

価値の高い情報を効果的に解析するためのデータアルゴリズム開発

東北大学 総合知インフォマティクス研究センター

生成情報の消失速度が急速に加速する中、知の創造に有効な高価値情報の抽出は喫緊の課題である。その解決には従来人間が評価している目的に合致した情報や普遍性の高い情報をコンピュータに置き換える科学技術が必須である。研究者により人間の情報価値評価の理解を実現することで、目的による学習可能なシステムを構築し、情報生成とその価値を評価するシステムへの展開を目指す。人文社会系の研究者と情報科学系の研究者をつなぐ総合知的な組織を設置し、総合知的な視点を持つ研究者によるサポート体制を確立している。

総合知により目指すビジョン / 解決する社会課題

総合知に基づく多様なデータの価値評価と保存優先付けが可能となり、社会全体の計算・伝送・蓄積コストの削減に貢献した。Society 5.0の実現と新たな価値創造にも寄与する。

ビジョン達成のための課題

情報の価値は利用者と目的により異なり、人間による評価と信頼性の考慮が必要だが、現代の巨大データ環境では限界がある。そのため、人間の価値評価を可能にし、個別対応と学習機能を備えたシステム構築が必要である。

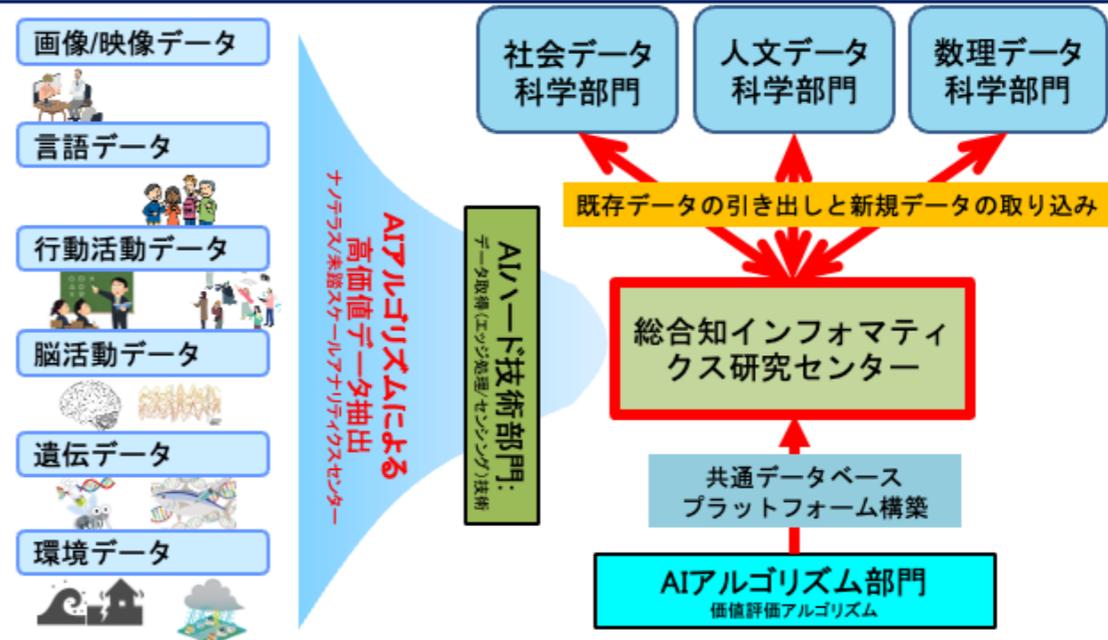
「矩」を超えた場づくり / 得られた新たな価値

多様な目的に応じた情報価値の実現に向け、人文社会科学を含む総合知による価値基準の設定を推進する。東北大学の16部局と国内外の機関が連携し、文理融合の5部門体制でSociety5.0に貢献。2024年度は17件の学際研究を助成している。

総合知を利用した高価値情報創出システムの開発と実装



東北大学16部局による知見に対し、ヒトをモデルとする価値化を行い、そのアルゴリズムに基づき、新規データの価値づけ（取捨選択）を自動的に行う高価値創造システムを構築する



AI So-Go-Chi

Tohoku University Advanced Institute of So-Go-Chi (Convergence Knowledge) Informatics

高価値評価に基づくデータの有効活用

社会データ
科学部門

人文データ
科学部門

数理データ
科学部門

総合知データ分析支援室：各分野のデータ分析支援・人材育成

AIアルゴリズム部門： 価値評価アルゴリズム

AIハード技術部門： データ取得（エッジ処理/センシング）技術

ヨタ (10²⁴) スケール
データ生成

生成情報量の加速度的増大
多くは保存できずに生成直後に消滅



多様多様の大量データの生成

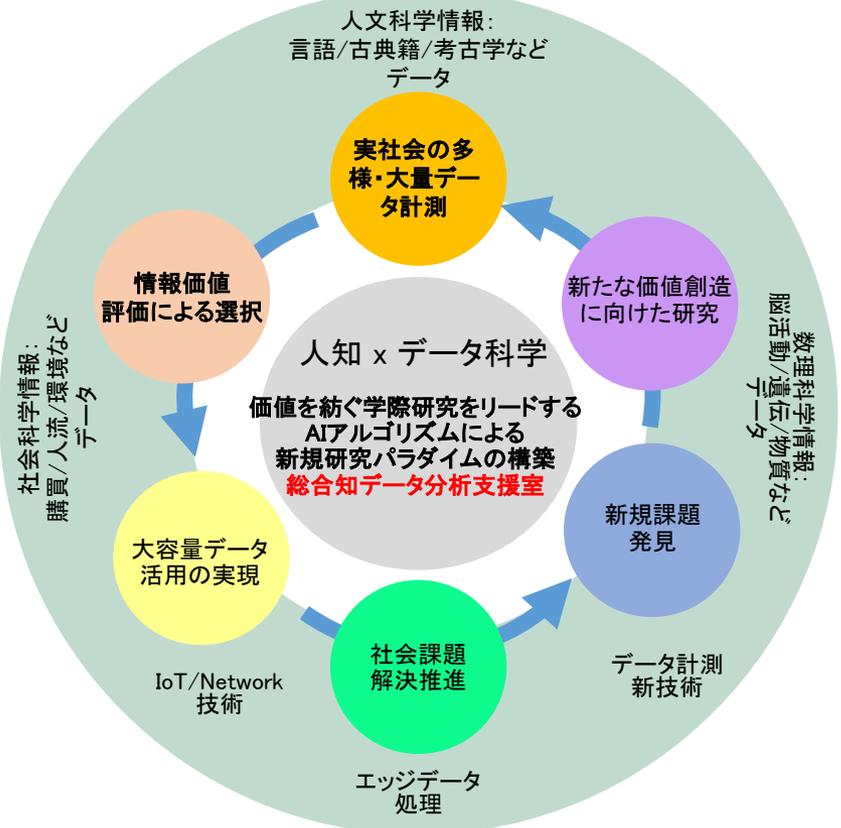
ナノテラス/未踏スケールアナリティクスセンターと連携

購買データ
人流データ
環境データなど

言語データ
古典籍データ
考古学データなど

脳活動データ
遺伝データ
物質データなど

価値評価/データ利用/価値創出サイクル



異なる分野のデータを共通の枠組みで捉え、価値の高い情報を効率的/効果的に解析するためのデータ科学的手法(アルゴリズム)を開発し、システム(AIハード)に実装することで多様な分野における価値創出を促進する。

AI So-Go-Chi

Tohoku University Advanced Institute of So-Go-Chi (Convergence Knowledge) Informatics

ヨタ(10²⁴)スケールに近づく生成情報を有効利用するためには、人間が評価している目的に合致した情報あるいは普遍性の高い情報の選択の、計算機による自動化が必須である。そのために、人文科学、社会科学、数理(自然)科学の分野のデータに対して、利用者や目的のために有効な情報の価値評価方法を開発する。また、情報生成と同時にその価値を評価するシステムへの展開を目指すとともに、情報アーキビストなど情報科学知識を持つ人文社会系を含む多様な分野の専門家育成を行う。

新学術領域

画像/映像データ



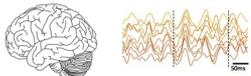
言語データ



行動活動データ



脳活動データ



遺伝データ



環境データ

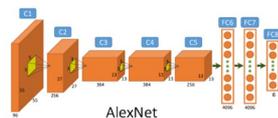


AIアルゴリズムによる
高価値データ抽出
ナノクラス/米踏スケールナノクラスセンター

プラットフォーム構築
データベース

AIハード技術

データ科学



データ解析
深層学習

情報価値評価
高価値情報活用

人間の評価
直感/分析

人間の価値判断



期待される成果事例:

1) 健康的食生活の維持

- 体調に合わせた適切な食情報の提供
- ・ 信頼できる情報に基づく科学的根拠を参照した選択肢の提示
- ・ 個人の生活記録から個人に合わせた食環境の自動推定
- ・ 顔表情、生体情報などエッジ評価による体調、嗜好に合わせた選択肢提供

2) 学習における主体性確保

- ・ 収録顔画像情報から受講者の興味/集中度/理解度の自動推定
- ・ 教育学的視点による評価付与
- ・ 受講者のビデオ画像を分析・情報を集約

3) 古典籍データの利活用

- ・ くずし字のテキスト化
- ・ 文書作成の経緯、作成後の情勢変化等の情報付与
- ・ 紙の原材料や付着物に係る報付与

4) 言語消失による損失軽減

- ・ 多言語共通の音声特徴の同定
- ・ 非言語情報の解析
- ・ 映像分析による環境情報取得

新たな価値

新たな視点

新たな課題

想定外
矛盾
不一致

課題解決