

データ連携基盤の構築に向けたデータ利活用ルールに関する方向性 (案)

1. 政策目標

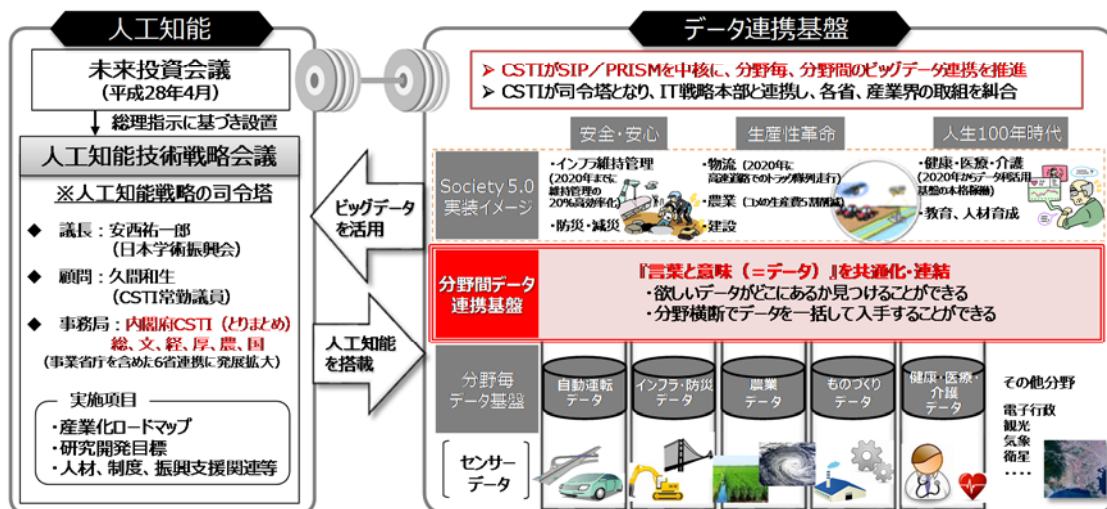
(新たな経済政策パッケージ 2017 年 12 月 8 日閣議決定)

国・自治体の各行政機関や企業等の民間機関の間で散在するデータをすべて連携することを旨とし、「横断的分野」(位置、時間等)と「固有分野」(農業、インフラ等)双方について、データ標準や共通語彙基盤 (IMI) の横断的なデータ活用を推進するための基盤を3年以内に整備することとし、そのためのシステム開発を開始する。並行して米国、欧州のデータ連携基盤とのデータ連携を検討し、日米欧 10 億人のデータ共通市場を創設することを目指す。

Society 5.0の本格実装 ～ データ連携基盤の整備～

2017.12.25
CSTI本会議資料

- Society 5.0の本格実装に向け、CSTIが司令塔機能を発揮し、データ連携基盤の取組みを推進
- 「人工知能」とビッグデータを共有・活用する「データ連携基盤」はSociety 5.0実現のための両輪
- ※ 欧米は、政府公共データを中心に様々なデータが繋がる仕組みを構築しつつあり、「データ連携大競争時代」が到来
- ※ 日本は、各省、産業界の取組むデータベースがバラバラなため、データ連携の仕組みの構築が急務



- 世界に先駆けて人工知能を搭載し、あらゆる分野のデータが垣根を越えて繋がるデータ連携基盤を構築
- あらゆる分野から融合したビッグデータと人工知能を活用し、安全・安心、生産性革命、人生100年時代に大きく貢献

2. データ連携基盤の基本的な整備方針

- 分野共通のコア語彙、分野ごとのドメイン共通語彙・固有語彙やデータ構造を整備することで、分野横断でのインターオペラビリティ(相互運用性)を可能とする。
- また、データカタログ(メタデータ)を用い、欲しいデータがどこにあるかを検索でき、標準 API を介して様々な分野のデータをワンストップで入手することを可能とするサービスプラットフォームを構築する。
- データ活用促進の観点から、個人情報保護、プライバシーに配慮しつつも過剰な保護により、利活用を阻害しないようにバランスをとり、これまでの発想や利害関係にとらわれない新規サービスを生み出せる制度、ルールを構築する。
- 分野毎のデータ基盤とは、準備が整い次第、順次接続を開始する。

3. データ連携基盤の利活用ルールに関する検討事項

(1) データ利用権限のポリシー

- 営業秘密と整理されるデータ、国家安全保障に関わる等機微なデータ、個人情報を含むデータ、オープンデータ等の様々なデータの分野間連携を想定した場合、様々な利用権限に応じた対応が必要となる。
- 営業秘密(データ)については、ビジネス上は、データ提供者と利用者との間の1対1での契約を中心に進められている。
- 個人情報保護、プライバシーへの配慮をしつつも、過度な保護による利活用を阻害しない多数対多数の分野間連携基盤を推進するにあたってのルール整備が求められている。

【方向性】

- オープンデータはフルアクセス(利用者登録不要)
- 利用者を制限すべき情報や機微な情報についてはアクセス制御(利用者登録要)
- 個人情報を含むデータについては、提供者側で匿名加工済みのデータを扱う、あるいは、PDS、情報銀行を活用
- 営業秘密データについては、データ連携基盤とは別に、契約プロセスを別途行った上で、双方のアクセス権を決定し、基盤上でアクセス権を設定・制御

(2) データ品質(完全性、正確性、有用性、安全性等)の確保

- 様々な利用者からの利用を想定した場合、データ品質に関するニーズがある。
- 他方で、より多くのデータを集めるためには、データの提供を妨げるような評価をすべきではない。

【方向性】

- データを扱う利用者からは、データの品質を正確に知る必要があり、データそのものを入手する前に、メタデータに適切な項目でデータ品質が明記されていることが望ましい。
- 一方で、分野毎にメタデータの指標が異なるため、メタデータにも共通的な項目(作成組織、作成日、最終更新日、ライセンス等)とそれ以外の任意項目を設けて、今後作成されるメタデータについては、記述の方向性を示すとともに、メタデータを登録する仕組みを設けることが必要。
- 基盤として評価の基準づくりをすることで、それぞれのデータ品質を統一した基準のもとで表現されることが望ましい。その上で評価の活用に関しては、ユーザー側に任せるべき。

(3) データ提供者へのインセンティブ付与

- 効果的なデータ連携基盤の活用のためには、イノベーションの源泉となる価値のあるデータをいかにデータ連携基盤に繋げることができるかが重要。
- そのためには、分野毎の基盤運営者、データ提供者の協力が不可欠である。
- 一方、メタデータの整備等、データ提供者の負荷が発生するため、メリットがないと協力が得られにくい。

【方向性】

- データ提供者への対価の還元の仕組み作りや、公的資金による研究開発成果の1次データの公開推奨(〇年以内、機械判読可能なデータ形式で)などのルール化。
- 分野間連携を促進するためには、外部に向かってメリットを発信することや、データ提供による対価が得られる仕組みの構築が重要。効果が見えやすい分野間連携の具体的な姿、仕組みを想定し、象徴的な案件で良好事例を作る。
- フォーマットやデータ形式等を標準化は重要であるが、既存のデータに関しては利活用の観点から相互運用性が優先され、変換機能の実装など、技術的に解決されるべき。

(4) データ連携基盤を取り巻くエコシステムの形成の促進

- 政府資金に頼らず、持続的に自立してデータ連携基盤を運用するためには、データ連携基盤の維持・管理・保守費用を賄うエコシステムの形成が必要。

【方向性】

- データ連携基盤整備のビジョンは、日本の強みであるセンサー／デバイスから吸い上げられるリアルデータをも活用し、政府・産業界のデータ利活用を促進するサービスプラットフォームとして機能させる。
- 内閣府が IT 戦略室と協力のもと、司令塔機能を発揮し、分野間の調整・合意形成ができる体制を構築する。また、実施にあたっては、SIP、PRISM 等の取組を活用し、PDCAサイクルを回しつつ、効果的な分野間連携ができるところから進めるとともに、将来的な基盤運営については、民間連携と拡張を目指す。
- センサー／デバイスなどのIoTデータなどの、足が速い(変化が多い)分野については、民間主導により、データ形式等のルール作りがなされることが基本。