

玄道 俊 松下 光範 山西 良典 (関西大学)

背景と目的

共同利用可能なデータセットを用いて研究を行うことがトレンド

研究者のメリット

コストがかからず大規模データを使用が可能



データ提供者のメリット

先進技術をいち早く事業に取り入れることが可能



論文を特徴化し、共同利用可能なデータセットの使われ方を可視化する

実現可能な未来

研究者の未来

- 自身の研究と目的が同様で、異なるデータセットを用いた研究を探ることが可能
- 研究を始める際、自身の立ち位置を明確化するために使用可能

データ提供者の未来

- 提供したデータセットの使用方法を把握することが可能
- これからデータを提供する提供者は、自身が保有するデータセットと類似しているデータセットを用いた論文から出力結果を予測可能

提案手法

論文内で使用したデータセットの**スキーマ**を記載したデータセットを作成

アブストラクト内で出現した単語を基に、**バイナリベクトル**を用いて論文を特徴化

作成した論文データセット

	アブストラクト	データ1	データ2	データ3	データ4
論文A	本研究では、...	レシピ タイトル	手順の 位置	手順の 内容	
論文B	近年、料理が...	レシピ タイトル	材料の 名前	材料の 分量	手順の 内容



実施例

提案手法を用いて論文を特徴化した実施例を QR コードのリンク先に掲載しています。



謝辞

この研究は 2021 年度国立情報学研究所公募型共同研究 (21S0501) の助成を受けています。