

令和 2 年 7 月 13 日現在

機関番号：32410

研究種目：基盤研究(C) (一般)

研究期間：2016～2019

課題番号：16K00702

研究課題名(和文)自らを通時的な対象とする情報蓄積フレームワークの構築

研究課題名(英文)Construction of an information accumulate framework with diachronic objects

研究代表者

田中 克明(Tanaka, Katsuaki)

埼玉工業大学・人間社会学部・准教授

研究者番号：80376657

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：本研究では、蓄積済みの情報をもとに、新たな情報を蓄積する際の対象となる「通時的対象」を設ける情報蓄積フレームワークの検討と構築を行った。これにより、1. 蓄積する情報と同様の形式をもち、2. 把握がほぼひと目で行え、3. プロセスを示す情報より結果を示す情報であることが、通時的対象として必要であることが明らかになった。この条件のもと、「通時的な対象には、異なるタイミングで類似した言及がなされる」と仮定し、蓄積済みの情報から通時的対象を抽出し、新たな情報を蓄積する際に提示する枠組みの構築を進めた。

研究成果の学術的意義や社会的意義

SNSの書き込みや散発的に行われる会議では、情報の記述対象と情報を蓄積する人間の両方が入れ替わるため、議論が発散する・堂々めぐりするなど、蓄積された情報において時間の経過が意味を持たないことがある。そこで本研究では、情報の記述対象として、情報の蓄積過程においてある一定の期間を通して対象となりうる「通時的対象」を、それまでに蓄積された情報から抽出し、新たな情報を蓄積する際に提示する枠組みの検討を行った。今後は、複数の人間が入れ替わり情報を蓄積する場において、建設的な議論を支援する仕組みへと応用する研究を進める。

研究成果の概要(英文)：In this study, we developed an information accumulate framework with "diachronic objects" which are targets of describing new information. It was found that 1. diachronic objects should be in the same type as the information to be described, 2. they should be easy to understand at a glance, and 3. they should be more indicative of results than process information. Under these conditions, we assumed that "a diachronic object has similar references at different times," and constructed a framework extracting diachronic objects from the accumulated information and presenting them when accumulating new information.

研究分野：知能情報学

キーワード：時系列文書処理 知能情報学 情報可視化 音声情報蓄積 テキストマイニング デジタルストーリーテリング

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

ニュース記事や論文など、ある程度一般的な内容を含む文書集合では、含まれる内容を時刻順に提示するだけでも、人間が時間経過に伴う内容変化を感じ取ることができる。これは、内容が一般的であるがゆえ、時間軸に沿った提示のみで、読み手となる人間の知識との対応付けが行えるためと考えられる。

これに対し研究代表者は、読み手となる人間の視点に沿い記述内容の抽出を行うシステムの研究を進めていた。このシステムは、文書集合から時間経過に沿った記述内容を抽出した後、人間が指定した単語を中心にそれらを再構成し、読み手の視点に沿った提示を行うものである。このシステムにより、さまざまなプロジェクトのメーリングリストや人工物の設計過程の記録など、情報を記述した人間以外には内容を把握しづらい文書集合において、「何がどのように行われたか」という時間経過に伴う内容変化の理解を支援することができた。

しかし、このシステムを、Twitter 上のツイートや外部からの話題提供者の講演を中心とした審議会の議事録など、文書の内容を記述する人間が常に入れ替わる文書に対して適用したところ、文書に書かれた内容が「その通りである」こと以上を示すことが難しく、文書集合に含まれることが期待される「何がどのように行われたか」という内容の理解を行うことが難しかった。

この原因として、ツイートや対象とした審議会の議事録が、メーリングリストや設計過程の記録と異なり、文書の生成に関わる人間が常に入れ替わることから、文書集合より継続的な意図を汲み取りづらいことや、各文書の記述内容が、事件など外的要因に重きを置き、それまでに蓄積された情報とは関係性が低い状態で記述・蓄積されていることなどが、考えられる。

なお、このような状況に対し、新しい記述内容の有用性を、記述を作成する人間の行う表記に着目し分析する研究などがなされていた。

### 2. 研究の目的

本研究では、人間が入れ替わり情報を蓄積される過程に「通時的な対象」を設け、この対象を意識して情報の蓄積をうながすことにより、時間経過に伴って蓄積される情報の集合に、継続的な意図が含まれるようにするフレームワークの構築を目指した(図1)。本研究における通時的対象は、それまでに蓄積された情報自体から構成することとした。そのために、さまざま情報蓄積の場面において、通時的対象としてどのようなものを設定することが適切かの研究を行った。

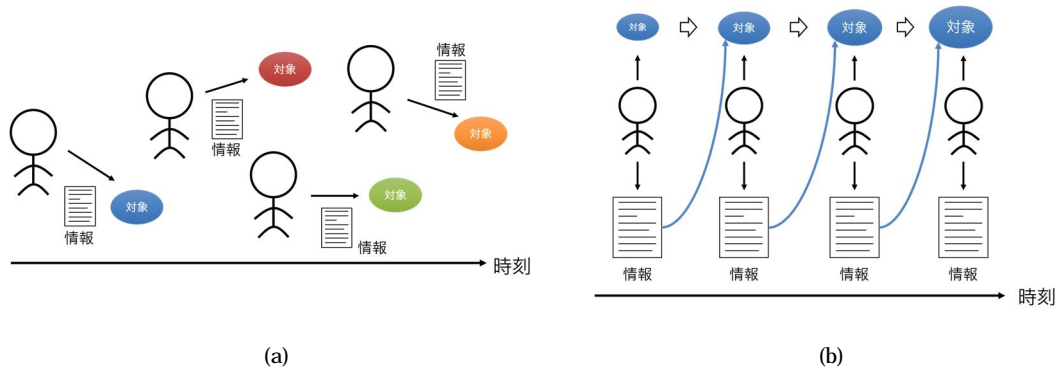


図1: 時間経過に沿った単純な情報蓄積(a)と提案フレームワークと用いた情報蓄積(b)

### 3. 研究の方法

(1) 研究代表者は、本研究開始と前後して、黒電話型のデバイスを用いた音声による語りの蓄積に関する研究を行っていた(図2)。語りの蓄積において、「黒電話」が本研究の「通時的対象」と同様に、人が新たな情報の蓄積を行う際に方向付けを与える役目を果たしていると考えたことが、本研究の端緒である。そこで、音声情報の蓄積に着目し、電話機以外にも、蓄積された語り(情報)が新たな語りを誘発するきっかけとなる方式を検討するところから、本研究を開始した。

新たに語りを蓄積する際に、黒電話付近において、蓄積された音声を自動再生する方式と、蓄積された音声をテキストに変換したものを表示するタブレットを設定する方式を併用して、実験を行った。この結果、テキストに変換して表示することで蓄積されている音声を能動的に聴取しようとする際の利便性の向上は認められる一方、新たに蓄積音声を聴取しようとする「きっかけ」を作るためには、音声を自動再生する仕組みが有用であることが確認できた。

(2) 時間の経過に沿って文書を蓄積する場面を想定し、それまでに蓄積された文書に含まれる単語を、平面上にレイアウトし、時間経過に沿って変化させるアニメーション表示を、通時的対



図2: 黒電話型デバイスによる音声収集システム

象とする実験を行った。このアニメーションは、対象とする文書集合に現れた単語や、文書集合からトピック抽出により抽出したトピックの代表語を平面上にレイアウトし、文書が作成されたタイミングごとに表示・非表示を切り替えることにより生成した(図3)。

この方式では、多くの単語を表示させると内容が把握しづらく、少数の単語に表示を絞ると、時間経過に伴う変化がなく特徴語を一覧表示させる形式とほぼ同様の結果に過ぎなくなることがわかった。これらの結果から、時間経過に沿った変化をほぼそのまま提示するのでは通時的対象として期待した効果は得られず、時間経過を意識させない形での表示が必要であると考えられる。

(3) その他、複数のグループにより同時に行われる議論をリアルタイム入力・提示し、通時的対象として議論を進める方式の研究を行った(図4)。この研究では、各グループの議論そのものの記録と、各グループの議論の仮のまとめの入力・提示を行うこととし、それぞれが通時的対象として議論を蓄積していく中心となるかの検証を行った。この結果、他グループの議論も参考になるが、自グループのまとめを構築していくことの方が議論を蓄積するために果たす役割が大きいと考えられることがわかった。

(4) 次に、すでに蓄積された文書集合から、通時的な対象を抽出する手法の研究を行った。この研究では、通時的な対象は、「文書集合に現れる通時的な対象には、異なるタイミングで類似した言及がなされる」と定義し、抽出を行った。抽出された結果が通時的な対象とみなせるか、出現率の高い単語と比較して評価を行い、この手法により、通時的対象の抽出が可能であることが分かった。

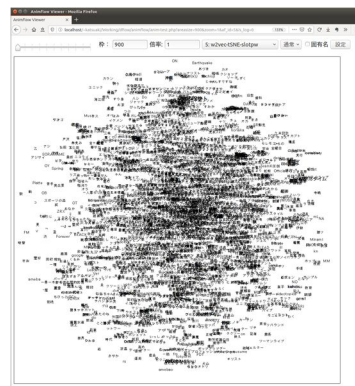


図3: 文書集合出現単語のアニメーション表示



図4: 複数グループによる議論の内容の提示システム

#### 4. 研究成果

研究方法(1)~(3)に述べた手法を通し、本研究で目指した、時間経過に沿って情報を蓄積する過程でその情報自体を通時的対象とする際には、通時的対象として、1. 蓄積する情報と同様の形式をとること、2. 対象の把握がほぼひと目で行えること、3. プロセスを示す情報より結果を示す情報であること、の3つが必要であることが示された。

そこで、文書集合を対象とし、研究手法の(4)に述べた単語を通時的対象とする手法の研究を進めた。現在までに、研究背景に述べた文書集合可視化システムの一部に組み込みを終え、SNS上での議論時に、対象を明らかにしつつ議論を進めるための仕組みに適用する手法の検討を行っている。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 田中克明	4. 巻 17
2. 論文標題 共起語の類似度を利用した文書集合からの変化記述の対象抽出の試み	5. 発行年 2019年
3. 雑誌名 埼玉工業大学人間社会学部紀要	6. 最初と最後の頁 15-22
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計8件（うち招待講演 2件 / うち国際学会 3件）

1. 発表者名 田中克明
2. 発表標題 共起語の類似度と時刻分布を利用した文書集合からの変化記述の対象抽出の試み
3. 学会等名 第33回人工知能学会全国大会
4. 発表年 2019年

1. 発表者名 Katsuaki Tanaka, Koichi Hori
2. 発表標題 Finding Diachronic Objects of Drifting Descriptions by Similar Mentions
3. 学会等名 2019 Pacific Rim Knowledge Acquisition Workshop（国際学会）
4. 発表年 2019年

1. 発表者名 Shin Mizukoshi, Rikutaro Manabe, Katsuaki Tanaka
2. 発表標題 Digital Storytelling
3. 学会等名 Doing Digital Methods: Interdisciplinary Interventions（国際学会）
4. 発表年 2018年

1. 発表者名 Shin Mizukoshi, Masako Miyata, Rikutaro Manabe, Katsuaki Tanaka
2. 発表標題 Telephonoscope: A Media Design Study of Technologies and Cultural Programs for Novel Microdigital Storytelling
3. 学会等名 67th Annual Conference of the International Communication Association (国際学会)
4. 発表年 2017年

1. 発表者名 田中克明
2. 発表標題 システム0708によるふり返し
3. 学会等名 『5』シンポジウム：次の一手はなにか（招待講演）
4. 発表年 2017年

1. 発表者名 水越伸、宮田雅子、真鍋陸太郎、田中克明、栗原大介
2. 発表標題 テレフォノスコープ：電話機型装置によるマイクロ・デジタル・ストーリーテリング
3. 学会等名 カルチュラル・タイフーン2016
4. 発表年 2016年

1. 発表者名 田中克明
2. 発表標題 時間の経過に沿った語りの蓄積と再構成
3. 学会等名 埼玉工業大学第14回若手研究フォーラム（招待講演）
4. 発表年 2016年

1. 発表者名 田中克明
2. 発表標題 黒電話により個人の語りを蓄積する仕組みの検討
3. 学会等名 人工知能学会 第1回市民共創知研究会
4. 発表年 2016年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----