

CD 理論による情報検索支援に関する研究

A Research on information retrieval support by CD theory

1w130513-6 牧山直堯
Naotaka Makiyama

指導教員 長 幾朗 教授
Prof. CHOH Ikuro

概要：スマートフォンやタブレットの普及に伴い誰もが手軽に検索エンジン、SNS に代表されるプルメディアによる情報収集が可能となった。しかし、多くのユーザが自身の関心の高い情報やコンテンツだけを検索するという現状がある。そこで本研究では気軽に情報収集できる SNS の特性を活用し、ユーザに対しあえて関心の低い情報を検索させるきっかけ作りをする情報検索支援システムの提案を行った。

キーワード：プルメディア、CD 理論、情報検索

Keywords : pull media, CD theory, information retrieval

1. はじめに

情報通信技術の発展により誰もが気軽に多種多様な情報を収集することが可能となった。何かを調べる際に「インターネットの検索サイトで検索する」という回答が約 7 割を占めるという(総務省,2014)[1]。一方で、検索エンジンによる情報収集は、能動的に関心のある情報を検索することが多い。従来のテレビやラジオに代表されるプッシュメディアでは当たり前であった「思わぬコンテンツとの出会い」の機会が失われていることを示唆している。

2. 文章理解

インターネットの情報の多くがテキストデータで書かれている。しかしインターネット記事のインタラクションを調査すると、閲覧者の 1 ページの閲覧時間は 15 秒未満であり、スクロールの深度（どの程度下までスクロールされているか）は大部分のユーザが 60%までスクロールするが、10%の人は全くスクロールしないという(トニー・ヘイル,2014)[2]。一方で、オンラインにおける「読み」の特徴に「外的な体験」と「内的な体験」がある。集中して読解する時は、内容を解釈しようとする内的な体験になる。しかし流し読みする際は、人はより視覚的な読み方になる。この視覚的な読み方では外的な体験になり内容を深

く理解しようとする機会は失われてしまう。前述したオンラインにおける記事の調査結果からもユーザはオンラインの記事を外的な体験により読解していることが多いことを示している。

3. CD 理論

認知科学の方法論の中で、知識表現の考えがある。知識表現とは、ある形式で表現された情報がある機構の適用によって情報の内容を変化させることである(橋田他,1995)[3]。一般的に、対象とする文章をノードとリンクで表現する手法である「意味ネットワーク」が主であるが、自然言語全体の意味表現を実現するのに、どの程度の意味要素を設定すればよいのかという課題を考えなければならない。そこで、ジャンク[4]によって CD 理論が提案された。CD 理論の優れている点として抽象的な意味により構成された基本行為概念にある。しかし、ジャンクが述べているように基礎行為概念が本当に CD 理論で用意されている 11 の概念で済むのかどうかという議論は今日まで、避けられてきている。

4. 提案の概要と目的

検索エンジンや SNS のように、自らの意思で情報を取得するプルメディアによる検索の効率性を高めることで「思わぬコンテンツとの出会い」の

実現を目的とした。特に情報ニーズに合ったサイトが少なく検索に相応の時間がかかるという背景[5]から、世間的に関心が急上昇し該当するサイトが見つからないと思われる「最新情報」に関する情報検索支援を本提案システムの機能とした。そこで、同じプルメディアである SNS の中で「知りたいことについて情報を探すため」という目的で利用しているユーザが多かった Twitter をプルメディアのサンプルとした。すなわち、最新情報に関する Twitter のツイートを用いることで前述した提案システムの提案を行った。

5. 提案手法

Google の検索エンジンで多くのユーザに検索されているトレンドワードを含んだツイートを収集し、形態素解析により解析を行った。その収集キーワードのカテゴリによって特徴を抽出し、各カテゴリにおける最もツイートされやすいテンプレート文を作成した。抽出した特徴から新たに基本行為概念を見直し、それを表すピクトグラムにより概念モデルを作成した。テンプレート文に対し CD 理論を適応しネットワーク表示を作成し実験用のツイートに対し図 1 のようなネットワーク表示を作成した。



図 1 ネットワーク表示文例

上記のネットワーク表示を以下の 3 つの条件で Processing にて実装した仮想情報収集環境の下で表示しシステムの評価を行った。

- 条件 i 原文そのまま
- 条件 ii ネットワーク表示を補助的に表示
- 条件 iii ii をより簡潔に表示

6. 実験結果

実験は健常者 15 名を対象とした。実験は条件 ii, iii それぞれで「ツイートの再現性」「被験者の

理解度」「関心度の変化」「実際に利用したいか」「被験者の満足度」の五段階評価と感性評価を SD 法により行った。実験の結果は条件 ii と比べ条件 iii の方が、被験者にとって分かりやすく伝わり、条件 iii ではツイートの原文と比べ 7 割以上の被験者が関心度に変化があったと回答した。

7. 結論

実験結果から「新たな品詞の拡張」「カテゴリ別の更なるネットワーク表示の細分化」「実装環境の見直し」の 3 つが主な課題として挙げられた。全体として、情報の種別による最適化されたネットワーク表示と提示方法が考えられた。今後、より汎用性の高い処理システム提案のために、更に情報を収集しより網羅性の高い概念モデルの提案が不可欠であると考えられる。

参考文献

- [1] 総務省平成 26 年度情報通信白書 ICT がもたらす世界規模でのパラダイムシフト(ネット利用の最大の目的)
<http://www.soumu.go.jp/johotsusintokei/whitepaper/jah26/html/nc141120.html>(2016/12/30 現在)
- [2] スーザン・ワインチェンク 著 武舎広幸 武舎るみ 阿部和也 訳 『続インタフェースデザインの心理学』 pp106-109
- [3] 橋田浩一 安西祐一郎 波多野誼余夫 田中啓治 郡司隆男 中島秀之著 『認知科学の基礎』 岩波書店 1995 pp10
- [4] 戸田正直 阿部純一 桃内佳雄 往住彰文 著 『認知科学入門「知」の構造へのアプローチ』 サイエンス社 p115
- [5] 総務省 情報通信政策研究所 平成 21 年 インターネット検索エンジンの現状と市場規模に関する調査研究 (検索サービスの将来展望)
<http://www.soumu.go.jp/iicp/chousakenkyu/data/research/survey/telecom/2009/2009-I-14.pdf>(2017/01/23 現在)

図表出典一覧

図 1 牧山(2017)