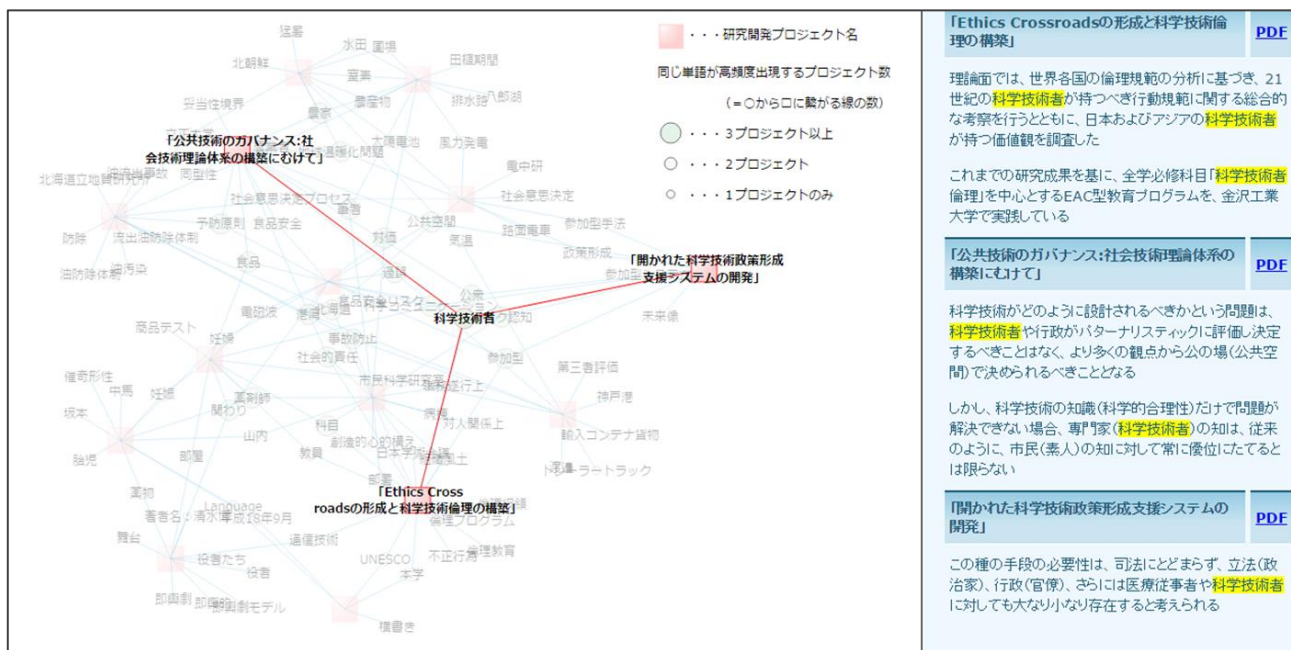


2015年3月31日  
株式会社アイズファクトリー

日本語テキストデータに挑む

大量のテキストデータを、関連づけて活用しやすく可視化  
必要なノウハウや興味のある文書がスムーズに見つかる！  
ドキュメント管理やQ&A サイト運営にも朗報

独立行政法人科学技術振興機構 社会技術研究開発センター（東京都千代田区、センター長：泉紳一郎、以下「RISTEX」）実施の『論文データ可視化』において、株式会社アイズファクトリー（東京都千代田区、代表取締役：大場智康、以下「アイズファクトリー」）の日本語テキストマイニング技術、『複合語解析』・『特徴度』・『共起スコア』を利用した Web での可視化サービスが開始されました。アイズファクトリーは、RISTEX で保有する過去の研究開発実施終了報告書を、組織内外の目的に合わせて活用しやすく開示するための研究、およびシステム開発に協力いたしました。



社会技術研究開発センターホームページより引用

「社会システム／社会技術論」研究開発領域の各プロジェクトで出現する単語と、該当する報告書『複合語解析』・『特徴度』・『共起スコア』の技術は、社内にある文書の関係把握や、データ活用、部門間で業務ノウハウをスムーズに共有するためのドキュメント管理にも活用できます。

国や地方自治体が保有する公共データを中心としたオープンデータの公開が始まり、企業活動や福祉に役立てることも期待されます。

本技術は、日本語テキストデータの利活用の際に必要な、データの検索利便性を高めるためのメタデータ<sup>\*1</sup>作成にも有効です。

## ■ 社会技術研究開発センターホームページ

<http://www.ristex.jp/result/network/index.html>

## ■ 活用例

- ・ 企業内の営業活動データ・業務日報・各種報告書等のテキストデータを、各部門の担当者が業務ニーズに合わせて検索もしくは探索する
- ・ Q&A サイト上の話題を集約・可視化し、ユーザーが用途に合わせた検索を容易に行う

## ■ アイズファクトリーの日本語テキストマイニング技術について

アイズファクトリー独自の『複合語解析』により、形態素（言語で意味を持つ最小単位）同士を結合し意味のある単位でのキーワード抽出が可能です。『複合語解析』によって、より人間の感覚に沿った、より理解しやすい分析結果を得ることができます。

例)「東京スカイツリー」は一塊で意味を持つ言葉だが、形態素解析を行うと、「東京」「スカイ」「ツリー」として認識される。『複合語解析技術』によって、「東京スカイツリー」を1つの単語として抽出できる。

また、文書間・単語間の関係性の強さについては、『特徴度』という指標が使われます。

『特徴度』とは、各文書内において、他の領域または他の文書と比較した時に現れる頻度の割合のことで、これにより各領域間における関連度合い・単語間の関係性の強さを定量的に表せます。

## ■ 概要

論文データ可視化には、アイズファクトリーの日本語テキストマイニングの基礎技術『複合語解析』・『特徴度』・『共起スコア』が用いられています。

研究開発実施終了報告書データの可視化にあたって、組織内外の活用目的を鑑み、可視化で実現すべきことを下記の通り整理しました。

要件を的確に掴み、設計に落とし込むことで、活用しやすいシステムを実現しました。

1. これまでの研究開発プロジェクトにおける成果の把握  
→各領域におけるプロジェクト成果を、可視化により把握する
2. 研究者が直面している、課題の可視化  
→将来研究すべき研究開発領域・プログラムを把握し、将来に向けた検討・選択を最適化する
3. 各研究開発領域における、関連可視化  
→これまでの研究成果を一貫性を持って把握し、社会に対して分かりやすい形で還元する



## ■ RISTEX について

社会の具体的な問題を解決するための研究開発推進のため、下記の活動を行っています。

- I. 研究開発の対象となる、具体的な社会問題の発掘
- II. 問題解決に寄与する、研究開発領域・プログラムの設定
- III. 明確な研究開発目標に向けた、研究開発の推進
- IV. 研究開発成果を社会に還元するためのプロトタイプの呈示
- V. 成果を社会で適用・定着させるための実装支援

社会技術研究開発センターホームページより引用

## ■ テキストマイニング技術『複合語解析』は下記よりご体験頂けます

<http://lab.bodais.jp/tmbssc/index.php>

\*<sup>1</sup>メタデータとは、情報検索システムの検索の対象となるデータを要約したデータのこと。

例えば文書であれば著者名や表題、発表年月日等のほか、関連キーワードなど。

<input type="checkbox"/> 会社概要 株式会社アイズファクトリー 設立：2000年4月25日 代表者：代表取締役 大場智康（理学博士） 資本金：130,971千円	<input type="checkbox"/> 本件に関するお問い合わせ TEL：03-5259-9004 E-mail： <a href="mailto:mrk@isfactory.co.jp">mrk@isfactory.co.jp</a> 担当：マーケティング部 生田 <a href="http://bodais.jp/company/">http://bodais.jp/company/</a>
--	--