検索エンジンを用いた記載欠落箇所の補完

野浪 尚哉 [*1]　村田 真樹 [*2]　馬 青 [*3]
[*1] 鳥取大学 工学部 知能情報工学科
[*2] 鳥取大学大学院 工学研究科 情報エレクトロニクス専攻
[*3] 龍谷大学 理工学部 数理情報学科
[*1*2]{s132043,murata}@ike.tottori-u.ac.jp
[*3] qma@math.ryukoku.ac.jp

1　はじめに

文章に重要な情報が記載されていない場合，読者の知りたい情報が欠落しているため読みにくい文章となる．そこで，書き漏れのある文章であることを指摘したり，書き漏れのある文章に書き足すべき情報を提示する技術があれば文章の修正がしやすくなると考える．

本稿では，まず Wikipedia から多くの記事で共通して現れる項目を重要項目として，それに関わる情報を取り出し, 表の形に整理して表示する．さらに，その表において空白になっている箇所があった場合は重要情報が欠落している記載欠落箇所とする．重要情報とは表の重要項目に入る具体的な情報のことである．例えば，表 1 において記載欠落箇所となっている「人名」と「組織名」には，対応する Wikipedia の城ページに「人名」や「組織名」に関する情報が記載されていないため，この記載欠落箇所を指摘することで不足していた情報を書き足すことができるようになる．表に記載欠落箇所があることを指摘し書き足すように促すだけでも文章作成支援になるが, 表 2 のように表の記載欠落箇所を埋めることができれば，それを参考にして文章の書き足しを行うことができる. 表 2 の括弧付きの箇所が補完された情報である．

そこで，本稿では検索エンジンを用いて記載欠落箇所に適切な情報を埋める研究を行う．記載欠落箇所に埋められた情報を参考にしながら文章の書き足しを行ってもらうことで書き漏れのある文章の修正支援に役立つと考える．

表 1: 表の空白を記載欠落箇所と定義

	地名	人名	組織名
宇和島城	宇和島		

表 2: 表の記載欠落箇所を埋める

	地名	人名	組織名
宇和島城	宇和島	(藤堂高虎)	(宇和島藩)

2　関連研究

文書からの情報抽出による文章作成支援として藤原 [1]，赤野 [2] の研究が挙げられる．藤原 [1] の研究では，Wikipedia の城に関するページを抽出し，その中から城に関する重要情報を CaboCha[3](固有表現抽出ツール) を用いた固有表現抽出に基づく手法と ALAGIN[4] の上位下位知識に基づく手法の 2 手法で抽出する．2 つの手法によって重要情報を抽出したものを表にまとめる．表に空欄箇所があった場合，情報が欠けている記載欠落箇所と判定し，記載欠落箇所をユーザに知らせて記載の追加を促すことで文章作成支援を行った．赤野 [2] の研究では，藤原 [1] の研究で行った重要項目の取り出し技術の改良を目的として，word2vec[5] を用いたクラスタリングに基づく手法で Wikipedia から重要情報を抽出し表にまとめた．藤原 [1] と同様に，表に空欄箇所があった場合，情報が欠けている記載欠落箇所と判定し，記載欠落箇所をユーザに知らせて記載の追加を促すことで文章作成支援を行った．これらの先行研究では表の記載欠落箇所の指摘は行うものの，記載欠落箇所に埋めるべき情報をユーザに提示する手法の検討は行われていない．そこで本稿では検索エンジンを用いて，記載欠落箇所に埋めるべき情報を取得する手法の研究を行うことで文章修正支援に役立てる．

質問応答システムの精度向上として村田ら [6][7] の研究が挙げられる．質問応答システムの精度向上を目的に，スコアを減らしながら複数の記事でのスコアを利用する方法を提案した．単純にスコアを加算するだけではシステムの性能が下がる場合があるため，この研究ではスコアの加算の際にスコアを減らしながら加算する手法を用いている．

Web を情報源とする質問応答システムとして北斗ら [8] の研究が挙げられる．質問文に含まれる「誰」や「どこ」などの疑問詞を手がかりに，人名，地名，組織名，人工名，数字の 5 種類の回答タイプの分類を行う質問解析と，Web ページ上の質問文に対する回答が書かれていそうな部分の抽出を行うための文書検索と，回答候補を固有表現抽出とスコアの加算処理によって抽出する回答選択の 3 つの手法を組み合わせたシステムを構築した．

論文に記載すべき情報の自動検出による文章作成支援として岡田ら [9] の研究が挙げられる．論文の研究成果や研究の有効性・必要性といった論文に記載必要な情報を「記載必要項目」として論文内で記載必要項目が欠落しているか否かを自動で検出することで文章作成支援を行った．

3　提案手法

本稿の手法は, 文書内における重要情報の抽出と, 検索エンジンを用いた文章修正支援の 2 つの段階からなる．

3.1　文書内における重要情報の抽出

Wikipedia の城名に関するページに対して，Cabocha(固有表現抽出ツール) を用いた固有表現抽出に基づく手法と，ALAGIN の上位下位知識に基づく手法の 2 手法により城に関する重要情報を抽出する．抽出は城のページ単位で行う．

3.1.1　固有表現抽出に基づく手法

Wikipedia の城ページから CaboCha を用いて，「地名」「人名」「組織名」に分類される表現を抽出し，表の行を城名，表の列を重要項目として表にまとめる．表には城データの中で単語が出現した頻度の多い上位 5 つの単語を出力する．この手法では城に関わる人物や城の所在地などの重要情報が抽出される．Wikipedia の記事から「地名」「人名」「組織名」を抽出し表にまとめた例を図 1 に示す．図 1 は Wikipedia の根添城のページから「地名」「人名」「組織名」の表現を抽出し，表にまとめたものを示している．根添城の記事中に「地名」の表現である宮城県，「人名」の表現である源頼義，「組織名」の表現である坪沼八幡神社があるため，これらの表現が抽出されて表に出力される．

3.1.2　上位下位知識に基づく手法

本稿では上位下位関係の抽出に ALAGIN の上位下位関係抽出ツールを用いる．上位下位関係抽出ツールは，

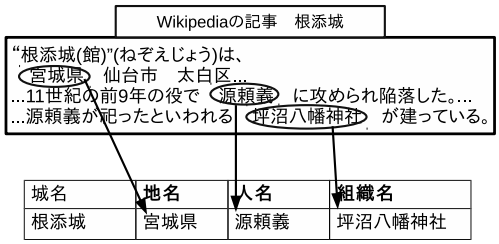

図 1: Wikipedia の記事に固有表現抽出を使用した結果の例

Wikipedia から上位下位関係となる用語ペアを数百万対のオーダーで抽出できるツールである．上位下位関係とは，“X は Y の一種 (一つ) である”と言える X と Y の関係を言う．X のことを下位語，Y のことを上位語と呼ぶ．上位下位知識の例を表 3 に示す．

表 3: 上位下位知識の例

上位語	下位語
仏像	七面大明神像
楽器	カンテレ
文房具	スティックのり
プログラミング言語	prolog

上位下位知識を用いて対象データで下位語の頻度分析を行い，頻度が高かった下位語の上位語を重要項目とする．対象データで重要項目の下位語を取り出し，表にまとめる．固有表現抽出を用いた方法では抽出できなかった情報を抽出できる可能性がある．本稿では「県名」「時代」「地名」「元号」の 4 つの上位語を重要項目として選定して，表の行を城名，表の列を重要項目として表にまとめる．表には城データの中で単語が出現した頻度の多い上位 5 つの単語を出力する．

Wikipedia の記事から「地名」「人名」「組織名」を抽出し表にまとめた例を図 2 に示す．図 2 は Wikipedia の根添城のページから「県名」「時代」「地名」「元号」の下位語となるものを抽出し，表にまとめたものを示している．根添城の記事中に「県名」の下位語である宮城県，「地名」の下位語である仙台があるため，これらの下位語が抽出されて表に出力される．図 2 では根添城の「県名」が宮城県，「地名」が仙台として情報抽出されているが，「時代」「元号」は空白になっている．このような空白がある場合は，Wikipedia のページに「時代」や「元号」に関する情報が記載されていないということであり，空白になっている箇所を埋めるように文章の書き足しを行えばより読みやすい文章になる．そこで本稿では次節のような手法による文章の修正支援を提案する．

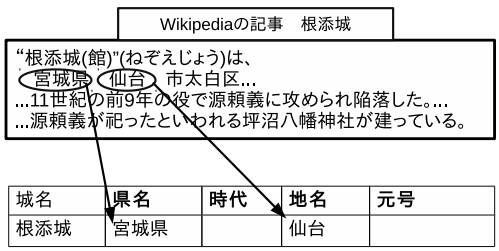

図 2: Wikipedia の記事に上位下位関係抽出を使用した結果の例

3.2　検索エンジンを用いた文章修正支援

上記 2 手法により作成した表に空白があった場合，対応する城ページにはその重要情報が記載されていない．そこで，検索エンジンを用いて表の空白を埋めるべき情報を取得し，表を補完する．

検索エンジンを用いた表の補完方法として，まず城名を検索クエリとして検索エンジンに入力し 50 件の記事を取得する．取得した記事 50 件をまとめた文書に対し 3.1.1 節や 3.1.2 節の手法を用いて城に関する重要情報の抽出を行う．抽出した重要情報のうち，記事 50 件の中で単語が出現した記事の数が多い上位 5 つの単語を表にまとめる．作成した表をユーザに提示することで文章の修正に役立つ．Web 文書から重要情報を抽出し表にまとめる例を図 3 に示す．図 3 では Web 文書からの情報抽出のために上位下位知識に基づく手法を用いている．根添城を検索クエリとして検索エンジンに入力し取得した Web 文書から「県名」「時代」「地名」「元号」の下位語となるものを抽出し，表にまとめたものを示している．図 3 中の文章は取得した 50 件の記事の中からランダムに選んだ記事 1 つを抜粋したものである．記事中に「県名」の下位語である宮城県，「時代」の下位語である平安時代，「地名」の下位語である仙台，「元号」の下位語である永承があるため，これらの下位語が抽出されて表に出力される．図 2 と見比べると，図 2 では表に出力できなかった「時代」「元号」を図 3 では出力できている．このように Wikipedia に記載されていなかった重要情報を Web 文書から取得して表にまとめることで，文章の修正支援に役立てることができる．

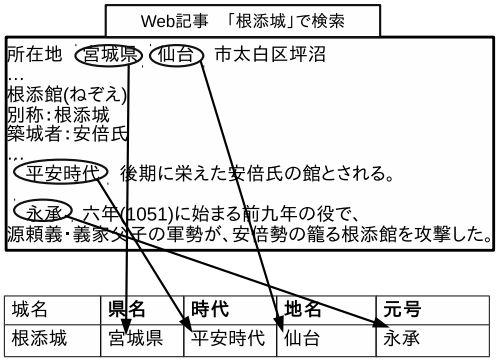

図 3: Web の記事に上位下位関係抽出を使用した結果の例

4　実験

4.1　実験条件

本稿では，Wikipedia の城に関する記事を入力として 3.1 節の手法で情報抽出を行い，記事中の記載不備を検出するための表を作成する．なお，本稿ではその表として，先行研究である藤原 [1] の研究結果を用いる．

3.2 節の手法では検索エンジンにより得られた Web の情報を用いて，Wikipedia の記事中の記載不備を修正しやすくするための表を作成する．3.2 節の手法の性能評価のために，以下の 2 種類の実験を行う．ただし，表の全ての箇所に対して検索エンジンを用いて情報抽出する実験は，Web 文書からの情報抽出の性能を見るために行う実験である．

- 表の全ての箇所に対して検索エンジンを用いた情報抽出
- 表の記載欠落箇所のみに対して検索エンジンを用いた情報抽出

4.2　実験データ

Wikipedia(2014 年 11 月) の城ページ (2,665 ページ) に対して 3.1.1 節や 3.1.2 節の手法を用いて情報抽出と表の作成を行った．また，2,665 件の城ページのうちランダムに選んだ 30 件の城名を検索エンジンにそれぞれ入力して記事を取得し，この 30 件の城データで 3.2 節の手法の実験と評価を行った．なお本研究で用いる検索エンジンは Microsoft 社の BingSearchAPI[10] である．

4.3 評価方法

3.2 節の手法の性能評価のために，Web 文書を対象に 3.1.1 節や 3.1.2 節の手法を用いて作成した表に対して正解率を求める．本稿では，1 位正解率，5 位正解率，MRR の 3 つの方法で評価実験を行った．

4.3.1 固有表現抽出に基づく手法の正解基準

3.2 節の手法で，Web 文書に対して固有表現抽出に基づく手法を用いて作成した表の評価方法について，n 位正解率，MRR で評価を行うための正解基準を説明する．

「地名」の項目は，県名または所在地が抽出された場合正解とする．「人名」の項目は，築城主，城主のどちらかが抽出された場合正解とする．「組織名」の項目は，城に関すると思われる組織が抽出された場合正解とする．また，表に出力する 5 つの単語のうち 1 つでも正解があれば正解とする．

4.3.2 上位下位知識に基づく手法の正解基準

3.2 節の手法で，Web 文書に対して上位下位知識に基づく手法を用いて作成した表の評価方法について，n 位正解率，MRR で評価を行うための正解基準を説明する．

「県名」の項目は，その城が存在する県名が抽出された場合正解とする．「時代」の項目は，築城されてから廃城するまでの時代のいずれかが抽出された場合正解とする．「地名」の項目は，城の所在地が抽出された場合正解とする．「元号」の項目は，築城されてから廃城するまでの元号のいずれかが抽出された場合正解とする．また，表に出力する 5 つの単語のうち 1 つでも正解があれば正解とする．

4.4 実験結果

4.4.1 表の全ての箇所に対して検索エンジンを用いた情報抽出

提案手法の 3.2 節より，Wikipedia(2014 年 11 月) の城ページ (2,665 ページ) を対象として，そのうちランダムに選んだ 30 件の城名を検索エンジンにそれぞれ入力する．検索エンジンより取得した Web 文書を用いて固有表現抽出に基づく手法と上位下位知識に基づく手法により表を作成する．

固有表現抽出に基づく手法を用いて得られた結果の例を表 4 に示す．表において太字で記載してあるものは正解と判断したものである．また抽出した表の正解率を 1 位正解率，5 位正解率，MRR で評価した結果を表 5 に示す．

表 4: 固有表現抽出に基づく手法による情報抽出の結果

	地名	人名	組織名
宇和島城	**宇和島** 日本 宇和島城 宇和島市 愛媛県	**藤堂高虎** 伊達 桑折 藤兵衛 高虎	宇和島城 MS 朝日新聞 **宇和島市観光協会** 二ノ丸
筑後十五城	龍造寺 **九州** 龍造寺隆信 筑後 肥前	**蒲池** 大友 筑後 隆信 龍造寺隆信	**龍造寺軍** 南朝 鎮実 龍造寺家 中日
岡崎城	**岡崎** 岡崎城 日本 岡崎市 愛知	岡崎 **家康** 徳川家康 岡崎城 三河武士	名鉄 備前曲輪 **能楽堂** 岡崎城 ブリタニカ国際大百科事典小項目事典
リンダーホーフ城	ドイツ ミュンヘン 日 日本 バイエルン	城内 **ワーグナー** ルイ 城内リンダーホーフ ルイ世	宿泊ホテル申込最大人数名予約期限 ツアー 中日 ファミリーマート シンデレラ城 サークル K サンクス
小田原城	**小田原** 小田原城 日本 小田原市 神奈川県	**北条** 小田原 豊臣秀吉 大久保 北条氏政	小田原城 日立 新越 日経 韮山

表 5: 固有表現抽出の表の正解率

評価方法	地名	人名	組織名
1 位正解率	0.63(19/30)	0.66(20/30)	0.16(5/30)
5 位正解率	0.80(24/30)	0.83(25/30)	0.50(15/30)
MRR	0.70	0.75	0.26

表 6: 上位下位知識に基づく手法による情報抽出の結果

	県名	時代	地名	元号
宇和島城	福井県 長野県 静岡県 香川県 愛知県	**江戸時代** 戦国時代 安土桃山時代 現代 戦前	石垣 **四国** 城山 名城 大洲	文化 **慶長** 明治 昭和 寛文
築後十五城	**福岡県** 佐賀県 大分県 三潴県 熊本県	**戦国時代** 南北朝時代 江戸時代 室町時代 現代	**田尻** 山下 毛利 天皇 河崎	天皇 文化 **天正** 平成 天文
岡崎城	**愛知県** 静岡県 長野県 新潟県 岐阜県	**江戸時代** 戦国時代 安土桃山時代 現代 室町時代	**愛知** 平成 名城 城内 城下	平成 文化 昭和 明治 **天正**
リンダーホーフ城	新潟県 福岡県 静岡県 神奈川県 ムジンバ	現代	中央 城内 東京 駅前 カナ	普通 文化 平成 天皇 大統
小田原城	**神奈川県** 葉県 千葉県 静岡県 茨城県	**江戸時代** 戦国時代 現代 室町時代 安土桃山時代	北条 石垣 **箱根** 平成 城内	平成 文化 明治 **天正** 昭和

表 7: 上位下位知識の表の正解率

評価方法	県名	時代	地名	元号
1 位正解率	0.53(16/30)	0.66(20/30)	0.36(11/30)	0.33(10/30)
5 位正解率	0.56(17/30)	0.86(26/30)	0.50(15/30)	0.86(26/30)
MRR	0.54	0.75	0.40	0.50

上位下位知識に基づく手法を用いて得られた結果の例を表 6 に示す．表において太字で記載してあるものは正解と判断したものである．また抽出した表の正解率を 1 位正解率，5 位正解率，MRR で評価した結果を表 7 に示す．

なお，表 5 と表 7 で求めた正解率は，単に Web 文書からの情報抽出の性能を見るために行った評価実験である．

4.4.2 表の記載欠落箇所のみに対して検索エンジンを用いた情報抽出

Wikipedia からの情報抽出によって得られた表の記載欠落箇所に対して，検索エンジンによって得られた文書から作成した表を用いて，記載欠落箇所に対応する箇所のみでの検索エンジンの情報抽出の正解率を集計した．その結果を表 8 と表 9 に示す．

表 8 と表 9 において，重要項目ごとではなく全ての重要項目に対しての正解率を求めている．「Wikipedia 内に正解がないもの」とは Wikipedia からの 3.1 節の手法に基づく情報抽出が 100% の正解率で行えており，正しく記載欠落箇所を過不足なく検出できた場合を想定した実験であり，すべての正しい記載欠落箇所での実験である．「表の記載欠落箇所かつ Wikipedia 内に正解がないもの」とは 3.1 節の手法に基づく情報抽出の失敗を考慮した場合の実験であり，3.1 節の手法で正しく特定できた記載欠落箇所だけでの実験である．3.1 節の情報抽出について，固有表現抽出に基づく手法で作成した表では，本稿で用いた 30 件の城データにおいて 20 個の正しい記載欠落箇所のうち 10 個の記載欠落箇所が正しく検出され，同様に上位下位知識に基づく手法で作成した表では，本稿で用いた 30 件の城データにおいて 37 個の正しい記載欠落箇所のうち 33 個の記載欠落箇所が正しく検出された．「表の記載欠落箇所かつ Wikipedia 内に正解がないもの」の実験では，5 位正解率では，固有表現抽出に基づく手法は 0.40 を検出できた．上位下位知識に基づく手法は 0.37 を検出できた．

表 8: 固有表現抽出の表の記載欠落箇所のみでの正解率

評価方法	表の記載欠落箇所かつ Wikipedia 内に正解がないもの	Wikipedia 内に正解がないもの
1 位正解率	0.10(1/10)	0.15(3/20)
5 位正解率	0.40(4/10)	0.45(9/20)
MRR	0.20	0.25

表 9: 上位下位知識の表の記載欠落箇所のみでの正解率

評価方法	表の記載欠落箇所かつ Wikipedia 内に正解がないもの	Wikipedia 内に正解がないもの
1 位正解率	0.21(7/33)	0.18(7/37)
5 位正解率	0.37(15/33)	0.45(17/37)
MRR	0.30	0.28

5 考察

5.1 情報抽出の実験考察

表 5 と表 7 の 2 つの表において，固有表現抽出に基づく手法の重要項目「組織名」と上位下位知識に基づく手法の重要項目「地名」を除いて MRR の値を求めると，およそ 0.50 から 0.70 の性能であることが分かる．そのため，固有表現抽出に基づく手法の重要項目「組織名」と上位下位知識に基づく手法の重要項目「地名」を除いた表をユーザに提示すれば，文章修正支援に役立てられると考える．ただし，表 5 と表 7 で求めた正解率は，Web 文書からの情報抽出の性能を見るために行った評価実験である．本稿の主たる目的は表の記載欠落箇所のみで行った評価実験結果の表 8 と表 9 が性能が良いものになることである．

表 8 と表 9 において，5 位正解率の結果を見ると，固有表現抽出に基づく手法での「表の記載欠落箇所」かつ「Wikipedia 内に正解がないもの」の 5 位正解率は 0.40 であった．上位下位知識に基づく手法での「表の記載欠落箇所」かつ「Wikipedia 内に正解がないもの」の 5 位正解率は 0.37 であった．この結果から，5 割には満たないものの 4 割程度は検索エンジンによって適切な情報を補完することができていると考える．性能が悪い理由としては，情報抽出によって得られた表の記載欠落箇所の中には，Web 文書にも正解候補が見当たらない場合もあり，補完することができた箇所よりも正解候補の推定が困難であったからだと考える．

本稿で行った 3.2 節の手法による情報抽出の実験では，Web 記事 50 件中で出現した記事の数が多い上位 5 つの単語を表の出力としているが，記事頻度以外のパラメータを用いていない．そのため頻度だけではなく，城名と重要項目との単語間の距離を求め頻度に足し込んでから評価を行うなど，新たに頻度以外のパラメータを増やすことによってさらに良い正解率が見込めると考える．

5.2 文章修正支援の成功例

文章修正支援の成功例について説明する．表 10 は，情報抽出した結果 Wikipedia 内に正解の記載がなかったため，記載欠落箇所がある城データである．この記載欠落箇所に対して，本稿の検索エンジンによる実験で得られた正解候補を埋めた結果を表 11 に示す．表 11 の括弧付きの箇所が記載欠落箇所に正解候補を埋めた部分である．このように記載欠落箇所を埋められた場合は文章修正支援に役立つと考える．さらに，表 11 のような記載欠落箇所を補完した表に加えて，該当する Wikipedia の城ページや，検索エンジンにより取得し記載欠落箇所を補完するために利用した文書群をユーザに提示することで，その文書を表の出力が正しいかの確認のために用いてもらうことができる．

表 10: 記載欠落箇所の例

城名	県	時代	地名	元号
溝口城	愛知県		愛知	

表 11: 文章修正支援の例

城名	県	時代	地名	元号
溝口城	愛知県	(戦国時代)	愛知	(天正)

6 おわりに

本稿では，文章の修正支援を行うことを目的に，Wikipedia の文書から重要情報を抽出し，その結果から書き漏れのある文章の検出と，書き漏れのある文章を修正するための情報の提示を行った．固有表現抽出に基づく手法と上位下位知識に基づく手法の 2 手法それぞれに対して表の記載欠落箇所のみでの正解率を求めた結果，5 位正解率の値は，固有表現抽出に基づく手法が 0.40，上位下位知識に基づく手法が 0.37 で検出できた．性能が低い理由としては，Wikipedia からの情報抽出によって得られた表の記載欠落箇所の中には，Web 文書にも正解候補が見当たらない場合もあり，補完することができた箇所よりも正解候補の推定が困難であったからだと考える．

また，3.2 節の手法として，Web 文書を対象とした固有表現抽出に基づく手法と上位下位知識に基づく手法の 2 手法による情報抽出の実験では，MRR の値が，固有表現抽出に基づく手法の重要項目「組織名」と上位下位知識に基づく手法の重要項目「地名」を除けば，およそ 0.50 から 0.70 の性能で検出できた．

今後の課題として，3.2 節の手法による情報抽出の実験で Web 文書から抽出した単語の正解率の向上を考えている．具体的には，本稿での 3.2 節の手法による情報抽出で抽出した単語は，検索エンジンにより取得した Web 文書の中で単語が出現した記事数が多いものであるため，記事頻度に単語間の距離を足し込んでから正解率を求めることで，正解率の向上を見込めると考える．正解率が向上することによって，文章の修正を行うための情報としても精度の高いものにできると考える．

参考文献

[1] 藤原隆太: “Wikipedia からの城情報の取り出しと文章作成支援”，鳥取大学工学部卒業論文，2015.

[2] 赤野北斗: “Wikipedia からの情報抽出における重要項目の選定と改良”，鳥取大学工学部卒業論文，2016.

[3] CaboCha/南 瓜: Yet Another Japanese Dependency Structure Analyzer http://code.google.com/p/cabocha/

[4] 上位下位関係抽出ツール: Hyponymy extraction tool http://alaginrc.nict.go.jp/hyponymy/

[5] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality.In C.J.C.Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pp. 31113119. Curran Associates, Inc., 2013

[6] 村田真樹, 井佐原均: “質問応答システムにおける再検索を用いた回答候補の抽出手法”，情報処理学会自然言語処理研究会，2004-NL-160 , pp115-122，2004.

[7] Masaki Murata, Masao Utiyama, and Hitoshi Isahara: “Use of Multiple Documents as Evidence with Decreased Adding in a Japanese Question-answering System”，Journal of Natural Language Processing，Vol.12，No.2，pp.209-247，2005.

[8] 北斗修哉, 村田真樹, 馬青: “Web を情報源とする日本語質問応答システムに関する研究”，言語処理学会，第 12 回年次大会，pp.939-942，2006.

[9] 岡田拓真, 村田真樹, 徳久雅人, 馬青: “論文における記載不備の自動検出と自動修正に向けた分析”，言語処理学会，第 22 回年次大会，pp.176-179,2016.

[10] Bing Search API :Microsoft Azure Marketplace http://datamarket.azure.com/dataset/bing/search