

課題名：生命科学系イベント抽出のための言語モダリティのAIの開発

実施時期：2025/04/15 - 2026/03/30

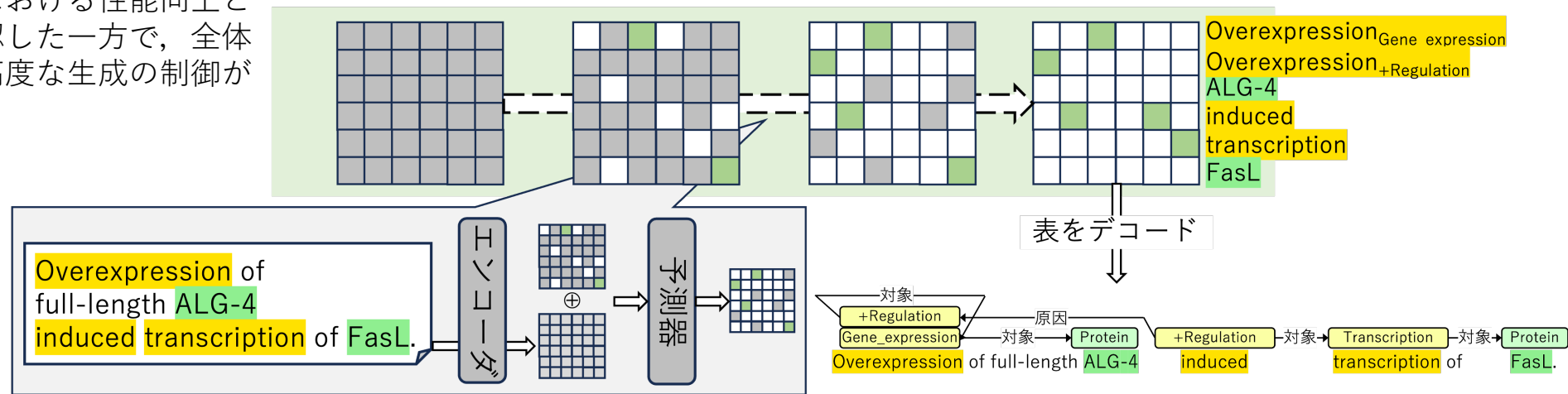
所属機関名：豊田工業大学

代表者氏名：三輪誠

成果概要：文献情報から生命科学系イベント抽出を行う言語モダリティのAI開発を目的に、拡散モデルを利用した表形式の抽出手法に基づいた用語抽出・イベント抽出モデルの開発を行なった。

成果のポイント：

従来のイベント抽出モデルはイベントを個別に順次抽出しており、一度抽出した結果を修正することはできなかった。この問題を解消するために、イベントをまとめて抽出した上で、その抽出した情報を修正し、複数イベント間で整合性のとれた抽出の実現を目指した。具体的には、文中のイベント全体を表形式で表現し、表を埋めることでまとめた抽出を実現し、その表を拡散モデルを用いて修正を繰り返す手法を提案した。反復によってイベントを表すトリガー抽出における性能向上と制約違反のイベントの減少を確認した一方で、全体的な性能には変化は少なくより高度な生成の制御が必要であることがわかった。



成果についてより詳細な情報を提供しているWebページ、発表論文などの情報：

金児 一矢, 井上 純大, 三輪 誠. 反復的な抽出により抽出結果を精練する生物医学イベント抽出. 言語処理学会第32回年次大会発表論文集. pp. 2765-2770. 2026年3月.

https://www.anlp.jp/proceedings/annual_meeting/2026/pdf_dir/C6-24.pdf