

今月の



隣に伝えたい 新たな言葉と概念

【テキストマイニング】

英 text mining

和 テキストマイニング

類 同義語（日本語，英語），類義語（日本語，英語）：テキストデータマイニング，text data mining

当該用語の解説：テキストデータ（文書・言語など）を解析して，有用な情報を探ること（マイニング：発掘）．構造化されていない自然言語をデータとして扱うためには，自然言語処理（Natural Language Processing：NLP），すなわち形態素・構文・意味・文脈の4つのプロセスで加工する必要がある．NLPで加工されたテキストデータをテキスト出現頻度（term frequency）や逆文書頻度（inverse document frequency）の解析に加え，センチメント・共起・対応・主成分分析を行う．テキストマイニングは，SNS分析や自由記載のアンケート分析などに利用されており，企業などでは，顧客アンケートやWeb情報などから，販売戦略を行うのに利用されている．医療も診療記録・患者 - 医療者，医療者間などコミュニケーションや記録など多くの言語を利用しており，近年テキストマイニングを用いた医学系論文は増加している．同じような概念の言語を提示するテキストマイニングの技術は，検索システムにも利用されており，私たちは，意識していないが実生活で汎用している（例：Google検索など）．

その他必要事項（本用語とつながりの深い専門分野，関連学会など）：情報科学