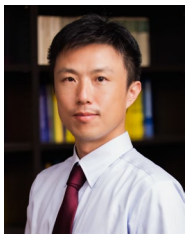
【Grant-in-Aid for Transformative Research Areas（A）】

Establishing data descriptive science and its cross-disciplinary applications（data descriptive science）



Head Investigator	Kyoto University, Kyoto University Institute for Advanced Study, Professor HIRAOKA Yasuaki　Researcher Number:10432709
Research Area Information	Number of Research Area : 22A201　Project Period (FY) : 2022-2026 Keywords : data descriptive science, mathematics, machine learning

Purpose and Background of the Research

●Outline of the Research

The modern world is inundated with data. However, as seen in the black box problem in AI technology, it cannot be said that the true value contained in such data is being fully utilized. In order to make effective use of such big data, it is significant to describe the essential structure of the data in an appropriate mathematical language, and to use that descriptive language in the process of giving meaning and of understanding mechanisms behind the phenomena. In this research area, we will solve this problem by developing descriptors (mathematical languages that express the essential structure of data) that focus on the "shape" and "movement" of data, using state-of-the-art mathematical and data science methods. To this end, our research area is organized as a trinity of mathematics, data science, and application to create a new fusion area, "data descriptive science".

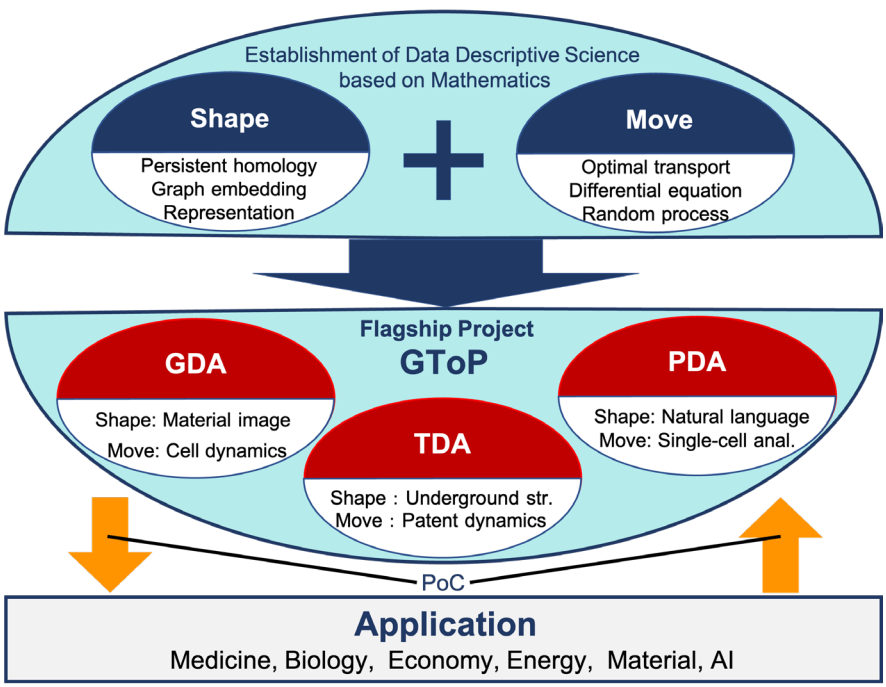


Figure 1. Activity on data descriptive science

●Organization

The Planned Research consists of three groups. Research Group A01 will conduct mathematical and data science research for building theoretical foundations of data descriptive science. Research Group A02 will focus on research exploring new areas of data descriptive science. Research Group A03 will conduct research applying our methodologies to materials science and life science.

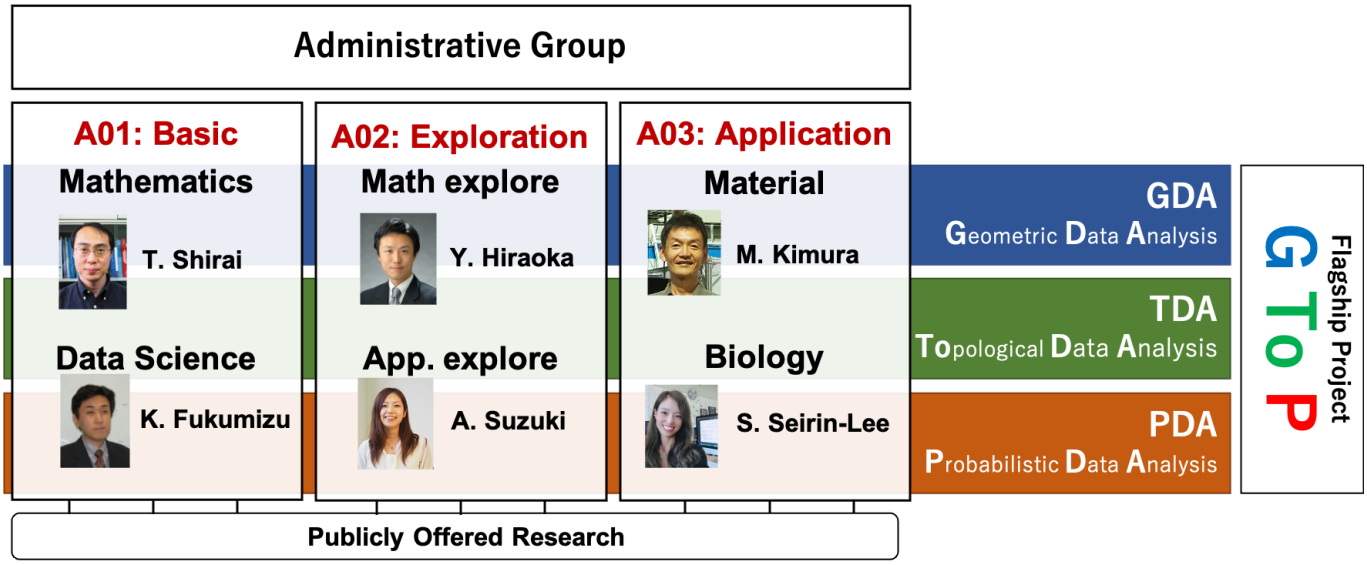


Figure 2. Overview of our research area

Expected Research Achievements

● **Flagship project "GToP"**

Towards the establishment of the data descriptive science, we set the flagship project "GToP", consisting of

- GDA (Geometric data analysis)
- TDA (Topological data analysis)
- PDA (Probabilistic data analysis)

Those schemes provide three axes in our data analysis methods, and all research projects are organized to contribute to at least one of them. The basic idea of GToP is to describe "shape" and "move" of data in geometric/topological ways (GDA, TDA), and to characterize random effects induced from the data on those geometric/topological structures (PDA). In this research area, by comprehensively integrating such diverse fields of mathematics with data science, we will develop robust and universal data descriptiors, and as a result, a wide range of applications will be possible.

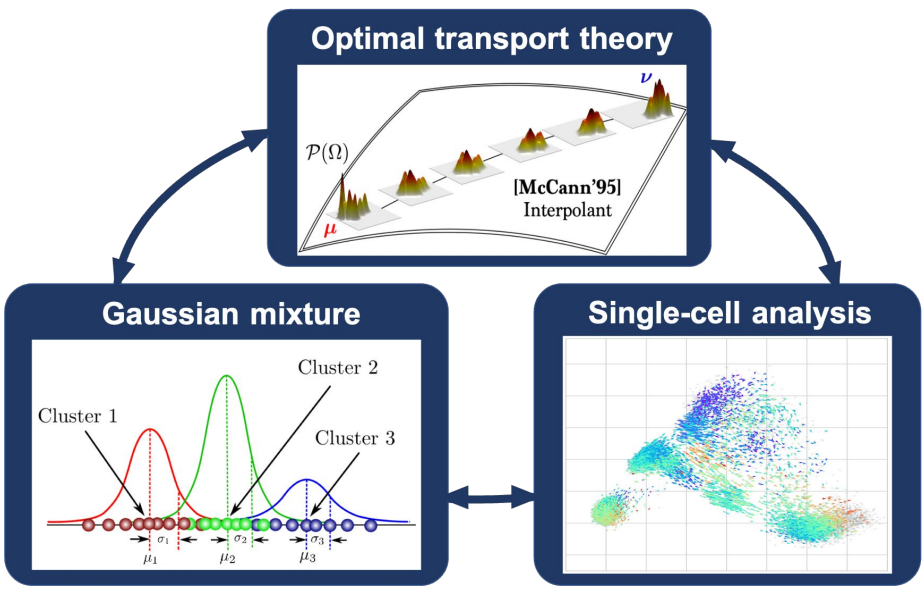


Figure 3. An example of PoC cycle

Homepage Address, etc.	https://data-descriptive-science.org Twitter : @gakuhen_a_dds