

- これまでの国内における取組事例に関するヒアリング等を踏まえると、ビッグデータの活用については、例えば、次の課題があると考えられる。

データの生成・収集・蓄積における課題(例)

- ☞ 同種のデータが生成されるセンサー等の管理者が様々である場合におけるデータ収集の在り方。
- ☞ 多種多様なセンサー等からのデータ収集等におけるセンサーの設計や統計的手法との連携等、フィジカルな領域におけるICT人材や統計等の専門家の活用の在り方。
- ☞ データの収集におけるGPSの普及やデータ入力に関するユーザインタフェース等の利用者からのデータ収集の在り方。
- ☞ 個人に関する情報を利用するサービスの提供における契約約款による同意の取得や重要事項の説明等の在り方。
- ☞ 例えば、スマートフォンにおけるユーザーからのパーミッションの取得時の説明等、通信の秘密や個人情報保護法上の「個人情報」に該当する場合における事前の利用目的の特定(同法第15条)、本人に対する利用目的の明示(同法第18条)や不正な手段による取得の禁止(同法第17条)への対応の在り方。
- ☞ センサー等で生成されるデータをリアルタイムに収集・処理するための帯域や遅延時間等のネットワークや処理システム等の基盤の在り方。
- ☞ センサー等からのデータ収集におけるネットワーク化するモノの多さとの費用対効果や、センサーをばらまくことによる100メートル以下における無線の競合の回避の在り方。
- ☞ 正確に特定された時間に関するデータの活用が必要とされる場合における利用者による確認の必要性等のネットワーク輻輳への対応の在り方。
- ☞ 企業等利用者におけるデータ活用の重要性の認識やIDによる紐付け等のデータ整備・管理の在り方。
- ☞ 複数主体からの多種多量のデータをクラウド等により取扱う場合における適切なデータ管理の在り方。

ビッグデータの活用に関する課題②

データの流通・連携における課題(例)

- ☞ 情報を的確に常時送る場合における放送メディア等の全国的なインフラ整備の在り方。
- ☞ 指数関数的に増加するデータトラフィック等ビッグデータを支える基盤となるサーバ、ストレージ及びネットワーク等について、ソフトウェアでフローをコントロールする技術やHadoop等の各種ストレージ技術等の活用や、電力等を考慮したアーキテクチャ等の在り方。
- ☞ 携帯電話ネットワークを活用したデータの送受信やサービス販売のプラットフォームとしての携帯電話の活用におけるスマートフォンへの対応の在り方。
- ☞ 例えば、契約約款の変更について事前同意がある場合における当該約款の内容による利用目的の変更や統計的利用のための完全匿名化等、個人情報保護法上の「個人情報」に該当する場合における第三者への提供時の事前の同意取得(同法第23条)への対応の在り方。
- ☞ 位置情報等の情報自体やそれに基づく一定の事実のみの提供等、情報の種類に応じたりリスクとビジネスチャンスとを踏まえた、情報の複数事業者間における授受等の取扱いの在り方。
- ☞ 例えば、ツイッターの内容の第三者への表示等、第三者に提供すべきデータの解析結果にユーザから取得した著作物が含まれている場合における表現上の本質的特徴が直接感得できないものに情報提供の内容を変更することや約款によりユーザから許諾を取得すること等の対応の在り方。
- ☞ 震災時等において、自動的に、様々な者からデータがアップされサーバ間で連携・加工されることにより、必要な情報が関係者に送信される公共的なプラットフォームや体制等の在り方。
- ☞ 民間企業において収集等されたデータを平時から公開等により流通・連携する場合におけるデータ形式の標準化やメタデータの定義化、先行者利益に配慮したビジネスモデルの在り方。
- ☞ 国民IDの将来的な民間利用等のIDによる個人に関する情報の紐付けの在り方。
- ☞ 各分野におけるルールや情報リテラシー等の相違等を踏まえた情報共有のための入力や表示等の在り方。

その他のデータの活用における課題(例)

- ☞ 海外事例や海外企業による日本国内でのサービス提供状況を踏まえた、利用者との明示的な契約の締結や統計的な匿名化処理等による、災害対応等の公共目的のための通信ログの利用の在り方。
- ☞ 通信サービスの提供にあたっての位置登録情報等の運用データの非識別化処理等の技術やまちづくり等の公共分野における活用の在り方。
- ☞ EUからのデータ移転や国内における医療分野等における機微情報の取扱い等、個人に関する情報の取扱いの在り方。
- ☞ EUのデータ保護指令の改正やアメリカにおけるDo Not Track等の全体的な国際動向を踏まえた対応の在り方。
- ☞ 例えば、プライバシーポリシーの統合によるサービス提供や児童からの同意取得の場合や統計的利用のための完全匿名化等、個人情報保護法上の「個人情報」に該当する場合における利用目的の範囲内での利用(同法第16条)への対応の在り方。
- ☞ 個人識別性がない情報が転々流通するうちに個人識別性を有する可能性やプライバシーを侵害する可能性があるライブログ活用サービス等の提供にあたっての透明性の確保や利用者関与の機会の提供等の配慮の在り方。
- ☞ データ内容の暗号化とIDの取扱いや、k-匿名化に関する技術的な研究等の個人に関するデータに関する秘匿化の在り方。
- ☞ 誤った情報が利用者へ送信される場合等における情報の受信端末等の更新や、契約における責任減免条項の規定等、利用者における判断と提供側の責任等の責任分界の在り方。
- ☞ データの正確性や観測器の精度等が制度上求められている気象データ等をインターネット上で公開等することにより他分野や他目的で活用する場合におけるデータの正確性の確保等の在り方。

ビッグデータの活用に関する課題④

その他のデータの活用における課題(例)[続き]

- ☞ リアルなサービスにおける各種ログの収集や、多様なデータを組み合わせて活用できる仕組みの在り方。
- ☞ 災害関連や人の流れ等の多様な情報について、個人情報保護やプライバシー、著作権に配慮しつつ、まずは情報を使えるようにするための提供フォーマットやインターフェース等、政府や地方自治体が保有するデータの一元的な公開の在り方。
- ☞ 公共データにおけるデータの所在の明確化やデータ取得時の形式等のプロトコルの在り方。
- ☞ Hadoop等のオープンソースソフトウェアに関する運用コミュニティが買収された場合やサポート主体の不在への対応の在り方。
- ☞ ホストコンピュータにあげる前処理におけるメタデータ化に関するフォーマット変換等、今後のよりリアルタイムなデータ処理等における高度かつ新たな処理基盤と分析技術の在り方。
- ☞ 現在の利用目的毎かつ様々なデータ形式によるアプリケーションセントリックなデータの活用とは異なり、今後のM2M等の様々なデータが活用される場合において個々のアプリケーションのベースとなるセキュリティやプライバシー等の水平部分の標準化等の在り方
- ☞ 収集等したデータの帰属先や、当該データに関する第三者によるフリーライドや流用への対応の在り方。
- ☞ 多大な投下資本が必要となる一方で、情報の選択又は体系的な構成による創作性を有しないために著作物として保護されず、また、損害賠償請求の対象にはなる場合があるが、差止請求の対象にはならないファクトデータベースに関する欧米等の諸外国の状況も踏まえた保護の在り方。
- ☞ 営業秘密としての管理と、守秘義務契約による個別企業への提供や第三者への提供の関係等、著作物として保護されないデータ解析結果に関する取扱いの在り方。
- ☞ 日本の大学に統計学部や学科が存在せず、工学部・経済学部・医学部等の各学部で統計関連のスタッフが在籍し、横のつながりが無い状況等における非構造化データ等の様々なデータの活用に関する数学的・統計学的・法学的な知識やビジネス管理における知識等を有する人材等の確保・育成の在り方。

その他のデータの活用における課題(例)[続き]

- ☞ 行政機関において民間から購入する等活用し得るデータの在り方。
- ☞ データの誤用・不適切利用におけるアルゴリズムの賢さ・間違いの取扱いの在り方。
- ☞ 量や多様性の中から価値ある情報を見つけるための分析ツール、データ分析のための継続的なノウハウの蓄積、そして、適切なデータの取得のための工夫を実施するためのビジネス等への総合的な理解と深い分析技術による仮説構築・検証・データ取得方法等に関するPDCAサイクルの継続的な検証の在り方。
- ☞ 国内におけるスモールサクセス等の成功体験や海外事例の共有等、データ活用の重要性に関する意識向上の在り方。
- ☞ 自社内で膨大にたまっているデータを新規事業に役立てることができる確率を高めるための他業種等との連携の在り方。
- ☞ 例えば、社会インフラや医療分野等、新産業の創出やビジネス領域の拡大をもたらす分野への投資を促進することによるマーケットセグメントを跨がるデータの融合を促すための仕組みの在り方。
- ☞ フィジカルとの融合等、GoogleやFacebook等が現在主に取り組んでいない分野への取組の在り方。
- ☞ 個人情報の活用におけるベネフィットと失われるものとのバランスの問題に関する実証の場等、国民の不安感の払拭やコンセンサス等の醸成の在り方。
- ☞ 実社会において実際に収集したデータを用いた分析結果や実証事業をアクション等するためのオープンイノベーションに向けた場の在り方。
- ☞ 海外プロジェクトにおける技術実証の進展も踏まえた、国内における技術の実装のためのソーシャルな利益をもたらすビッグデータの活用に関する社会実験等の在り方。
- ☞ スマートフォンの普及やテロ対策としてのソーシャルモニタリング等により増加するウェブにおける多言語化への統計的自然言語処理等による多言語解析の在り方。

3. ビッグデータの活用に向けた方向性と具体的方策