

# ウェブナビゲーションにおける日本語 LSA を用いた検索難易度の予測

羽瀨 由子<sup>†</sup> 竹内 晴彦<sup>†</sup> 北島 宗雄<sup>‡</sup>

<sup>†</sup>産業技術総合研究所 人間福祉医工学研究部門 〒305-8566 茨城県つくば市東 1-1-1

E-mail: † {habuchi.y, takeuchi.h, kitajima.muneo}@aist.go.jp

**あらまし** ウェブ閲覧におけるユーザビリティ問題を特定し、その重大度を途中ページにおけるリンク選択数により評価する方法として、ウェブ認知ウォークスルー (CWW) がある。本研究では、CWW に利用される日本語 LSA 意味空間を新たに構築し、階層メニューを利用して記事の探索を行うことができるオンライン百科事典のユーザビリティ評価を行った。32 のリンク選択予測値を、18 名の参加者を得て実施した実験結果と比較した。予測値と観測値の間の一致率は低く、CWW のユーザビリティ問題特定基準を見直して適用すると多少改善した。問題が大きいと評価するほど誤差が大きくなる傾向が示され、日本語サイトで階層メニューを持つユーザビリティ評価の信頼度を高めるためには、CWW のユーザビリティ問題特定基準の改良、日本語 LSA 意味空間の整備が必要であることが示唆された。

**キーワード** 情報香, 潜在意味解析 (LSA: Latent Semantic Analysis), 階層的情報構造, ウェブユーザビリティ評価法

## Predicting Website Navigation Problems by Using a Japanese LSA Semantic Space

Yoshiko HABUCHI<sup>†</sup> Haruhiko TAKEUCHI<sup>‡</sup> and Muneo KITAJIMA<sup>‡</sup>

<sup>†</sup> Institute for Human Science and Biomedical Engineering,  
National Institute of Advanced Industrial Science and Technology (AIST)  
Central 6 1-1-1 Higashi, Tsukuba, Ibaraki, 305-8566 Japan

E-mail: † {habuchi.y, takeuchi.h, kitajima.muneo}@aist.go.jp

**Abstract** Cognitive Walkthrough for the Web (CWW) is a usability inspection method for identifying navigational problems and predicting the number of link selections before reaching the target page. This study constructed a Japanese LSA semantic space and assessed usability of an experimental encyclopedia website. The predicted numbers of link selections for carrying out 32 article-searching tasks were compared with those from the human data. The results showed that the predicted numbers of clicks had a large margin of error. It is suggested that it is needed to refine the CWW usability problem identification rules and to build a Japanese LSA semantic space that represents actual users more adequately.

**Keyword** information scent, LSA: Latent Semantic Analysis, hierarchical information structure, Web usability evaluation method

### 1. 問題と目的

ウェブサイトのナビゲーション設計においては、ユーザのナビゲーションプロセスに関する理論に基づいたユーザ動線設計が必要である。それが適切に行われていれば、ユーザは迷うことなく求めている情報に到達できる。しかしながら、サイト設計者の主観や経験則に基づいて動線設計がなされている場合が多く、ユーザの情報検索行動が十分に支援されていないのが現状である。

近年、ウェブサイトのユーザビリティを評価する手

法の開発が進められてきているが、Blackmon ら [1] は、ユーザの動線設計を客観的におこなう方法として、問題箇所を形式的に定められた規則に基づいて特定し、検索の難易度を予測する「ウェブ認知ウォークスルー法 (CWW: Cognitive Walkthrough for the Web)」を提唱した。ウェブ認知ウォークスルー法は、認知モデルに基づくユーザビリティ問題の特定法であり、情報を求めてウェブサイトを訪れたユーザが情報にたどりつけるかどうか、たどりつけない場合の原因の特定、改善法の提案をおこなう方法である。

ウェブ認知ウォークスルーは、ユーザのウェブサイトのナビゲーション行動をシミュレートする認知モデル (CoLiDeS: Comprehension-based Linked model of Deliberate Search) に基づいている。CoLiDeS モデルが仮定するナビゲーション行動の核となる心的プロセスは、「領域の選択」「操作対象の選択」「操作の選択」という選択プロセスである。ユーザは、ゴールの心的表象に最も類似した領域 (サイトナビゲーション、項目リストなど)、操作対象 (テキストリンク、イメージリンクなど) を選択し、最も適切な操作 (通常、クリック) を選択するとされる。

ウェブのユーザビリティを考えたとき、検索ゴールへの動線上にあるハイパーリンクが、順次、確実に選択されるように配置されていれば、問題がない。

ウェブ認知ウォークスルー法では、CoLiDeS モデルに基づき、ユーザが選択する動線を予測し、それが、ウェブサイトが想定する動線と異なっているとき、あるいは、容易にそれを見つけられないときに、ユーザビリティの問題があると判定する。

以下に、ユーザビリティ問題が生じる3つの場合を挙げる。

#### 情報香の弱いリンクラベル (Weak scent correct link)

ゴールにつながるリンクであるが、ユーザの想定するゴールとの意味的な類似性が低いリンク。ユーザはリンクの意味を理解するが、それがゴールにつながるリンクであると判断される可能性が低い。結果的に、ゴールにつながるリンクが選択される可能性が低くなる。

#### 馴染みのないリンクラベル (Unfamiliar correct link)

ゴールにつながるリンクであるが、熟知度の低い単語が用いられているリンク。ユーザはリンクの意味を理解できず、それがゴールにつながるリンクであると判断される可能性が低い。結果的に、ゴールにつながるリンクが選択される可能性が低くなる。

#### 競合する見出しの下にある情報香の高いリンクラベル (Competing links nested under the competing headings)

ゴールにつながらないリンクであるにもかかわらず、ユーザの想定するゴールとの類似性が高いリンク。ユーザは、競合するリンクに注意を奪われ、正しいリンクに気づかない。結果的に、ゴールにつながるリンクが最初に選択される可能性が低くなる。

CoLiDeS モデルでは、検索ゴールの心的表象とハイパーリンクの心的表象の間の類似度を計算するために、潜在意味解析 (LSA; Latent Semantic Analysis [2]) によって構成される意味空間を用いている。LSA 意味空間は、以下の手順で構成される (詳細は、[3]を参照のこ

と)。まず、大規模コーパスを利用して、コーパス内の個々の単語が、個々の文脈で現れる頻度を計数し、単語間の共起関係を引き出し、文脈×単語の共起行列を作成する。そして、この行列に対して特異値分解を行う。さらに、次元縮減を行って、各単語、文脈を 300 次元程度のベクトルとして表現する。このベクトル空間が、利用したコーパスに含まれる単語と文脈の意味的関連性を表現する LSA 意味空間となる。LSA 意味空間では、複合語は、それを構成する単語のベクトル和となる。語間の類似度 (意味的類似度) は、語ベクトルのなす角度の余弦の値で与えられる。

<http://lsa.colorado.edu> には、さまざまなコーパスから構成された意味空間をインタラクティブに利用できる環境が提供されている。例えば、米国の 12 年生の一般的な語彙知識 (General reading up to 12th grade) を表現する意味空間を利用して、彼らが任意の 2 語の間に認知する意味的類似度を知ることができる。school と teacher を入力すると、意味的類似度として 0.61 が返される。このことにより、米国 12 年生は、school という言葉と teacher という言葉の間に、高い意味的な類似性を感じることがわかる。一方、school と actor を入力すると、意味的類似度が 0.08 となり、高い意味的な関連性を感じていないことがわかる。

ウェブ認知ウォークスルー法では、算出された類似度を用いてウェブサイトの問題点の特定をおこなう。基本となる考え方は、以下の通りである。

まず、ユーザは検索ゴールに対する答えが記載されているページにたどり着くために、ユーザがウェブページを探索しているという状況を仮定する。また、ウェブページは、見出しをつけられたいくつかの領域から構成され、各見出しの下には、いくつかのクリック可能なリンクが提供されているものとする。ユーザは、検索ゴールの心的表象に最も意味的に類似した見出しを選択し、続けて、見出しの下に提供されているリンクの中から、検索ゴールの心的表象に最も意味的に類似したリンクを選択すると仮定する。この仮定は、CoLiDeS モデルに基づくものである。語の間の意味的類似度を求める際には、まず、語の精緻化を行う。これは、語に関連した知識の活性化過程をシミュレートするものであり、CoLiDeS モデルが基礎を置いている文章理解の認知理論である Construction-Integration 認知機構[4]に基づいている。

以上の前提に立ち、Blackmon ら[1]は、問題点の特定の手順と問題点の基準を以下のように定めている:

1. 検索ゴールについての解説文を 100~200 語で作成し、その中で使用されている単語集 (ゴール単語集) を作成する。

2. リンクを精緻化し、意味を拡大する。各リンクについて、潜在意味解析によって類似度.50以上の類似語を選出する。その中からコーパス内での出現頻度が50回以上の単語だけを選出する（リンク単語集）。
3. 見出し語を精緻化して、意味を拡大する。見出しの下に含まれるリンクを精緻化した結果も加えた見出し語について、潜在意味解析によって類似語を選出する（見出し語単語集）。
4. 検索ゴールとリンクの類似度を計算する。ゴール単語集とリンク単語集との類似度を計算し、そのリンクとゴールとの類似度とする。
5. 検索ゴールと見出しの類似度を計算する。ゴール単語集と見出し語単語集との間の類似度を計算し、その見出しとゴールとの類似度とする。

算出された値に基づき、問題箇所を特定をおこなう。それぞれの基準については、以下のように定義されている[1]。

#### 情報香の弱いリンク(Weak scent correct link)の基準(WS)

ゴールと正しいリンクとの間の類似度が.10より小さく、それ以外にゴールとの間の類似度が.10以上のリンクがない場合、そのリンクを弱い情報香のリンクとする。

#### 馴染みのないリンク(Unfamiliar correct link)の基準(UF)

正しいリンクの単語長が1で意味空間におけるベクトル長が.55以下、あるいは単語長が2以上で合成ベクトル長が.80以下の場合、そのリンクを馴染みのないリンクとする。

#### 競合する見出しの下にある情報香の高いリンク(Competing links nested under the competing headings)の基準(CL)

##### 競合する見出しの特定

**規則1** 正しいリンクを含まない見出しであるにもかかわらず、ゴールとその見出しの類似度が、ゴールと正しいリンクを含む見出しとの類似度の.80倍以上であり、かつ、ゴールとの類似度が.10以上であり、かつ、見出しの下に張られているリンクとゴールとの類似度が.20以上ある場合、その見出しを競合する見出しとする。

**規則2** 正しいリンクを含まない見出しであるにもかかわらず、その下に張られているリンクがゴールと.30以上の類似度をもつ場合、そのリンクが含まれる見出しを競合する見出しとする。

##### 競合するリンクの特定

**規則1** 競合する見出しの下にあって、ゴールとの

類似度が.10以上であり、かつ、ゴールとの類似度が、正しいリンクとゴールとの間の類似度の.80倍以上であり、かつ、同じ見出しの下にあるリンクの中でゴールとの類似度が高い方から3番目まで、あるいはゴールとの類似度が.30以上である場合、そのリンクを競合するリンクとする。

**規則2** 競合する見出しの下にあって、ゴールとの類似度が.20以上であり、かつ、それが、これを満たす唯一のものであるとき、そのリンクを競合するリンクとする。

米国の大学生レベルのLSA意味空間を利用し、上に述べたユーザビリティ問題の特定基準に従って、多くの実験結果を編纂し、リンククリック数を予測する式が導き出されている[1]。

平均総クリック数＝

2.292

+1.516

(正しいリンクが情報香の低いリンクの基準を満たす場合)

+1.757

(正しいリンクが馴染みのないリンクの基準を満たす場合)

+ .655×N(競合する見出しの下にある情報香の高いリンクを満たすリンクの総数)

## 2. ウェブ認知ウォークスルー法の実施

従来、ウェブ認知ウォークスルー法は、階層メニューを含まない英語ウェブサイトを用いる対象としてきた。本論文では、階層メニューを含む日本語サイトに対してウェブ認知ウォークスルー法を適用し、適用の可能性、問題点の抽出を試みた。

### 2.1. 日本語意味空間の構築

意味的類似度の計算を行うために、日本語LSA意味空間を構築した。日本人成人を想定し、毎日新聞1997[5]をコーパスとして利用した。各記事の文字数が100字以上400字以内となるように文章量を調整し、47,878単語、258,512文脈を対象として、300次元の意味空間を作成した。

### 2.2. 評価サイトの選択と問題箇所の特定

評価対象ウェブページは、コロラド大学の自動的ウェブ認知ウォークスルー法のサイトに掲載されている階層メニューを含むオンライン百科事典模擬サイト

(Exp040423のA1

(<http://autocww.colorado.edu/~blackmon/Expt040423Home.html>)を翻訳して作成した日本語版であった。

日本語版百科事典サイトは、6つの人文分野の見出

し（芸術・言語・文学，歴史，地理，宗教と哲学，スポーツ・趣味・ペット，舞台芸術），3つの科学分野の見出し（生命科学，自然科学と科学技術，社会科学）から構成され，その下に4から15の下位リンクが置かれている．下位リンクの総数は64である．下位リンクの下に記事見出しが五十音順に配置され，それぞれの詳細な記事とリンクしている．

このオンライン百科事典記事検索サイトを利用して，分野による偏りがないように，人文分野の16記事，科学分野の16記事の合計32の記事を検索する状況を想定し，サイトが提供する階層メニューのユーザビリティ評価を日本語CWWによりおこなった．評価を行う際，検索ゴールを記述する必要があるため，本研究では，検索ゴールは，検索記事を短く要約したものを想定し，それは，実験者が与えるものとした．

問題箇所の特定には，まず，日本語潜在意味解析を用いて，各見出し語と検索ゴール，リンクと検索ゴールについて，余弦値およびベクトル長を手順に従って計算した．ただし，ゴール単語集，見出し語単語集，リンク単語集のサイズが膨大になったので，リンクと見出し語の単語集作成の際の類似語の選出基準は.80とした．

各検索ゴールにおける問題点の有無および数に基づいて，ウェブ認知ウォークスルー法のリンククリック数の予測式から，表1のように各課題のクリック数を予測した．分野の違いによる適合度の違いについても後で検討するために，以降の分析では検索記事の分野について，人文分野と科学分野に分けて検討をおこなうことにした．

表1 32の検索ゴールに対するユーザビリティ問題の所在とリンククリック数の予測値

	ページ数		予測クリック数	
	人文	科学	人文	科学
問題なし	0	2	-	2.29
問題あり				
低類似度 (WS)	1	0	3.81	-
低熟知度 (UF)	1	0	4.05	-
競合リンク (CL)	6	6	5.35	6.88
複合問題	8	8	7.70	8.71
計/平均	16	16	5.23	5.96

### 3. 検証実験

#### 3.1. 方法

**参加者** 日本人成人20名（男性7名，女性13名）．平均年齢は，30.1歳（SD=9.7）であった．

**装置** パーソナルコンピュータ（DELL Dimention/9150），

17インチTFTディスプレイ（1024×768 pixel），Tobii Clear view2.70を使用した．

**手続き** 個別実験であった．実験参加者は，画面に表示される検索記事について詳細な情報を得るためにオンラインの百科事典の中から記事の書かれたページを130秒以内に探し出すように指示された．

実験の流れは以下のものであった．最初にディスプレイ上にターゲット記事が呈示される．記事のハイパーリンクをクリックすると，記事を要約した説明文と9つの見出しを含むページが全画面表示される．見出しの一つをクリックすると，下位リンクが開いたページが全画面表示される．下位リンクをさらにクリックすると，そのカテゴリーにある記事見出しが50音順に並べられたページが開く．その中にターゲットのタイトルがあればそこをクリックする．すると，詳細な記事画面が表示される．課題画面でのハイパーリンクのクリックから，ターゲット記事名のクリックまでの反応時間，閲覧ページ，および，クリック数が記録された．記事の見出し画面，および，リンク画面ではターゲット記事を要約した説明文が常にページの上部に表示されており，いつでも確認できた．本試行に先立ち，練習が5試行おこなわれ，練習に続いて，32試行の本試行がおこなわれた．課題の呈示順序はブロック化してカウンターバランスがとられた．

#### 3.2. 結果

20名のうち機器トラブルによる欠損値が生じた2名を除いた18名のデータを分析の対象とした．ウェブ認知ウォークスルー法では，リンク画面のみのクリック数を予測の対象としているので，リンク画面以外のクリック数は分析の対象から除外した．

各問題箇所における平均課題達成率，リンクページの平均クリック数，予測値との平均誤差，予測クリック数と観測クリック数との平均一致率（予測クリック数±1の範囲内にある観測クリック数の割合）を分野別に表2に示す．

両分野における課題達成率，および観測クリック数は，予測した難易度の順を反映していないようであった．問題がないと評価された検索課題では，誤差が2クリック以下であった．しかし，問題があると評価された検索課題では，予測値との誤差が約2クリックより多く，予測値との一致率は低かった．

次に，問題の大きさの程度の観点から予測クリック数を3段階（問題なし： $N \leq 2.5$ ，中程度の問題： $2.5 < N \leq 5.0$ ，重篤な問題： $5.0 < N$ ）に分けて平均値を求めた結果を表3に示す．両分野において得られた結果は，問題が大きいと評価された検索課題ほど実際のクリック数との誤差が大きいことを示している．

表2 各条件の記事検索における平均値 (N=18)

	人文			
	課題達成率	観測クリック数	予測との誤差	一致率
問題なし	-	-	-	-
低類似度 (WS)	53%	5.94 (3.00)	2.72	17%
低熟知度 (UF)	61%	5.33 (2.40)	2.21	22%
競合リンク (CL)	76%	3.21 (1.10)	3.40	11%
複合	63%	4.69 (1.22)	3.91	10%
平均	63%	4.79	3.06	15%
	科学			
	課題達成率	観測クリック数	予測との誤差	一致率
問題なし	92%	2.64 (1.37)	1.76	14%
低類似度 (WS)	-	-	-	-
低熟知度 (UF)	-	-	-	-
競合リンク (CL)	82%	2.86 (.76)	4.24	15%
複合問題	82%	3.04 (.73)	5.91	4%
平均	85%	2.85	3.97	11%

(SD)

表3 問題の大きさの程度別の平均値 (N=18)

	人文			
	予測	観測	誤差	一致率
問題なし (N≤2.5)	-	-	-	-
中程度 (2.5<N≤5.0)	3.96	3.84 (1.21)	2.46	16%
重篤 (5.0<N)	7.78	4.50 (1.01)	4.19	9%
	科学			
	予測	観測	誤差	一致率
問題なし (N≤2.5)	2.29	2.64 (1.37)	1.76	14%
中程度 (2.5<N≤5.0)	3.60	1.92 (1.07)	2.02	25%
重篤 (5.0<N)	8.65	3.14 (.70)	5.72	6%

(SD)

### 3.3. 考察

全体的に誤差が大きく、一致度が低い結果が示された。特に、問題が大きいと評価された検索課題ほど実際のクリック数との差が大きかった。このような結果が得られた原因として、問題箇所が過剰に特定されている可能性が考えられた。すなわち、ウェブ認知ウオ

ークスルー法の予測式では、見出しと下位リンクが一覧表示されている平面構造のナビゲーションが想定されているが、平面構造のナビゲーションでは、見出しの下リンクがすべて一覧できる状態なので、リンクの選択時には他の見出しの下にある競合するリンクの影響は大きい。しかし、階層構造のナビゲーションでは、見出しをクリックして開くまでは下位リンクは見えないので、平面構造のナビゲーションよりも選択時に競合する可能性は低い。よって、予測で問題箇所が多くても、実際はあまり問題として影響しなかった可能性が考えられた。そこで、これらの推測を確かめるために、階層ナビゲーションでは影響が小さいと思われる規則を適用せずに評価を再度おこない、結果を比較することにした。

## 4. 結果の再分析

### 4.1. 方法

階層ナビゲーションの評価において、競合するリンクが関わる基準のうち、「競合する見出しを特定する規則2（正しいリンクを含まない見出しであるにもかかわらず、その下に張られているリンクがゴールと.30以上の類似度をもつ場合、そのリンクが含まれる見出しを競合する見出しとする）」について、階層ナビゲーションではリンク選択の回数のおよぼす影響が小さいと判断し、ナビゲーションの評価時に規則を適用せずに評価をおこなった（表4）。この手続きによって競合見出しとしてカウントされる見出しの数が減り、全体の65.63%にあたる21のナビゲーションで予測クリック数が減少した。

表4 32の検索ゴールに対するユーザビリティ問題の所在とリンククリック数の予測値

	ページ数		予測クリック数	
	人文	科学	人文	科学
問題なし	2	7	2.29	2.29
問題あり				
低類似度 (WS)	2	0	3.81	-
低熟知度 (UF)	2	4	4.05	4.05
競合リンク (CL)	4	1	5.40	4.91
複合問題	6	4	7.75	7.32
計/平均	16	16	4.66	4.64

### 4.2. 結果

再度算出した予測値と観測値とを比較した結果を表5、表6に示す。全体的に誤差が減少し、操作前よりも一致率は向上した。人文分野では問題の大きさの程度に従って、課題達成率が減少し、観測クリック数が増える傾向が示された。しかし、科学分野では問題

の大きさの程度にかかわらず全体的に課題達成率が高く、クリック数も少なかった。先の分析と同様に、全体として、問題が大きいと評価した検索課題ほど誤差が大きく、一致率は低かった。

表5 各条件の記事検索における平均値 (N=18)

	人文			
	課題達成率	観測クリック数	予測との誤差	一致率
問題なし	86%	2.31 (1.59)	1.52	33%
低類似度 (WS)	60%	5.47 (2.17)	2.40	19%
低熟知度 (UF)	58%	5.19 (2.03)	2.24	22%
競合リンク (CL)	71%	3.67 (1.31)	3.54	17%
複合問題	63%	4.57 (1.37)	3.79	10%
	科学			
	課題達成率	観測クリック数	予測との誤差	一致率
問題なし	86%	2.85 (.67)	1.81	25%
低類似度 (WS)	-	-	-	-
低熟知度 (UF)	83%	2.90 (.92)	2.40	14%
競合リンク (CL)	78%	2.50 (1.20)	2.42	22%
複合問題	81%	3.18 (1.05)	4.54	6%

(SD)

表6 問題の大きさの程度別の平均値 (N=18)

	人文			
	予測	観測	誤差	一致率
問題なし (N ≤ 2.5)	2.29	2.31 (1.59)	1.52	33%
中程度 (2.5 < N ≤ 5.0)	3.96	4.87 (1.35)	2.36	21%
重篤 (5.0 < N)	7.78	4.19 (1.22)	4.24	9%
	科学			
	予測	観測	誤差	一致率
問題なし (N ≤ 2.5)	2.29	2.85 (.67)	1.81	25%
中程度 (2.5 < N ≤ 5.0)	3.60	2.82 (.75)	2.41	16%
重篤 (5.0 < N)	8.65	3.18 (1.05)	4.54	6%

(SD)

### 4.3. 考察

予測クリック数が多くなる原因を再考し、問題箇所を特定する基準を調整することによって、全体的に予測と実際のクリック数の一致率が向上した。この操作

によって問題のない検索課題では2クリック以下、問題のある検索課題で5クリック以下の予測誤差でページが評価された。

課題分野の違いによる異同については、両分野で問題が大きいと評価された検索課題ほど誤差が大きくなる傾向が同様に示されたが、人文分野での一致率の方が科学分野の一致率よりも相対的に高いようであった。これは、日本語意味空間を新聞記事コーパスに基づいて構築した影響によるものと思われる。すなわち、新聞記事では、科学分野に多い領域固有の専門用語のような単語の出現頻度が低いので、実際の日本人成人の語彙知識を反映できていなかった可能性が考えられる。

今後、日本語環境においてウェブ認知ウォークスルー法による問題箇所特定と検索難易度の推定および評価の精度を向上させるためには次の2つの課題の克服が必要であろう。一つは、階層ナビゲーションに最適化された問題点の特定基準と予測式の構築であり、もう一つは、対象者の語彙知識を十分に反映しているコーパスを用いて意味空間を作成することである。

### 5. まとめ

ウェブ認知ウォークスルー法を用いて日本語の階層ナビゲーションサイトで問題箇所特定と問題の大きさの予測をおこない、行動実験データと比較した。比較の結果、階層ナビゲーションでは、競合見出しの影響を減少させて適用するほうが予測の精度が高まること示された。しかしながら、実際のデータとの一致率はまだ低く、さらなる調整の必要性が示された。一般日本人成人の意味空間を代表するようなコーパスを用いて、データ数を蓄積しながら予測の精度を高めていく必要があるだろう。

### 文 献

- [1] M. H. Blackmon, D. R. Mandalia, P. G. Polson, and M. Kitajima, Automating usability evaluation: Cognitive walkthrough for the web puts LSA to work on real-world HCI design problems., in Handbook of Latent Semantic Analysis, eds. T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch, pp.345-375, Lawrence Erlbaum, Mahwah.
- [2] T. K. Landauer, and S. T. Dumais, A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review, vol.104, no.2, pp.211-240, April 1997.
- [3] 北島宗雄, “潜在意味解析 (Latent Semantic Analysis: LSA)”, “日本知能情報ファジィ学会誌, vol.17, no.1, p.76, 2005.
- [4] W. Kintsch, Comprehension: A paradigm for cognition, Cambridge University press, Cambridge, 1998.
- [5] 毎日新聞社, CD-毎日新聞'97データ集, 毎日新聞社, 日外アソシエーツ, 東京, 1998.