Big Data and Big Control

D.A.Novikov

Abstract
This paper analyzes and structures modern challenges in the field of big data handling, arising before hardware developers, experts in applied mathematics and artificial intelligence, as well as researchers focused on different application domains. The author outlines further development of the theory and applications of big control, i.e., control based on big data.
Keywords Big data, Data analysis, Big control, Big data management

1 Introduction

I am a great believer in the simplicity of things and as you probably know I am inclined to hang on to broad & simple ideas like grim death until evidence is too strong for my tenacity.

C Sir Ernest Rutherford
From Letter to Irving Langmuir (June 10, 1919).

The term "big data" means unstructured data whose volume exceeds the available handling capabilities in required time. This term appeared just a few years ago [1], but became very popular among IT experts, scientists, business analysts and other specialists. For instance, a retrieval request to Google returns millions of links on the subject. What are the capabilities and risks carried by big data? What challenges and problems do they formulate before scientists, experts in different application domains (particularly, specialists in systems science and control theory-big data usage in control problems and big data management), the education system as a whole?

2 Big Data. Big Analytics. Big Visualization

In information technology, *big data* (perhaps, the term was first mentioned in the special issue of *Nature*[1]) represents a direction of theoretical and practical investigations on the development and application of handling methods and means for the big volumes of unstructured data.

Big data handling comprises their[1]:
- acquisition;
- transmission;
- storage (including recording and extraction);
- processing (transformation, modeling, computations and analysis);

[1] In certain classifications, big data handling is associated with 4D (data discovery, discrimination, distillation and delivery/dissemination).

- usage (including visualization) in practical, scientific, educational and other types of human activity.

In the narrow interpretation, the term “big data” sometimes covers only the technologies of their acquisition, transmission and storage. In this case, big data processing (including construction and analysis of corresponding models) is called *big analytics* (including big computations), whereas visualization of the corresponding results (depending on users cognitive capabilities) is called *big visualization* (see Fig. 1).

The universal cycle of big (generally, any) data handling is illustrated by Fig. 2. Here the key role belongs to an *object* and a *subject* (a “customer”); the latter requires knowledge on the state and dynamics of the former. However, sometimes there exists a chasm between the *data* acquired on an object and the *knowledge* necessary for a subject. Primary data must be preprocessed, i.e., transformed into more or less structured information. Subsequently, necessary knowledge is extracted from this *information* depending on a specific task solved by a subject.

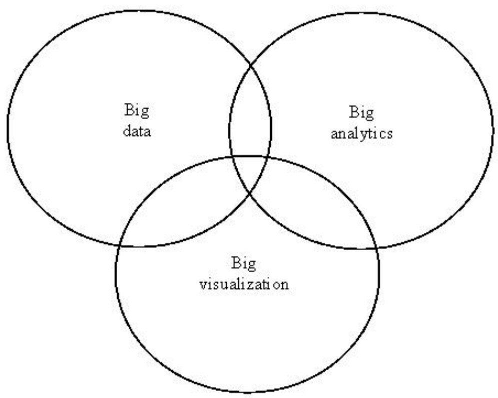

Fig.1 “The big triad”[2]: Data, analytics, visualization

Particularly, a subject may adopt this knowledge for object ***control***, *viz.*, exerting purposeful impacts on an object to ensure its required behavior. Control can be automatic in a special case (an inanimate subject). Perhaps, the term “*big control*”[3] will become common soon for indicating control based on big data, big

[2]We will not discuss another fashionable triad (big data, high-performance computations, cloud technologies

[3]For justice’ sake, note that in the recent fifteen years experts in control theory have tended to consider the problems of control, computations and communication jointly (the so-called C3 problem (Control, Computation, Communication)). According to this viewpoint, control actions are synthesized in real time taking into account the existing delays in communication channels and information processing time (including computations). There is another generally accepted term (large-scale systems control), but big data can be generated by “small” systems.]

analytics and, possibly, big visualization.[4]

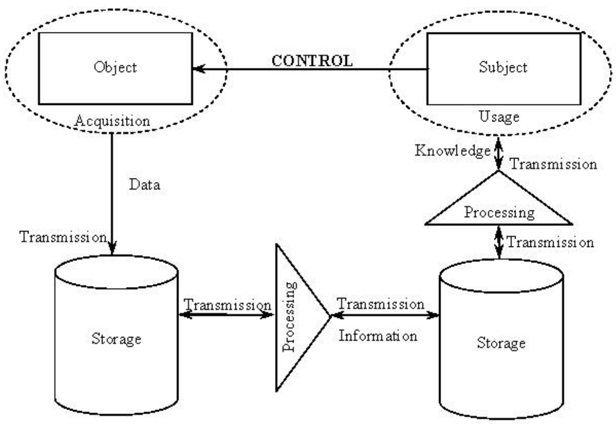

Fig.2 The universal cycle of big data handling

The qualitative analysis of numerous publications on big data leads to the author's subjective expert appraisal for the current distribution of attention paid by researchers and developers (but not users!) to big data handling problems. This expert appraisal is demonstrated in Fig.3.

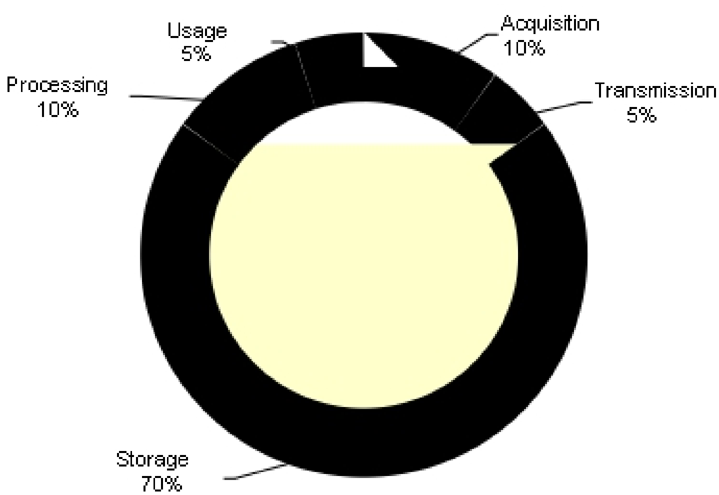

Fig.3 The current distribution of attention paid by researchers and developers to big data handling problems

[4]An alternative interpretation of "big control" concerns control of big data handling processes. Actually, this represents an independent and nontrivial problem.

In other words, the overwhelming majority of big data investigations create the technologies of big data acquisition, transmission, storage and preprocessing, whereas big analytics and visualization receive by far less consideration. Big control problems are almost not studied, despite that control objects become more and more complicated and multi-scale, while "networkism" represents a leading trend of modern control theory [2]!

3 Civilization-Scale Problems

However, is the current state of affairs (see the distribution in Fig.3) reasonable? On the one hand, the answer is affirmative. Indeed, technologies followed exactly this path of development; moreover, data analysis and visualization requires data acquisition and storage (no doubt, with the feasibility of rapid accessing and processing). On the other hand, the existing "disbalance" is the result of the following. Nowadays, mankind realizes the potential utility of any data, but does not completely understand what should be done with the growing avalanche of data.

This problem seems not novel, since a class of similar civilization-scale problems has appeared recently, formally called the *problems of anticipatory technologies development*. To elucidate this idea, consider the correlation among science, technologies and practice (see Fig.4). During different periods of mankind development, science often initiated creation and adoption of certain technologies; sometimes, the chain was "inverse" (presently, we observe exactly this picture!).

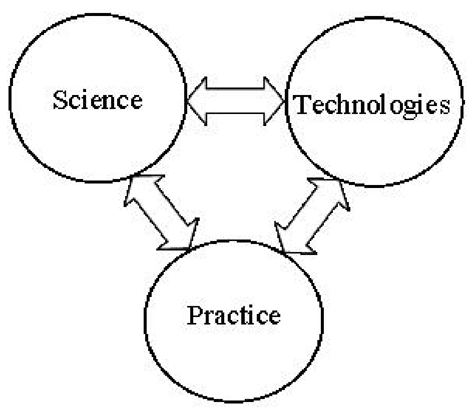

Fig.4 Science, technologies and practice

Really, appeal to history. Starting from the middle of the 19th century, for over 100 years mankind had experienced the triumphal development of science and the technological boom (science anticipated and predetermined technologies, the latter were widely adopted in practice; generally speaking, before science was stimulated by practice).In this context, we mention electricity, communications, nuclear power engineering, electronics, etc. Most technologies were clear and available to average men.[5]

Later on (approximately, since the 1970s) the situation gradually changedCthe accumulated fundamental scientific results were enough for technologies growth (i.e., some technologies even "anticipated" science). Consequently, their "demand for science" was reduced (perhaps, living systems made an exception).

Nevertheless, technologies demonstrate further development, and the pace of their development even increases. For instance, in the recent decade advances in information and communication technologies (ICT) have been even anticipating practice[6], including human recognition of new technological capabilities, development prospects and associated threats.[7]

In other words, the threshold of the third millennium was remarkable for a turning point: mankind had earlier mastered science and technologies to satisfy its needs, whereas afterwards technologies rather dictated the directions, restrictions and conditions of scientific development, at the level of individuals and at the level of states and mankind as a whole.[8] Exactly this effect is called "anticipatory technologies development". Mankind will have to realize the corresponding civilization-scale problems and learn to respond to them.

Similar things happen with big data: human beings have mastered technologies for accumulating giant volumes of data, but are still unable to process and utilize them. The major problem concerns the comprehension of why should big data processed (rather than how should it be done).

To have a clear view of the challenges faced by scientists and engineers, we

[5]Amusingly, the heroes of fantasy books on technological breakthrough, owing to isolation (e.g., Cyrus Smith from J. Verne's *The Mysterious Island*) or a trip to the past (e.g., the hero from M. Twain's *A Connecticut Yankee in King Arthur's Court*), are often engineers or average men but not scientists.

[6]As a positive feature, we acknowledge that advances in ICT were stimulated by the development of applied mathematics, particularly "network mathematics" (random graphs, large-dimensional graphs, network games and others, see the surveys [3-6]).

[7]Scattered manifestations of this phenomenon took place earlier; e.g., A. Noble and the participants of the Manhattan Project thought about the ethical responsibility of scientists.

[8]The problems of information security have become common. Today is the right time for thinking not just about technological security (cybersecurity), but about the social, economic and other security of information technologies, i.e., the security of ICT users, their groups and society as a whole against informational impacts (here a striking example concerns social media including online social networks [7]). Recall that all important decisions (at the level of an enterprise and at the level of a government) are made on the basis of information coming from certain sources and processed by certain methods (far from always, a decision-maker knows these sources and methods). Therefore, we have to acknowledge the relevance of the socioeconomic security of ICT, i.e., the security of an individual, an economy, a society and a country against the consequences of decisions made by modern ICS (including decision support systems applied in economics, military science, politics, etc.).

discuss a series of questions. Which data are actually big? Where do big data arise? What are the modern applications of big data? How can one improve the efficiency of their usage in future? What is the role of science in big data handling?

4 Sources and Customers of Big Data

They include large groups, namely:

- science (astronomy and astrophysics, meteorology, nuclear physics, high-energy physics, geoinformation systems and navigation systems, distant Earth probing, geology and geophysics, aerodynamics and hydrodynamics, genetics, biochemistry and biology, etc.);
- Internet (in the wide sense) and other telecommunication systems;
- business, commerce and finances, as well as marketing and advertising (including trading, targeting and adviser systems, CRM-systems, RFIDCradiofrequency identifiers used in sales, transportation, logistics and so on);
- monitoring (geo-, bio-, eco-; space, air, etc.);
- security (military systems, antiterrorist activity, etc.);
- power engineering (including nuclear power engineering), SmartGrid;
- medicine;
- governmental services and public administration;
- production and transport (objects, units and assemblies, control systems, etc.).

Numerous applications[9] of big data in these fields can be found in popular science literature (or even "glossy" journals) available at public Internet sources. We will not describe these applications here to avoid embarrassing "zettabytes" and "yottabytes".

In almost all fields cited, the modern level of automation is such that big data have automatic generation. Therefore, the following question gains growing importance. What is the volume of "lost" data flows (due to insufficient capabilities or time for their storage or processing)? This question seems correct for an engineer in ICT, but not for a scientist or a user of big data processing results. Rather, the former and the latter would ask "What are essential losses in this case?" and "What are the changes if we successfully acquired and processed all data?", respectively.

[9]The principal idea of using big data is revealing "implicit regularities", i.e., answering non-trivial questions: epidemic prediction based on information from social networks and sales in drugstores; medical and technical diagnostics; retention of clients by analyzing sellers' behavior in stores (the spatial movements of RFID-tags of products); and others.

5 Which Data Are Big? Scientific Challenges

Traditionally, big data are unstructured data whose volume exceeds the available handling capabilities in required time. However, this definition appears somewhat "cunning": data considered big today cease to be such tomorrow owing to the progress of data handling methods and means. Data that looked big several hundreds or even thousands of years ago (in the absence of automatic treatment) are easily processed today by home computers. The competition between the (hypothetic) computational demands of mankind and corresponding technical capabilities has been known very long ago. Of course, the capabilities have been always chasing the needs. And the gap between them represents a monumental stimulus for science development. Researchers have to suggest simpler (yet, adequate) models, design more efficient algorithms, etc.

Sometimes, the definition of big data includes the so-called 5V properties (Volume, Velocity, Variety, Veracity, Validity). Alternatively, the difference between the big volume of conventional data and big data proper is that the latter form the big flow of unstructured[10] data (in the sense of volume and velocity as the volume per unit time).

In the wide comprehension, the unstructuredness of big data (text, video, audio, communications structures, etc.) is actually their characteristic feature and a challenge for applied mathematics, linguistics, cognitive sciences and artificial intelligence. Creation of real-time processing technologies[11], including the feasibility of implicit information revelation, for large flows of text, audio, video and other information forms the mainstream of applications of the above sciences[12] to ICT.

Therefore, we observe a direct (and explicit) query from technologies to science. The second explicit query concerns adaptation of traditional statistical analysis, optimization and other methods to big data analysis. Furthermore, it is necessary to develop new methods with due consideration of big data specifics. A modern fashionable trend is boosting analytics tools (generally, business analytics) for big data. But their list almost coincides with the classical kit of statistical tools (or is even narrower, since some methods are inapplicable to big data). This is also the case for:

- machine learning methods (neural networks, Bayesian networks, fuzzy inference and other logical inference, etc.);

[10] Data unstructuredness can be the result of their omissions and/or different scales of studied phenomena and processes (in space and time, see the so-called multi-scale systems).

[11] In the first place, these technologies must perform data aggregation (e.g., detecting changes in technological data or storing aggregated indices). Really, one does not need all data (especially, "homogeneous" data).

[12] Mathematics rather easily operates structured data; and so, data structuring makes an important problem.

- high-dimensional optimization problems (in addition to traditional parallel computing, intensive research focuses on distributed optimization, l_1-optimization; e.g., see [8-13]);
- discrete optimization methods (here an "alternative" lies in application of multiagent program systems [14]) and others.

The common feature in the stated queries of technologies to science is adaptation or small modification of well-known tried-and-true methods. We have to be aware of the following. Generally, automatic modeling (by traditional tools[13]) based on raw data represents just a fashionable delusion.[14] We expect to suggest algorithms and apply them to bulky volumes of unstructured (often irrelevant) information, thereby improving the efficiency of decision-making. Such delusions occurred in the history of science at the early development stages of cybernetics and artificial intelligence.[15] They resulted in numerous disappointments and considerably impeded the development of these scientific directions. There exist no miracles in science: generally, new conclusions require new models and new paradigms (e.g., see the books on science methodology [15-16]).

6 General Challenge

The complexity of the surrounding world grows at a smaller rate than the capabilities of data detection ("measurement") and storage. Perhaps, these capabilities have exceeded the ability of mankind to realize the feasibility and reasonability of their usage. In other words, we "choke" with data, trying to find what to do with them.

However, there exists an alternative viewpoint of this situation as follows. Obtaining big data (having an arbitrary large volume) is possible and easy enough (obvious examples arise in combinatorial optimization, nonlinear dynamics or thermodynamics, see below). But we have to understand how to manage big data (and ask the Nature correct questions). Furthermore, it is possible to construct an arbitrary complex model using big data and then try to reach a higher accuracy within the model. But the associated dilemma is whether we obtain new results or not (in addition to very many new problems[16]). Long ago math-

[13]An additional encumbrance is the accumulated experience of a researcher/developer and the traditions of its scientific school. Successful solution of a certain problem leads to the conviction that same methods (only!) are applicable to the rest open problems.

[14]In some cases, additional information can be obtained by increasing the volume of data (under correct processing).

[15]The behavior demonstrated by a cybernetic system is always the result of embedded algorithms (stochastic, nondeterministic, etc.), despite the seeming generation of new knowledge or revelation of new ("unexpected") behavior. This is especially the case for interaction of very many elements.

[16]We recognize the importance of model's adequacy and stability of modeling results, but omit these problems.

ematicians and physics knew that increasing the dimensionality and complexity of a model (aspiration for considering more factors and relations among them) does not necessarily improves the quality of modeling results; sometimes, it even carries to the point of absurdity.[17]

Let us study a series of examples.

Example 1. The book [17] by Nobel Prize winner H. Simon considered the following example. Imagine an ant walking along a beach. The ant may try to minimize the efforts required for moving from one point to another; and so, it escapes rocks, sometimes turns back, etc. If we observe just the horizontal projection of the ant's path (without the knowledge of relief), explaining its behavior (a very winding, complex path) seems difficult. H. Simon arrived to an important conclusion. The existing variety and complexity of human behavior are explained not by their complex principles of decision-making (actually, these principles appear simple), but by the diversity of related situations. One would hardly disagree with this opinion. Really, nontrivial results can be provided by a complex model based on simple input data and by a simple model based on complex input data. Ideally, nontrivial results should be generated by simple models owing to the correct choice of relevant and simple input data (mathematicians are used to say that"simplicity is the sign of truth").

Example 2. Suppose that, as if by magic, scientists of the 18th century receive a modern laptop and a tomographic scanner of laptops (with user manuals). The scientists make the tomographic image of the laptop and save it in the memory. By analyzing these bulky and very detailed data on the physical structure of the laptop, they would hardly understand the operation principle of the laptop. A correct paradigm, a correct conceptual model forms a necessary (yet, not sufficient) condition of a success.

Example 3. Take any NP-hard problem of combinatorial optimization [18], e.g., the travelling salesman problem. There are about $n!$ variants to-be-analyzed for exact solution search; in the case of $n\sim100$, the number of variants exceeds the computational capabilities of mankind. The network excitation control problem for a real social network possesses the dimensionality of $n\sim 10^6$[19]. This property of NP-hard problems has been known for decades, motivating researchers to develop (generally, heuristic) methods of approximate solution search with estimated guaranteed accuracy at reasonable time. A similar example concerns nonlinear dynamics models: the"observations" of dynamic chaos demonstrated by a rather simple (medium-dimensional) nonlinear dynamic system may require the memory of all computers in the world, but generate no new knowledge.

Example 4. A classical example from the history of physics is the discovery of

[17] Not to mention situations, when existing scientific paradigms make it impossible in principle to model system behavior on a large time horizon (e.g., accurate weather forecasting).

the law of universal gravitation. For two decades, T. Brahe (1546−1601) observed planetary motion in the Solar System. For those days, his records represented big data. Based on Brahe's records, I. Kepler (1571−1630) formulated his empirical (!) laws of planetary motion:

- the orbit of a planet is an ellipse with the Sun at one of the two foci;
- the square of the orbital period of a planet is proportional to the cube of the semi-major axis of its orbit.

Keplers laws of planetary motion aggregated Brahes data, and the motion of any planet could be calculated by them (instead of Brahe's many-volumed recordings) with a high accuracy. In other words, Brahe learned to describe[18] planetary motion, whereas Kepler learned to describe and predict it. However, Kepler's laws did not explain why planets move so. The answer was later given using the law of universal gravitation established by I.Newton (1643-1727). Perhaps, Kepler's laws would be derived by any modern computer (with adequate algorithms embedded and initial data entered). At the same time, modern computers would be unable to obtain the law of universal gravitation without the corresponding model of mass interaction. Kepler's laws are the "corollaries" of the law of universal gravitation (can be deduced from it), just like Brahe's results follow from Kepler's laws. Therefore, the law of universal gravitation made useless (epistemologically superfluous) both Kepler's laws and Brahe's big data.

Today's experience in big data handling testifies that, in most cases, we are at the level of Brahe and apply titanic efforts to reach the level of Kepler. But a qualitative jump occurs only under the appearance of generalizations (at the level of Newton) that radically simplify the situation. Once again, note that the whole secret is"putting correct questions to the Nature."

Example 5. The second classical example from the history of physics relates to the development of molecular-kinetic theory. It shows that generation of a big data flow appears easy. The whole point is what to do with this flow and what questions to answer.

Consider the following mental experiment (the problem of ideal gas behavior description). One cubic meter of air is in the normal conditions. It contains approximately 10^{25} molecules. The motion of molecules and their collisions are completely described by kinematics and dynamics (within the framework of the ideal gas model). In other words, there exist no fundamental obstacles in such description. During a second, each molecule suffers from 10^9 collisions with other molecules. The real-time description of such a system (the coordinates and velocities of all molecules) requires the minimum data flow of 10^{35} bytes per second.

[18] Recall the basic functions of scientific cognition (particularly, modeling) [16]: description (the phenomenological function)- explanation- prediction (the prognostic function) - control (the normative function).

This data flow exceeds even the modern processing capabilities of mankind (to say nothing of the processing capabilities in the 1850s when the theory appeared)! Perhaps, physicians recognized the senselessness of such detailed analysis (even despite its possibility in principle) and passed to macrodescription in terms of aggregated characteristics (temperature, volume, pressure) and, subsequently, to the description of probabilistic distributions within the framework of statistical physics. But if physicians of that period were able to perform necessary calculations, science would never have statistical physics!

Today a similar effect arises, e.g., in informational control problems for online social networks. Transition from microdescription using graphs with tens of millions of nodes and billions of connections (no doubt, this is **a big control problem**) to macrodescription in terms of probabilistic distributions [20] leads to realistic theoretical study, still preserving the key properties of an object.

And finally, note the following. Depending on their source, big data can be natural or artificial. In the former case, data are generated by some independent object and we ("investigators") decide what should be "measured" (Examples 1, 2 and 4). In the latter case, the source of data is a model (Examples 3 and 5); complexity (data flow) is partially controlled and defined during simulation.

"Recipes." There exist four large groups of subjects (see Fig.5) operating (explicitly or implicitly) big data in their professional (scientific and/or practical) activity: - manufacturers of big data handling tools (software/hardware developers, suppliers, consultants, integrators, etc.);

- designers of big data handling methods (experts in applied mathematics and computer science);

- specialists in application domains (scientists focused on real objects or their models) that represent big data sources;

- customers utilizing or planning to utilize the results of big data analysis in their activity.

Representatives of the mentioned groups interact with each other (see the dashed lines in Fig.5). The normative ("ideal") division of "responsibility areas" is illustrated by Fig.6; here the thickness of arrows corresponds to the level of involvement.

Proceeding from sensus communis and not claiming to be constructive, we formulate the following general "recipes" for the listed groups of subjects.

For manufacturers of big data handling tools: with the course of time, it will be difficult to sell big data solutions (including analytical ones) without suggesting new adequate mathematical methods and stipulating for the feasibility of close cooperation between customers, the developers of appropriate methods and specialists in application domains.

For mathematicians (the author's "brothers-in-arms"): a topical query con-

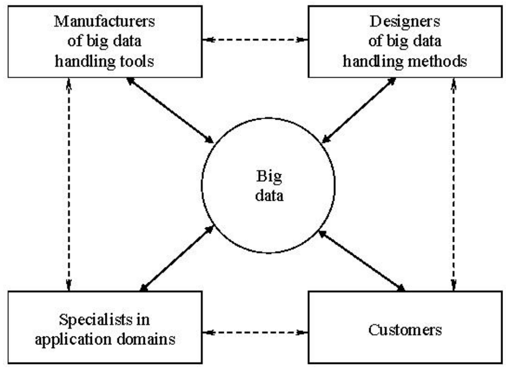

Fig.5 Subjects operating big data

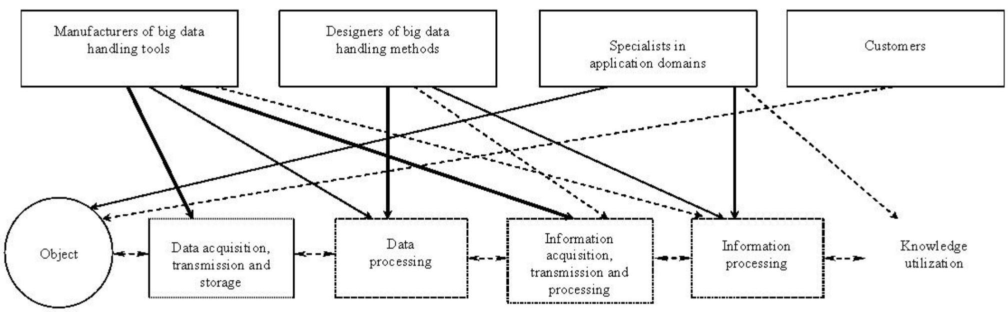

Fig.6 Division of "responsibility areas"

cerns adapting well-known methods and developing new processing methods (in the first place, with nonlinear complexity!) for the large flows of unstructured data representing a good testing area for new models, methods and algorithms (to the extent possible, at the expense of manufacturers and/or customers).

For specialists in application domains: big data technologies lead to new capabilities for acquiring and storing the bulky arrays of "experimental" information, conducting the so-called computing experiments; the associated methods of applied mathematics enable systems generation and rapid verification of hypotheses (revelation of implicit regularities).

For customers: the expensive technologies of big data acquisition and storage would hardly be economically sound without involving specialists in appropriate methods and subject areas (only if it is absolutely clear which questions a customer would like to answer using big data[19]).

Some threats. In addition to the emphasized necessity of searching for ade-

[19]Though, it is possible to store data de bene esse (e.g., to verify a certain hypothesis in future based on them).

quate simple models and the alerting trend of anticipatory technologies development, we expect the future relevance of the following problems (the list below is unstructured and incomplete).

• *The informational security of big data.* This requires adaptation of well-known methods and tools, as well as development of fundamentally new ones.

• *The energy efficiency of big data.* Even today, data processing centers represent a considerable class of power consumers. The bigger are data to-be-processed, the higher is energy needed.

• *The principle of complementarity* was established in physics long ago; it declares that measurements modify the state of a system. However, does it apply to social systems whose elements (people) are active, i.e., possess their own interests and preferences, choose their actions independently, etc. [21]?

A demonstration of this principle lies in the so-called information manipulation (strategic behavior). According to theory of choice [4,22], an active subject reports information by forecasting the results of its usage; generally speaking, an active subject does not adhere to truth-telling.

Another example concerns the so-called active forecasting: a system changes its behavior based on new knowledge about itself [23].

Are these and similar problems (e.g., crowdsourcing [4,24], conformity behavior [25], etc.) eliminated or aggravated in the case of big data?

• As far as we have mentioned the principle of complementarity, it is necessary to recall *the principle of uncertainty* in the following (epistemological) statement [16]: the current level of science development is characterized by certain mutual constraints imposed on results “validity” and results applicability. In the context of big data, this principle means the existence of a rational balance between the level of detail in the description of a studied system and the validity of results and conclusions to-be-made on the basis of this description.

• A traditional assumption in design and operation of information systems (corporate systems, decision support systems of governmental services, inter-agency circulation of documents, etc.) is that all information in such systems must be complete, unified and publicly available (under existing access rights). But it is possible to show the “distorting-mirror” reality to each person, i.e., to create an individual informational picture[20], thereby performing informational control [4, 7, 21, 23]. Should we strive for or struggle against these effects in the field of big data.

[20]At the very least, a fragment of the “objective” picture (hushing up the whole truth); at the most, an arbitrary inconsistent system of beliefs about the reality.

7 Instead of Conclusion

Thus and so, data have always been big. New, more and more perfected tools of big data acquisition, storage and processing appear intensively. We wish we managed to perform these operations in real time. This calls for developing appropriate directions of applied mathematics and computer science (a topical query from technologies and practice to modern science).

There also exists a pressing need for mass training of specialists in big data, big analytics and big visualization (with focus on concrete application domains).

However, this is not enough: we have to accumulate knowledge (in appropriate branches of science) and create models for compact and adequate description of studied phenomena and processes (with due consideration of a solved problem). In other words, it is desired to advance from the level of Brahe to the level of Newton in each possible area of big data applications. Otherwise, we are doomed to handle particulars, not seeing the wood for the trees (see E. Rutherford's citation above).

Furthermore, the anticipatory development of technologies has become the civilization-scale problem to-be-accounted by scientists and engineers (including the field of big data and big control), as well as by the consumers of appropriate methods and tools created by them.

References

[1] Special Issue. (2008), *Nature.*

[2] Forrest J. and Novikov D. (2012), "Modern trends in control theory: networks, hierarchies and interdisciplinarity", *Advances in Systems Science and Application*, Vol.12, No.3.

[3] Dorogovtsev S. (2010), *Lectures on Complex Networks*, Oxford University Press, Oxford.

[4] Gubanov D., Korgin N., Novikov D. and Raikov A. (2014), *E-Expertise: Modern Collective Intelligence*, Springer

[5] Jackson M. (2010), *Social and Economic Networks*, Princeton Univ. Press, Princeton.

[6] Novikov D. (2014), "Games and networks",*Automation and Remote Control*, Vol.75, No.6, pp. 1145-1154.

[7] Gubanov D., Chkhartishvili A. and Novikov D. (2010), Social Networks: Models of Informational Influence, Control, and Confrontation, Ed. by D.Novikov, Fizmatlit, Moscow, Russian.

[8] Noam Nisan and Tim Roughgarden. (2009), *Algorithmic Game Theory*, Ed. by N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani. - N.Y, Cambridge University Press.

[9] Boyd S., Parikh N. and Chu E. et al. (2011), "Distributed optimization and statistical learning via the alternating direction method of multipliers", *Foundations and Trends in Machine Learning*, No.3, pp.1-122.

[10] Granichin O.N. and Pavlenko D.V. (2010), "Randomization of data acquisition and $\ell 1$-optimization (Recognition with Compression)", *Automation and Remote Control*, Vol.71, No.11, pp. 2259-2282.

[11] Nesterov Y. (2012), "Efficiency of coordinate descent methods on Huge-scale optimization problems",*SIAM Journal on Optimization*, Vol.22, No.2, pp.341-362.

[12] Polyak B.T. and Tremba A.A. (2012), "Regularization-based solution of the pageRank problem for large matrices", *Automation and Remote Control*, Vol.73, No.11, pp.1877-1894.

[13] Shoham Y. and Leyton-Brown K. (2009), *Multiagent systems: Algorithmic, Game-Theoretical and Logical Foundations*, Cambridge University Press, Cambridge.

[14] Wooldridge M. (2002), *An Introduction to MultiAgent Systems*, John Wiley & Sons, N.Y.

[15] Kuhn T. (1962), *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.

[16] Novikov A and Novikov D. (2013), *Research Methodology: From Philosophy of Science to Research Design*, CRC Press, Leiden.

[17] Simon H. (1996), *The Sciences of the Artificial,3rd Edition*, The MIT Press.

[18] Garey M. and Johnson D. (1979), *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and company, San Francisco.

[19] Novikov D. (2014), "Models of Network Excitation Control",*Procedia Computer Science*, Vol.31, pp.184-192.

[20] Breer V., Novikov D. and Rogatkin A. (2014), "Micro- and Macro-Models of Social Networks", *Control Sciences*, No.5, pp.28-33.

[21] Novikov D. (2013), *Theory of Control in Organizations*, Nova Science Publishers, N.Y.

[22] Aizerman M. and Aleskerov F. (1995), *Theory of Choice*, Elsevier, Amsterdam.

[23] Novikov D. and Chkhartishvili A. (2014), *Reflexion and Control: Mathematical Models*, CRC Press, Leiden.

[24] Surowiecki J. (2005), *The Wisdom of Crowds*, Anchor, N.Y.

[25] Breer V. and Novikov D. (2013), "Models of Mob Control",*Automation and Remote Control*, Vol.74, No.12, pp.2143-2154.

Corresponding author
D.A.Novikov can be contacted at:novikov@ipu.ru