

平成 29 年（2017 年）2 月 8 日

産学連携で研究成果を社会実装へ アーキテクチャ科学研究系 坂本一憲助教の web 情報抽出技術

大学共同利用機関法人 情報・システム研究機構 国立情報学研究所（NII、所長：喜連川 優、東京都千代田区）アーキテクチャ科学研究系 助教、坂本 一憲（さかもと・かずのり）が開発したwebサイトコンテンツなどからの情報抽出技術について、情報・システム研究機構と株式会社ルーター（代表取締役CEO：前田 芳紀、東京都渋谷区）は2月7日付でライセンス契約を締結し、株式会社ルーターが本技術を同社の事業に活用していくことになりました。

坂本助教が開発した技術は、webページのように構造化された文書から個人情報などの特定の情報を機械的に抽出するとともに、抽出した情報を自動的に追跡することができます。本技術では、同じレイアウトが適用されている複数のwebページからページ間で変化する情報を特定して、情報を追跡し続けることを実現します。また、従来の技術ではwebサイトの仕様が変更された後は機械的に情報を抽出することが難しかったのに対して、本技術では、webサイトに仕様変更があっても抽出ルールや特定ルールを再度作成する必要がなく、仕様変更後も機械的に高い精度で情報を抽出し続けることができます。このため、本技術の利用により、情報抽出プログラムの保守コスト削減が期待されます。本技術は特許出願中^(*1)です。

株式会社ルーターは、web サイト上のデータをクローリング^(*2)してwebサイトのデータアグリゲーション^(*3)を行ったり機械学習の教師データを作成したりすることを主要な事業としています。同社は坂本助教が開発した web からの情報抽出技術を事業に活用することで、より運用コストが低く可用性が高いデータ分析や情報抽出サービスの提供を目指します。

NII は情報学分野で国内唯一の学術総合研究所として、情報学分野での未来価値創生を使命としています。その研究分野は幅広く、情報学における基礎論から、人工知能 (AI) やビッグデータ、IoT (Internet of Things=モノのインターネット)、情報セキュリティなど多くの最先端のテーマにも取り組んでいます。NII は今後も産学連携を積極的に進め、研究成果の社会実装により学术界や産業界のみならず日本の社会全体に貢献できるよう、研究活動に取り組んでいきます。

※本件につきましては、本日、株式会社ルーターも別途発表しています。

〈メディアの皆様からのお問い合わせ先〉

情報・システム研究機構 国立情報学研究所

総務部企画課 広報チーム

TEL : 03-4212-2164 E-mail : media@nii.ac.jp

株式会社ルーター

135-0064 東京都渋谷区代々木 4-28-8-303

E-mail : contact@router.co.jp

-
- (*1) 特許出願中： 国際出願番号 PCT/JP2015/084974「情報抽出装置、情報抽出方法、及び情報抽出プログラム」(出願人：大学共同利用機関法人 情報・システム研究機構、発明者：坂本一憲、本位田真一)
(*2) 「クロール」： ロボット型検索エンジンにおいて、プログラムがインターネット上のリンクを辿って Web サイトを巡回し、Web ページ上の情報を複製・保存すること。
(*3) 「データアグリゲーション」： 多種多様な情報源から情報を集約して一覧できるようにすること。