

文部科学省
「多メディアWeb解析基盤の構築及び
社会分析ソフトウェアの開発」
の活動進捗状況

平成21年12月1日

文部科学省 研究振興局 情報課

「情報の巨大集積化と利活用基盤技術開発」連携施策群における位置づけ

目標

Web上及び非Web上にある大量かつ多様な情報を、個人が簡便、的確、かつ安心して収集、分析できる

次世代の情報検索・情報解析技術基盤を構築する。

2012年までに、情報爆発時代に対応した超巨大データの戦略的活用を可能とする。

非Web空間(実空間)

Web(サイバー空間)

[2009年度からの連携枠組み]

必要性

- Web情報の価値の拡大
- Web情報の学術研究、経済活動への高度活用に対する期待

- Web情報分析のための技術の発展が必要

情報の活用

ビジネス、生活、教育・研究、医療等

Web社会分析プロジェクト

- Web情報(テキスト、動画、画像、音声等)を効率よく収集してアーカイブを形成
- 蓄積したWeb情報の高度な分析

過去の情報を含めた高度な分析

新たな分析手段

情報の分析

分析手段

連携を拡充

情報の信憑性を検証

[2008年度までの連携枠組み]

情報大航海プロジェクト

- 巨大集積化した情報を再利用できるように仕分け
- 整理整頓して器に格納

情報大航海の一部成果も活用し、研究開発を効果的に推進

センサ情報の社会利用コンテンツ化プロジェクト

- プライバシー情報を管理
- センサ情報のコンテンツ化

超高性能データベースプロジェクト

処理の効率を格段に向上

情報信憑性検証技術プロジェクト

- 格納された情報の純度を上げるべく精錬
- 純度の保証付きでユーザーに提供

多メディアWeb解析基盤の構築及び社会分析ソフトウェアの開発

目的

社会学、言語学、リスク管理、マーケティング等多様な社会分析ニーズに応じるために、膨大な多メディアWeb情報を収集、蓄積し多様な解析を可能とする多メディアWeb情報解析基盤の構築と社会分析ソフトウェアの研究開発並びに実証を行う

課題

爆発的な増加を続けるWeb情報は、画像・映像等への多メディア化が急速に進むと同時に、放送映像等の実世界情報と相互に及ぼし合う影響も拡大し続けており、人類社会の観測・調査・解析から新たな価値を創出するために必要不可欠な情報源

多メディアWeb情報の収集・蓄積、多メディア内容解析、高並列計算環境上での大容量・高スループット解析基盤、ならびにそれらに基づく有効な社会分析ソフトウェアの実現が必須



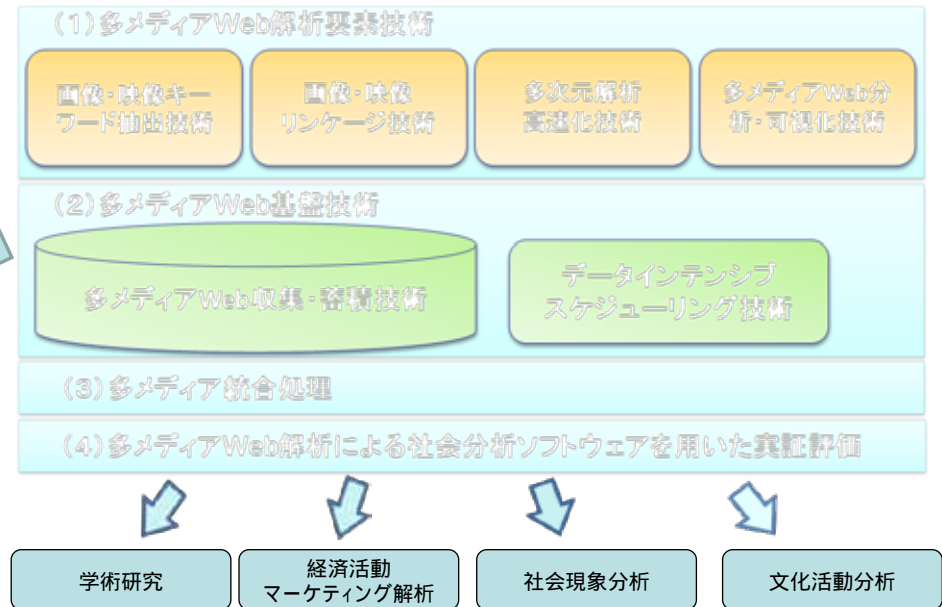
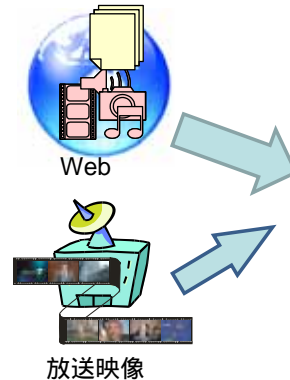
多様なWeb上で観測される社会現象
 (例: イラク大統領選挙への抗議活動)

Webによる社会現象の追跡調査
 (例: 「社会起業」の広がり)

研究開発内容・目標

研究開発内容

多メディアWeb解析基盤を構築し、それに基づく社会分析ソフトウェアの開発を行うため、(1)～(4)のサブテーマに基づき研究開発を実施する。



成果目標

- 平成22年度中には、多メディアWeb解析要素技術ならびに基盤技術の基本設計を終了し、社会分析ソフトウェアのプロトタイプ実装を行う。
- 平成24年度中には、多メディアWeb解析要素技術ならびに基盤技術の実用規模での実装を行い、多メディア統合処理プラットフォーム上で統合し、社会分析ソフトウェアの実証実験を行う。

多メディアWeb解析基盤ならびに社会分析ソフトウェアとは

特徴

- 多メディアWebの大規模・時系列的な収集・蓄積による社会分析の情報源確保
- 多メディアWeb解析による、特定の話題の出現の検出、話題の発生・発展・反響等の時系列的追跡(多メディアWeb情報のリーズニング解析)
- 解析結果ならびにその可視化等による社会分析ソフトウェアの実現



Webに投稿された画像から、テレビ報道や新聞報道などに話題が広がっていった例(ハドソン川での飛行機事故)

期待される効果

- 社会学、言語学、経営学等、多様な学術分野におけるこれまでにない新たな学術調査手法の創出が期待できる。
- また、これら調査により、従来手法では十分把握できなくなった国民生活の実態を捉えることが可能となり、有効な政策の立案や新規ビジネスの創出につながる。これにより業界の活性化や健全化がもたらされると期待できる。
- 多メディアWeb情報を用いたリーズニング解析により、テレビ番組やWeb上の動画を通じて企業やコンテンツ制作者から発信された情報がどのように視聴者・消費者に影響を与えたかを詳細に解析することが可能となり、高度なマーケティング手法、エンタープライズリスク管理手法等が現出する。
- 多メディア情報解析、大規模Web情報の時系列的収集・蓄積、高並列計算環境を用いた大規模解析処理、ならびにそれらに基づく社会分析ソフトウェアの実現を一貫して研究開発する事例は国内外において他に例を見ず、独自性が高く、国際競争力の向上が期待できる。

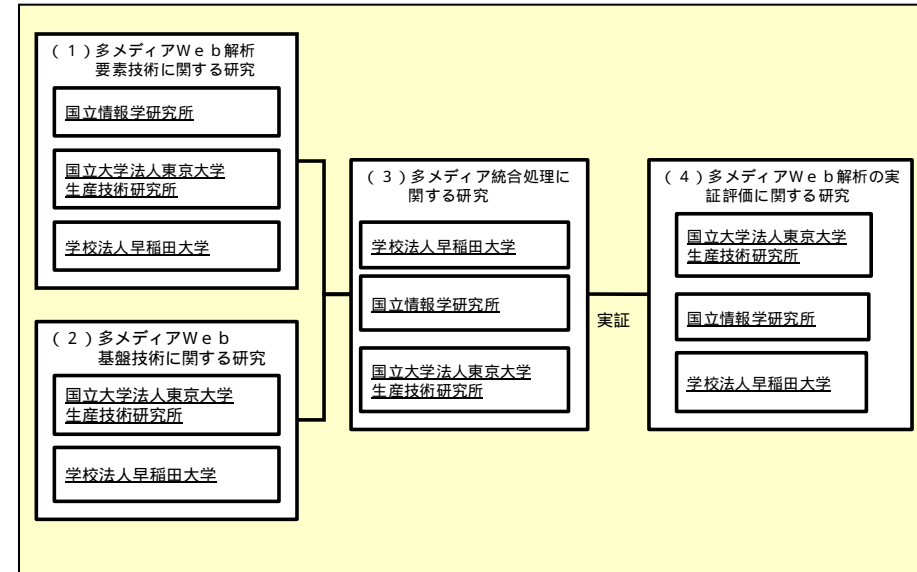
研究開発体制

実施体制

国立情報学研究所、東京大学、早稲田大学の三機関を研究拠点として実施する

研究代表者：佐藤真一 教授
(国立情報学研究所)

西尾章治郎教授(大阪大学 / 本事業プログラムオフィサー)を交えた進捗評価委員会を立ち上げる予定



Webアーカイブ構築活用助言委員会

産業、法曹、学術各分野より、有識者を迎えたWebアーカイブ構築活用助言委員会を設置し、成果を用いた学術調査への利用や実サービス実施に対する要請、法制度的問題点等に関する議論を行いつつ研究開発の方針を定めていく予定

構成員

長尾真 (国立国会図書館・館長)
 牧野二郎 (牧野総合法律事務所・弁護士)
 小島啓二 (株式会社日立製作所中央研究所・所長)
 有川節夫 (九州大学・総長)
 Calton Pu (ジョージア工科大学・教授)
 Nozha Boujemaa (INRIA Rocquencourt・Director of research)

研究開発計画

