

ウェブサイトユーザビリティアンケート評価手法の開発

仲川 薫*¹ 須田 亨*¹ 善方 日出夫*² 松本 啓太*²

The Development of Questionnaire for Evaluating Web Usability

Kaoru Nakagawa, Toru Suda, Hideo Zempo and Keita Matsumoto

Abstract – This document describes the development process of a questionnaire to evaluate the Web Usability in Japanese. This questionnaire is an effective methodology for evaluating Web Usability for the following reasons. First, evaluation of Web Usability involves cost and time effectiveness, so this questionnaire is designed to evaluate Web Usability quickly and effectively. Second, user satisfaction is also required to determine Web Usability. This questionnaire is also a means of gathering quantitative data regarding for satisfaction. However the existing questionnaire was developed in English and is difficult to carry out in Japan. Therefore, we chose to develop a Japanese questionnaire for this study.

Keywords: usability, Web, questionnaire, evaluation, subjective satisfaction

1. 背景

ウェブサイトの普及とその役割の広がりから、ユーザビリティの重要性が増している。現在はヒューリスティック評価やユーザーテストが中心に行われているが、これらの手法はコストと時間がかかるため、開発コストがハードウェアよりも低く開発期間の短いウェブサイトにおいては適応されにくい場面もある。従って、簡潔かつ定量的にウェブサイトのユーザビリティを評価することのできる手法の確立が望まれている。また、特にウェブサイトにおいては、ユーザーの主観的満足度を得ることが重要となる。主観的評価は定量的なデータにより裏付けされることが望ましく、ヒューリスティック評価やユーザーテストでは対応しきれない部分も大きい。

主観的満足度を定量的に把握し、短期間で簡潔にウェブサイトのユーザビリティ評価を行うための手法としてアンケート評価法の有効性が高いと考えられるが、現在日本においてはウェブサイトの評価するための汎用的かつ適正なアンケート項目や評価手法は未発達である。

2. 目的

本研究は、上記のような背景を踏まえ、日本語によるウェブサイトのカテゴリー横断的なユーザビリティ評価のためのアンケート項目を作成し、その運用方法を決定し、簡便かつ有用なアンケート評価法を確立することを目的として行われた。

3. ウェブサイトユーザビリティアンケート評価手法開発

本研究は以下の手順で行われた。

1. アンケート項目案の作成
2. アンケート試行/ヒューリスティック評価実施
3. アンケート結果分析
4. ヒューリスティック評価結果との比較分析
5. アンケートの妥当性の検討、アンケート項目決定
6. 運用方法の検討

3.1. アンケート項目案の作成

まず、既存のユーザビリティ評価のためのアンケート(QUIS、SUMI その他)の項目のレビューを行い、その後、富士通株式会社 デザイン企画部、株式会社イードのそれぞれのユーザビリティについての知見をレビューし、網羅的な評価項目の抽出を行った。

次に、抽出された評価項目を整理し、さらに内容の重複の排除を行い、評価項目を59項目に厳選した。

アンケートの回答方法は、運用場面における回答者の負担や精度を考慮して、59項目それぞれについて、「大変そう思う」～「全くそう思わない」までの5段階評価とした。

この際、開発段階のものを評価する場合、完成度が高くなるにつれて、評価が高くなる可能性があるため、実際に公開されているウェブサイトに対するアンケート項目であるということを前提とした。実際の運用も、この前提を考慮に入れた上で行うものとした。

*¹ 株式会社イード

*¹ IID, Inc.

*² 富士通株式会社

*² FUJITSU LIMITED

3.2. 評価サイト、タスクの選定

評価対象サイトの偏りを防ぐため、既存のサイトを3つのカテゴリー(情報探索系サイト/情報検索系サイト/相互取引系サイト)に集約し、各カテゴリーにつき、2サイトずつ評価対象を選定した。サイトを選定する際は、あらゆる年齢層が何らかの興味を持つことができ、認知度の高いことを条件とした。タスクについては、各サイトで最も汎用的に行われると思われる作業をタスクとして設定した。

3.3. ユーザーによるアンケート試行

アンケート項目の確定、妥当性の検証のために、抽出された59の評価項目を使用した会場アンケートを実施した。

アンケートの総回答者数は53名(男性27名、女性26名)。1人の被験者につき6サイトを評価してもらった。従って、各サイト毎のサンプル数は53、総サンプル数は318となった。

ユーザーには、自由にサイトを使用させるとともに、サイトの特徴に合ったタスクを1つ遂行させ、その後アンケートによる評価を行うよう指示した。その際、アンケート項目自体に対する評価(分かりにくさ、答えにくさ)も同時に行った。

タスク遂行とアンケート回答、アンケート回答チェックなどを含めて全体で1時間半を要した。(実際のアンケート部分は1サイトにつき15分程度であった。)

3.4. ヒューリスティック評価

アンケート実施につづいて、アンケート結果妥当性の検討のために、アンケート同様、6サイトについて専門家(3名)によるヒューリスティック評価を行った。

3.5. アンケート結果解析

3.5.1 解析データの事前処理

解析データの事前処理として、アンケート項目自体の評価(分かりにくさ、答えにくさ)の結果に基づいて、解析データとして不適当な被験者及び項目を除外した。まず、59問中9問に対して「意味がよくわからない」としていた被験者1名を、判断の信頼性の観点から、解析の対象から除外した。次に、解析の対象者52名中、3名(6%)以上が「意味がよくわからない」とした項目は、回答結果の信頼性が低いと判断し、解析の対象から除外した。結果、55項目をその後の解析に使用した。

3.6. 因子分析

3.6.1.1. 因子数の決定

相関行列の1以上の固有値の数は11、累積寄与率が50%を超えるのが第7因子であったので、因子分析を、11から6までの因子数で実行し、最も解釈のしやすかった因子数(=7)を、最終的に採用した。

的に採用した。

使用した手法は以下の通り。

因子負荷行列の推定 → 主因子法

因子の回転 → エカマックス回転¹

因子得点の推定 → アンダーソン・ルービン法²

3.6.1.2. 抽出された7因子の概要

抽出された7因子の概要は以下の通りである。

表 1.7 因子の概要

Table1. Tabular summary of 7 factors

第1因子 好感度	ウェブサイトに対して「いい感じ」を抱いたかどうか。特に「楽しさ」「親しみ」といった方向での個人的・主観的な満足感。
第2因子 操作のわかりやすさ	ウェブサイトを利用しようとするときの操作や手順はわかりやすいか。自分の思うとおりに操作できる感覚。
第3因子 役立ち感	ウェブサイトに対して「これは使える」「役に立つ」という感覚を抱いたかどうか。再訪意向や、他人への推薦意向に直結。
第4因子 構成のわかりやすさ	ウェブサイトの全体構成、階層構造といった空間的なわかりやすさや全体的な統一感はあるか。
第5因子 見易さ	ウェブサイトの視覚的な見易さは十分か。
第6因子 反応のよさ	ウェブサイト上での操作に対する反応や、ウェブサイトの動き具合が、適切でかつ素早い。
第7因子 内容の信頼性	ウェブサイトに掲載されている情報が、内容的にみて信頼できそうか。

3.6.1.3. 評価構造の理解

抽出された7因子は、「ウェブサイトについての主観的な印象や満足についての因子(第1(好感度),3(役立ち感),7(内容の信頼性)因子)」と「ウェブサイトを使うことについて(狭い意味での「使いやすさ」についての因子(第2(操作のわかりやすさ),4(構成のわかりやすさ),5(見やすさ),6(反応のよさ)因子)」の2つにカテゴライズでき、主観的な印象や満足についての因子を、クリアに抽出することができたといえる。

3.6.1.4. 7因子と「使いやすさ」との関係

今回の質問項目の最後に、質問項目と使いやすさの関係を考察するために、ユーザビリティの観点からみた総合的な評価項目となることを想定して、「このウェブサイトは使いやすい」という項目を設定した。

¹ バリマックス回転に比べて、各因子を平準化する(因子間の寄与率の差を小さくする)作用が強い。特に項目数が多いときに有効。

² 相互に無相関で、平均0、標準偏差1の因子得点を求める。

この項目と、7因子との関係を、因子負荷の大きさ(=項目と各因子との相関係数)で確認して見ると、最も大きいのは「構成のわかりやすさ」であり、次いで「操作のわかりやすさ」となっており、ユーザーにとっては、この2つが、ひとことで「使いやすさ」といった場合の、主要な中身であることがわかった。

3.7. 評価スケールの作成

3.7.1. 基本的な考え方

7つの「評価軸」

因子分析によって抽出された7因子を、ウェブサイト評価における、7つの基本的な「評価軸」として捉えることとし、7つの評価軸毎の得点化を行うことによりそのウェブサイトのユーザビリティを評価することとした。

得点化の考え方

7つの評価軸毎に、それを最もよく代表していると考えられる項目を選択した。選択する項目数は、実際の運用場面において、回答者の負荷を考慮し、各評価軸につき、3項目とした。従って総項目数は、7軸×3項目=21項目となる。

項目選択の考え方

項目を選択する際の基準としたのは、因子分析における「共通性の高さ」と、「共通性に占める最大負荷の比率」である。

その他の観点

候補となる項目が多い場合は、あまりに類似の内容を測定しているとみなせる項目からは、代表的な1つのみを採用した。

3.7.2 評価スケールとして使用する21項目(Web Usability Scale)

最終的に以下の21項目が選択された。(▼は逆転項目³の意。)

■好感度

- このウェブサイトのビジュアル表現は楽しい
- このウェブサイトは印象に残る
- このウェブサイトには親しみがわく

■役立ち感

- このウェブサイトではすぐにわたしの欲しい情報が見つかる
- このウェブサイトにはわからない言葉が多く出てくる ▼
- このウェブサイトを使用するのは時間の浪費である ▼

■信頼性

- このウェブサイトに掲載されている内容は信用できる
- このウェブサイトは信頼できる
- このウェブサイトの文章表現は適切である

■操作の分かり易さ

- このウェブサイトの操作手順はシンプルでわかりやすい
- このウェブサイトの使い方はすぐに理解できる
- このウェブサイトでは、次に何をすればよいか迷わない

■構成の分かり易さ

- このウェブサイトには統一感がある
- このウェブサイトはメニューの構成が分かりやすい
- 自分がこのウェブサイト内のどこにいるのかわかりやすい

■見やすさ

- このウェブサイトの文章は読みやすい(行間、文章のレイアウトなど)
- このウェブサイトの絵や図表は見にくい ▼
- このウェブサイトを利用していると、目が疲れる感じがする ▼

■反応の良さ

- このウェブサイトでは、操作に対してすばやい反応が返ってくる
- このウェブサイトを利用しているときに、画面が正しく表示されないことがある ▼
- このウェブサイトを利用しているときに、表示が遅くなったり、途中で止まってしまうことがある ▼

3.7.3. 評価スケールの得点化

評価軸毎の得点計算の考え方

各項目については、運用場面における回答者の負担や精度を考慮し、5~1の5段階尺度とし、得点計算は、運用上の「使い勝手」を考慮し、3項目の素点の単純平均とすることとした。

「逆転項目」の扱い

「逆転項目」については、方向を逆転するために、得られた5段階得点を6から引いたものを使用することとした。「通常の項目」と「逆転項目」の素点の分布形を比較すると、形状には特段の差異はなく、補正を要するほどではないと考えた。また、両者の平均点には、差があるが、尺度としての分かりやすさ、扱いやすさ(実用性)を重視して考えたとき、あえて、補正を要するほどではないと考えた。

評価軸毎の得点計算の方法

7つの評価軸毎に、選択された3項目(各5~1点、逆転項目は逆転)の平均点を、得点とした。更に、7つの得点の平均点を求め、これを「総合点」とし、総合的な評価指標とした。

3.7.4. 取得データによる7つの評価軸の得点計算の検証

上記の方法で、元のデータにさかのぼり、各サンプル各サイト毎の、評価軸毎の得点を計算し、その平均点、標準偏差を求めた。好感度、信頼性を除く5軸は、平均点、標準偏差の差も小さく、素点をそのまま用いている割には、安定している。

信頼性は、今回の6サイトがいずれもブランド力のある有名サイトであったことが、平均点のプラスに影響していると考えられる。好感度は主観的な要素が強く、今回の平均点が低い理由は、特定できなかった。

表2 評価軸毎の平均点と標準偏差

Table 2 Average Score and Standard deviation of each factors

³ 「逆転項目」とは、質問表現が、その評価軸の方向と逆になっている項目。同じ方向の場合は「正問」と呼んでいる。

	平均点	標準偏差
好感度	3.17	0.84
役立ち感	3.60	0.69
内容の信頼性	3.75	0.50
操作の分かりやすさ	3.60	0.76
構成の分かりやすさ	3.46	0.74
見やすさ	3.51	0.65
反応性	3.53	0.78

3.8. ヒューリスティック評価との比較分析

ヒューリスティック評価の結果をアンケート分析結果から抽出された7つの評価軸に沿って分類し、対象ウェブサイトの評価結果の傾向を比較した。

アンケート結果とヒューリスティック評価結果を比較すると、抽出された問題点の傾向は類似しており、アンケートの妥当性が検証された。また、出現した問題点の内容を比較してみると、アンケート結果においては、ヒューリスティックでは測ることのできないユーザーの主観的評価に基づく項目を測定できるということがわかった。

4. アンケート評価スケールの運用方法

アンケートの運用方法については、今後運用していく過程で修正などが必要となるであろうが、現時点では以下のような運用方法が考えられる。

4.1. 既存サイト評価を行う場合

実査方法

ウェブ上でのインターネットアンケート

想定サンプル数

500 サンプル以上

タスク

タスクは、特に評価したい項目がある場合は、その内容をタスクとして設定する。比較評価を行う場合は、可能な限り同様のタスクを課す。

タスク数は、被験者の負担を考慮し、1～2 までが望ましい。

評価対象ウェブ

評価対象ウェブ数は、被験者の負担を考え、3 サイトまでとするのが望ましい。

4.2. プロトタイプ評価を行う場合

実査方法

①ウェブ上でのインターネットアンケート（メンバー専用サイトを使用、パスワードでメンバーを認証）

②会場調査

想定サンプル数

インターネットアンケートの場合：500 サンプル以上

会場調査の場合：約 50 サンプル

タスク

既存サイト評価の際と同様、評価したい内容をタスクとして被験者に課す。

評価対象ウェブ

開発中のウェブサイト

4.3. 集計・分析方法

21 項目それぞれの得点を総合し、サイトの総合評価を行う。

今回抽出された 7 軸(好感度、有用感、信頼性、手順の分かり易さ、構成の分かり易さ、見易さ、反応のよさ)毎の得点を算出し、サイトの特徴を分析する。

サイトの総合評価と、7 軸毎の個別評価を基に、サイトの問題点を探る。

4.4. 効果とメリット

このアンケート項目は、ユーザーの考えるウェブサイトユーザビリティについての評価構造を表わすものであり、各項目毎に問題点を発見できると同時に、7 つの評価軸ごとのバランスを見ることにより、ウェブサイト全体としての問題点も発見できる。

また、本アンケートは、ユーザーが感じる問題点を短時間で抽出できるというシステムであるが、ヒューリスティック評価等の質的評価と組み合わせることにより、さらに効果的なものとなると考えられる。

5. 今後の課題

今後は、このアンケート評価法の運用の過程で、集計、分析の方法の微調整を行うとともに、より有用なアンケート評価法とするために、得られたデータを蓄積し、基準値の設定を行う必要があると考えられる。

6. 参考文献

[1] QUIS <http://www.lap.umd.edu/QUIS/index.html/>

[2] SUMI <http://www.ucc.ie/hfrg/questionnaires/sumi/>