISE to reprint text:

> "Major European science organizations have joined the "San Francisco Declaration On Research Assessment" which was released today by the American Society for Cell Biology (ASCB). Signatories in Europe include the European Mathematical Society, EUCheMS, European Sociology Association, European Education Research Association, FEBS, EMBO and other societies and organizations that are organized under the umbrella of the Initiative for Science in Europe (ISE).
>
> The increasing reliance on journal based metrics for research assessment, hiring, promotion or funding decisions has been criticized by experts for a number of years. The "San Francisco Declaration On Research Assessment" for the first time unites researchers, journals, institutions and funders to address the problems of an overreliance on the journal impact factor and to work for change of the current system of research assessment.
>
> The declaration formulates concrete recommendations for different stakeholder groups. It calls publishers to "greatly reduce emphasis on the journal impact factor as a promotional tool", funding agencies and institutions to consider "the value and impact of all research outputs" for purpose of research assessment, "including qualitative indicators of research impact" and researchers to make "decisions about funding, hiring, tenure, or promotion, [..] based on scientific content rather than publication metrics" when involved in assessment committees. It also invites organizations that supply metrics to "[b]e open and transparent by providing data and methods used to calculate all metrics".
>
> The San Francisco Declaration on Research Assessment was drafted by a group of editors and publishers of scholarly journals that met at the Annual Meeting of The American Society for Cell Biology (ASCB) in San Francisco in December 2012. It has since developed into a worldwide initiative welcoming all scientific disciplines including the social sciences and humanities.
>
> Scientists and institutions alike are invited to express their commitment and support for the initiative at http://ascb.org/SFdeclaration.html" [ISE, 2012].

Endorsing DORA, the Association for Computers and the Humanities (ACH) remarked that it is a set of recommendations for applying more nuanced, accurate ways to evaluate research than the Journal Impact Factor

(JIF). DORA makes eighteen recommendations for researchers, funders, research institutions, organizations that provide metrics, and publishers, such as focusing evaluation on the content of a paper, applying article-based rather than journal-based metrics, incorporating research outputs such as datasets and software in evaluating impact, and promoting the reuse of reference lists through the adoption of Creative Commons Public Domain Dedication licenses.

In addition, we have to underline that the variety of digital libraries and index data bases with open access to scientific publications and reviews is a crucial factor for further grow of the science. One may say that such practice will destroy the knowledge markets. This is only partially true because the societies invest in science by direct or indirect financing and further business with scientific results is not admissible

Following the considerations given above, this paper was aimed to present a new *usc-methodology* for evaluating the scientific contribution of a scientist or a scientific group (organization).

It consists in proposing three main groups of citations: ***Substantial citations***, ***Casual citations***, and ***Refuting citations***, which all have *temporal dimensions*.

In addition, due to existence of different types of values (for monographs, papers and citations), a common measurement unit (“idea” or “paper”) and four coefficients (for monographs, substantial, casual, and refuting citations) of correlation to measurement unit (paper) have been proposed.

The problem of automatic linguistic analysis of scientific publications, in accordance with usc-methodology and computing of its ***usc-index*** for different target scientific structures has been outlined.

Finally, we have to underline, that usc-methodology is aimed only to turn process of evaluation of scientific contributions back to human responsibility of authors, reviewers, and publishers. Modern science is distributed all over the world and concentration of any it’s part in one or two monopolies is absolutely inadmissible. To ensure growing of science we are obligated to provide for growing of variety of possibilities for doing science – financial resources, publishing opportunities, scientific indexing systems, and distributing organizations.

In addition to all printed universe we are obligated to take in account the variety of possibilities for direct contact between scientists in a single place like conferences, seminars, and workshops or distributed geographically like tele-conferences, electronic mailing lists, blogs, etc.

Special comment was done for *substantial self-citations*. They are indicator that the scientists provide longtime investigation and step by step publish new results. In addition, mentoring students and young researchers lead to publishing of co-authored papers which cause *substantial citations from co-authors* in further their independent work and publications. As the received knowledge is more qualitative so more important are the further citations from co-authors. Ignoring this means that we do not acknowledge the high level skills and leading ideas of the researchers and advisors.

This *usc-index* is *only a formal representation* of the understanding that the *scientific contributions have to be evaluated completely* taking in account as more parameters as possible. All types of publications as well as mentoring activities, learning materials, and all types of citations including substantial self-citations, substantial citations from co-authors, transitive citations, implementations, scientific projects, received funding, etc. have to be included in the evaluation of usefulness of scientific contributions.

Acknowledgements

Authors are indebted to acad. Alexander Palagin, prof. Levon Aslanyan and colleagues from the scientific seminar on “Information Modeling” for fruitful discussions and recommendations.

This paper is partially supported by the project **ITHEA-XXI**.

Bibliography

[Adler et al, 2008] Adler, R., Ewing, J. and Taylor, P. “Citation Statistics”, Joint IMU/ICIAM/IMSCommittee on Quantitative Assessment of Research, 2008 http://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf (accessed at 09.05.2013).

[AIRM, 2013] Association, Information Resources Management, "Data Mining: Concepts, Methodologies, Tools, and Applications (4 Volumes)." IGI Global, 2013, 1-2120. Web, 9 May. 2013, doi: 10.4018/978-1-4666-2455-9, http://www.igi-global.com/book/data-mining-concepts-methodologies-tools/68176 (accessed at 09.05.2013).

[Aksnes, 2005] Aksnes D.W. Citations and their use as indicators in science policy. // Dissertation for the doctoral degree of the University of Twente, March 2005.

[Alonso et al, 2009] S. Alonso, F.J. Cabrerizo, E. Herrera-Viedma, F. Herrera. h-Index: A review focused in its variants, computation and standardization for different scientific fields. // Journal of Informetrics 3 (2009) pp. 273–289. http://sci2s.ugr.es/hindex/pdf/JOI-3-4-273-289.pdf (accessed at 29.05.2013).

[AMA, 2013] American Marketing Association, “Dictionary - MarketingPower, inc. 2013”, http://www.marketingpower.com/_layouts/Dictionary.aspx?dLetter=A (accessed at 09.05.2013).

[Atanassov & Detcheva, 2012] V. Atanassov, E. Detcheva, Theoretical Analisis of Empirical Relationships for Pareto-Distributed Scientometric Data, Int. J. "Information Models and Analyses", Vol.1/2012, ISSN 1314-6416, pp.271-282.

[Atanassov & Detcheva, 2013] V. Atanassov, E. Detcheva, Citation-Paper Rank Distributions and Associated Scientometric Indicators – A Survey, Int. J. "Information Models and Analyses", Vol.2/2013, ISSN 1314-6416 (in print).

[Bergstrom, 2007] Carl Bergstrom, “Eigenfactor: Measuring the value and prestige of scholarly journals”. College & Research Libraries News, 2007, 68(5), 314–316, http://octavia.zoology.washington.edu/publications/Bergstrom07.pdf, (accessed at 18.05.2013).

[bibliometric, 2012] “On the use of bibliometric indices during assessment” EPS Statement // 5-11 June 2012 http://c.ymcdn.com/sites/www.eps.org/resource/collection/B77D91E8-2370-43C3-9814-250C65E13549/EPS_statement_June2012.pdf (accessed at 09.05.2013).

[Bornmann et al, 2008] Lutz Bornmann, Rüdiger Mutz, Hans-Dieter Daniel. Are There Better Indices for Evaluation Purposes than the h Index? A Comparison of Nine Different Variants of the h Index Using Data from Biomedicine, Journal of the American society for information science and technology, 59(5):830–837, 2008.

[BSC, 2013] Britain's Science Council, Definition of “Science” http://www.sciencecouncil.org/definition (accessed 09.05.2013)

[Carpenter & Narin, 1973] Carpenter, M. P and Narin, F. “Clustering of scientific journals”, J. Am. Soc. Inf. Sci., 1973, 24, 425–436.

[CBED, 2013] Definition of indirect advertising noun from the Cambridge Business English Dictionary, Cambridge University Press, 2013, http://dictionary.cambridge.org/dictionary/business-english/indirect-advertising?q=indirect+advertising (accessed at 09.05.2013).

[DORA, 2012] San Francisco Declaration on Research Assessment // American Society for Cell Biology (ASCB). http://am.ascb.org/dora/ (accessed at 17.05.2013).

[Egghe & Rousseau, 2008] Egghe, L. and Rousseau, R. “An h-index weighted by citation impact”, Information Processing & Management, 44, 2008, pp. 770–780, doi:10.1016/j.ipm.2007.05.003.

[Egghe, 2006] Egghe, L. “An improvement of the h-index: the g-index”. ISSI Newsletter, 2(1), 2006b, pp. 8–9.

[Elkana et al, 1978] Elkana, Y., Lederberg, J., Merton, R. K., Thackray, A. and Zuckerman, H., Toward a Metric of Science: The Advent of Science Indicators, Wiley, New York, 1978.

[Elsevier, 2011] Elsevier B.V. “Frequently asked questions” Research analytics redefined Copyright © 2011 All rights reserved. Scopus is a registered trademark of Elsevier B.V. http://www.journalmetrics.com/faq.php (accessed at 09.05.2013).

[ESF, 2011a] European Science Foundation, European Peer Review Guide, Integrating Policies and Practices for Coherent Procedures, Member Organisation Forum, 2011.

http://www.esf.org/fileadmin/Public_documents/Publications/European_Peer_Review_Guide.pdf (accessed at 26.05.2013).

[ESF, 2011b] European Science Foundation, Survey Analysis Report on Peer Review Practices, Member Organisation Forum, 2011, http://www.esf.org/fileadmin/Public_documents/Publications/pr_guide_survey.pdf (accessed at 26.05.2013)

[EzineMark, 2013] Ezine Mark. Direct Vs Indirect Marketing And Advertising. http://advertising.ezinemark.com/direct-vs-indirect-marketing-and-advertising-31ebf6babc8.html (accessed at 09.05.2013).

[Fingerman, 2006] Susan Fingerman "Web of Science and Scopus: Current Features and Capabilities" Electronic Resources Reviews Copyright 2006, Used with permission http://www.istl.org/06-fall/electronic2.html (accessed at 09.05.2013).

[Frascati Manual, 2002] Frascati Manual "The Measurement of Scientific and Technological Activities", Proposed Standard Practice for Surveys on Research and Experimental Development. 6th edition (2002), pp. 266, // Organisation for Economic Co-operation and Development (OECD), Paris, France. 2002. ISBN 978-92-64-19903-9. http://www.oecd.org/innovation/inno/frascatimanualproposedstandardpracticeforsurveysonresearchandexperimentaldevelopment6thedition.htm. Publié en français sous le titre: Manuel de Frascati 2002. Méthode type proposée pour les enquêtes sur la recherche et le développement experimental, OECD Publications, Paris, France.

[Garfield, 1977] Garfield, E. Can Citation Indexing Be Automated? // Essay of an Information Scientist, vol. 1 (Vol. 1). Philadelphia: ISI Press. 1977.

[Garfield, 2007] Eugene Garfield, The evolution of the Science Citation Index, International Microbiology, 2007, 10:65-69, DOI: 10.2436/20.1501.01.10, ISSN: 1139-6709, http://garfield.library.upenn.edu/papers/barcelona2007.pdf, (accessed at 09.05.2013)

[Gladun, 1994] Gladun V. P. The processes of knowledge creation, ("Процессы формирования знаний", Педагог 6), Sofia, 1994 (in Russian).

[Godin, 2005] Benoît Godin, From Eugenics to Scientometrics: Galton, Cattell and Men of Science. Project on the History and Sociology of S&T Statistics, Working Paper No. 32, Canadian Science and Innovation Indicators Consortium (CSIIC). 2006

[González-Pereira et al, 2009] Borja González-Pereira, Vicente P. Guerrero-Bote and Félix Moya-Anegón "The SJR indicator: A new indicator of journals' scientific prestige", 2009, http://arxiv.org/ftp/arxiv/papers/0912/0912.4141.pdf (accessed at 09.05.2013).

[GPM, 2010] State Government of Victoria State Services Authority, A guide to people metrics// The State Services Authority, Melbourne, 2010.

[Hamilton, 1991] David P. Hamilton," Research Papers: Who's Uncited Now?", Science, 251:25, 1991, http://garfield.library.upenn.edu/papers/hamilton2.html (accessed at 09.05.2013).

[Harzing, 2008] Anne-Wil Harzing "Google Scholar - a new data source for citation analysis", Research in International Management Products & Services for Academics, 2008, http://www.harzing.com/pop_gs.htm (accessed at 09.05.2013).

[Harzing, 2012] Harzing A.-W. How to become an author of ESI Highly Cited Papers? // University of Melbourne, 2012. http://www.harzing.com/esi_highcite.htm (accessed at 09.05.2013).

[Hawkins, 1982] Hawkins J.M., "*The Oxford Paperback Dictionary",* Oxford University Press, 1982, ISBN 0-19-281209-2.

[Hirsch, 2005] Hirsch, J. E., "An index to quantify an individual's scientific research output", Proceedings of the National Academy of Sciences of the United States of America, 102(46), 2005, pp. 16569–16572, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1283832/pdf/pnas-0507655102.pdf (accessed at 09.05.2013)

[Hood & Wilson, 2001] William W. Hood, Conceptión S. Wilson. "The literature of bibliometrics, scientometrics, and informetrics", Scientometrics, vol. 52, No. 2 (2001) 291–314, http://faculty.kfupm.edu.sa/MATH/kabbaj/Benchmarks/HoodWilson2001.pdf (accessed at 09.05.2013).

[Hornby et al, 1987] Hornby A.S., A.P. Cowie, A.C. Gimson, "*Oxford Advanced Learner's Dictionary",* Oxford University Press 1987, ISBN 0-19-431106-6

[IDF, 2011] Institut de France, Académie des Sciences, On the Proper use of Bibliometrics to Evaluate Individual Researchers, Rapport de l'Académie des sciences, 2011, www.academie-sciences.fr/activite/rapport/avis170111gb.pdf (accessed at 26.05.2013).

[ISE, 2012] Join Initiative to put an end to the misuse of the journal impact factor (JIP) // Initiative for Science in Europe (ICE). http://www.no-cuts-on-research.eu/news/dora.html (accessed at 17.05.2013).

[Ivanova et al, 2001] N. Ivanova, K. Ivanova, K. Markov, A. Danilov, K. Boikatchev. *The Open Education Environment on the Threshold of the Global Information Society,* IJ ITA, 2001, V.8, No.1 pp.3-12. (Presented at Int. Conf. KDS 2001 Sankt Petersburg, 2001, pp.272-280, in Russian, Presented at Int. Conf. ICT&P 2001, Sofia, pp.193-203)

[Ivanova et al, 2003] Kr. Ivanova, N. Ivanova, A. Danilov, I. Mitov, Kr. Markov, “Education of adult on the professional knowledge market”, (*Обучение взрослых на рынке профессиональных знаний,* Сборник доклади на Национална научна конференция “Информационни изследвания, приложения и обучение”), i.TECH-2003, Varna, Bulgaria, 2003, pp. 35-41, (in Russian).

[Ivanova et al, 2006] Kr. Ivanova, N. Ivanova, A. Danilov, I. Mitov, Kr. Markov. Basic Interactions between Members of the Knowledge Market. Int. Journal “Information Theories and Applications”, Vol.13/2006, No.:1, pp.19-30.

[Jeang, 2008] Kuan-Teh Jeang, “H-index, mentoring-index, highly-cited and highly accessed: how to evaluate scientists?” Retro virology, vol. 5, Article Number: 106, Nov 2008. BioMed Central, Springer. http://link.springer.com/article/10.1186%2F1742-4690-5-106 (accessed at 09.05.2013).

[Jeeves, 2013] Ask Jeeves. What is Indirect Advertising? http://uk.ask.com/question/what-is-indirect-advertising(accessed at 09.05.2013).

[Jin et al, 2007] Jin, B., Liang, L., Rousseau, R., & Egghe, L. “The R- and AR- indices: Complementing the h-index”, Chinese Science Bulletin, 52(6), 2007, pp. 855–863.

[Jin, 2006] Jin, B “h-index: an evaluation indicator proposed by scientist”, Science Focus, 1(1), 2006, pp. 8–9.

[Kosmulski, 2006] Kosmulski, M. A new Hirsch-type index saves time and works equally well as the original h-index. ISSI Newsletter, 2(3), 2006, pp. 4–6.

[Kumar, 2009] M. Jagadesh Kumar Evaluating Scientists: Citations, Impact Factor, h-Index, Online Page Hits and What Else? IETE TECHNICAL REVIEW, Vol. 26, ISSUE 3, 2009, DOI: 10.4103/0256-4602.50699; Paper No TR 81_09; Copyright © 2009 by the IETE, pp. 165 - 168

[Levine et al, 2005] Michele Levine, Gary Morgan, Angela Brooks, Marcus Tarrant, Howard Seccombe, “Advertising-Adverteasing-Advertiring?”, What sort of ROI can you expect from Print Advertising – unless your ad performs?” 12th Worldwide Readership Research Symposium, Prague - October 23-26, 2005, Roy Morgan International Melbourne, Australia, 2005

[Leydesdorff, 2005] Loet Leydesdorff, “The Evaluation of Research and the Evolution of Science Indicators”, Current Science, 2005, Vol. 89, No. 9, pp. 1510-1517, http://arxiv.org/ftp/arxiv/papers/0911/0911.4298.pdf (accessed at 09.05.2013)

[Maabreh & Alsmadi, 2012] Majdi Maabreh and Izzat M. Alsmadi “A Survey of Impact and Citation Indices: Limitations and Issues“, International Journal of Advanced Science and Technology, Vol. 40, March, 2012, pp. 35-54

[Markov et al, 2001] K. Markov, P. Mateev, K. Ivanova, I. Mitov, S. Poryazov. “*The Information Model”,* IJ ITA, 2001, V.8, No.2 pp.59-69 (Presented at Int. Conf. KDS 2001 Sankt Petersburg, 2001, pp.465-472)

[Markov et al, 2002] K. Markov, K. Ivanova, I. Mitov, N. Ivanova, A. Danilov, K. Boikatchev. *Basic Structure of the Knowledge Marke*t”, IJ ITA, 2002, V.9, No.4, pp.123-134 .

[Markov et al, 2003] Kr.Markov, Kr.Ivanova, I.Mitov, “*General Information Theory”, Basic Formulations,* FOI-COMMERCE, Sofia, 2003, ISBN 954-16-0024-1

[Markov et al, 2013] Kr. Markov, Kr. Ivanova, V. Velychko. “Evaluation of Scientific Contributions” // International Scientific Conference “Modern Informatics: Problems, Achievements, and Prospects of Development,” devoted to the 90th anniversary of academician V. M. Glushkov. Kiev, Ukraine, 2013 (in print).

[Markov et al, 2013a] Markov, Krassimir, Koen Vanhoof, Iliya Mitov, Benoit Depaire, Krassimira Ivanova, Vitalii Velychko and Victor Gladun. "Intelligent Data Processing Based on Multi-Dimensional Numbered Memory Structures", Diagnostic Test Approaches to Machine Learning and Commonsense Reasoning Systems, IGI Global, 2013. 156-184, Web. 7 May. 2013. doi:10.4018/978-1-4666-1900-5.ch007

[Markov et al, 2013b] Markov, Krassimir, Koen Vanhoof, Iliya Mitov, Benoit Depaire, Krassimira Ivanova, Vitalii Velychko and Victor Gladun. "Intelligent Data Processing Based on Multi-Dimensional Numbered Memory Structures." Data Mining: Concepts, Methodologies, Tools, and Applications, IGI Global, 2013, 445-473, Web. 6 May. 2013. doi:10.4018/978-1-4666-2455-9.ch022

[Markov et al., 2006] Kr. Markov, Kr. Ivanova, I. Mitov. "The Staple Commodities of the Knowledge Market", Int. Journal "Information Theories and Applications", Vol.13/2006, No.:1, pp.5-19.

[Markov, 1999] Kr. Markov. About harms from e-commerce (Относно вредите от електронната търговия), IECC'99: International e-Commerce Conference, ADIS & VIA EXPO, Bulgaria, Sofia, 1999 (in Bulgarian).

[Martin & Irvine, 1985] Br. Martin and J. Irvine, 'Evaluating the Evaluators: A Reply to Our Critics', Social Studies of Science, 1985, 15, pp. 558-75.

[Merton, 1957a] Merton, R.K. "Priorities in Scientific Discovery", American Sociological Review, Vol. 22, 1957, p. 635.

[Merton, 1957b] Merton, R. K. "Social and Democratic Social Structure", in Social Theory and Social Structure, Free Press, New York, 1957, pp. 550-561.

[Mooghali et al, 2011] A. Mooghali, R. Alijani, N. Karami, A. Khasseh "Scientometric Analysis of the Scientometric Literature" International Journal of Information Science and Management, Vol. 9, No. 1 January / June 2011, pp. 21-31

[Naidenova & Ignatov, 2013] Naidenova, Xenia and Dmitry I. Ignatov, "Diagnostic Test Approaches to Machine Learning and Commonsense Reasoning Systems", IGI Global, 2013, 1-367, Web 9, May 2013, doi:10.4018/978-1-4666-1900-5 http://www.igi-global.com/book/diagnostic-test-approaches-machine-learning/63869 (accessed at 09.05.2013).

[Narin, 1976] Francis Narin. Evaluative Bibliometrics: The use of Publication and Citation Analysis in the Evaluation of Scientific Activity. Computer Horizons, Inc. Project No. 704R, March 31, 1976. http://yunus.hacettepe.edu.tr/~tonta/courses/spring2011/bby704/narin_1975_eval-bibliometrics_images.pdf (accessed at 09.05.2013).

[National Science Foundation, 1989] National Science Foundation, "Science & Engineering Indicators" 1989, National Science Foundation, Washington, DC

[NLC, 2004] *Electronic Publishing: Guide to Best Practices for Canadian Publishers*. National Library of Canada, Ottawa, Created: 2001-10-03. Updated: 2004-03-0, http://www.nlc-bnc.ca/9/13/index-e.html, http://www.collectionscanada.ca/obj/p13/f2/01-e.pdf (accessed at 09.05.2013).

[Okubo, 1997] Okubo, Y. (1997), "Bibliometric Indicators and Analysis of Research Systems: Methods and Examples", OECD Science, Technology and Industry Working Papers, 1997/01, OECD Publishing. http://dx.doi.org/10.1787/208277770603 (accessed at 09.05.2013).

[Panaretos & Malesios, 2009] John Panaretos and Chrisovaladis Malesios, "Assessing scientist research performance and impact with single indices", MPRA Paper No. 12842, 2008, http://mpra.ub.uni-muenchen.de/12842/, Scientometrics, December 2009, Volume 81, Issue 3, pp 635-670. http://link.springer.com/article/10.1007%2Fs11192-008-2174-9 (accessed at 09.05.2013).

[Pendlebury, 1991] David A. Pendlebury, "Science, Citation, and Funding", Science, 251:1410-1411, 1991. http://garfield.library.upenn.edu/papers/pendlebury.html (accessed at 09.05.2013).

[Peper, 2009] Peper M. Publish or Perish. May 22, 2009. http://blogs.library.duke.edu/libraryhacks/2009/05/22/publish-or-perish/ (accessed at 09.05.2013).

[Pinski & Narin, 1976] Pinski, G. and Narin, F., Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. Inf. Proc. Manage., 1976, 12(5), 297–312.

[Price, 1963] Price, D.S., Little Science, Big Science, Columbia University Press, New York, 1963.

[RIN, 2010] Peer review, A guide for researchers, Research Information Network, UK, March 2010. www.rin.ac.uk/our-work/communicating-and-disseminating-research/peer-review-guide-researchers) (accessed at 26.05.2013).

[Roediger, 2006] Roediger III, H.L. The h index in Science: A New Measure of Scholarly Contribution, APS Observer: The Academic Observer, 2006, vol. 19, no. 4.

[SCI, 2013] Science Citation Index http://thomsonreuters.com/products_services/science/science_products/a-z/science_citation_index/#tab1 (accessed at 09.05.2013).

[Seglen, 1997] Per O Seglen "Why the impact factor of journals should not be used for evaluating research". BMJ 1997, 314:498–502.

[Small, 1982] Small, H.. Citation Context Analysis. // In B. Dervin & M.-. Voigt (Eds.), Progress in communication sciences (Vol. 3, pp. 287-310). Norwood: Ablex. 1982.

[Soler, 2007] Jose M. Soler. A Rational Indicator of Scientific Creativity. // Journal of Informetrics, Volume 1, Issue 2, April 2007, Pages 123–130. http://www.sciencedirect.com/science/article/pii/S1751157706000253. also available as eprint arXiv:physics/0608006v1 [physics.soc-ph]. (accessed at 29.05.2013)

[SRC, 2009] Swedish Research Council, Quality Assessment in Peer Review, Vetenskapsrådet, 2009, ISBN 978-91-7307-190-1, www.cm.se/webbshop_vr/pdfer/2011_01L.pdf (accessed at 26.05.2013)

[STC, 2011] Peer review in scientific publications, Science and Technology Committee, House of Commons, UK, 18 July 2011. www.publications.parliament.uk/pa/cm201012/cmselect/cmsctech/856/85602.htm (accessed at 26.05.2013).

[Teixeira da Silva & Van, 2013] Jaime A. Teixeira da Silva and Pham Thanh Van "The Impact of the Impact Factor®: Survey among Plant Scientists", Received: 1 October, 2011. Accepted: 29 October, 2012, the Asian and Australasian Journal of Plant Science and Biotechnology©2013 Global Science Books.

[Thelwall, 2006] Mike Thelwall, Liwen Vaughan & Lennart Bjorneborn: Webometrics in: Ann. Rev. Of Information Science & Technology (2006), p. 84, http://www.academia.edu/709949/Webometrics (accessed at 09.05.2013).

[Thompson, 2009] Dennis F. Thompson, Erin C. Callen, Milap C. Nahata, "New Indices in Scholarship. Assessment", American Journal of Pharmaceutical Education 2009; 73 (6) Article 111. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2769533/pdf/ajpe111.pdf (accessed at 09.05.2013).

[Tunger, 2007] Dirk Tunger, "Webometrics, Informetrics and Bibliometrics – How useful are these Indicators for Measuring Knowledge?"; ESSHRA June, 12-13, 2007 http://www.euresearch.ch/fileadmin/documents/events2007/ESSRHA07/ball_herget_tunger_esshra_june2007.pdf (accessed at 09.05.2013).

Authors' Information

***Markov Krassimir** – Institute of Mathematics and Informatics, BAS, Bulgaria; IJ ITA Editor in Chief. e-mail: markov@foibg.com*

Major Fields of Scientific Research: General theoretical information research, Multi-dimensional information systems

***Ivanova Krassimira** – University of National and World Economy, Sofia, Bulgaria; Institute of Mathematics and Informatics, BAS, Bulgaria;*

e-mail: krasy78@mail.bg

Major Fields of Scientific Research: Software Engineering, Business Informatics, Data Mining, Multidimensional multi-layer data structures in self-structured systems

***Velychko Vitalii** – Institute of Cybernetics, NASU, Kiev, Ukraine*

e-mail: Velychko@rambler.ru

Major Fields of Scientific Research: Data Mining, Natural Language Processing

RISKS IN USING BIBLIOMETRIC INDICATORS FOR PERFORMANCE EVALUATION OF SCIENTISTS

Douhomir Minev

Abstract: *The issues being discussed in this article are the consequences of the use of specific (journal – or article - and researcher-based) metrics (“bibliometric indices”) for assessment of the performance of scientists and research proposals.[1] The analysis is focused on the potential of the use of such indices to operate as a mechanism for control over the production of knowledge.*

The methodology is based on the complexity of relationships between sciences as systems for production of knowledge and their surrounding social environment. In these interactions arise motives for control and impact over knowledge production. The effects of these motives are expanding mechanisms for control over sciences and the knowledge they produce. The impact of the control mechanisms distorts knowledge and co-generates non-knowledge. When societies use distorted knowledge they face expansion of the so called “new risks”.

On this basis “bibliometric indices” are identified as components of larger (in many cases - supranational) system for control over knowledge production (sciences’ dynamics) and as generators of distorted knowledge and unexpected and negative consequences (new risks) for societies.

Keywords*: control (over sciences and knowledge); crisis of sciences; social knowledge; distorted knowledge, non-knowledge; new risks; bibliometric; academic assessment.*

Introduction

Although the use of journal-based and other (article- and researcher-based) metrics is not a radical innovation in assessing the work of scientists, the practice of assessments experienced important changes during the last decade – a trend of strengthening the impact of the metrics through: a) expanding the scope of the metrics’ application - use not only for the award of academic degrees and titles (the pursuit of individual scientific careers), but also in regularly conducted appraisals of scientists and evaluation of research projects; b) applying them often is compulsory (special standards have been set up and sometimes - by law); c) a key driver of this change are the governments (state institutions) etc.

The combination of these changes outlines a trend - the extensive introduction of a new system for performance evaluation of scientists.

As a result of the pressure of governmental institutions and other key actors, the trend spread in many countries in short period of time, despite the skepticism and criticism of scientists and their organizations[2].

---

[1] Two earlier versions of this article were presented at two seminars on the use of bibliometric indices during assessments of scientist’ performance – one of them organized by the Union of Scientists in Bulgaria and the other one – by a department at the Institute of Mathematics & Informatics at BAS, 2013.

[2] For instance, the American Society for Cell Biology (ASCB) together with a group of editors and publishers of scholarly journals launched a new initiative – the San Francisco Declaration on Research Assessment (DORA) - to make academic assessment less reliant on the impact factor.

Obviously, the system is given special importance because the objections were rejected and did not dissuade its introduction.

When, despite the doubts about the adequacy of the indicators per se and objections to the manner of their use, the institutionalization of the new system is imposed by dogged decision making centers, a number of questions arise: why the system is being so stubbornly introduced; what are the motives for introduction; what could be the implications thereof?

The State of Science and Production of Knowledge

"If there is such a thing as a Durkheimian conscience collective existing on a global scale, it is perhaps best represented by the widely held sentiment that we are living in a time of unprecedented danger. Although the chances of premature death or disability are probably no greater today than in any other periods of human history, the dangers we currently face are unique in two respects - they are caused by humans and their impact on us has a collective, rather than individual nature" [Lyng, Stephen, 2008, p. 106].

Many researchers search for the causes of these dangers in the existing knowledge, respectively - the state of science and the way they produce knowledge. Today is a widespread view that the systems of production of knowledge are experiencing crisis and produce uncertainty and ignorance rather than knowledge, and that if this trend continues, the future of societies will be determined more by ignorance and non-knowledge rather than knowledge. Among the proposals in response to this problem are identified two main ideas: either to slow down the production of knowledge or to reconstruct radically sciences and move on to post-academic (post-normal) science. The differences between the two proposals arise from the different diagnoses of the causes of the crisis in the production of knowledge.

The first proposal is based on the view that the cause of the crisis is the too rapid advancement of knowledge (too intensive production), therefore a slowdown in production is proposed and the most effective means of delay is reduction of resources or reduction in efficiency.

The second proposal (for reconstruction of sciences) stems from another cause of the crisis - the very 'means of production' of knowledge is already obsolete. The said obsolesence is due to a lack of connections with the moral values and the knowledge generated beyond the borders of science. It is therefore proposed a transition to a post-academic (post-normal) science, which operates based on the close relationship between scientists and the public [Funtowicz S. and Ravetz J., 1992]. Nevertheless, the question arises of the cause for breaking the links between moral values and knowledge that is created and exists outside of science itself? Many studies show that the problem is a fundamental and most common feature of today's production of knowledge - distorted relationships between power and knowledge, and more precisely - hypertrophied power control over the production of knowledge (some speak of vicious union between power and knowledge, other - of symbolic power, etc.) The reason for control is quite simple – as far as everything that people do intentionally depends on their knowledge, the control over creation of knowledge, allows to control everything that people do, and without them knowing that they are controlled (Note 1). Today this control mechanism has reached enormous proportions and complexity, but it is poorly studied, although remarkable minds are engaged in one or other sides of the power-knowledge relationship (mostly - Bourdieu, but also Foucault, Merton, Guldner and many researchers of risks). The mechanism has been extended and enhanced, especially intensively during the second half of 20th century, and it is what deforms sciences per se and the knowledge that they create. The problem is that control submits the creation of knowledge to one main goal - stability of the existing social order, in particular – preservation of positions of power elites who control society, and thus the created knowledge is inevitably distorted.

The Mechanism[1]

Central component of the mechanism are specific groups of scientists operating within science and influencing the way it functions, so that science creates knowledge, which meets certain requirements for the stability of the social order. Although they are difficult to distinguish from ordinary cliques commonly found in the sciences, these groups were noticed long ago (Alvin Guldner called them 'old people', P. Bourdieu – 'conservative ideologues', etc.) as a kind of 'proxies of power' in science. Influence that conservative ideologues have on knowledge production occurs along several lines: a) through influencing the basic elements constituting science itself as a system for production of knowledge (paradigms, theories (available or missing), approaches, principles, methods, sub-disciplinary structure, validation methods (recognition of obtained results as scientific), mechanisms for performance evaluation and therefore – also for shaping the scientific careers of researchers, etc.)[2]; b) impact on the choice of the main problematic areas of research interests; c) influence on the preferred interpretations of available data etc.

The second important component of the mechanism is the external support for the 'conservative ideologues' in science. The main channels through which these external effects run are science policies (defining research priorities and funding of sciences); special forms of support for 'conservative ideologues' especially - financial support[3] and occupying key positions in the hierarchy of research structures; support for private research centers and (albeit rare) private funding of research in public institutions.

Effects of Control Mechanism

Continuous effect of the control mechanism has given rise to a number of 'extra' effects, among which may be mentioned in particular:

- Distortion of incentives for research and the evaluation of the contribution of individual scientists;
- Establishment and a very strong influence of informal networks and groups where central position occupy the 'conservative ideologues';
- Formal internal hierarchies become too 'strong' and generate a deficit of democracy in the work of scientific organizations;
- Weakening of the impact of basic moral values on the modus operandi of scientific communities and organizations, etc.

The final effects of the above features are undermining the scientific communities and their work, and increasing dysfunctions of scientific organizations.

The effects of control and the relationship between them were noted long ago. Here is, for example, what Norbert Wiener wrote about some of these factors, already in 1947: 'It is clear that the demotion of the position of the scientist from an independent thinker to a servant who is employed in a science factory and is morally irresponsible, has happened much faster and in a much more devastating way than I expected. This

---

[1] Different sciences have different effects on the structures of power and domination. Some sciences contribute to consolidate the groups exercising power through the development of technologies in the economy, technologies of war and technologies of monitoring and control over their own population. The effect of social sciences is quite different. Therefore, the outlines of the control mechanism most clearly stand out exactly in social sciences. But this does not mean that there are no similar problems in other sciences.

[2] It is not hard to assume, and there are series of indications confirming the assumption, that such groups are also particularly active promoters of the new system for assessing the performance of scientists.

[3] The noticed by R. Merton 'Matthew effect' (accumulation of funding with the same scientists and groups) is usually an indicator for intensive financial support of conservative ideologues.

subordination of those who have to think to those who hold administrative power is destructive to the morale of the scientist' [Wiener, N.1947; Salomon, Jean-Jacques, 1970, pp. 308-309].

Moreover, demoralizing the scientists and hierarchies has far-reaching effects - it upsets scientific communities and their activities. Scientific communities are group (collective) mechanisms for production of knowledge and assessment of the scientific validity and relevance of results. The above effects of control (mainly - hierarchies and demoralization) weaken the capacity of communities to create knowledge.

Furthermore, 'conservative ideologues' and joining opportunists form their inner circles, groups, networks, cliques, which gradually conquer key positions in scientific organizations, gain significant influence and dominance in decision-making and submit the organization to the purposes of the dominant groups instead to the common organizational and social objective for which an organization is established and exists.

Under the influence of the above changes organizations lose the capacity to produce adequate knowledge, to regulate and govern themselves and become particularly susceptible to external influences flowing through informal networks and hierarchies.

Under the influence of the control mechanism (the key positions of 'conservative ideologues'; distorted incentive system; demoralization of communities, etc.) scientists quickly learn what efforts and results are the best rewarded and concentrate on to them. That is how variances occur in the process of knowledge creation.

Generator of Risks

Distortion of created knowledge has different aspects, but the main ones are: slow production of (for example - a strong lag behind of social sciences), creation of limited knowledge, creation of knowledge, whose scientific and social relevance are disputable [Simon, H., 1957, xxiv]; knowledge that may be more useful for society is not created.

A certain dynamics can be seen in these processes. For example, in the 50s, the Nobel Prize winner, Herbert Simon stated as an issue the production of irrelevant, unusable, unnecessary and useless knowledge. But today (the beginning of the second decade of the 21st century), knowledge is not just useless - Stefano Zamani already indicated the harmful effects of the research and not of peripheral researchers, but of the last ten winners of the Nobel Prize for economics.

It is not a coincidence that an increase of fraud in science is noticed (particularly evident for the award of academic degrees and titles, but also in the creation of new knowledge). And that exactly is the generator of hazards unknown to the society.

Search for a Solution

From this point of view, there may be two main alternatives to the creation of knowledge: maintaining and even expanding controls over the process or refusal of the old strategy, dismantling of control mechanisms and redeeming the production of knowledge. Supporters of the delay in the production of knowledge, tend to the first option. Supporters of the reconstruction of sciences are closer to the second alternative, because in essence they propose substitution of the knowledge - power alliance with a new one - between professional knowledge producers and public.

Where does the new scientists' performance evaluation system tend to?

Technology of Control

There are several signs that the system possesses capacity to strengthen and enlarge the control over production of knowledge.

First important sign, for example, is the identification of research work (i.e. creation of new knowledge) with publication activity, i.e. dissemination of already created knowledge. Although the publication activity of scientists has always been used to assess their work and performance, in the new system scientists can no longer present the results of their research that are not published or were not published in ‘prestigious’ journals, cannot submit the respective documentation proving that these results have been successfully tested for their scientific relevance and validity within a scientific community and cannot use these results in appraisal or for promotion in their career. Even if the results are very good, this does not affect directly the evaluation of scientists because of the compulsory requirement for a specific form of presentation - through publications and at that in certain (so-called prestigious) journals and editions. In fact, the main change is that the results obtain approval of another ‘validation center’.

However, the first objection is that in this way research work is not only identified with publication activity, but the one is substituted by the other. Such substitution is not only devoid of solid arguments, but also gives rise to a series of negative effects. Important questions in this regard remain unanswered: why dissemination of created knowledge is imposed as an obligation of the researchers (often - without asking their consent or just at least their opinion, and sometimes - despite their objections); why their performance appraisal should be more dependent on the fulfillment of this obligation incumbent upon them than by direct and substantive assessment of the way scientists perform their primary duty - to produce adequate new scientific knowledge; whether such substitution would not cause a decline in the usual, long tested and established procedures for verification of the scientific validity and relevance of the results achieved and how this (possible) decline would affect the knowledge produced and the societies that use it.

This substitution gives rise to several important effects. Firstly, is reduced the role of essential, direct and qualitative evaluations of scientific validity and relevance of the created knowledge, carried out in the relevant scientific communities. Direct substantive judgments of communities are replaced with indirect evaluations based on publication activity. Secondly, arises a kind of ‘disempowerment’ of scientific communities (respectively - the organizational structures in which they operate) because their judgments are substituted by journals (respectively – their editors and reviewers). Thirdly, emerge new ‘centers of influence’ with regard to the scientists and their work – the editors and their editorial apparatus. Thus, the substitution of substantive direct evaluation of scientific results with indicators based on publication activity actually shifts centers of judgment: from scientific communities (their organizational structures) to the editorial boards of the journals and then - not all editorial boards but mainly those with the highest rating[1]. Moreover, the location of these new centers of influence can be easily 'lifted' – at supranational level - by encouraging publications in foreign and international journals. Thus arises a sort of ‘globalization’ of the control (actually - centralization), whose reasons are not clear (scientists are increasingly sharing their results regardless of this system), but some of the consequences are pretty clear. Fourthly, where attributes (rating) of the journals in which research is published affect the performance evaluation of the scientist, this generates pressure to publish in appointed journals.

The biggest problem posed by these effects is that publication activity and publication success of scientists (i.e. the creation of knowledge) are in relationships that are (or at least - can be used as) technology for control over the production of knowledge. The technology is based on the possibility of the creation of knowledge to fall under the influence of factors of not strictly scientific nature. Some factors may even have anti-scientific nature.

---

[1] Major role here has the well known impact factor, but sometimes the scientist’s appraisal system includes also other indicators for ranking, for example whether the journals are national, foreign or international.

The falling of scientific work evaluation, and thus the scientific work per se, under the influence of non-scientific and even anti-scientific factors is not only evidence of the exercise of extra-scientific scrutiny, but inevitably leads to distortions in the production of knowledge.

Components of the Technology – 'Predispositions' of Selectors

Scientific journals can significantly influence the work of scientists. This possibility arises from the fact that in addition to the formal requirements of the journals to published materials there is also a wide range of informal requirements, which are not presented and are not formally put forward, but are applied quite strictly. These requirements are of cultural, ideological and political nature. Experiencing the pressure from the new evaluation system, scientists are beginning to adapt their work to the informal, unannounced requirements of the publication sphere (the journals). And in this way production of knowledge is influenced by non-scientific, even anti-scientific factors.

The system of informal requirements arises because the setting up of editorial boards of journals (respectively - circles of reviewers) and the way they work, stay far away from monitoring, participation and control by the scientific communities. In any case, it is possible (and this exactly happens) journals to selected authors and publications expressing certain views, predispositions, biases and even prejudices.

A remarkable fact illustrates the problem very well: many years ago, a legend in the economic analysis – V. Leontief (Nobel Prize winner), noted that the market theory has lost its connection with economic realities and that this poses a major risk for the very science of economics and economic policy and for the economies themselves. In protest against this distortion in knowledge, Leontief ceased to publish in the 'prestigious' economic journals because they were premeditated (uncritical) of the theory and especially contributed to the dominance of this mainstream of economic analysis. Thus, they contributed to the occurrence of the gap between theory and reality, through intensive dissemination of theoretical achievements of market theorists who are now more often called 'market fundamentalists'.

Another famed case also illustrates the above. Alan Sokal, a physicist at New York University, perceiving certain characteristics in the work of the journals, perpetrated a special hoax to highlight them. He submitted to a scientific journal (*Social Text*) an article entitled: "Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity".

The article was nonsense, deliberately compiled of fawning references structured around the silliest quotations he could find about mathematics and physics, but the magazine published it. The reasons for this, according to Sokal himself are two: "it sounded good and flattered the editors' <u>ideological preconceptions</u>". Knowledge researchers believe that this experiment of Sokal proves that 'the selection of articles to be published is highly dependent on political, social and cultural elements' [Bucchi, M. 2004, p. 95].

It is because of these dependencies, scientific journals of the highest rank often distort the most powerful tool for evaluating scientific results – the peer reviews. The above listed scientific organizations and their best practices in peer reviews show that the results of this evaluation procedure of scientific achievements are much better if the procedures are carried out within the scientific communities - especially where corresponding results are established.

However, particularly this adequate tool does not yield good results when applied in journals. A ruling on this issue was issued by no other than the U.S. Supreme Court. In one of its judgments it enacted that 'peer reviews' in journals can present deformed, distorted judgment. The same, only more emphatically, claim and scientists themselves. Chubin and Hackett point out a research, which shows that only 8% of the members of the Scientific Research Society (USA) have expressed the opinion that the method of peer review in journals gives good results. Obviously, arises the question of why the same tool gives different results in journals and in academic communities? The answer seems to lie in the fact that those who select works for publication in journals select them according to their ideological, political, and cultural predispositions.

Summarizing the above, it can be said that by the ratings of journals can be created (and are created) invisible hierarchies, hierarchical tiers in the publishing system, and these hierarchies are transferred to the very production of knowledge in the sciences themselves. When this happens, in practice are arranged hierarchically strands of research paradigms, theoretical constructs, methods, scientific communities and groups, and what is especially important - ultimately hierarchically are arranged also scientific results, various segments of created knowledge. This hierarchy is different and may have completely opposite structure to the real alignments based on the essential, direct evaluation of validity and relevance of research results. There are empirically established facts that confirm this statement.

The Australian Research Council has published the results of comparisons between the official (formal) values of the impact factor and the special expert assessments of the impact factor of one and the same journals. Large groups of experts have been involved for this purpose from various scientific organizations, including the Australian Academy of Science.

170 journals with impact factor in Applied Mathematics were subject to such appraisals. The results of the comparisons are quite eloquent.

It turned out that the impact factor values hardly corresponded to the evaluation of the experts. There were examples of journals that have a higher impact factor than others, but according to the expert assessment the values of the factor were found to be at a much lower level. Out of 10 journals with the highest values of the impact factor, only two journals have received the highest expert assessments. And the journal, which by its impact factor was considered the best in applied mathematics, according to expert assessment proved at a much lower level. This remarkable discrepancy between the impact factor of journals and expert assessment of their quality is usually explained by inaccuracies in the calculation and even deliberate falsification of the impact factor [Amin M. & M.Mabe, 2000, p. 3].

When the differences between the two evaluations of the validity and significance (usefulness) of created knowledge - formal bibliometric evaluation and substantial expert evaluation - go beyond a certain threshold, the normal, socially beneficial development of sciences themselves is blocked and deep distortions occur in the created knowledge.

Here are some more details about the risks of placing the performance of scientists, i.e. production of knowledge, under the strong influence of the above features of the publication sphere.

Promotion of Correct Scientists and Knowledge - "Conservative Ideologues", Networks and Self-Citations

The new system for performance evaluation of scientists gives special chance to the networks of 'conservative ideologues'. Studies of the new system often identify them as clusters (cartels) for self-citation and define them as follows: "Group of authors who have agreed on specific scientific or research methods, definitions or conclusions and cite only themselves or authors agreeing with them and ignore authors who disagree with their preferred methods, definitions or conclusions". These clusters are not simply mechanisms for trading on the deficiencies of a thoughtless and weak system of performance evaluation of scientists. They are much more dangerous phenomenon that has the potential to have a negative and a strong influence on the process of knowledge creation and on the result of this process - the knowledge itself. And there is evidence that they exert such influence and give rise to serious distortions in the structure of published knowledge, and through it - on the directions of scientific research.

Opportunities for such influences increase dramatically when the cartels establish close connections with (or conquer through their representatives) editorial boards and publishers of journals of 'high prestige in the publication market'. Over time, they inevitably begin to control not only publication, but the entire production of knowledge in a given field, directing it to the track of their group preferences.

For the increasing influence of self-citing cartels on the production of knowledge can be judged also by studies that show that the introduction of bibliometric evaluations intensifies the 'Matthew effect' – who follows the dominant 'fronts', receives the adequate reward.

Therefore, ultimately placing bibliometric indicators in the foundation of mandatory regular appraisals and evaluation of projects not only allows but also reinforces the capacity of organized networks (groups, clusters, cartels, cliques, etc., i.e. forms of organizational structures of 'conservative ideologues') to influence: the direction of research, dominant paradigms, scientific tools used to create knowledge, and therefore - to influence the process of knowledge creation, and hence - on the actual content and structure of the created knowledge. These are the outlines of a broad system for controlling knowledge and therefore - for its manipulation.

And this is a total failure in the creation of knowledge - control brings forth the change that was mentioned - institutionalized and specialized systems for the production of knowledge begin to produce indeterminacy and ignorance rather than knowledge. There are many signs for such dynamic.

The European Society of Life Sciences states that "the annually published impact factors of the referenced journals are averages based on many publications and publishing in a journal with a high impact factor does not guarantee that each publication is cited equally often". Critics of Hirsch index produce many data showing that the index stimulates the establishment of networks of authors who cite each other. But that is not all. Prof. Georgi Angelov emphasizes that a few years ago in the ranking of life sciences according to the indicator 'number of citations per one paper' Bermuda was at the top, followed by Panama, Gambia and Gabon at the leading places. This ranking may be contrary to common sense, but it shows what the rules for registration of citations are capable of. Moreover, even if the rules were perfect, it is evident that the ultimate judgment of the scientific relevance of performance can be manipulated in other ways. Therefore, deep doubts arise regarding the adequacy of registration rules and whether they reflect only the scientific relevance of certain research results or registration is influenced by other, poorly known (perhaps even - completely unknown) factors.

Professor Anne-Wil Harzing of the University of Melbourne focuses her analysis on the published by Thomson Reuters Highly Cited Papers List covering 1% most cited papers in a given discipline in a given period and illustrates the effect of the List with a case, which he called "super author". The latter has collected 512 citations, reflecting references to his work in 169 papers of other authors or networks of authors. Therefore, each paper that has cited the super author has cited at least three of his publications. Although repeated references in a paper to the work of one and the same author are not uncommon, more than three references seem pretty much. Even more interesting is that citing authors form narrow and highly self-referencing clusters. Moreover, the percentage of self-citation of the super author is also quite high - about 30%.

Professor Harzing wonders whether "this is a success story of a highly productive author or rather more complex and disturbing story of systemic impact of a series of 'innovative' and in a sense - abnormal decisions that seem to have the potential to change the very nature of the way in which scientists and the academic are perceived and evaluated?" The question is of course rhetorical. The truth is that this is a system that allows rapid scientific "successes" (respectively careers) regardless of the scientific validity and social relevance of the achieved scientific results. In the center of the system stands a consolidated network for self-citation. Harzing also points out the main components of the model of (self-) citations: a) **exceptionally high proportion of self-citation of journals** - 85% of the 512 super author citations are in the same journal in which he himself has published his works. As for his 10 most cited papers that percentage is 93. It is also quite interesting that his 10 most cited articles are not cited in other journals nor have only one citation; b) **exceptionally high proportion of self-citation of publishing companies**. 94% of the 169 papers citing the work of the super author are from the same publisher (*Academic Journals*). Only 10 citations were made in journals published by other publishers. Seven out of these 10 citations were made in a journal from the network ANSI Network's journals (another publisher with open access, which analysts consider "predatory"); c) **mutual citation authors - you cite me, I cite you**. Those 169 papers that cite the super author were cited in turn in other 165 papers. Only 7 papers out of those 165 do not cite papers by the super author. Thus, 158 of these 165 articles cite both the super author and the papers that have cited him.

It is especially important that ‘the cartel of citation’ managed to declare papers of the ‘super author’ an important line of research (a research front) and the effect is the resulting implicit requirement - if you want to be a successful scientist, follow the ‘front’ – either this one or some other that may be formed in the same manner.

Journals, Control and Deformation of Knowledge

It is often argued that the scientific quality of publications in journals with high impact factor is also very high because the results proposed there are subject to precise and highly qualified reviewing. We have seen that this statement can be deeply misleading and in many cases it really is. Indicators themselves really seem objective, but the factors that determine the values of these indicators are hardly ‘objective’ - they may be subject to targeted, ‘subjective’ effects i.e. to manipulation. Therefore, several observations suggest that, in fact, objective evaluation is not achieved and that achieving such an evaluation is difficult.

When the selection of material for publication is heavily dependent on “political, social and cultural elements", this simply means that the journals perform selection depending on the ideological, political and cultural predispositions of their editorial boards and reviewers they attract for their Peer Reviews. Hence, the judgment of the reviewers and editorial boards of journals in the U.S. is so distrusted that even the Supreme Court has registered this fact. But when the performance of scientists is evaluated according to its relevance to the ‘predispositions of publishers', the selection mechanism turns into a mechanism of control - censorship. Then the selection of materials accepted for publication will not depend on scientific validity and societal relevance of the results, but will depend on the ‘predisposition’ of selectors. This distorts the whole process of publication in which some results become public with advantages arising not from their scientific value but from other factors.

When this distortion of the publication process is combined with the forced inclusion of scientists in it (and the strong influence of successful inclusion on the evaluations scientists get and on their career), the result is clear - the work of scientists is ‘distorted’ and more precisely - the process of creating knowledge is distorted.

In general, distortions in the selection and dissemination of knowledge (publication process) become distortions in the production of knowledge - whoever wants to be published must produce results that meet certain extra-scientific requirements. Therefore, scientific journals (and the strong requirements to publish therein) can play a key role in the distortion of creating knowledge. And (especially in social sciences), the journals really play such a role, moreover, there is strong interest in the intentional distortion of knowledge about societies. How far this gap has gone, for example, between economic realities and dominant trends of economic analysis and the consequences thereof, can be seen very clearly in the economic crisis since 2008.

The Bibliometric ACTA

Obviously, the impact factor and other bibliometric indicators are not at all sound and ‘objective’ indicators for assessing the performance of scientists as it is claimed by their ‘promoters’. The latter are either ignorant or just cheat, pursuing certain goals of their own.

In fact, the above weaknesses in formation of bibliometric based evaluations indicate that the new system for performance evaluation of scientists has great potential to enhance control over the research and with this to cause an extremely negative trend for the substantive evaluation of scientific validity and relevance as well as of the social usefulness of scientific results to lose its significance and be replaced by indirect and subject to manipulation 'objective' indicators that can direct the creation of knowledge even at supranational level. As a result, scientists and scientific communities 'adapt' to the demands made on them (and take advantage of their weaknesses); research careers are less and less dependent on achievement and are more and more developed under the influence of surrogates, which in turn reflect the influence of other factors of non-scientific nature (mostly - expectations and predispositions of those who control the journals).

Sokal's hoax is not an isolated case and reveals the effect of a comprehensive mechanism for biasing the process of dissemination of knowledge.

It is not hard to notice three main tiers in the system of control over the production of knowledge. At the first tier (individual scientific institutions) a central role play the statutory systems for individual performance evaluation scientists – appraisals based mainly on bibliometric indicators related to their publication activity. The second tier comprises evaluation of the scientific organizations themselves. At this level governments (respectively - ministries) create a regulatory framework for a national system for performance evaluation of scientific institutions - universities and research centers. Both the national and the individual systems have the same pillars - bibliometric indicators. The third tier is supranational and covers all countries that have joined the system. At this level operate international scientific journals with various rating (especially those with high and very high impact factor). An important component of this level is a private rating agency, which gains tremendous opportunities to control the creation of knowledge as it ultimately decides what knowledge (which results) will receive the stamp of scientific validity and which will not be validated. Thus, a private corporation may determine the main directions in which knowledge will be created; the structure of this knowledge; the results that are acceptable or not – and against criteria only they are aware of.

Since practically they will issue certificates of scientific validity of research results and will control and direct the creation of knowledge, private rating corporations usurp (monopolize, 'privatize') a function with fundamental societal importance - the function to determine the development (direction of progress) of knowledge and the structure of general knowledge – the shares of individual sciences and the knowledge they create within this structure; the share of the different directions in various sciences and sub-disciplines and also - what will be the impact of different paradigms, theories, empirical data etc.

The three tiers are connected in a complete system through the same indicators. Since high scores at the first two levels (organizational and national) receive only those scientists and scientific organizations that have been approved at supranational level (i.e. publish in prestigious selected journals), it could be said that it is over-centralized system for controlling and directing the production of knowledge.

This system is essentially an analogue of the famous Anti-Counterfeiting Trade Agreement (ACTA) – a large scale system possessing capacity for control and restriction of freedom[1], but in our case – for control, direction and restriction of knowledge creation and especially – restriction of social knowledge. It is natural that such a system raises concerns.

Not only the components of the system described above raise concerns, but also the way in which it is being deployed - in parts, at the different levels, quietly, not described as a complete system, without an explicit formal

---

[1] The initiators of ACTA pretended to establish international standards for intellectual property rights enforcement. Additionally, this international legal framework aimed at **creation of a new governing body outside existing forums.**

document (as was the ACTA) that officially states: “we are setting a new system for evaluation and steering the development of science and knowledge, the reasons for its introduction are these, we expect to obtain the following effects of its introduction”. Instead, the system is developed silently, in the dark and the participants see only the parts of the whole, but do not link them together. At the ministry is developed a document for a national evaluation of research units (naturally - based on impact factors and citation indices); at universities and institutes is introduced a system of individual appraisal (again - impact factors and citation index) and at supranational level is created the mechanism that forms the indicators i.e. will issue certificates for proper science and correct knowledge.

Side Effects of the System - Deformations of the Publication Field

The bibliometric evaluations raise certain deformities in the publication field.

A) The emergence of "predatory journals”. One particularly deformed phenomenon that arose from a hypertrophied role of bibliometric-based evaluations is the emergence and rapid growth of ‘dark sector’ journals (*predatory journals*), whose main purpose is to profit from the pressure that is exerted on scientists to publish as much as possible in foreign journals. The system for performance evaluation of scientists and the results thereof, literally push scientists to these journals. A practice was born as a result, called "bait-and-switch" - scientists receive attractive invitations to publish in a journal, but then it becomes clear that they have to pay considerable sums for publication (reasons for asking for payment are different). Particularly easy victims are scientists from peripheral countries, where the pressure through requirements for publications in foreign (mostly - international and therefore ‘by definition’ - prestigious) journals is particularly strong. A typical example: “Nigerian scientists are particularly pleased with these invitations because the National Universities Commission (NUC) now require for promotion of lecturers to status of professors to publish some of their work in international academic journals”. And of course add that the requirement affects very badly the Nigerian academic community. The same could easily be written for the Bulgarian academic community and for any other community that is placed in similar circumstances.

Of course, these journals publish everything that has been duly paid for and are less concerned about the adequacy of the knowledge that comes through in their published articles. Scientists report the case of a teacher at the Benue State University, who in an interview with *The Guardian* (Guardian, July 28) announced that in an article published in a scientific journal he presented a solution to a 262 - year old math puzzle. As it turns out, the journal (its editorial board and reviewers) was not very concerned about the trustworthiness of the proposed solution.

However, even very sound journals that do not belong to the above group ask their authors to pay some amount to provide open access to their publications. This request has a very good reason and it can not be seen only as a ‘deviant behavior’ of journals. But this practice still shows how scientists are burdened with costs when publication activity becomes of too high importance for their valuation and becomes incumbent upon them. Moreover, as the number of citations depends on access, paying for open access ultimately affects the evaluation of the scientist. And this factor in the evaluation is definitely of no scientific nature. Naturally, the ‘most cited’ are the most solvent and solvent are those who have access to generous funding (projects). In turn, access to projects (Bulgarian experience is a good example) is often organized in a special way for special players, especially in the social sciences.

B) Scientific journals that are designed to meet the new requirements. Cases as the above are not found only in the journals - money making machines. The binding nature of ‘bibliometric evaluations’ in combination with their strong influence on the careers of researchers may give rise to similar effects in ‘normal’ journals designed for other purposes. For example, in the literature we find the case of Journal of Applied Pharmacy, published by

Intellectual Consortium of Drug Discovery & Technology Development from Saskatoon (Saskatchewan). The journal is not a money making machine, but was set up by Pakistanis living in Saskatoon to help Pakistani scientists to gain the necessary levels of ‘bibliometric indicators’ that are required by the Higher Education Commission in Pakistan. The journal became famous because of the case of a young and very active researcher who achieved very high ‘bibliometric values’ through articles published in that same journal. A review of her articles, however, showed extensive repetitions in "various articles" and "borrowings" whose sources were not indicated. In one of the articles the author in question claimed that some plants were not harmful and even had beneficial effect, but the problems in her own publications have led scientists to doubt whether her scientific ‘discoveries’ could be trusted. And what might have been the consequences if they were? Obviously, such practices can create a false scientific validity and relevance of the achievements of a scientist.

C) Prestigious journals and their rating. The main prerequisite for the occurrence of the above problems in publication of scientific results is that the impact factor of journals could be subjected to manipulation and, as shown by numerous observations, these opportunities are widely used. Even journals that do not fall into the above two categories and are considered prestigious use them - staggering facts of such falsifications have been found and at that for journals in the area of exact and life sciences. The above mentioned experiment conducted by the Australian Academy of Science is clear evidence of such deviations associated with the use of ‘objective’ bibliometric indicators in the field of mathematics. And what happens in the journals in the field of social sciences? The same, of course, but it is many times worse.

D) Forgetting the overall aim

The dissemination of research results through ‘high prestige journals’ is prone to adopt an increasing profit orientation that transforms the publication area into a large-scale private industry. The tendency is definitely in incompliance with the fact that many of the journals in one extent or another is supported by donations. This criticism is made by no other than *The Economist* [14 April 2012], which poses the question why funds from donations are used for the formation of a large private industry making profit. The criticism of "The Economist" highlights the important role played by private corporate body (e.g. Thomson Reuters Corporation) in the maintenance of database on registration (counting) of publications abroad and the citations networks.

Actually, *The Economist* points also to another, more important issue that ‘promoters’ of the performance evaluation system of scientists have forgotten - that the knowledge created in social sciences should be freely available to the public and not a commodity used for making huge profits because in this way it becomes difficult to access (R. Merton long ago paid attention to this already well-forgotten fact). Transformation of social knowledge into such commodity not only transforms the process of production of knowledge on society but also distorts the knowledge itself. Distorted knowledge raises inconceivable serious consequences because in its essence it is intentionally generated and maintained ignorance and non-knowledge.

The above stated explains why the striving to replace the system of direct and substantive evaluations of created knowledge with ‘bibliometric’ surrogates gives way also to the usual effect (and indicator of problems in science) - a growing number of scientific fraud. These have always existed, but in the recent decades they grew explosively. There is an abundance of cases like the one with A. Sokal or the teacher, who solved the 262 - year-old mathematical puzzle, or the mentioned assertive worker on the scientific front, lavishly publishing the inventions of the medicinal properties of various plants.[1] But an increase in fraud is still the lesser trouble - these are subject to relatively easy detection. The big trouble is another one – creation of social mechanisms for systematic control

[1] Disclosure of such fraud began to affect individuals with academic degrees from the highest ranks of political power – such scandals caused a sensation in Germany.

and distortion of knowledge continues - particularly in the social sciences, and these mechanisms are difficult to identify and the consequences are severe.

Social Sciences - a Particularly Vulnerable Field for Bibliometric Evaluation

Undoubtedly, the above-described system presents a serious threat to all sciences, to the knowledge created and to society that use that knowledge. This threat is particularly strong for social sciences and the knowledge created by them, because of their high vulnerability to such a system. And it is no chance that social sciences are considered the most inappropriate field for use of bibliometric evaluations - that has been emphasized both by scientific organizations, and individual researchers. As main reasons for social sciences'high vulnerability can be pointed as follows:

A) Unlike other sciences the knowledge generated in the social sciences can influence directly and strongly to social change and sustainability of the social order. Therefore, these sciences are exposed to particularly strong interest in the control of the created knowledge and control inevitably and automatically distorts the created knowledge. Thus emerged a specific model of development of social sciences, which may be called Preventive Model, because its nature is to adapt the creation of knowledge to requirements for stability of the social order by controlling and limiting the creation (and dissemination, transfer) of knowledge on certain aspects of social realities. The model covers all social sciences but it is especially noticeable in sociology, the most general social science. Preventive Model includes several components: the internal structure of science, the activity of specific groups of scientists, external influences on the process of knowledge creation (scientific policies turned the social sciences into Cinderella in the family of Sciences).

The traces of interventions through which the model was built are particularly evident in the internal structure of sociology (the set of paradigms, theories (present or missing), approaches, principles, sub-disciplinary structure, the rules of scientific work, including methods of validation (recognition of results of research), mechanisms for evaluation, and therefore - for the development of the scientific careers of researchers, etc.). These components comprising the science itself have gradually been constructed in such a way that science create knowledge maintaining the protective shell of non-knowledge about key aspects of social realities. Thus, through the very instruments for creating knowledge, knowledge has been actively restricted (ignorance supported) on the central aspects of societies.

B) Due to the above, in social sciences are particularly well developed and are unusually active groups of scientists whose primary role is to support the creation of "appropriate" knowledge, which stabilizes the social order. It is they who steer the mainstream of research and they are the main reason that made Norbert Elias to notice long ago that sociologist experiencing some 'inexplicable love' to the existing social order. Today these groups form particularly strong 'clusters (cartels) of citation' and are particularly clearly visible in the face of so-called 'think tanks'. No less dangerous form of their existence are also the informal networks in research organizations. Bibliometric evaluations are especially beneficial for these groups as they are able to not only act as cartels of citation, but to use the support of special scientific journals and external financing bodies. Academics, who do not belong to the network of cartels will be quickly forced to join in the game, otherwise they will end up with lower grades than the members of these structures;

C) Compared to other sciences, social sciences have less possibilities (approaches, criteria, procedures) for verification of the scientific validity of created knowledge and are therefore less 'protected' against major distortions thereof. This gives an unusual freedom of all 'guiding influences' - policies, publishing institutions, networks of cartels of citations etc.

D) The above characteristics have given rise to specific "Standard Model" of social research, spotted by Reynolds, Turner and many others. Steven Pinker also emphasized on the existence of standard model of social research and indicated the main effect of it: "Leading social researchers can say any nonsense as long as they conform to the standard model of the social sciences. .... hard to believe that the authors believe in what they say. Statements are made without regard to whether they are true. They are part of the Catechism of the century. ... Modern social comments remain based on archaic concepts ..." [Pinker, S. 1997; p. 57]. The model expands and 'normalizes' an effect, which the economist (and Nobel Prize winner) Herbert Simon had long ago noticed and discussed - the production of knowledge about secondary, minor issues i.e. irrelevant, useless, unnecessary, contradictory and uncertain knowledge (Note 2). Recently (April 19, 2012) *The Guardian* identified the same problem again. Looking through the site of the British Sociological Association, the author of the article found that there was not a single message related to the crisis in the EU and UK. Instead, the Association has published the 'fundamental results' from a study that "older bodybuilders can change the way young people perceive those who are older than 60 years." The author of the article has also checked the sites of three journals with quite high impact factor: American Sociological Review, Sociology (the leading sociological journal in the UK) and the British Journal of Sociology, having run a search on the keywords "finance", "economy" and "markets" for the entire last decade. In ten years, the first magazine has published 9 articles (another question is what exactly they contained) in the second were published 3 articles containing these keywords and in the third - one. In short, the three journals and have not been able to publish anything substantial on the issue that is essential for whole Europe. The central problem is referred to in the title of the article: "The Crisis is a Failure of Academic Elites".

Exactly here are the roots of other problems noted by risk analysis – the already mentioned generation of ignorance and the 'politicization' of knowledge; the falling behind of social sciences in comparison with others; the crisis of sociology; the emergence of a rift between the social realities and their scientific representations.

Therefore bibliometric evaluations are not able to restrict the creation of low quality, socially irrelevant, lacking in capacity for positive social functionality (even meaningless) knowledge. Moreover, such knowledge can be considered as knowledge with a reasonable degree of scientific validity and positive social relevance precisely because of the hypertrophied use of 'bibliometrics'. As pointed out by the said article into *The Guardian* - scientists embroider pieces for prestigious journals while a deep crisis shakes societies that scientists are called to study.

The above leads to the conclusion that bibliometric evaluations expand and institutionalize the Preventive Model for control over the created knowledge and therefore have the potential to further deteriorate the situation of social sciences. This in turn gives rise to unknown threats to society.

Conclusion: Threats to Societies

The widespread opinion that the financial and economic crisis is a failure of the academic elites was confirmed by the same academic elites on a special occasion, which quickly became known throughout the world. Visiting the London School of Economics the Queen of Britain asked why scientists have failed to foresee the crisis. Professors from the School failed to answer at the moment, but in the ensuing debates, one of them gave the obvious answer: "People do what they are paid for." This brings to the front the issue of the elaboration of science policies and research programs that fund research irrelevant to the most important risks to society.

The case is just a small example of a bigger trouble - these policies, together with the factors considered above have caused the 'discrepancy' between social research and knowledge on the one hand, and social realities (including the most acute problems of societies) on the other hand. This societal irrelevance and futility of the leading mainstream of social analysis is extremely dangerous phenomenon - it simply means that the connection with social realities is lost, i.e. societies do not have enough knowledge about themselves. And the loss of such a connection contributes to the stabilization of the social order, but on the other hand gives rise to decline in the rationality of societies, understood as the capacity of societies to identify risks in time and to set up adequate systems to neutralize them, to achieve development and to expand the boundaries of dignified human life. Therefore, the natural result of the decline in capacity to cope with risks is a blast of risks and the damage they cause to individuals and societies (Note 3).

Introduction of bibliometric evaluations in social sciences seems to be further enlargement of the control over sciences and knowledge and its effects will further aggravate the already very serious situation in the social sciences and the threats it poses to the societies themselves.

Therefore, it is the duty of the organizations of scientists, especially in the field of social sciences to warn societies about these dangers and make the necessary efforts to reduce them.

Notes

Note 1 The historical milestones that marked the recognition of the fact that control over knowledge is essential to control societies are well visible. Only a few names are sufficient to outline the gradual awareness of the potential practical use of control over knowledge to stabilize the social order: from Machiavelli (who clearly understood the importance of social differentiation of knowledge - the secret to maintain and exercise power); through Bulenvile (who expanded the idea, stressing the importance of control over knowledge to preserve power), Nietzsche (who was convinced that the "will to knowledge" is only a "will to power") to the explicit statement of social scientists published in scientific journal in the middle of the 20th century that "ignorance can be useful and potentially positive for maintenance of the social order" [Wilbert and Tumin, 1949].

Note 2 H. Simon gives the example of establishing a correlation between the number of unmarried older women in rural areas and yield clover crop. It was found that older unmarried women in rural areas often keep cats, cats hunt field mice and field mice feed on bumblebees that pollinate clover. Thus a larger number of older single women is associated with a larger number of cats, less mice and higher yield of clover seed. The conclusion was that possible decline in the production of clover must be assessed before decisions are made for payment of benefits on marriage or family allowances in rural areas. Simon points out that such knowledge should be rejected and restricted because it creates unnecessary 'noise' hampering and even misleading the making of adequate decisions [Simon, 1957].

Note 3 A typical examples: almost immediately prior to the crisis of 2008, *Citigroup* ordered a large-scale survey aimed to determine whether the ongoing concentration of income is not in any way a threat to the stability of the financial system. Researchers noticed that in many countries about 20% of the population receive a significant portion of the income and have a decisive influence on the dynamics of saving, investment, on the structure of consumer spending, and therefore – on the market and production, i.e. - on the whole economic dynamics. Yet the conclusion from these observations was that there was no danger to the banking business, and the researchers explicitly acknowledged that they were not guided by any moral judgments.

The financial crisis occurred shortly after the completion of the survey. When the property market crashed, venture securities suffered billions in losses - 27.7 billion for *Citigroup* and had to ask for 45 billion from the Federal Reserve. The corporation shares collapsed by 77% in one year. Shareholders suffered losses of 700 million dollars. In October 2007, one share of *Citigroup* was worth 47 dollars. But in 2009 the price was already 2 dollars. Because of the losses, the shareholders brought a claim and sentenced *Citigroup* to pay them 590 million dollars.

APPENDIX

EUROPEAN PHYSICAL SOCIETY RECOMMENDATIONS

On the use of bibliometric indices during assessment, V - 11 June 2012

Recent years have seen quantitative bibliometric indicators being increasingly used as a central element in the assessment of the performance of scientists, either individually or as groups, and as an important factor in evaluating and scoring research proposals. These indicators are varied, and include e.g. citation counts of individual papers published by researchers; the impact factors of the journals in which they publish; and measures that quantify personal research contributions over an extended period such as the Hirsch Hindex, and variants with corrections such as the G-index.

Although the use of such quantitative measures may be considered at first glance to introduce objectivity into assessment, the exclusive use of such indicators to measure science “quality” can cause severe bias in the assessment process when applied simplistically and without appropriate benchmarking to the research environment being considered. Funding agencies are aware of this, nevertheless experience shows that the reviewing of both individuals and projects on the national and European level is still relying excessively on the use of these numerical parameters in evaluation.

This is a problem of much concern in the scientific community, and there has been extensive debate and discussion worldwide on this topic (see for instance [ARC, 2010]).

Since the very first applications of bibliometric indicators in this way, scientists and science organizations have taken strong positions against such purely numerical assessment. Various organizations in Europe have published studies on their potential adverse consequences on the quality of funded scientific research. A prime example is the publication of the /Académie des Sciences of the Institute de France /that has presented clear recommendations on the correct use of bibliometric indices [Bibliometrie, 2011]. Other publications have addressed the role of peer review in the assessment of scientists and research projects e.g. the European Science Foundation /Peer Review Guide /published in 2011 [ESF, 2011] with recommendations for good practices in peer review following an extensive European survey on peer review practices [ESF, 2011a]. Other recent examples are a study of peer review in publications by the Scientific and Technology Committee of the House of Commons in the UK [STC, 2011], the peer review guide of the Research Information Network in the UK [RIN, 2010] and the recommendations formulated at a workshop dedicated to quality assessment in peer review of the Swedish Research Council [SRC, 2009].

A common conclusion of these studies is the recognition of the important role of PEER REVIEW in the quality assessment of research, and the recommendation to apply bibliometric performance indicators WITH GREAT CAUTION, and only by peers from the particular discipline being reviewed.

The European Physical Society recognizes and takes note of these recommendations for unbiased assessment procedures, and emphasizes in the following those aspects that are particularly important (in some cases unique) in the context of the assessment of the performance of the work of physicists, and of the quality and originality of physics research projects.

1. Evaluation should exclusively be carried out by peers, who must be independent and must have no conflict of interest with the evaluation process. They must strictly respect a published code of conduct. Whilst recognizing the role of confidentiality in some forms of peer review, the names of evaluators should normally be made public, either before or after the assessment procedure as appropriate to the evaluation being carried out.

2. An unbiased assessment of the scientific quality of individual researchers or their projects using bibliometric indices must take into account many factors such as: the scientific content; the size of the research community; the economic and administrative context; and publishing traditions in the field. Publishing habits and traditions

significantly vary between different fields of physics research, and are reflected for example in areas such as the name order in the list of authors and the particular choice of the journals in which to publish. A special example is publishing in the field of physics with large facilities where traditions are very different from many other fields. For example, accelerator physicists publish their work essentially in conference proceedings, while only a small percentage of their work appears in peer-reviewed journals. Another example is the publication policy of the large collaborations of physicists in the field of experimental particle and astroparticle physics. These collaborations apply strict procedures for the assessment and endorsement of results by every member of the collaboration prior to the internal publication of results. The external publication of results is also endorsed by the full collaboration. As a consequence of this policy, their articles in refereed journals often have long author lists published uniquely in alphabetical order.

3. The annually-published impact factors of refereed journals are averaged over many papers, and publishing in a high impact journal does not guarantee that every individual article is equally highly cited. Such quantitative measures based on the number of publications and/or citation statistics of researchers are one aspect of assessment, but they cannot and must not replace a broader review of researchers' activities carried out by peers.

The European Physical Society, in its role to promote physics and physicists, strongly recommends that best practices are used in all evaluation procedures applied to individual researchers in physics, as well as in the evaluation of their research proposals and projects.

In particular, the European Physical Society considers it essential that the use of bibliometric indices is always complemented by a broader assessment of scientific content taking into account the research environment, to be carried out by peers in the framework of a clear code of conduct.

Bibliography

[Adler, Ewing, Taylor, 2009] Adler, R., J. Ewing, and P. Taylor, Citation statistics. Statistical Sciences 24, 2009, pp. 1-14.

[Amin M. & M.Mabe, 2000] M. Amin & M. Mabe, Impact factors: use and abuse. Perspectives in Publishing 1, pp. 1-6, Arnold, D., Kristine Fowler. Nefarious Numbers. arXiv.org, Cornell University Library, 2000.

[ARC, 2010] Australian Research Council, Ranked Journal List Development, 2010, http://www.arc.gov.au/era/journal_list_dev.htm.

[Begley, S. 2006] Begley, S., Science journals artfully try to boost their rankings. Wall Street Journal, 5 June 2006, B1.

[Bibliometrie, 2011] Institut de France, Académie des Sciences, Du Bon Usage de la Bibliometrie pour l'Évaluation Individuelle des Chercheurs, 17 January 2011, http://www.academie-sciences.fr/activite/rapport/avis170111gb.pdf - English version.

[Bucchi, 2004] M. Bucchi, Science in Society. An Introduction to Social Studies of Science. Routledge, 2004.

[ESF, 2011] European Science Foundation, European Peer Review Guide, Integrating Policies and Practices for Coherent Procedures, March 2011, http://www.esf.org/activities/mo-fora/peer-review.html.

[ESF, 2011a] European Science Foundation, Survey Analysis Report on Peer Review Practices, March 2011, http://www.esf.org/activities/mo-fora/peer-review.html

[Ewing, J.] Ewing, J. Measuring journals. Notices of the AMS

[Funtowicz & Ravetz, 1992] S. Funtowicz & and J. Ravetz, Three Types of Risk Assessment and the Emergence of Post-Normal Science. In: Sheldon Krimsky and Dominic Golding, Social Theories of Risk, Praeger Publishers, 1992.

[Golubic, R., M. Rudes, N. Kovacic, M. Marusic, and A. Marusic, 2008] Golubic, R. M. Rudes, N. Kovacic, M. Marusic, and A. Marusic, Calculating impact factor: how bibliographical classification of journal items affects the impact factor of large and small journals. Sci. Eng. Ethics 14 (2008), 41-49.

[Grant J. 2007] Grant J. Corrupted Science. Fraud, ideology and politics in science. AAPPL, Church Farm House, Wisley, Surrey, 2007

[Gross M. 2000] Gross M. The Unknown in Process. Dynamic Connections of Ignorance, Non-Knowledge and Related Concepts, Canadian Journal of Sociology/Cahiers Canadiens de sociologie, Vol. 25, N2, 2000, Spring.

[Gross, M. 2010] Gross, M. 2010. Ignorance Surprise. Science, Society and Ecological Design. MIT.

[Harzing, Anne-Wil] Harzing, Anne-Wil; Web: http://www.harzing.com/index.htm

[Kasperson, R. 1992] Kasperson, R. The Social Amplification of Risk: Progress in Developing an Integrative Framework, In: S. Krimsky and D. Golding (eds) Social Theories of Risk, Praeger, 1992.

[Lyng, 2008] Stephen Lyng, Edgework, Risk, and Uncertainty. In: Zinn, Jens O (ed.) Risk and Uncertainty. An Introduction. Blackwell Publishing, 2008

[Macdonald S., J. Kam, Aardvark et al, 2007] Macdonald S., J. Kam, Aardvark et al.: Quality journals and gamesmanship in management studies. Journal of Information Science 33 (2007), 702-717.

[Monastersky, R., 2005] Monastersky, R., The number that's devouring science. Chronicle of Higher Education 52 (2005). ScienceWatch.com, April 2008, http://sciencewatch.com/inter/aut/2008/08-apr/08aprHe/

[Pinker, S. 1997] St. Pinker, How the Mind Works, Penguin Books, 1997

[RIN, 2010] Peer review, A guide for researchers, Research Information Network, UK, March 2010, http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/peer-review-guideresearchers

[Seglen, P., 1997] Seglen, P.O. Why the impact factor of journals should not be used for evaluating research. BMJ 314 (1997), 498-502.

[Simon, 1957] H. A. Simon, Administrative Behavior. A Study of Decision Making Process in Administrative Organization. The Macmillan Company, New York, 1957

[SRC, 2009] Swedish Research Council, Quality Assessment in Peer Review, 5 November 2009, www.cm.se/webbshop_vr/pdfer/2011_01L.pdf

[STC, 2011] Peer review in scientific publications, Science and Technology Committee, House of Commons, UK, 18 July 2011, http://www.publications.parliament.uk/pa/cm201012/cmselect/cmsctech/856/85602.htm

[Strathern, M.1997] Strathern, M., "'Improving Ratings': Audit in the British University System", European Review 5 (1997), 305-321.

[Wiener, N.1947] N. Wiener 1947, Open letter by N. Wiener published in the Bulletin of the Atomic Scientists, Vol. 3, N1, Jan. 1947. Quoted after Jean-Jacques Salomon, Science et politique, ed. Du Seuil, 1970, p. 308-309.

[Wilbert and Tumin, 1949] E. Wilbert E. & M. Tumin, „Some Social Functions of Ignorance", American Sociological Review, 14 (6); 787-795, 1949, cited by G. Matthias, Ignorance Surprise. Science, Society and Ecological Design. MIT, 2010

Authors' Information

Douhomir Minev *– Prof., chief of department "Public Policies and Social Changes" at the Institute for the Study of Societies and Knowledge, BAS, 3 A "Moskovska" str., Sofia-1000, Bulgaria; e-mail: perspekt@tradel.net*

Major Fields of Scientific Research: Sociology of Knowledge, Economic Sociology, Social Policy

TOWARDS A NOVEL DESIGN OF CREATIVE SYSTEMS

Vladimir Jotsov

Abstract: *The topic of the presented investigation is the automation of creative processes via one Synthetic Mata Method (SMM) and few analytic methods considered in the paper and used under SMM control. The prevention of contemporary web threats is discussed at an agent/application level. Advantages and disadvantages of synthetic data mining methods are investigated, and obstacles are revealed to their application in contemporary systems. Novel results for juxtaposing statistical vs. logical data mining methods aiming at possible evolutionary fusions are described. Recommendations are made on how to build more effective applications of classical and/or presented novel (meta) methods: SMM, KALEIDOSCOPE, FUNNEL, PUZZLE, and CONTRADICTION. The usage of ontologies is investigated with the purpose of information transfer by sense. Practical aspects of agent applications, intrusion detection, intrusion prevention, cryptography applications, multiple software and other research results are mentioned aiming to show that intelligent and classical technologies should be carefully combined in one software/hardware complex to achieve the creative goals. It is shown that all the demonstrated advantages may be combined with other known methods and technologies.*

Keywords: *automation of creative processes, human-machine brainstorming methods, knowledge discovery, data mining, web mining, ontology, information security systems, intrusion detection, intrusion prevention, human-centered systems, knowledge management, agent, collective evolution*.

1. Introduction

There exist a plenty of philosophy - or psychology-based research on creative systems [1] but the technically-oriented papers like [2] are rather scarce. On the other hand, creative systems are welcome to the field where contemporary statistical, data mining, web mining systems are still not efficiently used, especially in the dynamic environments of information security systems. Methods which in my opinion are essential to every creative system or to next generation knowledge discovery system are discussed. One of them, a conflict resolution method, allows the agent or other application element to be really autonomous or reinforces the possibilities of the contemporary systems.

Contemporary Information Security Systems (ISS) and especially the web-based systems represent a wide field for applications of modern creative methods and technologies. The need to create sufficiently effective and universal tools to protect computer resources grows every year in systems for detection and prevention from intrusions (Intrusion Detection Systems IDS, Intrusion Prevention Systems IPS). For this reason different applications of intelligent data processing are initiated based on a combination of methods from statistical and logical information processing [3, 4]. Other elaborations with growing influence in this domain are artificial immune systems and multiagent systems [5, 6]. The unifying factor is the longer life cycle, elaborations require bigger teams and time for introduction. Due to the complex structure the prevention from direct attacks against these systems is very challenging.

Modern applications of statistical methods are effective and convenient to use at the expense of information encapsulation. In other words, it is impossible to construct tools to acquire new knowledge or to solve other problems of logical nature in this area. If we split methods in two groups (quantitative and qualitative) then statistical methods belong to the first group and logical ones belong to the second group. For this reason, their mechanical union is of no perspective. We do not attempt to propose any isolated solutions; instead we offer a combination of novel methods that is well adjustable to the existing ones. Our research includes a new evolutionary meta-method for joint control of statistical and logical methods where the statistical approach is widely applied on the initial stage of the research when the information about the problem is scanty and it is possible to choose the solution arbitrarily [7]. The accumulation of knowledge makes logical applications more and more effective and more universal than the probabilistic ones, as well as fuzzy estimates and similar applications. The paper introduces the SMM (Synthetic Meta-Method) meta-method to control the process of consecutive replacement of applications by other ones and is synthetic by nature. The difference from the classical analytic methods is in the fact that the design of systems controlled by synthetic (meta) methods is not just science, sometimes it is an art. If we make an analogy with the set of traditional methods and fashion clothes then the synthetic method will apply the design of the display window with the fashion clothes. In the common case during intelligent data processing, there is no convergence of the results but this does not hamper practical applications of these systems. In other words, bad and good designers will arrange the display window in quite different ways and there is no guarantee that every user understands the technology and that his access to the system will have positive results.

The cited innovations are made and demonstrated here for the following reason. The problems presented above show that there is a need to introduce elements of machine creative work in ISS. It is demonstrated in the paper that this goal is accessible, if the usage of possibilities for human-machine contact and a set of comparatively simple intelligent technologies are done in the right way. On the other hand, the innovations can hardly be described in a single paper using the traditional academic style. For this reason, in the paper we avoided when possible the technical details and formalizations, and for the sake of the contents reduction descriptions by analogy are used, illustration visual aids and other nonstandard approaches.

How can security agents operate autonomously? In the first place this is because of the usage of neural networks where the agent most conscientiously copies the acts of the teacher. At that the agent itself does not understand the sense of teacher's transfer of knowledge; it operates as universal approximator instead. It is shown schematically at Fig. 1 in the following way.

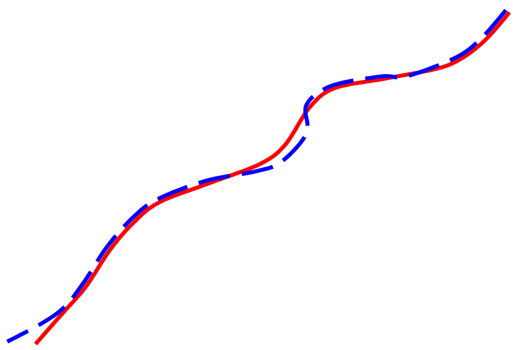


Figure 1. Approximation of teacher's activities

Teacher's acts are principally presented as a dotted curve. The neural network approximates this curve via a continuous line on the same figure; the deviation between the two lines must not exceed 3%. One of the main disadvantages of this approach is related to the necessity the agents in ISS to apply the learned knowledge in a rather creative way because frequently they operate in unexpected by the teacher situations. And as they are poorly trained or they are not at all trained how to act in unexpected situations, this approach as a whole is not very effective in its classical appearance. The presented approaches in anomaly ISS are combined with statistical applications which count for the normal traffic and other mean statistical values related to the operation of the guarded place. In Fig. 2 this process is schematically shown in the following way.

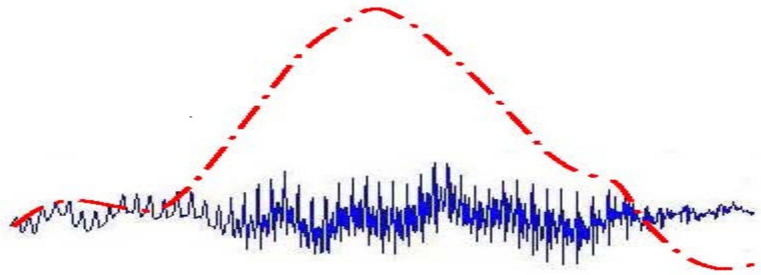


Figure 2. Anomaly detection

The normal traffic is presented via a continuous line and the dotted line shows a case with drastic deviation from the normal traffic which is considered as an anomaly and it is analyzed whether it is a consequence from an intrusion. One of the disadvantages of this approach lies in the fact that intruder(s) know the principles of operation of such systems and they can ‘fit’ in the constraints of the normal traffic.

There is a variety of other applied approaches when the system is overloaded by heuristic information but they are not discussed here due to their evident weakness for ISS.

Here we introduce a new way for agents’ operation. Best of all is the usage of ontologies to model the domain but this is not obligatory. On the other hand, as it is shown in the next section, the multiagent system functions more effectively if a system of ontologies is included.

But ontologies also do not contribute to a great extent in order to understand the sense of matters by the agents or with respect to transmit the sense of matters during communications between the agents. On the way to produce an analogy to how agents think we offer evolutionary methods to process data or knowledge because thinking (and understanding) is an evolutionary process. The problem is how to direct evolutionary methods with no exaggerations of heuristics, statistical information and other relative methods. Our approach is rather untraditional. We have elaborated for more than twenty years methods to detect and to resolve contradictions. A method for machine learning is built based on them. Searching and solving conflicts and contradictions the agent improves its knowledge and at the same time it may solve other problems. Detection of contradictions is based on using models of contradictions that can also be improved gradually. The agent may request an external help to solve the conflicts but this takes place only in extraordinary situations. At the same time it is shown how to change the reasoning component of security agents. Different logical methods are used that are rather analogous to means-ends analysis, constraint satisfaction, variable fitness function, brainstorming, and cognitive graphics. The combination of new methods to a great extent mechanizes creative efforts and also it serves agents’ operation to improve the abilities of security experts working with similar types of systems.

Suggested innovations serve the more effective application of data mining, Web mining, collective evolutionary components in multiagent systems. They are very well combined especially for applications of evolutionary approaches with classical neuro-fuzzy, statistical applications, genetic algorithms, and etc. methods for the domain. For example such systems critically accept teacher’s acts in cases of supervised learning (Fig. 3): they may precise or argue teacher’s acts and in this way they can learn more effectively and deep.

A wide application of intelligent agents is forecasted in the field of information security systems. This will lead to the situations when the agent has no possibility to learn from the expert (teacher) but should swiftly learn from other agents or should self-learn without teacher. Then the role of the above considered critical learning will significantly increase.

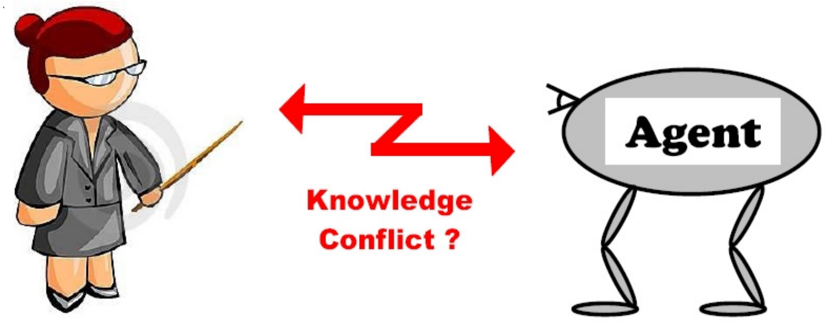


Figure 3. Example of supervised learning by using critical analysis

2. Synthetic Methods: a Necessary Background to Future Creative Systems

A synthetic meta-method (SMM) is elaborated and applied aiming at application of a set of ‘creative’ elements in agent environments. They work most effectively as a system, but even particular elements of them, let’s say, applied in semantic reasoners, are proved to be very useful. Their principles are easy-to-be explained: bind the unknown knowledge from the goal with the knowledge from the knowledge base, apply a flexible contraint system to manage a system of variable fitness/goal functions, make the goals automatically via self-improvement of the existing knowledge, etc. The consize goals are better applied in the intelligent agents. The set of methods necessary to every creative system is considered in this Section. The coordination activities under the control of SMM is discussed in Section 3. Ontologies and other applications are represented in Sections 4 and 5.

2.1 PUZZLE Method

The basic methods of the suggested meta-method SMM are presented below. Let the constraints connected to the defined problem form a line in the space described by the equation (1).

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1} \quad (1)$$

For example if a bachelor who has graduated our university SULSIT lives in Sofia and he/she does not want to work anywhere else, then the line restricts the search space and in this way a lot of unnecessary work is avoided. It is also possible to inspect a case when the constraint is defined as a type of surface but as a result a more general solution is obtained where a special interest is provoked by the boundary case of the crossing of two or more surfaces. When the common case is inspected in details, then in the majority of cases the problem is reduced to exploring the lines of type (1) or to curves with complicated forms obtained as a result of crossing surfaces by constraints. Therefore, below we investigate the usage in systems of constraints by lines of first or higher orders.

If the mentioned curves have common points of intersection and if they lie in a common plain so that a closed figure (triangle, tetragon, etc.) is formed then the search space is significantly reduced and it is searched much easier. The practical usage of the classical technology, as well as the constraint satisfaction, is complicated by the following. The viewed plains are not only nonlinear in the common case but they also include fuzzy estimates. The usage of fuzzy logic significantly raises the algorithmic complexity of the problem and it can make the application ineffective. Even when the usage of constraints significantly reduces the number of the inspected solutions, for example up to 10, this does not mean that the problem is solved and that all that must be done is to explore the possibilities one by one.

Let's admit that $k_i \in K$ are the elements of the solution, or, same, of the objective goal G, and $w(k_i)$ are the weights of the elements on the condition that the weight of the goal equals to one. In different situations a concrete measure is chosen and that is why the order of weights' processing is different. A popular example is measuring the length to the goal and then $w(k_i)$ equals to the length of the piece k_i from the segment G. Then S from the next formula shows the reliability of the inference 'G is true'. If $k_i \in K$ are continuous values then there integral calculus must be applied:

$$S = \frac{\sum_i k_i}{1 - \sum_i k_i}$$

It follows from the above formula that the bigger part of the goal G is confirmed by the knowledge, the easier is the goal confirmation. The detection of point of intersection between the available knowledge and the unknown goal uses an original method of the PUZZLE that binds the unknown knowledge (the goal) to the known knowledge (the accumulated up to the moment information).

The following example below shows how the search process is reduced via using ontologies. Let's admit that the search space is presented on fig. 4 where statistical data about ISS are generalized about the regions depending on their price and quality. It is necessary to select an acceptable ISS to our project.

In fig. 5 a subset of feasible solutions is chosen without ISS designed outside Europe. The space of feasible solutions is to the left of the separating surface that is depicted on the figure in blue color.

In fig. 6 another surface in green is shown delimiting the search space of the solutions. In our case it is 'systems with unknown principles of operation'. It is accepted that in the data bases there is no clear distinction related to the presented criteria so the search of the feasible solutions is nonlinear of high dimensionality and practically it cannot be solved using traditional methods. Nevertheless, by applying ontologies analogous to the ones from the previous section the problem is solvable via the PUZZLE method.

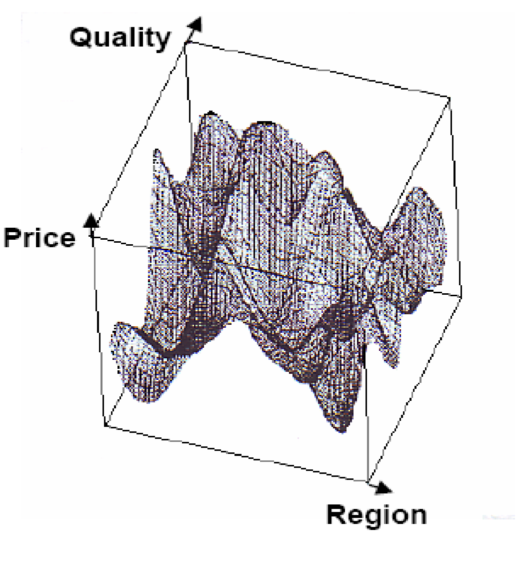


Figure 4. Example of a space of solutions

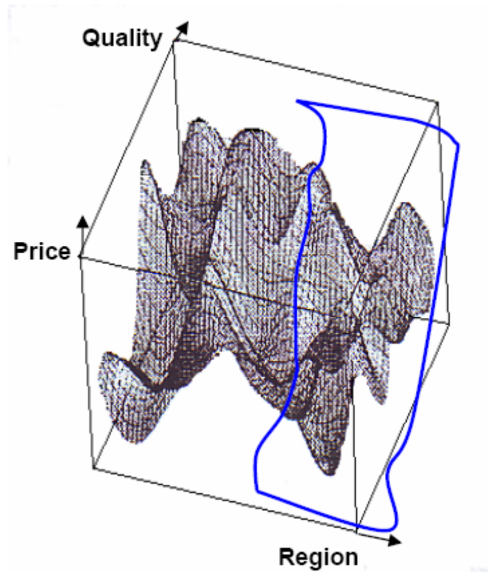


Figure 5. Nonlinear space division(s) of the region

There are two red dots on the same fig. 6 in its left corner. Each of them is a kind of constraint but of another type which we name a binding constraint and it is introduced by us. Its semantics is the following: it is not a solution but it resides close to the searched solution. For example we have the information that Fensel’s elaborations are a good solution to the problem and that they define the left dot; the right dot has semantics of some other type. By introducing new constraints, our goal is to show that it is possible to use causal links that are different from implications.

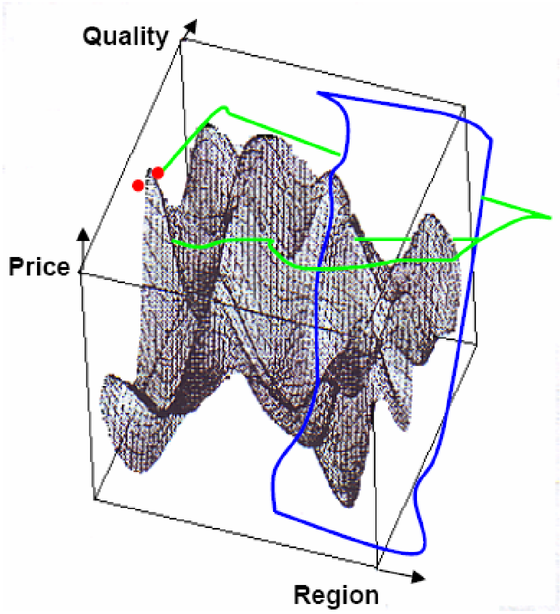


Figure 6. Binding and other constraints

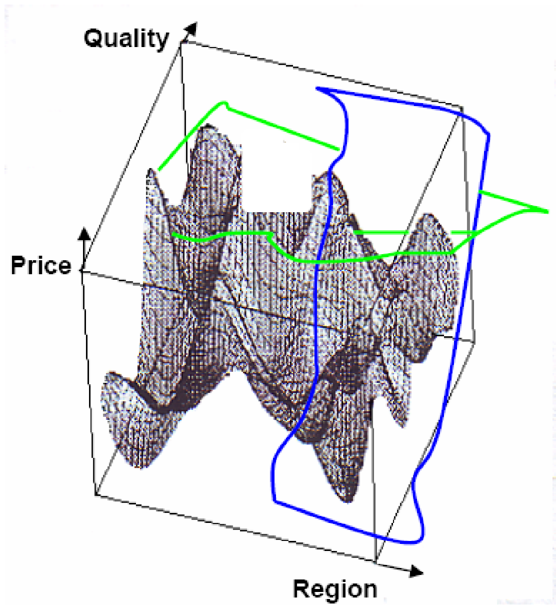


Figure 7. Two nonlinear intersections best executed by using ontologies

The same situation is presented on fig. 7 but some of the solutions are absent and this is evident in comparison to the images from the previous figures. It will be demonstrated below that the pointed incompletenesses are often met often and, even in this situation which is an obstacle for other existing methods, we offer an effective solution. In Section 3 it is discussed that the nonlinear surfaces could be efficiently represent by ontologies.

2.2. FUNNEL Method

Below, we discuss in brief the next proposed FUNNEL method. Fig. 8 presents the main elements of the method: a system of constraints in the form of a funnel around a center which is a goal (fitness) function which points to the desired direction for information output or to search for new knowledge.

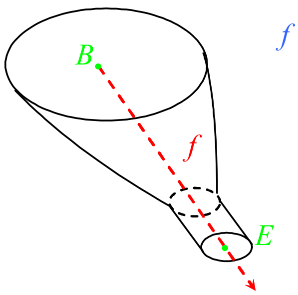


Figure 8. A funnel-type system of constraints

As it is evident, the goal of this method is the gradual narrowing of the space of the feasible solutions, together with the progress of the dynamic information processes. Usually the FUNNEL method operates properly when combined with the other methods introduced here and that is why its peculiarities are viewed in detail in the next section where the interactions between the methods are examined. For example, it is convenient to concentrate on fig. 7 shown above over one of the peek values of the diagram by fixing a funnel above it.

2.3. Conflict Resolution Method and its Machine Learning Applications

Any lack of collaboration in a group of agents or intrusion could be found as an information conflict with existing models. Many methods exist where a model is given and every non-matching it knowledge is assumed as contradictory. Let's say, in an anomaly intrusion detection system, if the traffic has been increased, it is a contradiction to the existing statistical data and an intrusion alert has been issued. The considered approach is to discover and trace different logical connections to reveal and resolve conflict information. The constant inconsistency resolution process gradually improves the system DB and KB, and leads to better intrusion detection and prevention. Models for conflicts are introduced and used, and they represent different forms of ontologies.

Let the strong (classical) negation be denoted by '¬' and the weak (conditional, paraconsistent) negation [8, 9, 10] be '~'. In the case of an evident conflict (inconsistency) between the knowledge and its ultimate form–the contradiction–the conflict situation is determined by the direct comparison of the two statements (the conflicting sides) that differ one from another by just a definite number of symbols '¬' or '~'. For example: A and ¬A; B and not B (using ¬ equivalent to 'not'), etc.

In the case of implicit (or hidden) negation between two statements, A and B can be recognized only by an analysis of preset models of the type of (2)

$$\{U\}[\eta: A, B], \quad (2)$$

where η is a type of negation, U is a statement with a validity including the validities of the concepts A and B, and it is possible that more than two conflicting sides may be present. It is accepted below that the contents in the figure in brackets U is called unifying feature. In this way, it is possible to formalize not only the features that separate the conflicting sides but also the unifying concepts joining the sides. For example, the intelligent detection may be either automated or of a human-machine type but the conflict cannot be recognized without the investigation of the following model

{detection procedures}[¬: automatic, interactive].

The formula (2) formalizes a model of the conflict the sides of which unconditially negate each another. In the majority of the situations, the sides participate in the conflict only under definite conditions: $\chi_1, \chi_2, \ldots\chi_z$.

$$\{U\}[\eta: A_1,A_2,\ldots A_p] <\tilde{\chi}_1 * \tilde{\chi}_2 * \ldots * \tilde{\chi}_z>, \quad (3)$$

where χ~ is a literal of χ, i.e. $\chi\sim \equiv\chi$ or $\chi\sim \equiv\eta\chi$, * is the logical operation of conjunction, disjunction or implication.

The present research allows a transition from models of contradictions to ontologies [11] in order to develop new methods for revealing and resolving contradictions, and also to expand the basis for cooperation with the Semantic Web community and with other research groups. This is the way to consider the suggested models from (2) or (3) as one of the forms of static ontologies.

The following factors have been investigated:

T – time factor: non-simultaneous events do not bear a contradiction;

M – place factor: events that have taken place not at the same place, do not bear a contradiction. In this case, the concept of place may be expanded up to a coincidence or to differences in possible worlds;

N – a disproportion of concepts emits a contradiction. For example, if one of the parts of the contradiction is a small object and the investigated object is very large, then and only then it is the case of a contradiction;

O – identical object. If the parts of the contradiction are referred to different objects, then there is no contradiction;

P – the feature should be the same. If the parts of the contradiction are referred to different features, then there is no contradiction;

S – Simplification factor. If the logic of user actions is executed in a sophisticated manner, then there is a contradiction;

W – Mode factor. For example, if the algorithms are applied in different modes, then there is no contradiction;

MO – contradiction to the model. The contradiction exists if and only if (*iff*) at least one of the measured parameters does not correspond to the meaning from the model. For example, the traffic is bigger than the maximal value from the model.

Example: We must isolate errors that are done due to lack of attention from tendentious faults. In this case we introduce the following model (4):

$$\{\text{ user : faults }\}[\sim\text{: accidental, tendentious }] <T,\neg M,O;\neg S>. \quad (4)$$

It is possible that the same person does sometimes accidental errors and in other cases tendentious faults; these failures must not be simultaneous on different places and must not be done by same person. On the other hand, if there are multiple errors (e.g. more than three) in short intervals of time (e.g. 10 minutes), for example, during authentications or in various subprograms of the security software, then we have a case of a violation, nor a series of accidental errors. In this way, it is possible to apply comparisons, juxtapositions and other logical operations to form security policies thereof.

Recently we shifted conflict or contradiction models with ontologies that give us the possibility to apply new resolution methods. For pity, the common game theoretic form of conflict detection and resolution is usually heuristic-driven and too complex. We concentrate on the ultimate conflict resolution forms using contradictions. For the sake of brevity, the resolution groups of methods are described schematically.

The conflict recognition is followed by its resolution. The schemes of different groups of resolution methods have been presented in Fig. 9 to Fig. 12.

In situations from Fig. 9, one of the conflicting sides does not belong to the considered research space. Hence, the conflict may be not be immediately resolved; only a conflict warning is to be issued in the future. Let’s say, if we are looking for an intrusion attack, and side 2 matches printing problems, then the system could avoid the resolution of this problem. This conflict is not necessary to be resolved automatically; experts may resolve it later using the saved information. In Fig. 10, a situation is depicted where the conflict is resolvable by stepping out from the conflict area. This type of resolution is frequently used in multi-agent systems where conflicting sides step back to the pre-conflict positions and one or both try to avoid the conflict area. In this case a warning on the conflict situation has been issued.

The situation from Fig. 11 is automatically resolvable without issuing a warning message. Both sides have different priorities, say side 1 is introduced by a security expert, and side 2 is introduced by a non-specialist. In this case, side 2 has been removed immediately. A situation is depicted on Fig. 12 where both sides have been derived by an inference machine, say by using deduction. In this case, the origin for the conflict could be traced, and the process is using different human-machine interaction methods.

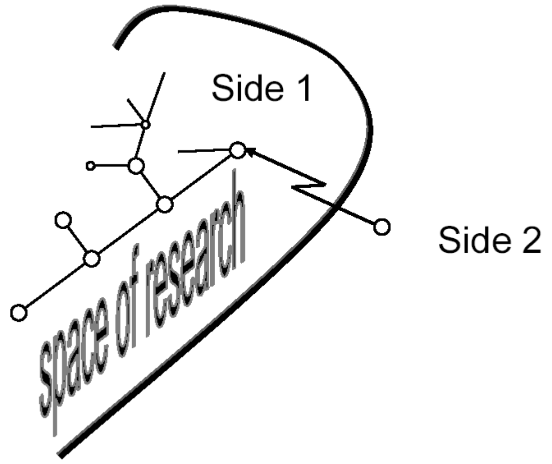


Figure 9. Avoidable (postponed) conflicts when Side 2 is outside of the research space.

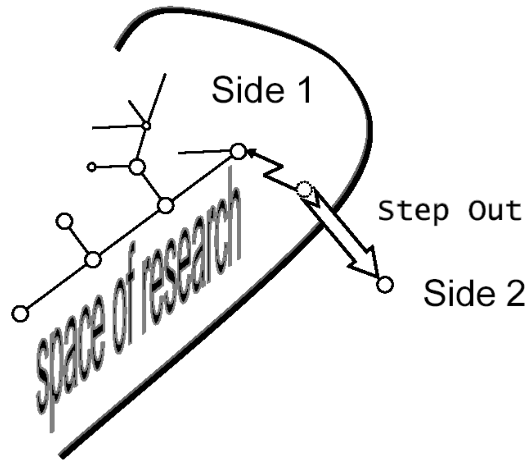


Figure 10. Conflict resolution by stepping out of the research space (postponed or resolved conflicts).

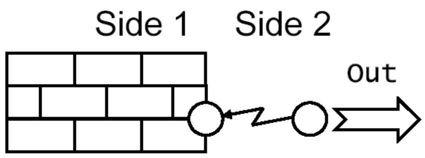


Figure 11. Automatically resolvable conflicts

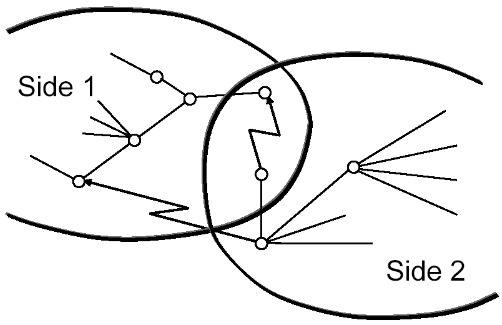


Figure 12. Conflicts resolvable using human-machine interaction

Knowledge bases (KBs) are improved after isolating and resolving contradictions in the following way. One set is replaced by another while other knowledge is supplemented or specified. The indicated processes are not directed by the elaborator or by the user. The system functions autonomously and it requires only a preliminary input of models and the periodical updates of strategies for resolving contradictions. Competitions to the stated method may be methods for machine supervised – or unsupervised – learning. During supervised learning, for example by using artificial neural networks, training is a long, complicated, and expensive process, and the results from the applications outside the investigated matter are unreliable. The 'blind' reproduction of teacher's actions is not effective and it has no good prospects except in cases when it is combined with other unsupervised methods. In cases of unsupervised training via artificial neural networks the system is overloaded by heuristic information and algorithms for processing heuristics, and it cannot be treated as autonomous. The presented method contains autonomous unsupervised learning based on the doubt-about-everything principle or on the doubt-about-a-subset-of-knowledge principle. The contradiction-detecting procedure can be resident; it is convenient to use computer resources except for peak hours of operation.

The unsupervised procedure consists of three basic steps. During the first step, the contradiction is detected using models from (2) to (4). During the second step, the contradiction is resolved using one of the resolution schemes presented above, depending on the type of conflict situation. As a result from the undertaken actions, after the second stage the set K is transformed into K' where it is possible to eliminate from K the subset of incorrect knowledge $W \subseteq K$, to correct the subset of knowledge with an incomplete description of the object domain $I \subseteq K$, to add a subset of new knowledge for specification $U \subseteq K$. The latter of cited subsets includes postponed problems, knowledge with a possible discrepancy of the expert estimates (problematic knowledge), and other knowledge for future research which is detected based on the heuristic information.

In cases of ontologies, metaknowledge or other sophisticated forms of management strategies, the elimination of knowledge and the completion of KBs becomes a non-trivial problem. For this reason the concepts of orchestration and choreography of ontologies are introduced in the Semantic Web and especially for WSMO [12, 13]. The elimination of at least one of the relations inside the knowledge can lead to discrepancies in one or in several subsets of knowledge in *K*. That is why after the presented second stage, and on the third stage, a check-up of relations is performed including elimination of modified knowledge and the new knowledge from subsets *W*, *N*, *I*, *U* are tested for non-discrepancies via an above described procedure. After the successful finish of the process a new set of knowledge *K’* is formed that is more qualitative than that in *K*; according to this criterion it is a result from a machine unsupervised learning managed by models of contradictions defined a priori and by the managing strategies with or without the use of metaknowledge.

2.4. KALEIDOSCOPE Method

The next SMM method is called KALEIDOSCOPE [6]. It is used to visualize the results. The difference from other analogical methods is the usage of cognitive approaches to human-machine communication. Similarly to the classical caleidoscope, the machine makes things that it performs better than humans: it arranges knowledge, it searches repeated elements and other regularities in the data, it presents the results to the humans. It is the user that makes the results creative and interesting for other researchers in the domain; the machine just assists him/her analogously to the caleidoscope which cannot estimate the beauty and the interest in the drawings produced during rotations and the arbitrary movements of the glass pieces inside of it. In both cases the human estimates the results and the developers apply methods to attract the user’s attention and to make the work with the system equipped with elements from games, less boring and consequently more effective. The below application is from Number Theory, it is used to reveal the way of automatic detection of regularities, their research, automatic or manual, leads to very deep theoretic knowledge.

KALEIDOSCOPE consists of a visualization approach, a method for information transfer-by-sense and applied approaches to maintain a natural-style dialog. It is shown in [6] how different machine-done visualized patterns lead to student’s perceptions that cannot be represented by the machine and when this is helpful during the educational process. Aiming to show domain independent approaches, the examples are introduced from Number Theory to language expressions and nonclassical logic applications. For example, on fig. 7 and 10 are shown multiplication cycles for numbers 9 and 11 in a set of 8 arithmetic series leaving after filtering out the first three prime numbers and their co-multipliers as well. The figures provoke questions why the images are repeated periodically across 11, or, respectively 19 rows, why their tour sequence is in a reverse order related to an imaginary center – the series $15+30k$ and many other questions the answers to which led us to interesting theorems from Number Theory together with cryptology applications. In the case of CALEIDOSCOPE the machine performs mechanical arrangement searches of repeated elements, etc. by presenting the results in a dialog with the human; it is expected that the human will see the unformal elements of the picture and will draw his/hers conclusions. The correct organization of the dialog and the presentation of the results is the essence of the CALEIDOSCOPE method.

The last one from the methods discussed here is the inference from contradictory information. This method is presented in details in [16]. On the first stage, based on the check for preliminary input knowledge as *models for contradictions,* the system determines whether there is a contradiction. If the sides of the contradiction are more than two, then the problem is reduced to the resolution of a set of 2-sided contradictions. Depending on the situation, the methods for the solution are automatic or interactive.

It is the inference method from contradictory information that allows to eliminate the incompletenesses from situations similar to the one depicted on fig. 13, and also to correct insignificant elements and/or to present the situation more clearly.

In this chapter basic methods of the synthetic meta-method SMM are presented. The next chapter is dedicated to the coordination of their operation under the control of SMM.

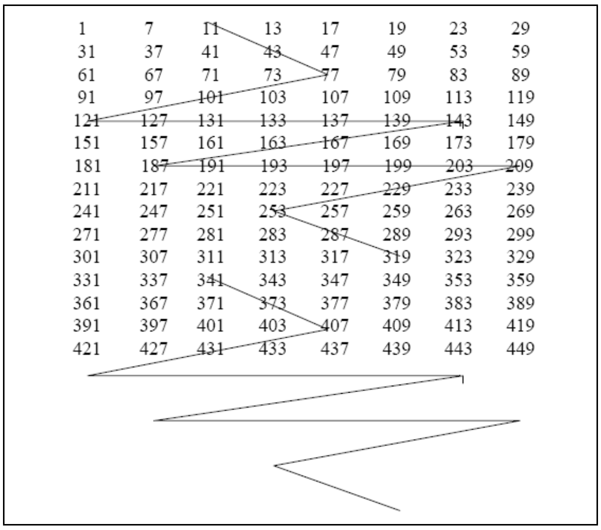


Figure 13. Cycle 11 in a set of 8 arithmetic progressions containing all primes ≥7

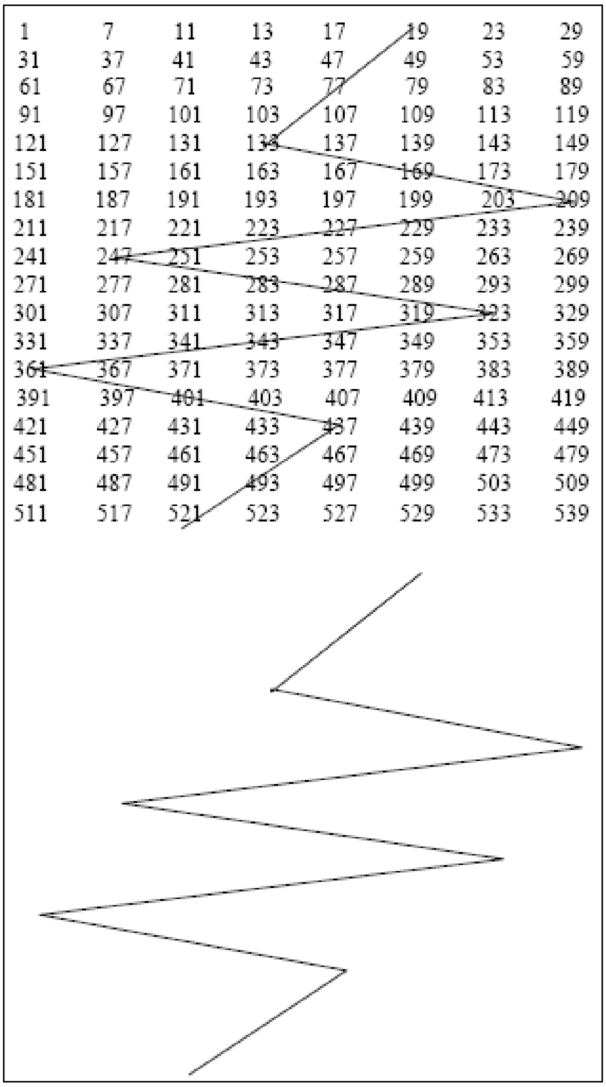


Figure 14. Repeating regularities similar to represented on Figure 13

3. Method coordination

Comparisons to well-known analogs may be given. Superiority of PUZZLE over classical case of constraint satisfaction, of CALEIDOSCOPE over cognitive graphics and other human-centered methods, of SMM/CONTRADICTION/FUNNEL applications over evolutionary fitness functions (variable or not, or a system of fitness functions), and of CONTRADICTION over conflict resolution methods may be discussed via comparisons of the above presented information to the well known cases. Altogether the advantages will allow construction of creative systems in future.

SMM is a synthetic method. Results from synthetic methods are more influenced by the coordination and the control among the components than by the components themselves. Usually the components of the synthetic methods are various algorithms from traditional analytic methods but we also investigated communications between several synthetic methods under the control of one synthetic meta-method. It is possible to make an analogy with the design of a display-window: if clothes are really good but the designer is not skillful enough, then the effect from the display-window dress make fail. The effect from an unqualified control in the domain of information security is even more significant and it leads to a complete crash.

It is recommended that at any time during the operations with SMM a direct control of the user's activities should be possible. For example, to detect the causes that led to a breakthrough in the system, the user inputs a {PUZZLE, FUNNEL, KALEIDOSCOPE, CONTRADICTION} and in this way he/she engages all resources to detect the cause. The cited sequence of commands is applied by default. In this case, every method operates independently, computer resources are allocated equally and it is the user intervention that may change the course of the investigation. It is provided that the results can be collected in two ways. The first way is the *blackboard* option. Here results are transferred between different methods and also to the user via a central node that is called the *blackboard*. In the second option methods from multiagent systems are used and results are transferred via the so called pheromones. This is information left by the agent which is kept intact for some time and which may be read only by agents that are close enough in space. The second option is especially good for transfer of operative information. Multiagent applications are discussed at the end of this section.

Another version for coordination is based on the use of statistical data and the personalization of information. For example, let's assume that statistics show that user X works most effectively with the PUZZLE method and also that he/she is interested least of all of the data visualization. In such case for him/her it is just the PUZZLE method that must be executed with the top priority, and the KALEIDOSCOPE must be activated only on special demands. On the other hand, as based on experts' estimates there must be a common system of priorities tied to the presented methods where the priorities are linked to the number of successive runs of the respective method. In this case the methods with higher priorities will occupy more computational resources.

The role of tests for contradiction(s) and the KALEIDOSCOPE method play a special role in the system. KALEIDOSCOPE is a visualization tool for results by constructing special modes for the human-machine contact. Therefore it must be activated after terminating the execution of any other method under the control of SMM. If the results do not require special technologies for their representation, then the method may be disabled manually. In the example presented below the set for coordination {FUNNEL-KALEIDOSCOPE} is used in a way shown on fig. 15 and 16. Let's assume that the goal function or the fitness function which are equivalent in the terms of evolutionary computation be marked by a central (red) line on fig. 15. Let's assume that the solution process of the defined problem is presented by a dotted trajectory on the figure. The intelligent data processing from the example is a type of approaches that are data driven. One of the properties of these approaches is the one that the solution may deviate from its way to the goal when lateral unexpected solutions are sufficiently interesting. A case with a deviation from the predicted trajectory is presented in the following way. Let the function G be denoted by the central (red) line on fig. 15.

The set of data accumulated in the system and knowledge M is filled via applying various logical inference rules, machine learning and/or other ways for knowledge acquisition. Each inference and any other change in the set M is traced via the gradient of information where the coordinate system is tied to the goal direction G.

The sum of gradients is shown on the figure via the dotted line. In the case when the change of M is inconsistent, for example the gradient calculus often change their direction because there are not enough significant intermediate results, then the dotted trajectory will change its direction when it reaches the funnel edge or, in other words, new knowledge will be searched only in the direction delimited by the position of the solution in the funnel which is presented for the point – an intermediate solution Y in (5):

$$A \le grad(Y) \le B\,. \tag{5}$$

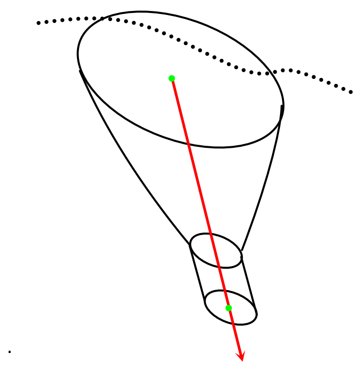


Figure 15. Funnel is not working

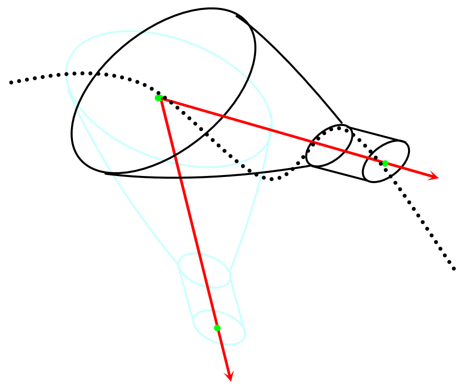


Figure 16. Funnel adjustments to work

The influence of G over the process of changes in M may be compared to the one of the gravity. Gravity is by no means the strongest force on Earth and we overcome it by walking, etc. On the other hand elevating (a motion along the antigradient) requires a lot of energy. If energy is not saved then the first cosmic speed is reached i.e. the earth attraction is completely overcome. Therefore limitations A and B from (5) are useless if in the flow – the dotted trajectory in fig. 15 there is much accumulated moment of inertia. There is an equivalent of the inertia in the case linked to the dotted trajectory in fig. 15. The following relations are introduced in relation to the problem defined. Let s_i be an intermediate solution that is interpreted as a point from the trajectory and then by analogy with (5) the following constraints (6) are introduced:

$$A \le \sum_{j=1}^{i-1} grad\left(s_j\right) \le B\,. \tag{6}$$

This is based on the following explanations. Gradients were the necessary interval between A and B for steps from the first up to $i-1$-th. Then, the bigger $i-1$ is, the bigger the inertial property ad hoc the solution s_i will be. When *i* is greater than an a priori threshold value then the constraints of the FUNNEL change their effect.

If the sum of gradients is bigger than an a priori set threshold value T_1, then the cited limitations may be overcome (7):

$$\sum_{j=1}^{i} grad\left(s_j\right) \ge T_1\,. \tag{7}$$

The next version includes a new constant T_2 and a solution s^* which is different from the defined goal G but which is marked by experts as sufficiently interesting. It can also be marked by the machine because for example, it is the missing link to solve other earlier defined problems. s^* does not participate in the series of s_i but if the distance between them is calculated via an additional p-adic metrics, then it is possible to avoid the limitations given by the FUNNEL:

$$\left\| s^* - s_i \right\| \le T_2\,. \tag{8}$$

In the three cases, cited above the ‘inertia stream’ pierces the FUNNEL under conditions given in the formulas as it is shown in fig. 15. Otherwise, all consecutive solutions will be inside the FUNNEL as it is presented in the description of the method above. In the case of a break, the FUNNEL stops running. Therefore, if the FUNNEL method does not adapt to the new conditions of the information flow as it is shown on fig. 16, then the method ceases to perform its main function and starts to gradually narrow the limitations to direct the set of dynamically changing information to the required place. During the tuning from fig. 16, this problem does not exist. On the other hand, by solving the problem from fig. 15 and 16 we had to construct more than one system of fitness functions where the first (classical) function is directed downwards and the second one operated temporarily only between the beginning and the end of the depicted FUNNEL. In this way, conflicts are avoided between classical evolutionary methods and the requirements of contemporary intelligent data processing.

4. Ontologies

Intelligent data processing by itself is a complex process of transforming data into knowledge. If data and the knowledge accumulated in the system are not structured in advance, then in the majority of cases their processing will be ineffective or it will produce poor results. For this reason, ontologies are used [11]. The tool to create and manage ontologies is Protégé OWL [14]. In the majority of cases ontologies are presented as graphs. For example, the fragment of ontology from fig. 17 illustrates the types of ISS: intrusion detection systems, intrusion prevention systems and others. The presented ontologies give an idea about the links between the objects and the relations between them.

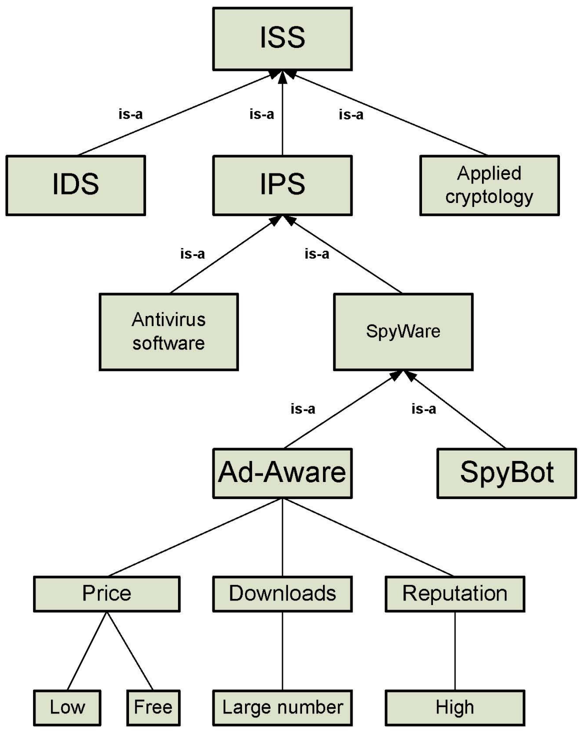


Figure 17. An example of ISS ontology

The relations on fig. 17 use the standard type ‘is-a’, and the taxonomy of the ‘ancestor-successor’ type is built based on them. The presented graphs are widely used to search information on the Internet, e.g. if the keywords are ISS then there is no need to write ‘anomaly detection’ because the machine will make the necessary links between the concepts. Another widely used relation is ‘has-a’ meaning ‘the object possesses the properties …’.

Examples of these relations are the meanings of the attributes in the lower part of the figure, e.g. ‘low cost’. Based on these two relations there are many applications of descriptive logics for the needs of the Semantic Web [13] and others. Recently various other relations are introduced in the KAON projects and other initiatives [15] that include operations with quantifiers and other logical operations.

The proposal to structure information is as follows. We propose the usage of tools close to the so called means-ends analysis. For this reason we introduce relations like WHY, HOW, WHAT FOR. In this way, new possibilities are formed to operate with knowledge that can be applied by traditional tools. For example, let’s propose an input in the system with a large number of differential equations. Let’s have a relation ‘WHAT FOR’ applicable to this block: ‘descriptions of trajectories during signal processing’. Using the relation WHY, we shall determine that ‘this way is more effective than traditional methods because …’; via the relation ORIGIN we shall be able to compare the offered fragment of knowledge with other popular methods, tools and approaches and so to discover the reasons for the problem. Applying the relation ‘WHAT IS THE WAY’, we shall discover texts with explanations how to use the differential method and other technical information. Using the enumerated tools, there will be a better structuring of information and prerequisites will be created for its more qualitative processing: in this way, we shall be to able more effectively operate with large blocks of formalized knowledge without going in technical details.

The usage of ontologies is not only building large sets of them. In the Web services, it goes through application of tools of the W3C consortium like WSMO and others [12]. The technology of these applications is more complicated. It includes specific concepts like orchestration and choreography. However, this is the way to apply a contemporary ontology processing - addition, deletion or update without breaking the links with other knowledge that is input in the system. In this way, the problems connected to the semantics of the problem are taken under consideration. Any other knowledge may change its meaning depending on the context, the point of view, etc. The version of Protégé-OWL is used in the creation of ontologies.

The necessity to introduce the following new relations in ontologies for intelligent applications of security is demonstrated. For example there is a large number of differential equations introduced in the system. A new relation ‘**what for**’ is introduced for this block: ‘to describe trajectories for signal processing’. The relation ‘**why**’ determines that ‘this way is more effective than traditional methods because …’; the relation ‘**origin**’ will be able to compare the objective fragment of knowledge with other known methods, tools and approaches to determine the cause for the problem; the relation ‘**in what way**’ will assist during localizations of explanatory texts about the usage of the concrete differential method and other information. It will be possible to structuralize information in a better way in the presented tools, and prerequisites will be created for its more qualitative processing. So we may operate with large blocks of formalized knowledge more effectively without deepening in technical details; at the same time we shall avoid much too detailed formalizations, i.e. here ontologies contribute to use good practice in educational processes. The more relations are examined, the more precise categorization of objects will be. Substantially all other used for the moment relations just illustrate this: in what way the relation arguments are connected , otherwise ‘$\mathbf{R}_i$ links the arguments (the objects) from A to N in a way determined in the definition’:

$$\mathbf{R}_i\ (A, B, C, D \dots N). \tag{9}$$

The set of $\mathbf{R}_i$ is used to determine that ‘A is connected with C’, etc. and the way of connecting is set in the definition of $\mathbf{R}_i$. It is suggested to introduce such $R_i^*(A, B, C, D \dots Z)$, the usage of which will form automatically answers not only of the type ‘B is connected with T’ but **why** ‘B is connected with T’, **in what way**, **when** the pointed link is actual, etc. Innovations do not deny previous ontological elaborations, they just supplement them; so the sense of matters is shown much more clearly, unambiguously that can be changed by the context. Also another ontology is introduced: ‘**important**’. At the first sight this type of information is subjective: the importance of knowledge is determined by the situation, the object, etc. But if information is not graded by importance then the agents will ‘drown’ in this sea of information provided the knowledge is not well structured; in the case structuring is performed best of all via the introduced new relation ‘important’.

Same as proposed ontologies may be used in the opposite direction of human-computer interaction: **what** is the reason to the question, in **what way** it is connected to the investigated goal G, why the user thinks the answer is **important**, etc. The new relations improve the human-machine interations and lead to human-machine brainstorming methods [6], one of the best ways to apply creative systems.

Part of ontologies may be succesfully shifted by PUZZLE and/or other applications of methods from Section 2. In the represented research is introduced that in some cases the graphic itself of the partly received resolve from figure 18 is enough to take decision from the side of the consumer or the expert.

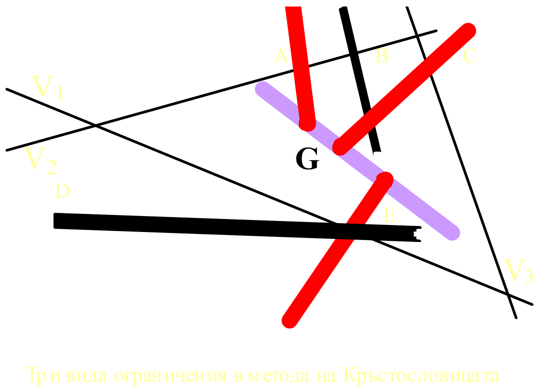


Figure 18. *Graphic presentation of the linguistic material*

The different variants of the modeling knowledge, consisted in the figure 18, represent resources for presentation of the things which are difficult to discover verbally. It is not unusual for the Chinese to use the aphorism “a picture replaces thousands words”. Here should to be added: “…and it is understood and remembered better”. Below are given different linguistic equivalents of the combinations of interpretations similar to that one from the figure 18 and received through the proposed and graphically described method of the puzzle.

- The things are moving to the way that…
- It is neither…nor…but it is too close to… (Without dilutions!)
- Look from the other angle … Other conclusions…
- The object can be represented as crossroad of the two dynamically developed areas…
- In the future these two concepts have to be used together…

The pointed naturally linguistic fragments can be represented and in direct text to the students but below is shown that trough the method of the puzzle this can be done significantly better. Modeling through the use of the method of the puzzle helps for the emotional contact with the audience. For example on the fig. 19 and fig. 20 is represented other example for the use of the method of the puzzle when are clarified particular processes and connections between the objects.

On the fig. 20 the system from different types of limitations, described in the second chapter, helps for discovery of the casual connection between M and N at the same time between them not existing any connection on the fig. 19. The connection is displayed in the impose of additional limitations on the fig. 20. The process is dynamic as nature and it is almost impossible to be correctly transferred by words only.

5. Other applications

The presented system source codes are written in different languages: C++, VB, and Prolog. It is convenient to use the applications in freeware like RDF, OWL, Ontoclean or Protégé. Many of the described procedures rely on the usage of different models/ ontologies in addition to the domain knowledge thus the latter are metaknowledge forms. In knowledge-poor environment the human-machine interactions have a great role, and the metaknowledge helps make the dialog more effective and less boring to the human. The dialog forms are divided in 5 categories from 1=’informative’ to 5=’silent’ system. Knowledge and metaknowledge fusions are always documented: where the knowledge comes from, etc. This is the main presented principle: every part of knowledge is useful and if the system is well organized, it will help us resolve some difficult situations.

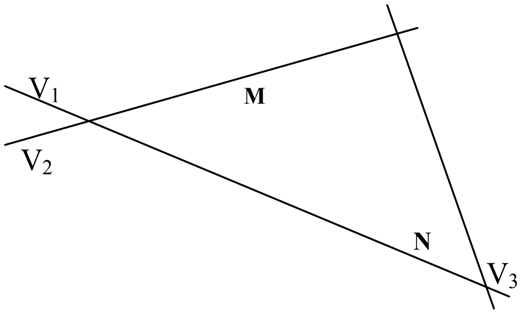


Figure 19. *M and N are not connected but located using linear contraints*

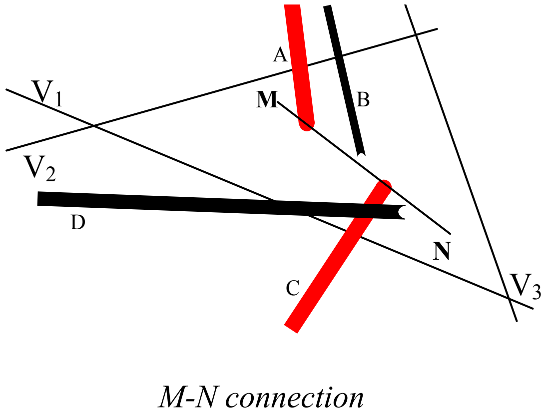


Figure 20. *Discovery of the connection between two objects M and N*

We rely on nonsymmetrical reply ‘surprise and win’, on the usage of unknown codes in combination with well known methods, and on the high speed of automatic reply in some simple cases e.g. to halt the network connection when the attack is detected. If any part of ISS is infected or changed aiming at reverse engineering or other goals, then the system will automatically erase itself and in some evident cracking cases a harmful reply will follow. The above presented models of users and environment are used in the case. Therefore different SMM realizations are not named IDS but ISS because they include some limited automatic reply to illegal activities.

Conflict situations occur when the agent can’t execute its goal in an acceptable time or in other situations. Say, the agent’s way to the goal is interrupted by a route for many other agents. The standard way frequently quoted in multiagent systems is ‘step out’ for all conflicting agents. Then another routes should be found, otherwise same conflicts will repetitively occur. Meanwhile some goals may be lost. In contemporary dynamic environments consisting of hundreds and thousands of agents the old-fashioned ‘step out’ strategy leads to creation of new conflicts rather than to their resolution. On the other hand, one or more of conflict resolution methods considered above resolve the problem using well known class Agent methods halt(); suspend(); activate(); getState(); move(Location where); and sometimes clone(Location where, java.lang.String newName). Several libraries for class Agent had been used. All of them comply to FIPA standards.

Upon the formulation of goal *G*, few variants of control strategies could be applied. Most of the presented methods are data driven, and the user may trace the intermediate results and interrupt or change the resolution process if necessary. Also he may apply his personal plans because the goal is seldom formulated without any plan or rough idea how to resolve it. It is possible to write the plan like ‘use the FUNNEL method in direction *X*, then use PUZZLE and in the end KALEIDOSCOPE to represent the result’. This example of methods interaction is interpreted in Fig. 4 where the solution is located in spot E, the dashed lines are linear constraints, and the spotted lines give us the parts of the solution as described in the PUZZLE section above. If we compare the results represented on figures from Section 2, we may conclude that PUZZLE method works better in combination with FUNNEL because we need less constraints using the fitness function f, and narrow search in FUNNEL helps to locate the goals. The same conclusion is to be applied to combinations with other existing methods from the scope: combining any of the above methods with other statistical or data mining methods is a way to better applications. The only obstacle here is the overall complexity of the obtained combinations. Almost the same result could be reached under SMM automatic control. The FUNNEL usage narrows the resolution process direction but in the beginning it doesn't exclude unpredictable results, hence the best possibilities for data-driven methods are obtained. If no interesting results have been obtained, then PUZZLE is activated to bind the goal to knowledge related to it.

6. Conclusions

The main conclusion is that to overcome the shortcomings of contemporary creative-related methods and systems, methods and applications are considered concerning the logical parts of knowledge discovery and/or human-machine creation of a brainstorming type. Special attention is paid to methods for identification and resolution of conflicts, and to machine (self-) learning based on them. The role of the above methods to the creation of more autonomous agents is discussed.

Analysis is represented for methods used for machine learning in intelligent agents, for sending information by sense, and for understanding the semantics of the information, all the parts essential to every creative system. Common disadvantages for different existing groups of contemporary applications are revealed. Ways for method coordination under a synthetic method SMM are discussed.

References

[1] M. Boden. The Creative Mind: Myths and Mechanisms. London: Routledge, Taylor and Francis Group, 2004.

[2] G. Wiggins. Creative Systems. A preliminary framework for description, analysis and comparison of creative systems. Knowledge-Based Systems, Vol. 19, No. 7, November 2006, p. 449-458.

[3] V. Jotsov. Intelligent Information Security Systems. Sofia: Za bukvite-o pismeneh, 2010, 211 p.

[4] B. Thuraisingham. Data Mining Technologies, Techniques, Instruments and Trends. NY etc.: CRC Press, 1999.

[5] D. Dasgupta. Artificial Immune Systems. Moscow: Fizmatlit, 2006.

[6] V. Jotsov. Novel Intrusion Prevention and Detection Systems. *In: Proc. 4th International IEEE Conference on Intelligent Systems*, Yager R., Sgurev V and Jotsov V. (Eds.), Vol. II, Varna, Bulgaria, September 6-8, 2008, pp. 14.20-14.27.

[7] V. Jotsov. “Evolutionary parallels,” *In: Proc. First Int. IEEE Symp. ‘Intelligent Systems’*, T. Samad and V. Sgurev (Eds.), Varna, Bulgaria, vol. **1**, pp. 194-201, 2002.

[8] A. Arruda, “A survey on paraconsistent logic,” *in: Math. Logic in Latin America*, A. Arruda, C. Chiaqui, N. Da Costa, Eds. North-Holland, Berlin NY, pp. 1-41, 1982.

[9] V. Jotsov. Semantic Conflict Resolution Using Ontologies, Proc. 2nd Intl. Conference on System Analysis and Information Technologies, SAIT 2007, RAS, Obninsk, September 11-14, 2007, vol. 1, pp. 83-88.

[10] V. Jotsov. “Knowledge acquisition during the integer models investigation,” *Proc. XXXV Int.Conf. "Communication, Electronic and Computer Systems"*, Tecnical University of Sofia, pp. 125-130, 2000.

[11] T. Gruber. A translation approach to portable ontologies. *Knowledge Acquisition*, Vol. 5, No. 2, pp.199-220, 1993.

[12] Web Service Modeling Ontology *(WSMO).* http://www.wsmo.org/TR to date.

[13] D. Fensel. Ontologies: A Silver Bullet for Knowledge Management and Electronic Systems. Berlin Heiderberg New York: Springer-Varlag, 2004.

[14] The Protégé Ontology Editor and Knowledge Acquisition System. http://protege.stanford.edu to date

[15] Welcome to KAON — KAON - The KArlsruhe ONtology and Semantic Web Open Source Management Infrastructure. http://kaon.semanticweb.org/ to date

[16] V. Jotsov. Emotion-Aware Education and Research Systems. J. Issues in Informing Science and Information Technologies, USA, vol. 6, pp. 779-794, 2009 (online version available).

Author information

Vladimir S. Jotsov (В.С. Йоцов): e-mail: bgimcssmc@gmail.com

State University of Library Studies and Information Technologies (SULSIT);

Intsitute of Information Communication Technologies of the Bulgarian Academy of Sciences (IICT-BAS);

P. O. Box 161, Sofia 1113, BULGARIA, EU;

ИНФОРМАЦИОННЫЕ МАШИНЫ: НЕКОТОРЫЕ КАТЕГОРИИ ФУНКЦИЙ И КОМПОНЕНТ

Мержвинский Анатолий Александрович

Abstract: *Отмечено, что развитие теории информатики пока не привело к единству определений отличающегося наибольшим объемом понятия информационные машины (ИМ). Цель статьи – выделение категорий актуальных компонент и разработка обобщенной схемы ИМ, овершенствование онтологии. На основе анализа взаимодействий объектов приводятся концепция структуры, основные компоненты и определение ИМ как родового понятия компьютерных и информационных систем. Показано, что для всех типов взаимодействий компонент ИМ характерным есть феномен присутствия носителей взаимодействий в виде вещественных или энергетических потоков, названных* ***коммуникатами.*** *В концепцию структуры ИМ введено универсальное понятие* ***оперант****: реализатор операций любого уровня, начиная от простейших - операций с коммуникатами - до самых сложных - со знаниями. По аналогии с понятиями «пиксель», воксель и др. введено и определено понятие* ***иксели*** *- простейшие материальные элементы, реализующие преобразование коммуникатов в информационные объекты и наоборот, а также выполняющие операции фиксации, хранения, отображения и передачи информации. По аналогии с икселем определена универсальная структура: агрегат "****физический объект – информационный объект»,*** *кратко ФИОб.* ***М****ножество агрегатов представлено диаграммой доменов объектов категорий R материального мира и отражений материальных объектов и ментальной деятельности. Завершенные акты взаимодействия коммуникантов, по аналогии с логической единицей работы с данными, определены как транзакции. В соответствии с ипостасью агрегата, материальной R или информационной, определены категории транзакций как <RxI>, классы информационных (содержащих I-объекты), неинформационных машин (не содержащих I-объекты) и компонент ИМ.*

Ключевые слова: *ИНФОРМАЦИОННАЯ МАШИНА, ФУНКЦИОНАЛЬНАЯ ЦЕПЬ, ВЗАИМОДЕЙСТВИЯ, КОММУНИКАТ, ОПЕРАНТ.*

ACM Classification Keywords*: Theory of the Information. Philosophy and Methodology of Informatics.*

Введение

Создание вычислительных средств нового поколения включает разработку онтологий баз знаний, онтологоуправляемых компьютерных систем, создание категориальных каркасов предметных областей (ПдО) верхнего уровня [Палагин, 2012]. Успехи теории пока не привели к единству определений отличающегося наибольшим объемом *понятия информационные машины*. В одних случаях понятие ИМ связывается с автоматизацией мыслительной деятельности, в других - с функциями и задачами обработки больших объемов информации [БСЭ, 1976]. Известные понятия *машина Тьюринга*, абстрактные автомат Мили и Мура (англ. *Mealy machine*, *Moore machine*) – абстракции, содержащие только информационные входы и выходы и не отражающие материальных воздействий.

Цель статьи – выделение категорий актуальных компонент и разработка обобщенной схемы ИМ, совершенствование онтологии и таксономии ИМ на основе общего анализа взаимодействий материальных и информационных объектов.

Согласно [ГОСТ, 1984] компонентами машин являются *функциональные части* и *функциональные цепи*. Исторически сложилось так, что в технической и научной литературе преимущественно уделялось внимание теории функциональных *частей* ИМ; теория же функциональных *цепей* ограничилась мало связанными друг с другом теориями электрических и оптоволоконных линий связи, теориями радио и гидролокации, описаниями конкретных интерфейсов. Термин *процессор* позиционируется в первую очередь как функциональная часть обработки программы ЭВМ, а не как устройство обработки информации или реализатор какого-либо процесса вообще. Отсутствует общая теория устройств ввода-вывода, связанная с вводом образов материальных объектов внешнего мира и преобразованием информации в воздействия на материальные объекты. Используемые в практике понятия *электронно-вычислительная машина*, *програмно-аппаратный комплекс, компьютерная система* мало связаны с родовым понятием ***машины***; *машины,* как средства выполнения определенных действий с целью уменьшения нагрузки на человека или полной замены человека при выполнении конкретной задачи. Так в толковом словаре по информатике абстрактная машина – *представление об ЭВМ в терминах информационных ресурсов и операций, доступных программе* [Першиков,1991].

Системная модель взаимодействий объектов в природе и технике

При рассмотрении сущности связей макрообъектов кроме материальных взаимодействий актуальны также информационные взаимодействия [Markov, 2007], в которых вещественные и энергетические потоки являются носителями "отражений" реального мира [Мержвинский, 2009]. Важнейшим свойством макрообъектов является способность обмениваться со средой V и другими материальными объектами веществом G и энергией E, в частности излучением Г разной природы [Мержвинский, 2011].

Непосредственные взаимодействия и связи. Взаимодействие макрообъектов в физике рассматривается как обменный процесс, как передача от одного объекта к другому кинетической и потенциальной энергии и вещества [Канарев, 2007]. Возможности непосредственного действия одного материального объекта на другой определяются физическими *связями* между материальными объектами - характеристиками исходящего материального потока и ограничениями среды в окрестности объектов. Формально связь может быть описана моделью механизма передачи воздействия одного объекта на другой. Таким образом, непосредственная *связь* это пара: *испускаемый поток* и *среда*, определяющая воздействие *испускаемого потока* на объекты. Каждый объект имеет свой уровень чувствительности, то есть внутренние изменения происходят, когда внешнее влияние лежит в пределах области чувствительности объекта. Связи с одной стороны обеспечивают, а с другой - ограничивают свободу взаимодействия объектов системы.

Опосредствованные взаимодействия возникают с учетом воздействий, сформированных некоторым объектом-посредником. Согласно природе и роли при действии одного материального объекта на другой сущность процесса опосредствованных взаимодействий в пространстве и времени может быть отображена в виде функциональной схемы на рис.1 [Мержвинский, 2009]. Стрелки обозначают направление действия.

Взаимодействие, в частности информационное, между отправителем и получателем возможно на разных уровнях структуры потоков и с помощью отличающихся средств, например, с помощью электромагнитных или звуковых сигналов. Понятие *сигнал* очень широкое (сигналы транспорта и др.), но в теории информации и связи - достаточно узкое: освещающий фотографируемый объект видимый свет, поток чернил струйного принтера на бумагу, твердые носители текстов не именуются сигналами. Ниже предлагаются изложенные в [Мержвинский, 2009] следующие понятия и специальные термины, позволяющие однозначно трактовать участников взаимодействий согласно ролям объектов.

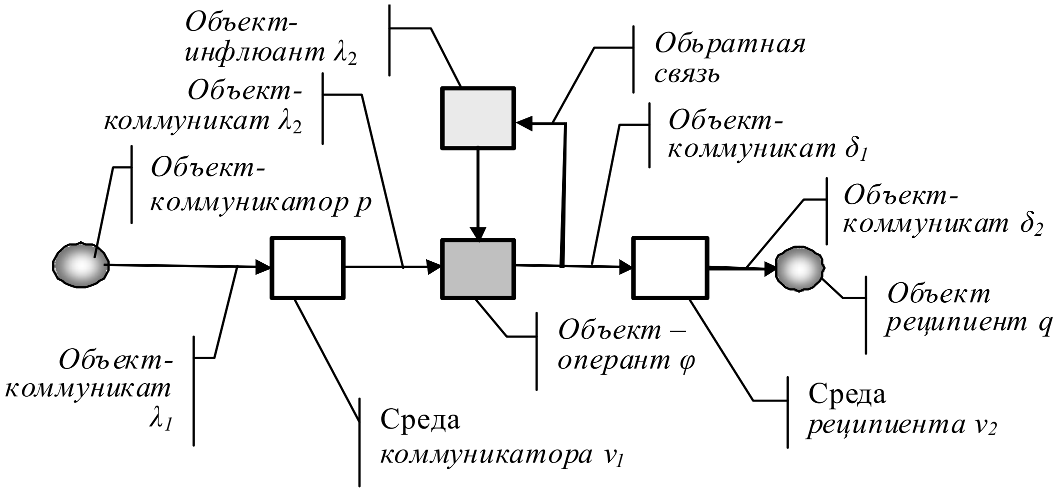


Рис.1 Схема опосредствованного воздействия *коммуникатора* на *реципиент*

Коммуникант - участник коммуникации, задействованный в коммуникативном акте взаимодействия между коммуникатором и реципиентом. ***Коммуникатор*** (*Источник*) - участник взаимодействия, который может порождать вещественные, энергетические или полевые потоки (*коммуникаты).* ***Коммуникат*** - вещественный, энергетический или полевой потоки, участвующие во взаимодействии объектов. В общем случае это могут быть потоки любой природы (материальные, информационные), которые способны осуществлять воздействие на реципиента структурными элементами потока. ***Реципиент*** – характеризуется состоянием и совокупностью реакций. ***Оперант*** - материальный объект, который обеспечивает взаимодействие между *коммуникантами*, в частности, опосредствованное воздействие *объектов- коммуникаторов* на *объекты-реципиенты.* ***Актанты*** – действующие участники взаимодействий. Среда – то, что окружает объект и оказывает на него влияние

Информационные процессы. Материальный объект может рассматриваться в ипостаси физического носителя информации, где информация - это "отображение" некоторой материальной неоднородности (объекта-*прототипа*) в совокупность элементов другой материальной неоднородности (объекта-*отображения*). В технике структура и физические свойства объекта-*отображения* выбираются в соответствии с технологией формирования *отображения,* природой объекта-*отображения* и исходя из требований *фиксации* либо *визуализации* отображаемой информации. Для описания материальных неоднородностей структуры объекта-*отображения* на нижнем уровне обычно используются такие термины как *пиксель, воксель*, *периодические колебания носителя*; в других случаях это более сложные конструкции – элементы памяти, регистры и матрицы памяти***.***

В литературе отсутствует обобщающий термин, именующий простейший *отображающий элемент* объекта-*отображения* независимо от того, что является носителем: *коммуникат,* преобразователь или элемент памяти. По аналогии с упомянутыми терминами *пиксель* (наименьший материальный элемент визуализации) и *сенсель* (от *sensor element* - чувствительный элемент) для именования элемента неоднородностей независимо от носителя введем обобщающий термин ***иксель*** (ixel – сокращение от *icon-element*). Конструкции ***икселей*** могут включать вспомогательные материальные элементы: защитные слои, повышения надежности и др. Состояние ***икселя*** формируется в результате действия на него ***коммуниката. Иксель*** *отражения* определим также как материальный элемент, реализующий функции *фиксации, хранения, отображения и передачи* информации (рис. 2, табл. 1).

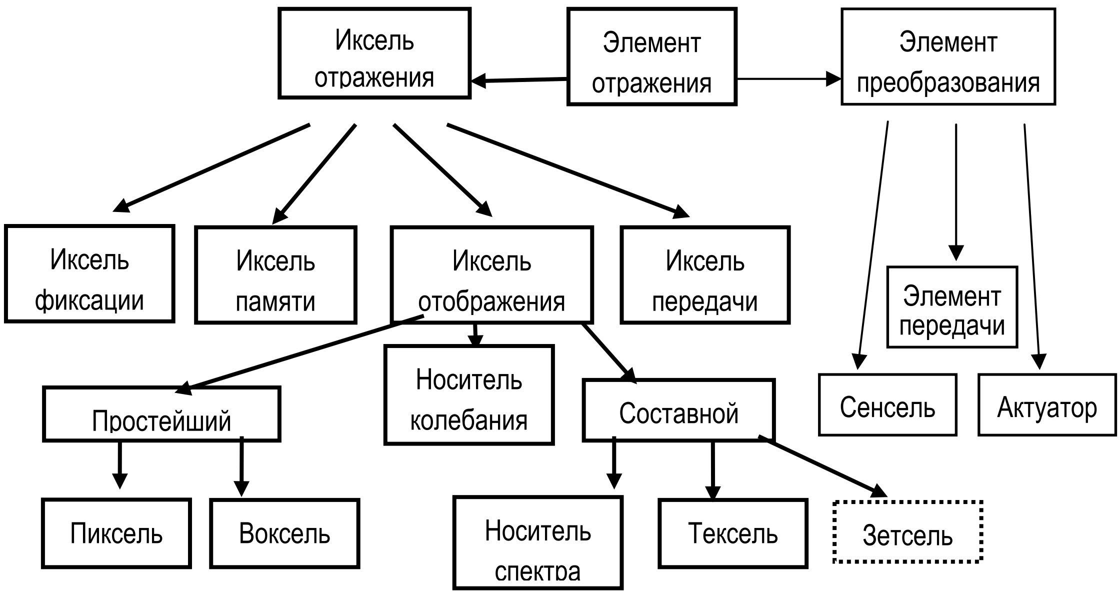


Рис. 2 Иксели – материальные носители информационных объектов

Таблица 1. Примеры и операции икселей отражения

Наименование	**Пример реализации**	**Функция отображения***
Иксель фиксации	Эмиттерный повторитель, элемент изображения видеосенсора	$i(t) = f(r(t))$
Иксель памяти	Триггер, конденсатор, ферритовая ячейка	$i(t)=f(r(t_0))$
Иксель отображения (пиксель)	Элемент текста или изображения на бумаге, на экране	$r(t)=f(i(t_0))$
Иксель передачи статический	Элемент регистра сдвига	$i(t_{j+1})=f(i(t_j))$
Иксель передачи динамический	Элемент сигнала	$s(x,t)=f(s(x-x_0,t-t_0))$

*где *i- ИО, t-* время*, r-* материальный объект*, t_0* – исходный момент времениI*, j-* номер разряда регистра*, x-*координата линии передачи*, x_0* – начальная координата

Иксели отображения – простейшая материальная структура некоторого носителя, который характеризуется непрерывным либо дискретным значением физической величины воспринимаемой органами чувств (яркости, прозрачности, спектра…). Визуально воспринимаемые *отображающие иксели* обычно реализуются на твердых носителях (бумаге, полотне, керамике, люминесцентных слоях). *Иксели передачи* – простейшая единица структуры *коммуниката* – материальной структуры переноса отражений (поля, потока частиц, в общем случае материальных объектов). Из *икселей* могут быть сформированы более сложные структуры, например *текселы* (сокращение от англ. Texture element) — минимальные единицы текстуры.

Будем исходить из того, что воспринимающий воздействие ***иксель,*** как элемент *отражения,* имеет две ипостаси (формы проявления) материальную и информационную. В первом случае он может быть описан физическими свойствами:

- быть статическим либо динамическим (в зависимости от носителя);
- способностью при физических воздействиях принимать одно из нескольких различающихся состояний.

Во втором случае - информационными:

— содержать один *информационный элемент* (ИЭ), который, как и материальный ***иксель***, может иметь то же количество состояний (в случае двух состояний ИЭ является *бит*);

— изменять состояние под влиянием информационных воздействий.

Из изложенного вытекают две категории отличающихся по природе функций *икселей*:

— А - Преобразование материального воздействия на *иксель* в состояние *икселя*, т.е. в состояние ИЭ;

— Б - Выполнение операций над ИЭ: *фиксация, хранение, отображение и передача ИЭ*.

С учетом естественности обратного преобразования функций категории А, можно дать такое определение икселя как агрегата:

Иксели – простейшие материальные структуры преобразования воздействующих материальных объектов (входных *коммуникатов*) в информационные объекты и, наоборот, преобразования информационных объектов в материальные объекты (выходные *коммуникаты*).

Абстрактная схема информационной машины

Реальные ИМ функционируют в среде из материальных объектов, взаимодействие с которыми обеспечивается определенными выше *коммуникатами*, поэтому компоненты и функциональные цепи ИМ могут быть представлены в виде обобщенной схемы как на рис. 3.

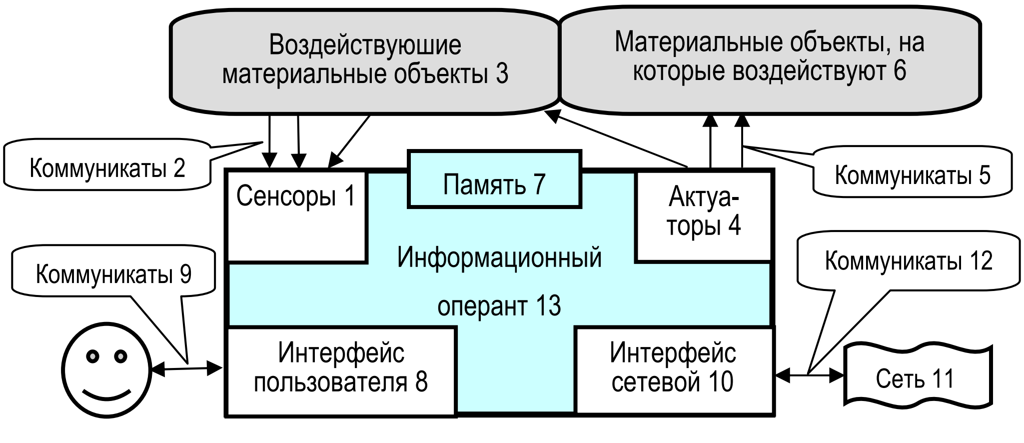


Рис. 3 Абстрактная схема информационной машины

Носители свойств материальных объектов, например, потоки частиц, электромагнитные либо звуковые излучения формируемые самим объектом, или излучаемые в результате зондирования объекта другим источником излучения (рис.3, 4), согласно изложенной выше терминологии участников взаимодействий являются *коммуникатами.* Параметры *коммуниката* 2, выходящие от объекта-*прототипа* 3 - амплитуда, частота, фаза, - могут зависеть, например, от таких свойств объекта-*прототипа*, как температура, скорость, способность отражать, перемещение, и т.п. Формально *коммуникатами* могут быть и потоки вещественных объектов-переносчиков, например, почтовые конверты с письмами.

Первичными преобразователями неоднородностей объектов материального мира (свойств объектов-*прототипов*) в неоднородности чувствительных элементов объекта-*отображения* являются *сенсели*, входящие в состав *сенсоров* 1. *Сенсели* превращают определенные параметры *коммуниката* 2 в аналоговый или дискретный сигнал (электрический, оптический, звуковой), удобный для воздействия на другие чувствительные элементы, например, усилители, *пороговые элементы. Сенсоры* снабженные, например, аналого-цифровым преобразователем, преобразовывают сигнал к удобной для работы цифровой форме. Каждый элемент неоднородности объекта-*отображения* в простейшем случае может характеризоваться численным значением соответствующей физической величины. Дальнейший переход от чисел к *символам* - в свое время шаг развития ИМ - позволил ставить в соответствие символам любые объекты, определять операции над группами объектов, разработать машинные алгоритмы решения задач, привел к созданию аппаратных и программных средств их реализации.

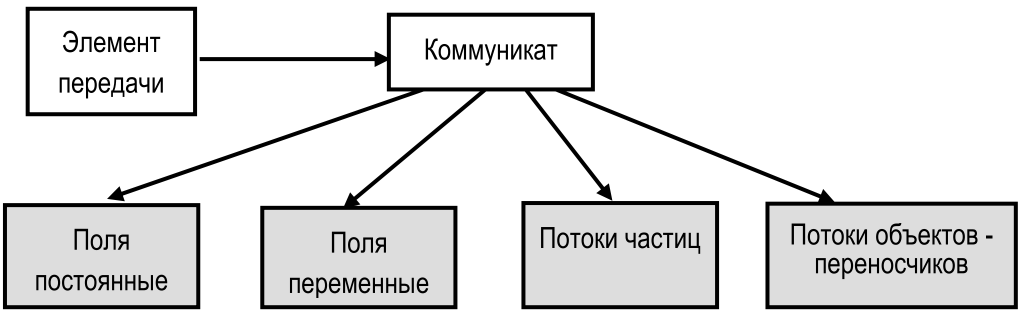


Рис. 4. Коммуникаты – материальные носители передаваемого информационного объекта.

Актуаторы 4 (*эфекторы, исполнительные механизмы, орудия труда*) - любая структура или вещество, которые в соответствии с состоянием ИЭ формируют поток *коммуникатов* 5 изменяющий состояние материального объекта 6.

Память 7 конкретной ИМ может содержать, например, базы знаний конкретных ПдО и другие встроенные структурированные информационные *ресурсы* накопленные обществом; чтобы отразить такую возможность, память частично вынесена за пределы операнта 13.

Интерфейс пользователя 8 включает средства ввода данных и задач, средства визуализации статической и динамической информации с помощью ***коммуниката*** 9.

Линии, каналы интерфейсов 8 и 10, сети связи 11 передаче сигналов классифицируются на физическом уровне по типу *коммуникатов* 9, 12 и среды распространения: электронные; акустические; оптические, инфракрасные, радио, почтовые.

Информационный ***оперант*** 13 - общий функциональный компонент ИМ, структура которого может состоять из совокупности последовательно и параллельно соединенных функциональными цепями *оперантов* нижних уровней. Важнейший параметр *оперантов* – скорость выполнения операций.

Очевидно, что на уровне икселей может быть реализована обработка потоков ***коммуниката*** с помощью таких операций:

- *объединение* потоков ***коммуниката,*** например, потоков электронов, фотонов, электромагнитных полей для реализации функций *И* и *ИЛИ*;
- *управление параметрами* ***коммуниката*** (*количественное**,** коммутация, инверсия функция НЕ*);
- *обнаружение* и *определение количественных характеристик* ***коммуникатов*** с помощью чувствительных элементов: фильтров, усилителей, пороговых элементов (квантор существования);

- *Обработка* информации под воздействием *коммуниката* (запись, хранение, стирание) на некотором *икселе.*

Одноименные приведенным операциям элементы можно отнести к элементной базе ***коммуникатного*** уровня. Очевидно, что эти элементы реализуют логические функции И ИЛИ, НЕ входящие в состав функционально полной системы элементов, и на них могут быть реализованы более сложные функции.

Аппаратно-информационный агрегат – базовый компонент многослойного представления информационных устройств

Из приведенных элементов И, ИЛИ, НЕ строятся, например, компьютеры структуры фон Неймана, в которых входная информация – информация, представленная в символах *входного алфавита* (числовая, текстовая, графическая, электрические сигналы и т. п.) [Электронный учебник, 2013]. Связывать через взаимодействия элементы структуры компьютера с материальными объектами не принято. Чтобы уйти от множества коннотативных значений понятий "информация" и "информационный объект", рассмотрим единицу структуры ИМ в виде: «*агрегат* ***физическийй объект - информация, содержащаяся в нем***», как усложненный аналог *икселя*. Этот *агрегат*, как и *иксель*, в простейшем случае может быть представлен как двухслойный объект, составленный из материального и информационного слоев.

Материальным слоем агрегата являются материальные компоненты агрегата – *иксели* и связанные с ними элементы (обеспечивающие доступ к ним и выполнение операций смены состояний). Этот слой, обычно именуемый как аппаратная часть, характеризуется *структурой:* входы, выходы, состав *икселей*, *связи, коммуникаты.*

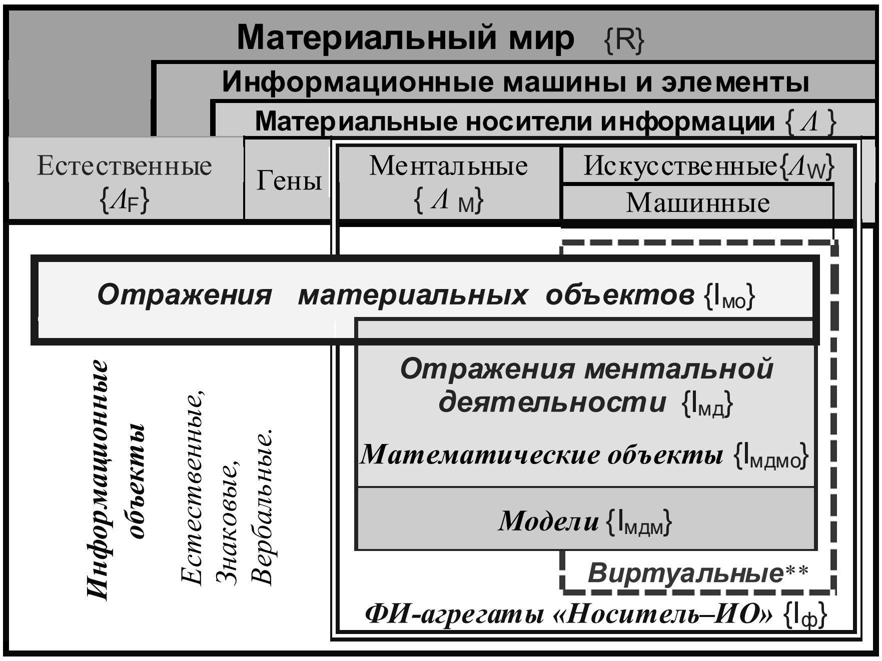
Информационным слоем агрегата являются "отображения" объектов в виде совокупности актуальных во времени или пространстве *состояний* ***икселей***. И элементарные и сложные структуры *состояний* совокупности *икселей агрегата* в общем случае будем называть "информационные объекты" (ИО). Информационная часть агрегата характеризуется формой представления ИО (аналоговая, цифровая), форматом (элемент, линейка, массив, таблица,текст, файл ...), отношениями между ИО.

Множества *агрегатов* из материальных и информационных компонент могут быть определены категориями по сочетаниям классов носителей и отображений. На диаграмме рис. 5 представлены устоявшиеся имена и взаимосвязи доменов объектов *категорий R* материального мира и отражений материальных объектов и ментальной деятельности человека. Поскольку функциями агрегата являются информационные операции и сам агрегат есть физический носитель информации, определим такие *агрегаты* как ***физические информационные объекты*** (сокращенно ***ФИ-объекты***, ***ФИОбы***). Множество ***ФИ-объектов*** обозначено *Iфа*, множество носителей информации - *Λ*, а подмножества ИО, которые содержатся в них, на диаграмме рис. 5 обозначены *I* с соответствующим индексом. При рассмотрении диаграммы уже на уровне *икселей* могут быть выделены такие слои ИО: бит, цифра, знак, icon-объект, массивы и гораздо более сложные структуры. Например, отображающие конкретные модели объектов внешнего мира, сложно организованные информационные структуры: знания, которыми мы определяем как математика, машинная математика и др.

Таким образом, и аппаратная и информационные части ***ФИ-объек***тов могут иметь *иерархическую* структуру; каждому уровню иерархии могут отвечать свои структуры и слои, то есть может иметь место многослойное представление как относительно постоянной аппаратной части ***ФИ-объек***тов, так и оперативно изменяющееся многослойное представление содержимого информационной части ***ФИ-объектов***.

Из изложенного выше может быть дано такое определение ИО:

Информационный объект - это отражение материального объекта-прототипа или другого ИО в информационной машине либо на некотором носителе в виде совокупности *состояний* определенных *икселей объекта-отображения*.



*Созданные в результате деятельности субъекта

* *Созданные в технических средствах

Рис. 5 Диаграмма доменов объектов категорий реального мира и информационного мира

Очевидно, что в зависимости от *природы* формирования информации ИО может:

- отображать некоторую свою сущность или других объектов (прообразов) множества *U* и отношения между ними, т.е. быть информационной моделью отображаемого объекта;
- представлять самостоятельную сущность (идею, конструкт, продукт деятельности), созданную творцом на основе двух или больше объектов, которая может непосредственно или через объекты-посредники влиять на третьи объекты.

При отображении других объектов ***ФИОб*** может содержать разные формы информации, которая касается как количественных, так и качественных характеристик объекта-прототипа. Отметим, что в случае отображения некоторой самостоятельной сущности ***ФИОб*** может содержать "код", "программу", созданные, например, ее творцом.

В информационных процессах у человека принимают участие периферийные рецепторы и нейронные структуры нервной системы. Естественно, в них принимают участие и "ментальные" (относящиеся к уму) объекты, будучи составляющей частью ***ФИ-объектов,*** *теорий, баз знаний*.

Категоризация взаимодействий информационных и материальных объектов

Операнты – материальные объекты, реализующие опосредствованное воздействие *объектов - коммуникаторов* на *объекты-реципиенты.* Завершенные *акты* взаимодействия коммуникантов, по аналогии с логической единицей работы с данными, определим как *транзакции.* В нашем случае ***транзакция*** – единица процесса опосредствованного взаимодействия. *Операнты* могут быть классифицированы по типу выполняемой *транзакции*, и по типу операции над *коммуникатами*. В первом случае в зависимости от ипостаси коммуникантов - как материальный объект *R* или как ***физический информационный объект*** *Iфа* (далее будем обозначать *I*) - можно выделить такие пары транзакций: *R→R, R→I, I→I, I→R* (рис. 6).

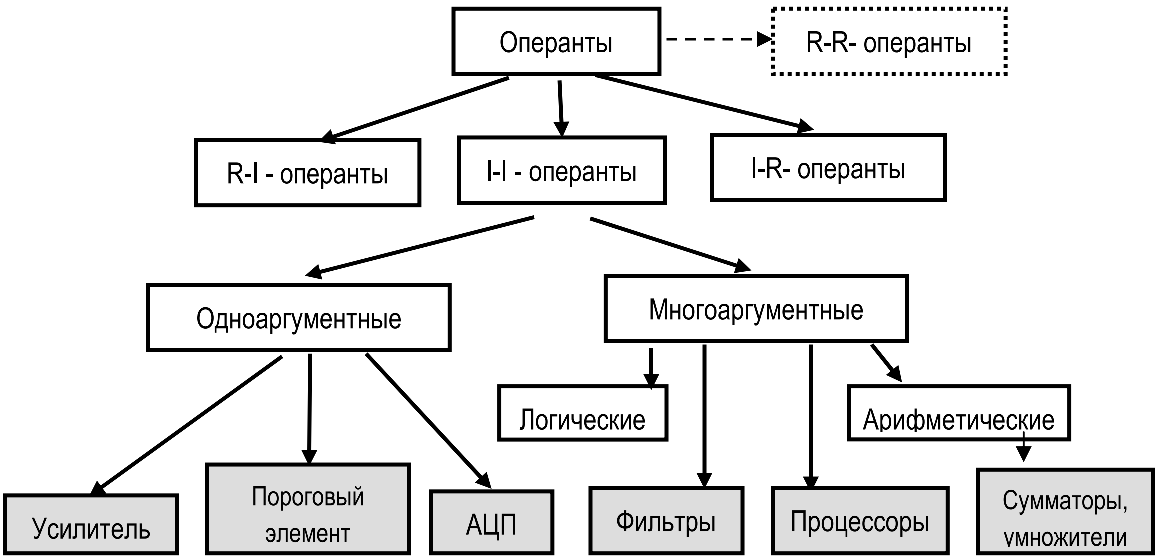


Рис.6 Примеры оперантов – материальных реализаторов транзакций

Классификация *оперантов* по комбинациям ***коммуникатов*** на входе и выходе *операнта r→r, r→i, i→i, i→r* определяет вид выполняемой операции. В математике *операция* - это двухместное отображение, или сопоставление двум элементам одного или разных множеств элемента этих же или третьего множества. В нашем случае ***оперант*** может иметь на входе один или несколько входных объектов и, соответственно функции преобразования, формирует один или, в общем случае, несколько выходных объектов. Поэтому преобразования <*rxi*> (входных объектов в выходные) удобнее толковать как *операцию технологического процесса* в отличие от *транзакций* <*RxI*>. Поскольку важнейшими компонентами последовательности ***транзакций*** кроме *коммуникантов* есть *операнты,* выполняющие ***операции***, при неформальном изложении последовательность ***транзакций*** может быть названа также "последовательность операций".

Отметим, что при опосредствованном взаимодействии результат может зависеть от влияния на ***оперант*** *φ* объектов, которые управляют, координируют или другим способом влияют на них, и которые можно отнести к категории ***объектов-инфлюантов. Объекти-инфлюанты*** могут включать среды, временные синхронизаторы, программы, модели объектов, модели желательных и реальных процессов.

Применительно к взаимодействию ***ФИ-объектов*** в множестве категорий ***транзакций*** можно рассматривать оба слоя взаимодействующих объектов и выделить такие виды взаимодействий: *материальные* (*R→R*) и с участием информационных объектов (*R→I. I→I, I→R*). Суть процессов взаимодействий на материальном слое определяется природой входных и выходных ***коммуникатов*** и ***связями***. При этом на объект влияют входные ***коммуникаты*** и сам ***ФИ-объект*** является источником выходных ***коммуникатов***.

Пример *коммуникатной* модели *материального слоя* объектов на рис. 7*:* Электрические сигналы λ_1 адреса, который хранится в постоянной памяти ФИ-1, адресуют некоторую ячейку памяти на магнитном диске ФИ-2, а инфлюант ФИ-4 формирует лазерный сигнал считывания λ_2. ФИ-2 формирует электрический сигнал *Δ*, переводящий ФИ-3 в состояние, отвечающее состоянию ФИ-2.

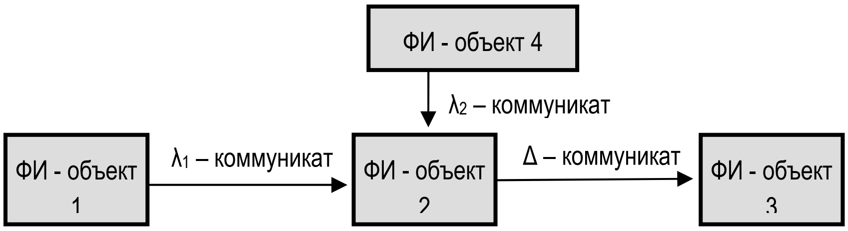


Рис. 7 Блок-схема процесса взаимодействий ФИ-объектов

Пример *информационной* модели *информационного слоя* объектов на рис. 7. На информационном уровне сущность ***ФИ-объекта*** определяется содержанием "отображения" и описание операции выглядит так. Код адреса ячейки памяти ФИ-2 с ФИ-1 передается в ФИ-2 и по команде считывания ФИ-4 код содержимого ячейки памяти ФИ-2 поступает в реципиент ФИ-3. Типичные операции памяти ФИ-2 такие, "запись", "стирание" или "считывание" конкретного ИО.

Ниже приведены особенности *транзакций* разных категорий.

R→I-транзакции. При *R→I - транзакциях* ***коммуникатор*** *R* рассматривается как материальный объект, а реципиент ***I*** как информационный. Такие *транзакции* выполняются с помощью сенсоров и могут быть определены как "сенсорные" *транзакции*; *I→R* - транзакции, выполняющиеся с помощью исполняющих устройств (*актуаторов)* могут быть определенные как "актуаторные" *транзакции* (по-другому *I→R* - реализации). *R→I - транзакции* более широкое понятие, чем общепринятые "операции ввода - вывода". В информатике, *ввод/вывод* (англ. *I/O - input/output*) обычно означает взаимодействие между обработчиком информации (например, компьютером) и внешним миром, которым может быть как человек, так и любая другая система обработки информации. Однако *Ввод* трактуется как получение системой сигнала или данных, а *Вывод*, как посылка системой сигнала или данных. Как и во многих определениях других литературных источников, процесс преобразования сущности объекта прототипа в его "отражение" в определениях не фигурирует. В нашем случае *R→I - транзакции* включают взаимодействие любого R-объекта (***коммуникатора***) и ***реципиента*** I с помощью ***коммуниката***. В российском и украинском языках отсутствуют обобщенное имя или понятие, означающие этот *R→I* процесс, как компоненты технологии добывания знаний о материальном мире. В то же время *R→I транзакция* - важнейшая составляющая когнитивного процесса. Наиболее близким представляется термин "отображение" как процесс. Техническими средствами *R*→I преобразование реализуется такими способами:

- Анализ с помощью соответствующего сенсора первичного *коммуниката* δ, (например, электромагнитного излучения или вещественного потока, которые создаются и эмитируются ***коммуникатором*** (рассматриваемым как активный объект); Это могут быть ***коммуникаты*** эмитируемые в результате процессов происходящих в объекте прототипе, например, собственное электромагнитное излучение объекта-прототипа имеющего температуру выше абсолютного нуля или излучение радиотехнического средства;
- Зондирование ***коммуникатора*** с помощью некоторого источника ***коммуникатов*** λ (который бомбардирует пассивный ***объект-коммуникатор***) и анализ с помощью соответствующего сенсора вторичного коммуниката δ, (который эмиттируется ***коммуникатором***). Это могут быть ***коммуникаты*** эмиттируемые объектом-прототипом в результате влияния на него внешних материальных потоков, например, звукового или электромагнитного зондирующего излучения;
- *Зондирование вспомогательных тест-объектов*, свойства которых известным образом изменяются под действием отображаемого объекта.

Вид ***коммуниката*** (фотонный, электронный или ионный потоки) для отображения в *искусственных* объектах выбирается исходя из особенностей материального объекта, необходимой разрешающей способности, удобства передачи, преобразования, хранения и регистрации информации.

В случае восприятия информации человеком может иметь место как процесс преобразования объекта-*прототипа* в его *морфный* образ, так и ментальный процесс кодирования явлений действительности *языковыми объектами*. Процесс кодирования отображается с помощью схемы Означаемое ↔ Знак. *Эта схема* - знаковая форма представления и отображения содержания смысла мышления или объективного содержания (смысла). Отношения значения раскладывается на две составляющих: от знака до означаемого - отношения номинации, от означаемого до знака - отношения референции.

I→R - транзакции. В простейшем случае осуществляется действие - преобразование "образа", "кода", "программы" или "модели" актуатором (иногда называют *интерпретатором*, а *I→R* преобразования называют *I-реализацией*). Классический пример такого преобразователя - пьезопривод. При изменении управляющего напряжения осуществляется перемещение незакрепленной стороны пьезокристалла. В более сложном случае орудие труда, средство для выполнения какой-либо работы. Для человека - формирование им графических или трехмерных объектов.

I→I - транзакции - выполняются средствами информатики (например, передача изображения из одного компьютера на другой, запись, считывание и обработка информации). В тех случаях, когда принципиально участие во взаимодействии с ***оперантом инфлюанта***, содержащего некоторую модель *m* (например, операции распознавания), *I→I транзакции* может быть обозначена как $I \rightarrow {}^{m}\varphi \rightarrow I$. Важнейшие параметры операнта - скорости выполнения операций и транзакций могут зависеть от модели *m*.

Из рассмотренного следуют следующие определения: *Информационная машина* - средство выполнения определенных действий с целью уменьшения нагрузки на человека или полной замены человека при выполнении конкретной задачи, в которой место операции с информационными объектами.

Приведенные категории *транзакций* могут быть положены в основу таксономии теории ИМ, в частности теория *R-I-R* транзакций – теория роботов, теория *I-I* транзакций - теория информационных оперантов, теория *R-I* и *I-R* –транзакций - теория устройств ввода/вывода.

Теория категорий – средство описания компонент ИМ

В отличие от конкретных *R→R транзакций* материальных объектов *I→I транзакции* информационных объектов естественно определять как отношения. Хотя *отношения* - достаточно общая категория, которая может быть применена и для случая *R→R транзакций*. ИО могут быть представлены в виде совокупности ИЭ, а сопоставления одних элементов совокупности других элементов этой же или другой совокупности - *отображениями*. Другое название *отображения* - это *функция*. Для формирования подкатегорий *коммуникатов* и ИО может быть также использован такой раздел математики, как *теория категорий*, в которой изучают свойства отношений между математическими объектами не зависящие от их внутренней структуры. В теории категорий вместо слова "функция" используют большее нейтральное слово "стрелка" (а также слово «морфизм") [Голдблатт, 1983]. Функция определяется как тройка (рис.8) f=<A,B,R>, где - $R \subseteq A \times B$ бинарное отношение между A и B, такое, что для каждого x ∈ A существует ровно один y ∈ B с $<x,y> \in R$.

Для каждой пары объектов *x* и *y* может быть задано множество *морфизмов,* которые в нашем случае могут быть реализованы *оперантами*, однозначно преобразующими вход в выход.

Аксиоматическое определение *категория* содержит в себе:

1. Совокупность предметов, которые названы Р-объектами (в нашем случае *коммуниканты*);

2. Совокупность предметов, которые названы *Р-стрелками* (здесь - функциональные цепи и *операнты*).

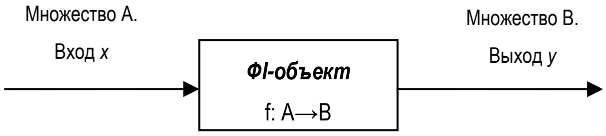


Рис.8 Блок-схема процесса взаимодействий ИО ***ФИ-объекта***.

Это дает возможность определить:

— Операции, которые ставят в соответствие каждой стрелке начало и конец: *f : a→b;*

— Операцию, которая ставит в соответствие каждой паре <g,f> *Р-стрелок* композицию $g \circ f$.

Композиция двух отображений (*стрелок*) φ_1 и φ_2 также есть отображение (*стрелка*) и обозначается символом "∘", например $\Phi = \varphi_1 \circ \varphi_2$. Причем, операция композиции ассоциативна: $h \circ (g \circ f) = (h \circ g) \circ f)$. Что позволяет:

— представить *транзакцию* некоторого *операнта* как последовательность *транзакций* его структурных компонент;

— перейти от отображения одной транзакции к отображению цепочек транзакций из разных вариантов взаимодействий *коммуникантов*, в частном случае приводяшим к тому же самому результату.

Заключение

Показано, что в информационных машинах для всех типов взаимодействий материальных и физических информационных объектов характерным есть феномен присутствия *носителей* взаимодействий. *Носители* в виде *вещественных или энергетических потоков* названы ***коммуникатами.*** В отличие от традиционного анализа компьютеров различных поколений, связанных с совершенствованием операций над *кодами*, в основу анализа функций элементарных компонент ИМ взяты операции над *коммуникатами*.

Введено и определено понятие **иксели** - простейшие материальные элементы, рализующие преобразование *коммуникатов* в информационные объекты и наоборот, а также выполняющие в ИМ операции *фиксации, хранения, отображения и передачи* информации. По аналогии с *икселем* определена универсальная структура: агрегат “***материальный объект – информационный объект”***. Множество агрегатов представлено диаграммой доменов объектов категорий *R* материального мира и отражений *(*материальных объектов и ментальной деятельности).

Завершенные а*кты* взаимодействия коммуникантов, по аналогии с логической единицей работы с данными, определены как *транзакции.* В соответствии с ипостасью агрегата, материальной R или информационной, определены:

- категории транзакций в ИМ как <*RxI*>;
- цепочки операций ИМ для случая многослойного представления их работы;
- классы информационных (содержащих *I*-объекты), неинформационных машин (не содержащих *I*-объекты) и их компонент.

Приведены определение ИМ как родового понятия компьютерных и информационных систем, ее абстрактная схема. В концепцию структуры ИМ ведено универсальное понятие *оперант* - реализатор операций любого уровня: начиная от простейших - операций с коммуникатами - до самых сложных - со знаниями. Показано, что в концепцию описания взаимодействий оперантов, материальных и информационных объектов могут быть положены основные понятия теории категорий: *бъект* и *стрелка.*

Введенные в статье понятия структуры компонент ИМ были использованы при разработке системы категорий для построения целостностной картины мира и модели универсума [Мержвинский, 2009] и могут использоваться при дальнейшей таксономии ИМ.

Ссылки

[Markov, 2007] Krassimir Markov, Krassimira Ivanova, Ilia Mitov. BASIC STRUCTURE OF THE GENERAL INFORMATION THEORY. International Journal "Information Theories & Applications" Vol.14 / 2007

[БСЭ, 1976] Информационная машина http://dic.academic.ru/dic.nsf/bse/90904/

[Голдблатт, 1983] Р.Голдблатт. Топосы. Категорный анализ логики: Пер. с англ.- М.: Мир, 1983. – 488 с.

[ГОСТ,1984] ГОСТ 2.701-84. Единая система конструкторской документации. СХЕМЫ. ВИДЫ И ТИПЫ. ОБЩИЕ ТРЕБОВАНИЯ К ВЫПОЛНЕНИЮ.

[Канарёв, 2007] Канарёв Ф.М. Начала физхимии микромира. 8-е изд. Краснодар, 2007

[Мержвинский, 2009] Мержвинский А.А. Модель универсума. Informations Models of Knowledge, XVI-th International Conference Knowledge – Dialogue – Solution, N. 15. – ITHEA, Sofia, 2009. p.31- 39. http://wiki.ithea.org/tiki-read_article.php?articleId=571

[Мержвинский, 2011] Мержвинский А.А. ФИЗИКО-ОНТОЛОГИЧЕСКИЙ ПОДХОД К ПОСТРОЕНИЮ ЦЕЛОСТНОЙ КАРТИНЫ МИРА// International Journal "Information Technologies & Knowledge" Vol.5, Number 3, 2011 с.259-273.

[Палагин, 2012]Палагин А.В., С.Л. Крывый, Н.Г.Петренко. Онтологические методы и средства обработки предметных знаний»/Луганск 2012, с. 323.

[Першиков, 1991] Першиков В.И., Савинков В.М. Толковый словарь по информатике.М: Финансы и статистика, 1991 – 543 с.

[Электронное пособие, 2013] Информатика. Логика и функционирование. Электронное учебное пособие. http://ien.izi.vlsu.ru/teach/books/003/.

Благодарности

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

Authors' Information

***Мержвинский Анатолий Александрович** – Институт кибернетики им. В.М.Глушкова НАНУ, 03680 МСП Київ-187, пр-т Академіка Глушкова,40, Украина;*

e-mail: merjv@mail.ru

Major Fields of Scientific Research: физико-технологические проблемы кибернетики, микро-оптоэлектроника, биосенсорика

THEORY OF NON-VIOLENT INTERACTION

Iurii Teslia

***Abstract**: The basic of the hypothetical non-violent theory of interaction is provided here. Introformational method is proposed to calculate the response to non-violent influence among intellectual property rights underlying the reflex construction of intelligent systems.*

***Keywords**: non-force interaction, introformation, reflex systems, self-organization, information theory.*

***ACM Classification Keywords**: I.2 Artificial Intelligence – Philosophical foundations.*

Statement of the problem

Despite the enormous investment of labor and financial resources there is no general theory till now and its practice of supporting the functioning of the human mind. There is still unknown mechanisms of the human brain as a complex structure consisting of a large number of interacting objects. Exploring the brain, or the results of its operations, it is difficult, perhaps even impossible to see those fundamental laws, which led to its establishment. Not in the narrow biological sense [Redko, 2007; Caparres, 1995], but in the context of those laws that underlie the development of nature and in the end led to the brain genesis as the best tool to process information. And for this, we first need to understand the role that nature gives to the information.

One of the most common problems in modern science is the knowledge of the role of information, not only in the life of biological objects, but also at the level of the existence of any material objects of nature. The emergence and development of scientific disciplines and areas of research which are information processing in a variety of natural and artificial systems, has led to the need to review the entire outlook on the essence and fundamentals of the laws of nature. Many scientists agree with the opinion that it exists at the level of non-living matter, that information is not only a vital part of living beings, but is present in all the processes of interaction in nature, and at all material entities and its laws, along with physical, shape our universe [Kamshilov, 1979].

Almost no discipline can do without the concept of information. It is not just about science, technical, biological, or philosophy. This is manifested in the natural sciences [Brillouin, 1972; Uspensky, 2010]. The classic example is quantum physics and its integral part - quantum information theory. Information is not just a concept used by different sciences. Based on it, the various sciences have long started to converge. But where is the limit of this convergence? If we consider some of the concepts, laws, or laws that are taking place in various sciences, we are able to interpret the essence of the understanding which has developed in other scientific field. Hence it is possible to construct a theory, based on the concept of information and explaining the basis for the unity of the laws of interaction in nature. Further it can be used to create artificial intelligent systems.

Analysis of the main research and publications

The emergence of information is impossible without the cooperation and reflection of the results of this interaction in the structure of the material objects. Interactions in nature are inseparable from the information. And information is inseparable from the process of interaction. Hence, if to seek the role that nature gave to the information we cannot ignore the issues of interaction. Everything interacts in nature. Interaction of different physical nature (gravitational, electromagnetic, weak and strong nuclear) are implemented between any of the material objects. A lot of attention is paid to the modern science of information interaction [Kuznetsov, 2011; Kuznetsov2, 2003]. A huge amount of work on this topic focus mainly on the formalization of the interaction in computer networks (especially the Internet), in social structures, in education, etc. As part of this research information is understood as a message of knowledge, data, intellectual resources, etc. [Kuznetsov, 2011]. Semantic and axiological parameters of the information are recognized, the view is accepted that the information inherent in a self-managed systems, which make biological and social form of motion. Today, most scientists accept the view that the information is inherent not only at the level of self-managing systems; it is objective and is inherent in all the processes of interaction in nature [Redko, 2007; Kuznetsov2, 2003]. Then there is a need to expand the scope of theories on information interaction to all forms of movement in nature, not just the biological and technical aspects of its existence.

This is done in the theory of non-violent interaction [Teslia, 2005; Teslia, 2012].

Unresolved part of the problem

The results obtained in the theory of non-violent interaction results, expanding the scope of its use requires a formal and systematic localization subject of research relating to the theory, accessible representation of hypotheses, ideas and conclusions that led to the formation of mathematical tools to solve many practical problems. Spreading the ideas of the theory of non-violent interaction and confirmation of its truth by showing the benefits based on it artificial intelligent systems brings an objective need for research in this area. The lack of scientific work which fully represents the essence of the theory of non-violent interactions, was the source of this writing.

The wording of the purposes of Article

The main task of the non-violent interactions theory can be formulated as follows: "through our understanding of the world try to find a reasonable start in his laws, which may be completely unlike what we see, but it is expressed in it". The theory suggests one possible implementation of the mechanism of interaction in nature. The results obtained in the theory of non-violent interaction reinterprets physical laws and allow to create more accurate and precise in operation intelligent systems, and most importantly, create a unified picture of the implementation of the laws of interaction in Nature, suggest the importance of the dissemination of ideas called the theory in the scientific world. This is the focus of this paper.

Basic material research

Non-violent nature of the interaction

The author of works came to an interesting and somehow crazy assumption, based on the studies for many decades, issues of cooperation related to physical science (gravitational, electromagnetic, weak and strong nuclear), and interaction at the self-managing systems (IT), spending thousands of computer experiments in the search for analogies [Teslia, 2005; Teslia2, 2010]. And what if the laws of responses to the impact of the external environment in the living and non-living matter are one and the same? And, just as a person can handle the information, and on this basis, forms his behavior (defines the trajectory in the medium of existence), and all material objects process information (coming through the interactions of different physical nature - gravity, electromagnetic, weak and strong nuclear) and based on that change their trajectory. From this assumption that the objects do not bend the space-time continuum, forcing other objects go to the "hole". They "convince" other objects move in a certain path (to or from exposure). We can say that the interaction of different physical nature is not by force, but is non-violent (information). That is, the "word was first". Word provides information about the existence of some objects to other objects and changes something in the internal organization of these objects, which leads to a change in their behavior. Hence the name of the theory is theory of **non-violent interaction**.

The internal organization of material objects in the theory of non-violent interaction called **introformation**. Introformation (internal organization, their own functional material objects) forming their attitude to the truth (reality).

If the motion of any material formation is determined by its internal organization (introformation), the changes of the laws of motion can be obtained by changing the laws introformation. And this has to lead to the numerical measure of introformation. It should be such that the actual "amount" of the movement of material corresponds to its introformational filling. We can go even further. If the laws of interaction in nature are one, it is likely that the resulting introformational laws changes will work at the level of living matter at the level of the man and his intellectual apparatus.

Based on the research on this scheme a new theory was created. Let's have a look at it.

VIP-interpretation of the motion

The theory of non-violent interaction provides a new interpretation of the VIP-motion (linking: V-speed traffic, I-introformation, P-probability). The essence of it in the following. Mechanical motion is characterized by the direction and speed. It is traditionally believed that the laws of nature must provide formation in the interaction of material formations of different directions and different speeds of movement. Suppose, the nature laws of motion are implemented more simply than we had so far? And there is only one **(absolute) velocity** of the matter! The speed of light in vacuum is "c". And at this rate moves all matter. And all observed or not observed but the existing example in microcosm, the diversity of the relative velocities of the internal organization formed themselves moving objects as follows. Suppose there is a quantization of space and time, at each time slot material formation is shifted to one quantum of space in one direction. The possibility of bias in each direction is given its own (for physical education) probability, which in turn generates an internal organization (internal relation to reality) of this formation, its introformation. As a result of the formationeach material will drift in that direction, which probability is higher.

Expected drift rate for one-dimensional motion will be:

$$V_{drift} = (p - (1 - p))c = (2p - 1)c \ , \qquad (1)$$

where V_{drift} – the observed velocity (drift velocity);

p – The probability of displacement;

c – Absolute speed.

If displacement occurs only in one direction, the drift velocity of the material formation is equal to the absolute velocity of matter – c.

$$p = 1 \Rightarrow V_{drift} = (2p - 1)c = (2 \cdot 1 - 1)c = c$$

$$p = 0 \Rightarrow V_{drift} = (2p - 1)c = (2 \cdot 0 - 1) = -c$$

The light in vacuum moves with the speed $V_{drift} = |c|$. Hence, the movement of the light can be seen as a one-way.

Introformation measures

Internal organization (introformation) of the material entities in work [Teslia, 2005; Teslia2, 2010] is presented with a geometric model - areas which determine their displacement in space. For one-dimensional motion – it is two directions of displacements determination (DDD) (Fig. 1).

The probability of selecting the direction is determined by the size relation DDD. If there is a choice between one-dimensional motion to the direction of Z, and the direction opposite to Z, then the following relation

$$\frac{p^+}{p^-} = \frac{p}{1-p} = \frac{i^+}{i^-}\left(given\, i^- \neq 0\right),$$

where $p^+ = p$ – the probability of displacement in the direction Z;

$p^- = 1-p$ – The probability of displacement in the opposite direction Z;

i^+ – The size of the area that forms the shift in direction Z;

i^- - Size of the area that forms the displacement in the opposite direction Z.

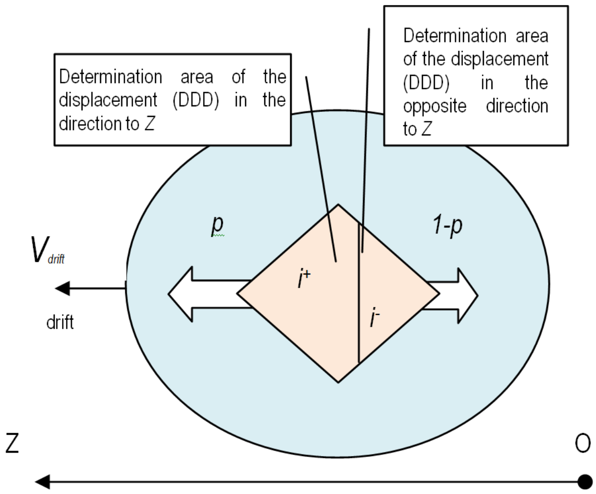


Figure 1. *VIP*-motion interpretation in nature

In this model, the movement is the property (functionality) of the object, and not the result of force or other material objects curvature of space-time continuum. Only the aspect ratio DDD sets its own probability of displacement in one direction or another.

The essence of **non-violent interaction** is that the interactions between material formations lead, initially, to change of the dimensions DDD, and the change in size DDD leads to a change of direction and speed of drift in space.

In this case, the drift velocity of material formation can be represented by the rate of motion of matter in nature and probability of displacement

$$V_{drift} = \left(2p^+ - 1\right)c = \left(2\frac{i^+}{i^+ + i^-} - 1\right)c = \frac{i^+ - i^-}{i^+ + i^-}c\,.$$

To simplify the model, the difference in the size of DDD in non-violent interaction theory was called certainty, and the amount – material formation awareness.

$$d = i^+ - i^-$$

$$i = i^+ + i^-$$

where d - the certainty of material formation, relative to the movement in the direction Z;

i - Awareness of the material formation, relative to the movement in the direction Z.

But the question is: what should be the dimensions of DDD, to determine the speed of 1 m / s? This question is linked with the other. After all, even Albert Einstein showed that all traffic except light traffic is relative. Where as the relative movement linked with the proposed VIP-interpretation? To get the size of DDD in the works [Teslia, 2005; Teslia2, 2010] a model is proposed, and the corresponding physical laws, and the intuitive understanding that the frequency of manifestation of material objects must conform to the size of their DDD. Consider the motion of two objects - X and Y. Of the special theory of relativity, the velocity of the object Y equal to the object X .

$$V_{XY} = \frac{V_Y - V_X}{1 - \frac{V_Y \cdot V_X}{c^2}} \tag{2}$$

where V_{XY} - the velocity of the object Y relatively to the object X ;

V_Y - the velocity of the object Y relatively to an observer located at a point O ;

V_X - the velocity of the object X relatively to an observer located at a point O ;

c - the speed of light in vacuum.

Substituting (1) into (2) we obtain

$$p_{XY} = \frac{p_Y(1-p_X)}{p_Y(1-p_X)+p_X(1-p_Y)} \tag{3}$$

where p_{XY} - the probability of displacement Y in the direction of the object Z relative to the object X ;

p_Y - the probability of displacement of the object Y in the direction Z ;

p_X - the probability of the object displacement X in the direction Z .

If we look at the denominator of (3), it can be far-reaching conclusion. Objects X and Y move only when they move in different directions. It is clear. But more interesting is the following. The denominator is not the amount that reflects the displacement of these objects in the same direction; we find that there is no displacement of different objects in the same direction at all? Or in such displacements different objects "transformed" into a single object?

From (3) that the number of times the subject Y moves in the same direction, the same amount of time the object X moves in the opposite direction (Fig. 2) and vice versa. But if the displacements "generate" internal organization of material formations, it means that the internal organization of the material objects must be related as follows:

$$\frac{i_X^+}{i_Y^+} = \frac{i_Y^-}{i_X^-}\left(given\ i_Y^+ \neq 0; i_X^- \neq 0\right), \tag{4}$$

where i_X^- - the size DDD of the object X in the direction opposite to Z ;

i_X^+ - the size of DDD of the object X in the direction Z ;

i_Y^- - the size DDD of the object Y in the direction opposite to Z ;

i_Y^+ - the size DDD of the object Y in the direction Z .

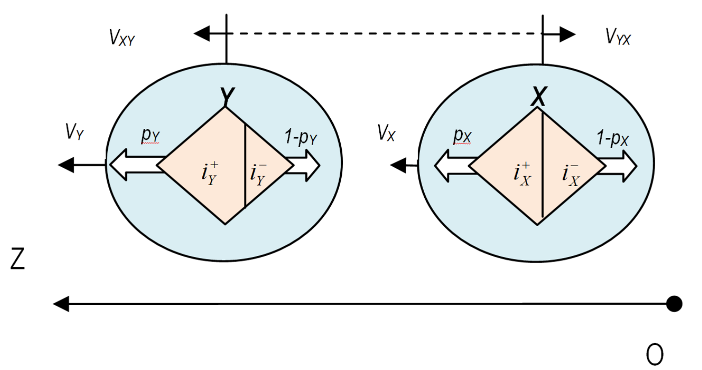


Figure 2. *VIP*- interpretation of the motion of two objects

This ratio provides correspondence between the probability of actual displacement and shifts frequency with respect to any observer of statistically independent manifestations. Correlation (4) is set to the same choice of different displacements of material formations in different directions. It is because of this relationship follows the formula of the relativistic velocity addition (2).

In non-violent interaction theory is suggested that if the ratio (4) was not satisfied in anytime, then more than 13 billion years could be eliminated [Teslia, 2012]. Inanimate matter has evolved too!

From VIP-interpretation of motion in the theory of non-violent interaction the relationship between the speed of the drift, the probability of displacement, definition and awareness of material objects was received:

$$i = \frac{1}{2\sqrt{p \cdot (1-p)}} \tag{5}$$

$$d = \begin{cases} 0{,}5\sqrt{\dfrac{p}{1-p} + \dfrac{1-p}{p} - 2}, & p \geq 0{,}5 \\ -0{,}5\sqrt{\dfrac{p}{1-p} + \dfrac{1-p}{p} - 2}, & p < 0{,}5 \end{cases} \tag{6}$$

$$i = \sqrt{d^2 + 1} \tag{7}$$

$$p = 0{,}5 + \frac{d}{2i} \tag{8}$$

$$V = \frac{d}{i}c \tag{9}$$

where V - the observed speed.

These relationships allowed a new perspective on a number of physical concepts and quantities [Teslia2, 2010; Klapchenko, 2011]. Thus, it became clear that the Lorentz factor of the material is identical to the material formation awareness. Indeed, from (5) and (9)

$$i = \frac{1}{2\sqrt{p(1-p)}} = \frac{1}{2 \cdot \sqrt{\dfrac{c+V}{2c} \cdot \dfrac{c-V}{2c}}} = \frac{1}{\sqrt{\dfrac{c^2 - V^2}{c^2}}} = \frac{1}{\sqrt{1 - \dfrac{V^2}{c^2}}}$$

$$\frac{1}{\sqrt{1 - \dfrac{V^2}{c^2}}} - \text{Lorenz factor.} \tag{10}$$

Then the formula for the relativistic mass and time are simplified

$$m = \frac{m_0}{\sqrt{1 - \dfrac{V^2}{c^2}}} = m_0 i \tag{11}$$

where m_0 - the rest mass of material formation;

m - relativistic mass of material formation.

$$\tau \frac{\tau_0}{\sqrt{1 - \dfrac{V^2}{c^2}}} = \tau_0 i \tag{12}$$

where τ_0 - the time of the material formation if it is at rest;

τ - relativistic time of the material formation.

Operations on introformation

The interaction of material objects changes the direction and speed of movement. So, it brings to the change of the introformation. Numerical measures that is specific and informed. How operates Nature of a certainty and awareness? The answer can be derived from physical laws. In particular, the law of conservation of momentum.

Substituting (9) and (11) into the momentum

$$P = m \cdot V = \frac{m_0}{\sqrt{1 - \frac{V^2}{c^2}}} \cdot V = m_0 \cdot i \cdot (\frac{d}{i} c) = m_0 \cdot d \cdot c \ , \tag{13}$$

where P - the momentum of the material formation.

Then the law of conservation of momentum sets the conservation law of certainties in a closed system

$$\sum P = const \Rightarrow \sum d = const \tag{14}$$

So, in a closed system the total certainty (the sum of the size difference DDD) does not change. Therefore, the law of conservation of momentum can be the basis for the formulation of the law unchanged definition of a closed system of material entities. In addition, based on the expression (14) the addition operation certainties can be offered. Namely, all material formation of the closed system can be replaced with a material formation with certainty, equal to the amount of certain material structures of the closed system

$$d_\Sigma = \sum_i d_i \ , \tag{15}$$

where d_Σ - the total certainty of material objects of a closed system;

d_i - certainty of the material formation M_i.

Another operation of the definition may be obtained from the formula of the relativistic velocity addition. From (2) and (9) comes up operation refilling of certainty

$$d_{XY} = d_Y \cdot i_X - d_X \cdot i_Y \ , \tag{16}$$

where d_{XY} - supplement certain (value of the certainty that reflects the difference in the definition of material objects M_X and M_y);

d_Y - certainty of the material formation M_y;

d_X - certainty of the material formation M_X;

i_Y - awareness of the material formation M_y;

i_X - certainty of the material formation M_X.

Or

$$d_Y = d_X \cdot i_{YX} + d_{YX} \cdot i_X \tag{17}$$

What is the purpose of these operations? Build of the certainties - is the sum of non-power impacts on all of us, which is provided by our friends, acquaintances, the media, etc. Operation of the certainty refilling gives the difference in the amount determined by different subjects. And it is equal to the magnitude of the impact, which is needed for the subjects were equally determined.

The conclusions from the theory of non-violent interaction

As can be seen from (11) - (17) *non-violent theory of interaction* not only allows us to simplify some of the physical expression, it still gives a qualitatively new interpretation of known physical concepts and laws.

1. It becomes clear why the speed of light is limited and maximum. After all, it is always offset with probability 1. And probably more than one does not happen.

2. It becomes clear why the speed of light is absolute (unchanged relative to the movement at any speed). From (1) $V = c$

$$p = \frac{V + c}{2c} = \frac{c + c}{2c} = 1$$

In addition, from (9) comes that, for material objects whose velocity is less than the speed of light ($V < c$) probability of displacement is less than one ($p < 1$)

$$V < c \Rightarrow p = \frac{V + c}{2c} < 1 .$$

Then, substituting the value of $p_y = 1$ in (3) and obtain

$$p_{XY} = \frac{p_Y \cdot (1 - p_X)}{p_Y \cdot (1 - p_X) + p_X \cdot (1 - p_Y)} = \frac{1 \cdot (1 - p_X)}{1 \cdot (1 - p_X) + p_X \cdot (1 - 1)} = \frac{1 - p_X}{1 - p_X} = 1, \quad (p_x < 1)$$

Substituting in (1) we will get

$$V_{XY} = (2 \cdot 1 - 1) \cdot c = c$$

Therefore, with respect to any object, the direction of displacement is not set with probability ($p_X < 1$), one-way movement will have the velocity *c*!

The reason is very simple. Since the material formations exist in relation to each other only when they are moved in opposite directions (3), the light we see only at those times when we displace by ourselves in the direction opposite to the direction of displacement of the light.

3. Becomes clear the core (if you want - intelligence) of certain physical laws. Thus the expression for the relativistic time (12) and weight (11) is simplified. And the increase in time and weight, with significant relative velocities can be seen as the result of increasing the size of DDD, i.e. increasing certainty (confidence) of one object relative to another on the right direction. The essence of this increase may be due to the greater awareness (increased areas DDD) produces greater confidence in how this formation acts (where it moves). Indeed, the likelihood that a significant certainty (confidence) is formed incorrectly or accidentally is insignificant. Therefore, a significant size of DDD shows not just the area of the corresponding relation to reality. But also greater confidence in the formation of this material (it is informed, mean very smart, so it can be trusted).

It also appeared that the momentum of the material is proportional to its certainty (can be generated by its definition?) (13). Comes up the following analogy. It is difficult to change the direction and speed of the massive and rapid material formation. Similarly, it is difficult to convince the informed person in something that does not coincide with his opinion.

If this study ends at this point, then many would have had the view of the next "setting up" the known laws. But these studies only begin. In any case, in the use of the theoretical results obtained for the creation of artificial intelligent systems. If we recall the origins of the theory of non-violent interaction, it should be recognized that the theoretical model was not the result of consideration. It was obtained by computer experiment with the natural language text. It turned out that the statistical regularities in texts (the experiments were conducted on different texts of the Russian language) correspond to the above equations [Teslia, 2005]. And only then, the author has found a theoretical explanation for this coincidence. In any case, the only criterion of truth is practice. Therefore, we consider the application of the theory of non-violent interactions for construction of artificial intelligent systems.

Applications of non-violent interaction to build reflex intelligent systems

As of now a number of "smart" programs and systems that solve intellectual problems are developed. But there is no significant progress in establishing "artificial intelligence". Enormous capital investment often ends with scientific rather than beneficial to the business practical results. One of the reasons is seen in the fact that there are no simple tools for implementing the basic intellectual functions of living organisms. No tools to develop reflexes to external stimuli [Caparres, 1995; Teslia, 1998]. We must learn to store statistical information about the necessary reactions to certain external influences the same way as does the human brain and develop on the basis of information, the correct response to new (including those that occur for the first time) combination of external influences. Under the proper response is understood to satisfy the intellectual system by reaction to external stimuli.

The central assumption of the theory of non-violent interaction is the assumption of the unity of the laws of interaction for any form of existence of matter. In non-violent interaction theory the category of the internal organization of matter – introformation is introduced as the root cause (source) displaying material formations in Nature. Some of the physical laws were presented through this category. Is it possible to do the opposite: to Use introformational representation of physical laws to describe the information interaction of people. Expressions for refilling of certainty (16) - (17) follow from the addition of relativistic velocity. And based on the law of momentum conservation the law of the closed certainty of the material formations is obtained (15).

But the question arises. Do these formulas show deeper patterns in the construction of the laws of nature? Laws of wisdom of the Universe, which are embodied in particular in physical laws. If you imagine the reaction of material objects on collision is a reflex, the formation laws of which lie in the depths of nature, than it could be that the laws of reflex making in the objects of nature are the same? And then these formulas can be used to calculate the "human behavior"? If nature has constructed this way the physical laws, by transforming introformational content of matter, it may be, that, the human brain works on that basis? Maybe each neuron shows its internal organization based on the same laws that show material formation in motion of its introformation? Perhaps the realization of the above operations with introformation underlies the workings of the brain? And maybe they can be used to create artificial intelligent systems.

Using VIP-interpretation the method was proposed (hereinafter - introformational method) for constructing intelligent reflex systems, which react to exposure based on the same laws that all matter responds to the impact of different physical nature. More precisely, the calculation of the effect of the reaction is the a new introformational interpretation calculation of material object speed after simultaneous collision (impact) with a variety of other objects, if you know how to change the speed of the object after the collision with each of these items separately. Method is a sequential execution of the following calculations (with italic the physical nature of the calculations are described) [Teslia2, 2010]:

1. On known probabilities of the reactions (actions) of the system its definition (6) with respect to these reactions is calculated. Let's determine

p_0 - The unconditional probability of the reaction x;

$p_j = p(x/y_j)$ - The probability of the reaction x, if the action was y_j.

$$d_j = \begin{cases} +0{,}5\cdot\sqrt{\dfrac{p_j}{1-p_j}+\dfrac{1-p_j}{p_j}-2}, & p_j \ge 0{,}5 \\ -0{,}5\cdot\sqrt{\dfrac{p_j}{1-p_j}+\dfrac{1-p_j}{p_j}-2}, & p_j < 0{,}5 \end{cases}, \quad j=\overline{0,n},$$

where d_j, $j=\overline{1,n}$ – is certainty of the reaction x, if the action was done to the system y_j (d_0 –certain reactions in the absence of actions to the system).

At this point, a transition from the speed of the material objects to their certainty, if the object impacted on y_j. For example, a collision with an object, the impact of which is referred to as y_j.

2. On known probabilities awareness of (5) is calculated with respect to these reactions

$$i_j = \frac{1}{2\sqrt{p_j \cdot (1-p_j)}}, j = \overline{0,n} ,$$

where i_j - the awareness of the system relative to the reaction x, when exposed to y_j (i_0 – awareness of the system relative to the reaction x, in the absence of action on the system).

And here is a transition from speed of the material object to its knowledge, if the object y_j was impacted.

3. The calculation of the total, based on all impacts to the system, the increment of certainty of the system. Introformational representation of momentum conservation (15) and (16) and the formula of the addition of relativistic velocity are used

$$\Delta d = \sum_{j=1}^{n}(d_j \cdot i_0 - d_0 \cdot i_j) = \sum_{j=1}^{n} d_j \cdot i_0 - \sum_{j=1}^{n} d_0 \cdot i_j = i_0 \sum_{j=1}^{n} d_j - d_0 \sum_{j=1}^{n} i_j , \quad (18)$$

where Δd – total increment of certain reactions you received from the law of conservation of momentum

Performed calculation increment certainty if the total exposure to the object (the clash happens with all objects at the same time).

4. The calculation of the increment of awareness of (7)

$$\Delta i = \sqrt{\Delta d^2 + 1} ,$$

where Δi – the increase of awareness of the system.

It is calculated for the specified increment of awareness in terms of the method in point 3.

5. Calculation of a new determination of the reaction x. The identity of the calculation of the new relative speed of the object after its collision with all objects. Using (17)

$$d_\Sigma = \Delta d \cdot i_0 + d_0 \cdot \Delta i ,$$

where d_Σ – a new determination of the reaction x, resulting from the law of conservation of momentum.

Calculation of a new determination of the motion of the object x.

6. Calculation of a new awareness of the system action. Use the formula (7)

$$i_\Sigma = \sqrt{d_\Sigma^2 + 1} ,$$

where i_Σ – a new awareness of the system derived from the law of conservation of momentum.

Calculation of a new awareness of the movement of the object x.

7. The calculation of the relevant physical laws of probability of reaction x (8)

$$p_\Sigma = p \ (x/Y) = 0{,}5 + \frac{d_\Sigma}{2i_\Sigma} ,$$

where $p_\Sigma = p(x/Y)$ – derived from the law of conservation of momentum reaction probability x, in operations $Y = \{y_j\}, j = \overline{1,n}$.

The calculation of a new probability of a shift in the direction of the object (which determines the speed of the movement after a collision with all objects).

The idea of the above method is that it points to the expected "reaction" to the impact, the adequacy of which comes up from the well-known and experimentally verified physical laws. By assumption, the interaction of neurons is based on the same laws and implemented in accordance with the proposed model non-violent interaction. And on this basis it is possible to create artificial introformational processors operating as neurons. Such neurons are not the same as a well-known in classical cybernetics formal neurons, which are also similar to the natural, like a paper boat on the ocean ship. A more advanced and complex structures that respond to stimulation (effects) as well as material formations and react in inanimate nature, and natural neurons. All this is embodied in a number of **reflex intelligent systems** that can store information about the environment and to develop an adequate functioning of the reaction (reflexes) on everything in this environment [Teslia, 2005; Teslia2, 2010; Teslia, 1998; Teslia3]. This is one of the arguments in favor of recognizing the wisdom of the laws of Nature. And pretty much confirms the hypothetical theory of non-violent interaction.

What applications of the theory of non-violent interactions have been implemented? The theory of non-violent interaction and its application to the development of artificial intelligence systems were examined, in particular, in [Teslia, 2005; Teslia2, 2010; Teslia, 1998; Teslia3]. Statistical analysis of the significant amount of text in Russian was performed. It was found that the probability of the different length of text fragments completely correspond to equations (15) - (17), which may be indicative of the unity of the laws of interaction in nature. In any case, such a correspondence, at least, is quite interesting. Besides theory has practical development in the design and operation of the reflex intelligent systems: evaluation of investment proposals in development, natural language access to databases, evaluate the impact of harmful substances in the water resources of the region on public health; predicting results of sports events, voice control by technical means. The main advantage of these systems is an ease of development (development cost is lower than creating traditional expert systems), and the effectiveness of solutions of various intellectual tasks. More details with these systems can be found at the web site *introformatika.org.ua.*

Conclusions and prospects for further research

The theory of non-violent interaction is based on the hypothesis of the primacy of internal organization (introformation) of material objects in the processes of interaction and movement. In this case, the expression obtained in the theory is mathematically beautiful, simple, and give a reasonable explanation of many physical laws and paradoxes. During the study the author founds similarities in the interaction in the physical and self-managed systems, similarity, which formed the basis of this work, and which may be of interest not only to specialists in the field of computer science, but also to experts in the field of theoretical physics, philosophy, biology, computer science and computer engineering. And that similarity, in my opinion, is a consequence of the unity of the laws of interaction in nature. The important thing is that based on the non-force model of interaction in nature, you can create a fundamentally new system of artificial intelligence for many areas of human activity. After all, the theory of non-violent interaction reveals the root causes and the laws of interaction including the basic elements of the human brain - neurons. It is hoped that this article will help many researchers, engineers apply their knowledge to further progress in solving the basic problems of cybernetics - the construction of the disclosure laws and mechanisms of the brain and on this basis to create artificial systems that are not inferior to their "intelligence" of man.

The next article will talk about the reflex intelligent systems built using the mathematical apparatus of the theory of non-violent interaction.

Bibliography

[Ay, 2008] N.Ay, N.Bertschinger, R.Der, F.Guttler, and E.Olbrich Predictive information and explorative behavior of autonomous robot. The European Physical Journal B - Condensed Matter and Complex Systems, 63 (3): 329-339, 2008.

[Brillouin, 1972] Brillouin L.new look at the theory of relativity. - Wiley, 1972, p. 142 (In Russian).

[Caparres, 1995] Gonzalo Joya Caparres, Francisco Sandoval Hern ndez. A Neural Paradigm for Controlling Autonomous Systems with Reflex Behaviour and Learning Capability. In Jos Mira, Francisco Sandoval Hern ndez, editors, From Natural to Artificial Neural Computation, International Workshop on Artificial Neural Networks, IWANN 95, Malaga-Torremolinos, Spain, June 7-9, 1995, Proceedings. Volume 930 of Lecture Notes in Computer Science, pages 283-290, Springer, 1995.

[Hartmut , 1989] Hartmut H. The application of information theory in physics. - Academic Press, 1989, p. 347 (In Russian).

[Kamshilov, 1979] Kamshilov N.M. The evolution of the biosphere. - Moscow: Nauka, 1979, p. 256 (In Russian).

[Klapchenko, 2011] Klapchenko V.I. Probabilistic interpretation of mechanical motion / V.I. Klapchenko, Y.M.Teslya / / Cornell University Library. - [Electronic resource] - [arXiv:1102.0441], 2011

[Kuznetsov, 2001] Kuznetsov N.A. Informational cooperation in technical and living systems. Information Processes, Vol.1, № 1, 2001, pp. 1-9 (In Russian).

[Kuznetsov, 2006] Kuznetsov N.A., Baksan O.E., Grechishkina N.A. The fundamental importance of science in the modern scientific worldview. Information processes, T.6, № 2, 2006, p.81-109 (In Russian).

[Kuznetsov, 2011] Kuznetsov N.A., Baksan O.E., Zholkov S.Y. The origins and pragmatic knowledge base. Information processes, T.11, № 4, 2011, pp. 428-447 (In Russian).

[Kuznetsov1, 2003] Kuznetsov N.A., Lyubetsky V.A., Chernavskii A.V. On the concept of information exchange, 1: underphysical level. Information processes, V.3, № 1, 2003, pp. 1-22 (In Russian).

[Kuznetsov2, 2003] Kuznetsov N.A., Lyubetsky V.A., Chernavskii A.V. On the notion of information interaction, 2: underphysical level. Information processes, V.3, № 2, 2003. - P.154-172 (In Russian).

[Oleksienko, 2009] Oleksienko M.M. The method of forecasting the number of illnesses from exposure to harmful substances, based on the model of non-violent interaction. East European journal of advanced tehnology.-Kharkov. - 2009. - № 1.-P.34-38 (In Ukrainian).

[Redko, 2007] Redko V.G. Computer science and biology - the science of the 21st century. That at the interface? / V.G.Redko / / Information Processes, V.7, № 3, 2007, pp. 214 - 247 (In Russian)

[Teslia, 1998] Teslia I.N. Reflex system for processing natural language text in the ACS building complex energy facilities. Electronics and Computer Science. - 1998. - № 4.-P.52-55 (In Russian).

[Teslia, 2005] Teslia I.N. Non-force interaction. Monograph. Kiev: Condor, 2005, pp. 196 (In Russian).

[Teslia, 2009] Teslia I.M. Application of the reflexive approach to building intelligent systems for evaluating investment proposals. - Ukrainian national collection of scientific papers: mining, construction, road and agricultural machines. № 73-2009. - P.82-87 (In Ukrainian).

[Teslia, 2012] Teslia I.N. Dynamic properties of introformation. Managing the development of complex systems, 2012, №. 9, pp. 44 – 49 (In Russian).

[Teslia1, 2010] Teslia I.N., Teslia O.V. Conceptual apparatus of the theory of non-violent interaction. Managing the development of complex systems, 2010, №. 1, pp. 46 – 52 (In Russian).

[Teslia2, 2010] Teslia I.N. Introduction to computer science of nature. Monograph. Kiev: Maklaut, 2010, p. 256 (In Russian).

[Teslia3] Teslia I.N. Demonstration of GUT / I.N.Teslia - [Electronic resource]. - Mode of access: http://www.youtube.com/watch?v=m_pYXVndpbc (In Russian).

[Timchenko, 1991] Timchenko A.A., Rodionov A.A. Fundamentals of Computer Systems Engineering machinery projects. - Naukova Dumka, 1991, p. 152 (In Russian)

[Uspensky, 2010] Uspensky V.A., Vyugin V. Formation of algorithmic information theory in Russia. Information Processes, Vol.10, № 2, 2010. - P.145-158.

Authors' Information

Iurii Teslia *- doctor of technical sciences, professor, laureate of the State Prize of Ukraine in Science and Technology, Head of the Department of Information Technology of the Kyiv National University of Construction and Architecture, e-mail: teslyas@ukr.net, tel.: +38 067 503 39 16, work tel: +38 044 241 55 07.*

Major Fields of Scientific Research: Theory of non-violent Interaction, General theoretical information research, Multi-dimensional information systems

TABLE OF CONTENTS