

異分野・社会との連携のための共通言語「データサイエンス」の 学際的な研究・教育拠点の形成

① ビジョンの概要

統計検定、統計エキスパート人材育成コンソーシアム等が運用され、滋賀大学など複数の大学でデータサイエンス系学部が設立された。これらの人材育成を加速させ、データサイエンスの統計数理基盤を担うべく、2つを促進する。(1)学際的な研究・教育拠点の形成とシンクタンクの構成。(2)科学行政官、チーフ・サイエンス・オフィサーの育成。

② ビジョンの内容

(1)学際的な研究・教育拠点の形成とシンクタンクの構成

科学が細分化された現在、分野間を越境できるデータサイエンティストの育成、そして、データサイエンスを教える教員の育成は鍵となる。学際的な研究・教育拠点を拡大・整備して、米国国立科学財団等と連携し、地球規模、緊急性の高い課題を解決に導くためのプラットフォームとする。英国オックスフォード大学に設立されるシュワルツマン人文学研究センターとも連携し、AI倫理について法律学や公共政策学などの社会科学・人文学の研究者と共同研究を促進する。データサイエンティストが集うシンクタンクを構成し、教育・学術・科学技術の長期的展望を描き、新規政策を打ち出し、その意義を社会に正しく伝える役割を果たす。

(2)科学行政官、チーフ・サイエンス・オフィサーの育成

データサイエンスによる分析能力と批判的思考力は、大学院における知の生産を意識した教育によって身につく。現状、文部科学省は博士号をもつ職員が少なく、企業の研究者に占める博士号取得者の割合も米国と比べ日本の低さが目立つ。日本にはイノベーションを俯瞰し未来への展望を構想できる人材の育成が不可欠である。科学行政官にはデータサイエンスを心得た目利き力と、博士号をもち分野横断的な研究をプロデュースし新たな価値を創造する構想力が求められる。科学行政官が育成されれば、政策形成の場においてデータサイエンスによるビジョンが提供できるだろう。データサイエンスが諸科学・産業の基盤として広範に利用されるように、信頼性や質保証に関わる技術的な基準や利用上の倫理的・法的・社会的な基準が整備されるだろう。また、民間企業には、チーフ・サイエンス・オフィサーとして活躍するデータサイエンティストを育成する。データサイエンスを駆使した意思決定をすることで、価値の創造と共鳴が生まれるだろう。

③ 学術研究構想の名称

異分野・社会との連携のための共通言語「データサイエンス」の学際的な研究・教育拠点の形成

④ 学術研究構想の概要

(1)データサイエンスの主要技術の一つは統計数理である。本構想は、「統計エキスパート人材育成コンソーシアム」（統計数理研究所、複数の大学で構成）を拡大・整備し、データサイエンスの学際的な研究・教育拠点を形成する。データサイエンスの理論や方法を開発する研究者と応用に関わる異分野や産学官の研究者・実務者が参加する学際的な拠点とし、国際連携推進、及び地球規模、緊急性の高い課題解決のための基盤とする。「統計検定」（日本統計学会で開発）などを活用し、データサイエンティストや教員を育成し、データサイエンティストが集うシンクタンクを構成する。データサイエンスが諸科学・産業の基盤として広範に利用されるように、信頼性や質保証の技術的な基準、倫理的・法的・社会的な基準を研究開発する。

(2)データサイエンスを心得た目利き力を有し、イノベーションを俯瞰し未来への展望を構想できる科学行政官及び民間のチーフ・サイエンス・オフィサーを養成する。科学行政官は新たな価値を創造し、政策形成の場においてデータサイエンスの教育・学術・科学技術政策の意義を社会に正しく伝えることが期待される。

⑤ 学術的な意義

データは21世紀の石油とも呼ばれ、GAFAなどのプラットフォーム企業はデータを資源としたビジネスをおこなっている。また学術研究においても大量のデータに基づくデータ駆動型の研究の重要性が増している。このようなデータ革命の中で統計科学、さらにその基礎をなす数理統計学の重要性も増している。

ビッグデータが得られるようになったことにより、深層学習の手法など複雑な予測モデルを構築することが可能となった。一方で、ビッグデータは、しばしばバイアスを含むことが指摘されており、ビッグデータにそのまま予測モデルを当てはめても、予測モデル自体がデータのバイアスを反映したものとなる。正しい

分析をおこなうためには、データの生成過程を適切に反映したモデルを構築することが必要である。統計科学の方法論では、実験計画法や標本調査法によるデータの生成、データ生成モデルの定式化、モデルを推定した後の信頼性の評価、などの一連の過程を重視している。科学のあらゆる分野でデータ駆動型研究が進められる中で、データ分析の信頼性を確保する統計科学の役割は大きい。

データ革命の中で応用技術は企業研究者によって行われている現状がある。日本では、博士号取得者数が少ないのみならず、就職先はアカデミックである場合が多い。特に統計学の教員不足が顕在化している。日本の課題を解決するには、企業と大学の間の人的交流を含め、データサイエンスを横串とする分野融合的な研究推進体制が必要である。博士課程においては、データサイエンスを心得た目利き力を持つ科学行政官を養成することが不可欠である。企業においてもチーフ・サイエンス・オフィサーとして活躍が期待できる。

⑥ 国内外の研究動向と当該構想の位置付け

2012年、オバマ政権がビッグデータイニシアティブを発表し、世界にデータサイエンスの重要性を強く印象付けた。以来データサイエンティストの育成が加速し、我が国も「数理・データサイエンス教育強化拠点コンソーシアム」が実現し、統計検定制度の運用、データサイエンス系学部の設立と、高等教育での人材育成が始まっている。しかし、企業のニーズは遥かに高く、世界基準で未来を展望できる人材が決定的に不足している。本構想では今後10年の間に、これらの人材育成を加速させる国家戦略を提示する。

⑦ 社会的価値

1. データサイエンスの学際的な研究・教育拠点の拡大・整備
2. データサイエンスの科学行政官及びチーフ・サイエンス・オフィサーの養成
1. では「統計エキスパート人材育成コンソーシアム」を拡大・整備し、データサイエンスの理論と応用に関わる異分野、産学官の研究者や実務者の学際的な拠点を形成する。
2. では、我が国に決定的に不足している、データサイエンスの学位をもってイノベーションを俯瞰し未来への展望を構想できる科学行政官及びチーフ・サイエンス・オフィサーを養成する。

⑧ 実施計画等について

本構想は次の4つの段階を踏んで実施する。

第1段階

統計エキスパート人材育成コンソーシアムを研究組織に発展させ、テーマごとに研究チームを作る。これに企業・官庁の実務家を含む外部参加者を交え、学会・研究会よりも深く議論できる環境を構築する。

第2段階

上の各テーマに対応する大学院プログラムを設定し、実務家からも広く大学院生を募集する。外部組織からも最適な指導者を招聘し、国内最高水準の指導体制を構築する。

第3段階

研究成果の社会還元のために、大学教育の枠をこえたアウトリーチ活動、統計検定を発展させたオンデマンド教材やオンライン教材の開発、資格の認証などを行う。

第4段階

大学院生のインターン活動や実務家の学生・研究者としての受け入れ、行政・産業界との交流を活性化し、実務界がデータとアイデアと目的意識を研究者と共有できる体制を構築する。これにより、テーマの発見から、研究開発、アプリやビジネスモデルの開発と試行までを一貫して実施できる体制を提供する。

本構想は、統計エキスパート人材育成コンソーシアムの加盟機関を実施機関とし、統計数理研究所が取り組み機関となり、情報・システム研究機構傘下の国立情報学研究所・データサイエンス共同利用基盤施設、SIP 分野間データ連携基盤と連携する。

教員5名、研究員15名、事務員2人、プロジェクトマネージャ1名の給与：2億2,000万円

教員・研究員に対する研究費：5,000万円

建物・計算機・ネットワーク賃料：7,000万円

総額 3億4,000万円/年、10年間で34億円

⑨ 連絡先

青嶋 誠（筑波大学）