

Information Technology Center, The University of Tokyo

東京大学情報基盤センター

Toward a Trustworthy Information Society



〒277-0882 千葉県柏市柏の葉 6-2-3
6-2-3 Kashiwanoha, Kashiwa-shi, Chiba 277-0882, Japan
Phone: 04-7133-4658 / Fax: 04-7135-5532 (G3)
<https://www.itc.u-tokyo.ac.jp/> (日本語)
<https://www.itc.u-tokyo.ac.jp/en/> (English)

2022年12月発行



信頼できる情報社会

A Trustworthy Information Society

センサーによって得られたデータを解析して工場、医療現場などの安全性・生産性を向上する、観測データをシミュレーションと連携させて農漁業の収穫量を上げるなど、データ活用型社会への期待が高まっています。一方で、個人情報を含むデータがごく少数の営利企業に独占され、暗黙の偏見や差別につながりかねない利用のされかたをすること、また権力機関によって不当な検閲に利用されることへの懸念・疑念なども広がっています。そのような時代において、大学に職を持つ我々がその公共的使命を強く自覚しつつ、データ処理やシミュレーションのための透明な情報基盤を自ら構築し、信頼に足るデータ活用型社会の実現に中心的な役割を果たす高い技術や倫理感を身につけた人材を育成し続けることは、これまでになく重要になっていると考えます。

情報基盤センターはその根幹を担う機関であり、4つの研究部門「情報メディア教育研究部門」「データ科学研究部門」「ネットワーク研究部門」「スーパーコンピューティング研究部門」が連携して研究をしています。そして、全学の研究・教育に欠かせない情報基盤を設計・導入し、安定運用することを重要な使命としています。また、全国8大学のセンターを連携させたネットワーク型拠点「学際大規模情報基盤共同利用・共同研究拠点(JHPCN)」の中核拠点として機能しています。

国内有数の先進性と規模を持つコンピュータシステムを構築し、本学および全国の研究者に最先端の情報基盤を提供する。さらに、最先端の基盤であるがゆえに得られる知見や技術的課題を次世代の情報基盤の創出へと活かしていく。こうして継続的に、相乗的な効果を上げていく取り組み自体も、本センターにとつての重要な研究テーマとなります。

最先端の基盤を安定運用し続けることは大きなチャレンジです。そして、このチャレンジを続けることこそが、技術や人を育てるのだと確信しています。

Expectations for data utilization in society are rising, with applications such as improving security and productivity in factories and hospitals using data from sensors, increasing harvests in agriculture and fishery using simulations that assimilate data, and more. However, concerns and suspicions also abound. Among them is the fear that important data, including those which are private, are increasingly monopolized by a handful of for-profit corporations and used in a way that might lead to implicit biases and discrimination. There is also the fear that data may be used by authorities for unjust censorship. It has never been more important for us academics to recognize our public roles and transparently manage our own IT infrastructures for data processing and simulations. In doing so, we thereby take on the role of educational experts who are central to developing a trustworthy data-utilizing society with our strong technical backgrounds and high ethical standards.

The Information Technology Center (ITC) at the University of Tokyo is an institution that acts as the foundation for this work. The ITC consists of four cooperating divisions actively engaged in fundamental and practical research: the Campus-wide Computing Research Division, Data Science Research Division, Network Research Division, and Supercomputing Research Division. The ITC's mission is to design, procure, and stably operate the university- and nation-wide infrastructures critical for research and education. In addition, it acts as the core institute of the Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructure (JHPCN), a network of eight academic centers that contribute massive computing resources extensively to Japan and the world.

We have been building cutting-edge computer systems that are among the largest in the world and making them available to researchers both inside and outside the university. We have been taking the problems we face and insights we draw along the way and carrying them on to the next development cycle. Such activities are possible only because we make an effort to stay on the cutting-edge of technology. Continuing this synergy between development and feedback from service operations is our important theme.

Operating large-scale, cutting-edge infrastructure is always a challenge, and we firmly believe it is our continued commitment to this challenge that allows us to develop future technologies and experts.

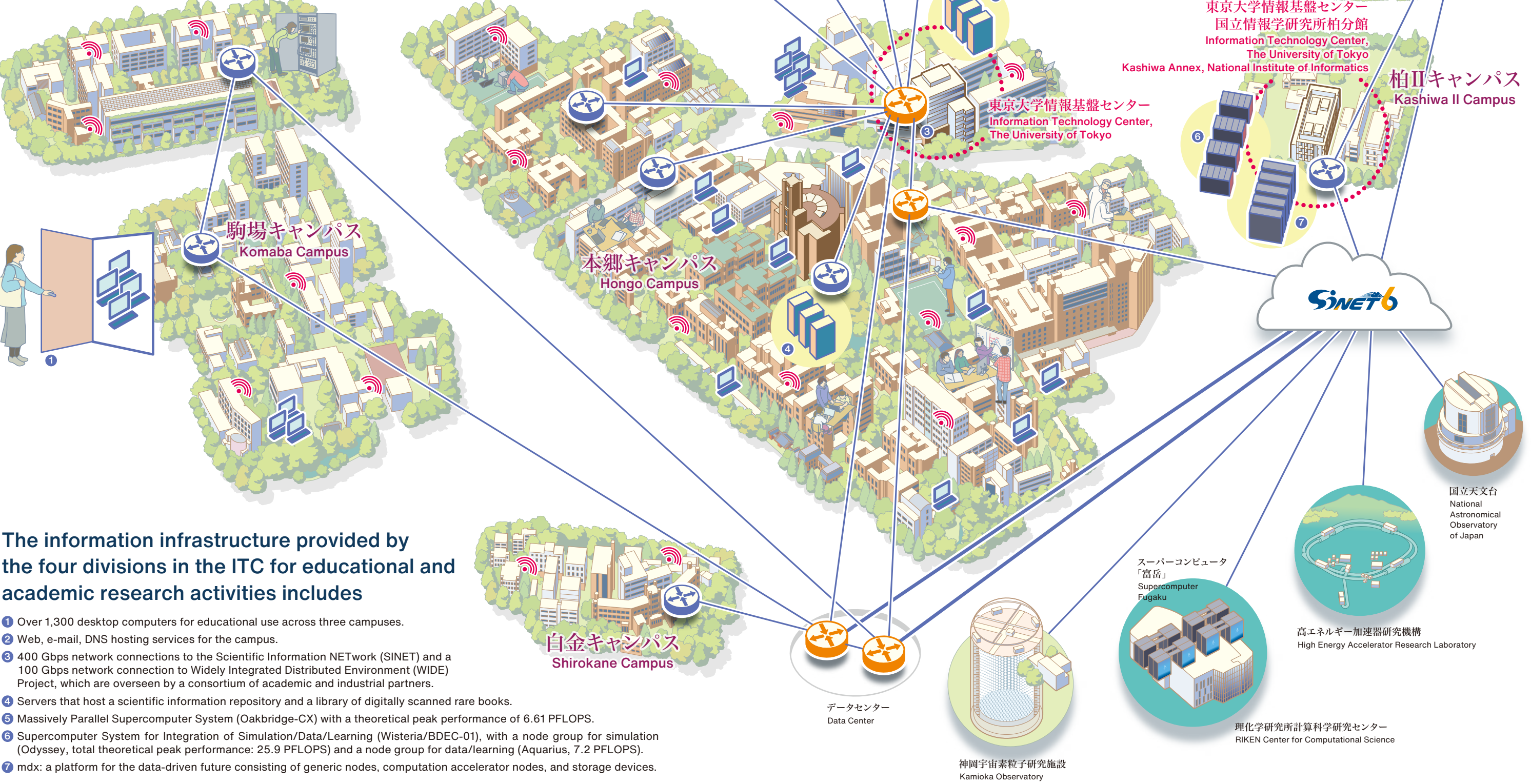


東京大学 情報基盤センター長 田浦健次郎

Professor Kenjiro Taura
Director of Information Technology Center
The University of Tokyo

教育と学内外の研究活動を支える情報基盤を4つの部門が提供しています。

- ① キャンパスに分散配置された1,300台を超える教育用コンピュータ。
- ② 学内向けのウェブ、メール、DNSのホスティングサービスを提供。
- ③ キャンパスネットワークは、学術情報ネットワーク「SINET」と400Gbpsの回線で、産学連携の研究コンソーシアム「WIDEプロジェクト」のネットワークと100Gbpsの回線で接続し、学外と通じている。
- ④ 学術情報リポジトリや電子化された貴重書を公開するサーバー群。
- ⑤ 総理論演算性能6.61ペタフロップスの大規模超並列スーパーコンピュータシステム (Oakbridge-CX)。
- ⑥ 「計算・データ・学習」融合スーパーコンピュータシステム (Wisteria/BDEC-01)。シミュレーションノード群 Odyssey (総理論演算性能25.9ペタフロップス)とデータ・学習ノード群Aquarius (同7.2ペタフロップス)からなる。
- ⑦ 汎用ノード、演算加速ノード、ストレージ群で構成された「データ活用社会創成プラットフォーム mdx」。



The information infrastructure provided by the four divisions in the ITC for educational and academic research activities includes

- ① Over 1,300 desktop computers for educational use across three campuses.
- ② Web, e-mail, DNS hosting services for the campus.
- ③ 400 Gbps network connections to the Scientific Information NETwork (SINET) and a 100 Gbps network connection to Widely Integrated Distributed Environment (WIDE) Project, which are overseen by a consortium of academic and industrial partners.
- ④ Servers that host a scientific information repository and a library of digitally scanned rare books.
- ⑤ Massively Parallel Supercomputer System (Oakbridge-CX) with a theoretical peak performance of 6.61 PFLOPS.
- ⑥ Supercomputer System for Integration of Simulation/Data/Learning (Wisteria/BDEC-01), with a node group for simulation (Odyssey, total theoretical peak performance: 25.9 PFLOPS) and a node group for data/learning (Aquarius, 7.2 PFLOPS).
- ⑦ mdx: a platform for the data-driven future consisting of generic nodes, computation accelerator nodes, and storage devices.

人の活動にどのように情報基盤を活かすか？

情報メディア教育研究部門

3万人が利用する大規模な教育用システムを企画・構築するとともに、情報基盤をとりまく多彩なテーマを研究しています。

Providing Information Infrastructure for the Students and Faculty Campuswide Computing Research Division

The Campuswide Computing Research Division designs and builds the educational computing infrastructure throughout the university campus. The division also conducts research on important information infrastructure topics.



【写真：駒場キャンパスの演習室】

本郷・駒場・柏キャンパスには、教育用計算機システムとして約1,300台のPCを配置。プログラミングなどの講義に使われているほか、日常的な道具としてワードプロセッサや表計算などにも使われている。また講義に合わせて、数式処理や統計処理、設計用のCADなどの専門的なソフトウェアパッケージが用意されている。

Photo: Educational Campuswide Computing System at Komaba campus
Over 1,300 personal computers are deployed across the Hongo, Komaba, and Kashiwa campuses. These computers are used for programming classes and for word processing and spreadsheet applications. In addition, mathematical and statistical analysis, computer-aided design, and other specialized software packages are also available.

人が力を発揮できる情報基盤をつくる

コンピュータの力を最大限に引き出して自らの活動に活かす——東京大学はこのための教育に力をいれ、教育に使用する情報基盤を1970年代前半から構築してきました。本郷・駒場・柏キャンパスには現在、1,300台を超えるPCからなる教育用計算機システム(ECCS: Educational Campuswide Computing System)が分散して配置され、登録されているユーザーはどこでも利用することができます。教養教育や専門教育で実施されるプログラミングの講義や演習に、数値計算やCADなどの専門的で多彩なアプリケーションを使った実習に、そして学生の自習の場として、ECCSは使われています。加えて、学内の情報発信や情報交換の基盤として、ウェブ、メール、DNSのホスティングサービスを提供しています。

広範にわたる研究分野と多彩な国籍の3万人の学生と教職員、そして地理的に離れている複数のキャンパス。その多様な利用者に、変化の激しいIT社会のなかでどのようなシステムをつくり、いかに効率よく安定して運用していくか。それを検討し、蓄えた知恵をオープンにしていくことは、私たちの活動テーマのひとつです。

その一方で、人に寄り添い、コンピュータの性能を引き出せるように、情報基盤をめぐる多彩な研究を進めています。

研究テーマ

・プログラミング言語処理系

ソフトウェアの信頼性、安全性、性能を高めるプログラミング言語処理系の技術を研究しています。プログラムの間違いを自動的に検出する手法は、その後、通信プロトコルの実装の正当性を検証する方法にも発展しました。

・ゲームプログラミング

ルールがはっきりしていて評価しやすいゲームは、情報処理の題材としてよく取り上げられ、アルゴリズム、探索、機械学習など、多方面の技術を発展させてきました。2013年、共同研究してきたGPS将棋がトッププロ棋士に挑戦し、好成績を収めました。

・システムソフトウェア

BitVisor——コンピュータのハードウェアとオペレーティングシステムの間を介するハイパーバイザー（仮想マシンを制御するソフトウェア）を研究・開発しています。システムのセキュリティを高めるセキュアVM（仮想マシン）のほか、クラウド構築の要素技術の研究も始動しました。

・教育支援システム

学習管理システムをはじめ、授業を支援するシステムを研究・開発しています。変わったところでは、学科のシラバスから機械学習の手法で教育の内容を抽出し、学科の教育の特徴を可視化する研究もしています。

・エッジコンピューティングにおけるネットワーク技術の研究

端末により近い場所でサービスを提供する仕組みとしてエッジコンピューティング技術が注目されています。本部門では、エッジの計算機資源に効率よくトラフィックを誘導する仕組みなどを研究しています。

Information Infrastructures that Empower Users

To enable everyone to utilize the power of computers to the maximum of their abilities, the University of Tokyo provides rich computer education and a large-scale information processing infrastructure on campus since the early 1970s. Currently, the Educational Campuswide Computing System (ECCS) comprises more than 1,300 personal computers distributed across the Hongo, Komaba, and Kashiwa campuses, which are also accessible to registered users from anywhere. The computers comprising the ECCS are used not only during programming and computer science classes, but also for computer-aided design and numerical processing applications, and by students studying on their own. The ECCS also provides web, e-mail, and DNS hosting services for publishing and exchanging academic information via the campus intranet.

With the ECCS providing computing power for a wide range of academic projects and a student body and faculty of over 30,000 individuals covering five campuses, maintaining an efficient, solid platform in a dynamically changing information technology environment provides various unique challenges. Therefore, one of the ongoing missions of the division is to review and assess solutions to these challenges and then share our findings with the general public. We also conduct research into various aspects of information infrastructure with the goal of providing enhanced computing performance with a more user-friendly interface.

Research Subjects

・Programming Languages

This division conducts research on how to design and implement software programming languages and on programming environments to improve the reliability, security and availability of programmed software. Developing testing and verification methods to confirm the completeness and correctness of secure networking communication protocols, is also another important area of our research.

・Game Programming

Games that have a concrete set of rules, such as chess and shogi (Japanese chess), are ideal study topics for data processing. By trying to solve these games, we have contributed to improvements in algorithm design, search logic, and machine learning. In 2013, through a joint research project into computer shogi software, we produced GPS Shogi, which subsequently beat a human master-level shogi player in a tournament setting.

・Systems Software

BitVisor is a hypervisor (computer software that creates and controls virtual machines between a computer's operating system and its hardware) that has been developed under the system software research group. In addition to Secure VM, a virtual machine that provides system security, additional research on fundamental VM technologies in building a cloud platform environment has recently been initiated.

・Education Support Systems

The division is currently developing learning management systems and various other education support systems. Our unique developments include systems that use machine learning methodologies to automatically extract the abstracts of what students can expect to learn in a specific division or major from the list of individual course syllabuses provided for that division.

・Network Technologies for Edge Computing

This division conducts research on networking technologies for Edge Computing. Edge Computing aims to provide multiple services on the edge of mobile networks to improve user experiences.

データに潜在している価値を取り出す

今日、天候や地震活動の自然観測データ、人や車の移動データ、商取引、医療などの社会活動データを含む、さまざまなデータが自然および人工の情報源から集められ、膨大な量のデジタルデータとして蓄積されています。

デジタルデータはまた、散逸の危険性がある歴史的な文書および記録をデジタルアーカイブすることによっても生成されます。そこではデジタル化そのものが貴重な財産を生み出し、物理的な距離に関係なくそれらにアクセス可能にすることによって、それらの価値を高めています。さらに、Webページ、ソーシャルネットワーク、学術論文など、初めからデジタルで生まれたデータもあります。

データ分析とモデリング技術、特に機械学習の発展により、データから、より意味があり解釈可能な情報を抽出することが可能になり、さらにネットワーク技術の発展によってさまざまな情報源からの情報を組み合わせることが可能になっています。データ科学は、数字の並びにすぎない生データから価値ある洞察や知識を抽出することについての広大な分野です。データ科学はまた、高性能プロセッサ、ストレージ、ネットワーク、大規模データ分析、深層学習を含む数値計算アルゴリズムなど、高性能計算技術の進歩とも深く関わっています。

データ科学研究部門は、旧学術情報研究部門を改め2018年末に設立されました。データ科学に関する研究に加え、データ科学研究コミュニティのための基盤を設計、構築するために、中心的な役割を果たしていきます。また、附属図書館と密接に協力し、デジタルアーカイブプロジェクトと学術データベースの運用サービスを遂行します。

データを価値に、財産に

データ科学研究部門

歴史的史料のデジタル化、データから価値ある知見を生み出す方法論を研究しています。

From Data to Values and Human Heritage

Data Science Research Divisionon

We research digital archives of historical materials and methodologies to identify valuable insights from data.

Mining for Value Concealed within Data

Nowadays, a huge volume of digital data is gathered from both natural and artificial sources; for example, weather and seismic monitoring data, human and vehicle mobility data, and social activity data such as business transactions and medical care records. Digitally archiving historic documents and records at risk of dissipation also produces digital data. In these cases, digitization secures and enhances the value of knowledge by making it accessible regardless of physical distance. New digital data is also frequently being created, such as web pages, social networks, and academic papers.

Advances in data analysis and modeling techniques, most notably machine learning, enable us to extract more meaningful and interpretable information from data and networking technology. This makes it possible to combine information from various sources. Data science is about turning raw data from a stream of digits into valuable insights and knowledge. Data science is also closely related to advances in high performance computing technologies, including high performance processors, storage and networking, big data analytics, deep learning numerical algorithms, and so on.

The Data Science Research Division was established at the end of 2018, replacing the Academic Information Science Research Division. In addition to research on data science, the division will lead the designing and building of a national infrastructure for the data science research community. It will also continue to collaborate closely with the General Library on the digital archiving project and scholarly database services.

研究テーマ

・データ統合と機械学習による人の流れの予測

携帯機器の位置情報から得られる人の位置と、デジタル化された都市交通網データなどを統合、分析し、新世代の人工知能技術(深層学習、強化学習、アンサンブル学習など)と結び付けて、複数の交通手段を使用して刻々と移動していく人の流れを予測する研究をしています。予測情報は、交通システムの制御、緊急事態発生時の管理、災害発生時の支援、伝染病の拡大予防対策、医療資源配置の最適化などに役立てることができます。このような予測を実現するための、モデリングとシミュレーションの手法に力を入れています。

・大規模グラフニューラルネットワーク

デジタル空間と現実世界の「モノ・コト・ヒト」のつながりは、グラフ理論では、ノードとエッジによって構成される大規模な動的グラフ構造として表現できます。この大規模グラフ構造とノード、エッジの役割・性質を、ニューラルネットワークを用いて深層学習し、解析するための手法「グラフニューラルネットワーク(GNN)」を研究しています。GNNの応用研究では、電子商取引、ニュース記事の推薦問題、モビリティにおける交通予測、金融領域における不正検知に取り組んでいます。また、データ科学・機械学習と材料開発・研究の融合分野であるマテリアルズインフォマティクスへの応用にも取り組んでいます。シミュレーションによる計算データや実験機器からの実データの効率的・効果的な収集方法、それらのデータを活用したGNNによる物性値予測の研究を行っています。

〈データ活用社会創成プラットフォームmdx〉

研究環境を用途に合わせてオンデマンドで短時間に構築・拡張・融合できる、データ収集・集積・解析のためのプラットフォームを提供しています。

データ利活用やその社会実装には、複数の分野、複数のセクターでの協働が不可欠です。大学に限らず、企業や研究機関などが保有するデータに対し、課題解決のための個々の分野の専門的知見、プログラミング、アルゴリズム、機械学習などの情報科学・データ科学の知見などが、これまで以上に広範に、密に合わさる必要になっています。mdxはそのための大きな一歩となるよう作られたプラットフォームです。大学や国立研究機関、産業界、自治体などが組織の垣根を越え、迅速に、密に、効率よく連携できるようにすることをミッションとしています。

mdxは、仮想化技術を用いてプロジェクトごとに分離された、プライベート環境(仮想プラットフォーム)を提供します。プライベート環境はプロジェクトごとに柔軟に環境構築・設定が可能で、プロジェクトそれぞれに必要なソフトウェアスタックを導入できます。そしてSINETを活かしてリアルタイムに収集・集積したデータを、mdxで解析することを可能にします。このようなプラットフォームがさまざまな分野のデータ保持者、解析者、利用者のコミュニティを形成し、新たな価値創造につながっていきます。

2022年現在、mdxは国内の9大学・2研究機関によって共同で運営され、東京大学情報基盤センターが統括・事務局を担当しています。

Research Subjects

・Machine Learning on Big Human Mobility Data

Our research integrates and analyzes people's locations, obtained from mobile phone location data, with digitized urban transportation network data. We combine these data with next-generation AI technology (deep learning, reinforcement learning, ensemble learning, etc.) to predict the flow of people moment by moment using numerous means of transportation. The predicted information can be useful for various purposes such as transportation system control, emergency management, aid in the event of a disaster, strategies for preventing the spread of infectious diseases, and allocating medical resources. We are particularly focused on modeling and simulation techniques that will enable this sort of prediction.

・Large-Scale Graph Neural Networks

All entities in digital space and the real world—including objects, facts, and human beings—and their relationships can be represented as nodes and edges, leading to large-scale dynamic graphs in graph theory. We work on graph neural networks or GNNs that can learn graph structures and the roles of nodes and edges via deep learning. We target various GNN applications including recommender systems in e-commerce, news platforms, transportation as well as fraud detection in financial systems. We have also been investigating Materials Informatics, the interdisciplinary research field between data science/machine learning and materials science/engineering. We are working on developing an effective method to store big materials data, such as the theoretical results of physical simulations and real data from experimental instruments. Moreover, we are using these data to develop machine learning methods for material property predictions which use GNNs to effectively handle molecular graphs.

〈mdx: a platform for the data-driven future〉

mdx aims to be a platform that provides the functionality to collect, store, and analyze data that can then be rapidly created, extended, and integrated on demand for specific uses.

Collaboration between multiple fields and sectors is required to be able to utilize data and apply it to benefit society at large. There is a need to bring together organizations that possess and provide data, particularly corporations and research institutions in specific fields; experts in various fields needed to resolve issues; and specialists in information and data sciences, including in programming, algorithms and machine learning. As a major step toward that end, mdx aims to facilitate collaboration between universities, national research institutions, industry, and the government.

Using virtualization technology, mdx provides private environments (virtual platforms) that are unique to each project. A private virtual environment can be created and configured individually for each project in a flexible manner, making it possible to install the necessary software stack for the project. mdx utilizes the Japanese academic backbone network called the Science Information NETwork (SINET) and generates output from input in real time by connecting remote sensors and storage devices with the computational resources on the data platform. A community of data owners, analysts, and users in a variety of fields who collaborate across the boundaries of academia and industry is formed to create new values.

As of FY2022, mdx is jointly operated by nine universities and two research institutes, led by the Information Technology Center at the University of Tokyo.

ネットワークとコンピューティングの融合による
情報基盤の高度化

日本のインターネットは、1986年、東京大学の情報基盤センター（当時大型計算機センター）、慶應義塾大学、東京工業大学の間を実験ネットワークでつなぐことから始まりました。今では、インターネットは日々の生活に欠かせないものとなり、その安定性、信頼性に対する要求もたいへん大きくなっています。また、ネットワークとコンピューティングの融合が進み、情報基盤の高度化にはこの2つを切り離して考えることはできなくなりました。さまざまな脅威が増えるなか、セキュリティも重要な課題です。私たちは、ますます重要度を増すネットワークとコンピューティングについて、基礎から応用まで幅広い研究を進めています。研究は、学内ネットワークシステムUTNETの設計・運用・構築と一体になっており、研究の成果を運用に活かすとともに、運用で得られた知見を研究につなげています。

社会の様相と密にリンクしている情報大動脈
ネットワーク研究部門

情報ネットワークの基礎技術から応用までを研究するとともに、5つのキャンパスと複数の研究施設を結ぶ学内ネットワークシステムを構築しています。

The critical Link of Information between the University and Society
Network Research Division

The Network Research Division conducts research on the fundamental aspects and advanced applications of network technologies. The division also manages the university's network infrastructure which connects several research facilities and five campuses.

Integrating Networking and Computing for
Advanced IT Infrastructure

The Internet was introduced to Japan in 1986 as an experimental network connecting three universities—the University of Tokyo Information Technology Center (previously known as the University of Tokyo Computer Centre), Keio University, and Tokyo Institute of Technology. Since then, the Internet has been an essential part of our daily lives, and demand for its availability and reliability has been growing. Moreover, advancement in information infrastructures requires further integration of networking and computing. Security has also become a growing concern as cyber threats become more varied.

The Network Research Division conducts research on networking and computing by examining both fundamental and advanced research topics. Our mission involves designing, constructing, and operating the university's campus network system, UUNET. We apply our knowledge of network operation to our research while also applying the results of our research to enhance network operation further.

研究テーマ

- ・インターネットトラスト技術
信頼できるネットワークとコンピューティングを実現するため、数理論理に基づく解析的なインターネットセキュリティモデルの構築、トラストフレームワーク、それに関連する暗号技術に基づいた方法論を研究しています。また、機械学習の数理的な評価に関する研究も行っています。
- ・ネットワークの高速化・高度化
汎用オペレーティングシステムにおけるネットワーク処理の高速化や、バックボーンネットワークの柔軟な制御手法について研究しています。また、リアルタイムIoTを支える情報基盤のアーキテクチャに関する検討も行っています。
- ・日常生活の体験を拡張するユーザーインタフェース
新しい体験の創造を目指し、AR/VR環境におけるインタフェースや感覚間の相互作用を用いたインタフェースを研究しています。また、スマートフォンなどの携帯デバイスを用いたインタラクション技術や行動認識、認証技術に関する研究も行っています。
- ・ASANOシステム
オーバーレイネットワーク技術(SD-WAN)を活用し、あらゆる場所で、必要とされるリソースとセキュリティ機能を備えたネットワークを即座に提供できるシステムを研究開発しています。

Research Subjects

- ・Internet Trust Engineering
We are engaged in network/information security science and engineering, including formal and analytical internet modeling of security, trust engineering of Internet trust frameworks and related methodologies based on cryptography. We are also investigating mathematical models and applications of machine learning.
- ・Advanced Networking
We are researching two fundamental parts of computer networking: network stacks in host operating systems for high-speed networking and backbone networks for flexible traffic engineering. We are also investigating architectures for information systems specialized for real-time IoT systems.
- ・User Interface for Enhancing Daily Life
We are researching interfaces in AR/VR environments and cross-modal interfaces focused on interactions between the sensory modalities, with the aim of creating new experiences. We are also investigating 3D interaction techniques, context recognition, and user authentication using mobile devices.
- ・ASANO System
We are researching and developing systems that can immediately provide network access and security functions at any place using overlay network technology (SD-WAN).

〈情報セキュリティ研究体〉

2018年10月1日、東京大学のネットワークの運用によって得られた知見を基に実践的なサイバーセキュリティを研究するため、情報基盤センターに情報セキュリティ研究体が設置されました。情報基盤センターでこれまで開発してきた、トラフィックの異常検知技術や、サイバー攻撃からの回復性のある多層的な防御技術を基に、サイバー攻撃の予兆を検知し、システム運用での早期警戒に役立てる研究に取り組んでいます。また、情報セキュリティ研究体は連携研究機構として情報セキュリティ教育研究センターに参加し、東京大学の各部署でのセキュリティ担当者の人材育成に必要な実務教育を提供しています。

〈Security Informatics Initiatives〉

The Security Informatics Initiatives (SII) was established in the Information Technology Center (ITC) on October 1st, 2018 to conduct practical cybersecurity research based on knowledge obtained through the UUNET. Current research areas include traffic anomaly detection technology for preemptively identifying signs of cyberattacks during system operation and multi-layered defense technology which is resilient against cyberattacks. The ITC also participates in the Security Informatics Education and Research Center and provides practical training programs for operators at the University of Tokyo and outside.



計算で「できない」を「できる」に変える

スーパーコンピューティング研究部門

次世代のスーパーコンピュータの姿、科学技術計算の手法を研究し、日本の計算基盤の中核として大学・研究機関にスーパーコンピュータを提供しています。

Supercomputing – Making the Impossible Possible Supercomputing Research Division

The Supercomputing Research Division is conducting research on hardware, software, algorithms, and applications of next-generation supercomputers. The division also provides supercomputing services as the core computing platform for universities and research facilities throughout Japan.

「計算・データ・学習」の融合による 革新的スーパーコンピューティング

Society 5.0とは、デジタル技術によってコンピュータ上に形作られた仮想空間と現実空間を高度に融合させ、経済発展と社会的課題解決とともに実現する、人間中心の社会です。狩猟・農耕・工業・情報社会に続く新たな社会として提唱されました*1。これまでスーパーコンピュータは、主として計算科学・工学シミュレーションに使用されてきましたが、これにデータ科学、機械学習の手法を融合させることにより、Society 5.0のような新たな社会の実現が可能となります。

我々は、「計算・データ・学習(S+D+L)」融合を実現するスーパーコンピュータシステム(通称Big Data & Extreme Computing: BDEC)の構築を目指して、2015年頃からさまざまな研究開発を進めてきました。2021年5月に運用を開始した「Wisteria/BDEC-01」はBDEC構想に基づく最初のシステムで、国内最大級の規模を有しています。「富岳」と同じ汎用CPU(A64FX)を搭載したシミュレーションノード群(Odyssey)とGPUクラスタによるデータ・学習ノード群(Aquarius)で構成されています。シミュレーション向け、データ解析・機械学習向けの、異なるアーキテクチャを組み合わせたスーパーコンピュータシステムは、世界でも初めてです。Wisteria/BDEC-01は大容量の共有ファイルシステムとSSD搭載高速ファイルシステムを備えています。さらにAquariusの一部は大容量の通信回線で外部ネットワークに直接接続可能であり、SINETなどを介してセンサーネットワークからリアルタイムに観測データなどを取得することが可能です。

*1 Society 5.0: https://www8.cao.go.jp/cstp/society5_0/
https://www8.cao.go.jp/cstp/english/society5_0/index.html

Wisteria/BDEC-01のような複合的システムにおいて、最大限の能力を引き出しつつ計算量・消費エネルギーを最小限に抑え、「計算・データ・学習」融合を実現するのが、革新的ソフトウェア基盤「h3-Open-BDEC」です。センター内外の計算科学、計算機科学、数値アルゴリズム、データ科学、機械学習の専門家が協力して開発しています。Wisteria/BDEC-01とh3-Open-BDECによって、観測データとシミュレーションを組み合わせたデータ同化、機械学習などによるシミュレーションモデルやパラメータの最適化などが、容易になります。実現象の解明に使用される非線形シミュレーションでは、パラメータを変えて何度も実行する必要がありますが、機械学習によるモデル・パラメータ最適化によって、実行回数を大幅に削減可能です。さらに変動精度演算を取り入れてシミュレーション時間の短縮を図り、従来と同等の正確さを保ちながら、大幅な計算量・消費エネルギー削減が達成可能です。我々は、h3-Open-BDECによる「計算・データ・学習」融合手法の普及を目指しています。

■ Wisteria/BDEC-01の構成

	Odyssey	Aquarius
総理論演算性能	25.9 PFLOPS	7.2 PFLOPS
ノード数	7680	45 (90 CPU, 360 GPU)
総メモリバンド幅	8.38 PB/秒	
ノード構成	Fujitsu A64FX × 1	Intel Xeon Platinum 8360Y × 2 NVIDIA A100 Tensor Core GPU × 8
共有ファイルシステム	25.8 PB	
SSD搭載高速 ファイルシステム	1.0 PB	

Innovative Supercomputing by Integration of Simulation/Data/Learning (S+D+L)

Society 5.0 is a human-centered society that balances economic advancement with the resolution of social problems by means of a system that highly integrates cyberspace and physical space. It follows the hunting, agricultural, industrial, and information society. While supercomputers have been mainly used for simulations in computational science and engineering (CSE), the integration of CSE with the methods of data science and machine learning will enable the realization of such a new society.

We have been conducting various research and development since around 2015 to build a supercomputer system that realizes the integration of simulations in CSE, data science, and machine learning (S+D+L). The project is called Big Data & Extreme Computing (BDEC). Wisteria/BDEC-01, which began operation in May 2021, is the first BDEC system, and it consists of a node group for simulation (Odyssey) using general-purpose CPUs and a node group for data/learning (Aquarius) using GPU clusters. This is the first supercomputer system in the world to combine different architectures for simulations and data analysis/machine learning. Wisteria/BDEC-01 is equipped with a shared file system with large capacity as well as a fast file system with SSDs. In addition, a part of Aquarius can be directly connected to an external network via a high-capacity communication line, making it possible to acquire observation data in real-time from sensor networks via SINET.

h3-Open-BDEC is an innovative software platform that integrates S+D+L by minimizing computation and energy consump-

tion and maximizing effectiveness on heterogeneous supercomputers such as Wisteria/BDEC-01. h3-Open-BDEC is a collaborative development effort between computational science, computer science, numerical algorithms, data science, and machine learning. Wisteria/BDEC-01 and h3-Open-BDEC facilitate data assimilation combining observation data analysis/utilization and simulation, and optimization of models and parameters for simulations. While nonlinear simulations for investigating real phenomena require multiple runs with different models and parameters, the number of runs can be dramatically reduced by optimizing the models and parameters through machine learning. In addition, incorporating adaptive-precision computing can reduce simulation time and energy consumption. Our goal is to promote the integration of S+D+L through h3-Open-BDEC.

■ Wisteria/BDEC-01 Configuration

	Odyssey	Aquarius
Total theoretical calculating performance	25.9 PFLOPS	7.2 PFLOPS
Total number of nodes	7680	45 (90 CPU, 360 GPU)
Total memory bandwidth	8.38 PB/s	
Node configuration	Fujitsu A64FX × 1	Intel Xeon Platinum 8360Y × 2 NVIDIA A100 Tensor Core GPU × 8
Shared file system	25.8 PB	
Fast file system with SSD's	1.0 PB	

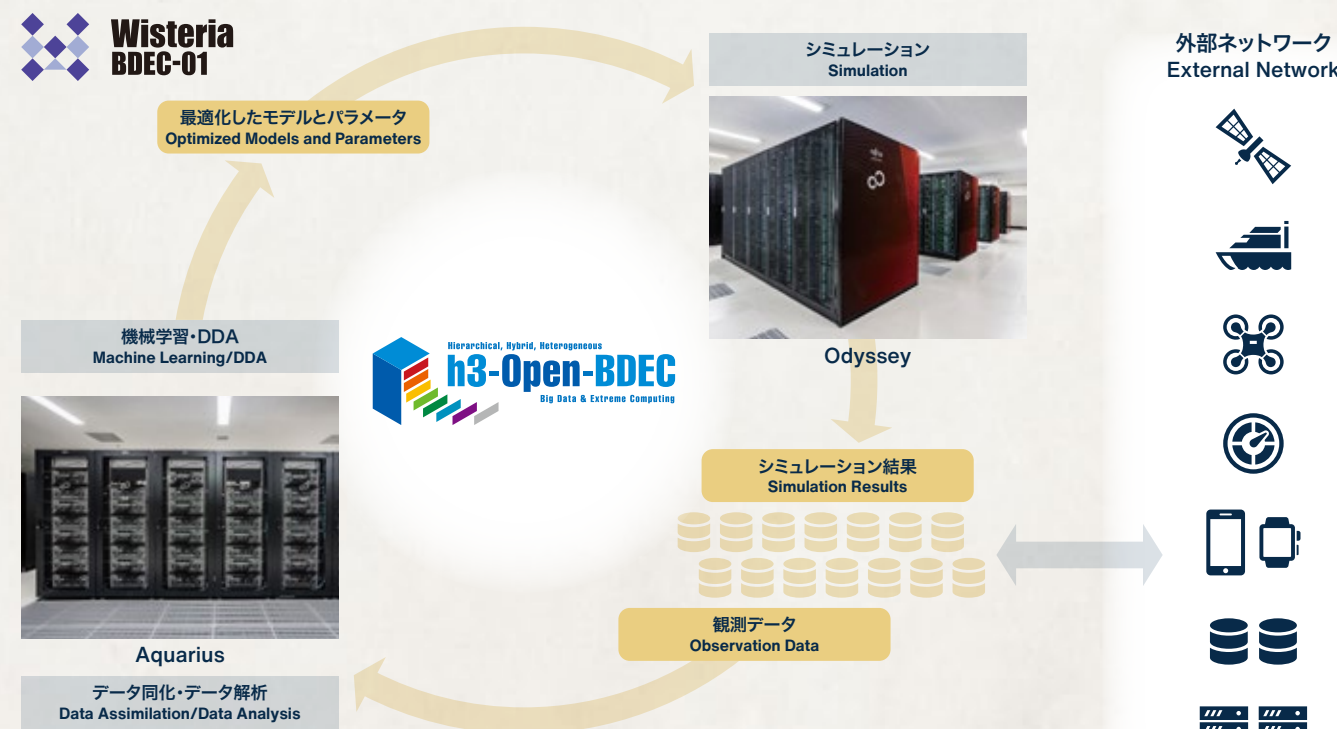


図: Wisteria/BDEC-01はシミュレーションノード群のOdysseyとデータ・学習ノード群のAquariusで構成されている
Figures: Wisteria/BDEC-01 consists of Simulation Nodes Odyssey and Data/Learning Nodes Aquarius

JHPCN(学際大規模情報基盤共同利用・共同研究拠点)

8大学センターの専門家と計算機資源が全国の研究者と連携、計算科学・データ科学・計算機科学の学際研究フィールドを切り拓く

JHPCN : The Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures

Eight University centers and their staff collaborate with researchers nationwide to expand interdisciplinary research fields in computational science, data science, and computer science.

計算科学・データ科学・計算機科学を 広範な学際研究に融合する

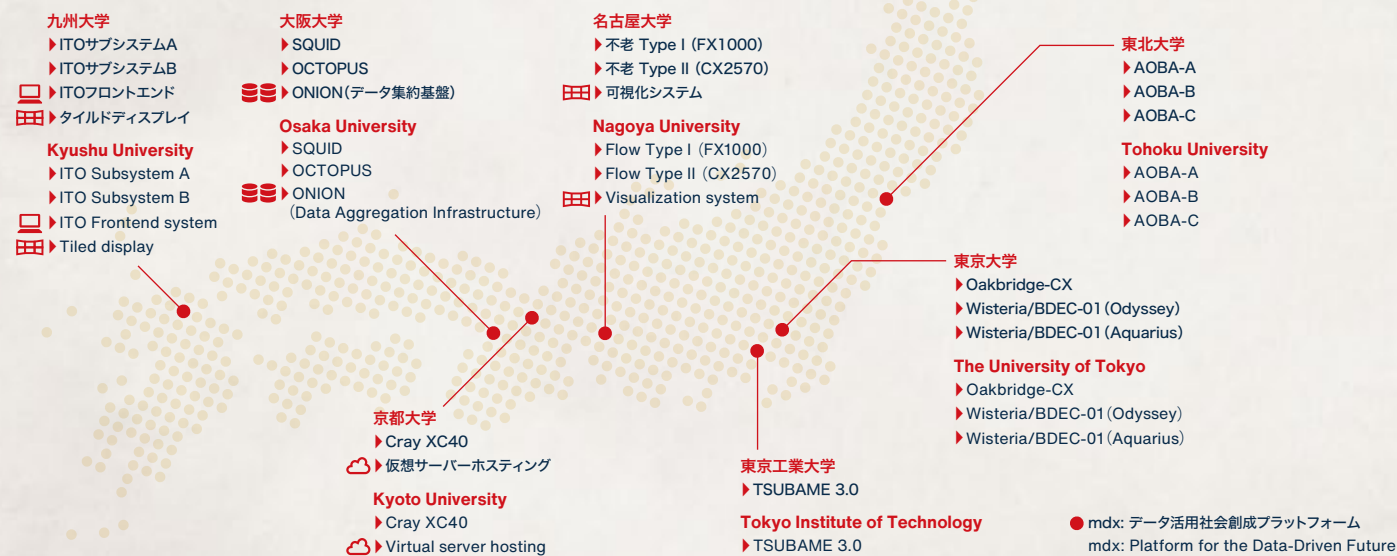
計算とデータは、さまざまな学術研究をより深く、より速くへ進化させる大きな力となります。「学際大規模情報基盤共同利用・共同研究拠点」(JHPCN)は、計算科学・計算機科学の専門家の知見と大規模な計算資源をもって、自然科学、工学から人文社会科学までの広範な分野と、学際的な共同研究をしていく活動です。それぞれ異なるタイプのスーパーコンピュータを運用している8つの大学センター(北海道大学、東北大学、東京大学、東京工業大学、名古屋大学、京都大学、大阪大学、九州大学)が協力して運営し、その中核拠点を東京大学が担っています。2022年度からは、データ科学・データ活用に主眼をおいた計算資源「mdx」(8大学、国立情報学研究所、産業技術総合研究所、筑波大学人工知能科学センターが共同運用)が加わりました。

■ JHPCNの拠点と多彩な計算機資源(2022年度・抜粋)

汎用CPU、GPU、ベクトル型、「富岳」と同系統など、種類の異なるスーパーコンピュータ群、仮想サーバーホスティング、ソフトウェア、可視化システムに加えて、データ科学・データ活用に最適化されたmdxを利用できます。

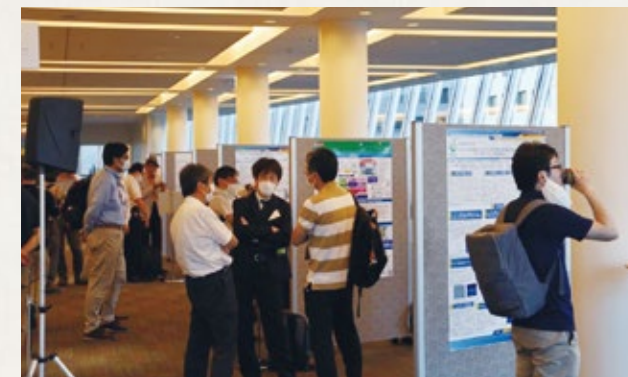
■ JHPCN sites and diverse computing resources as of FY2022

Users can use the different types of supercomputers, softwares, virtual servers and visualization systems, as well as “mdx”, which is optimized for data science and data utilization.



学際研究 — バックボーンの異なる 研究者が交差するところ

JHPCNの共同研究には、年に1回公募される一般・国際・企業共同研究課題(審査あり)と、各大学センターが推薦する萌芽型共同研究課題の2種類があります。提案課題が採択されると、JHPCNの計算機資源を無償で利用でき、各センターの先導的な研究者と共同で研究が進められます。また、論文出版、海外出張、関連シンポジウム、国際課題については海外の共同研究者との打ち合わせ費用がサポートされます。各課題代表者に発表いただく年次シンポジウムでは、研究手法・成果の共有を通じ、JHPCNの精神である学際研究へとつながるネットワークを促進しています。



Interdisciplinary research—where researchers with different backbones intersect

There are two types of JHPCN joint research themes: general, international, or corporate joint research proposals, and exploratory joint research themes recommended by each university center. If a proposal is accepted, the research group will have access to the JHPCN's computational resources free of charge and conduct joint research with the leading researchers of the other centers. In addition, publication fees, participation in international conferences, and meetings with overseas collaborators (in the case of international proposals) will be supported.

At the annual symposium, each project leader is invited to give a presentation. In addition, they can network and share their research methods and results, as the encouragement of interdisciplinary research is the spirit of the JHPCN.

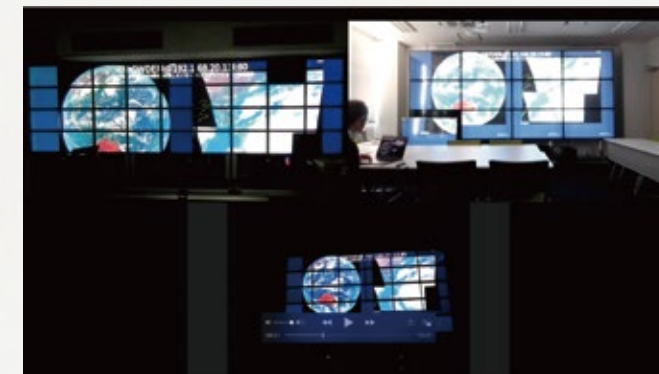
JHPCNシンポジウムのポスターセッション(2022年度)。採択された研究課題のポスター発表に、バックグラウンドの異なる研究者が足をとめて話し込んでいく。終了した研究課題は口頭発表で広く共有される。

Poster session at the JHPCN Symposium (FY2022). Researchers from different backgrounds meet to discuss the poster presentations of the accepted research themes. Completed research themes will be shared in the form of oral presentations.

研究課題紹介 Example of Joint Research Project

全国の拠点を結び、高速・大容量のネットワーク技術の開発を目指す(村田健史／情報通信研究機構)

各分野で著しく進展しているビッグデータサイエンスの加速を目指して、超高速・大容量通信技術をJHPCN設置当初から研究しています。全国に所在するJHPCNの拠点を高速ネットワークで結び、大規模シミュレーションの分散化、多様なセンシングデータのリアルタイム収集、超高解像度データの可視化など、さまざまな分野で活用されるネットワーク関連技術の基礎開発や、実データを使った検証を可能としてきました。今後はJHPCN8拠点からさらにスケールアップした広域分散クラウドを構築し、次世代のビッグデータサイエンスを牽引する通信技術研究に取り組んでいきます。



Developing high-speed, high-capacity network technology by connecting sites across Japan (Takeshi Murata/NICT)

Towards accelerating the progress of big data science in various fields, we have been researching ultra-high-speed, large-capacity communication technology since the establishment of JHPCN.

By linking JHPCN sites located throughout Japan with high-speed networks, we have developed basic network-related technologies for various fields such as distributed large-scale simulations, real-time collection of diverse sensing data, and visualization of ultra-high resolution data, as well as verification using actual data.

In the future, we will build a wide-area distributed cloud that is further scaled up from the eight JHPCN sites and engage in communication technology research that will drive the next generation of big data science.

超高解像度かつ時系列のひまわり衛星画像の3拠点同時表示実験の様子
Simultaneous display experiment of ultra-high resolution and time-resolved images taken by Himawari satellite at three different sites.



公募、シンポジウム、最新の計算機資源、これまでの採択課題については、JHPCNのウェブサイトを参照してください。
Please refer to the JHPCN website for information on open calls, symposia, the latest computing resources, and previously accepted proposals.

日本語 <https://jhpcn-kyoten.its.u-tokyo.ac.jp/ja/> English <https://jhpcn-kyoten.its.u-tokyo.ac.jp/en/>