

研究テーマ: 情報コンテンツの信頼性評価プロジェクト(1/2)

(プロジェクト番号 JGN2P-A20002)

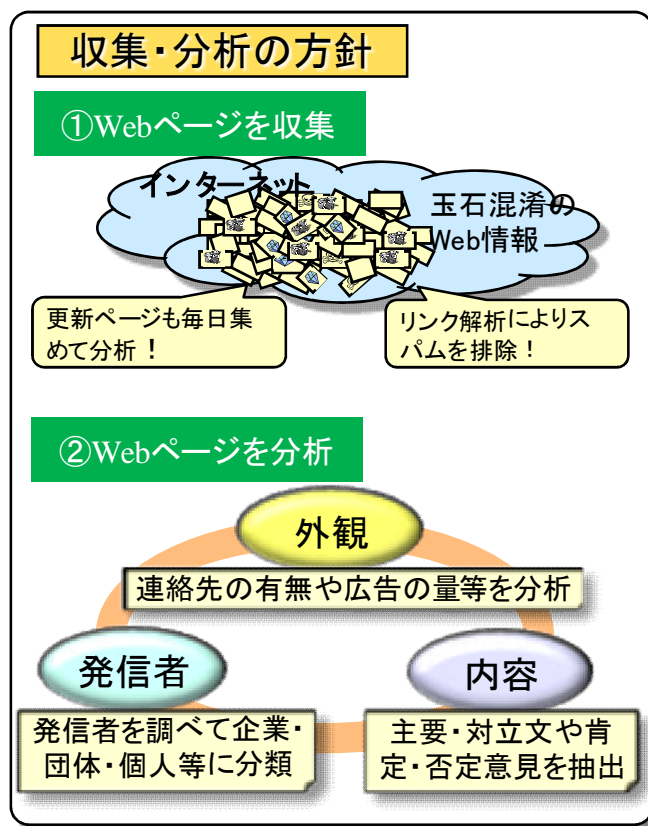
研究機関: 情報通信研究機構知識創成コミュニケーション研究センター
知識処理グループ

研究の概要:

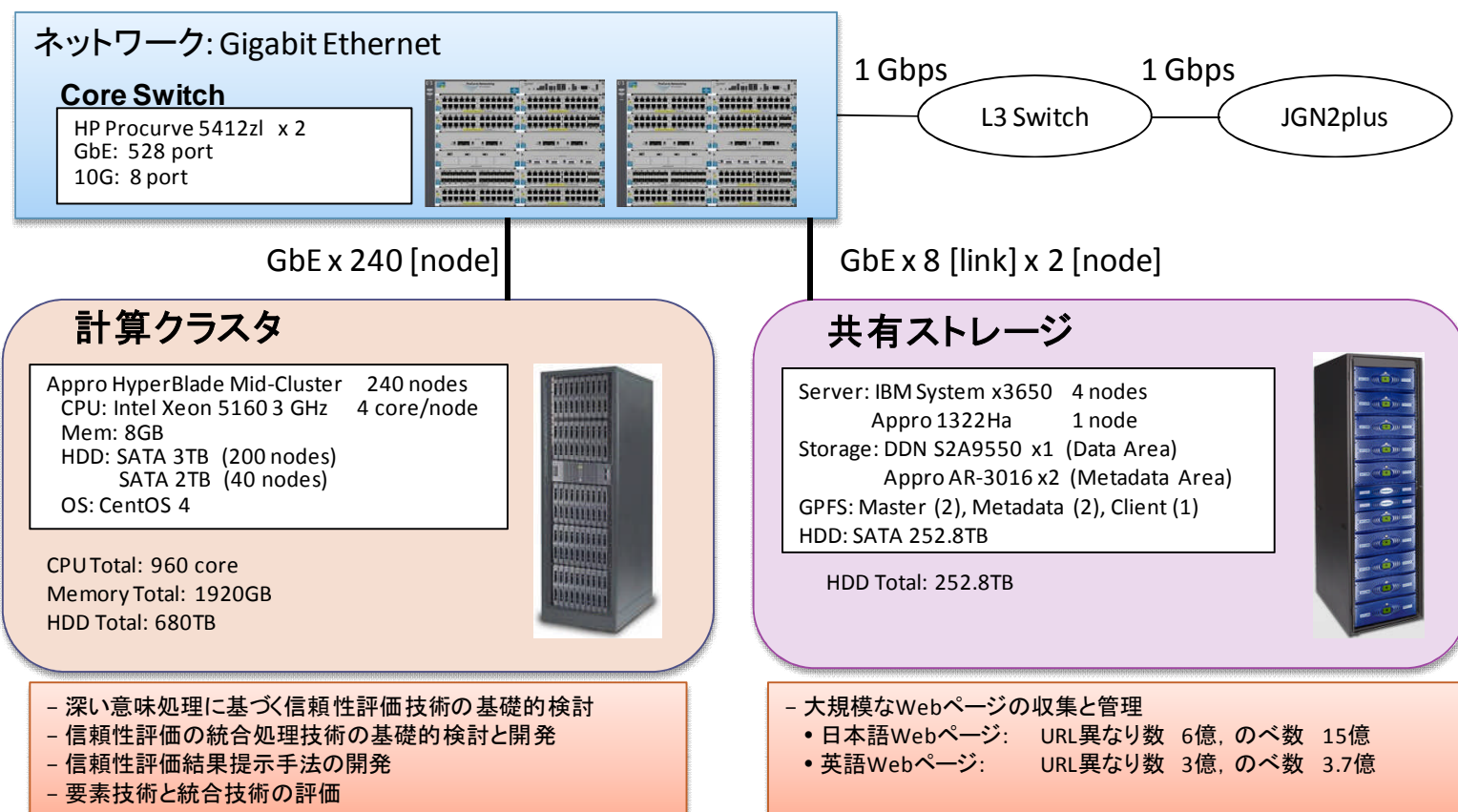
- 多種多様なWebページから、「誰が」「何を」「どのように述べている」のかを自動抽出して、多面的に分析システムを開発.

研究の目的:

- インターネットを安心して利活用するために、膨大なWeb情報から、信頼でき、役立つ情報を見つけることを支援.



実験機器構成:



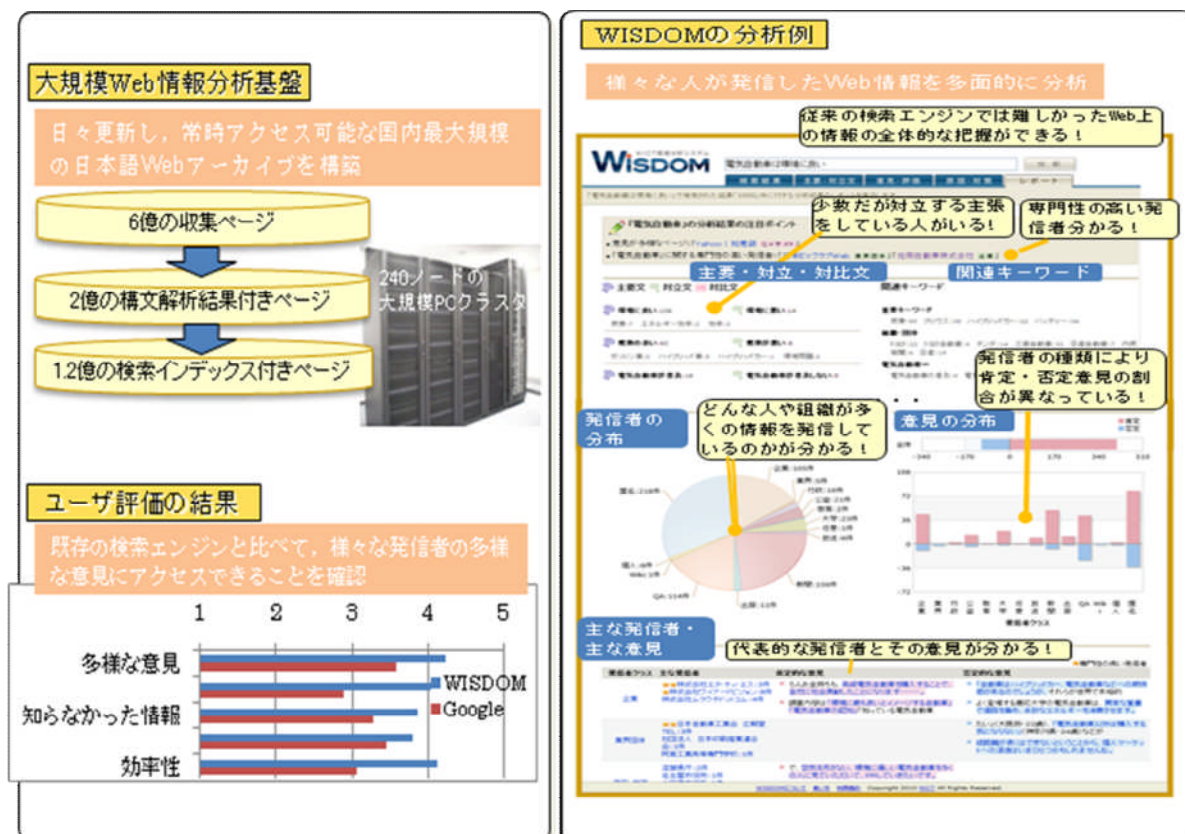
研究テーマ: 情報コンテンツの信頼性評価プロジェクト(2/2)

(プロジェクト番号 JGN2P-A20002)

研究機関: 情報通信研究機構知識創成コミュニケーション研究センター
知識処理グループ

研究開発成果:

- 大規模Webページ収集・検索基盤の開発と運用
 - URLの異なる日本語Webページを6億件, のべ数で15億件を収集
 - 比較的質の高いページを選別し, 言語解析済みの1.2億件が常時検索・分析対象に
- 情報分析システムWISDOM
 - 大規模Webページ収集・検索基盤に登録されたWebページを対象に, 任意の話題についてWeb上の情報を分析することが可能
 - 発信者分析, 評価表現抽出, 主要・対立・対比文抽出などの機能を備え, それらの結果をレポートしてユーザに提示



プロジェクトのアピールポイント

- Toy Levelにとどまらない10億ページ規模で日本語Webページを収集・検索・分析するためのWeb情報分析基盤を構築
- Web情報を様々な観点から分析した上で, 整理し組織化して表示するための各種要素技術を開発
- 要素技術を統合したWeb情報分析システムWISDOMを開発
 - 報道発表をし, メディアやWebで多数の反響
 - ユーザ評価で, 従来の検索エンジンと比べて優れているという結果
 - 一般向けに公開中 <http://wisdom-nict.jp/>

プロジェクトの自己評価

- 研究計画に基づき, 着実に成果を出すことができた
- 今後の課題: 多言語化(英語, 中国語), 分析精度の向上, 分析対象の規模拡大