

CGMテキストを対象とした自然言語処理: 今後の展望と課題

鍛治 伸裕

東京大学生産技術研究所

kaji@tkl.iis.u-tokyo.ac.jp

概要 CGM の出現によって、人々の行動や思考内容がテキストデータとして大量に発信、蓄積される時代が到来した。その結果、例えば評判分析など、テキストデータの解析を通して社会の動向を分析することに対する期待が高まっている。こうした背景から、近年の自然言語処理において、CGM テキストは研究の潮流を形成する軸の一つとなりつつある。すなわち、CGM テキスト解析に対する社会的なニーズの高まりを受けて、新しい課題が研究者に認識され、それに対する解析技術が進展するというサイクルを見ることが出来る。こうした現状を踏まえて、本講演では、CGM テキストが自然言語処理にどのような影響を与えてきたのか、そして、自然言語処理が CGM テキストと関わっていく中で、今後どのように発展すべきかについて議論を行う。

キーワード CGM テキスト, 自然言語処理