2023 年 11 月 6 日

報道関係各位

国立大学法人大阪大学
日本電気株式会社

# 大阪大学サイバーメディアセンターと NEC、国際会議・展示会 SC23 にて オープンサイエンスを加速するデータ基盤 RED-ONION のデモ展示

大阪大学サイバーメディアセンターと日本電気株式会社(以下、NEC)は、オープンサイエンスを支えるデータ基盤の実現に向けた取り組みを進めています。大阪大学サイバーメディアセンターと NEC が 2021 年に大阪大学サイバーメディアセンターに設立した高性能計算・データ分析融合基盤協働研究所（所長：伊達 進教授、注 1）は、オープンサイエンスを支えるデータ基盤 RED-ONION を構成する(1)データ集約基盤-データ公開基盤連携、(2)データ基盤-スーパーコンピュータ連携、(3)高速データ転送機能の三つの要素技術を研究開発しています。この研究成果を、2023 年 11 月 12 日から米国コロラド州デンバーで開催されるハイパフォーマンスコンピューティングに関する国際会議・展示会 SC23 にて展示します。

＜背景＞

近年、大学などの研究機関には研究データを社会全体で共有するオープンサイエンスの推進が強く求められています。2023 年 5 月に仙台で開催された G7 仙台科学技術大臣会合においても、G7 各国および EU の共通認識として「G7 科学技術大臣の共同声明」が出され、その中でオープンサイエンスの推進が挙げられています。このためには、日々の学術研究で産み出される研究データを様々な研究機関と共有可能とし、利活用可能な形で管理し、また社会に広く公開するデータ基盤の構築が不可欠となっています。

このような背景から、大阪大学サイバーメディアセンターでは、学内外の研究者が保有する研究データを蓄積・共有可能とする、データ集約基盤 ONION (Osaka university Next-generation Infrastructure for Open research and innovatioN)の試験運用を 2021 年 5 月から行ってきました。高性能計算・データ分析融合基盤協働研究所では、この試験運用中のデータ集約基盤 ONION をさらに高度化・発展させ、オープンサイエンスを加速するデータ基盤 RED-ONION(Research EnhanceD ONION)構想を掲げ、研究開発に取り組んできました。RED-ONION 構想では、学内の研究開発を支えるデータ集約基盤のあるべき形として、研究データを蓄積し研究者間で共有可能な機能を提供することに加えて、信頼性のある研究データおよび研究プロセスの適切な管理・利活用

を促進するデータ基盤の実現を目標としています。

今回のデモ展示では、この RED-ONION 構想の下で取り組んだ研究開発の成果を報告します。

＜研究成果＞

今回、高性能計算・データ分析融合基盤協働研究所において、RED-ONION を構成する、以下の三要素技術の研究開発成果を創出しました。

（1） データ集約基盤-データ公開基盤連携

データ集約基盤に蓄積された未公開の研究データの公開を支援するシステム技術です。研究データを学術成果として社会に還元し利活用を促進することの重要性が増していますが、研究者への負担が大きいという問題がありました。本技術は、データ集約基盤 ONION において、研究者が慣れ親しんでいる研究者間でのデータ共有と同じユーザインタフェースで、大阪大学附属図書館の管理・運用する機関リポジトリ OUKA(注２)で公開可能とすることを特徴とします。文部科学省「AI 等の活用を推進する研究データエコシステム構築事業」の支援をうけ、本技術のプロトタイプシステムを大阪大学附属図書館と共同で開発、学内での実証実験を 2023 年 7 月に実施し、研究者の負担を大幅に軽減できることを確認しました。本研究開発成果は、論文誌学術情報処理研究に採択され、2023 年 11 月に出版される予定です（注 3）。

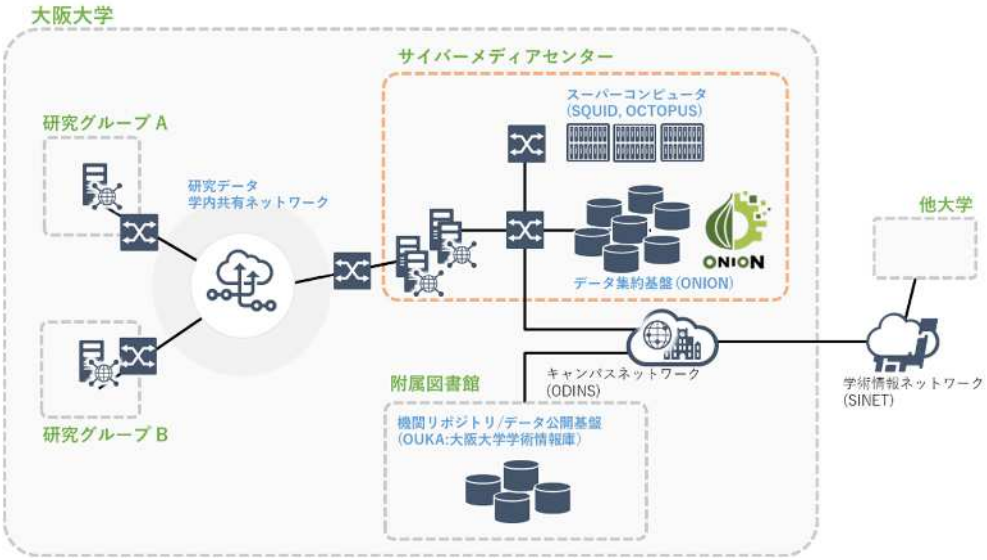
（2） データ基盤-スーパーコンピュータ連携

研究データとあわせて研究プロセスを管理可能とするデータ基盤-スーパーコンピュータ連携技術を開発しました。研究データと紐づいた形でその研究データを生み出した過程を示すことが研究成果の利活用促進において重要な役割を果たします。しかし、それには研究者による記録を必要とし作業負担が大きく、また正確に行うことが困難でした。本技術は、スーパーコンピュータを用いて生み出される研究データについて、それらがどのような研究プロセス、即ちスーパーコンピュータにおけるどの計算処理を経て生成されたデータかを示す来歴情報をシステムが自動的に記録することを特徴とします。これにより、研究者の負担なく正確な記録が可能となりました。本研究開発成果は、2023 年 10 月にキプロスで開催された国際会議 IEEE eScience’23 にて発表しました（注 4）。

（3） 高速データ転送機能

大量の研究データを研究グループ間で高速にデータ共有可能とする

高速データ転送システムのプロトタイプを開発しました。この高速データ転送システムは、学内に研究データ専用の高速な共有ネットワークを構築することを想定したものです。大量の研究データをデータ集約基盤に保存可能とし、またその研究データを複数の研究グループで利活用可能とすることを目的としています。また将来的には海外も含む学外とのデータ共有への適用も検討しています。




大阪大学のオープンサイエンスを支えるシステムの概略

＜SC23 での展示＞

SC23 では、上記三つの研究成果について、(1)データ集約基盤-データ公開基盤連携と(2)データ基盤-スーパーコンピュータ連携を、大阪大学サイバーメディアセンターが出展する研究展示ブースで展示します。また、(3)高速データ転送機能を国立研究開発法人情報通信研究機構（以下、NICT）が出展する研究展示ブースで展示します。(3)については、本プロトタイプを用いた国際間のデータ共有を想定した日米間のデータ転送実験を NICT が提供する日米間の 100Gbps のネットワーク回線を用いて実施する予定です。

以上

注 1: 高性能計算・データ分析融合基盤協働研究所
(https://www.nri.cmc.osaka-u.ac.jp/ja/)

注 2: 大阪大学学術情報庫 OUKA
（https://ir.library.osaka-u.ac.jp/repo/ouka/all/)

注 3: 田主 英之, 山下 晃弘, 細見 岳生, 並木 悠太, 甲斐 尚, 松浦 かんな, 伊達 進, “研究データ管理を支える学内情報基盤連携の実現に向けて”, 学術情報処理研究（採択)。

注 4: Yuta Namiki, Takeo Hosomi, Hideyuki Tanushi, Akihiro Yamashita and Susumu Date,

"A Method for Constructing Research Data Provenance in High-Performance Computing Systems", 2023 IEEE 19th International Conference on e-Science (e-Science), Limassol, Cyprus, 2023, pp. 1-10, doi:10.1109/e-Science58273.2023.10254932.