

高速な類似文字列検索アルゴリズムと自然言語処理への応用

岡崎 直観, 辻井 潤一

- ▶ 辞書（外部知識）の参照は、様々なNLPタスクで重要
 - 例：固有表現抽出
 - 機械学習手法を良くすることも重要だけど…
 - 人名リスト，タンパク質リストなどの外部知識を効率よく詰め込むことも重要

▶ 苦勞して作った辞書（知識）
と，文章中の表現が微妙に一
致せず，歯がゆい思いをする
ことがある

- polymerase chain reaction
 - polymerase-chain-reaction
 - polymerized chain reaction
 - polymer chain reaction
 -

- ▶ 知識を上手に活用しましょう
 - 似ている表現を瞬時に検索し、ちよつとの表層表現のずれを許容できるようにになれば、NLPとエコに貢献できるのではないか