

データ利活用の推進に向けた 情報基盤整備の現状と課題等について

SINET（学術情報ネットワーク）の概要

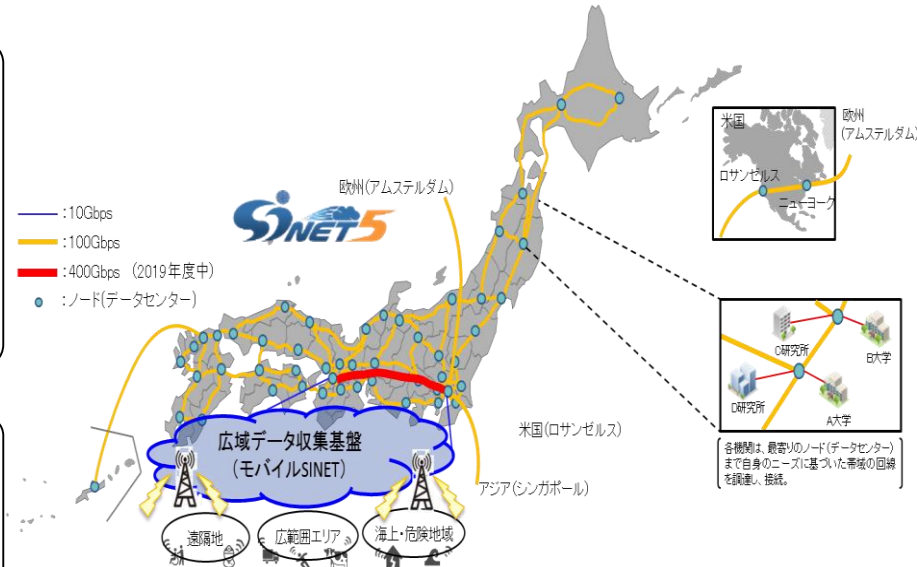
- ・日本全国の国公立大学、公的研究機関等を結ぶ超高速・セキュアなバックボーンネットワークを国立情報学研究所(NII)が、効率的に整備・運用。
- ・国内のみならず、海外学術ネットワークとも相互接続し、国内外の最先端研究等に大きく貢献。

SINETの変遷

- 1992年： SINET運用開始(29拠点、6~50Mbps)
 ※前身の学術情報ネットワークは1987年運用開始
 スーパーSINET(14拠点、最大10Gbps)
 SINET3(34都道府県、1~40Gbps)
 SINET4(47都道府県、2.4~40Gbps)
 2016年-: SINET5(47都道府県、100Gbps)

SINET5の概要

- ・運用期間： 2016年4月～2022年3月まで。
- ・全都道府県を**100Gbps**で網目状に構築。
東京-大阪間に400Gbps回線を増設予定(2019年度中)。
海外(米国、欧州、アジア)の学術ネットワークとも100Gbpsで接続。
- ・モバイル網とSINETを直結し、広範なエリアから収集したデータの利活用を促進する「**広域データ収集基盤(モバイルSINET)**」を運用。



【加入機関数と内訳】

	国立大学等	公私立大学	短大・高専	国立研究法人	民間	その他 (独法、財団・社団法人、 公立機関、等)	計
加入機関数 (2019.3.31)	102	481	136	26	9	156	910

今後の課題

- ・大学をハブとした**地方創生・産学連携の基盤としての活用促進**(利用多角化・民間開放に伴う制度設計含む)
- ・**広域データ収集基盤(モバイルSINET)の高度化**(5G対応等)
- ・遠隔教育等の基盤としての**初等中等教育への開放**(全国の小中高校(約36,000校)のうち希望する機関をSINETに接続)

上記課題を踏まえつつ、年々増加するデータ流量や、学術コミュニティ等からの要望に対応するための**SINETの高度化を推進**
 (→次期SINETの構築に向けた検討、基盤の増強、コンテンツの充実等に向けた検討等)

Society5.0を実現するためのデータ活用による知識集約型社会の創成

ーデータ活用社会創成プラットフォームの構築ー

データ活用社会における現状認識

- ICT機器の爆発的な普及や、AI、ビッグデータ、IoT等の社会実装が進むなど競争が激化、一部の企業や国のデータの囲い込みにより経済社会システムの健全な発展が阻害される懸念。
- 我が国が成長していくためには、デジタル新時代において、データを我が国全体の共同資産として、スピード感をもったデータ利活用環境の整備が急務。
- Society5.0が目指すインクルーシブな社会を実現するためには、地域における知識集約の中核を担う大学を起点としてイノベーションの創出を図り、知識集約型社会を構築することが重要。
- サイバー空間とフィジカル空間が融合するデジタル新時代において、我が国に蓄積された農業、医療・健康の分野、教育データを含む多くの有用なビッグデータを共同で活用する上で、人材と技術を有する全国の大学を超高速・高信頼で網目状につなぐ国際的優位性をもつSINETを最大限活用することが重要。
- 異種データや異種知識の融合・活用を促進するための「場」として、様々な分野のデータ保持者、解析者、利用者が参画するコミュニティを形成するとともに、データ活用を目指す利用者へのコンサルティングやアプリケーション開発支援が不可欠。

文部科学省における取組

- 経済財政運営と改革の基本方針2018や未来投資戦略2018において重要性が指摘されているリアルデータの利活用を念頭に、データ活用社会創成プラットフォームを推進するため、SINETを通じて収集されるリアルデータの集積や、解析結果を速やかにフィードバックする機能を備えたシステムを整備(2019年度予算)
- 文部科学省と大学コミュニティ、地域社会等が一体的に連携し、全国の国立大学等をハブとしたデータ活用社会創成プラットフォームの実現促進に向けた検討を行うための「データ活用社会創成プラットフォームの推進に関する有識者会合」を設置。
- 地域・産業・社会基盤を支える拠点となる大学を中心として、民間への利用拡大も視野に我が国全体の知識集約型社会の実現に向けた環境「データ活用社会創成プラットフォーム」を構築

データの高度利活用環境(NII・東大に先行して整備)

【設備整備】

IoT接続(モバイル)
AI特化サーバ
リアルタイム処理対応サーバ
高速/セキュアストレージ 等



SINETを通じて、全国のデータ収集・通信・解析環境をオンデマンドで活用。
高度・多様なデータ利活用により新たな価値を創出。

利活用ニーズを踏まえたシステム整備・ソフトウェア開発

【大学等におけるデータ利活用の潜在的なニーズ】

- ・地域農業・漁業・観光業のスマート化
- ・認知症・生活習慣病などの早期発見、予防方法の提案
- ・スポーツ科学への応用
- ・初中段階から高等教育、社会人教育に至る一貫した教育データの利用 等

文部科学省と大学コミュニティ、地域社会等が一体的に連携し、プラットフォームの実現に向けて整備・検討を加速

IoTデータ接続
(モバイルSINET)



データ活用社会創成プラットフォームの推進に関する有識者会合

リアルタイム処理対応基盤社会創成プラットフォームの実現に向けた実務的な検討を行う場

【主な検討課題】

- ・リアルタイムデータの解析・活用を目的とした基盤ソフトウェアの研究開発や技術の実証のための基盤システムの整備のあり方
- ・産学連携体制(コミュニティ)の構築・強化、その中核としての大学の役割等一体的な連携を確保する仕組み 等



大学等連携コンソーシアム

大学を中核としたデータ活用実務機関が連合したコンソーシアム

【主な取組】

- ・データプラットフォームの活用促進、データ活用ニーズ調査
- ・コミュニティ間連携の強化・促進等



データ活用社会創成プラットフォームの推進に向けた議論の観点

1. データプラットフォームが求められる背景

- (ア) 地方の産業、自治体を含む我が国全体でのデータ利活用推進の重要性
- ・データ活用社会は、**地域格差などの課題を解消し、地域の特色を活（い）かした多様な産業の活性化に貢献するSociety 5.0が目指す社会**であり、知・人材が集積する大学等を中核とした知識集約型拠点を形成・実現していく必要がある。
 - ・これからの社会では**あらゆる分野でデータの利活用が必須であり、学術コミュニティにおいても、社会実装を指向した、いわゆる実学をベースとした研究分野が重要**である。
 - ・しかし、データ利活用においては、必要となる異分野のデータの把握、それらのデータを利活用する解析ノウハウの獲得、更に解析に用いるITインフラの構築を行う必要があり、**データ利活用に至る障害の大きさから、様々なデータが広く有効活用されていない**。
- (イ) データ利活用の推進に向けた早急な利活用環境の整備の必要性
- ・データの利活用を推進するためには、**Proof-of-Concept (PoC) によりその効果を実証することが重要**であり、PoCをいかに迅速に行えるかが活用の成否を左右する。
 - ・データ利活用を促進するためには、データ量の大小に関わらず、自身が有するデータと利用可能な公共データを合わせて様々な解析を行い、**より高い技術を習得する「データ工房」として位置づけていくことが重要**である。

2. データプラットフォーム推進における課題

- (ア) 地方創生、地方活性化に向けたデータ活用のための人材に関する課題
- ・データの利活用を推進していくためには、データを実際に生成する地方・地域社会の現場の知識と、データ活用の知識の両者が必要であるが、**アイデアがあっても実際にデータを収集・集積・解析する情報技術が分かる人材がおらず、データ活用の知識・ノウハウを持つ人材とのつながりが少ない**ことも多い。
 - ・地方・地域社会でのデータ活用は、**活用の範囲をそれらの地方・地域内で閉じて考えがち**なため、波及効果が限定的でエコシステムも成立しないことが多く、**広く展開するシステムがない**。
- (イ) データ利活用のインフラ（基盤）の未整備と方法論の確立に関する課題
- ・現時点において、**大量データの処理、リアルタイム処理に利用可能なインフラがない**。
 - ・地方・地域社会の現場では、データを活用することで何が実現できるかについて十分な知見がなく、解決したい課題に対し、どのようなデータと解析技術があれば実現できるのかが分からない。
- (ウ) データ流通における課題
- ・データは、量・種類ともに多く、大量のデータを分析するためには専門的な知識を有した人材が必要であるが、**人材が不足していることもあり、どのようなデータが利用可能であるか、そのデータをどうすれば入手できるのかといった情報が広く共有されていない**。
 - ・データを有する研究者や企業等が**データを他に活用させることへのインセンティブがない**ため、有用なデータが流通しないことや、個人情報などの問題によりデータの活用などが進んでいない。
 - ・従来の産学官連携の仕組みは、どのような技術が必要か明確な場合にのみ機能するものが多く、潜在的な需要を拾い上げるのには適していない。
- (エ) エコシステムの機能に関する課題
- ・エコシステムへの期待が高まる中、**潜在的なデータ活用の需要を拾い上げるエコシステムが存在しない**。また、データ活用にコストをかけることによって得られる利益について、全事例が成功するとは限らずリスクがあり、当初のデータ利活用を担う主体が、広範囲に広がるデータ利活用を最後まで担うとは限らず、利益を生み出すエコシステムが確立されていない。

データ活用社会創成プラットフォームの推進に向けた議論の観点

3. 実現すべきデータ活用を促進する長期的な仕組み

(ア) 基盤整備と先導アプリケーション開発

- データ利活用の方法論を確立するためには、データ提供者・利用者が利用可能な共通基盤を整備するとともに、**先導的なデータ利活用アプリケーション開発プロジェクトを推進**し、成功事例を蓄積、再利用することが必要である。具体的には、共通基盤の整備やアプリケーション開発の推進を通じて、全国に分散した設備や人材を適材適所で活用することにより、オールジャパンの体制で取り組むべきである。

(イ) データ利活用を推進する体制の構築

- 全国的なデータ利活用体制を構築するためには、**データの整備・流通を図るとともに、データ活用のコンサルティング機能、データ提供者/データ解析技術者/データ利用者のマッチング機能を全国に展開する必要がある**。全国展開に当たっては、先導的機関による取り組みから事例を蓄積するとともに、**全国の大学がハブとなり、地方自治体や企業も含め展開するコミュニティを形成していくことが重要**である。

(ウ) セキュリティの確保の仕組みの検討と、セキュリティルールの形成

- 医療情報等、機微なデータを扱う場合等も想定し、扱うデータに応じ最適なレベルのセキュリティを確保する方策や、基盤運用上のセキュリティを高める方策を検討し、**常にセキュリティレベルの向上に努めていく必要がある**。
- 実際に機微な情報をデータ保有者が提供するためには、セキュリティ要件に関して広く認められた基準を設定する必要がある。このため、法制化等、社会的に認められるセキュリティレベルの基準を定める方策が必要である。

(エ) エコシステムの形成（全国的なデータ利活用のエコシステムを形成する）

- 共通基盤やデータ利活用推進体制を維持するための予算施策と、全国的なデータ利活用サービスを新たな産業として安定させるための施策を検討すべきである。
- データ利用の許諾に関わる手続をデータ利用者・データ提供者双方が速やかに進められる仕組みが必要であり、また、**データ利用中の事故や不適切なデータ利活用等が発生した場合の訴訟等のリスクや、その保障制度等の対策について検討し、利用者に提示する仕組みが必要**である。

4. 整備の現状

- データ利活用を推進していく上での現状のデータ利活用の共通基盤として、①DIAS（Data Integration and Analysis System）や、② ABCI（AI Bridging Cloud Infrastructure）が運用されているとともに、全国利用を前提としたデータ活用社会創成プラットフォームの先導的システムの整備が予定されており、今後は、本設備を運用するためのアプリケーション開発やデータ利活用を推進する体制の構築について検討を進めていく必要がある。
- これらの共通基盤（プラットフォーム）は、**全国の900以上の機関を100Gで接続し、広帯域かつ、セキュアな通信環境を実現するSINET（学術情報ネットワーク）に接続されるものであり、今後の有効な利活用環境の構築が期待される**。なお、SINETへの新たなアクセス環境として、モバイル網による広域的なデータ収集基盤の構築も進められている。

5. 当面の方策

- 課題や現状を踏まえ、データ活用のためのアイデアを持っていてもそれを実現するための**データ収集・集積・解析する仕組みを実際に作ることが困難な地方大学や自治体等に、それを可能にする「場」を提供し、解析されたデータから新たな「知」の創出を容易にするシステム**としての「データ活用社会創成プラットフォーム」の実現を目指し、**リアルタイムデータの解析・活用を目的としたリアルタイム処理対応基盤ソフトウェアの研究開発や、技術の実証のための基盤システム整備、産学連携体制（コミュニティ）構築・強化、その中核としての大学の役割等について、重点的、集中的に検討・整備していく必要がある**。