

## 数値情報を用いたテキスト検索方式の提案と評価

山田洋志 福島俊一

{h-yamada, fuku}@hml.cl.nec.co.jp

NEC ヒューマンメディア研究所

本論文では、テキスト中から数値情報を抽出し、単語と組み合わせて検索する方式とその精度評価結果について述べる。テキスト中の数値情報を抽出し登録することで、数値によるテキスト内容の検索や絞り込みができる。本方式では、数値の係り先を特定しなくても、単語条件との単純な組み合わせで数値条件による絞り込み効果を実現した。一方、一般の文章中の数値表現には、概数表現や範囲表現が含まれており、表現もさまざまである。本方式では、数値を修飾する表現を6種類に分類して、抽出した数値とともに利用することで検索もれを防ぐようにした。筆者らが開発した多角的検索システム OTROS に数値情報検索機能を組み込み、BMIR-J2 を用いて検索精度を評価した。その結果、構文解析を使わずに単語と AND 演算するだけでも再現率をあまり下げずに適合率を9%から66%向上でき、本方式の効果を確認できた。

### A Proposal and Evaluation of a Text Retrieval Method Based on Numerical Value Extraction from Text

Hiroshi YAMADA Toshikazu FUKUSHIMA

NEC Human Media Research Laboratories

This paper describes a Japanese text retrieval method based on numerical values extraction. Numerical values in text are useful to decrease unnecessary retrieval results. To select proper text, our method doesn't use sentence structure analysis but uses simple combination of key words and numeric conditions. On the other hand, numerical expressions have various modifiers which show numerical range or approximate values. Authors classify such modifiers into six types, and translate them into suitable numeric ranges. Authors use BMIR-J2, benchmark data set for Japanese text retrieval system, for evaluation of our method. With numerical conditions, precision rate improves 9%–66%.

## 1 はじめに

電子化されたテキストの量は増大する一方であり、それに伴って、ユーザが求めるテキストを適切に選び出す技術がますます重要になっている。現在の検索システムは、単語の検索を中心として発達してきた。類似テキスト検索やクエリ・リファイン、タイトルやURL文字列の検索などは、単語による検索の精度を高める、あるいは、容易にする効果をあげている。しかし、ユーザの求める情報がすべて単語の組み合わせで表現できるわけではなく、それ以外の情報を利用することで、より高精度に検索できる。本論文では、テキスト中の数値に着目して、より高精度な検索を実現する方式を示す。

数値の記述は多くの文書で頻繁に使われ、内容にも強く関わるため、製品の値段など、数値の指定でより正確に検索できるものがある。このことから、

従来、数値を検索に利用するシステムが開発されている [3, 4]。これらのシステムでは、構文解析を利用して数値が表す対象を判断し、数値を含む文脈での検索を実現している。しかし、現状の構文解析・意味解析の精度は必ずしも十分ではなく、解析の誤りや曖昧性のため、かえって検索もれや誤検索を生ずる側面がある。現在、実用システムではテキスト中の数値を対象とした検索機能を持つものはなく、テキスト作成日時など本文とは別に登録される書誌・属性項目としての数値の利用に留まっている。

筆者らは、構文解析や意味解析なしでも、数値条件を組み合わせることで、単語だけを指定する場合よりも適合率を向上させられることに着眼して検索方式を開発した。本方式で扱う数値表現は、数値を表す文字列と単位の組み合わせであり、意味解析や文脈解析と比較して高い精度で抽出でき、構成が

単純で幅広い分野のテキストに対して頑健なシステムを実現できる。

しかし、テキストから数値部分を取り出すだけでは、もとの意味が失われて検索もれを起こす場合がある。たとえば、「2万から3万」のように範囲を示す場合や、「約2万」のように概数を表す場合である。この場合、「2万」や「3万」といった数値だけでなく、範囲内の数値や概数の前後の数値も検索できることが望ましい。本方式では、テキストから数値を抽出する際に、概数や範囲の表現も取り出し、意味別に分類した上で適切な数値範囲に換算することで検索もれを防ぐようにした。

本論文では、2節で数値情報を用いた検索方式を説明し、3節でテキストから数値情報を抽出する方法を説明する。さらに、4節で、構文や意味を使わずに十分な適合率が得られることを示すことで、本方式の有効性を示す。検索精度の評価には、情報検索用テストコレクションBMIR[5, 6]を用いた。

## 2 数値情報を用いた検索方式

本節では、提案する数値検索方式の全体像を説明する。

数値検索は、検索語による条件に加えて、検索語に関連する数値の条件を指定することによって、検索結果を絞り込む機能である。数値検索を実装した検索システムでは、ユーザは検索語に加えて数値の条件を入力できる(下例)。

検索語: パソコン AND モバイル

数値条件 数値: 20万 単位:円 範囲:以下

この例では、数値条件を数値、単位、範囲の組み合わせで指定する。“範囲”は、数値に付加して数値の範囲を拡張する表現で、“丁度”、“以上”、“以下”、“程度”などからユーザが選択する。また、2つの数値を指定して明示的に範囲を指定することもある。

本論文では、上記のような数値検索機能を、構文解析・意味解析は行わず、検索語と数値条件のAND検索で検索結果を絞り込むことで実現する。すなわち、検索語の条件と数値条件はまったく独立な条件としてAND演算を行う。上記の例でいうと「20万円以下」という数値条件は「パソコン AND モバイル」の価格が20万円以下という意味ではなく、ひとつのテキストが検索語と数値の条件を同時に満たすという意味で検索することにする。

このような数値情報検索方式を実現するための

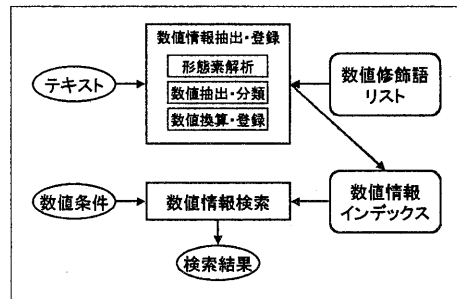


図 1: 数値検索方式の構成

構成を図1に示す。

図1を参照して説明する。“数値情報抽出・登録”部では、テキスト中の数値情報を、検索時の数値条件と同様に、数値、単位、範囲(修飾語の分類)の組み合わせで抽出する。数値抽出の際には、概数や範囲表現などできるだけ多くの種類の数値表現を正確に把握することで検索の漏れを減らすようにした。抽出手順を以下に示す。

1. **形態素解析**: テキストを単語単位に分割し、品詞を添付。
2. **数値情報の抽出**: 品詞をもとに数値を抽出し、数値を表す文字列を数値に変換。さらに、前後の単位・数値修飾語を抽出(3節)。
3. **数値表現の分類**: 数値修飾語を分類・登録したリストを参照して、数値表現を分類(表2)。
4. **数値情報の換算**: 概数や範囲表現は分類に応じて数値の範囲に換算。たとえば、「約1万円」を前後10%の範囲(9000円~11000円)に換算する、「千円弱」は、前側5%まで拡張する(950円~1000円)などの処理を行う。
5. **データベース登録**: 抽出した情報を“数値情報インデックス”に登録。

“数値情報検索”部では、ユーザが指定した数値条件と、テキスト中から抽出されインデックスに登録された数値情報を比較して、範囲が一部でも重なるものを検索結果として選択する。たとえば、インデックスに下記の情報が登録されているとき、

ID	下限値	上限値	単位	元の表記(参考)
0067	9000	11000	円	約一万円
0088	14800		円	1万4800円
0169	27	32	kg	27から32kg
0347	0	9500	円	9500円以下
0360	50	55	kg	50kg~55kg

検索条件が「1万円以上」の場合は、検索結果として文書IDの0067,0088が得られ、「10～30キログラムの範囲」に対しては、0169が得られる。

### 3 数値情報抽出処理

本節では、形態素解析後のテキストから、数値情報を取り出す処理の詳細を述べる。前述のように、品詞をもとに数値を抽出し、次に、前後の単位、修飾表現を抽出する。以下、個々の抽出処理を説明する。

#### 3.1 数値

品詞を調べ数値の文字列を抽出して、大小や範囲の演算ができる数値型に変換する。検索もれを防ぐためには、できるだけ広範囲の数値を対象とすることが望ましい。数値の表記には算用数字、漢数字の組み合わせ方(下表)や、分数や指数、「1万円の50%」のように句や文全体で数値を表現する場合など多くの形式がある。

算用数字	「1234」「1,234」「1.23」
漢数字	「一二三四」「百二十三」
混在	「12万3千」「12.3百万」

本方式では、上の表のような算用数字・漢数字のバリエーションを問わず、単独あるいは範囲表現の整数と小数を抽出対象とする。分数、指数はほとんど出現しないため現在の実装では抽出対象にしていない。前述の「1万円の50%」の場合、計算結果にあたる数値(この例では5000円)は抽出しない。個別の数値(「1万円」、「50%」)は抽出する。

#### 3.2 単位

数詞の直前・直後にある単位を、品詞をもとに抽出する。検索時には金額や大きさなど特定の数値が対象となると考え、単位の付いていない数値は抽出しない。なお、表形式のように単位が数字と離れて記述されている場合があるが、現在は対応していない。「円」と「¥」のような異表記についてはリストを用意し、抽出後に書き換えて統一している。

抽出する単位に方式上の制限はないが、数値検索の対象としては、金額や具体的なものの大きさの需要が大きいと考え、通貨の単位と長さ・重さをシステムに実装している(表1)。「人、年、期」については、後述の精度評価のために抽出するようにした。

表 1: 抽出対象単位

前置:	¥,\$
後置:	円,元,ウォン,セント,ドル,パーツ,フラン,ペニー,ペンス,ポンド,マルク,メートル,m,センチ,センチメートル,cm,キロメートル,km,ミリメートル,mm,マイル,フィート,グラム,g,キログラム,kg,ミリグラム,mg,トン,t,人,年,期

表 2: 数値表現の分類

分類	修飾語
数値単独	(修飾語無し)
概数の表現	前後,程度,くらい,程,おおむね,おおよそ,およそ,ほぼ,大体,約
下限指定	以上,最低,最小,最少,少なくとも
下限(概数)	強,余り
上限指定	以下,以内,未満
	最大,最高,最長,多くとも
上限(概数)	弱,近く

#### 3.3 修飾表現

数値をひとつ指定して範囲を表現する場合,(a)数値の前後に広げる,(b)数値より大きな方へ広げる,(c)小さな方へ広げる,の3とおりがある。本方式では,(b)と(c)について、広げる度合いの多いものと少ないものに分割し、範囲ではない表現を加えて6通りに分類する。表2で「下限(概数)」は、範囲が指定値より大きな方へ広がるが、上限があまり指定値から離れていないものを表す。この判断は、筆者らの主観によった。「上限(概数)」も同様。

日本語の教科書やテキストから数値に付く修飾語を57種類を収集して分類した。WWWページと新聞記事による評価では、これらで数値の修飾表現の95%以上を網羅している[2]。

今回、収集した表現のうち使用頻度の多いもの30種類を抽出対象として実装した(表2)。

#### 3.4 範囲表現

範囲を表す記号の直前直後が数値表現のときは、まとめて範囲表現とする。現在、「から」、「～」、

「-」を抽出対象としている。

数値は単位付きで抽出するのを原則としているが、範囲の開始を表す数値に単位が付いていないときは、終了の数値に付いている単位を用いる(例:「1万~3万円」→「1万円~3万円」)。

#### 4 検索用テストコレクションBMIR-J2による検索精度の評価

数値情報を用いる検索機能を、多角的検索システムOTROS[1, 2]に実装し、具体的な検索条件に対して検索精度を評価した。評価に利用したOTROSシステムでは、検索語の条件としてANDとORの組み合わせが利用でき、数値に対しては、“ちょうど”、“以上”、“以下”、“くらい”、“範囲”が選択できる。

検索精度評価には、BMIR-J2を利用した。

BMIRは、(社)情報処理学会・データベースシステム研究会の下部組織である「情報検索システム評価用データベース構築ワーキンググループ(BMDBWG)」が作成した、検索用のテストコレクションである[5, 6]。

新聞記事と自然言語で記述された検索要求と検索結果の正解となる記事集合からなり、BMIR-J1, BMIR-J2の2セットが作成されている。

BMIR-J2は、毎日新聞CD-ROM'94データ版を基に構築されている。5080件の新聞記事と60件の検索要求/正解記事集合からなる。今回は検索要求のうち数値条件を必要とするものを使用した。

記事集合には、Aランク(検索要求を主題とする記事)とBランク(主題ではないが、検索要求の内容を少しでも記述している記事)の区別ランクが与えられているが、今回は両者とも正解として扱った。

##### 4.1 検索要求

BMIR-J2の検索要求は、正解数が5記事以上ある基本セットと、5記事未満の追加セットからなる。今回は基本セットから、数値を条件として含む検索要求3件を使用した。検索要求は以下のとおり。

115: 「1ドル=100円を超える円高」

116: 「3期以上連続の減益企業」

117: 「千人以上の人員削減を計画している企業」

##### 4.2 検索条件の設定

検索条件は、BMIR-J2の検索要求の記述を参考にして手作業で作成した。作成に当たっては、BMIR-

表 3: 検索に用いた語と数値条件(BMIR-J2)

ID	検索語 1	検索語 2
115	円高	
116	減益	減益 & 連続
117	人員 & 削減	(人員 or 社員) & 削減 & (予定 or 計画)

ID	数値条件 1	数値条件 2
115	100円以下	(50-100)円
116	3期以上	3期以上 or (3-10)年
117	1000人以上	

J1に含まれる類似の検索要求を用いた予備実験を行い、その結果に基づいて検索条件を修正するという手順を経た。

設定に当たって考慮したのは、本方式では、検索語による検索でのみれを数値条件では回復できないため、検索語で絞り込みすぎない方がよい。また、ユーザが指定するのは単純な検索語が多い、という点である。そこで、比較的単純な条件(検索語1, 数値条件1)と、より精度よく絞り込みができそうな条件(検索語2, 数値条件2)を用意した(表3)。評価は、検索語のみの場合と、検索語と数値条件の組み合わせ(2通りまたは4通り)とで行った。

現在のシステムでは期間を表す「年」と年号を区別しないため、「3年以上」を指定すると西暦の「94年」も検索される。そこで、検索要求116の数値条件2については「3-10年」の範囲指定とした。

##### 4.3 検索精度

適合率と再現率、F値<sup>1</sup>を表4, 図2, 図3, 図4に示す。各検索要求ごとの詳細は付録Aで述べる。

いずれの検索要求についても数値条件を指定することで適合率が向上しており、その差は9%から66%になる。数値条件の利用が無駄な検索結果の削減に有効であることが分かる(図2)。それに対して、再現率は検索要求116以外では変わらず、数値条件指定による検索もれが少ないことが分かる(図3)。再現率・適合率を総合すると全体に改善効果が大きく、検索要求116の検索語2の場合でもやや劣る程度と言える(図4)。

<sup>1</sup>適合率は、検索結果中に占める正しい検索結果の割合、再現率は正しい検索結果のうち実際に検索できたものの割合である。F値は適合率、再現率を重み付きで統合した値である。ここでは同じ重み( $\beta = 1$ )で計算した。

表 4: BMIR-J2による検索精度

検索ID	検索条件	検索語のみ			数値条件1			数値条件2		
		適合率	再現率	F値	適合率	再現率	F値	適合率	再現率	F値
115	検索語1	19.5%	90.0%	32.0	67.9%	90.0%	77.4	72.0%	90.0%	80.0
116	検索語1	34.1%	100.0%	50.8	100.0%	60.0%	75.0	61.9%	86.7%	72.2
	検索語2	83.3%	100.0%	90.9	100.0%	60.0%	75.0	92.9%	86.7%	89.7
117	検索語1	20.0%	45.5%	27.8	45.5%	45.5%	45.5			
	検索語2	29.4%	45.5%	35.7	55.6%	45.6%	50.0			

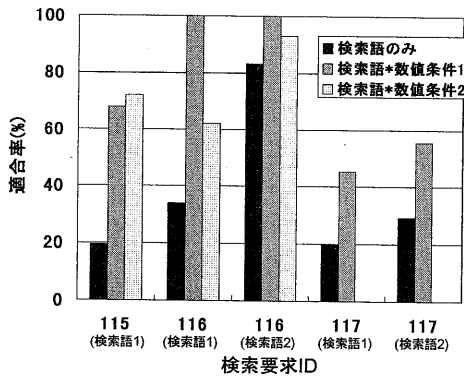


図 2: BMIR-J2による適合率

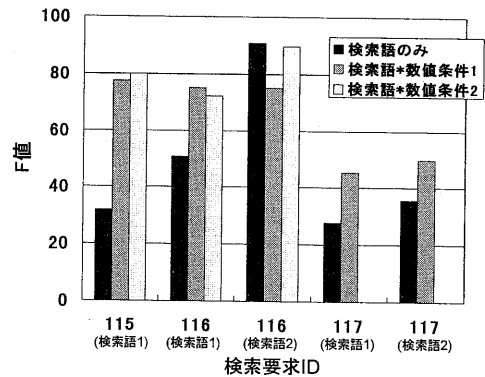


図 4: BMIR-J2によるF値

評価実験の結論として、構文や文脈を利用しなくても、数値条件を指定することで適合率が大幅に向上した。一方、数値条件の指定による再現率低下は少なく、本方式の有効性を確認できた。

#### 4.4 検索誤りの分析

以下では、検索誤りの原因について、検索もれ、検索過剰に分けて分析する。

##### 4.4.1 検索もれの原因

検索要求115と117では、数値指定による検索もれは無い。しかし、検索語だけの段階でもれが多い。

検索要求115では、「円高」ではなく「ドル安」「円相場」「為替相場」などの表現が使われていることでもれが生じた。検索要求117では、「人員」「社員」ではなく、「従業員」が使われている。また、「予定」や「計画」を使わずに表現しているものもある。これらの多くは、同義語辞書やシソーラスを併用することで解決する。

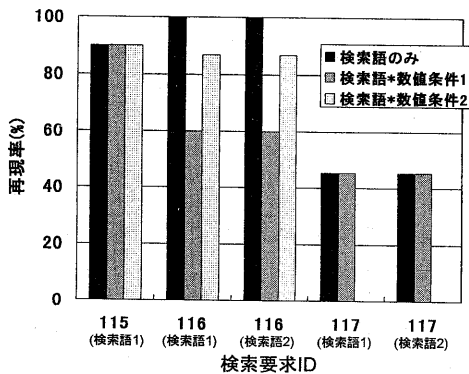


図 3: BMIR-J2による再現率

検索要求116では、数値条件1を使用したときに6件の検索もれがあった。その原因を以下にあげる。

**単位の同義語** 数値条件1では、期数を「年」で記述している記事が検索もれになる(4件)。同様に、「ドル」と「米ドル」、「\$」のように同じ単位の別表記がある場合に、指定を忘れると検索洩れになる。重要な単位についてはシステム側で対応する必要がある。

**内容理解が必要** 内容を理解しないと「3期連続」なのが分からない(1件)。

例:「九四年三月期決算は、……経常利益は、……二期連続の二ケタ減益……九五年三月期も……経常利益は一千億円(同九・四%減)と予想している。」

**形態素解析誤り** 形態素解析結果から数詞と単位を抽出しているため、解析を誤ると抽出もれに通じる(1件)。「三年」が名詞として辞書登録されていたために検索もれが起こった。

#### 4.4.2 検索過剰の原因

数値条件の指定によって適合率が大幅に向上しているが、まだ、過剰な検索結果が含まれている。その原因を分析する。このうち、始めの2種類は本方式に起因し、残りは文章理解の問題である。

**無関係な数値** 検索要求とは無関係の数値が同じテキスト内にある場合に検索されてしまう(14件)。記事を調べると、構文・意味解析なしでも、検索語と数値表現との近接演算を使うことでさらに精度を上げられるケースが多い。

例:「昨年八月二十四日以来の一マルク=六〇円台を記録、」

例:「電気料金が……月額九十八円程度、」

**修飾表現** 本方式では、「約」などの修飾表現を、一定の割合で数値の範囲に置き換えている。(例:「105円前後」→「94.5~115.5円」)

そのため、置き換えた範囲が広すぎて絞り込みに失敗する場合がある(2件)。表現の意図を正確に推測することは困難なため、検索もれが起こらないことを重視して広めの範囲設定にしたことによる。

**省略表現** 単位や数値の省略で抽出誤りを起こす(1件)。今回、「3~7万」を「3万~7万」と判断できなかったために検索過剰を起こした。

**内容理解が必要** 単なる検討や仮定など、数値を記述していても検索要求に沿わない記事や、「100円に迫る」など文脈なしでは数値の大小関係が判定できない表現がある(23件)。

例:「円高をあおることがあっても、1ドル=100円に再び接近することはなさそうで、」(予想、接近方向の判定必要)

また、今回の結果には影響しなかったが、以下の例については、削減数が1000人以上であると判断するために文章を理解して計算することが必要である。

例:「従業員二千八百人の半数を削減」

例:「2万1650人を1万7400人にする」

例:「現在二十一万五千人の社員を、一九九六年度末までに二十万人に絞り込む」

## 5 おわりに

数値の条件指定と検索語とを組み合わせて検索する手法を提案し、検索精度を検索用テストコレクションBMIR-J2を利用して評価した。

本方式では、数値の係り先を特定しなくても数値条件による絞り込み効果を実現でき、また、数値を修飾する表現を6種類に分類し、抽出した数値とともに利用して検索もれを防ぐことができる。

評価の結果、数値条件を追加することで適合率が大幅に向上し、構文解析などの高度な文書処理を行わなくても数値を適合率向上に利用できることが確認できた。数値の抽出は、構文解析などに比べると、高精度に実現でき、テキストの変化の影響を受けにくいため、本手法は種々のテキストを扱うシステムに特に有効である。

今後、WWWページなど各種のテキストで評価するとともに、単位表記の拡充や数値の修飾語の扱いの精密化など、実用化のための改良を行う。また、近接演算との組み合わせなど、多くの選択肢を用意することで、さらに有効な絞り込みを実現する。

本研究には、社団法人 情報処理学会・データベースシステム研究会・情報検索システム評価用データベース構築ワーキンググループが、株式会社日本経済新聞の協力によって、1993年9月1日から12月31日の日本経済新聞記事を基に構築した情報検索評価用データベース(テスト版)、および、同ワーキンググループが新情報処理開発機構との共同作業により、毎日新聞CD-ROM'94データ版を基に構築

表 5: 115: 「1ドル=100円を超える円高」の検索結果

記事ID	正解	検索語1	数値1	数値2	備考
323800	A	もれ	もれ	もれ	「円相場」
479840	A	もれ	もれ	もれ	「円の急騰」
479860	A	もれ	もれ	もれ	「突破」
716410		×	×	×	BMIRの正解判定に疑問(→Bランク正解)
754390		×	×	×	「3~7万円」誤解析
857460	A	もれ	もれ	もれ	「円相場」
890300	B	もれ	もれ	もれ	「為替相場」
922270		×	×	×	BMIRの正解判定に疑問(→Bランク正解)
その他	A,B	○72	○72	○72	
		×7	×7	×7	予想,計画など
		×5	×5	×5	別の価格
		×9	×9	×9	修飾表現の対応不足
		×4	×4	×4	「一〇五円程度」「一〇〇円前後」
		×6	×6		
		×264			
	A	もれ3	もれ3	もれ3	「ドル安」
正解数	80	72	72	72	
検索数		370	106	100	

した情報検索システム評価用テストコレクション BMIR-J2を利用した。

## 参考文献

- [1] 山田,松田,竹元,赤峯,福島,“インターネット多角的検索システムOTROS-全体の概要と構成-”, 情処57回大会,3L-01
- [2] 山田,福島,“インターネット多角的検索システムOTROS-数値情報の抽出と検索-”, 情処57回大会,3L-02
- [3] 岸本,須之内,塚田,千葉,石川,“テキストの構造化に基づく検索システム”, 情処論文誌,Vol.35,No.5,1994
- [4] 斉藤,迫田,中江,岩井,田村,中川,“数値情報をキーとした新聞記事からの情報抽出”, 情処,NL125-6,pp.63-70,1998
- [5] 木本ほか,“日本語情報検索システム評価用テストコレクションの構築”, 1998年情報学シンポジウム講演論文集,pp.103-119,1998

- [6] 木谷ほか,“日本語情報検索システム評価用テストコレクションBMIR-J2”, 情報処理学会研究報告,DBS-114-3,pp.15-22,1998

## A 検索結果の詳細

検索要求別の検索結果を示す(表5-表7)。

各表で“記事ID”は,記事につけられた固有の番号である。“正解”はBMIRによって正解とされている記事とランクの区別(AまたはB)を示す。“検索語1(検索語2)”は検索語1(検索語2)だけを使用した場合の検索結果である。“数値1(数値2)”は検索語に加えて数値条件1(数値条件2)を使用した場合の検索結果である。

検索結果欄で,“○”はBMIRによる正解を検索した,“×”は正解以外を検索した,“もれ”は正解を検索できなかった,空欄は検索しなかった(しないのが正解)ことを示す。記事IDが“その他”の場合は,同じ結果の記事をまとめてあることを示す。○または×の後の数字がまとめた記事数である。

備考欄は,検索誤りの原因について示した。「」で記述してあるのはテキストの中で誤りの原因に関連する箇所である。

表 6: 116: 「3期以上連続の減益企業」の検索結果

記事ID	正解	検索語1	数値1	数値2	検索語2	数値1	数値2	備考
138580		×		×				「十年来言われてきた」 要計算
394620	A	○	もれ	もれ	○	もれ	もれ	
407170		×			×			
837970		×			×			
その他		×19						
		×7		×7				「〇〇年ぶり」
	A	○9	○9	○9	○9	○9	○9	
	A	○5	もれ5	○5	○5	もれ5	○5	「〇〇年連続」
正解数	15	15	9	13	15	9	13	
検索数		44	9	21	18	9	14	

表 7: 117: 「千人以上の人員削減を計画している企業」の検索結果

記事ID	正解	検索語1	数値1	検索語2	数値1	備考
37380	A	○	○	○	○	
38840	A	○	○	○	○	
67430	A	○	○	もれ	もれ	「予定・計画」なし
109260	A	もれ	もれ	○	○	
152160		×	×			従業員数
176400	A	○	○	もれ	もれ	「予定・計画」なし
195190	B	もれ	もれ	もれ	もれ	「従業員」
224140	A	もれ	もれ	もれ	もれ	「社員」あり, 「予定」なし
284890	A	もれ	もれ	もれ	もれ	「従業員」
464070		×	×	×	×	「三千万人いるといわれる核被害者」
616450	A	○	○	○	○	
628020		×	×			全従業員数
831020	A	もれ	もれ	もれ	もれ	「削減」なし
922360	A	もれ	もれ	○	○	「社員」
1005650		×	×	×	×	検討のみ
1007390		×	×	×	×	原案
1008000		×	×	×	×	5社合計
その他		×10				
				×4		
				×4		
正解数	11	5	5	5	5	
検索数		25	11	17	9	