

知識ゴミは知的資源となり得るか？ ～棄却文章断片の活用環境構築に向けて～

生田泰章^{†1} 高島健太郎^{†1} 西本一志^{†1}

概要：これまで、知識を効率的に活用する研究が数多くなされてきた。これらの研究が対象とする知識は、有用であるという判断のもとで形式化された知識である。一方、これらの知識が創造される過程では試行錯誤が行われることが一般的であり、この過程において不用と判断された知識断片は知識ゴミとして棄却され、活用対象としてみなされない。しかしながら、知識ゴミは、ある知識の創造において不用と判断されただけであり、その他の知識創造においても一律に不用と判断されるのは適切でない。筆者らはこれまで、知識ゴミの活用可能性に着目し、文章作成過程において知識ゴミとして棄却された文章断片（棄却文章断片）を活用すべく研究を行ってきた。具体的には、棄却文章断片を活用可能性の高いもの（R-DTF）と低いもの（F-DTF）とで分別収集可能な文章作成支援システム Text ComposTer を開発した。本稿では、Text ComposTer で収集した R-DTF が新たな文章作成において実際に活用されるかどうかを検証し、その検証結果を踏まえて棄却文章断片を知的資源として活用するための環境構築について検討する。

Can Knowledge Trash be Utilized as Knowledge Resource? - Toward Utilizing Deleted Text Fragments -

HIROAKI IKUTA^{†1} KENTARO TAKASHIMA^{†1} KAZUSHI NISHIMOTO^{†1}

Abstract: There have been many researches to efficiently utilize knowledge so far. These focused on knowledge that is formalized and is determined useful. On the other hand, it is common that trial and error is carried out in the process of creating this knowledge. During this process, knowledge fragments judged useless are discarded as knowledge trash and are paid no attention. However, knowledge trash is only judged useless in the creation of certain knowledge. It is not appropriate to judge that it is also unnecessary in other knowledge creation. The authors have focused on the possibility of utilizing knowledge trash and have studied to utilize Deleted Text Fragments (DTFs) which are text fragments deleted as knowledge trash in document writing process. We developed a document composition support system named “Text ComposTer,” which can separately collect R-DTFs with high possibility to be utilized and F-DTFs with comparatively low possibility to be utilized. In this paper, we investigate whether R-DTF collected by Text ComposTer is actually used in new document creation, and discuss the design of environment to utilize the DTFs as intellectual resource based on the experimental results.

1. はじめに

情報技術の発展・普及に伴い知識創造社会を迎えつつある現在において、新たな知識の創造と、創造された知識の活用を促進することは非常に重要な課題である[1]。これまで、エキスパートシステムや検索・推薦システムなど、知識の活用を促進する研究が活発に行われてきた。これらの研究では、主たる活用対象は、創造された完成形の知識であった。例えば、これまで数多くのエキスパートシステムが提案されており[2]、様々な分野における専門家の有用な知識の活用が試みられている。

一方、従来、創造活動支援システムや発想支援システムなど、新たな知識の創造を促進する研究が数多く行われてきた。これらの研究は、新たな知識を自動的に生成するというアプローチではなく、人間の知識創造における試行錯誤のプロセスを支援するアプローチをとっているものが多い。この試行錯誤のプロセスにおいていったんは創造されたものの最終的には不用と判断される知識断片（例えば、

不用と判断されたアイデア）が数多く発生する。これらは、一般的に知識ゴミとして棄却され、最終的にアクセス不能な状態となる。

しかしながら、上述の知識ゴミは本当に不用として棄却されてよいのであろうか。知識ゴミは、ある知識を創造するためには不用と判断されたが、別の新たな知識を創造するためには有用である可能性がある。実際、ポスト・イット®の開発や本稿第3著者によるドラム奏者のトレーニング支援システムに関する研究では、他の知識創造における知識ゴミが有用な知識として活用されている[3]。

このような背景から、筆者らはこれまで、知識ゴミの活用可能性に着目し、文章作成過程において知識ゴミとして棄却された文章断片（棄却文章断片）を活用すべく研究を行ってきた[3][4][5]。これらの研究の中で、DTF（Deleted Text Fragment）を粒度別に分別して収集することが可能な Text ComposTer を開発し、収集された棄却文章断片のうち、意味的内容を多く含む粗粒度の DTF（R-DTF：Rough-grain DTF）の活用可能性が高いという示唆を得ている。しかしながら、R-DTF が新たな文章作成に実際に活用されるかどうかの検証は未実施であった。そこで、本稿では、新たな

^{†1} 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

創造的文章作成において R-DTF が実際に活用されるかどうかの検証実験を行い、その実験結果を踏まえ、DTF が有効に活用可能であるかどうかの検討を行う。さらに、その実験結果を踏まえ、棄却文章断片を知的資源として活用するための環境構築について検討する。

2. Text ComposTer

2.1 システム概要

図 1 は、本稿の検証実験で使用する Text ComposTer の操作画面である。Text ComposTer は、文章断片が記入されたカード状の要素の配置位置にしたがって本文全体を形成することにより、執筆者の文章作成を支援するシステムである。Text ComposTer は、本文全体を表示する表示領域と、要素の配置が可能な配置領域を備える。また、配置領域は、反映領域と非反映領域を有しており、反映領域に配置された要素内の文章断片は表示領域に反映され、非反映領域に配置された要素内の文章断片は表示領域に反映されない。ここで、表示領域に反映される文章断片の順序は、Nakakoji らの Art#001[6]と同様に、反映領域の上下方向における要素の配置位置に対応している。すなわち、反映領域の最上部に配置された要素内に記入された文章断片が表示領域の最先に表示され、反映領域の最下部に配置された要素内に記入された文章断片が表示領域の最後に表示される。ユーザは、要素の生成、要素内への文章断片の記入、反映領域または非反映領域への要素の配置を行うことによって、本文全体を作成する。

Text ComposTer は、以下の各機能を有している。

- (1) Generate 機能：要素の生成
- (2) Split 機能：要素の分割
- (3) Save 機能：操作画面内の作業環境の保存
- (4) Load 機能：Save 機能で保存された作業環境の再現
- (5) Done 機能：表示領域内のテキストを外部ファイルとして出力（作業完了時等に行われる機能）

図 1 の要素は、メモ欄と本文欄の 2 つのテキスト入力可能な領域を有している。本文欄に記入されたテキストの内容は、要素の配置位置によって表示領域への反映・非反映を変化させる。メモ欄への記入内容は、要素が反映領域、非反映領域のいずれに配置されていたとしても、表示領域に反映されない。メモ欄と本文欄を 1 つの要素に設けることによって、ユーザがメモ欄に文章作成のためのアイデア・キーワードを記入し、そのアイデア・キーワードを基にした本文を本文欄に記入する



図 1. Text ComposTer の操作画面

Fig. 1. Snapshot of Text ComposTer

とを可能とする。

2.2 R-DTF とその特徴

Text ComposTer は、粗粒度と細粒度の 2 つの粒度別に棄却文章断片を収集する機能を有している。粗粒度の棄却文章断片 (R-DTF) は、Done 機能が実行されたときに非反映領域内に配置された要素それぞれに記入されている文章断片である。細粒度の棄却文章断片 (F-DTF: Fine-grain DTF) は、各要素内の文章断片の編集集中に削除された文字列のことである。

前述のような GUI および機能を備える Text ComposTer は、一般的なテキストエディタのように単に文章作成過程の最終状態を表示するだけでなく、文章作成過程全体を支援する機能を持つ。Text ComposTer を用いることで、文章作成の上流工程（文章の構想・構成段階）で創造されたものの、最終的に本文に採用されなかった棄却文章断片を R-DTF として収集し、文章そのものに対する誤字の訂正や表現の修正等の編集操作によって生じる棄却文章断片を F-DTF として分別収集することができる。

これまでの被験者実験から、R-DTF は、文章構成においていったんは書かれたものの最終的に削除された文章塊であることが多く、後に活用される可能性が F-DTF に比べて大幅に高いという結果を得ている[4]。一方、Text ComposTer の使用方法によっては、R-DTF の一部に、本文に採用された文章断片の内容とほとんど同じ内容が含まれる場合があるという示唆も得ている[4]。

このように R-DTF の特徴を知るための被験者実験を行ってきたが、新たな創造的文章の作成時に、実際に R-DTF が活用されるかどうかの検証は未実施であった。そこで、本稿では 3 章および 4 章で示す実験を行うことで R-DTF の

活用可能性を検証する。具体的には、R-DTF を収集するための事前実験（3章参照）を経た後、収集された R-DTF が新たな文章作成時に活用されるかどうかを検証する活用実験（4章参照）を行う。

3. 事前実験

3.1 実験設定

4名の被験者（S1～S4）それぞれに図1の操作画面を有する Text ComposTer を用いて文章作成を行ってもらい、R-DTF を収集する実験を行った。被験者はすべて、筆者らの所属する大学院の修士課程の学生であり、以下の4つのテーマについて文章作成を行ってもらった。本稿においては、作成される文章が執筆者のアイデアが含まれるものとなるように各テーマを設定した。

テーマ1：10年後の公園がどのようになっているか。

テーマ2：50年後の公園がどのようになっているか。

テーマ3：未来のレストランがどのようになっているか。

テーマ4：未来のファッションがどのようになっているか。

また、各被験者には1日に1テーマの文章作成を行ってもらった。このとき、カウンターバランスを考慮して各被験者にテーマを割り振った。文章作成時間は30分程度に設定し、各文章の作成文字数を100字以上400字以内となるように被験者に教示した。

以上の設定のもとで被験者それぞれが各テーマに関する文章を執筆した後、本文に関する情報および R-DTF に関する情報を収集した。具体的には、反映領域に配置されたエレメントの情報を本文に関する情報として収集し、非反映領域に配置されたエレメントの情報を R-DTF に関する情報として収集した。ここで、上述したように、収集した R-DTF の中には、本文に含まれる文章断片とほぼ同内容のものが含まれている場合がある。そこで、本稿では、本文 d に対する R-DTF の文字列 s の含有率 $R_{s,d}$ を下記式によって算出し、 $R_{s,d}$ が閾値未満の R-DTF を真の R-DTF、 $R_{s,d}$ が閾値以上の割合の R-DTF を偽の R-DTF として区別する。なお、R-DTF の文字列 s は、本文 d の執筆中に生成されたものとする。

$$R_{s,d} = \frac{|LCS(s,d)|}{|s|}$$

ここに $|LCS(s,d)|$ は、R-DTF の文字列 s と本文 d の最長共通部分列（LCS: Longest Common Subsequence）の長さであり、次式で定義される[7]。

$$|LCS(s,d)| = \max_{\forall \vec{i}, \forall \vec{j}: s[\vec{i}] = d[\vec{j}]} |\vec{i}|$$

ただし、 \vec{i} は、R-DTF の文字列長 $|s|$ 以下の長さのインデックスベクトルであり、 $s[\vec{i}]$ は、R-DTF の部分列を表す。また、

表1. 収集された R-DTF の数。

Table 1. Number of collected R-DTFs

被験者	テーマ1 (真/偽)	テーマ2 (真/偽)	テーマ3 (真/偽)	テーマ4 (真/偽)	計
S1	2(1/1)	3(2/1)	1(0/1)	0(0/0)	6(3/3)
S2	4(4/0)	4(4/0)	7(7/0)	6(4/2)	21(19/2)
S3	8(5/3)	5(3/2)	4(1/3)	11(9/2)	28(18/10)
S4	11(5/6)	5(5/0)	11(10/1)	9(5/4)	36(25/11)
計	25(15/10)	17(14/3)	23(18/5)	26(18/8)	91(65/26)

\vec{j} は、本文の文字列長 $|d|$ 以下の長さのインデックスベクトルであり、 $d[\vec{j}]$ は、本文 d の部分列を表す。なお、部分列は、記号列に対して順序を保持した部分的な記号列を表し、隣接性は問わない[7]。また、本実験では、閾値を 0.7 として R-DTF の真偽を区別した。

3.2 実験結果

表1は、非反映領域に配置されたエレメントの数、つまり R-DTF の数を被験者およびテーマ別にまとめた表である。また、表1には、真の R-DTF および偽の R-DTF の内訳も含まれている。表1に示されたように、計 91 個の R-DTF が収集され、そのうち真の R-DTF は 65 個であった。また、テーマ間で R-DTF の数に統計的な傾向はなかったが、被験者間においては、Tukey 法による多重検定を実施したところ、被験者 S1 と S3 および、被験者 S1 と S4 の間で収集された R-DTF の数に有意差があった。

4. 活用実験

本章では、3章で収集した R-DTF が新たな文章作成時に実際に活用されるかどうかの検証実験について説明する。活用実験では、R-DTF を活用する主体は、事前実験の被験者とは異なる被験者を対象とする。事前実験の被験者は、当該実験にて不用と判断した文章断片を R-DTF として棄却している。事前実験とは異なる新たな文章を作成する場合であったとしても、その被験者が過去に R-DTF を不用なものであるとして棄却した記憶がまだ強く残っている状態では、その R-DTF が活用される可能性が低いことが考えられる。そのため、事前実験の被験者を活用主体とする場合、事前実験の実施後から十分な期間が経過後、活用実験を実施する必要があると思われる。そのため、本稿では、まず R-DTF を生成した本人以外が活用主体となるかどうかを検証する実験を行う。

4.1 実験設定

活用実験においては、事前実験の被験者とは異なる4名の被験者に収集実験のテーマ2について文章作成を行ってもらう実験を行った。被験者は、筆者らの所属する大学院の修士課程の学生であり、収集実験で採用した被験者とは異なる。このとき、R-DTF が参照可能な状態で文章作成を行う。具体的には、被験者には、図2の操作画面を用いて文章作成を行ってもらった。操作画面には、R-DTF リスト

と編集領域が設けられている。R-DTF リストには、事前実験で収集された 91 個の R-DTF が順に記載されており、各 R-DTF には識別番号とチェックボックスが付されている。被験者は、R-DTF リスト内の R-DTF を参照しながら編集領域に文章を記入することができる。被験者には、文章作成時に R-DTF を参考にしてもよい旨教示し、さらに R-DTF を参考にした場合にはその R-DTF に付されたチェックボックスにチェックを入れるように教示した。

本実験においては、文章作成時間を無制限に設定し、最低 1000 文字以上の文章を作成することを被験者に要求した。ただし、各被験者には、最大 3 時間分の謝金が生産時間に依りて支給される旨伝えられている。このような実験設定において、各被験者が文章作成を行う様子を実験用システムの画面を録画することで観察した。また、文章作成後の各被験者にインタビューを実施した。インタビューにおいて、どのように操作画面を使って文章作成を行ったかということと、どのような場面で R-DTF を参考にしたかということを中心に質問した。

4.2 実験結果

以上の実験設定において活用実験を行ったところ、被験者 1 は 1427 文字の文章を作成し、被験者 2 は 1050 文字の文章を作成した。また、被験者 3 は 1421 文字の文章を作成し、被験者 4 は 1366 文字の文章を作成した。各被験者が作成した文章の長さには大差なく、統計的な傾向も見られなかった。

4.2.1 文章作成プロセスについて

次に、活用実験における被験者の文章作成プロセスについて説明する。各被験者へのインタビュー結果と、実験中の操作画面を記録した動画の分析の結果から、すべての被験者が、本文を作成する前に構成を思案する様子が観察された。具体的には、各被験者は、最終的な本文の内容を書き始める前に、まず文章構成のためのメモ（構成メモと呼ぶ）の作成に取り掛かっていた。また、各被験者は構成メモの取り方にそれぞれ特徴があったため、本文の作成を開始するまでのプロセスを中心に、各被験者の文章作成プロセスの概要を説明する。

被験者 1 および被験者 2 は、実験開始後、まず構成メモの作成に取り掛かった。その際、適宜 R-DTF リストを参照しながら、構成メモの内容の拡充を行っていった。構成メモの完成後、被験者 1 および被験者 2 は本文の作成を開始した。被験者 3 は、まず R-DTF リストを参照しながら、テーマ 2 に関する自身の意見を構成メモとして箇条書きし、構成メモを書き終えた後、本文の作成をいったん開始した。



図 2. 活用実験の被験者が使用する操作画面

Fig. 2. Snapshot of the system for subjects in the experiment of R-DTFs utilization

しかしながら、本文作成開始後しばらくして、これまで作成した本文をすべて削除し、もう一度新たに本文の作成を開始した。このとき、新たな本文に記載のテーマ 2 に対する被験者 3 の意見は、削除された本文の内容に記載されていたものとは異なる内容であった。インタビューより、被験者 3 は R-DTF リストを参照しているうちにテーマ 2 に対する意見が変わったため、一から本文作成を開始したと実験を振り返っていた。このとき、被験者 3 は特定の R-DTF によって意見が変わったのではなく、複数の R-DTF に影響されたと述べていた。被験者 4 は、まず R-DTF リストを参照しながら気になったキーワードを構成メモとして抽出した。その後、抽出したキーワードを組み合わせる短文を複数生成し、自身がテーマ 2 について書き出したいことの探索を始めた。そして、その探索の完了後に本文作成を開始した。

4.2.2 文章作成時に参照された R-DTF について

次いで、各被験者が文章作成時に参照した R-DTF について説明する。図 3 および図 4 はそれぞれ、被験者が文章作成の際に活用した R-DTF の数を、事前実験の被験者別に集計した棒グラフと、テーマ別に集計した棒グラフである。また図 3 および図 4 は、真の R-DTF と偽の R-DTF を区別して集計してある。

図 3 に示すように、被験者は、少なくとも 3 つは真の R-DTF を活用していた。また、各被験者は、事前実験における被験者 S4 が生成した R-DTF を多く活用していた。被験者 S2 が生成した R-DTF のうちで活用されたものは、すべて真の R-DTF であった。さらに、事前実験で収集した真の R-DTF の割合は、約 71%であったのに対し、条件 3 にて活用された真の R-DTF の割合は約 72%であり、事前実験における R-DTF の真偽の割合と同様の割合にて R-DTF が活用されていることが分かった。

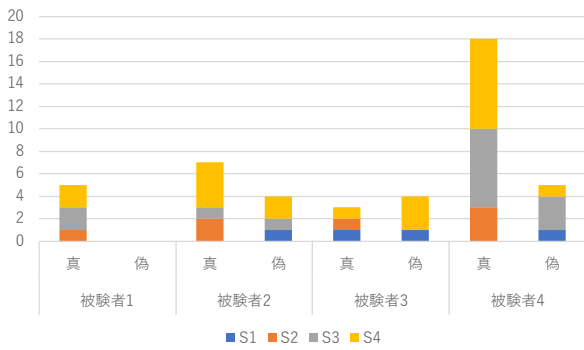


図 3. R-DTF の活用個数 (被験者別に集計).
Fig. 3. Number of utilized R-DTFs by subjects

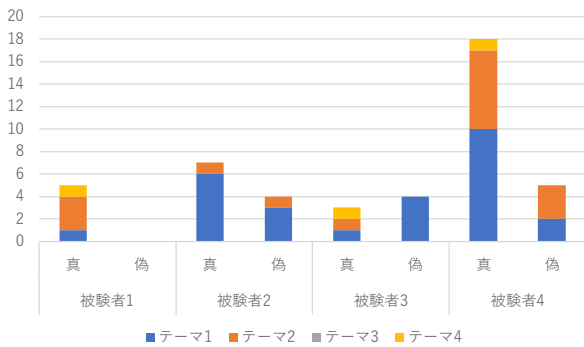


図 4. R-DTF の活用個数 (テーマ別に集計).
Fig. 4. Number of utilized R-DTFs by themes

表 2. 文章間類似度と活用した R-DTF の重複数の対応表.
Table 2. Correspondence table of the pair of similarity of documents and duplicate number of utilized R-DTFs

	被験者1	被験者2	被験者3	被験者4
被験者1	-	(0.319, 1)	(0.649, 0)	(0.397, 4)
被験者2	(0.319, 1)	-	(0.408, 2)	(0.415, 9)
被験者3	(0.649, 0)	(0.408, 2)	-	(0.407, 3)
被験者4	(0.397, 4)	(0.415, 9)	(0.407, 3)	-

ただし、(類似度, 重複数)を表す。

図 4 に示すように、事前実験におけるテーマ 1 とテーマ 2 の関する文章作成時に生成された R-DTF が多く活用されていた。テーマ 3 に関する文章作成時に生成された R-DTF は 1 つも活用されなかった。テーマ 4 に関する文章作成時に生成された R-DTF は、ほとんど活用されなかったが、すべて真の R-DTF が活用されていた。

表 2 は、各被験者が作成した文章間の類似度と、R-DTF の重複数の対応表である。表 2 に記載の類似度は、対応する 2 つの文章のコサイン類似度を計算したものである。具体的には、対応する 2 つの文章それぞれについて名詞、動詞、形容詞、未知語を Mecab[8]を用いて抽出し、各文章に対応する Bag-of-words を生成する。そして、この 2 つの Bag-of-words のコサイン類似度を計算することで、2 つの

文章間の類似度を算出する。なお、本稿では各文章で頻出する「公園」の文字列をストップワードとした。また、表 2 における R-DTF の重複数とは、被験者それぞれが活用した R-DTF のうち、被験者間で重複しているものの数を表す。

表 2 に示すように、被験者 1 が作成した文章と被験者 3 が作成した文章の類似度は約 0.65 と比較的高いが、R-DTF の重複数は 0 である。一方、表 2 に示すように被験者 2 と被験者 4 が活用した R-DTF の重複数は 9 であるにもかかわらず、類似度は 0.5 を下回っている。したがって、作成される文章は、R-DTF の活用に起因して似通うわけではない。また、Spearman の順位相関係数で類似度と重複数に相関があるかどうかを検証したところ、相関係数が約-0.086 で相関がないことが判明した。

また、インタビューの結果から、被験者 1 および被験者 2 は、自身の構成メモの作成に当たり欠けている視点を、R-DTF を活用することで補っていた。このとき、両者は主に R-DTF に記載の単語を補う視点として活用していた。さらに、被験者 1 および被験者 2 は、文字数が多すぎる R-DTF は、活用しづらいと回答した。また、被験者 2 は、文章作成時には自身の文章表現と同内容の R-DTF を探し、活用できないかどうかを検討していた。さらに、すべての被験者が、R-DTF を活用していたのは、主に構成メモを作成する段階であった。また、すべての被験者は、テーマ 2 とは明らかに異なるテーマ、すなわちテーマ 3 および 4 の R-DTF は参照時に活用可能性が低いとして読み飛ばしていたと回答した。このインタビュー結果は、上述の図 4 の実験結果とも整合する。

5. 議論

3 章の実験結果から、新たな創造的文章の作成時における R-DTF の活用可能性について議論する。まず、図 3 および図 4 に示したように、新たな創造的文章作成時において R-DTF は実際に活用された。活用実験の被験者は真の R-DTF についても偽の R-DTF についても同様に活用していることから、他者が本文に採用するのに有用と判断した文章断片 (偽の R-DTF) であろうが、不用と判断した文章断片であろうが (真の R-DTF)、分け隔てなく活用されることが分かった。

一方、R-DTF が活用されたからと言って、作成される文章の内容に関して、必ずしも特徴が出るわけではない。上述したように、文章間の類似度と R-DTF の重複数には相関関係がない。これは、R-DTF が活用される主なタイミングが構成メモを作成する段階であったことに起因すると考えられる。つまり、R-DTF を本文作成に当たって直接的に活用したわけではないため、R-DTF が有する内容が直接的には反映されない。

しかしながら、インタビューの結果から、R-DTF を参照した場合、執筆者の文章作成行動に影響を与えることが分

かった。具体的には、被験者3がR-DTFによって自身の意見を変化させ、被験者4がR-DTFを用いることによって自身の意見を探索していた。このように、被験者3および被験者4は、R-DTFの内容に自身の考えが影響されて文章作成を行っていることが判明した。

このような事象が生じた原因の一つに、参照する外部知識の液状化[9]の程度が関係すると思われる。液状化は、堀らの研究グループが提唱した概念[9]であり、知識がより大きく分節している状態を指す。本稿の事象で言えば、R-DTFは、文章断片であるため、完成形の文章に比べて知識が液状化している。したがって、R-DTFは、執筆者が有する知識と結合しやすいと思われる。つまり、知識を液状化した状態で提示されたR-DTFは、完成した文章を提示するよりも、より活用されやすい状態であると考えられる。

以上より、R-DTFは新たな創造的文章を作成するに当たり、有効に活用可能であると結論付けることができる。

次に、以上の実験結果および考察を踏まえて、棄却文章断片の活用環境の構築に向けて有効と思われる方法について議論する。まず、R-DTFは文章構成段階で活用されることが主であった。したがって、執筆者に効率的なR-DTFの活用を促すためには、Text ComposTerをはじめとする、新たな文章の構想・構成段階を支援可能な文章作成支援システムにR-DTFの提示機能を追加することが効果的であると思われる。また、文章作成時のテーマとかけ離れたテーマから生成されたR-DTFは、活用対象から活用されにくいことが、4章の実験より判明した。このようなR-DTFが活用されるためには、かけ離れたテーマであることを執筆者に気付かせないようにする処理が必要である。例えば、テーマ固有の単語(トピック)をR-DTFから除外することで、より有効に活用されるものと思われる。さらに、R-DTFを有効に活用可能とするためには、液状化の粒度を調整し、R-DTFと執筆者の知識が結晶化して新たな知識創造を行うことができるようにする方策を検討する必要があると思われる。

6. 関連研究

従来、生成されたコンテンツを効率的に再利用するためのContents Reuseの実態を調査する研究が行われている。Mejovaらは、Contents Reuseが企業内のどの部署においてどの程度行われているかの実態を調査している[10]。また、Jensenらは、17人のナレッジワーカーを対象に、個人におけるContents Reuseの実態を調査している[11]。

これらの調査は、コンテンツ生成の効率化を図るために、生み出されたコンテンツの一部または全部が直接的に他のコンテンツに活用されている事象を観察対象としている。対して、本稿における検証実験は、DTFがそもそも新たな文章作成において、どのように活用されるかということ調査する実験であり、直接的、間接的問わずにR-DTFが活

用される事象を詳細に観察するものである。

7. まとめ

本稿では、知識ゴミの一事例であるR-DTFについて、活用可能性の検証を行った。結果として、R-DTFが新たな文章作成において活用可能であることを確認した。また、R-DTFは、一般的な活用対象である文章よりも液状化の程度が高いために、執筆者の文章作成行為に影響を及ぼすという示唆を得た。さらに、R-DTFを有効に活用可能とするための活用環境構築について検討した。

以上より、文章作成において生成される知識ゴミは、知的資源となり得ることが本稿によってより明確になった。今後は、以上の実験と検討の結果を踏まえ、実用的な棄却文章断片の活用環境の構築を目指す。また、新たな知的資源として活用すべく、棄却文章断片以外の知識ゴミについても包括的に検討していく予定である。

謝辞 本研究は、JSPS 科研費 JP15K12093 の助成を受けたものです。本稿の執筆に当たり実験に協力下さった被験者の方々に謝意を表します。

参考文献

- [1] 湯川 抗 ITを活用した知識創造社会の実現にむけて-プラットフォームとしてのコミュニティ- 研究レポート / [富士通総研経済研究所] [編] 142 1~35, 巻頭 2p, 2002.
- [2] Liao, S.. Expert system methodologies and applications—a decade review from 1995 to 2004. *Expert Systems with Applications*. vol. 28, Issue 1, p. 93–103, 2005.
- [3] 生田泰章, 才記駿平, 西本一志. 文章作成過程における棄却文章断片の活用に関する一検討. *インタラクション2016論文集*, 1B35, pp. 302-305, 2016.
- [4] Ikuta, H. and Nishimoto, K.. Wasting “Waste” is a Waste: Gleaning Deleted Text Fragments for Use in Future Knowledge Creation, *Proc. The Tenth International Conference on Advances in Computer-Human Interactions (ACHI 2017)*, pp.193-199, 2017.
- [5] 生田泰章, 西本一志: Con-Text ComposTer: 棄却テキスト断片の活用機会を創出する知識創造活動支援システム, *インタラクション2017論文集*, 2-508-41, pp. 529-534, 2017.
- [6] Nakakoji, K., Yamamoto, Y., Reeves, B.N., Takada, S.. Two-Dimensional Positioning as a Means for Reflection in Design, *Proceedings of Design of Interactive Systems (DIS'2000)*, ACM Press, New York, NY, pp.145-154, 2000.
- [7] 浅原 正幸, 加藤 祥. 文書間類似度について. *自然言語処理*. 23 巻 5 号 pp. 463-499, 2016.
- [8] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto: Applying Conditional Random Fields to Japanese Morphological Analysis, *Proc. of the 2004 Conference on Empirical Methods in Natural Language Processing*, pp.230-237, 2004.
- [9] 堀浩一: 創造活動支援の理論と応用, オーム社, 2007.
- [10] Mejova, Y., Schepper, K. D., Bergman, L., & Lu, J.. Reuse in the wild: An empirical and ethnographic study of organizational content reuse. *Proc. of the 2011 Annual Conference on Human factors in Computing Systems* pp. 2877–2886, 2011.
- [11] Jensen, C., Lonsdale, H., Wynn, E., Cao, J., Slater, M., Dietterich, T. G. The Life and Times of Files and Information: A Study of Desktop Provenance. *Proc. of the 2010 Annual Conference on Human Factors in Computing Systems*, pp.767-776, 2010.