Research area in Strategic Objective *"Toward scientific discoveries through DX in life science research"*

Innovation of Life Science through Digital Transformation Focused on Data-Driven and AI-Driven Technologies

Research supervisor: Yasushi Okada
(Team Leader, Center for Biosystems Dynamics Research, RIKEN)

Overview

Introducing information technology with the aim of comprehensive and innovative progress that focuses on venturing into the state of operational processes, the way organizations work, and even the culture of each field, rather than simply making operations more effective, is known as digital transformation (DX). Scientific research, especially that of the life science fields, is one of the areas DX is focusing on for innovation. In fact, since the introduction of deep learning, there has been remarkable integration between the biology field and AI research, and research in image analysis, predicted risks for disease, searches for new medicines, and estimations of protein 3D structures are being accelerated around the world.

One of the reasons that DX is particularly anticipated in life sciences is because life systems range from a single protein molecule to ecosystems; characteristically, the focus of life science research is large-scale, complex systems in which thousands of types of components or more interact with each other. While the progress of computational technologies and the automation of measuring and experiments means that we can obtain huge amounts of digital data from these complex systems at high speed, we have already reached a position in which it is no longer possible to analyze everything with the thinking capabilities of humans.

Therefore, this research area aims to bring together the "trinity" of information science, engineering, and life science, and to make advanced scientific discoveries that were not previously within our reach using "data-driven and AI-driven" research that promotes DX in life science research.

More specifically, we aim to (1) Carry out research and development that will overcome the qualitative and quantitative limits of data acquisition and analysis from varied and large data sets by making use of digital processing technologies such as AI, making new life phenomena and their models possible, and providing model cases for next generation life science research. Furthermore, we will (2) Support research and development that will establish innovative techniques for data-driven research and techniques for AI-driven research, which are necessary elemental challenges to accomplish this.

By doing the above, we aim to realize a society in which we can create a tide leading to a paradigm

shift in life science research, as well as clarify complex life systems; a society in which researchers can focus on truly creative activities.

This research area will be operated as part of the Ministry of Education, Culture, Sports, Science and Technology (MEXT)'s Artificial Intelligence/Big Data/IoT (Internet of Things)/Cybersecurity Integration Project (AIP Project : Advanced Integrated Intelligence Platform Project).

Research Supervisor's Policy for Application, Selection, and Management of the Research Area

1. The aims of this research area and specific examples of research and development

This research area aims to **provide model cases for next generation life science research** by promoting data-driven and AI-driven research based on large datasets and high-quality datasets in life science research, and **repeating a research cycle of hypothesis-verification-analysis** to come up with scientific discoveries that we may not have been able to reach with human cognitive abilities. Moreover, we will support the development of the technologies and techniques related to digital transformation (DX) that we need to accomplish this, and consider the importance of creating technologies that will **form a tide leading to a paradigm shift in life science research** so we can fulfill the abovementioned aim. However, the development of technologies and techniques for DX itself is not the objective of this research area. When it comes to projects that focus on developing technology and techniques, we anticipate research proposals that are clearly aware of their contribution to research and development, with an eye to the abovementioned scientific discoveries.
The following are specific examples of research projects, but you are not limited to these.

(1) Discovering new life phenomena and the rules behind them using varied and/or large datasets
We promote research that will obtain homogenous, real-time, multimodal data through the use of automated equipment and advanced analysis equipment, and use AI to overcome the qualitative and quantitative limits of data acquisition and analysis by humans, making it possible to discover new life phenomena and model these.

A. Research that models life phenomena that occurs by chance using large-scale repeated testing
B. Research that automatically measures the activity of cells and tissues in real-time, and, on top of this, leads to clarification of life phenomena, such as a mathematical understanding of disease, by adding perturbation, etc.
C. Research that contributes to the advancement of materials production and the creation of medicines, etc. by modeling metabolic systems within cells and exploring target molecules, to be achieved through the automation and advancement of DBTL (design-build-test-learn) cycles and DMTA (design-make-test-analysis) cycles
D. Research that develops technologies that make it possible to acquire real-time data from cells, tissues, and individuals in a non-invasive fashion through integrated analysis of imaging information and omics information, which will lead to the clarification of living phenomena, such as the structural principles of multicellular living systems

(2) Developing innovative technology for data acquisition and analysis that will contribute to data-driven research

We carry out research on foundational technologies that will contribute to the acceleration of data-driven research, by developing innovative technologies to acquire homogenous and large amounts of data, such as multimodal analysis or automating analysis equipment.

A. Developing intelligent measuring technologies that will make it possible to acquire large-scale, high-dimension data that overcomes conventional limits, to be achieved based on the premise of combining advanced information processing technologies
B. Developing technologies that will make it possible to acquire large-scale and advanced data that was previously impossible to realize, by automating experimental and/or measuring techniques required for the manual processes of skilled researchers/making these techniques capable of high throughput and/or intelligence
C. Developing new foundational technologies that cover everything from the synthesis of compounds to assaying and evaluating them, by combining microfluidic technologies with robots/AI

(3) Research on and the development of data processing and analysis, modeling, and knowledge processing techniques to establish techniques for data-driven and AI-driven research

When working towards discoveries that essentially overcome human cognitive capabilities and bias in research processes within life science, completeness of analysis and modeling based on experimental data and evaluations of their validity are important. The development of a variety of AI technologies is indispensable to these approaches, and will establish a new way of making scientific discoveries based on the vast amount of knowledge we have accumulated in life science fields to date.

A. Developing systems that extract knowledge of a given area from previous research, based on natural language processing technologies, and using this to help make research policy decisions
B. Research on active inference techniques, such as the automatic design of experiments that will generate the next hypothesis to test and the optimal experimental conditions to implement this, from previously-obtained experimental data and literature data
C. Developing systems that will integrate experimental data and the knowledge of a given area that has been accumulated in databases, and make it possible to automatically assemble, improve, and evaluate models that will lead to new scientific discoveries

2. Management policies for this research area

This research area will be managed based on the following policies; please ensure you have gained a sufficient understanding of their meaning before you apply.

(1) Team organization

In this research area, we expect the challenges of young researchers who will lead the next generation, so we call for not only the conventional CREST but also mini- CREST, which is positioned between PRESTO and CREST.

A. CREST

This research area aims to realize the fusion of life science research and AI technology, and the innovation of life science research through the fusion. To achieve this goal, we encourage you to form a team that covers the fields of informatics, engineering and life sciences and allows young researchers to actively participate in the research.

B. mini-CREST

We will also set up application slots by individuals or small groups to encourage the challenges of outstanding young researchers in each field. We also welcome research aimed at solving specific technical issues, depending on the size of the team. Even in that case, in view of the purpose of this area, we expect to promote data-driven and AI-driven technologies and propose technological developments aimed at innovation in life science research.

(2) Accelerating research outcomes and reviews based on progress

A. Please be advised that we will review research costs once each academic year, and these may increase or decrease according to the progress of your research.

B. We will consider additional measures for supplementary budgets for research projects that are making excellent progress and research projects that contribute to this research area (providing technology, data analysis, etc.).

C. You may be asked to review and/or rearrange your research organization and budgets based on your research progress and the outcomes of interim evaluations. When it comes to reviewing your research organization, we may request that you rearrange it to span multiple research teams or that you add an external researcher, rather than just reviewing the organization within your team. Notably, when it comes to mini-CREST, we expect that you will integrate with research teams according to the development of your research, and ultimately achieve great outcomes.

(3) Sharing data and technologies, and collaboration within the research area

A. We encourage participants in this research area to share scientific knowledge (data, etc.), technologies, protocols, and more within the research area. To accomplish this, if participants are actively promoting data sharing within the research area, we may take measures to supplement their budget to accelerate these activities.

B. We will promote joint research, etc. through the creation of a network within the research area to maximize research outcomes. To make this happen, we will hold discussions in research area meetings, as well as meetings and study groups with comparatively few people based on the advice of research supervisors and advisors, and ask for your active participation.

C. From the perspective of efficiently making use of research costs in the research area as a whole, we ask that you share the use of large pieces of equipment, etc. purchased using research costs as necessary with other research teams in the research area.

(4) Collaboration between research areas

We aim for collaboration with relevant CREST and PRESTO research areas so as to maximize our outcomes, not just as a research area but also as a JST Strategic Basic Research Program , and will jointly hold workshops, symposiums, etc. as necessary. In addition, initiatives for the automation of life science experiments are being carried out within the JST-Mirai Program "Common Platform Technology, Facilities and Equipment" area "Accelerating Life Sciences By Robotic Biology," and so we aim to promote effective and efficient research through technical collaboration and research exchange with this research area.

(5) International collaboration, etc.
We will collaborate with research institutions and research projects in Japan and overseas that are highly relevant to this research area, and implement collaboration and exchange regarding technology, data, and more. Through these activities, we will promote the overseas development of technology originating in Japan and the training of young human resources. Moreover, each research team will have an active point of contact overseas, and we expect that this will serve as a link between international projects and this research area.

3. Points to note regarding applications
When making a research project application to this research area, please pay attention to the following points.

(1) Eligibility
In this research area, we are focused on research that aims to clarify life phenomena; we do not mind what field or classification is covered, as long as the research falls into this category. We welcome proposals not only for basic life science research, but also for medical and industrial applications and social implementation. Moreover, the development of data acquisition technologies and analysis technologies that contribute to DX, which is necessary to achieve the abovementioned goals, is also eligible. The creation of technology that will lead to a paradigm shift in research is also a focal point of evaluations within this research area, but this should ultimately be a means to the aim of clarifying life phenomena, so please be careful that this means does not become the goal of your proposal. Moreover, research proposals with objectives that can be approached through conventional life science research or that merely acquire Big Data are not desirable.
In mini-CREST, we also welcome proposals that specialize in one-point breakthrough type technology development based on free ideas. However, in that case, please understand the purpose of this area and clarify the prospect of how you can contribute to the realization of Bio DX if your technology is realized.
We expect to gather many creative and innovative research proposals.

(2) Team organization
Please put together an appropriate team from the perspectives of research fields, expertise, technology, etc., so you can work towards realizing the content of your proposal. We especially recommend collaboration between life science, engineering, and information science fields. Please make sure to create a team that is made up of members who are necessary to and sufficient for the achievement of

your proposal content. We expect the research director to have a cross-field and interdisciplinary mind that can compensate for elements that are lacking at the time of application, so as to achieve the objective of this research area. We also expect that you have a clear vision of how your proposal will play a part in the new life science research style. We will also place emphasis on whether the research director has adequately demonstrated their leadership, and whether we can expect effective synergy within the team.

Moreover, it is necessary to train human resources for research who can lead this research area in the medium- and long-term, enabling us to strongly push DX research centered on the use of AI and Big Data. In light of this, we encourage applications from teams that have young researchers at their core, rather than just researchers with proven track records.

In the case of Mini CREST, your team does not necessarily have to cover the three fields of life science, engineering, and information science, as you will form a team of individuals or small groups needed to realize the proposal. However, we expect you to have a concrete vision of how to contribute to the goals of this research area by integrating the three fields in the future.(3) Research period and research costs

The research period should be no more than five years and six months. You should apply for the research costs (direct costs) that you will need to achieve the content of your proposal; there is an upper limit of 300 million yen. Mini-CREST has an upper limit of 150 million yen. However, please make sure that the budget plan is commensurate with the size of the team and the content of the proposal.

In addition, please be aware that research costs may be adjusted upon selection as a result of careful investigation by research supervisors. An interim evaluation will take place approximately three years after your research begins. Depending on the results of the evaluation, we may review your research costs, and/or instruct you to reorganize your research team.

4. Other

(1) From CREST's perspective of a top-down research and development program with an eye to realizing social and economic needs, we will also emphasize the impact of any outcomes and their ripple effect.

(2) We welcome creative and ambitious proposals from young researchers and female researchers.

(3) We recommend proposals that integrate life science, engineering, and information science. We also welcome proposals that integrate these with other fields.

(4) Please make specific note of your proposal's superiority with regard to past research and similar and competing technologies.

(5) As one research area of the AIP Network Lab , which organizes MEXT's Artificial Intelligence/Big Data/ IoT/Cybersecurity Integration Project (AIP Project), this research area will also contribute to the integrated management of the AIP Project, including focusing on research projects while collaborating with relevant research institutions, such as the RIKEN Center for Advanced Intelligence Project .

(6) This research area is part of an initiative to train and ensure that there is a place for data scientists in the health and medical care fields, actions required by the Drug Development Committee , etc. under the Headquarters for Healthcare Policy; as such, during team formation, we require that researchers that can promote data-driven and/or AI-driven research with a view to application to life

science in the future participate in teams as research directors or major collaborating researchers.