

## 情報信憑性判断のための自動要約に向けた 人手による要約作成実験とその分析

中野 正寛

横浜国立大学 大学院 環境情報学部

宮崎 林太郎

横浜国立大学 大学院 環境情報学部

森 辰則

横浜国立大学 大学院 環境情報研究院

E-mail: {nakano,shib,rintaro,ishioroshi,mori}@forest.eis.ynu.ac.jp

渋谷 英潔

横浜国立大学 大学院 環境情報研究院

石下 円香

横浜国立大学 大学院 環境情報学部

本稿では情報信憑性の判断に役立つ要約について扱う。ブロードバンド化の進展やブログの普及に伴って爆発的に増加する情報の中には利用者にもたらす情報も含まれており、情報の信憑性判断を支援する技術の実現は重要な課題である。我々は情報信憑性の判断に役立つ要約の自動生成に向けて、複数の作業者の人手により情報信憑性判断のための要約を作成する実験を行った。そして、得られた要約文章の安定性や複数作業者間の一致度を分析し、情報信憑性の判断に役立つ要約を作成する際に人間が重要視する情報を調査した。

実験結果では、要約対象として収集した文書から要約に必要な記述を抜き出すまでの何段階かの絞り込みで中程度の一致が期待できる事がわかった。事後の作業者へのアンケート調査によれば、要約として適当な長さと考えたのは1,000から3,000文字程度であり、作業者が情報信憑性の判断に役立つ情報として重要だと考えたのは情報の発信者及び具体的な記述がまともなものである。また、情報信憑性の判断に役立つ要約の自動生成に向けて、作業者が作成した要約を参照要約とし、それに対応する原文からの文の抜粋に関するデータを整備した。

## Experiment and analysis of manual summarization aiming at automated summarization that supports the judgment of the information credibility

Masahiro NAKANO Hideyuki SHIBUKI Rintaro MIYAZAKI  
Madoka ISHIOROSHI Tatsunori MORI

Graduate School of Environment and Information Sciences, Yokohama National University

E-mail: {nakano,shib,rintaro,ishioroshi,mori}@forest.eis.ynu.ac.jp

In this paper, we investigated processes of text summarization that supports the judgment of the information credibility. The realization of technology that supports the judgment of the information credibility is one of important issues, because information harmful to users creeps into the huge amount of available information in the era of information explosion.

Aiming at automated summarization, we conducted experiments in which multiple subjects manually created text summaries from the viewpoint of the judgment of the information credibility. We studied the stability of the summarization and the degree of agreement in the results of extraction of important information. We also investigated the information that subjects supposed to be important in the process of the creation of summaries, by using questionnaires after the experiments.

The experimental results show that moderate agreement can be expected in the result of extraction of important information. The length of summaries was between about 1,000 and 3,000 characters. According to the questionnaires, the documents that were well-organized and information about information holders were supposed to be important. Aiming at the automated summarization, we also prepared the information of the extracted sentences that correspond to the created summaries.

### 1 はじめに

Web上に存在する情報は、ブロードバンド化の進展やブログ等の普及に伴い、爆発的に増加し続けている。これらの情報の中には、出所が不確かな情報や利用者にもたらす情報などが含まれてお

り、信頼できる情報を利用者が容易に得るための技術に対する要望が高まっている。しかしながら、情報の内容の真偽や正確性を検証することは困難である上に、その情報が意見などの主観を述べるものである場合には、利用者により考え方や受け止め方が異なることから、その真偽や正確性を検証すること

はさらに困難なものとなる。そのため、情報の信憑性は、最終的に個々の情報利用者が判断しなければならず、利用者による信憑性の判断を支援する技術の実現が優先して解決すべき課題であると考えられる。我々は、主観的な意見や評価だけでなく、疑問の表明や客観的事実の記述を含めたテキスト情報を広く言論と呼ぶこととし、ある言論集合における個々の言論の相対的な位置づけを提示することで情報信憑性を表現することを目指している。

信憑性の判断を支援する技術には幾つか考えられる。まず、利用者が着目するトピックに対して、言論間に存在する根拠情報や矛盾・対立情報といった論理的関係を提示することで、利用者が情報を多角的に俯瞰することを支援する技術である。赤峯ら[5]や河原ら[8]は、述語項構造に基づき、あるトピックに関する主要・対立表現を俯瞰的に提示するためのシステム WISDOM を開発しており、村上ら[6]は、扱える論理的関係の種類を拡張した言論マップの生成課題を論じている。

これらの論理的関係を解析するための技術と並んで、関連言論群や背景情報を論理的関係に基づいて要約・整理するサーベイレポート生成技術も信憑性判断において重要である。なぜならば、爆発的に増加する Web 文書の多様性を考えると、個々の言論に関係づけられる周辺言論が往々にして膨大な数となり、利用者の理解が困難になることは想像に難くないからである。背景情報のような弱い関連情報まで提示することを想定すれば、利用者にとって情報過多となることは必至である。そこで、まず、関連する言論間の論理的関係の全体像を把握するのに有用な、論理的関係の要所に位置する言論を重要な言論とみなし、それらを優先的に提示することによって情報量を抑えることを考える。また、言論間の論理的関係の把握が容易になるようにこれを簡潔な文章に自動要約し、サーベイレポートとして利用者に提示できれば、着目する言論に関連する言論群の全体像が把握しやすくなると考えられる。

要約を目的とした文間の論理的関係を扱った研究として、Radev ら [1, 2] により提唱された、RST (Rhetorical Structure Theory) を文書間関係に拡張した CST (Cross-document Structure Theory) が存在する。しかしながら、Web 文書ではなく新聞記事を対象としており、信憑性判断という観点から行われたものではない。こういった背景から、我々は、Web 文書を対象とした情報信憑性判断のための自動要約の研究に取り組んでいる。

本稿の構成は以下の通りである。2 章では、我々が目的とする情報信憑性判断のための自動要約に関して説明する。3 章では、本実験の目的とそれを実現する手段について述べる。4 章では、人手による情報信憑性判断のための要約作成実験の作業について説明する。5 章では、要約作成作業の結果と、その結果についての考察を行う。6 章では、本実験についてのまとめを行う。

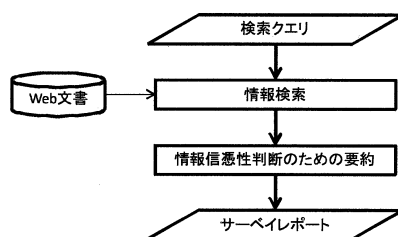


図 1: サーベイレポート生成までの過程

## 2 情報信憑性判断のための自動要約と本稿の位置づけ

我々が目的とする、情報信憑性判断のための自動要約は複数文書要約 [4] の一種である。しかしながら、従来の新聞記事等を対象とした複数文書要約 [7] と比較して以下のような特徴がある。

複数文書要約において情報の発信者は複数存在するのが普通であるが、対象文書が新聞記事等の場合には、どの発信者も一貫した論旨で記述していることに加え、発信者が明示されることが少ないため、記述内の矛盾点や発信者の存在を意識することはない。したがって、従来の複数文書要約では、どの情報も同等に信じられるとしており、言論の間にも矛盾はないとしていた。一方で、情報信憑性判断のための自動要約が対象とする Web 文書では、原文書の情報が全て信じられるとは限らず、どの言論が本当に正しいのか分からない場合がある。その結果、言論間に矛盾や対立関係が存在しうることとなり、これらを明らかにするために情報発信者による情報の区分が重要となってくる。

また、情報利用者による信憑性判断は一般的な情報アクセスの枠組みの中において行われることを想定しているため、その判断を支援する自動要約も、図 1 に示すように、情報検索まで含んだ過程の中でとらえる方がより実際の状況に即していると考えられる。すなわち、利用者が着目するトピックに関する文書を検索するためのクエリを入力し、その結果となる文書集合を信憑性判断の観点から要約してサーベイレポートを生成するという流れである。それゆえ、以下の議論においても、クエリを用いて Web 文書を収集することから始めている。このように、情報信憑性判断のための自動要約は今まで広く行われてきた複数文書要約と異なる部分があり、参照要約の作成においても上記の点を考慮して行う必要がある。

したがって、本稿では、情報信憑性判断のための参照要約を手手で作成し、以下の点を明らかにする。

- 人間が情報信憑性判断のための要約を作成する際に重視する情報は何か
- 要約の出力として適当な形式や長さほどの程

度か

- 複数の作業員で要約を作成した場合の一致度はどの程度か

また、本稿では、情報信憑性判断のための自動要約を以下単に要約と呼ぶこととする。

### 3 要約作成過程に関する考察

2章で述べたように、本実験は最終的に情報信憑性判断のための要約を自動的に行うために、人間がどのように要約を作成し、結果としてどのような要約が作成されたかを観察し、考察することを目的としている。情報信憑性判断のための要約とは、ある特定の話題についての言論を集めて言論間の論理的関係を明らかにし、利用者の情報信憑性判断を支援する目的で取りまとめ、人間が読みやすいように要約したものと同様では位置づける。

また、情報信憑性判断のための要約作成を自動的に行うために、その「正解」となる参照要約を準備することも目的のひとつである。自動要約システムの理想的な出力は人間が作成した要約と同様のものであるが、人間と違ってシステムが最初から文章を書き起こすことは困難であるため、通常は対象文書から得られた抜粋を基にして要約文章を生成する手法が採用される。そこで、人間が文章を書き起こして作成した要約と、要約対象となる文章からの抜粋を中心として作成した要約とを関連付けておくことで、自動要約の際の「正解」として利用できるようにする。これは、平尾ら [9] における参照要約の作成と同様の考え方に基づいている。

#### 3.1 人間が要約を作成する過程に関する考察

本実験では、人間が情報信憑性判断のための要約を作成する際にどのような情報を重視して要約を作成するのか、またどのような知識が要約の作成に必要なだったのかといったことを調査する。通常の要約では、要約の対象となる文書はタスクの開始時に与えられることが多いが、本実験では、人間が要約の対象となる文書を Web からどのように収集することも調査の対象とする。

人間が実際に情報信憑性判断のための要約を作成するには、要約の対象とする文書に目を通し、要約に必要な記述が含まれる文書を選別し、選別した文書に含まれる重要な記述を中心に要約を作成すると考えられる。言い換えると、要約の対象となる文書から要約に必要な記述を得るためには、文書の選別や重要な記述の抽出など、単位を変えて何度かの絞り込みを行っていると考えられる。また、要約の前段階である文書の収集も、絞り込みの一種と考えられる。従って、本実験では人間が実際に要約を作成

する際に行う絞り込みを観察し、複数の人間が行う選別や抽出の一致率を調査する。絞り込みの結果が作業員によって大きく異なるならば、重要だと考える基準が作業員によって大きく異なるということであり、安定した自動要約を実現するのが困難になると考えられる。

また、適切な絞り込みを行うには、要約の対象となる文書に記述されているトピックに関するさまざまな背景知識が必要になると考えられる。トピックに関する背景知識は、一般的には要約を作成する際に必須のものではないが、本実験では作成する要約は情報信憑性判断のために役立つものであるため、要約の作成者には要約対象のトピックに関する問題点や論点を俯瞰できるような背景知識がもって必要になると考えられる。従って、本実験では人間が実際に要約を作成する際にこのような背景知識をどのように獲得し、どれを重要とするかを観察する。

先に述べたように、人間が実際に要約を作成する際には、要約の対象となる文書から要約の作成に必要な記述を抜き出すことで絞り込みを行い、それを中心としてトピックに関する要約を作成すると考えられる。この要約の作成に必要な記述とは、要約の作成者が重要だと考える記述であると考えられる。この判断は要約を作成する各人に委ねられるが、情報信憑性判断のための要約において特に重要な情報になると考えられる情報として、たとえば根拠のような言論間の相対的な位置づけを明らかにする論理的関係を表す情報が挙げられる。情報信憑性判断のための要約では、たとえば同じ内容の意見や情報でもその根拠が明示されている場合とそうでない場合とでは、その情報の信憑性の判断のしやすさが大きく違うと考えられるからである。

また、人間が情報信憑性判断のための要約を作成する際には、同じ内容の言論であってもそのトピックにかかわりの深い個人や団体の発信した言論が重要であると考えられる。また矛盾した言論のどちらを信頼するかという判断を人間が行う際にも、言論を発信した個人や団体の情報は重要になると考えられる。本実験では言論を発信している個人や団体の名称を情報発信者名とし、情報発信者が情報信憑性判断のための要約において重要な役割を持つと考えて調査を行う。

#### 3.2 要約作成過程を観察するための実験の設計

前節に挙げた事を踏まえて、本実験における、人手による情報信憑性判断のための要約を作成する作業を設計する。

本実験の要約作成は、次の4つのトピックを対象としてそれぞれのトピックに対する作業を複数の作業員が行う。要約の対象としたトピック名およびトピックごとの作業員数は、以下の表1の通りである。

表 1: トピック名と作業者数

トピック名	作業者数	
	前半	後半
「還元水」	5	2
「裁判員制度」	4	2
「メタボリックシンドローム」	4	2
「女性専用車両」	4	2

表 2: 検索クエリの実例

トピック名:「還元水」	
検索クエリ	検索文書件数
『還元水』	500
『還元水 効果』	250
『還元水 用途』	250
合計	1,000

本来ならば一人の作業者が全作業を通してひとつの要約を作成するべきであるが、作業量と時間的な制約によりそれが一部実現できなかった。作業者が文書単位で選別を行うまでの作業を前半、それ以降を後半として、前半だけに参加した作業者、後半だけに参加した作業者、及びどちらにも参加した作業者が存在する。

なお、トピック名「メタボリックシンドローム」は以下の表中では「メタボリック」と簡略に示す。

まず、要約作成作業に必要な背景知識は、要約対象となる文書の収集の前に web から獲得し、獲得した背景知識を元にして収集する文書を決定できるようにする。

要約対象となる文書の収集は、検索エンジン基盤 TSUBAKI[3] を用いて作業者が行い、収集する文書の数は本実験では 1,000 件とする。収集する 1,000 件のうち 500 件は、収集する文書に一般性を持たせるために要約対象となるトピック名そのものを検索クエリとして収集する。しかし、文書の収集は作業者が行う最初の絞り込みであると考えられるので、残りの 500 件の検索に使用する検索クエリは、獲得した背景知識に基づいて作業者が決定できるものとする。その際に使用する検索クエリは、要約対象となるトピック名そのものに作業者が決定した語句を加えたものとし、複数の検索クエリを使用しても良い。その場合は、収集する文書の合計が 500 件になるように、作業者が自由に検索クエリごとの収集件数を決定する。検索クエリの実例を表 2 に示す。

次に、作業者は検索エンジンによって収集された文書から、要約に必要な文書を選別し、その文書から要約に必要な記述を抜き出す。要約を作成する際には、要約の対象となる文書が与えられるのが一般的であるが、2 章で述べたように本実験では情報信憑性判断のための要約を作成するために、要約の対象となる文書の収集から作業を開始している。しか

し、これ以降の要約過程について一般的な文書要約としての分析も行えるように、本実験では文書単位の選別の結果をトピックごとにマージして、作業者の判断が必要だということによって一致した文書について実際に要約に必要な記述を抜き出す作業を行う。本来ならば、作成される要約の網羅性が高くなる各作業者の文書単位の選別結果の和集合から、段階的に扱う範囲を絞っていき検討を行うべきであるが、本実験では各作業者の文書単位の選別結果の積集合について、要約に必要な記述を抜き出す作業を行う。

次に、情報発信者は情報信憑性判断のための要約において重要だと考えられるので、作業者は文書の著者を情報発信者として、その名称を調査する。情報発信者の名称は文書内の記述の一部を参照して作業者が決定できるものとして、参照した箇所と情報発信者の名称を文書ごとに記録する。また、本来ならば文書中に他の発信者の言論が引用される場合には、引用の情報発信者と文書の情報発信者を別のものとして扱うべきである。しかし、文書中の引用は必ずしもその引用元である発信者についての記述があるとは限らず、作業者がそれを調査することは非常に時間が掛かることがあるので、本実験では文書の情報発信者だけを考慮するように統一する。

最後に、抜き出された重要な記述を参考にして作業者が自由に情報信憑性判断のための要約を作成し、要約中に用いた記述がどの文書のどの文を参考にしたものかという情報を付与する。

また、絞込みが文書の収集、文書単位での選別、要約に必要な記述の抽出の三回行われるので、それらについての結果を保存して分析できるようにする。

## 4 情報信憑性判断のための人手による要約作成実験

本章では、人手による情報信憑性判断のための要約を作成する作業の手順について述べる。

要約の作成作業は、以下の手順に従って行った。

### 1. 背景知識の獲得

対象トピックに関する知識や概念を、web から収集する。概念の収集は、検索エンジンなどを用いて作業者各人が行う。

また、この段階で獲得した背景知識を記録しておき、後の作業に役立てる。

### 2. 文書の収集

対象トピックに関する文書を、検索エンジン基盤 TSUBAKI を用いて収集する。

収集に使用する検索クエリは前節で述べたとおりとする。検索クエリは、対象トピック名としたものを共通クエリとし、それ以外の検索クエリを個別クエリとする。文書はそれぞれ文書 ID を持ち、以降の文書管理は文書 ID を用いて行う。

ここで収集された文書集合を、作業者ごとの収集文書集合とする。

### 3. 文書の選別

作業者ごとの収集文書集合に含まれる文書を対象に、要約の作成に必要な記述を含む文書を選別する。

ここで選別された文書集合を、作業者ごとの選別文書集合とする。

### 4. 要約対象の文書の決定

作業者ごとの選別文書集合の積集合を要約対象文書集合とし、和集合を全選別文書集合とする。また、作業者ごとの収集文書集合の和集合を、全収集文書集合とする。

本実験ではここまでの作業を前半とし、これよりの作業を後半とする。

### 5. 情報発信者の名称の抽出

要約対象文書集合に含まれる文書について、文書の著者である情報発信者の名称を抽出する。

情報発信者の名称は情報発信者の名称について記述された文書内の一部分を参照して作業者が決定できるものとし、情報発信者の名称と参照した部分を文書ごとに記録する。また、文書内に発信者の名称についての記述が複数ある場合は、その全てを情報発信者の要素とする。

文書内に発信者についての記述がない場合は、発信者の名称は文書の URL を手がかりとして関連する文書を Web から探し出して発信者の名称を決定する。文書の URL を手がかりとした関連する文書の探索は、

- URL に示されるディレクトリ構造を判断し、末尾のファイルやディレクトリを順次削除して得られたドメイン内のページに移動して調査する
- ページ内のハイパーリンクを利用して移動した先のページを調査する

を行う。リンク切れなどの要因で、この探索でも発信者の名称が決定できなかった場合は、発信者は不明とする。

### 6. 重要な記述の抽出

要約対象文書集合に含まれる文書を対象に、作業者が要約の作成に必要なと考えた記述を抽出する。抽出は文単位で行い、抽出された文の集合を重要文集合とする。また文は複数文書に亘って識別される文 ID を持ち、以降の文管理は文 ID を用いて行う。

### 7. 要約の作成

対象トピックに関する要約を作成する。作成は次の段階を踏んで行う。

はじめに、重要文集合を参照しながら、形式を指定しないプレテンテキストで要約を作成する。次に、作業者が作成した要約とできるだけ同じものになるような抜粋要約を、同じ作業者が重要文集合から作成する。ここで、作成した要約のある記述と対応する抜粋要約がどの文書のどの文であったかの情報を、抜粋の参照元とする。

### 8. 作業者へのアンケート調査

最後に、要約の作成に関して留意した点を要約とは別に作業者に記録してもらう。

また、作業の際に作業者に提示したガイドラインを以下に示す。

- 文書の選別では、収集された文書の内容を見て文書単位で要約の作成に必要であるか、不要であるかをに分類する。
- 文書の選別では、文書の内容が既出・未出に関らず、要約の作成に必要な含まれるかを文書内の記述から判断する。
- 情報発信者は、その文書の著者だけを考慮する。文書中に引用があり、それが重要であるとして抜き出す場合、引用であると判断できる部分までを抜き出す。
- 重要な記述の抽出では、同じような、もしくはまったく同じ記述であっても要約に必要な記述は全て抽出を行う。
- 本作業では、出力とする要約に対して文字数や記述形式の制限を行わない。
- 要約の作成では、情報の網羅性を考慮しながら冗長性を排除したものを作成する。
- 抜粋要約の作成では、重要文集合に含まれる記述を並べ替え、要約に不用な部分を削除し、作業者が自由に語句を追加できるものとする。
- 要約に複数の文書に含まれる記述を用いた場合は、最も適当であると思われるひとつをその参照元とする。

## 5 要約作成作業の結果と考察

本章では、要約の作成作業の結果を示すとともに、その結果に対しての考察を行う。

### 5.1 文書の収集と選別

文書の収集の段階では、検索クエリは基本的には対象トピック名そのもの及び、対象トピック名に自由に語句を追加したものであると作業者に説明したが、たとえば「裁判員制度」のトピックでの『死刑制度』のように、作業者が決定した語句だけを単独で検索クエリとした作業者も存在した。

表 3: 作業者ごとの収集文書集合数

トピック名	最低	最高	平均
「還元水」	488	795	638
「裁判員制度」	515	748	623
「メタボリック」	644	853	768
「女性専用車両」	635	744	668

表 4: トピックごとの文書集合数

トピック名	全収集 文書集合	全選別 文書集合	要約対象 文書集合
「還元水」	1,248	504(273)	41(36)
「裁判員制度」	1,095	555(307)	87(84)
「メタボリック」	1,498	803(342)	143(130)
「女性専用車両」	1,090	654(363)	161(147)

括弧内は共通クエリによって収集された数

作業者ごとの収集文書集合は 500~ 650 程度となり、収集した 1,000 に対してかなり小さくなった。これは、類似した検索クエリで収集した文書には同一の文書が多く含まれたためだと考えられる。特に、自由に語句を追加した検索クエリの種類が多いと、同一の文書が複数回出現する可能性が高くなり、結果として作業者ごとの収集文書集合が小さくなる事があった。各トピックごとの作業者ごとの収集文書集合の最低値と最高値、及び平均値を表 3 に示す。

また、全収集文書集合は 1,000 から 1,500 程度となり、作業者ごとに検索クエリに自由に語句を追加した検索で収集を行った場合でも、全体として追加された文書はそれほど多くないことがわかった。そこで、選別された文書を共通検索クエリによって収集されたものと、個別クエリによる検索によって収集されたものに分類した。ただし、どちらにも含まれる文書は、前者に属するものとした。全選別文書集合は 500 から 800 程度で、全収集文書集合のおよそ半分程度となった。共通クエリによって選別された文書は、300 から 400 程度であり、全選別文書集合のさらに半分程度となった。また、要約対象文書集合は 40 から 160 程度となり、トピックによって数倍の差があった。要約対象文書集合のうち、共通クエリによって収集された文書は、どのトピックでも 9 割程度となり、かなり高い結果となった。

トピックごとの全収集文書集合、全選別文書集合、要約対象文書集合の数を表 4 に示し、共通クエリによって収集されたものを、表中の括弧内の数字として表す。

作業者ごとの文書の選別、およびその結果の積集合である要約対象文書集合の決定では、収集文書集合約 1,000 に対して要約対象文書集合は 100 程度であり、およそ一割に絞り込みが進んだということになる。個別クエリによって収集された文書に含まれる要約対象文書は全体の二割程度であることから、個別クエリの追加は要約の作成にあまり効果は

表 5: 文書の選別における一致率 1( $\kappa$  値)

トピック名	収集文書集合中		
	最低	最高	平均
「還元水」	0.25	0.65	0.39
「裁判員制度」	0.24	0.61	0.39
「メタボリック」	0.27	0.53	0.40
「女性専用車両」	0.41	0.54	0.48

表 6: 文書の選別における一致率 2( $\kappa$  値)

トピック名	共通クエリによって 収集された文書中		
	最低	最高	平均
「還元水」	0.38	0.76	0.49
「裁判員制度」	0.37	0.66	0.49
「メタボリック」	0.38	0.62	0.49
「女性専用車両」	0.36	0.70	0.49

なかった可能性がある。

次に、作業者ごとの文書の選別の一致について述べる。文書の選別の段階は要約の対象となる文書を決定する段階であるので、3章で述べたように、自動要約を実現するためには作業者ごとの選別に揺れが少ないことが望ましい。

ここでは、選別の一致率を  $\kappa$  値で表す。 $\kappa$  値は主観に拠る判定が偶然ではなく一致する割合であり、0.4 から 0.6 の間ならば中程度の一致、0.8 を超えるとはほぼ完璧な一致と考えられる。一致率は二人の作業者間について計算し、各トピックについて最も一致度が低かった組と最も一致度が高かった組の値および全組の平均の値を表 5 に示す。

結果を見ると、女性専用車両では  $\kappa$  値が最低でも中程度の一致となっているが、他の 3 トピックでは  $\kappa$  値が最も低い場合には 0.25 付近となっている。これはあまり良い一致であるとは言えないが、平均値はいずれも 0.4 前後で、中程度の一致となっている。

また、表 4 から、要約対象文書集合は共通クエリによって収集された文書に集中しているの、共通クエリによって収集された文書だけを対象として選別の一致度を計算しなおしたものを表 6 に示す。

結果を見ると、どのトピックにおいても最も低い  $\kappa$  値が中程度の一致と考えられる 0.4 付近となっている。従って、要約対象文書を作業者ごとの選別文書集合の積集合とするならば、共通クエリだけで文書の収集を行うことによって要約対象文書集合をあまり減らさずに、選別の結果を中程度の一致と考えることができる。

## 5.2 情報発信者の名称の抽出

本実験では、全トピックの要約対象文書集合 498 に対して情報発信者が不明であったものは 71 であ

表 7: 文抽出の一致率

トピック名	一致率	要約対象 文書集合	含まれる 文の合計
「還元水」	0.69	36	3,044
「裁判員」	0.67	87	6,288
「メタボリック」	0.56	143	13,916
「女性専用車両」	0.26	161	35,941

り、約 14%であった。

### 5.3 重要な記述の抽出

要約対象文書集合を対象に行われた、重要な記述の抽出についての一致率を  $\kappa$  値で表 7 に示す。ここでは作業者は各トピックごとに 2 名で、4 章で述べたように重要な記述の抽出は文単位で行われている。また、本実験では意味内容が同一な文同士であっても、文 ID が違う場合は違うものとみなして一致率を計算しているため、意味内容が同一な文間の一致まで考慮した場合には、一致率の値はさらに高くなる。

結果を見ると、「女性専用車両」以外のトピックは中程度以上の一致となっているが、「女性専用車両」では一致率が他のトピックに比べて著しく低い。この原因として、「女性専用車両」は要約対象文書集合が大きいことに加えてひとつの文書あたりの文数が他のトピックに比べて 2 倍程度多く、作業者間の判断が一致しないことが多くなる可能性がある。また、「女性専用車両」の要約対象文書集合だけに含まれた文書の種類として大規模な掲示板があり、この文書における判断が難しかったという可能性がある。さらに本実験では意味内容が同一の文をマージしておらず、全て別のものとして扱っているために、この一致率は実際よりも低下しているものと考えられ、特に文が多い「女性専用車両」ではこの低下が他のトピックよりも大きかった可能性がある。

### 5.4 要約の作成

本稿で行った作業では、要約の作成について記述形式や文字数を制限していない。これは人間が要約として適当だと考えるものがどのようなものであるかを分析するためである。

本作業ではまず人手で要約を作成しているため、作成された要約の文字数を表 8 に示す。表中で形式が「箇条」となっているものは、要約の記述形式が箇条書きであったものである。また、それ以外の形式は全て作業者が文章を書き起こして要約を作成したものであり、表中では「書起」として示す。この結果から、人間が箇条書きではない形式で要約を作成すると、1,000 から 3,000 文字程度になるということがわかった。

抜粋要約の作成では、要約の作成の際に箇条書き以外の記述形式を選択した作業者は、最初に作成し

表 8: 要約の文字数

トピック名	要約 1		要約 2	
	形式	文字数	形式	文字数
「還元水」	書起	987	書起	1,824
「裁判員」	書起	812	書起	2,616
「メタボリック」	箇条	238	書起	1,727
「女性専用車両」	箇条	321	書起	638

表 9: 抜粋要約の文字数

トピック名	要約 1		要約 2	
	文字数	参照	文字数	参照
「還元水」	1,336	5	1,701	5
「裁判員」	913	8	2,003	19
「メタボリック」	5,244	12	1,644	17
「女性専用車両」	2,076	3	2,171	7

た要約と抜粋要約とは完全に同じものとなっていないが、抜粋要約の文字数は、最初に作成した要約とあまり変わらなかった。一方、箇条書きで要約を作成した二人の作業者は、抜粋要約を作成する際に具体的な記述を書き足していた。本来の作業は、4 章で述べたように人手で最初から要約を書き起こし、その後でその要約と同じものになるように抜粋要約を作成することとしていたが、単に自由な形式で要約を作成するという説明を行ったために、このように参照元の情報を付与する際に大幅に修正することを前提とした要約を作成したものと考えられる。抜粋要約の文字数および抜粋の参照元となった文書の数を表 9 に示す。箇条書きから大幅に修正を行って作成した抜粋要約は、最初から文章を書き起こす場合に比べて、文字数が大きく増加し、他の抜粋要約の長さと比較しても比較的長いものとなった。これは作業者が選択した要約の作成についての考え方の違いだと思われるが、具体的な考え方については調査できなかった。

また、抜粋の参照元とされた文書の数は、最低で 3 文書、最高で 19 文書となり、要約対象文書集合よりもかなり少なくなった。これは、要約対象文書から抜き出された重要な記述に冗長なものが多く含まれるためであると考えられる。また、同じ文書に複数の参照元が含まれていることも多かった。

最後に、要約の参照元とされた文書のうち、情報発信者が不明であるものは 16 であり、約 18%であった。これは要約対象文書集合全体の情報発信者が不明である割合とほぼ同じであった。

### 5.5 作業へのアンケート調査

最後に、作業者に要約の作成に際して留意した点のアンケート調査では、以下のような回答が得られた。

- 情報発信者がどのような個人、団体であるかということ
- 内容がひとつの文書にまとまっていること
- 読みやすく具体的に記述されていること

従って、参照元とされた文書数が要約対象文書よりも少ない原因として、冗長を排除した際に選択されなかった文書の存在したということだけではなく、まとまった記述が存在する文書を優先して使用したことも重要だと考えられる。また、最初から文章を書き起こした場合は比較的短い要約が多かったことから、要約として読みやすい長さを考えた際に排除された文書が存在したということも原因として考えられる。

## 6 おわりに

本稿では、情報信憑性の判断に役立つ要約の作成を複数の作業員で行い、得られた要約文章の安定性や複数の作業員間の一致度を分析した。

複数の作業員で作業を行う際には、文書の選別の結果を共有することにより、文抽出では中程度の一致を期待できることがわかった。何段階かの絞り込みの結果、要約に使用された文書は収集した文書の1%程度であり、作成された要約はおよそ1,000から3,000文字であった。作業員へのインタビューから、要約に使用された文書は、必要な記述が多く含まれたものを代表として使用することが多く、また情報信憑性判断に重要な情報としては、情報発信者や記述の具体性が重要であるということがわかった。

最後に情報信憑性判断に役立つ要約を自動化するための準備として、人間が文章を書き起こして作成した要約と、要約対象となる文章からの抜粋を中心として作成した要約とを関連付けておくことで、自動要約の際の「正解」として利用できるようにした。本実験では、人間が作成した要約と、それと同じものになるように重要文からの抜粋で作成された要約を作成し、抜粋の参照元の情報が保持されている。

今後は要約の作成に参加する人数を増やし、多段階の絞り込みにおける各段階で情報を共有することにより出力がより安定するかどうかの実験を行いたい。また、意味内容が同じものである文のマージを行い、より正確な文抽出の一致率を計算して、検討を行いたい。

## 謝辞

本研究は、独立行政法人情報通信研究機構の委託研究「電気通信サービスにおける情報信憑性検証技術に関する研究開発」プロジェクトの成果である。

## 参考文献

- [1] D. Radev. A common theory of information fusion from multiple text sources step one: Cross-document structure. In *Proceedings of the 1st SIGdial workshop on Discourse and Dialogue*, 2000.
- [2] D. Radev, J. Otterbacher, and Z. Zhang. Cst bank: A corpus for the study of cross-document structural relationships. In *Proceedings of the 4th International Language Resource and Evaluation (LREC'04)*, 2004.
- [3] Keiji Shinzato, Tomohide Shibata, Daisuke Kawahara, Chikara Hashimoto, and Sadao Kurohashi. TSUBAKI: An open search engine infrastructure for developing new information access methodology. In *Proc. of IJCNLP2008*, 2008.
- [4] 難波英嗣. 情報抽出を利用した複数文書要約. 知能と情報, Vol. 18, No. 5, pp. 16-22, 2006.
- [5] 赤峯亨, 宮森恒, 加藤義清, 中川哲治, 乾健太郎, 黒橋禎夫, 木俣豊. Web 情報の信頼性検証のための情報分析システム WISDOM. 言語処理学会第 14 回年次大会発表論文集, pp. 721-724, 3 月 2008.
- [6] 村上浩司, 松吉俊, 隅田飛鳥, 森田啓, 佐尾ちとせ, 増田祥子, 松本裕治, 乾健太郎. 言論マップ生成課題: 言説間の類似・対立の構造を捉えるために. 情報処理学会研究報告, 自然言語処理・言語理解とコミュニケーション合同研究会, 信学技報 Vol.108 No.141, 2008-NL-186, pp. 55-60, 7 月 2008.
- [7] 吉岡真治, 原口誠. イベントの参照関係に注目した新聞記事の複数文書要約. 言語処理学会第 10 回年次大会発表論文集, pp. 257-260, 3 月 2004.
- [8] 河原大輔, 黒橋禎夫, 乾健太郎. 主要・対立表現の俯瞰的把握 - ウェブの情報信頼性分析に向けて. 情報処理学会研究報告, 自然言語処理・言語理解とコミュニケーション合同研究会, 信学技報 Vol.108 No.141, 2008-NL-186, pp. 43-54, 7 月 2008.
- [9] 平尾努, 奥村学, 福島孝博, 難波英嗣, 野畑周. 抜粋による複数文書要約を評価するためのコーパスと評価指標. 情報処理学会論文誌データベース, Vol. 48, No. SIG 14, pp. 60-68, 2007.