

Web 上の情報収集行動の振り返りを支援する Web ページ閲覧履歴の提示方法

佐藤 千尋

本研究は Web 上の情報収集行動の記録である Web ページ閲覧履歴に着目する。現代では Web ブラウザを利用できるスマートフォン等の普及が進み、あらゆる場面で必要な情報収集を Web 上で行える。情報収集の過程で閲覧した Web ページの履歴はブラウザによって自動的に記録される。この Web ページ閲覧履歴を情報収集行動の振り返りに役立てることができれば、Web ページ閲覧履歴のライフログとしての価値が高まる。さらに情報収集行動を振り返ることによりユーザは自身の情報収集プロセスやその背景にある当時の思考や興味、生活を把握することが可能になる。

本研究の目的は Web 上の情報収集行動の振り返りを支援する Web ページ閲覧履歴提示方法の提案である。

本研究の提案手法は、(1) Web ページ閲覧履歴のメタデータ抽出、(2) Web ページの本文取得、(3) 分かち書きによる特徴語群取得、(4) Web ページの特徴を表すベクトルを作成、(5) Web ページを内容ごとにまとめるクラスタリング、(6) 内容を表す主題ラベルを作成、(7) 主題と時系列を組み合わせた可視化、の 7 段階で構成される。(4) の Web ページの特徴を表すベクトルは TF-IDF 値を用い、(5) のクラスタリングは K-means 法を用いる。(7) の可視化のツールとして Google Charts から二次元空間にデータをプロットする Scatter Chart を使用する。提案手法による提示画面では、閲覧した Web ページがそれぞれ横軸は主題、縦軸は時系列に従って二次元空間上に表示される。このとき同一主題の Web ページが同一の横軸目盛り上に並び、閲覧した Web ページの主題が時系列に沿って移り変わる様子が可視化される。提案手法の有効性を確かめるため、15 人の参加者に対して評価実験を行った。実験参加者の閲覧履歴データから古い順に抽出した 100 件のデータを提案手法と既存手法 (Google Chrome) のそれぞれによる提示画面で見てもらったのち、当時の情報収集プロセスや背景をどの程度思い出すことができたかを明らかにする主観評価アンケートを実施した。また、実験の前後に実験参加者の Web ブラウザや閲覧履歴の利用状況を明らかにする事前アンケートと提案手法システムの問題点を調査する事後アンケートを実施した。

評価実験の結果では提案手法と既存手法との間に有意な差が見られず、提案手法の有効性を確かめることができなかった。提案手法システムの構築と評価実験を通して、特徴を抽出できない Web ページが存在する・クラスタリング精度が低い・適切でない主題ラベルが生成される・提示画面で主題が見分けづらい・実験用データが 100 件では提案手法の効果が十分に発揮されない場合がある、など本研究の課題が明らかになった。提案手法や評価実験の改善に取り組み、提案手法の評価実験を再度行うことを今後の展望とする。

(指導教員 高久雅生)