

タイトル	Webサイトの構造的特徴に関する探索的検討
著者	五十嵐, 祐; Igarashi, Tasuku
引用	北海学園大学経営論集, 9(2): 23-34
発行日	2011-09-25

Webサイトの構造的特徴に関する探索的検討

五十嵐

祐

インターネットを通じて接触する情報量の多さは、人々の認知能力や精神的・心理的健康に過大な負担をかけている。日本では、インターネット普及後の1996年から2006年にかけて、個人が接触する情報量が530倍にも拡大する、いわゆる「情報爆発」が起きている（総務省，2008）。1997年にはすでに、「死ぬほど情報が欲しい；情報患いによる死」（“dying for information”）と題するレポートがロイター社によって報告されており、電子メールやオンライン会議など、多様なメディアの普及によって、イギリスで企業のマネージャーが社内や社外におけるコミュニケーション量の増大に直面した結果、受け取った情報を処理しきれずにストレスを抱えていることが指摘されている（Waddington, 1997）。多くの情報がデジタル化・アーカイブ化された社会では、検索エンジンなどを用いて主体的な情報探索を行うことで、自分の望む情報にたどり着ける可能性が高まる（Vise & Malseed, 2005）。その一方で、多すぎる情報はかえって人々の選択を惑わせ、選択の結果に対する満足感を低めてしまうことにもつながる（Iyengar, 2010）。また、情報の過度の氾濫は、目的の情報以外のいわゆるノイズ、例えば、特定の企業や商品の広告・宣伝、信頼度の低い情報、ウソの情報などに人々を接触させる可能性も高めてしまう（Carr, 2010）。このように、インターネット上に溢れる情報は増大の一途をたどる一方、

それらの価値を吟味し、適切に処理することに関して、人々の負担はますます大きくなっているといえる。

Edmunds & Morris (2000) は、過剰な量の情報への接触が心理的なストレスを引き起こすことを、情報オーバーロード (information overload) と呼んでいる。情報オーバーロードに広く認められた明確な定義はないが、大量の情報を受け取ることで、情報をストレスやエラーなく効率的に処理できる限界を超えた状態を指す。Butcher (1998) は、マネージャーが情報オーバーロードに陥ってしまう要因を、以下の7点にまとめている (Edmunds & Morris, 2000)：(1)よりよい意思決定をもたらすと思われる合理性や能力への強い関心を示す目的で情報の収集を行う、(2)不必要な情報を膨大に受け取る、(3)すでに持っている情報を確認するためにさらに情報を集める、(4) (情報を集めることで) 決定の正当性をアピールする、(5)必要かもしれない情報を集める、(6)安全のためにできる限り多くの情報を集める、(7)同僚に取り残されないように情報を「通貨」のように扱おうとする。これらのことは、情報オーバーロードが主に心理的な要因によって引き起こされることを示している。

インターネット上のサイトがもたらす過剰な量の文字情報の処理は、情報の受け手にとって大きな負担となる。Jones, Ravid, & Rafaeli (2004) は、ニュースグループのコ

コミュニティにおける投稿内容を分析し、コミュニティの情報量が増大して受け手が情報オーバーロードを感じるようになると、コミュニティではシンプルなメッセージが目されやすくなり、参加者もよりシンプルなメッセージを投稿するようになることを明らかにした。一方で、コミュニティでの情報量が多くなりすぎると、グループへの参加をやめる人も出てくることも明らかとなった。

また、インターネット上の情報には文字だけでなく、動画や音声なども含まれる。こうしたコンテンツのマルチチャネル化は、インターネットの利用者が接触・選択可能な情報の量を飛躍的に増加させているといえよう。実際、複数のメディアを通じた過剰な量の情報への接触は、注意力を散漫にし、認知能力を低下させてしまう。Ophir, Nass, & Wagner (2009) は、12の異なるメディア（印刷物、テレビ、ネット上の動画、ネット上の音楽、ネット上の音声メディア（ネットラジオなど）、コンピュータゲーム、音声通話（固定電話、携帯電話）、携帯メール、PCメール、オンラインチャット（MSN、Skype など）、ネットサーフィン、コン

ピュータによる他の作業（ワープロ、表計算など）の同時利用（マルチタスキング）が認知課題の処理方略に与える影響を検討している。コンピュータ課題を用いた記憶実験では、マルチタスキングを行う人ほど課題に関連しない妨害刺激の影響を受けやすく、異なる課題間の切り替えに手間取る傾向がみられた。マルチタスキングは情報オーバーロードをもたらす重要な要因のひとつであるが、動画や音声といった内容の解釈に高い負荷がかかる情報を含むサイトは、それ単体でも利用者にマルチタスキングと同等の影響を及ぼすと考えられる。

以上のように、情報オーバーロードをもたらす要因は、利用者の心理的要因とサイトに含まれる情報の2つに大別される。Germonprez & Zigurs (2003) は、インターネット上のサイトの複雑性を構成する次元として、(1)認知、(2)コンテンツ、(3)形式の3つを挙げている。それぞれの次元の規定因は Table 1 の通りである。

認知の次元は、利用者の志向性やサイト利用の経験、サイト利用に関して個人・集団が持つ信念によって決定される。コンテンツの

Table 1. サイトの複雑性の次元（Germonprez & Zigurs, 2003 を一部改変）

次元	規定因（先行研究で検討された内容）
(1)認知	認知マップ 目印やルート、査定による情報探索 ページ同士の関係 内容の整合性や認知的負荷への志向 達成度と志向との関係
	合理的行為理論に基づく信念 利便性や有用性の知覚
	集団に対する魅力
(2)コンテンツ	情報オーバーロード 情報のあいまいさ 情報の複雑性
	(3)形式

次元は、情報の量や複雑性によって決定される。形式の次元は、サイトのデザインや構造といった、サイト作成者が規定する情報の伝達の様式によって決定される。インターネット上のサイトは、静的な文字情報、画像情報を提供するとともに、リンク（ハイパーリンク）によって、サイト内外の関連情報の参照・移動が簡単に行えるという特徴を持つ。また、従来の HTML では表現が難しかった動画や音声のストリーミング、情報のリアルタイム更新、インタラクティブなコンテンツなどは、Flash アプリケーションやブラウザに実装された JavaScript によって実現されている。

本研究では、特にサイトのコンテンツ・形式の次元に注目し、情報オーバーロードを生み出すような Web サイトの構造的特徴について探索的な検討を行う。具体的には、アクセス数の多い Web サイトが持つ構造的特徴を HTML タグの分析を通じて分類し、サイトの構成要素の縮約を試みる。

方 法

分析対象 ㈱翔泳社と㈱NTT データキュービットが運営する donnamedia（ドンナメディア：<http://donnamedia.shoeisha.jp/>）を利用して、2010 年 8 月の月間閲覧数（PV 数）が多かった日本国内の Web サイトを抽出した。donnamedia では、約 15000 のサイトを、カテゴリー（系）別に「企業・経済」、「インターネット」、「エンターテインメント」、「コンピュータ・技術」、「生活」、「ショッピング」、「食と健康」、「地域・交通・旅行」、「その他」の 9 種に分類して PV 数を掲載している。本研究では、それぞれのカテゴリーで PV 数が上位のサイトを抽出し、計 140 サイト（donnamedia 掲載サイト全体の約 1%）について分析を行った。PV 数は最も多いサイトで月間約 8 億、最も少ないサ

イトで月間約 3500 万であった。

なお、本研究では各サイトの内容を吟味した上で、改めて 12 種類のカテゴリー分けを行った（サイトのリストは Table 5 を参照）。「サービス」は、日本郵便や Google マップなどの公共性の高い情報を提供するサイトを含む。「ショッピング」は、楽天市場やアスクルなどのオンラインショッピングに関するサイトを含む。「スポーツ」は、阪神タイガースや日本ゴルフツアー機構など、スポーツに関する情報を提供するサイトを含む。「ソーシャル」は、FC2 ブログや Twitter などのソーシャルサービス全般を含む。「ポータル」は、Yahoo! JAPAN やアット・ニフティなどのポータルサイトを含む。「企業」は、Apple や東芝などの企業のサイトを含む。「検索」は、Google や bing などの検索サービスを含む。「商品」は、爽健美茶やジョージアなど、商品のキャンペーンサイトを含む。「情報」は、アサヒドットコムやウィキペディアなど、ニュースや特定のトピックに関する情報を提供するサイトを含む。「動画」は、YouTube やニコニコ動画などの動画配信サービスを含む。「旅行」は、JAL や楽天トラベルなどの旅行に関する情報を提供するサイトを含む。それぞれのサイトの数および割合は、Table 2 の通りである。以下では、

Table 2. カテゴリーごとの分析対象サイト数（割合）

カテゴリー	サイト数	(割合)
サービス	6	(4%)
ショッピング	31	(22%)
スポーツ	8	(6%)
ソーシャル	22	(16%)
ニュース	14	(10%)
ポータル	6	(4%)
企業	17	(12%)
検索	5	(4%)
商品	3	(2%)
情報	16	(11%)
動画	4	(3%)
旅行	8	(6%)
計	140	

これらのカテゴリーを用いて分析を行う。

サイトの構造指標 2010年11月3日時点の各サイトのトップページについて、(1)文字数、(2)リンクの数、(3)画像の数、(4)Flashファイルへのリンクの数、(5)<SCRIPT>タグの数、の5つの指標を算出した。(1)の指標は、VBScriptのInternetExplorer.Applicationオブジェクト（Internet Explorer 7）を用いて、document.body.innerTextプロパティからテキストのみを取得し、制御文字と空白を除いた文字数をカウントした。同様に、(2)、(3)の指標はdocument.links.length、document.images.lengthの各プロパティから取得した。(4)、(5)の指標は、トップページのhtmlファイルまたはaspファイルから、swfファイルへのリンク数、<SCRIPT>タグの数をそれぞれカウントした。

結 果

構造指標によるサイトの分類 サイト構造の5つの指標の平均値、標準偏差、および各指標間の相関係数をTable 3に示す。5つの指標のうち、Flash数のみが他の指標と有意な関連を示さなかった。次に、これらの指標がサイトごとの特徴をどのように表すのかを明らかにするために、対数変換後の構造指標に対して主成分分析（バリマックス回転）を行った。分析の結果、累積寄与率が80%を超えていること（85.1%）、および主成分の解釈可能性から、第Ⅲ主成分までを解釈の対象とした。主成分分析の結果をTable 4に示す。

第Ⅰ主成分は、文字数、リンク数、画像数と高い負荷を示し、全体の分散の41%を説明していた。これらは文章や画像など、サイトに含まれる静的なコンテンツについての構

Table 3. 各構造指標の記述統計量および相関係数

	実測値	対数変換後の値	相関係数			
			1	2	3	4
1 文字数	3338.1 (2645.1)	3.41 (0.35)				
2 リンク数	226.5 (144.5)	2.26 (0.31)	.849**			
3 画像数	81.7 (63.8)	1.76 (0.44)	.447**	.500**		
4 Flash数	0.58 (1.05)	0.13 (0.21)	.078	.105	.157	
5 <SCRIPT>タグ数	16.1 (13.10)	1.13 (0.31)	.347**	.389**	.361**	.107

$n=140$; ** $p<.01$.

Table 4. 各構造指標の主成分負荷量（バリマックス回転）

	I： 静的コンテンツ	II： ミックスコンテンツ	III： リッチコンテンツ	共通性
文字数	.94	.13	.00	.89
リンク数	.93	.20	.04	.90
画像数	.54	.47	.19	.55
Flash数	.05	.05	.99	.99
<SCRIPT>タグ数	.18	.95	.02	.93
寄与（分散）	2.07	1.17	1.02	
寄与率	41%	23%	20%	
累積寄与率	41%	65%	85%	

造指標と考えられるため、「静的コンテンツ得点」と命名した。また、第II主成分は、画像数および<SCRIPT>タグ数と高い負荷を示していた。これらは静的な画像とインタラクティブ広告や動画、新着情報などを組み合わせた複合的な構造指標と考えられるため、「ミックスコンテンツ得点」と命名した。第III主成分は、Flash数のみに高い負荷を示していたため、「リッチコンテンツ得点」と命名した。主成分得点の分布と主成分得点間の

散布図を Figure 1 に示す。以下では、Table 5 をもとに、カテゴリーごとのサイト構造の特徴について述べる。

■サービス：全体に静的コンテンツ得点が低く、文字数やリンク数、画像数が多いとはいえなかった。また、佐川急便や日本郵政など、リッチコンテンツ得点の高いサイトがいくつか含まれていた。ミックスコンテンツ得点については、いずれのサイトも目立った特徴は見られなかった。

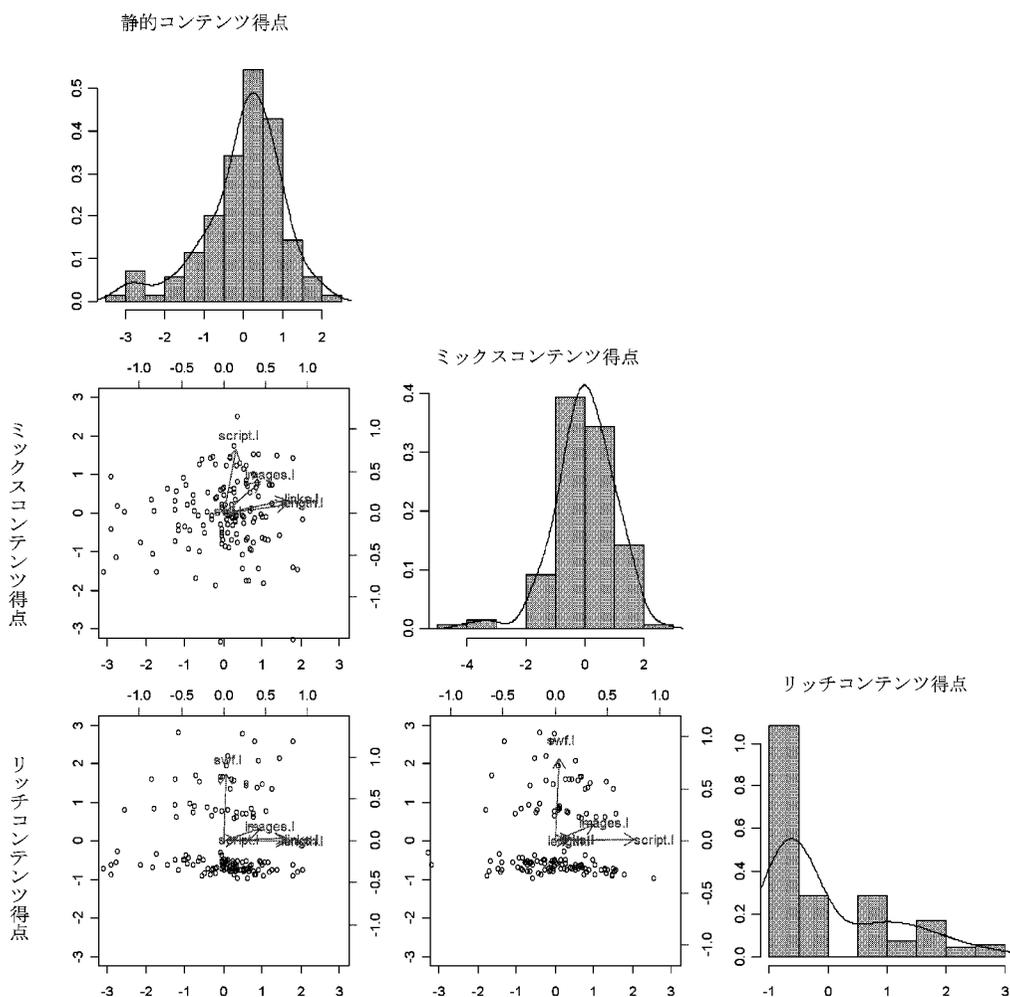


Fig. 1. 各主成分得点の分布と散布図 (矢印 (length.l, images.l, links.l, swf.l, script.l) はそれぞれ対数変換後の文字数, 画像数, リンク数, Flash 数, スクリプト数のベクトルを表す)

■ショッピング：静的コンテンツ得点の高いサイト（ヨドバシ.com, ビックカメラ.com など）、ミックスコンテンツ得点の高いサイト（nissen, セシールなど）、リッチコンテンツ得点の高いサイト（DHC, ネットプライスなど）、というように、サイト毎に特徴的な構造がみられた。

■スポーツ：静的コンテンツ得点が高いサイトは少なく、トップページで文字や画像による情報が多く提供されているわけではなかった。ただし、ミックスコンテンツ得点の高いサイト（ゴルフダイジェスト）や、ミックスコンテンツ得点の高いサイト（Goal.com）は、動的なコンテンツを中心とした情報提供を行っていると考えられる。

■ソーシャル：静的コンテンツ得点の低いサイトが多く、2ちゃんねる（VIP）などを除いて、トップページで文字情報が多いサイトは少なかった。また、リッチコンテンツ得点、ミックスコンテンツ得点が高いサイトは、いくつかの例外（ハンゲームなど）を除いて多くはなかった。

■ニュース：静的コンテンツ得点はいずれも正の値であり、文字、画像、リンクといった情報がサイトの中心であった。REUTERSや時事ドットコムなど、通信社のニュースサイトはミックスコンテンツ得点が高く、Yahoo! ニュースやGoogle ニュースなど、ニュースの二次配布サイトはミックスコンテンツ得点が低かった。また、リッチコンテンツ得点が高いのはNHK オンラインやAFPなど一部のサイトのみであり、全体にはニュース動画がそれほど多用されているわけではなかった。

■ポータル：ミックスコンテンツ得点が高いサイト（@nifty）、リッチコンテンツ得点が高いサイト（BIGLOBE, livedoor）など、サイト毎の特徴が見られた。PV数が最も多いYahoo! JAPANは、ミックスコンテンツ得点が低く、スクリプトをあまり用いずにサ

イト構成を行っていた。

■企業：他のカテゴリーと比較して、いずれのサイトも静的コンテンツ得点が低く、トップページに掲載される文字や画像、リンクは多くはなかった。また、ミックスコンテンツ得点が高いサイト（Apple, TOYOTA）、リッチコンテンツ得点が高いサイト（Adobe, Sony など）もいくつかみられたものの、全体に情報の分量、型とも明確な特徴は見られなかった。

■検索：GoogleやYahoo! 検索など、いずれの検索サイトもすべての主成分得点が低く、他のカテゴリーのサイトと比べて、トップページは多くの情報を含めずにシンプルに構成されていた。

■商品：いずれの商品サイトも静的コンテンツ得点が負の値を示す一方で、他の主成分得点もそれほど高い値ではなく、他のカテゴリーと比較した場合、トップページの情報の分量は多いとはいえなかった。

■情報：いくつかの例外（@cosme, 食べログなど）を除いて、ミックスコンテンツ得点が高いサイト、リッチコンテンツ得点が高いサイトはほとんどなかった。全体としていずれの主成分得点も高いとはいえず、トップページは他のカテゴリーのサイトと比較して平均的な情報の分量・型であった。

■動画：いずれのサイトも静的コンテンツ得点は低く、ミックスコンテンツ得点、リッチコンテンツ得点のいずれか、もしくは両方が正の値を示しており、主に動的な情報によってサイトを構成していることが示された。

■旅行：いずれのサイトも静的コンテンツ得点、ミックスコンテンツ得点が正の値を示し、文字や画像、リンクだけでなく、動画やインタラクティブな検索を含めた情報が総合的に提供されていた。リッチコンテンツ得点については、積極的に使用しているサイト（近畿日本ツーリストなど）とそうでないサイト（JAL など）が同数程度であった。

Web サイトの構造的特徴に関する探索的検討(五十嵐)

Table 5. 対象サイトのカテゴリと主成分得点

サイト名 (URL)	I	II	III
	静的コンテンツ 得点	ミックスコンテンツ 得点	リッチコンテンツ 得点
〈サービス〉			
佐川急便 (http://www.sagawa-exp.co.jp/)	-1.873	0.403	1.639
日本郵政 (http://www.japanpost.jp/)	-1.276	0.365	1.645
Google マップ (http://maps.google.co.jp/)	-0.536	-0.282	-0.440
ヤマト運輸 (http://www.kuronekoyamato.co.jp/)	-0.432	0.078	0.925
Map Fan (http://www.mapfan.com/)	-0.197	0.470	-0.652
Mapion (http://www.mapion.co.jp/)	0.007	0.513	-0.724
〈ショッピング〉			
エコポイント (http://eco-points.jp/)	-0.929	-0.865	-0.377
HOME'S (http://www.homes.co.jp/)	-0.799	0.337	0.777
nissen (http://www.nissen.co.jp/)	-0.286	1.496	-0.669
ショップチャンネル (http://www.shopch.jp/)	-0.076	0.638	1.708
amazon (http://www.amazon.co.jp/)	-0.053	0.212	-0.669
OIOI web channel (http://www.0101.co.jp/)	-0.030	0.686	1.692
www.japanet.co.jp (http://www.japanet.co.jp/)	0.028	0.271	-0.468
セシール (http://www.cecile.co.jp/)	0.061	1.284	-0.679
楽天 BLOG (http://plaza.rakuten.co.jp/)	0.097	0.023	-0.556
au one オークション (http://auok.auone.jp/)	0.107	-0.025	-0.462
モバオク (http://www.mbok.jp/)	0.217	-0.095	-0.484
bidders (http://www.bidders.co.jp/)	0.227	-0.290	-0.555
Goo-net (http://www.goo-net.com/)	0.248	0.668	0.611
coneco.net (http://www.coneco.net/)	0.265	0.584	1.599
X BRAND (http://xbrand.yahoo.co.jp/)	0.323	-0.498	0.797
Yahoo! オークション (http://auctions.yahoo.co.jp/)	0.469	-1.382	-0.521
ネットプライス (http://www.netprice.co.jp/)	0.492	-0.025	2.808
楽天オークション (http://auction.rakuten.co.jp/)	0.514	-0.067	-0.753
EC ナビ (http://ecnavi.jp/)	0.523	-0.201	-0.674
dinos (http://www.dinos.co.jp/)	0.574	0.795	0.740
楽天ブックス (http://books.rakuten.co.jp/)	0.624	0.753	-0.672
価格.com (http://www.kakaku.com/)	0.659	0.589	-0.613
楽天市場 (http://www.rakuten.co.jp/)	0.763	1.563	0.762
DHC (http://www.dhc.co.jp/)	0.788	-0.080	2.642
FELISSIMO (http://www.felissimo.co.jp/)	0.953	0.020	-0.530
Yahoo! ショッピング (http://shopping.yahoo.co.jp/)	1.050	-0.638	-0.518
GDOSHOP (http://www.gdosshop.com/)	1.095	-0.041	0.828
ビックカメラ.com (http://www.biccamera.com/)	1.136	-0.664	-0.525
ASKUL (http://www.askul.co.jp/)	1.186	0.788	-0.783
ZOZORESORT (http://zozo.jp/)	1.797	-1.343	2.638
ヨドバシ.com (http://www.yodobashi.com/)	2.036	-0.111	-0.719
〈スポーツ〉			
日本女子プロゴルフ協会 (http://www.lpga.or.jp/)	-0.634	-0.934	-0.393
JAPAN GOLF TOUR (http://www.jgto.org/)	-0.045	-0.529	-0.338
Yahoo! スポーツ (http://sports.yahoo.co.jp/)	0.135	-0.851	-0.699
ゴルフダイジェスト (http://www.golfdigest.co.jp/)	0.156	1.483	1.386
Yahoo! スポーツ プロ野球 (http://baseball.yahoo.co.jp/)	0.187	-0.668	-0.753
Goal.com (http://www.goal.com/jp/)	0.583	-0.650	1.471
GIANTS (http://www.giants.jp/)	0.589	-1.694	-0.494
阪神タイガース (http://hanshintigers.jp/)	1.435	-0.525	2.190

サイト名 (URL)	I	II	III
	静的コンテンツ 得点	ミックスコンテンツ 得点	リッチコンテンツ 得点
〈ソーシャル〉			
モバゲーTOWN (http://yahoo-mbga.jp/)	-2.144	-0.712	-0.568
facebook (http://www.facebook.com/)	-1.844	-0.991	-0.661
pixiv (http://www.pixiv.net/)	-1.164	-0.285	-0.710
Woopie (http://www.woopie.jp/)	-0.774	0.516	-0.581
@peps! (http://peps.jp/)	-0.735	-1.636	1.742
HANGAME (http://www.hangame.co.jp/)	-0.642	1.323	1.591
FC2 (http://fc2.com/)	-0.494	-0.452	-0.799
yaplog (http://www.yaplog.jp/)	-0.401	0.580	-0.682
Ameba (http://www.ameba.jp/)	-0.311	-0.036	0.696
魔法のiらんど (http://ip.tosp.co.jp/)	-0.226	-1.817	0.859
人気プログラミング (http://blog.with2.net/)	-0.046	-0.368	-0.527
teacup (http://www.teacup.com/)	-0.008	0.710	-0.652
goo ブログ (http://blog.goo.ne.jp/)	0.062	-0.114	-0.554
Twitter (http://twitter.com/)	0.201	-0.046	-0.673
komachi.yomiuri.co.jp (http://komachi.yomiuri.co.jp/)	0.278	-0.143	-0.749
教えて!goo (http://oshiete.goo.ne.jp/)	0.303	0.387	-0.630
mixi (http://mixi.jp/)	0.355	-0.715	-0.487
ウィメンズパーク (http://women.benesse.ne.jp/)	0.529	1.196	-0.696
Yahoo! 知恵袋 (http://chiebukuro.yahoo.co.jp/)	0.644	-1.694	-0.742
はてなダイアリー (http://d.hatena.ne.jp/)	0.705	0.179	-0.574
FC2 ブログ (http://blog.fc2.com/)	0.986	-0.591	-0.576
2ちゃんねる (VIP) (http://yuzuru.2ch.net/news4vip/)	1.026	-1.774	-0.854
〈ニュース〉			
Yahoo! ニュース トピックス (http://dailynews.yahoo.co.jp/)	0.087	-0.476	0.862
NHK オンライン (http://www.nhk.or.jp/)	0.100	-0.246	2.249
時事ドットコム (http://www.jiji.com/)	0.150	1.516	-0.732
Sponichi Annex (http://www.sponichi.co.jp/)	0.292	-0.019	-0.615
REUTERS (http://jp.reuters.com/)	0.370	2.552	-0.931
SANSPO (http://www.sanspo.com/)	0.405	1.334	-0.724
goo ニュース (http://news.goo.ne.jp/)	0.686	-0.869	-0.687
nikkansports (http://www.nikkansports.com/)	0.779	-0.168	-0.741
asahi.com (http://www.asahi.com/)	0.783	1.049	0.664
NIKKEI NET (http://www.nikkei.co.jp/)	0.846	0.723	-0.839
AFP BB News (http://www.afpbb.com/)	0.889	0.501	2.121
Yahoo! ニュース (http://headlines.yahoo.co.jp/)	0.902	-1.400	-0.721
YOMIURI ONLINE (http://www.yomiuri.co.jp/)	1.383	1.551	-0.724
Google ニュース (http://news.google.co.jp/)	1.897	-1.417	-0.777
〈ポータル〉			
@nifty (http://www.nifty.com/)	-0.576	1.451	-0.868
au one (http://auone.jp/)	-0.317	-0.115	-0.715
BIGLOBE (http://www.biglobe.ne.jp/)	0.057	0.078	1.998
goo (http://www.goo.ne.jp/)	0.489	-1.035	0.743
Yahoo! JAPAN (http://www.yahoo.co.jp/)	0.913	-4.044	-0.493
livedoor (http://www.livedoor.com/)	1.002	-0.048	1.529
〈企業〉			
Apple (http://www.apple.com/jp/)	-2.921	1.009	-0.576
TOSHIBA (http://www.toshiba.co.jp/)	-2.568	0.072	0.852
資生堂ウェブサイト (http://www.shiseido.co.jp/)	-1.811	0.112	0.892
FUJITSU (http://www.fmworld.net/)	-1.452	0.685	-0.579
Adobe (http://www.adobe.com/jp/)	-1.180	-0.401	2.834
TOYOTA (http://toyota.jp/)	-1.040	0.961	-0.464
EPSON (http://www.epson.jp/)	-1.000	-0.309	-0.465
三菱東京UFJ (http://www.mufg.jp/)	-0.989	0.772	-0.317

Web サイトの構造的特徴に関する探索的検討(五十嵐)

サイト名 (URL)	I	II	III
	静的コンテンツ 得点	ミックスコンテンツ 得点	リッチコンテンツ 得点
Honda (http://www.honda.co.jp/)	-0.939	0.261	-0.552
KIRIN (http://www.kirin.co.jp/)	-0.899	-0.410	1.008
suntory (http://www.suntory.co.jp/)	-0.744	0.113	0.927
国土交通省 (http://www.mlit.go.jp/)	-0.661	-0.030	-0.488
BUFFALO (http://buffalo.jp/)	0.075	-0.106	-0.492
三井住友銀行 (http://www.smbc.co.jp/)	0.208	0.249	1.651
Sony Japan (http://www.sony.co.jp/)	0.289	0.029	-0.548
みずほ銀行 (http://www.mizuhobank.co.jp/)	0.325	-0.021	-0.614
Microsoft (http://www.microsoft.com/japan/)	1.790	-3.214	-0.586
〈検索〉			
Yahoo! 検索 (http://search.yahoo.co.jp/)	-3.096	-1.475	-0.671
Bing (http://www.bing.com/)	-2.921	-0.369	-0.848
Google 日本 (http://www.google.co.jp/)	-2.788	-1.105	-0.526
goo 検索 (http://search.goo.ne.jp/)	-1.727	-1.468	-0.310
スペースアルク (http://www.alc.co.jp/)	-0.219	1.317	-0.817
〈商品〉			
爽健美茶 (http://www.sokenbicha.jp/)	-2.769	0.235	-0.240
日本コカ・コーラ (http://www.cocacola.co.jp/)	-1.255	-0.689	0.989
ジョージア (http://georgia.jp/)	-0.063	-0.472	-0.494
〈情報〉			
JR 東日本 (http://www.jreast.co.jp/)	-1.275	0.125	-0.452
ekitan (http://www.ekitan.com/)	-0.330	1.480	-0.835
クックパッド (http://cookpad.com/)	-0.229	0.429	-0.668
気象庁 (http://www.jma.go.jp/)	-0.098	-3.298	-0.252
at home web (http://www.athome.co.jp/)	-0.046	-0.760	-0.517
JRA (http://www.jra.go.jp/)	-0.024	-0.658	-0.561
NAVITIME (http://www.navitime.co.jp/)	0.030	-0.803	-0.530
@cosme (http://www.cosme.net/)	0.245	1.791	-0.829
HotPepper (http://www.hotpepper.jp/)	0.278	-0.181	-0.674
ORICON STYLE (http://www.oricon.co.jp/)	0.349	0.858	-0.717
ウィキペディア (http://ja.wikipedia.org/)	0.621	-0.588	-0.940
Walkerplus (http://www.walkerplus.com/)	0.685	0.118	-0.679
Vector (http://www.vector.co.jp/)	0.711	-0.897	0.883
ぐるなび (http://www.gnavi.co.jp/)	0.766	-0.728	-0.756
食べログ (http://r.tabelog.com/)	0.865	0.835	1.422
carview (http://www.carview.co.jp/)	1.235	0.292	-0.703
〈動画〉			
USTREAM (http://www.ustream.tv/)	-1.221	0.619	0.644
ニコニコ動画 (http://www.nicovideo.jp/)	-0.093	-0.248	1.596
YouTube (http://www.youtube.com/)	-0.051	0.277	0.819
Fooooo (http://www.fooooo.com/)	-0.036	0.664	-0.608
〈旅行〉			
宿ふらぎ (http://www.yadoplaza.com/)	0.148	0.644	-0.610
JTB (http://www.jtb.co.jp/)	0.334	1.289	0.664
一休 (http://www.ikyuu.com/)	0.564	1.277	-0.722
近畿日本ツーリスト (http://www.knt.co.jp/)	0.578	0.891	1.537
楽天トラベル (http://travel.rakuten.co.jp/)	0.811	1.025	-0.824
ANA (http://www.ana.co.jp/)	0.906	1.577	-0.701
じゃらん (http://www.jalan.net/)	1.256	0.770	1.392
JAL (http://www.jal.co.jp/)	1.780	1.457	-0.853

考 察

本研究では、PV数の多いサイトのトップページから5つの構造指標を抽出し、その特徴を要約する3つの主成分得点を算出したのち、カテゴリ毎にサイトの構造的特徴を分析した。サイトの構造はHTMLタグで表現されるさまざまな要素から成り立っているが、第I主成分は文字数、画像数、リンク数といったサイトの内容や性質を示す基本情報、すなわち「大きさ」を表す成分(size component)として解釈可能である。また、第II主成分は画像・JavaScript数、第III主成分はFlash数にそれぞれ対応しており、これらはサイトの視覚的な特性やインタラクティブ性、すなわち「形式」を表す成分(shape component)として解釈可能である。大きさ、形式というこれらの主成分の次元は、一般的な主成分分析で見出される傾向と一致する(Jolliffe, 2002)とともに、Germonprez & Zigurs (2003)の示すWebサイトの複雑性の規定因のうち、コンテンツと形式の次元にそれぞれ対応する。したがって、構造指標からWebサイトの分類を行うという本研究のアプローチには、先行研究との理論的な整合性が認められたといえる。

インタラクティブなコンテンツを提供する技術であるJavaScriptとFlashは、それぞれ独立した主成分を構成していた。このことは、動的なコンテンツの扱いで、既存技術の連携ではなく対立が生じていることを示している。第III主成分得点が負であったサイト数は96(全体の69%)にのぼっており、これらのサイトではFlashが1つも用いられていなかった。2011年6月現在、Apple社のiPhoneなど、一部のスマートフォンでは正規の方法でFlashコンテンツを閲覧することができず、サイト作成者がスマートフォンからのシームレスなサイト閲覧を考慮して、バナー広告などでのFlashの採用を控える

動きが出ている。また、企業の検索エンジン最適化(SEO; Search Engine Optimization)対策で、Flashのみを用いたコンテンツは検索エンジンにクロールされにくいとの指摘もある。したがって、Flashはコンテンツの複雑性やユーザビリティという観点よりも、プラットフォームとしての汎用性やサイトの効果性の側面からその利用が減少していると考えられる。一方で、次世代のHTML規格であるHTML5を用いることで、サイト作成者はFlashプラグインを用いずに、タグによって動画や音声、グラフィックを用いることが可能になる。例えば、Adobe社が提供するツールであるWallabyは、FlashコンテンツをHTML5に変換することが可能であり、Tumult社の開発したツールであるHypeは、HTML5でFlashと同等レベルのインタラクティブなコンテンツを簡単に作成できるように設計されている。旧来のソフトウェア資産の継承やブラウザ間の互換問題などもあり、HTML5への完全な移行がいつ行われるかは決まっていないものの、こうした技術革新がサイトの構造や利用者の情報処理過程に与える影響については、今後より詳細に検討を行う必要がある。ただし、現時点では動画を再生するコーデックの規格(H.264, ogg/theora, WebM)が統一されおらず、サイト作成者には利用者の利便性を考慮した設計を行うことが望まれる。

主成分得点には、サイトのカテゴリ毎にいくつか特徴的な傾向が見られた。まず、ショッピング系や旅行系など、利用者が主体的に情報を検索することの多いサイトでは、他のカテゴリのサイトに比べて豊富な情報が提示されていた。静的コンテンツ得点の高いサイトは文章と画像、リンクによる商品説明などで熟慮型の情報を提供し、ミックスコンテンツ得点、リッチコンテンツ得点の高いサイトは、検索などのインタラクティブ性の高いサービスや、動画広告を通じて複雑度

の高い情報を提供していると考えられる。これに対して、ニュース系のサイトでは、型を表すミックスコンテンツ得点、リッチコンテンツ得点の高いサイトが少なかった。これらのサイトは、新聞メディアの特徴を踏襲し、動的なコンテンツで表される複雑な構成ではなく、むしろ文字やリンクを中心とした熟慮型の情報を提供していると考えられる。また、企業系や商品系のサイトは静的コンテンツ得点が低く、シンプルなトップページによってコーポレートイメージや製品イメージを中心とした情報を呈示していると考えられる。情報系のサイトでは、そのほとんどでリッチコンテンツ得点が負の値を示しており、Flashを使わないサイト設計が行われていた。このことは、コンテンツの大部分がFlashで作成されたサイトが減りつつあることを示しており、また上述のように、手軽で即時性の高いスマートフォンからの利用を想定したサイト設計が行われているためとも考えられる。一方で、情報系のサイトはミックスコンテンツ得点も低く、全体として文字やリンクを中心とした情報提示が行われていた。

以上のように、本研究ではWebサイトが呈示する情報の複雑性について、サイトの構造指標から3つの主成分を構成して検討を行った。分析の結果、静的なコンテンツ、動的なコンテンツで表されるサイトの複雑性には、カテゴリによって違いがみられた。こうしたサイト構造の基礎的な検討は、サイトの構造的要因が利用者に情報の過負荷をもたらすプロセスの理解において、重要な意味を持つ。ただし、本研究のアプローチには一定の限界があることにも留意する必要がある。まず、HTMLタグから抽出可能な構造指標は、必ずしも利用者の視覚的特徴と一致しているとは限らない。例えば、画像に含まれる情報量は、画像の大きさや内容によって異なる。また、トップページのみを分析の対象としたことで、ソーシャル系や検索系のサイト

の情報量が割り引かれている可能性もある。実際、ソーシャル系のサイトは主成分得点に目立った特徴が見られなかった。したがって、実際に利用者が閲覧している画面の構造的な特徴を客観的な基準で抽出し、内容の検討を行う必要がある。さらに、コンテンツの内容は日々変化するものであり、こうしたサイト構造の分析は継続して行っていく必要がある。

また、サイトからの情報提供のスタイルとして、Googleなどによる機械検索型のコンテンツの提供とともに、FacebookやTwitterなどのソーシャルメディアを通じたクチコミ・評判ネットワークによる情報提供の流れが近年加速している。ソーシャルメディアを介したネットワークでは、リンク情報を提供する主体はサイトに参加している個人であり、個人間のつながりを通じて、利用者個々の特性や嗜好を反映した情報が提供される。本研究で情報系に分類されたサイト(クックパッド、食べログなど)も、特定のカテゴリの情報をインデックスして提示するだけでなく、クチコミや評判などソーシャルメディアによる情報を取り入れて構成されており、今後は目的の情報を検索して手に入れるスキルだけでなく、情報の価値や発信者の信頼度をいかに見極めるかがますます重要になると考えられる。特に孤独感の高い人にとっては、検索型の情報探索は対人的な懸念を高めないものの、ソーシャルメディアで対人的な情報源の信頼性を適切に判断し、その内容を客観的に評価することは、強い懸念を引き起こす可能性がある。こうした情報系のサイトの「ソーシャル化」がもたらす情報活用のあり方の変化については、孤独感との関連からさらなる検討を行う必要がある。

謝 辞

本研究は、平成22年度電気通信普及財団研究調査助成(研究代表者:五十嵐祐)を受

けて行われた。

引用文献

- Carr, N. (2010). *The Shallows: What the Internet is doing to our brains*. W. W. Norton & Company. (カー, N. G. 篠儀直子 (訳) (2010). ネット・バカ: インターネットが私たちの脳にしていること 青土社.)
- Edmunds, A., & Morris, A. (2000). The problem of information overload in business organisations: A review of the literature. *International Journal of Information Management*, 20, 17-28.
- Germonprez, M., & Zigurs, I. (2003). Causal factors for Web site complexity. *Sprouts: Working Papers on Information Systems*, 3. Retrieved June 20, 2011, from <http://sprouts.aisnet.org/3-13>
- Iyengar, S. (2010). *The art of choosing*. Twelve. (アイエンガー, S. 櫻井祐子 (訳) (2010). 選択の科学 文藝春秋)
- Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.): Springer.
- Jones, G. Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. *Information Systems Research*, 15, 194-210.
- Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, 106, 15583-15587.
- 総務省. (2008). 平成18年度情報流通センサス報告書 <http://www.soumu.go.jp/johotsusintokei/linkdata/ic_sensasu_h18.pdf> (2011年6月25日)
- Vise, D. A., & Malseed, M. (2005). *The Google Story*. Delacort Press. (ヴァイス, D. A., & マルシード, M. 田村理香 (訳) (2006). Google 誕生: ガレージで生まれたサーチ・モンスター イースト・プレス)
- Waddington, P. (1997). Dying for information? A report on the effects of information overload in the UK and worldwide. Retrieved June 21, 2011, from <http://www.cni.org/regconfs/1997/ukoln-content/repor~13.html>