

## 分野融合型科学創出のための試行プロジェクト AMIDER

# 梅村 宜生 [1]; 田中 良昌 [2]; 阿部 修司 [3]; 中野 慎也 [4]; 新堀 淳樹 [1]  
[1] 名大・宇地研; [2] 国立極地研究所/総研大; [3] 九大・ICSWSE; [4] 統数研

### Pilot project AMIDER for data-driven science

# Norio Umemura[1]; Yoshimasa Tanaka[2]; Shuji Abe[3]; Shin'ya Nakano[4]; Atsuki Shinbori[1]  
[1] ISEE, Nagoya Univ.; [2] NIPR/SOKENDAI; [3] ICSWSE, Kyushu Univ.; [4] The Institute of Statistical Mathematics

ROIS (Research Organization of Information and Systems) has established the AMIDER project in Apr 2018 for the purpose of acceleration of data science business and establishment of its technical procedure. The AMIDER project was started by some researchers in ROIS-DS (Joint Support-Center for Data Science Research), Nagoya University, Kyushu University, National Institute of Polar Research (NIPR) and Institute of Statistical Mathematics (ISM), and is now building a technical method to suggest the possibility of scientific relationship on the basis of the statistical calculation with the several sample data (climatology data, environmental data, geology data, earth-space data, biology data and so on) managed by NIPR. The steps of the statistical calculation are formed by adding several functions of Data Conversion and Data Fusion on the Data Archiving represented by IUGONET and DIAS so far.

Data Conversion converts research data from research field-dependent format to general ASCII format, so that more users can analyze the data and read the data for computer easily. Furthermore, Data Fusion reads general-format data converted by Data Conversion, and calculates and determines the relationship between these data. For example, we may be able to derive the truth of how interaction in the adjacent or entire space several sun-earth phenomena (upper atmospheric cooling due to greenhouse effect, ionospheric variation associated with earthquake and volcano activities) are generated through, how these natural phenomena affect the ecology of plants and animals, and so on. These data-driven-specific functions are expected to not only support advanced discovery across multiple fields and elucidation of their relationships in research scene, but also suggest the possibility of the occurrence of the preceding chained events in social scene.

As of July 2019, the AMIDER project achieved the step of Data Conversion (CDF and NetCDF to be tested and its provision). Then, in Nov. 2019, we have a plan to achieve the step of Data Fusion, and to provide data owners with a list of correlation analysis results between submitted data and other registered data.

情報・システム研究機構は、データ駆動型科学事業の加速とその手法の確立を目的として、2018年4月、同機構のデータサイエンス共同基盤施設、名古屋大学、九州大学、国立極地研究所、統計数理研究所の研究者らによる試行プロジェクト AMIDER を発足した。AMIDER は、機構下の国立極地研究所が所有する極域の観測および採取データ（気候・環境、地質、宇宙、生物データなど）をサンプルとして、統計手法から導かれるデータ間の数値的關係から科学的關係の可能性を示唆する仕組みの構築を進めている。そのプロセスは、これまでの IUGONET や DIAS に代表されるデータアーカイブ機能に、データ変換とデータ融合の機能を付け加えて形成される。

データ解析支援は、バイナリ形式など研究の都合で採用されることの多い特異なフォーマットを普遍的なアスキー形式に変換する仕組みであり、専門外の研究者にも解析機会を与える効果があるほか、高度計算などの機械処理への入力を実現するものでもある。データ融合は、機械可読化されたフォーマットのデータを突き合わせて、その間の關係性を導出し、類似性を判定する仕組みである。例えば、太陽地球系科学分野では、地球温暖化による超高層大氣の寒冷化や地震・火山活動による電離圈変動といった現象が隣接または全空間にてどのように相互作用を通じて発生しているか、分野を拡大すると、地球物理現象の変動が、植物や動物の生態にどのような影響を与えているか等の広義的な關係を見出す手がかりとなる可能性がある。これらのデータ駆動科学特有のステップは、特定の分野に留まった科学的探究に留まらず、複数の分野を跨いだ高度な発見や、その因果關係解明のための支援、また、社会の場面では、連鎖關係にある先の事象の発生の可能性を示唆する仕組みとして期待される。

AMIDER は、2019年7月時点ではデータ解析支援（テスト対象である CDF と NetCDF の変換とその提供）を達成した。同年11月にデータ融合を達成する予定であり、同時期より、データ所有者に対して、提出されたデータと他の登録データの關係性の一覧を提供する予定である。