

# Web リサーチのパラダイムシフト - パネル調査から SNS クローリングへ -

小野田哲弥\*1・石神康陽\*2・高津雪乃\*3・小野真穂\*3

Email: onoda@mi.sanno.ac.jp

\*1: 産業能率大学 情報マネジメント学部 教授

\*2: 積水ハウス株式会社

\*3: 産業能率大学 情報マネジメント学部 4年

◎Key Words パネル調査, クローリング, データマイニング, Instagram, ハッシュタグ

## 1. はじめに

『Web2.0<sup>(1)</sup>』が脚光を浴びた2005年頃には、まだ「ネットリサーチ」を邪道と見なし、いわゆる「アンケート」は郵送法や留置法でなければならぬとする説も有力であった。しかしインターネットの普及がさらに進み、社会インフラとしての地位を確固たるものにした今日、マーケティングリサーチのみならず、社会調査においても、異議を唱える研究者は少数派となっているように思われる。

第一筆者は、2006年度から産業能率大学において、3年次生・4年次生を対象にした科目『Webマーケティング』を担当している。大教室の講義型ではなく、備え付きPCが50台ほど設置される情報化教室において行われる演習型授業だ。年度によって若干の増減はあるが、おおむね45人程度の学生が受講するクラスを2クラス担当している。5~6人程度を1グループとしたグループワークを実施し、学期冒頭の「研究計画プレゼン」、中盤の「中間プレゼン」そして最終週の「最終プレゼン」の3回の発表機会を設け、それらをマイルストーンあるいはゴールとして、学生たちがデータ分析や解釈結果の解釈に勤しんでいる。

当該分野はイノベーションが頻繁で、トレンドの移ろいも激しい。そのため、大幅なシラバス改定をこれまでに2回行った。それらを機に、当該授業内容は3つの時期に分けられる。第一期(2006~2011年度)の6年間は、学内共同研究において開発した独自の「SNS<sup>(2)</sup>」や「Webアンケートシステム<sup>(3)</sup>」を用い、授業履修者自身が、調査者と被調査者を兼務するスタイルで運営した。価値観やライフスタイルなどを「リッカート尺度」を用いて調査し、カイ二乗検定や主成分分析・因子分析などを行っていた。

第二期(2012~2018年度)の7年間は、リサーチ会社に調査を委託し、調査会社が保持する登録パネル1万人を対象にした大規模調査を実施し、納品されたデータに対してデータ解析演習を行うスタイルで運営してきた。ここでは、第一筆者の専門である「ロングテールマネジメント<sup>(4)</sup>」を意識し、グループごとに担当ジャンルを決め、数百のアイテムを分析対象として全体構造を可視化した上で、その中身を掘り下げる。そして考察の過程で気づいた問題点やチャンスについて、課題解決やさらなる成功の施策についてプレゼンテーションを行う。

しかし、熊坂<sup>(5)</sup>が2015年頃からすでに指摘しているように、リサーチ会社が行うパネル調査は、従来の質問紙調査の形式をほぼ維持したまま、単に回答と集計の効率化を図っただけに過ぎない面も否めない。そのようなオー

ルドファッションな調査方法で、『Webマーケティング』という科目名に相応しい現代的な消費者の実態を把握できるのかという問題意識を痛切に感じるようになった。

結果として2019年度に3度目となるシラバスの大改定を行い、分析対象データを「Instagramのクローリングデータ」に変更した。本稿執筆中の現段階では、まだ当該授業を完了しておらず、経過報告に留まるが、第二期(2012~2018年度)の「パネル調査」と、現在実施中の第三期(2019年度~)「SNSクローリング」の概要について、シラバス変更の転機となった第二筆者の卒業論文<sup>(6)</sup>を交えながら論じたい。

## 2. 2018年度までの授業内容

### 2.1 第二期(2012~2018年度)のシラバス概要

第二期(2012~2018年度)『Webマーケティング』全14週の授業項目は概ね表1のようであった。

表1. 第二期『Webマーケティング』のシラバス

段階	週	授業項目
I	1	オリエンテーション
	2	調査テーマの検討
	3	調査票 Layout の入稿
	4	研究計画プレゼン準備
	5	★研究計画プレゼン
II	6	集計結果の考察
	7	レイヤー分割
	8	中間プレゼン準備
	9	★中間プレゼン
III	10	相関ルール
	11	自己組織化マップ
	12	系統図の作成
	13	最終プレゼンの準備
	14	★最終プレゼン

学期初めの段階I(第1~5週)は「研究計画プレゼン」をゴールとする。分析対象ジャンルを決め、当該ジャンルの現状を文献調査等によって把握し、数百のアイテム(固有名詞)を用意する。第3週に調査票をリサーチ会社に入稿するが、入稿後すぐに調査結果が得られるわけではない。よって調査実施中の期間を利用して、調査票の設計意図や調査結果に対する仮説探索などを行う。

中期の段階Ⅱ(第6~9週)は「中間プレゼン」をゴールとする。第6週の前にはリサーチ会社から調査結果が納品されるため、それらを用いて、第6週は全体・性別・世代別・地域別ランキングを作成し、属性間の比較を行う。つづく第7週では、ロングテール論<sup>(7)</sup>における「ヘッド」と「テール」の二元論を複数階層に発展させた「レイヤー分割」を実行する。4レイヤーの場合は、レイヤー1は「ヘッド」、レイヤー2は「ややヘッド」、レイヤー3は「ややテール」、レイヤー4は「テール」を意味する。そしてそれらの考察の過程で気づいた当該ジャンルの問題点やチャンスについて、中間プレゼン資料にまとめていく。

終盤の段階Ⅲ(第10~14週)は「最終プレゼン」をゴールとする。前段階Ⅱでは「量」的な分析がメインであったが、段階Ⅲでは「質」的な側面を重視し、どのアイテムとどのアイテムが類似した傾向にあるのか(支持層が同じなのか)を検証する。1対1のマイクロレベルでの関係性は第10週の「相関ルール<sup>(8)</sup>」によって、集合体としてのマクロレベルでの関係性は第11週の「自己組織化マップ<sup>(9)</sup>」によって把握し、両者をもとにグループワークでの解釈を踏まえながら、クラスタリング結果をブラッシュアップし、第13週にはジャンルの俯瞰図である系統図に落とし込む。そして、系統図のどこに問題点やチャンスがあるのかを具体的に示し、当該ジャンルの課題解決や、さらなるビジネス成功に向けた施策提案をスライドに盛り込んで、最終プレゼンに臨む。

2.2 第二期(2012~2018年度)の具体事例と意義

前節が第二期『Webマーケティング』の概要説明であるが、抽象的でイメージしづらいとも思われる。そのため、具体的なジャンルを例に、各プロセスの意義も含めて、改めて説明したい。具体例として取り上げるのは、2018年度に第三筆者が担当した「海外旅行」と第四筆者が担当した「カップ麺」の2ジャンルである。

まず段階Ⅰにおいて重要な調査票の作成だが、「海外旅行」を例にとると、一問目の質問は「あなたがこれまでに海外旅行で訪れた中で特に良かった場所を、10ヶ所までチェックしてください。」であった。当該質問では選択肢として、6つのエリア(アジア、オセアニア、中東・アフリカ、北米、中南米、ヨーロッパ)に分けて作成した計255のアイテムリストを掲載した。つづく二問目では「前問で選んだ地域の良かった理由について、当てはまる項目すべてにチェックしてください。」とし、図1のような複数選択可能なチェックボックス式のグリッド質問とした。行には前問の選択アイテムを表示するようにし、列には10項目(ひと、ことば、買い物、街並み、格安、レジャー、大自然、文化、グルメ、その他)がランダム表示されるように設計した調査票を作成し、リサーチ会社に画面実装を依頼した。

無事に調査が完了すると、リサーチ会社から調査結果が納品される。現在では静的な調査結果データだけでなく、クロス集計を対話的に行うことのできる簡易集計アプリも同時に提供されることが多い。段階Ⅱでは、当該アプリなども活用しながら、属性別の特徴を把握し、レイヤー分割を行う。このプロセスも「海外旅行」を例に説明すると、表2のように属性によって異なるランキング結果

を抽出することができ、学生たちはそこから様々な気づきを得ることができる。



図1. ジャンル「海外旅行」の二問目の例

表2. ジャンル「海外旅行」の属性別ランキングの例

順位	女性	街並み	格安
1	ハワイ	ベネチア	ホーチミン
2	ソウル	バルセロナ	ソウル
3	パリ	パリ	バンコク
4	グアム	ウィーン	釜山
5	台北	フィレンツェ	台北

その上で、基本的には総合支持率をもとにレイヤー分割を行う。ロングテールをグラフ化したものが図2であり、上位レイヤーほど少ないアイテムが、下位レイヤーほど多くのアイテムが該当するようにレイヤーを分割する。

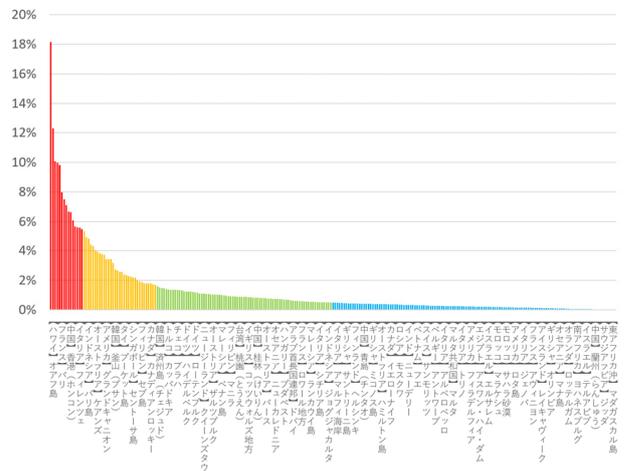


図2. ジャンル「海外旅行」のロングテールグラフ (横軸ラベルの表示間隔は4)

最後の段階Ⅲでは、俯瞰的な系統図を作成する。1対1のマイクロレベルでの関係性は表3の「相関ルール」によって、マクロレベルでの関係性は図3の「自己組織化マップ」によってそれぞれ把握できる。しかしながら両者には矛盾点も存在する。したがってアイテムの最終的な帰属クラスターや、系統のネーミングは、グループワークの議論によって擦り合わせを行い、決着させる必要がある。

表3. ジャンル「海外旅行」の相関ルールの例

前提条件アイテム				共起アイテム				相関ルール				
C.ID	カテゴリー	レイヤー	アイテム名	支持率	C.ID	カテゴリー	レイヤー	アイテム名	支持率	信頼度	リフト (通常の〇倍)	信頼度差 (通常との差)
1	アジア	3	【台湾】桃園(とうえん)	0.88	1	アジア	1	【台湾】台北(タイペイ)	10.07	88.64	8.802	78.57
1	アジア	2	【中国】澳門(マカオ)	2.00	1	アジア	1	【中国】香港(ホンコン)	6.68	79.00	11.826	72.32
1	アジア	2	【シンガポール】セントーサ島	2.24	1	アジア	1	【シンガポール】シンガポール	7.50	79.02	10.536	71.52
6	ヨーロッパ	3	【フランス】ロワール地方	0.59	6	ヨーロッパ	1	【フランス】パリ	9.80	77.97	7.956	68.17
1	アジア	2	【台湾】高雄(たかあ)	2.29	1	アジア	1	【台湾】台北(タイペイ)	10.07	76.42	7.589	66.35
1	アジア	3	【カンボジア】シェムリアプ	0.68	1	アジア	2	【カンボジア】アンコールワット	1.79	67.65	37.792	65.86
3	中東・アフリカ	3	【トルコ】アヤソフィア	0.50	3	中東・アフリカ	3	【トルコ】イスタンブール	1.33	66.00	49.624	64.67
3	中東・アフリカ	3	【トルコ】アヤソフィア	0.50	3	中東・アフリカ	3	【トルコ】カッパドキア	1.39	66.00	47.482	64.61
1	アジア	3	【インドネシア】ウブド	0.61	1	アジア	2	【インドネシア】バリ島	4.94	68.85	13.938	63.91
6	ヨーロッパ	2	【イタリア】ミラノ	3.43	6	ヨーロッパ	1	【イタリア】ローマ	7.10	70.55	9.937	63.45
6	ヨーロッパ	2	【フランス】モンサンミッシェル	2.35	6	ヨーロッパ	1	【フランス】パリ	9.80	73.19	7.469	63.39
3	中東・アフリカ	3	【エジプト】アブシンベル神殿	0.60	3	中東・アフリカ	3	【エジプト】ルクソール	0.88	63.33	71.970	62.45
1	アジア	2	【台湾】九フン(きゅうふん)	3.75	1	アジア	1	【台湾】台北(タイペイ)	10.07	72.00	7.150	61.93
6	ヨーロッパ	2	【イタリア】ナポリ	1.74	6	ヨーロッパ	1	【イタリア】ローマ	7.10	67.24	9.471	60.14

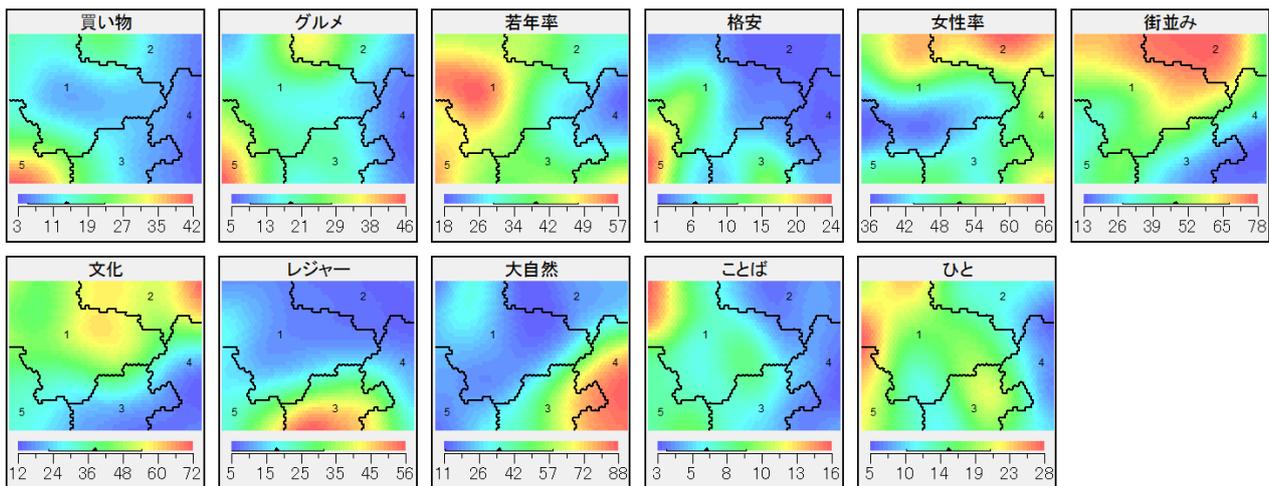


図3. ジャンル「海外旅行」の自己組織化マップの例

以上によってデータ解析は完了である。最終プレゼンで問われるのは着眼点の鋭さや発想の斬新さだ。ジャンル「海外旅行」では、「大学生がまだ知らない卒業旅行にオススメの観光スポット」と題し、「格安」を大前提として、世界各地で「グルメ」「街並み」「レジャー」「大自然」などで好評の観光地を下位レイヤーから抽出し、魅力的なツアープランとして紹介した。

ジャンル「カップ麺」では、「カップやきそば」の地域性、特にゆで汁をスープに利用している商品が人気を博している点(関東圏ではほぼ知られていない「やきそば弁当」が北海道では全アイテムの中で1位、東北でも「焼そばバゴーン」が7位)に注目し、神奈川県西部のご当地カップやきそばを2つ提案した。一つは、キャンパスが立地する伊勢原市の観光名所「大山」で有名な豆腐に着目し、それを具とした味噌汁を付けた「冬季限定大山焼きそば」、そしてもう一つは、釜揚げしらすが有名な湘南発の「湘南塩やきそば」である。

いずれの提案も「ロングテールマネジメント」の趣旨に合致した優れた提案だと、担当教員の第一筆者のみならず、クラスメイトたちからも称賛を得たプレゼンテーションであった。

### 3. 2019年度からの授業内容

#### 3.1 契機となった2018年度卒業論文

第二期の授業内容の学生からの満足度が高く、その点ではシラバス変更の必要性は低かった。だが、冒頭の熊坂

の指摘にあるように、このままではマンネリ化してしまい、時代性との乖離も広がってしまうという危惧を覚えていた。なかなか踏み切りがつかない第一筆者の背中を強く押したのが、二人三脚で研究を進めた第二筆者の卒業論文であった。当該卒業論文の発表を第一筆者が所属する日本マーケティング・サイエンス学会の部会で行ったところ、実務に通じた参加者からも高い評価を得ることができたため、それを自信に、思い切ってシラバス変更には舵を切ることにした。

当該卒業論文は、すでにハウスメーカーに就職を決めていた第二筆者が、注文住宅の営業マンとして、顧客ニーズを把握するための方法論開発を主眼に進めた研究成果である。文献調査として住宅メーカー9社が共同で実施した『イエノミカタ』(<http://ienomikata.com/report/6>)を参照し、ソフト面への拘りが強いのが「妻」であることから、女性を調査ターゲットとして研究を進めることにした。

利用したリソースは、女性利用率の高いクチコミサイト『ガールズちゃんねる』(<https://girlschannel.net/>)と『インスタグラム』(<https://www.instagram.com/>)である。前者では「注文住宅」に関する書き込みをデータベース化してテキストマイニングし、そこからこだわりのポイントが「住宅のパーツ」ごとに異なることを発見した。ゆえに後者の『インスタグラム』では、10種類のハッシュタグ(#リビング、#キッチン、#ガレージ、#トイレ、#バルコニー、#屋上、#玄関、#洗面所、#庭、#和室)でWebク

ローリングを行い、「共起ハッシュタグ」に注目して、前章の『Web マーケティング』同様のプロセスでロングテールの構造化を行い、女性のニーズを8タイプ(①インテリア巡り、②カフェでくつろぐ、③カスタマイズ好き、④趣

味充実、⑤ガーデニング女子、⑥キッチン女子、⑦自分で制作、⑧和を好む)にクラスタリングすることに成功した。図4はその抜粋である。



図4. 「注文住宅」に対する女性ニーズのクラスタリング結果の例

### 3.2 第三期 (2019 年度)のシラバス概要

前節をもとに、現在実施中の第三期(2019 年度～)の『Web マーケティング』の概要が表4になる。段階としての大きな変更点は、段階Ⅱまでで解析は完了し、段階Ⅲはそれまでのデータ分析結果をもとに「施策提案」に集中するようにした点である。

表4. 第三期『Web マーケティング』のシラバス

段階	週	授業項目
Ⅰ	1	オリエンテーション
	2	Web 上の文献調査
	3	第1回 SNS クローリング
	4	研究計画プレゼン準備
	5	★研究計画プレゼン
Ⅱ	6	第2回 SNS クローリング
	7	ハッシュタグのクラスタリング
	8	ユーザのクラスタリング
	9	ライフスタイル分析
	10	中間プレゼン準備
	11	★中間プレゼン
Ⅲ	12	具体的施策の検討
	13	最終プレゼンの準備
	14	★最終プレゼン

基本的には第二期における「アイテム」が第三期では「ハッシュタグ」に変わっただけで、解析としては、ロングテールのグラフ化、レイヤー分割、クラスタリングという同様のプロセスを踏む。SNS クローリング自体は2回行い、第3週の1回目はいわゆる「予備調査」的な位置づけであり、そこで得られた反省や知見を段階Ⅱの第2回クローリングに活かす。

一番大きな違いは第8週から第9週にかけての「ユーザクラスタリング」と「ライフスタイル分析」にある。これまでの「パネル調査」では、基本属性(性別・年齢・都道府県)以外の情報は調査票に盛り込まない限り得ることはできなかった。しかし「SNS クローリング」ではユーザを特定でき、とりわけ『インスタグラム』では、投稿さ

れた他の画像データをもとに、カスタマージャーニーを想起しやすいビジュアル情報が入手できる。

### 4. おわりに

「パラダイムシフト」という大仰なタイトルを付けてしまったが、「パネル調査」「SNS クローリング」双方に強みがあり、補完的に利用されることが望ましい。だが1学期限定の1科目でどちらを重視するかという観点で、本稿では「パネル調査」から「SNS クローリング」への乗り換えについて論じた。まだ初年度も終了していない段階のため教育効果については評価が難しいが、今後も状況の報告を重ね、CIEC 会員からのコメントやアドバイスを踏まえて科目内容の改良を図っていきたい。

### 参考文献

- (1) O'Reilly, Tim (2005), "What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software," <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.
- (2) 斎藤文・盛屋邦彦・長岡健・小野田哲弥 (2008), 「大学授業における双方向参加型学習環境構築の試み」, 『産業能率大学紀要』, 28(2), pp.23-40.
- (3) 郷孝憲・盛屋邦彦・長岡健・小野田哲弥 (2009), 「授業用 Web アンケートシステムの開発と運用」, 『2009 PC Conference 論文集』, 2009, pp. 39-42.
- (4) 小野田哲弥 (2007), 「柔軟な構造化モデル: ロングテール事象のカテゴリーマネジメント支援」, 博士学位論文, 慶應義塾大学大学院政策・メディア研究科.
- (5) 若林靖永・熊坂賢次 (2019), 「研究者に不可欠なベンチャー・スピリット」, 『CIEC 会長インタビュー#1 熊坂賢次さん』, <https://www.cicc.or.jp/special/entry-1190.html>.
- (6) 石神康陽 (2019), 「注文住宅に対する多様な顧客ニーズの構造的把握 - Instagram の共起ハッシュタグ分析に基づいて -」, 卒業論文, 産業能率大学情報マネジメント学部.
- (7) Anderson, Chris (2004), "The Long Tail", Wired Magazine, Oct. 2004, <http://www.wired.com/wired/archive/12.10/tail.html>.
- (8) Agrawal, R., T. Imielinski, and A. N. Swami (1993), "Mining Association Rules between Sets of Items in Large Databases", Proceedings of the ACM SIGMOD Conference on Management of Data, pp.207-216.
- (9) Kohonen, Teuvo (1998), "The self-organizing map," Neurocomputing, 21(1-3), pp.1-6.