



「次世代人工知能・ロボット中核技術開発」
(人工知能分野) 中間成果発表会
— 人間と相互理解できる人工知能に向けて —

AI for 科学技術研究の概要

平成29年 3月29日

国立研究開発法人 産業技術総合研究所

人工知能研究センター

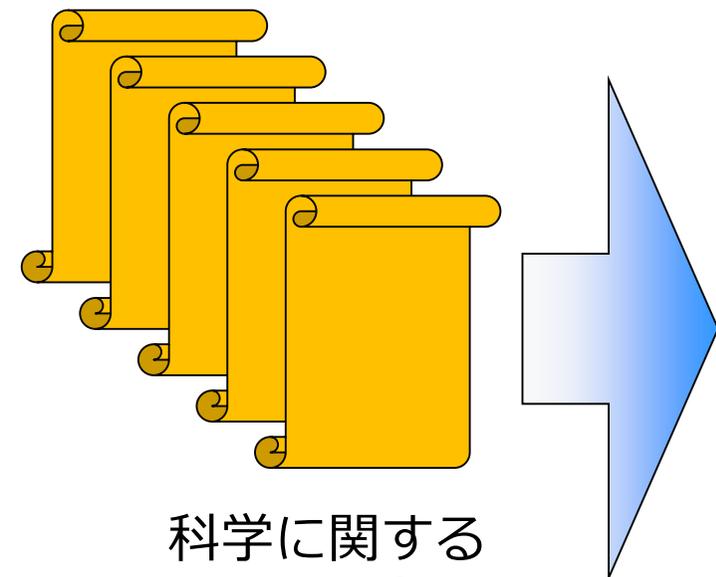
麻生 英樹

国立研究開発法人 産業技術総合研究所

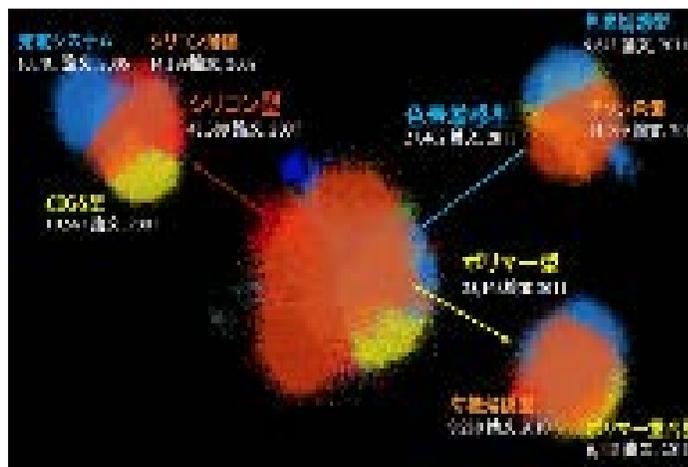
国立研究開発法人 新エネルギー・産業技術総合開発機構

- 人間の知識を、AI 技術を使ってより高度に利活用する
 - 科学技術の研究を加速し、ブレークスルーを支援する
 - 事故の原因究明や予防を支援する
- 科学技術文献の量が膨大になり、すべてを見ることは不可能
- 文献等から構築される知識も複雑になり、人手だけでは構築・メンテナンスが困難
- 科学技術研究を加速する人工知能技術の研究開発
特に、膨大な科学技術に関するテキストに含まれる人間の知識の利活用のための人工知能技術の研究開発を行う
- 技術分野の俯瞰、技術の未来予測、事故予防、酵素反応のデータベース化、など具体的なユースケースに即しながら、共通的な技術の研究開発を進める
- 将来的には、科学実験データなどのテキスト以外のデータとも統合して仮説生成や実験計画なども含めた展開も考えられる

- ネットワーク分析と言語処理の融合による大規模文献データからの技術の未来予測プラットフォームの研究開発
- 事故情報テキスト解析・事故予防技術の研究開発
- 酵素反応データベースに向けた文献キュレーション支援技術の研究開発

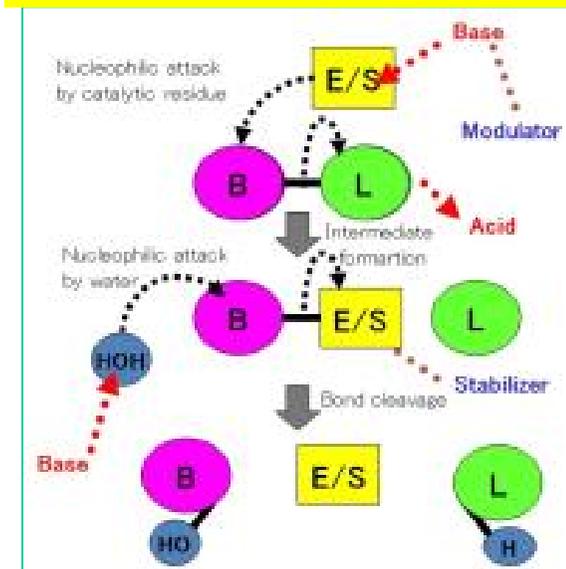


科学に関する
テキスト
(論文、特許、等)



技術分野の動向の俯瞰

酵素反応データベース

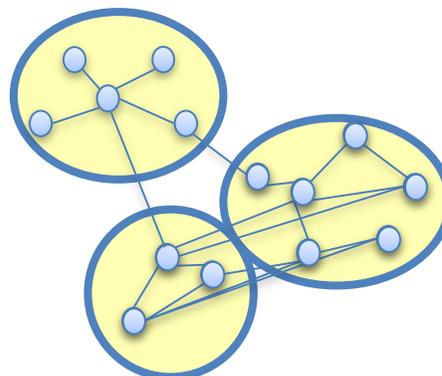




書誌情報
入手



引用ネットワーク
生成

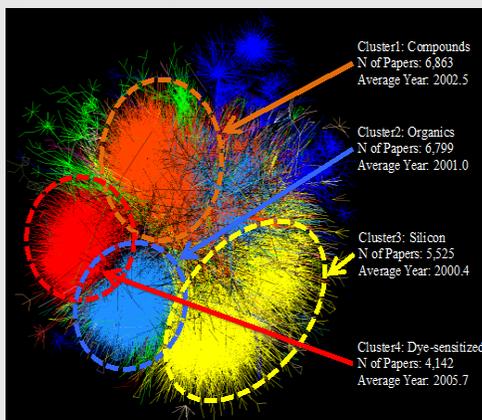


Thomson Reuters
Web of Science



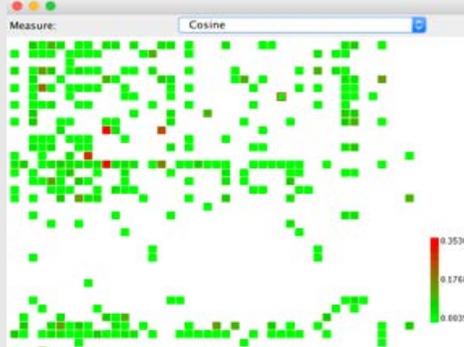
主な機能

④ 俯瞰分析



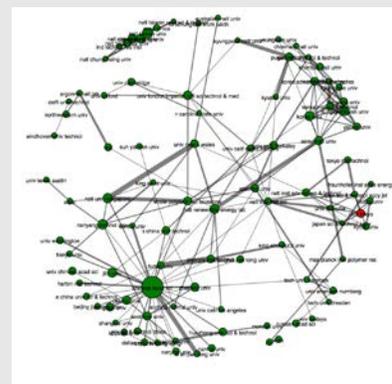
学術領域を俯瞰

⑤ ヒートマップ



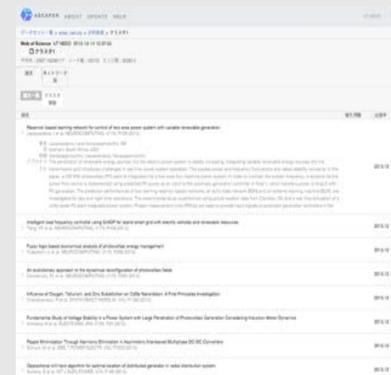
分野間の関係性を理解

⑥ 共著ネットワーク



中心となる組織や人物を特定

⑦ 萌芽論文特定



将来の高引用数文献を予測

酵素反応データベースと 自然言語処理とのコラボ



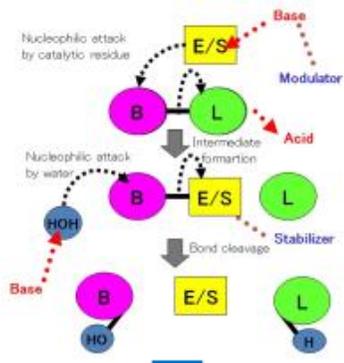
ARCO

英国マンチェスター大学
テキストマイニング・システム

膨大な酵素関連
英語文献情報

情報抽出

複雑な酵素反応機構



分類・登録

酵素反応データベース・EzCatDB



新たに
酵素反応分野に
適用

酵素反応に関する
専門知識をAIに移す
“AI for 科学技術研究”
の実例

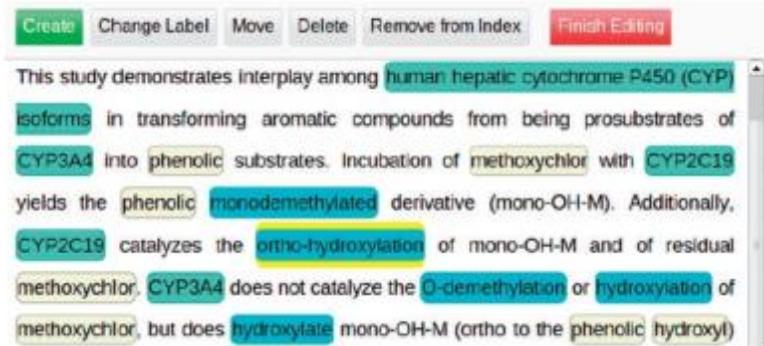
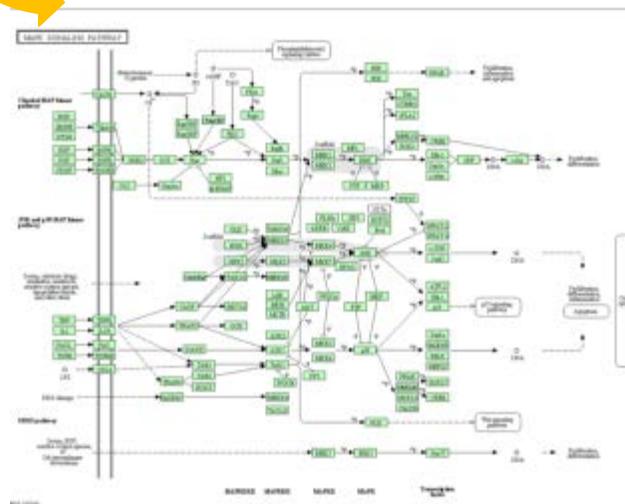


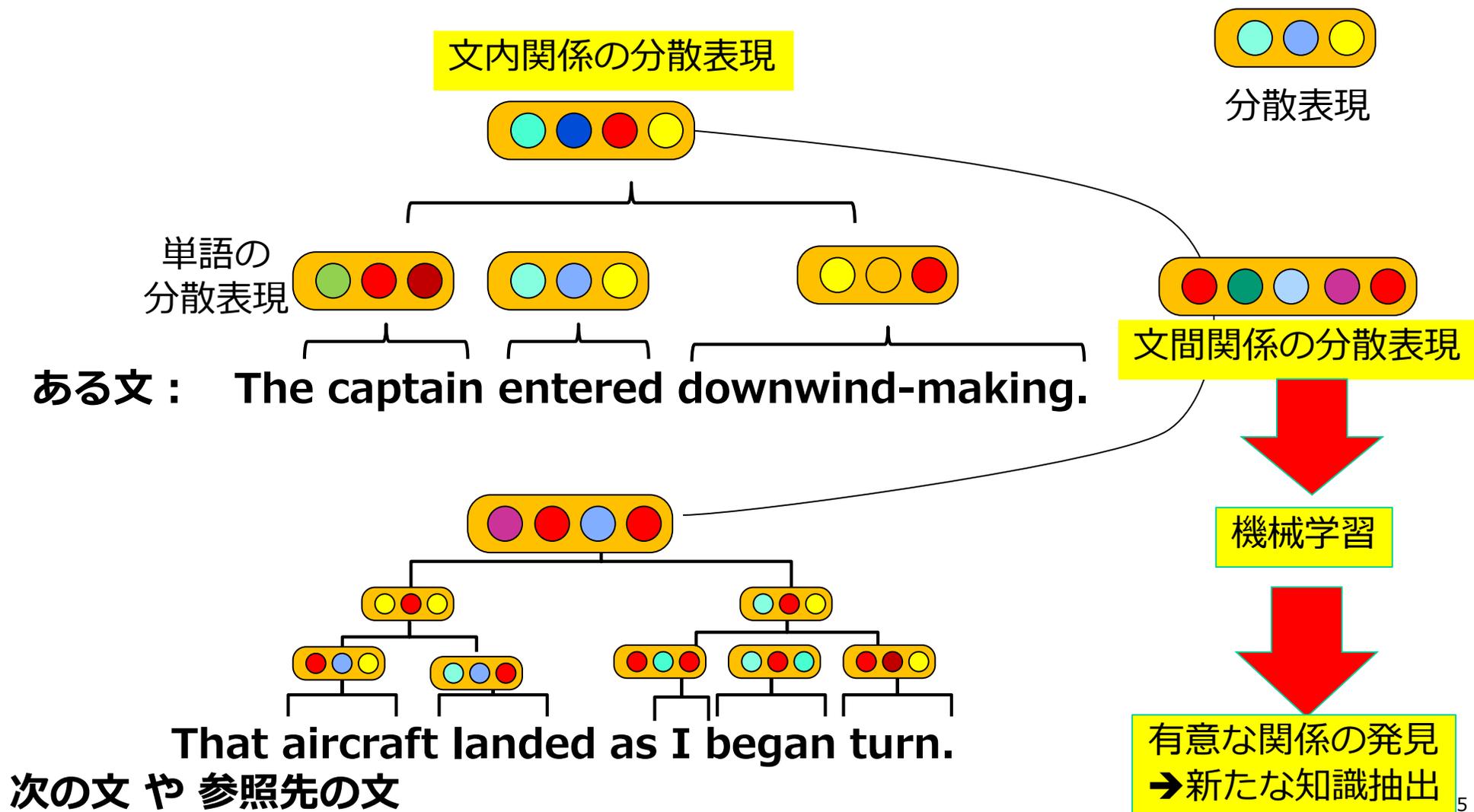
Fig.3 of Rak, *et al.*, (2014) *Database (Oxford) 2014*, pii:bau070.

情報伝達系、代謝経路のキュレーション
に適用実施



MAPK情報伝達系データの例 @KEGG

- 単語、単語共起、文間関係、参照関係、文献引用関係、等の分散表現(embedding)を用いて、テキストに潜在する関係構造を機械学習する → 新たな知識の発見



対象	目的	方法・ツール
科学文献大規模コーパス	重要論文、科学技術分野の予見	<ul style="list-style-type: none"> 引用ネットワークとテキストの統合的な関係構造解析
酵素反応論文データ	酵素反応のデータベース化と新規反応の発見	<ul style="list-style-type: none"> ARGO(マンチェスター大) テキストマイニングとの連携
GPCR情報伝達論文データ	シグナル伝達パスウェイの解明	<ul style="list-style-type: none"> 自然言語意味理解モジュールとの連携