Formation of research strategy and synthetic research evaluation based on the strategy

— Toward research program evaluation as a creative activity —

Naoto KOBAYASHI[1]*, Osamu NAKAMURA[2] and Kenta OOI[3]

[Translation from *Synthesiology*, Vol.4, No.1, p.11-25 (2011)]

Formation of research strategy, and synthetic research evaluation based on the proposed strategy have been considered. The importance of a setup of targets and a scenario of the research program to achieve the targets as a part of strategy formation, and the importance of research evaluation consistent with the research strategy are emphasized. Research evaluation should be performed in three aspects – the research progress, the research depth and the research phase. In the individual evaluation aspect, comparison of the research performance with the research strategy framework is essential and synthetic evaluation appropriately composed of deductive inference, inductive inference and abductive inference[Term 1] is recommended. To make the final integrated evaluation, the synthetic method is very crucial. Examples of research unit evaluation at AIST, and the research strategy formation and evaluation of public research organizations in Nagasaki prefecture are compared with the synthetic evaluation method. The method is thought to be a creative activity that can contribute to extract the value of research and accelerate the future evolution of research programs.

Keywords: Research strategy formation, research program construction, synthetic research evaluation, abductive inference, reflection and chain of evaluation, logic model

1 Introduction

In the 21st century, the environment where the earth and human community reside is subject to more pressure than in the 20th century. For the humankind to survive into the future, there are many issues to be solved by science and technology. Therefore, humankind needs a strategic approach in terms of how to promote science and technology. The evaluation of research and development (R&D) is extremely important when performing the R&D along a research strategy. Particularly, the evaluation must be performed to make the comparison with the research strategy, and to adequately present the significance and direction of the R&D being based on analysis and integration. Such research evaluation may spread further than the assumption made in the strategy, and has the potential of making the evaluation into a creative activity. Therefore, it is beneficial to propose a synthetic methodology for evaluation. It will be greatly useful if such evaluation method can be used to link the technology to an innovation.

As the social significance of R&D increases, the work on strategy is becoming active in various countries. The "Energy Innovation Hub"[Note 1] [1] that shows the strategic direction of the R&D on future energy in the United States, the "Construction of the knowledge-based European Economic Society" at the Lisbon Strategy[Note 2] [2], and the "New Growth Strategy" in Japan[Note3] [3] are examples. However, the "research strategy study" that provides the academic foundation of research strategy has not been established, and that is a major issue of research and its evaluation.

In the past researches on research evaluation, the application of the "logic model"[Term 2] that indicates the logical sequence of input, output, outcome, and impact that were used for program evaluation for nearly 20 years has become active, and recently research on the method of performing evaluation according to the logic model has been conducted[4][5]. This has been effective as the evaluation method for research programs conducted through public grants in the United States and Canada, and is an excellent method because it clarifies the external logical structure that surrounds research. However, it does not evaluate the content of the research. On the other hand, peer review and bibliometric methods are available as means to evaluate the research itself. The former is a method in which the peers, or the experts of the research field, evaluate the contents and results of the research, while the latter is an evaluation through measurable values such as the number of papers, citations, or patents. In many cases these methods are combined to perform the actual evaluation[6][7]. However, research on the reaseach evaluation from the basic perspective of how to watch the research has not been so far sufficiently performed.

In this paper, we present the outline of the thoughts on how to synthesize research evaluation, starting from the perspective of properties of research, and the formation of research strategy and the logical construct of the elements

1. Center for Research Strategy, Waseda University 513 Wasedatsurumaki-cho, Shinjuku-ku 162-0041, Japan * E-mail : naoto.kobayashi@waseda.jp, 2. Science and Technology Promotion Bureau, Nagasaki Prefectural Government 2-13 Edo-machi, Nagasaki 850- 8570, Japan, 3. Evaluation Department, AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba 305-8563, Japan

Original manuscript received February 25, 2010, Revisions received December 24, 2010, Accepted December 24, 2010

related to the research evaluation based on such strategy. Particularly, the "synthetic research evaluation" is important in evaluating the essence of research and in advancing the research, regardless of whether a research is basic or applied, or analytic or synthetic.

2 Properties of research

Research is thought to possess "intrinsic properties". These properties include: 1) novelty, 2) originality, 3) logical completeness, and 4) influence. 1) Novelty is to add new academic findings not limited to a certain academic field, and 2) originality is that the research itself provides unique findings and introduces new theses. Research that gives totally new explanation to a known phenomenon may be low in novelty but is high in originality. 3) Logical completeness means that the research is a complete expression upon accumulation of solid logic. 4) Influence expresses the strength of the effect to the external field. It includes influences that affect the discipline itself, and those that affect other disciplines (these are effects on the academic field); and the influences that affect society. The latter influence will be called "practical properties".

The *Type 1 Basic Research*[Term 3] proposed by Yoshikawa is basic research that produces new knowledge in a certain discipline, and the influence is mainly contained within the discipline. *Type 2 Basic Research* (and the *Product Realization Research*) is research with practical properties that affect society, but both can be discussed within the same intrinsic property[8][9]. However, these two types of basic researches are not always clearly separated, and the elements of both can be found within a research project. Also, there is research where the influence of the intrinsic properties may stay within a discipline for a short period, but may become socially influential after some time. For example, it is well known that the various spatial and temporal compensations are done in devices that receive signals from the GPS (Global Positioning System) satellites, and this is based on the theories of special and general theory of relativity presented by Albert Einstein at the beginning of the 20th century.

3 Research strategy and research program

3.1 Significance of the research strategy and its formation

By defining strategy as a "way in which an objective is set; various elements such as people, resource, time, and information are allotted appropriately; and these are organically combined and activated in order to make the whole system function properly", the research strategy can be defined as "the strategy deployed to set and achieve the objective of research and its influence".

In forming the research strategy, it is desirable to set the specific research program for achieving the goal of the strategy, as well as the targets, scenario to attain the targets, and the targets for the individual research project that comprise the research program. The research program, according to Hirasawa, can be defined as "the unit of execution, development, and management of the policy that is structuralized and logically constructed, linking the policy and the research project[6]". Here, a research program will be defined widely as the "unit of research development that is structuralized and logically constructed to link the objective of the research strategy and the research project". Therefore, the research program may be also applied in the *Type 1 Basic Research* such as the experimental elementary particle study.

To what extent the targets and the scenario of the research program are set must be agreed upon by the research promotion group and the research sponsor (state or society in case of public research), and a preliminary contract should be signed. It is also important to incorporate a review process of the research strategy during the progress of the research program.

Also, assuming unforeseen circumstances that may occur in the research progress, some degree of redundancy should be included in the contract. The scenario must include several options and flexibility in schedule. Even in *Type 1 Basic Research* that is research within a discipline, for example, the research strategy can be formed and the potentials of the results may be cosmic and the influences may last for a long term. Even if a result different from the major assumption is obtained, the value of the research strategy will be determined based on how much it contributed in increasing the scientific knowledge of the discipline.

One example is given. The main objective of the Kamiokande, for which Dr. Masatoshi Koshiba received the Noble Prize in Physics in 2002, was to prove the proton decay by detecting the collision of the neutrinos that were released in proton decay. However, Koshiba *et al.* fortuitously detected, for the first time in the world, the neutrino produced by the supernova explosion that occurred in the Large Magellanic Cloud with the Kamiokande in February 1987. This demonstrated the correctness of the theoretical model of supernova explosion, and the era of neutrino astronomy started. While Koshiba had initially indicated the possibility of observing the neutrinos from space, the proton decay has not been observed even in the succeeding Super Kamiokande. As this example indicates, it is normal in science that an expected result is not necessarily obtained. However, the research strategy, in which a neutrino detector was made with a tank with 3,000 ton of pure water and 1,000 photomultiplier tubes under the Kamioka Mine, turned out to be very significant because it added new knowledge to physics[10].

In forming the research strategy, it is necessary to organize

the crucial issues structurally into global issues that affect the wide-ranging field over long-term, social or domain issues that are mid-term issues for a nation, region, or an academic discipline, and the research programs that are conducted in relatively short-term in some specific field.

An example of the global issues is the "realization of a society with sustainable development". It is possible to consider a layer structure containing various issues within this subject. Four issues of sustainability – "environment, energy, and natural resource", "human, organism, and food", "society, economy, and industry", and "information, culture, and education" – are identified as shown in Fig. 1. In these layers, the subject of sustainability changes from natural to artificial as it moves from the lower to the upper layer. The items in the lower layers need urgent attention, but an integrated effort must be made to cover all layers in the national, social, and international policies.

What is important in forming the research strategy is to clarify the individual social issues that are reflected in the real world, to define and to break them down into research programs and individual research projects, and then to visualize their relationships. Figure 2 shows an example. The four points of sustainability are given as global issues, and the examples of related social issues and research programs are also shown. This method is a top-down or deductive method. The reason why it is called "deductive" is because the items that are assumptions of the strategy, such as the preservation of environment (realization of low-carbon society) in the context of the aforementioned sustainable society, can be deployed by the social agreement. However, complete deductive inference is not possible, and abduction[11]-[13] will always be in action[Note 4].

On the other hand, in the formation of the research strategy, there is also the bottom-up strategy formation from experience, findings, and future vision of the researchers in the specific research field. For example, if it is possible to develop optical switch devices with ultra-low energy consumption, and an optical path communication network using these devices, the power consumption of the current Internet can be reduced by three orders, and a scenario that can make a great contribution to the realization of a low-carbon society can be written[14]. While this is one of the elemental technologies, the research program can be based on the realization of a group of elemental technologies. Although this can be called the inductive strategic formation where the main proposition is formed from the individual facts and certain logical inferences, this also includes the inference by abduction. That is because several hypotheses are necessary to create a specific system by bundling the current technologies. The abductive research strategy formation is necessary as the juncture of deductive top-down synthesis and inductive bottom-up synthesis.

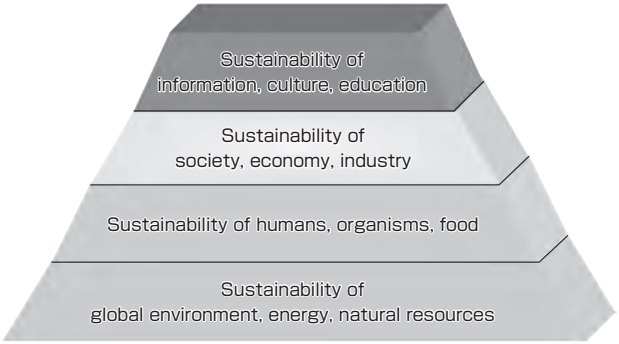

Fig. 1 Layer structure of sustainable development

3.2 Synthesis of the research program

For the synthesis of the research program, it is necessary to consider on which domain the center of the research should be set within the academic framework. To solve the global and human issues (for example, industrial development and environment), it is impossible to rely on a single discipline, and we must muster knowledge of various fields including those of humanities and social sciences. Therefore, the domain will have a wide research field. In setting the research program, it is necessary to indicate the target of the research program set by the research strategy and the specific scenario to achieve the target. A roadmap is such a presentation along with the milestones that should be achieved along the time axis. It is necessary to set the individual research projects that comprise the research program α, as shown in Fig. 3. In this case, the program is composed of a group of research projects (A, B, and C) that are topics of the several research domains.

4 Structure of the research evaluation

4.1 On research evaluation

In considering the research evaluation, ideally it should help the researchers and the research advancement, and, on the other hand, it should be used as an effective method to obtain

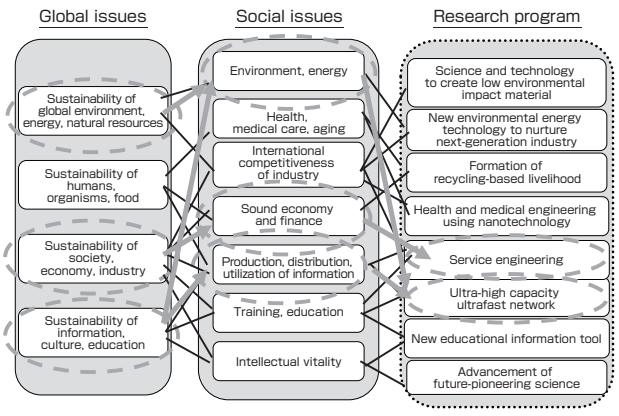

Fig. 2 Example of strategy formation

the understanding and cooperation of society, by making the "value" of the research visible in society even if it is basic research. The research evaluation should be performed based on the following way of thinking.

(1) As the basic thinking of the research evaluation method: 1) whether it is basic research, applied research, or experimental development[Term 4 [15]] or *Type 1 Basic Research*, *Type 2 Basic Research*, or *Product Realization Research*[8][9], the research and the evaluation should be done under one integrated way of thinking; 2) the research evaluation should be based on a contract [in which the purpose of the research is written as the strategy and scenario and shared by the parties (research fund provider, research group, and research evaluator), and the evaluation should be performed accordingly]; and 3) the evaluator and the evaluated should cooperate on the same plane. Also, (2) the aim of the research evaluation should be: 1) to bring out the value of the research; 2) to push the research forward; 3) to be the source of motivation for the researchers and their group; and 4) to be accountable to the sponsors and stakeholders.

The following is an explanation of the "synthetic evaluation", after due consideration of the characteristic of the research evaluation. The "synthetic evaluation" can be defined as the evaluation method that synthesizes the comprehensive evaluation of the research by clarifying the properties of the various aspects of research evaluation (this is called elemental evaluation) and by clearly positioning their relationship structurally. For the methods of research evaluation, refer to the comprehensively written article by Ohtani[16].

4.2 Research program/project and its evaluation

For the evaluation of R&D, it is necessary to understand the characteristics of the evaluation at each stage of the research. The main ones are: 1) appraisal, 2) process evaluation, 3) output evaluation, 4) program evaluation, and 5) outcome evaluation. Figure 4 shows what kind of evaluation and feedback are done in the series of R&D processes based on the strategy, from the strategy formation, construction/ execution of program, creation of output, achievement of program, and creation of direct outcome. This series of processes is almost the same as the program evaluation method known as the ROAMEF[Term 5 [6][16]].

In performing the research evaluation, the appraisal based on foresight is particularly important. In the appraisal, the adequacy of the research development scenario and the research program based on the strategy, the plan and content of each research project in accordance with the research program, and the system, resource, time, place and others for the execution are considered carefully. Particularly, it is important to clarify the aim of the research program, and to clarify the dynamic relationship among the plans, resources, expected outputs of several research projects included in the research program, and how the individual research projects interact with each other (sharing and utilization of the project results).

For the appraisal, it is necessary to perform an inductive and abductive evaluation to determine the strategy along a highly probable inference. One example is, when creating a research strategy for reduced greenhouse gas, low-cost production, and high export competitiveness by developing a new technology using a certain organic material. While it can be inductively estimated that the potential for realization is high, judging from the current material and device performance, there is a counterargument that the durability is an issue. Then, it is possible to perform the appraisal of the strategy based on the inference of the hypothesis that the durability can be dramatically improved by developing a technology to avoid contact with moisture and oxygen.

On the other hand, the flexibility of the scenario to respond appropriately to the changes in the situation will be important. However, achieving both the strictness and flexibility of the strategy and scenario may not be easy, and how to embed such flexibility is one of the issues.

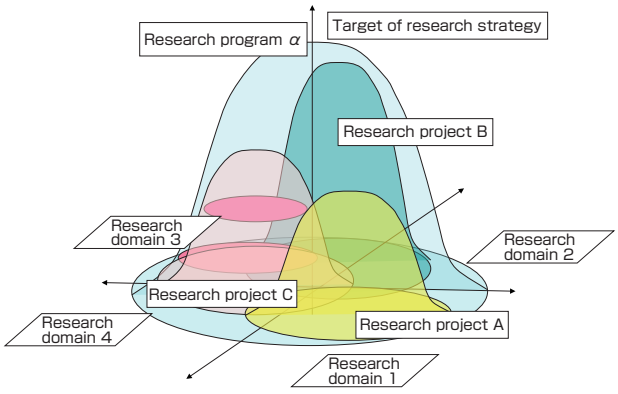

Fig. 3 Research projects that comprise the research program and their properties

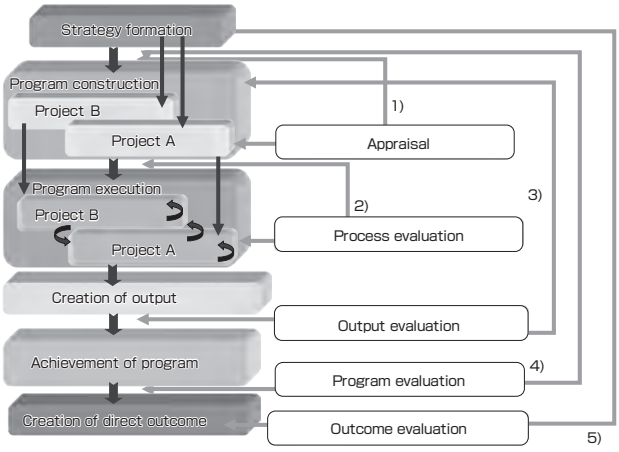

Fig. 4 Evaluation during the process from strategy formation to outcome

In the process evaluation, it is necessary to have a dynamic response such as feedback where the correction is made if there are problems and the recommendation to collaborate with other projects, along with checking the progress of the individual project. In the output evaluation, how the results obtained by the achievement of the program compare with the initial program target is checked. Here, the peer review that will be explained later is important in the *Type 1 Basic Research*, while the evaluation by the experts and stakeholders will be important when the main topic is social effect.

In program evaluation, the target of the research program set in the strategy and the execution of the scenario are investigated. In direct outcome evaluation, the direct outcome produced when the output of the research program is handed to the external party is compared with the target of the strategy. However, the creation of the direct outcome normally takes time after the completion of the program.

The feedback loop (FBL) is important in the evaluation process. In the FBL1, the issues extracted in the appraisals are fed back and reflected in the program construction. FBL2 is one of the PDCA (plan-do-check-act) cycles[Term 6] at the project level during the program execution. Here, the progress check at the individual project level is reflected in the course adjustment of the project and review of the invested resource. FBL3 and 4 are loops where the contents of the program evaluation and the output evaluation are reflected in the program formation in the next step. FBL5 is the process that uses the direct outcome evaluation in the adjustment of the research strategy and the formation of new strategies.

The research project has a simpler structure or function compared to the research program. Therefore, the research objectives, methods, results, and the expected outcomes are bound within a small range, and the above evaluation process can be applied since they possess a fractal structure upon the research program. However, the research project is positioned as an element of the research program, and the appraisal can be simplified.

4.3 Synthetic research evaluation and its application

4.3.1 Overview diagram

Figure 5 shows the overview diagram of research and evaluation for the execution of the research program based on the research strategy formulation. First, the research evaluation is considered analytically by breaking it down into elemental evaluations.

The X-axis is time that shows the progress of the research. Here, the process from program building to output creation will be simplified, and a research program will be composed of three blocks of plan, process, and results. The evaluation along this line mainly determines whether the research progressed according to the process assumed by the strategy, in terms of plan, process, and results. Here, the evaluation will certainly be on the content of the research, but even more so on whether management was done for the effective progress of the research. Not only is it necessary to deductively determine along the rule of the agreed strategy, but it is also necessary to conduct evaluation that encourages various trials and devising. The appropriate evaluators in this case will be the peers and experts with experience in research progress, and it is particularly desirable to have someone experienced as a research program leader.

The Y-axis shows the depth of research. The depth of research in terms of results means the quality of 1) novelty, 2) originality, 3) logical completeness, and 4) influence, which are the four properties of research explained in chapter 2. For the plan and process, they are the density and the vastness of the prospect of the plan for which high expectation is expected and the progress of the research that may lead to important results. Here, the evaluators must be peers in the same discipline or in multiple disciplines. Different evaluators are needed for each step for the Z-axis (phase) that will be explained below. In the case of the pure basic research phase, a good evaluator is a peer within the same discipline, but as the phase approaches the outlet to society, experts of industry and journalism will be necessary. The influence will be the scale of social effect and the potential for having such effect.

The Z-axis shows the phase of research. Phase is an index that shows where the research is positioned from the basic research to the social exit. For example, the research can be categorized into basic research, applied research, and experimental development[15], or can be categorized as the aforementioned *Type 1 Basic Research*, *Type 2 Basic Research*, and *Product Realization Research*. The evaluator must have knowledge about the content of the research and the significance of the strategy for each phase, and must be able to consider the potential for realizing the outcome. With the evaluation of this axis the results alone are not evaluated, but the results and the processes for arriving there, as well as the road to utilization of the results expected in the future are

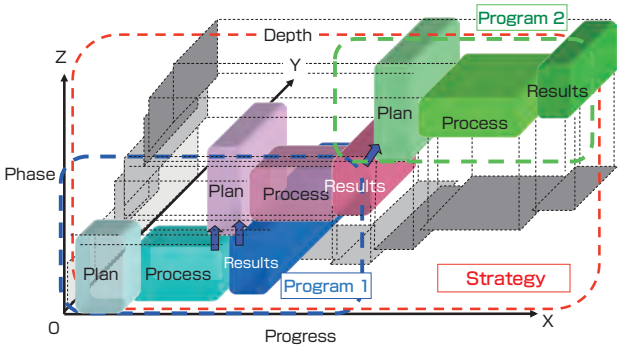

Fig. 5 Overview diagram of the synthetic research evaluation along with strategy formation and research program execution

also evaluated. In that sense, the basis of evaluation would be the targets and scenario set down in the strategy and the roadmap that specify them, and the evaluation would be the significance of the results in each phase (*Type 1 Basic Research*, *Type 2 Basic Research*, and *Product Realization Research*). In the case where the influence of the result stays within one or several disciplines, the influence is in the lowest level of the Z-axis. In this case, it is evaluated according to the depth axis by the peers. In summary, this overview shows the axes of the elemental evaluations, and the relationships of various elemental values in the synthetic evaluation are positioned structurally.

On the other hand, the positioning of the research becomes clear by applying the mode of the R&D. Figure 6 shows the phase-time-space arrangement composed of the XYZ axes of R&D. To synthetically link the evaluation of each axis to the final evaluation, it is necessary to refer to the primary strategy. In the formation stage of the strategy, it is important to clearly state which part of the phase-time-space the result was intended to cover. The three-dimensional structure shown as the transparent block in Fig. 6 is the estimate diagram of the result of research program considered in the strategy. This estimate is obtained as a result of the deductive, inductive, and abductive inferences as mentioned before. On the other hand, in the same diagram, the solid blocks in various colors indicate the actual results of research. The comparison between the transparent and the solid blocks can be linked to the final evaluation.

4.3.2 Actual practice of synthetic research evaluation

The evaluation at the individual evaluation axis is conducted through comparison with the target and scenario indicated by the strategy for that axis. For the progress evaluation at the X-axis, the evaluation index will be how the progress of the research matches or departs from the schedule intended and planned strategically. For example, if there is a departure from the plan within the expected time schedule, and if acceleration is necessary through the "selection and concentration" of research resources or by narrowing down the product through management, the abductive process

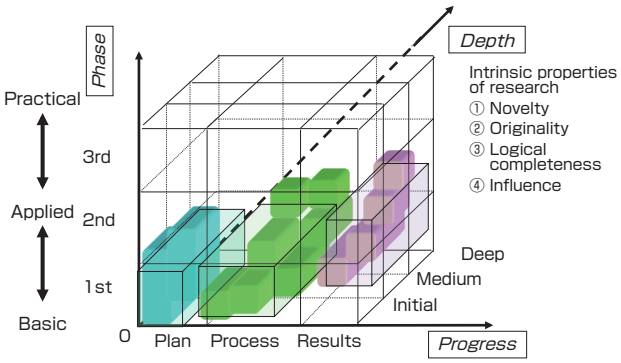

Fig. 6 Conceptual diagram of the research evaluation synthesis

is required in the sense that the effects must be estimated abductively.

For the depth of research in the Y-axis, the intrinsic properties of the research are evaluated, while the inductive decision based on the knowledge and experience of the peers plays a major role for the individual elements. In general, the "excellence" of the research will be determined, and abductive inference is necessary. Of the intrinsic properties of the research, novelty, originality, and logical completeness can be evaluated fairly objectively by peers, while the influence of the research depends largely on the evaluator's abductive inference or imagination. This is because the evaluation of the intrinsic value of research is determined only after the evaluator performs the abductive inference including thinking about the influence.

In many cases of research, serendipity or unforeseen results may play a major role. This is not foreseen in the strategic plan, and it is extremely high in 1) novelty and 4) influence, among the four properties of research stated in chapter 2. The evaluation of such results may be high as they often surpass the planned range in terms of the depth of research.

For the evaluation axis of the research phase in the Z-axis, the evaluation index will be the social effect (in case of *Type 1 Basic Research*, it is the impact in the academic field, and this overlaps with the Y-axis evaluation), and further abductive inference and evaluation are necessary. That is because the social influence requires values that are accepted by society in addition to the intrinsic values of the science and technology research. It can be said that for this determination, there are more abductive elements than in the evaluation of progress (X-axis) or depth (Y-axis) of the research.

For the overall synthetic evaluation, it is necessary to comprehensively understand the elemental evaluations as mentioned above, and then synthesize the integrated evaluation. In that case, as mentioned earlier, the deductive inference that depends on the logical conclusion, the inductive inference that derives the conclusion from several specific examples, as well as abduction must be utilized, and the combination of the inferences to investigate the potential of the value of the results is important.

4.3.3 Integrated research evaluation

In conducting the synthetic evaluation, it is important for the research promoting group and the evaluators, to consider the property of research, share common understanding of the strategy and result index including the goal to be achieved, engage in deep discussion on the results and how the research should progress, check the target indicated in the strategy and distance from the actual results, and finally abductively discuss and investigate the significance and the effect of the research program execution. This entire process

can be considered the "research evaluation = abduction and its expression". This is closely related to the research that is a creative activity, and research evaluation can be considered to be part of this creative endeavor.

To apply the synthetic evaluation to actual evaluation, some devising is necessary. This must be designed upon agreement by the three parties consisting of the research promotion group, research fund provider, and the research evaluator. While the proposal of the paper was not directly applied, a general evaluation was conducted in the example of "Evaluation from the viewpoint of outcome" that was conducted during the Second Mid-term Research Goal Period (FY 2005~2009) at AIST, by introducing the elemental evaluation of management evaluation, output evaluation, and roadmap evaluation that could be thought to be related to progress (X-axis), depth (Y-axis), and phase (Z-axis)[17]. However, it was not necessarily an elemental evaluation by comparison with the strategy as proposed in this paper. There was a lack of deep discussion with the evaluators, and it was still in a developing stage as a synthetic research evaluation. Specifically, it is necessary to design a general evaluation system including the incorporation of the constructive, deep discussion between the research promotion group and the evaluators, in addition to the elemental evaluation and appropriate synthesis.

5 Example of the synthetic research evaluation

5.1 Characteristics and issues of the research evaluation at AIST

5.1.1 Evaluation from the viewpoint of outcome

In the Second Mid-term Research Goal Period (FY 2005~2009) at AIST, "Evaluation from the viewpoint of outcome" was performed upon proposal of the Research Evaluation Investigation Committee (Chairman Rei Hirasawa) in 2004, to focus on the perspective of contribution of the R&D activities to industry and society[17]. In the process of design, 1) the roadmap evaluation, 2) the output evaluation, and 3) the management evaluation were set as the three items of elemental evaluation. The major part of the synthetic evaluation of this paper is the idea based on this experience. As written in chapter 4, 1) the roadmap evaluation corresponds to the Z-axis (phase) evaluation, 2) the output evaluation to the Y-axis (depth) evaluation, and 3) the management evaluation to the X-axis (progress) evaluation. Figures 7 shows the example of the input, output, outcome, and impact in the R&D at AIST as shown as the logic model related to research strategy formation.

The research strategy at AIST is formed from both the top-down and the bottom-up viewpoints, as already mentioned[18]. The outcome is defined at the research unit considering the relationship with the strategy. Looking at the work at AIST from the viewpoint of outcome, of course, the main work is the R&D activity for the creation of advanced knowledge and technology, but the activity for outcome creation by providing the results to the external parties can also be positioned as important work. At AIST, the latter is called the innovation hub function[19], and an innovation hub strategy is established with the research strategy.

5.1.2 Specific examples

At AIST, the target to be achieved as the result of strategy formation is separated into seven items, and the R&D is conducted systematically by breaking down each target into the strategy target, strategic issues, and priority issues. The process from the strategic targets to the research unit issues is designed in a top-down style based on the external environmental factors such as the social demand and marketability, and on the internal factors such as the technological portfolio and strength of the core technology. On the other hand, the results expected from the R&D are designed in a bottom-up style, looking at the process from output to outcome in the R&D from the individual research issues.

As a specific example, the "Research of biomarkers for brain cell functional molecules", a topic of the Neuroscience Research Institute, AIST is explained. This research is characterized by the clarification of the molecular behavior of the ion channel, receptor, and intercellular cell-signaling molecules in neurological diseases, and the search and identification of the biomarkers that specifically recognize the functional proteins. The roadmap for this topic is shown in Fig. 8. The roadmap shows the chronological development of the technologies from the R&D conducted at AIST (search of the marker, functional evaluation, and sensing core technology), to the development of the diagnostic and prevention system of the cerebral disease (development of sensor, cerebral disease risk diagnosis technology) through collaboration with companies.

The Evaluation Committee of the Research Unit is composed of external members (about five people) from universities,

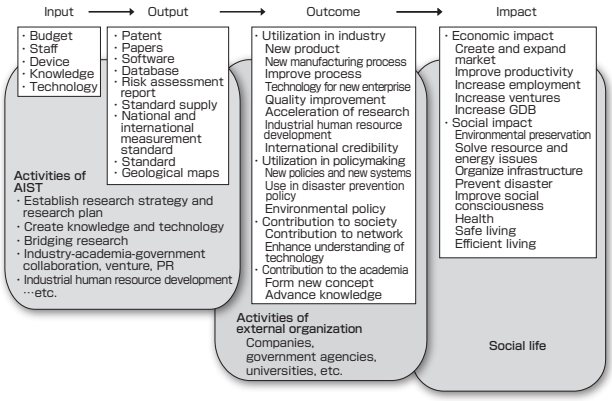

Fig. 7 Model of R&D and result diffusion at AIST

industries, and journalism, and internal members (about two chief evaluation officers from the Evaluation Department), to evaluate the roadmap and output, and the adequacy and appropriateness of the management. At the Evaluation Committee of FY 2008 (the fourth year of the five-year Second Mid-term Research Goal Period), there were positive comments for the roadmap such as "clear" and "shows advantages", and the research plan was judged to be appropriate. However, some members requested "clarification of the positioning in the overall roadmap" and "clarification of the milestone"[20]. For output, "discovery of new peptide through original molecular evolution technology", "results for the unique bioactive peptides such as ant and spider toxins", and "development of signaling substance receptor ligand sensor" were highly evaluated as new technologies that could be utilized for clinical application and as analysis tools for neurological function.

As other examples, the heads of research units including the Nanoarchitectonics Research Center and the Correlated Electron Research Center have reported that they obtained beneficial indications and guidance for the research unit activities from the evaluation from the viewpoint of outcome. The details are described in Reference [21]. In the former case, a clear scenario was written by focusing on the viewpoint of outcome, and the opinions of the evaluation committee members helped establish the strategic R&D plan. In the latter case, the task of forming a roadmap in *Type 1 Basic Research* helped to build the logical framework for the research progress, and it was deemed appropriate to set the "construction of new theory" as one of the outcomes. More beneficial results are expected through synthetic research evaluation in the future.

5.1.3 Phase evolution of the R&D and reflection of the synthetic research evaluation

Figure 9 shows the modeling of the above case from the perspective of phase evolution of the R&D. In practice, the R&Ds with different phases such as the "search of biomarker", "sensor development", and "application to diagnostic technology" progressed concurrently at different temporal development. The R&D topics can be modeled as the cycle of "accumulation of knowledge", "accumulation of elemental technologies", and "accumulation of product realization technology". The key technology linked the R&Ds of different phases, and the quality of this technology greatly influenced the level of R&Ds that developed in the new phase. For the movement between the phases, the fusion of technology through collaboration with the external parties (external knowledge, elemental technology, use of product realization technology, etc.) became the important management items in outcome creation.

In the evaluation from the viewpoint of outcome, it is necessary to aptly grasp the overall picture of the R&D and the result, and to have an integrated evaluation of the results with different qualities. It is also important to perform evaluation from the perspective of the adequacy and effectiveness of the overall composition of the R&D, as well as the effectiveness of the management to smoothly turn the R&D cycle of different phases.

In the evaluation at AIST as shown above, the synthetic research evaluation is being employed as a result, but there are still issues in the research evaluation based on the research strategy which is one of the major characteristics. Particularly, to evaluate the R&Ds at AIST in general, it is necessary to have a fine formation of the research strategy and the synthetic evaluation based on the comparison with the strategy. The issues that must be solved include: 1) the clarification of the final goal of the research strategy and the scenario to get to the goal, and the clarification and the evaluation axis (XYZ axes) based on the comparison; 2) the matching of the evaluation from the bottom-up perspective (inductive evaluation from the viewpoint of outcome for the result) and the evaluation from the top-down perspective (setting of the topic from target to the issue and the deductive evaluation of the roadmap); 3) the clarification of the correlation between the different phases; and 4) the clarification of the evaluation method on the phase axis (Z-axis) considering the characteristics of the research (such as basic, industrial application, policy, etc.). These are related

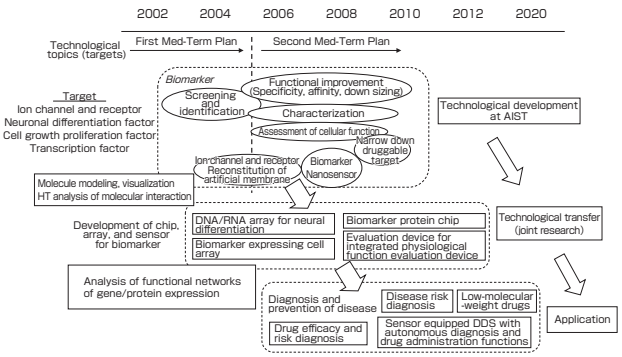

Fig. 8 Roadmap for the research of biomarker for brain cell functional molecule

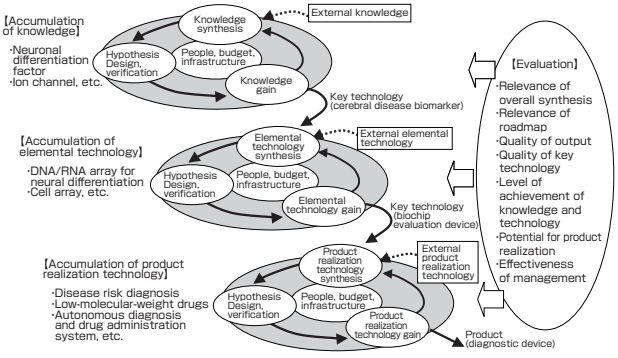

Fig. 9 Phase evolution of R&D and reflection of synthetic evaluation

to the issues of research strategy formation, composition and execution of the research program, and the elemental and synthetic evaluations, described above, and they should be fine-tuned for the evaluation of the research units for the Third Term[22].

5.2 Example of Nagasaki prefecture- Strategy formation and evaluation of the promotion of science and technology for local revitalization

5.2.1 Strategic vision and logic model

The public research institutes in the local government play a role in the important missions to promote the regional industry, and have unique issues as well as common issues shared with the R&D conducted by the government, universities, and companies. Here, an example of the research strategy formation and evaluation of the public R&D of the Nagasaki Prefectural Government, in which one of the authors (*i.e.* Nakamura) was involved, will be presented.

The companies and product districts of Nagasaki suffer from various issues such as population decline and low income. As the recovery from the global economic crisis triggered by the Lehman Shock in 2008 is slow, it is necessary to utilize the unique local resources unseen elsewhere to win the fierce competition and to build a sustainable society. Therefore, it is mandatory that the research institutes quickly capture the social needs, and consider them in the selection of the R&D topics. To do so, it is necessary to build strategic visions, set the clear goals, and to write the scenario for achieving them. Here, the case of promotion of strategic R&D[23] applying the logic model[4] will be introduced. The main point to employ the logic model is to set the strategic R&D in order to produce the results that can be accepted by the customers. The strategic R&D scenario to create the outcome must be completed through the series of tasks mentioned above[24].

5.2.2 Mission and strategy formation of the Science and Technology Promotion Bureau of Nagasaki Prefectural Government[25]

(1) The role of prefectural research institutes

The Science and Technology Promotion Bureau of Nagasaki Prefectural Government is an organization that governs the five research institutes: Institute for Environmental Research and Public Health, Industrial Technology Center, Ceramic Research Center, Institute of Fisheries, and the Agricultural and Forestry Technical Development Center[26]. The missions of the Promotion Bureau are 1) to nurture competitive and strong industry, and 2) to realize a safe and comfortable life through the application of science and technology, hence to build an energetic Nagasaki prefecture where people can live with dreams for the future. To do so, as the long-term outcome, the creation and aggregation of new industries in Nagasaki through the changes in the industrial structure are necessary. As elements, it is necessary to develop policies to create new businesses and new industries, strengthen the existing industries by utilizing the local resources, increase productivity of the prefectural industries, and expand employment. As the mid-term outcome, the current companies must take one step forward to advance into new fields, develop original products, establish the brand, and increase shares. As the short-term outcome, the companies must increase technological prowess, save energy and cost, and increase ability in marketing and design. The prefectural research institutes must collaborate with the universities and others to advance and increase precision of the core technologies, engage in R&D and technological supports for systemization, as well as provide technological development through interdisciplinary researches, marketing and design supports and others according to the needs of enterprises. The above items were organized through the logic model (Fig. 10) and this was shared with the related departments.

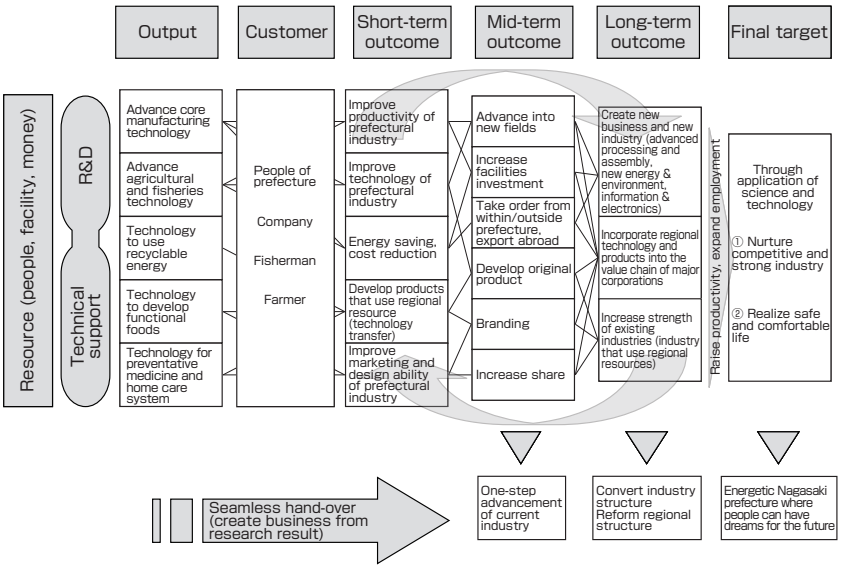

Fig. 10 Logic model for establishing how a research institute should function
(created by the Science and Technology Promotion Division, Nagasaki prefecture)

Using this diagram as the template, the application to the strategy formation at the Ceramics Research Center[27] will be explained below.

(2) Strategy formation at the Ceramics Research Center
The Ceramics Research Center of Nagasaki was established in Hasami-cho, Higashi-Sonogi-gun in 1930. Its mission is the development and promotion of the ceramics industry in Nagasaki. The Center works on the R&D of new materials and recycling of waste, the product development of new fields through fusion with new technology, and the advancement of the core manufacturing technology to support the ceramics industry. Of the logic models organized for the industrial supports of the ceramics field, inorganic material field, and the design field, the diagram for the ceramics field is shown (Fig. 11). In this logic model, the "Development of ceramics core technology and new products" is one program, and the research topics for individual output (for example, "technology for developing a new light-weight porcelain body") correspond to the individual research projects. The mid- to long-term outcome corresponds to the mid- to long-term strategy.

The missions of the Ceramics Research Institute is the vitalization of the ceramics industry, and sets the mid- to long-term outcome to increase the brand power, to pioneer new markets through ceramic products with new functions, and to obtain domestic shares that beat the other product districts in the competition. As the short-term outcome, it is necessary to reduce production cost, develop high quality and high-value added products, develop new fields, develop products in response to changing lifestyles, and provide advanced supports. Figure 11 is a detailed summary that shows what is demanded as R&D outputs to be handed over to the customers in order to gain the short-term outcome. To produce such outputs, it is necessary to engage in the strategic R&D to develop the ceramics core technologies and new products, and to provide technological supports matching with the production sites.

5.2.3 Reflection of the research enterprise evaluation[28] and future issues

For the research enterprise evaluation at Nagasaki, the strategy formation for which the necessity is stated in this paper is just beginning, and the introduction of the thoughts of synthetic research evaluation is a future issue. Currently, the research enterprise evaluation by external committee members is conducted by the ordinance, to ensure researches that reflect the needs of the people and industry of Nagasaki with close observation of the market. This is also utilized for changing the thinking of the prefectural workers. As the evaluation scheme, for the individual research conducted by the research institutes of Nagasaki, the appraisal, interim, and ex-post evaluations are performed from the perspective of necessity, efficiency, and efficacy. The evaluation is done by the Research Evaluation Subcommittee (six external evaluators) set for each research institute, and the Research Enterprise Evaluation Committee (eight external evaluators) engage in the meta-evaluation (or evaluation of the evaluation itself) as the parent committee based on the reports from the subcommittees. In FY 2009, the logic model was started to be used to create an overview diagram of all projects handled by the research institutes. The positioning of each project based on the mission of each research institute, and the scenario of how the research result is handed over to the customers to form the outcome are clarified, whether the projects are progressing appropriately along the overall projects of the research institute are explained, and definite evaluations are obtained.

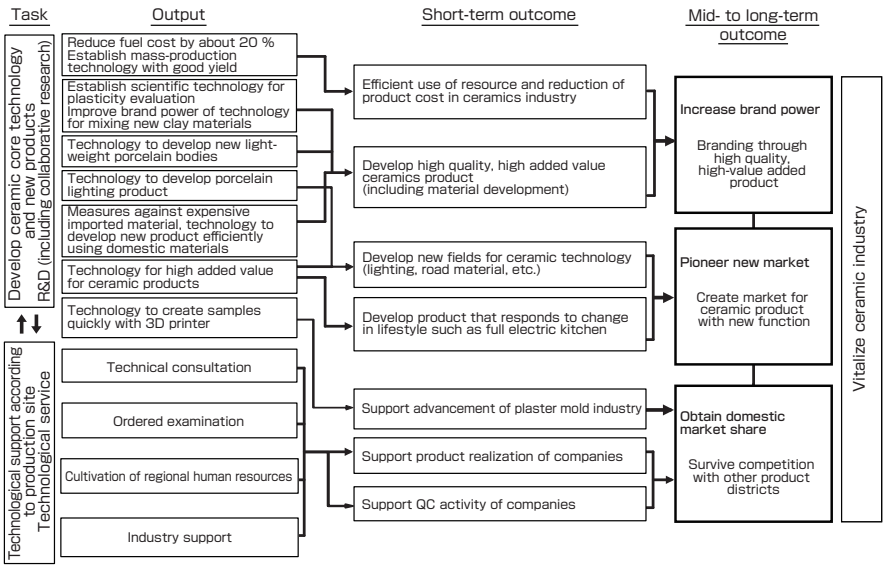

Fig. 11 Logic model for the strategy formation of the Ceramic Research Center
(created by the Ceramic Research Center of Nagasaki)

The evaluation results of the Research Enterprise Evaluation Committee are fed back to the research institutes and used to improve the projects, and are utilized in the discussions for the management of the specific policies indicated in the "Visions of the Promotion of Science and Technology in Nagasaki", proposal of new policies for the promotion of science and technology, and the proposal of strategic promotion fields.

Excavating the hidden needs from wide-ranging fields and creating new technologies that the market needs will lead to long-term and continued economic effects, and may lead to more expansion of job opportunities. For this purpose, it is necessary to maintain strong collaboration with the related departments, and to engage in interdisciplinary strategic R&D by drawing a scenario to achieve the goal. To organize the role and the future image of the R&D and technological supports by the research institutes, it is necessary to optimize the synthesis of the R&D to be undertaken by the research institutes as programs, and to continue the synthetic research evaluation from the perspective of whether the programs are based on a long-term strategy.

6 Reflection and linkage of the synthetic evaluation

One of the greatest responsibilities of the synthetic research evaluation based on the strategy formation is the reflection of the evaluation results. At AIST, the research unit evaluation is performed every two years as mentioned before, and the evaluations are performed for the purpose of (1) encouragement of research in the research units, (2) feedback to the management of AIST, and (3) execution of accountability to external and internal sectors. It is important that the evaluation results are reflected effectively.

Particularly in reflecting the appraisal, the evaluation must be utilized for the optimization of the resources, environment, and conditions necessary for the R&D when establishing research units and starting projects, and in some cases, the targets must be totally revised.

For the research evaluation during the progress, it is important to carry over that evaluation at that point to the next step. To do so, it is important to establish the methodology for rotating the PDCA model, and the most desirable situation is that the evaluation is fed back to the strategy spirally and is carried on to the formation of a new strategy. Moreover, in advancing the R&D, it is necessary to consider where the results are handed over and outcomes are directly produced.

As an issue of research evaluation, it is necessary to have an optimal overall strategic system in which the chain of PDCA from the project to the policy level is utilized effectively by each other. The PDCA cycle with insufficient linkage cannot be considered a functioning strategic research evaluation. For a public research institute, it is important that the evaluation is always linked as a chain from the evaluation of the institute level of whether it fulfills the function expected of the government including the mission and the resource invested, to whether the government (or local government) has a policy to utilize it effectively, all the way to the policy level of whether it is positioned clearly in the innovation policy[28].

7 Conclusion

In this paper, focus was placed on the formation of the research strategy based on the intrinsic properties of research and the R&D program to realize the strategy, and the research evaluation was viewed from the aspects of strategy formation and the synthetic evaluation based on the strategy. In the research evaluation until now, appraisal, interim, ex-post, and follow-up evaluations were performed, and the elements of the synthetic evaluation discussed in this paper had been incorporated. What should be emphasized here are: 1) in the research evaluation, the research strategy is extremely important, and the evaluation by comparison with the strategy should be set as the basic; 2) evaluation from the three aspects - progress, depth, and phase of the research - is necessary; and 3) a synthetic evaluation that summarized these ideas by taking abductive inference is important.

Figure 12 shows the summary of the intrinsic properties of the research, formation of the research strategy, and the synthetic research evaluation based on them that were discussed in this paper, and the aim of research evaluation. The research strategy shows the targets to be achieved by the R&D and the scenario. By conducting a synthetic research

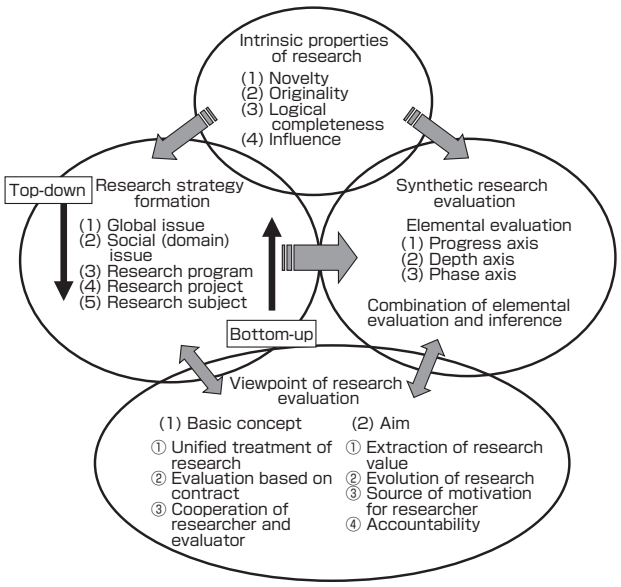

Fig. 12 Research strategy formation and synthetic research evaluation considering the intrinsic properties of research

evaluation based on them, the extraction of the value of research, the evolution of research, the source of motivation for the researchers, and the fulfillment of accountability can be performed effectively.

It would be very significant if the research program evolves through the research strategy formation and the synthetic evaluation proposed in this paper, and then heads toward a new development.

Acknowledgements

To perform this research and to write this paper, we have received great support from the staff of the Evaluation Department, AIST. The central concept of this research is based upon the knowledge and experience in the research unit evaluation performed by the Evaluation Department over nine years, since AIST was reorganized as an independent administrative agency, after the integration of the former Agency of Industrial Science and Technology. We are deeply grateful to Dr. Hiroyuki Yoshikawa, former President of AIST (currently Grand Emeritus Advisor, AIST; Director-General of The Center for Research and Development Strategy, Japan Science and Technology Agency), who provided valuable insight on abduction and synthetic evaluation as well as continuous support. We are also thankful to the staff of the Science and Technology Promotion Bureau of Nagasaki Prefectural Government and the Ceramics Research Center of Nagasaki who created the logic model for the role of public research institutes in Nagasaki prefecture etc..

Notes

Note 1) This project supports the series of activities for the research, development, and practical implementation of clean energy, led by Secretary Steven Chu of the U.S. Department of Energy in December 2009.

Note 2) This strategy was set for the establishment of Framework Program 7 (FP7), an R&D program for the entire EU started in 2007. It was declared at the Council of European Union at Lisbon in 2000.

Note 3) The strategy (basic policy) of Japan was approved by the Cabinet in June 2010. The policy emphasizes green innovation and life innovation as the growth field in which the strength of Japan can be applied.

Note 4) Deduction derives the result (this fruit is delicious) from the rule (an apple is delicious) and the fact (this fruit is an apple). Induction derives the rule (an apple is delicious) from the fact (this fruit is an apple) and the result (this fruit is delicious). Abduction is a method to infer the fact (this fruit is an apple) from the rule (an apple is delicious) and the result (this fruit is delicious). While the deductive inference has a logical structure not dependent on the content, the inductive inference has logic based on several specifics and experiences. Abduction is weaker in logic, and therefore various limitations are set in the conditions by which inferences are made. Yet unlike induction and deduction, abduction is a method with great inference potential. Historically, the discovery of the universal gravitation by Sir Isaac Newton and the elliptical orbit of heavenly bodies by Johannes Kepler are considered typical examples of abduction[11].

Terminologies

Term 1. Deduction, induction, and abduction: Deduction or deductive inference is “to logically derive the true conclusion from the true assumption only through inference without considering the content of inference”. Induction or inductive inference is “to derive a general proposition or law from individual specific facts”. On the other hand, abduction or abductive inference is the third way of inference proposed by an American philosopher C. S. Peirce (1839~1914)[11]-[13], and it is “to derive the individual fact from a certain result and the proposition or law that may cause that result”.

Term 2. Logic model: The logic model is a tool developed to visualize the scenario that is an important element in the strategy formation, to clarify the logical linkage by which the research program achieves its goal. It is employed when the American agencies apply for financing to the Office of Management and Budget (OMB). To write the scenario, an issue is broken down into resource, R&D, output, customer, short-term outcome, mid-term outcome, and long-term outcome. Then, starting from the long-term outcome of the program, the direct outcome that will be produced when the customer receives the result of the R&D is clarified, and the target of the R&D must be spelled out. The logic model is a tool to summarize this process into one big picture[4][5][23][24].

Term 3. *Type 1 Basic Research*, *Type 2 Basic Research*, *Product Realization Research*: According to the definition by Hiroyuki Yoshikawa, *Type 1 Basic Research* is the research to analyze the unknown phenomenon through observation, experiment, and theoretical calculation to establish universal principles and theories. *Type 2 Basic Research* is the research to realize a social value by integrating the knowledge of multiple disciplines, and it also includes research that derives a generalized methodology. *Product Realization Research* is a research that uses the result and knowledge obtained from *Type 1 Basic Research* and *Type 2 Basic Research* and the actual experience to realize the use of the new technology in society[8][9].

Term 4. Basic research, applied research, experimental development: According to the definition of OECD, basic research is an "experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundations of phenomena and observable facts, without any particular application or use in view". Applied research is an "original investigation undertaken in order to acquire new knowledge. It is, however, directed primarily towards a specific practical aim or objective". Experimental development is the "systematic work, drawing on existing knowledge gained from research and/or practical experience, that is directed to producing new materials, products or devices; to installing new processes, systems and services; or to improving substantially those already produced or installed"[15].

Term 5. ROAMEF: Acronym for R (rationale), O (objective), A (appraisal), M (monitoring), E (evaluation), and F (feedback, or review of the cycle)[16].

Term 6. PDCA cycle: A kind of management cycle proposed in the 1950s by Walter A. Shewhart and W. Edwards Deming (U.S.A). It is an acronym for P (plan), D (do), C (check), and A (act). When one cycle is completed, the final "act" is linked to the "plan" of the next cycle so the project will improve spirally. The PDCA cycle can be implemented for a small group or for the entire organization. Ideally, the individual PDCA cycles should link to the higher PDCA cycle[24][28].

References

[1] U.S. Department of Energy website: http://www.energy.gov/hubs/index.htm

[2] European Commission FP7 website: http://ec.europa.eu/research/fp7/index_en.cfm

[3] http://www.kantei.go.jp/jp/sinseichousenryaku/sinseichou01.pdf (in Japanese)

[4] J. A. McLaughlin and G. B. Jordan: Logic models: A tool for telling your programs performance story, *Evaluation and Program Planning*, 22 (1), 65-72 (1999).

[5] G. B. Jordan: A theory-based logic model for innovation policy and evaluation, *Research Evaluation*, 19 (4), 263-273 (2010).

[6] R. Hirasawa: Hyoka kenkyu no sokatsu (Summary of evaluation research), *Material for the 3rd MEXT Mutual Training Session for Policy Evaluation FY 2008* (2009) (in Japanese).

[7] AIST Evaluation Department: *Heisei 17 Nendo Kenkyu Yunitto Hyoka Hokokusho (FY 2005 Research Unit Evaluation Report)* (2005) (in Japanese).

[8] H. Yoshikawa: *Honkaku Kenkyu (Full Research)*, The University of Tokyo Press (2009) (in Japanese).

[9] H. Yoshikawa: A journal for *Type 2 Basic Research*, *Synthesiology*, 1(1), 1-6 (2008).

[10] M. Koshiba: *Nyutorino Tentai Butsurigaku Nyumon (Introduction to Neutrino Astrophysics)*, Kodansha (2002) (in Japanese).

[11] Y. Yonemori: *Pasu No Kigogaku (Symbology of Peirce)*, Keiso Shobo, 173-200 (1981) (in Japanese).

[12] H. Yoshikawa: *Kagakusha No Atarashii Yakuwari (New Role of Scientists)*, Iwanami Shoten, 13-29 (2002) (in Japanese).

[13] Y. Yonemori: *Abudakushon (Abduction)*, Keiso Shobo (2007) (in Japanese).

[14] H. Ishikawa: Technologies for all-optical path routing toward huge capacity and low power consumption networks, *International Symposium on Ultra-high Capacity Optical Communication and Related Optical Signal Processing and Devices*, Sept. (2010).

[15] *OECD Frascati Manual*, Sixth edition (2002).

[16] R. Ohtani: "Science for society" and evaluation of research and technology development, *Synthesiology*, 3 (1), 264-274 (2010) (English version available).

[17] N. Kobayashi: Sangyo gijutsu sogo kenkyusho ni okeru autokamu e no shiten karano senryaku teki kenkyu hyoka to sangakukan renkei e no kadai (Strategic Research Evaluation in AIST from the Perspective of Outcomes and Its Critical Issues for Industry-Academia Collaboration), *Journal of the Japan Society for Intellectual Production*, 4 (1), 14-24 (2007) (in Japanese).

[18] AIST: *Dai Niki Kenkyu Senryaku (Second Term Research Strategy)*, (2005) (2006) (2007) (2008) (2009) (in Japanese).

[19] AIST: *Sansoken No Keiei To Senryaku (Management and Strategy at AIST)* (2005) (in Japanese).

[20] AIST Evaluation Department: *Heisei 20 Nendo Kenkyu Yunitto Hyoka Hokokusho (FY 2008 Research Unit Evaluation Report)* (2008) (in Japanese).

[21] O. Nakamura, S. Ito, K. Matsuzaki, H. Adachi, T. Kado and S. Oka: Using roadmaps for evaluating strategic research and development: Lessons from Japan's Institute for Advanced Industrial Science and Technology, *Research Evaluation*, 17 (4), 265-271 (2008).

[22] AIST Evaluation Department: *Dai Niki Chuki Mokuhyou Kikan Kenkyu Yunitto Hyoka Hokokusho (Second Medium Term Target Period Research Unit Evaluation Report)* (2010) (in Japanese).

[23] O. Nakamura: Imakoso senryaku teki puranningu ni motozuku kenkyu kaihatsu o!- rojikku moderu no katsuyo- (R&D based on strategic planning now! – Use of logic model), *Letter from Industrial Technology Center of Nagasaki Challenge*, 145, Prefatory Note (2009) (in Japanese).

[24] O. Nakamura: R&D evaluation, *Kenkyujo Keiei (Research Institute Management)* (K. Kurata and K. Kawajiri, eds.), Kogyo Chosakai Publishing, 311-336 (2010) (in Japanese).

[25] Science and Technology Promotion Bureau, Nagasaki Prefecture: http://www.pref.nagasaki.jp/kagaku/ (in Japanese)

[26] O. Nakamura, S. Iwai, K. Nishimura and N. Inada: Chiiki kasseika o mezashita nagasaki ban kagaku gijutsu shinko no jitsurei to kosatsu (Case of Nagasaki's promotion of science and technology for regional vitalization), *Collection of Lectures, 24th Annual Academic Conference, The Japan Society for Science Policy and Research Management*, 442-445 (2009) (in Japanese).

[27] Ceramic Research Center of Nagasaki: http://www.pref.nagasaki.jp/yogyo/index.php (in Japanese)

[28] O. Nakamura, O. Nakamura, M. Takagi, T. Sawada, S. Kosaka, M. Koyanagi, I. Matsunaga, K. Mizuno and N. Kobayashi: Strategic evaluation of research and development in a Japan's public research institute, *New Directions for Evaluation*, 118, 25-36 (2008).

Authors

Naoto KOBAYASHI

Completed the doctoral course at the Graduate School of Engineering, Kyoto University in 1978. (Ph.D.) Joined the Electrotechnical Laboratory (ETL), Agency of Industrial Science and Technology (AIST), Ministry of International Trade and Industry (MITI) in 1978, and worked as director of Planning Division and director of Quantum Radiation Division from 1997. Director of Photonics Research Institute, AIST in 2001; vice president and director of Evaluation Department, AIST in 2003; vice president, director of Safety and Environmental Management, and director of General Affairs Headquarters in 2007. Professor of Center for Research Strategy, Waseda University since April 2009. Specialties are optical device engineering, semiconductor material engineering, quantum beam engineering, and theories of research strategy and evaluation. For this paper, built the basic concept, considered the overall strategy formation, and discussed the synthetic research evaluation.

Osamu NAKAMURA

Completed the master's course at the Graduate School of Agriculture, Kyushu University in 1979. Engaged in education and research as assistant professor of Oral Biochemistry, Kagoshima University Dental School. Doctor of Dentistry (Osaka University) in 1987. Visiting researcher, Case Western Reserve University, Cleveland, Ohio, U.S.A.; senior researcher, Kyushu National Industrial Research Institute; deputy director and manager of Bioresource Division, Biotechnology and Food Research Institute, Fukuoka Industrial Technology Center; senior researcher, Evaluation Department, AIST; and general investigator of industrial technology, Technology Evaluation and Research Division, METI. Appointed deputy director of the Evaluation Department, AIST in 2007, and was involved in the evaluation of R&D management, and built human network of evaluation in Japan and abroad. Director General of the Science and Technology Promotion Bureau, Nagasaki Prefectural Government from April 2009. Specialties are biochemistry and R&D management evaluation. For this paper, contributed to build the system of synthetic research evaluation, and wrote the section on the strategy formation and evaluation of the science and technology promotion for local revitalization of Nagasaki.

Kenta OOI

Completed the master's course at the Graduate School of Science, Nagoya University in 1976. Joined the Shikoku National Industrial Research Institute in 1977, as manager and chief of Planning Division. Principal researcher of the Institute for Marine Resources and Environment, AIST in 2001. Worked as manager of Technological Policy Survey Division, Technology Information Department; coordinator for industry-academia-government collaboration, AIST Shikoku; deputy director, AIST Shikoku; principal surveyor of Industrial Technology Division, Shikoku Bureau of Economy, Trade and Industry; and deputy director of Evaluation Diepartment, Shikoku Bureau of Economy, Trade and Industry in 2008. Specialties are inorganic chemistry, marine resource engineering, and technology evaluation. For this paper, wrote the section on the research evaluation at AIST.

Discussions with Reviewers

1 Synthetic evaluation versus analytic evaluation

Question (Motoyuki Akamatsu, Human Technology Research Institute, AIST)

When we use the term "synthetic evaluation", we think of "analytic evaluation" as its antonym. For example, when you break things down in elements as in Fig. 6, it seems you are doing analytical evaluation. What is the characteristic of synthetic evaluation when seen from the stance of synthetic versus analytic?

Answer (Naoto Kobayashi)

As you indicated, evaluation by breaking down into elements as shown in Fig. 6 can certainly be called analytic evaluation. Particularly, the elemental evaluation (progress, depth, and phase) in this paper corresponds to the analytic evaluation. For example, in the depth evaluation, the evaluation performed along novelty, originality, logical completeness, and influence is an analytical process. On the other hand, the characteristic of synthetic evaluation is to take the result of such analytical evaluation, and 1) to synthesize along the direction shown in the strategy (what did the strategy emphasize to begin with?), and 2) to form a progressive evaluation through deep discussion (here, abductive inference is important) between the evaluator and evaluated side.

2 Abduction in synthetic evaluation

Comment (Motoyuki Akamatsu)

You state that abduction is important in research strategy formation, and that abduction is necessary when selection and concentration are required in an evaluation. I see that abduction becomes most important in strategy formation. However, I don't think there has ever been a good discussion on this subject. I imagine that the readers won't understand how abduction or hypothesis forming is done, so I think they can understand better if you describe what kind of hypotheses are made for the strategy formation using a specific example. Also, you state that abductive inference is necessary in the evaluation for the depth of research in the Y-axis, but what do you mean specifically?

Answer (Naoto Kobayashi)

Thank you very much for pointing out the central tenet of this paper. The abduction or hypothesis forming needed in the strategy formation is not a factual hypothesis, but is a hypothesis of what "ought to be". The issue depends on how to form such a hypothesis. We added the description of how abduction is actually done. We also added the description of abduction using a specific example. Also, for the evaluation of the depth of research in the Y-axis, we added the explanations in places where abductive inferences are required.

3 Timing of when the evaluator should step in

Question (Kazunori Nakamura, Biomedical Research Institute, AIST)

The appraisal is important in the research evaluation based on research strategy formation, Does the evaluator need to be involved from the process of research strategy formation?

Answer (Naoto Kobayashi)

Basically, since the strategy formation and research evaluation are inextricably associated with each other, the evaluator should ideally be involved from the time of research strategy formation. However, the viewpoints of research strategy formation and evaluation are slightly different, and it is not desirable for the same entity to execute the P and C of the PDCA cycle. Therefore, I think only some members should overlap in the committees for research strategy formation and research evaluation.

4 Feedback in evaluation

Comment (Motoyuki Akamatsu)

In subchapter 4.2, you discuss the importance of feedback, and I think everyone will agree. However, if the feedback is provided offline in writing, I think you loose the opportunity for abduction based on the discussion between the research executant and the evaluator. If possible, please include a discussion on how this feedback should be done.

Also, a program is expected to run for five to seven years, and I imagine that the feedback loop is in the order of five years. In general, I don't think it is easy to reflect the result of a review from five years ago in the current program. That is because in a program of five to seven year scale, it is difficult to determine whether the objective was finally achieved immediately after the completion of the program, and the immediate feedback to the next step is difficult.

Answer (Naoto Kobayashi)

As you indicate, the evaluator and the research promoting group should not be antagonistic, but they should "walk side by side". Since the act of research is a repetition of abduction and validation, a feedback without the discussion on abduction loses its meaning.

For the feedback cycle, as you indicate, it is difficult to evaluate a program unless about three years have passed after completion. However, when the R&D, technological development, and system development are done in society, they inevitably become contiguous due to the social demand. Specifically, the FP7 framework in Europe is a seven-year (2007~2013) program. An intermediate evaluation was conducted in autumn 2010, and the appraisal of the FP8 (2013~2019), the next framework, will be started in 2011 based on this evaluation. In practice, the programs are sequential and the feedbacks are done in extremely short span of time.

5 Evaluation of the research program and evaluation of the research project

Comment (Kazunori Nakamura)

In order to apply this research evaluation method, it is necessary to apply synthetic research evaluation in the process of strategy formation of the research program. Therefore, I don't think it is readily applicable to a general research project that has not passed through this process.

Answer (Naoto Kobayashi)

A research project has, in a sense, a simple structure compared to a research program. There, the research objective, method, results, and expected outcome are contained in a small area, but has a fractal structure to the research program. For example, it is possible to apply the evaluation using the aforementioned three aspects (1) progress, (2) depth, and (3) phase, and it is also possible to apply the evaluation method that synthesizes the three. However, the execution of a research project is positioned as one element of the research program within the research strategy, and the appraisal of the research strategy can be simplified.

6 Verification of the case of research evaluation

Comment (Motoyuki Akamatsu)

Since this is a paper on research evaluation, it is desirable if you present an evaluation of an actual research. If there are reviews and discussions from the members of the evaluating committee, Evaluation Department, and the evaluated entity, the paper itself will become abductive and interesting.

Answer (Kenta Ooi)

In the *Second Medium Term Target Period Research Unit Evaluation Report* published in May 2010, the comments by the external evaluation committee members, the research unit heads, and the coordinators were analyzed, the characteristics and issues of the evaluation system in the Second Term were summarized, and the improvement points were organized. For the evaluation system, there were many who highly evaluated the current system, such as the "evaluation from the external evaluation committee members" and the "introduction of the viewpoint of outcome taking into account the exit to industry and society". On the other hand, many issues that must be improved were pointed out, for example, the "need for flexibility in handling of various types of R&D such as the bottom-up research or the long-term research", the "reduction of evaluation load", and "further utilization of the evaluation results". Being aware of these improvement issues, the evaluation from the viewpoint of outcome will be continued for the Third Term evaluation system, and efforts will be made to raise the immediacy of effect. It is not even 10 years since the research evaluation started at AIST. As you indicated, it is important to make improvements based on the verification of hypotheses, toward a better evaluation system.

For specific examples at AIST in this paper, we discuss the issues of the evaluation system currently used at AIST from the perspective of an ideal synthetic evaluation based on the research strategy. To actually apply the synthetic evaluation system, I think it is necessary to design the system as a whole including the strategic research advancement, rather than cutting out the evaluation system only. I think it is necessary to conduct the modeling and the hypothesis verification for the AIST system in a larger framework.

7 Logic model and synthetic research evaluation

Comment (Kazunori Nakamura)

The example of Nagasaki prefecture is described as a case where the logic model was applied to advance the strategic R&D, and the Research Evaluation Subcommittee and the Research Enterprise Evaluation Committee were held utilizing the logic model. I think you should clarify the basic differences between the evaluation based on the logic model in this case and the synthetic research evaluation described in this paper.

Answer (Osamu Nakamura)

It was presented as one of the examples of "R&D evaluation that seemed to have incorporated the thinking of synthetic research evaluation by trial-and-error", and so there is no basic difference.

As mentioned in this paper, the mission of the Science and Technology Promotion Bureau of Nagasaki Prefectural Government is to contribute to create energetic Nagasaki prefecture where people can live with dreams for the future by utilizing science and technology. To evaluate whether the research institutes quickly recognize the needs of the local companies and product districts, and set the research topics that can achieve the demanded results, we asked to review all the running projects and summarize them into the logic model at first. That is because the application of the logic model is effective to clarify the strategic logic of the scenario.

Each evaluation committee checked the positioning of each research topic based on the logic model, evaluated whether the

strategic R&D were conducted according to the mission of each research institute, checked whether the structure of the overall projects was strategic, and checked whether the program was a strategic program. They also evaluated whether the individual project had a long-term vision, had clear targets, and whether it was producing adequate results toward the targets. However, this attempt has just started, and so it is necessary to evolve the strategic evaluation system.

8 Specific content of the integrated evaluation

Comment (Motoyuki Akamatsu)

You write that synthetic evaluation is to "integrate" the evaluations, but I think it is hard to understand because there is no example on what exactly is integrated evaluation. Do you add the three axes, or do you evaluate by changing the weight of other axes? Please provide a specific example.

Answer (Naoto Kobayashi)

This indeed is the pillar of the evaluation design. In the case of AIST, various things were done such as weights were added or averages were taken, and this is where the evaluation organization can become creative. Considering this point, the following description was added to "4.3.3 Integrated research evaluation".

"In conducting the synthetic evaluation, it is important for the research promoting group and the evaluators, to consider the property of research, share common understanding of the strategy and result index including the goal to be achieved, engage in deep discussion on the result and how the research should progress, check the target indicated in the strategy and distance from the actual result, and finally abductively discuss and investigate the significance and the effect of the research program execution. This entire process can be considered the "research evaluation = abduction and its expression." This is closely related to the research that is a creative activity, and research evaluation can be considered to be part of this creative endeavor.

… Specifically, it is necessary to design a general evaluation system including the incorporation of the constructive, deep discussion between the research promotion group and the evaluators, in addition to the elemental evaluation and appropriate synthesis."

9 Conclusion

Comment (Motoyuki Akamatsu)

For research evaluation, you give as requirements: (1) unification under *Type 1*, *Type 2*, and *Product Realization*, (2) to be based on a contract, (3) the evaluator and evaluated to stand on the same ground, (4) value of research to be brought out, (5) research to evolve, (6) to be the source of motivation, and (7) accountability. Can you please describe in the conclusion, how these and the research strategy formation and synthetic evaluation discussed from chapter 3 are connected. I think the strategy formation satisfies (1)~(3) and (7), and (4)~(6) can be realized by synthetic evaluation. I think the paper would become organized and comprehensive if you have a summary that shows the relationships among the four intrinsic properties of the research, seven points of research evaluation, research strategy formation, and synthetic evaluation, and diagrams to explain them.

Answer (Naoto Kobayashi)

Thank you for indicating a very important point. We added Fig. 12 at the end, and also added the following description: "Figure 12 shows the summary of the intrinsic properties of the research, formation of the research strategy, and the synthetic research evaluation based on them that were discussed in this paper, and the aim of research evaluation. The research strategy shows the targets to be achieved by the R&D and the scenario. By conducting a synthetic research evaluation based on them, the extraction of the value of research, the evolution of research, the source of motivation for the researchers, and the fulfillment of accountability can be performed effectively."