

# 図書館におけるレファレンス質問と Q&A サイトにおける質問の差異

古澤 智裕<sup>†</sup> 松林 麻実子<sup>††</sup> 佐藤 哲司<sup>†††</sup>

<sup>†</sup>筑波大学情報学群知識情報・図書館学類 〒305-8550 つくば市春日 1-2

<sup>††</sup>筑波大学図書館情報メディア系 〒305-8550 つくば市春日 1-2

E-mail: <sup>†</sup> tomohiro.furusawa.629@gmail.com, <sup>††</sup> mamiko@slis.tsukuba.ac.jp, <sup>†††</sup> satoh@ce.slis.tsukuba.ac.jp

あらまし 本研究ではレファレンスサービスと Q&A サイトの性質の違いを明らかにすることを目的として、質問タイプと主題の 2 つの軸による質問の分析を行った。質問タイプの分析には質問者が回答に期待する内容に基づく分類指標を用い、主題の分析には質問文の中心的内容を把握するための分類法を用いた。レファレンスサービスの分析には「レファレンス協同データベース」を、Q&A サイトの分析には「Yahoo!知恵袋」を用いた。分析の結果、両者の特徴として、レファレンスサービスでは資料を求めるような質問が行われており、歴史に関する主題が人気である一方、Q&A サイトでは個人の考えや思いつきを回答に求めるような質問が行われており、恋愛や人間関係に関する主題が人気である、といったことが明らかになった。

キーワード QA サイト, レファレンスサービス, 質問分類, 主題分析

## 1. はじめに

現在、Web の発展に伴って、Web 上でますます多くの情報を入手できるようになってきている。情報の入手先としての Web の存在感は今後もさらに拡大していくと考えられ、図書館はそうした Web の発展による影響を受けつつある状況にある。2003 年、世界でも最大規模の図書館員の組織である OCLC のレポートに “Google is disintermediating the library. (Google が図書館を無用化しつつある)” と掲載された[1]。Web サービスが図書館に取って代わってしまうのではないかという問題意識が顕在化されると、以下に述べるように様々なところで活発に議論が行われるようになった。例えば、2005 年当時の中国国家図書館の副館長である陳は google が図書館にとって代わる可能性について考察している[2]。BROPHY, BAWDEN は大学図書館と google について、資料の質やアクセス率、カバー率などを比較し[3]、JOINT は Web2.0 時代における図書館の問題点を整理しどのように対処すれば良いのかを考察している[4]。

このように図書館と Web についての議論が盛んに行われてきたなかで、図書館業務におけるレファレンスサービスだけは Web にとって代わることはないだろうということが言われてきた。ここで、レファレンスサービスとは「何らかの情報あるいは資料を求めている図書館利用者に対して、図書館員が仲介の立場から、求められている情報あるいは資料を提供ないし提示することによって援助すること、およびそれにかかわる諸業務」[5]であり、レファレンスサービスの過程で利用者から行われる質問がレファレンス質問である。例えば、TENOPIR は Web の発展に伴ってレファレン

スサービスの質問内容は高度になったと述べている[6]。つまり、簡易な疑問は Web で解消できるようになったが、Web で解決できないような高度な疑問はレファレンスサービスが担うとされていた、と考えられる。このようなレファレンスサービスへの信頼は、徐々に「レファレンスサービスこそが図書館の存在意義である」というような意識へと変化していった。そのような意識は例えば、国際基督教大学図書館の「インターネットの発達で図書館の業務は大きく変化した、レファレンス・サービスはインターネットでは置き換えることのできない「最後の砦」と言われていた」[7]といった表現にみることができる。

一方、近年では「Yahoo!知恵袋」や「教えて goo!」などをはじめとする Q&A サイトが急成長している。なお、Q&A サイトとは、「利用者が質問を公開し回答を募って疑問を解消したり、また他の利用者の質問に対して回答したりできるウェブサービス」のことである。この Q&A サイトは、一般の人々が様々な観点から質問するという点において公共図書館におけるレファレンスサービスと機能が類似していると言える。レファレンスサービスと Q&A サイトについての機能の類似性から、多くの研究者によって両者の性質の違いや優劣を明らかにすることが試みられてきた。例えば、齋藤は安易に回答が得られるような情報要求であれば、人々は Q&A サイトを選択することを示唆している[8]。また、辻、楳原、木川田、原らは公共図書館のレファレンスライブラリアンと Q&A サイトに対して同じ質問を行い、結果として両者の正答率に大きな差がないことを示した[9]。さらに、渡邊は公立図書館におけるレファレンス事例の分析結果から、インターネットの登場によって情報入手チャンネルとしての図書館の地

位は低下していると推測し、Q&A サイトの発展に伴い図書館に寄せられるレファレンス質問が難化しているという通説は疑わしくレファレンスサービスへのリソースの投入が今後も継続可能だというのは楽観すぎるとしている[10]。ただし、斉藤は、全国の都道府県立図書館の統計をもとに、Web が発達してもすべての図書館でレファレンス件数が減少傾向にあるわけではないと述べている[11]。

上記のように、特に公共図書館におけるレファレンスサービスと Q&A サイトについては様々な比較がなされてきているが、これまでは全体的に優劣をつけてどちらが生き残るかといった発想に基づいた研究が多く行われてきたと言える。しかし、人々はレファレンスサービスと Q&A サイトに別々のニーズを持って質問を行っているのではないかと考えられる。例えば、レファレンスサービスにおいては図書館が保持する資料を前提とする質問などが行われているのではないかと考えられる。一方で Q&A サイトにおいては回答者・質問者ともに匿名性があることや、質問を広く Web 上で公開できることなどを背景とした質問が行われているのではないかと考えられる。そこで、本研究では、レファレンスサービスと Q&A サイトでは質問者がそもそも異なったものを期待しているのではないかという仮説に基づき、「質問」の側面から両者の違いについて明らかにすることを目的とする。なお、前述の通り、Q&A サイトと特に機能が類似していると思われるのは公共図書館におけるレファレンスサービスであるので、本研究では公共図書館におけるレファレンスサービスを対象とするものとする。

本研究では利用者の期待は質問タイプと質問の主題に現れると考え、この二つの観点に着目して、Q&A サイトとレファレンスサービスの双方の性質の違いを明らかにする。ここで、質問タイプとは質問者が回答に期待する内容を把握するための分類を言い、主題とは質問文における中心的内容のことを言う。このように質問タイプと主題に着目し、Q&A サイトとレファレンスサービスを横断的に比較した研究は、これまで数えるほどしか行われていない。

## 2. 関連研究

### 2.1. レファレンスサービスにおける質問分析

図書館では古くからレファレンス質問の特徴を捉えるための分類指標について議論がなされてきている。長澤は類別のために、受付方法、主題、探索に利用したツールの種類、情報源の種類、質問者、解答様式、処理経過、難易度などの観点があるとしている[12]。近年、利用者のニーズを捉えることを目的とした分類方法も考案されてきている。代表的なのは KATZ によ

るものであり、質問の難易度によって (1) directional reference questions (2) ready reference questions (3) search questions (4) research questions 4 タイプへの分類を提唱している[13]。その後 KATZ の分類を元に、SEARS によって(1) reference questions (2) policy and procedural questions (3) directional questions という分類が作られた[14]。さらに、WARNER によって (1) Nonresources-based (2) Skill-based (3) Strategy-based (4) Consultation といった分類指標が作成され、図書館レファレンスでは(1) Nonresources-based が最も多く、(2) Skill-based が次いで多いということが明らかになった[15]。NUMMINEN, VAKKARI は SEARS による分類指標を用いてフィンランドにおけるデジタルレファレンスの質問タイプについて 1999 年から 2006 年にかけて分析を行った結果、ready reference questions が 33% から 45% に上昇し subject-based research は 57% から 47% へ減少したことが明らかになり、このことから公共図書館の主題検索を担う伝統的な役割が低下しているとしている[16]。TSUJI らは e-mail によるレファレンスと伝統的な対面でのレファレンスの双方を分析することを目的として(1) Ready reference (2) Specific search questions (3) Research questions (4) Policy and procedural questions (5) Directional questions (6) Others の 6 タイプの分類を行った結果、e メールにおける質問では research question 及び地元に関する質問の割合が増えることを明らかにしている[17]。このようにレファレンスサービスでは多くの質問分類指標が用いられているが、これらの分類においては、回答者である図書館員の作業を前提とした、回答者視点の主題や回答の難易度などを基準としたものが多いと言える。

ここ数年では、レファレンスサービスと Q&A サイトの性質の違いに着目し、レファレンスサービスの分析に加え、Q&A サイトの分析も行う研究がなされるようになってきた。SHAH, KITZIE はオンラインの Q&A サイトにおける質問を「social Q&A」と「virtual reference」の 2 カテゴリに分類し、図書館員と学生が両カテゴリにおける質問をどのように評価するか実験した[18]。ZHANG, DENG は Q&A サイトのユーザーに対し調査を行い、Q&A サイトのユーザーによる図書館レファレンスの認識を探った[19]。RADFORD, CONNWAY, SHAH はバーチャルレファレンスと Q&A サイトの両ユーザーに対してインタビュー調査を行い、専門家とエンドユーザーが Q&A サイトに対してどのような認識を持っているかを明らかにし、両者のより良い関係性について考察している[20]。このように、Q&A サイトと図書館レファレンス双方を比較した研究は行われてきているが、質問の特徴に着目し、それらの違いを元に両ユーザーの期待を明らかにするとい

う研究は行われていないといえる。

## 2.2. Q&A サイトにおける質問分析

Q&A サイトの質問を分類し、その傾向をとらえるという研究についても、図書館でのレファレンス質問のタイプ分類と同様に多くなされてきている。KIM らは良い回答を選択するための基準として Q&A サイトの質問に対して内容分析を行い (1) information (2) suggestion (3) opinion (4) others の 4 タイプに分類した結果、(3) opinion タイプの質問が最も多く、次いで (1) information タイプの質問が多いことを明らかにし、そこから質問タイプと回答の関係性について分析している[21]。IGNATOVA らは Q&A サイトの質問に対して (1) Concept Completion (2) Definition (3) Procedural (4) Comparison (5) Disjunctive (6) Verification (7) Quantification (8) Causal (9) General Information Need の 9 つのタイプを用いてアノテーションを行い、(1) Concept Completion タイプの質問が最も多く、(1) Concept Completion (2) Definition (3) Procedural が全体の 8 割以上を占めることが明らかになった[22]。渡邊らは回答者が質問文を選択することを考慮に入れ、質問者が回答に期待する内容を把握するための分類指標として、Q&A サイトにおける質問を (1) 事実 (2) 根拠 (3) 経験 (4) 提案 (5) 意見 の 5 タイプに分類した[23]。

以上のように Q&A サイトをめぐるても、様々な質問の分類基準が考案されてきているが、代表的と言える基準は未だ定まっていない状況にある。

本研究では質問者の意図を捉えることによって、レファレンスサービスと Q&A サイト双方の特徴を捉えることを目的とすることから、質問者が回答に期待する内容の観点から質問を分類する渡辺らの質問タイプを使用する。

## 3. 研究手法

### 3.1. データ

本研究では、レファレンス質問の傾向を探るためにレファレンス協同データベース、Q&A サイトの傾向を探るために Yahoo!知恵袋データをそれぞれ用いた。

#### 3.1.1 レファレンス協同データベース

レファレンス協同データベース[24]は国立国会図書館によって作成、運営されている。このデータベースは多数のレファレンス事例の収集をしており、現時点では日本で最大規模のものである。レファレンス事例、調べ方マニュアル、特別コレクション、参加館プロファイルの 4 種類のデータが登録されている。レファレンス事例では、レファレンス協同データベース参加館

で行われたレファレンスサービスの記録をみることで、計 51,208 件のデータが登録されている (2014 年 11 月末現在) [25]。レファレンス協同データベースには、登録されているデータを検索し取得するためのインタフェース (API) が公開されている。API を用いて公共図書館におけるレファレンス事例の収集を行った結果、45,035 件のデータが得られた (2014 年 11 月末現在)。

#### 3.1.2 Yahoo! 知恵袋データ

Yahoo! 知恵袋とはヤフー株式会社が運営する、インターネット上の質問応答サービスである。サイト上で、日常のあらゆる疑問を他の利用者に質問したり、他の利用者の質問に対して回答したりすることができる。Yahoo!知恵袋データ (第 2 版) [26]は国立情報学研究所がヤフー株式会社から原データの提供を受けて、2011 年 2 月から研究者に提供を行っているデータセットである。データ範囲は 2004 年 4 月 1 日から 2009 年 4 月 7 日で、解決済みの質問が 16,257,413 件、回答が 50,053,894 件含まれている。

上記の Yahoo!知恵袋データから、質問タイプの傾向を把握するために質問を抽出した。今回使用する図書館におけるレファレンス質問の件数は全体で約 45,000 件であり、比較しやすくするため、Yahoo!知恵袋からも同程度の規模のデータを抽出した。データの信頼性を高めるため、できるだけ長い期間からデータを抽出することに注意し、2004 年の 4 月 1 日から 2009 年の 1 月 1 日まで 3 ヶ月ごとに 2000 件ずつ抜き出した計 40,000 件の質問を用いた。

## 3.2 質問タイプによる差異の比較

### 3.2.1. 質問タイプの定義

渡辺らによって作成された質問タイプをもとに、同様にレファレンス質問にも適用可能な形で質問タイプを定義した (表 1)。なお、以下に掲載する質問例の末尾についている「QA」や「図」といった文字列は、その質問が「Q&A サイトで行われた」ものか「図書館で行われた」ものかを示し、番号はそれぞれのデータベースにおける ID を示している。

渡辺らは質問を 2 段階に分けて分類している。質問はその質問者が要求する回答が唯一の正解を持つ場合と、複数の回答が存在する場合に大きく二分することができる。本研究において、前者を「ファクトイド型」質問、後者を「ノン・ファクトイド型」質問と呼ぶ。渡辺らはこの「ファクトイド型」質問を、「事実」と「根拠」の 2 タイプに分類している。それぞれの質問タイプは以下のような例となる。

(例)「オニオンは英語でどう書くのですか?」(事実) [QA2707330]

(例)「手術で全身麻酔をする場合、禁煙しなくてはならないらしいのですが、どうしてなのか教えてください?」(根拠) [QA6224078]

「事実」は質問例のような正解となる事実を含む回答を要求する単純なファクトイド型質問であり、質問例では、“オニオンは英語でどう書くか”という単純な事実を求めている。「根拠」は唯一の回答に加えその根拠も同時に求めるような質問であり、質問例では、“禁煙しなくてはいけないか”という事実に加え、「それがなぜなのか」という根拠も同時に求めている。

一方、「ノンファクトイド型」質問は、「経験」、「提案」、「意見」の3タイプに細分することができ、それぞれの質問タイプは以下のような例である。

(例)「広島県近隣で、心霊スポット知りませんか? そこで、心霊体験ありませんか? ぜひ教えてください。」(経験) [QA1431102]

(例)「私は上半身に比べて下半身が太いので、足を細くする方法を教えてください!」(提案) [QA3897711]

(例)「40万円もする時計を酔っ払って落とす男性って、どう思いますか?」(意見) [QA4887890]

「経験」は回答者の経験がなければ回答ができないような質問であり、質問例の場合“心霊体験”という経験について聞いている。「提案」は問題の解決方法や、情報提供を期待するような質問であり、質問例では“足を細くする方法”を求めているが、回答については複数存在すると考えられる。「意見」は推測や嗜好など、主観に基づいて自由に回答することが可能な質問である。質問例では“どう思うか”ということ求めているが、一般的に正解は存在せず、自由に回答することが可能な質問である。

レファレンス質問についても、Q&Aサイトにおける質問と同様に「ファクトイド型」と「ノン・ファクトイド型」に分けることができる。前者の「ファクトイド型」質問は、レファレンス質問において、特定資料の提示ができる質問と捉えることができる。これは、レファレンスサービスにおいて唯一の正解が定まるようなレファレンス質問は、利用者が特定の資料を探すような質問と対応付けられると考えるからである。「ファクトイド型」質問は「事実」、「根拠」、「経験」の3つに細分され、それぞれの質問例は以下のようになる。

(例)「『日本の市区町村役所・役場経緯度一覧』はあるか。」(事実) [図 1000000076]

(例)「地球をとりまくバンアレン帯(放射能帯)が核実験で破壊されているということが書かれている資料を探している。」(根拠) [図 1000019450]

(例)「『三式弾』について書かれた資料を探している。以前、埼玉県立浦和図書館で見た記憶がある。」(経験) [図 1000020346]

「事実」は単純に資料を求めるような質問であり、質問例では“『日本の市区町村役所・役場経緯度一覧』という資料の所在”を求めている。「根拠」は特定の回答(資料の所在)に加えて根拠も求めるものである。ここでは根拠というものを質問者自身が持つ知識についての裏づけと捉える。質問例で質問者は“地球をとりまくバンアレン帯(放射能帯)が核実験で破壊されている”という知識についての裏づけを求めており、そのような特定の資料を求めている。「経験」は「以前読んだことがある」と言うような利用者の経験を含む質問である。回答者の経験がなければ回答ができず、その経験が直接回答になるような質問は基本的に図書館には存在しえないと考え、「経験」タイプについては回答者の経験ではなく、利用者の経験についての質問とした。そのため、レファレンス質問における「経験」はQAサイトにおける分類とは異なり、「ファクトイド型」の質問に含まれる。

「ノン・ファクトイド型」質問では複数の回答が存在しうるような質問であることより、レファレンス質問においては資料よりも特定主題の知識を求めるような質問に対応すると考えることができる。このような「ノン・ファクトイド型」の質問は、レファレンス質問では「提案」と「意見」に細分することができる。それぞれの質問例は以下のとおりである。

(例)「作曲家のアンリ・デュパルクについて知りたい。」(提案) [図 1000015284]

(例)「個人で事業を始めようと考えているが、参考になる本はあるか。漠然とした構想はあるが、事業の内容ははっきりとは決めておらず、決定に役立つ本を希望している。」(意見) [図 1000139937]

「提案」は質問の中に資料についての言及がなく知識そのものを求めるような質問である。質問例では“作曲家のアンリ・デュパルク”についての情報を要求しており、このような質問は回答者によって回答の幅ができると考えられる。「意見」は図書館員が主観的に回答をすることが可能な質問である。質問例では“参考

になる本”，“役立つ本”といった表現が用いられ，これらは，図書館員による主観的な判断を求めている質問であると考えられる。

表 1 質問タイプの定義

タイプ	Q&A サイト	レファレンスサービス
事実	事象の定義，真実，客観的な理由や手法を問う質問	質問者の中に明確な資料のイメージが存在し，書誌事項などを元にそれらの所在を問う質問
根拠	客観的な根拠，理由を問う質問	質問者が自身の持つ知識について記載された資料を求める質問
経験	回答者の経験や体験がなければ回答できない質問	質問者が以前に経験した事柄に基づき行われる質問（レファレンス質問への解答は図書館員の経験が根拠となることはないため，Q&A サイトとは異なる基準を用いる）
提案	問題の解決方法を問う質問や情報提供を依頼する	資料よりも特定の主題について知識を問う質問
意見	推測，嗜好など，主観的に回答してよい質問	図書館員の主観的な判断を求めている質問

### 3.2.2. 分類器による質問の分類

質問の分類はサポートベクトルマシンを用いて行う。教師データについて，Yahoo!知恵袋の質問分類には CQA 質問評価データ，レファレンス協同データベースにおけるレファレンス質問の分類にはアノテーションデータを用いる。CQA 質問評価データ，アノテーションデータについては次項で説明する。実際の分類には，ツールとして R の e1071 パッケージ[27]を用い，素性として単語の出現頻度を利用した。

### 3.2.3 CQA 質問評価データ

CQA 質問評価データとは，渡邊らの研究において作成されたもので，Yahoo! 知恵袋データセットから抽出した 489 件に対して質問タイプを付与したものである[24]。付与されている質問タイプは，事実，根拠，経験，提案，意見の 5 タイプである。

### 3.2.3. レファレンス質問に対するアノテーション

レファレンス協同データベースの質問を分類するための教師データが存在しないため，本研究において作成した。

ヤフー株式会社の Yahoo!クラウドソーシング（Web 上で行われるタスク解決サービス）を用いて，レファレンス協同データベースにおけるレファレンス質問に対するアノテーションを行った。レファレンス協同データベースの全データをうまく反映するよう，全

45,035 件のレファレンス質問から，1,000 件を無作為抽出した。なお，サンプリングには R による sample 関数を用いた。この 1,000 件のレファレンス質問それぞれに対して 20 名のタイプ判定結果が付くようにタスクを設定し，Yahoo!クラウドソーシングに依頼した。なお同一人物の作業量は 200 件までとした。レファレンス質問のタイプ判定は Web 上の Yahoo!クラウドソーシングのプラットフォームで行ってもらった。判定者は登録しているユーザーである。作業は，表示された質問文に対して 5 つの質問タイプ（3.2.1 質問タイプの定義参照）から必ず 1 つを選択してもらうという形で行った。

各質問に対しては判定者数が最も多かった質問タイプを付与し，最も妥当だと判断された質問タイプを一つに決められない場合，すなわち判定者が同数である質問タイプが複数存在するような場合は「複合型」とした。表 2 にアノテーションの結果を示す。

表 2 アノテーション結果

タイプ	件数	割合
事実	446	44.6%
根拠	41	4.1%
経験	19	1.9%
提案	462	46.2%
意見	4	0.4%
複合型	28	2.8%
合計	1,000	100%

### 3.3. 主題による差異の比較

両者の質問タイプによる比較と平行して，主題による比較も行う。レファレンス協同データベースの各レファレンス事例には「日本十進分類法（NDC）」に基づく主題が付与されている。ただし，中には付与されていないデータもあり，それは本研究では欠損値として扱った。一方，Yahoo!知恵袋においても，各質問に「カテゴリ」と呼ばれる主題が付与されている。これらの主題を元に，レファレンスサービスと Q&A サイトの両者における質問の主題分布を調査した。「NDC」と Yahoo!知恵袋の「カテゴリ」は両者とも 3 段階の階層構造を持っている。今回の研究では両者の 2 階層目である「項」と「中カテゴリ」による主題の比較を行い，さらに主題ごとの質問件数の偏りを探った。主題の集計に際して，レファレンスサービスでは API を用いて取得した質問計 45,035 件，Q&A サイトでは Yahoo!知恵袋データに収録されている質問計 16,257,413 件を対象に行った。

## 4. 結果

### 4.1. 質問タイプによる質問の分類

アノテーションデータを教師データとした分類器を基に、レファレンスデータ 45,035 件を分類した結果を表 3 に示す。なお、分類器の精度は 82.51% となった。なお、精度は、正しく分類された質問の件数を全質問の件数で割ることにより算出した。

表 3 分類器によるレファレンス質問の分類

質問タイプ	件数	割合
事実	21,997	48.8%
根拠	1,054	2.3%
経験	521	1.6%
提案	21,422	47.6%
意見	41	0.1%
合計	45,035	100%

図書館の質問タイプでは「事実」が 48.8% と最も多く、次いで「提案」が 47.5% となっており、これらの 2 つのタイプが全体のほとんどを占めている。一方、「根拠」と「経験」もわずかであるが存在している。「意見」についてはほとんど見られない。

一方、CQA 質問評価データを教師データとして作成した分類器を基に Yahoo!知恵袋のサンプリングデータを分類した結果が以下の表 4 である。なお、分類器の精度は 59.10% となった。

表 4 分類器による Q&A サイトの質問分類

質問タイプ	件数	割合
事実	18,563	46.4%
根拠	727	1.8%
経験	5,945	14.9%
提案	5,000	12.5%
意見	9,765	24.4%
合計	40,000	100%

Q&A サイトの質問タイプでは「事実」が 46.4% と最も多い。次いで「意見」、「経験」、「提案」の順に多くなっている。一方、「根拠」は全体の 1% であり、ほとんどみられなかった。

### 4.2. 主題による質問の分類

#### 4.2.1. レファレンスサービス

レファレンス協同データベースから抽出した公共図書館のレファレンス事例 45,035 件において、NDC 番号が付与されていたレファレンス事例は 38,863 件であった。NDC 番号の付与されていなかったレファレ

ンス事例 6,172 件に関しては、欠損値として扱った。NDC 番号の付与されたレファレンス事例について NDC 番号の件数を集計したところ、51,007 件となった（ひとつのレファレンス事例に複数 NDC 番号が付与されている場合がある）。項では 892 種、目では 100 種の主題が存在した。以降、「001」や「210」といった数値ではなく「歴史」や「日本文学」といった見出しを用いて分析を行っていく。

二次区分集計の結果、レファレンスサービスでは「日本史」に関する質問が 6,447 件であり、2 番目に多い「日本文学」の約 2 倍となっており、圧倒的に多く行われていることが分かった。さらに、「日本文学」や「伝記」、「風俗習慣・民俗学・民族学」、「地理・地誌・紀行」といった主題が多く行われており、人文学系の質問が多くなされていることが明らかになった。これらの人文学系の主題においては提案タイプの質問が多くなされていた。一方で、「音楽・舞踏」や「医学・薬学」といった主題に関する質問も多く行われており、「音楽・舞踏」では提案タイプの質問が、「医学・薬学」では事実タイプの質問がなされる傾向にあることが明らかになった。

人気の主題 1~10 位までで 24,161 件となり、これは全体の 48.3% となり、約半数を占めることが分かった。一方、残りの半数には 90 の主題が含まれており、これらは全 100 主題のうち 90% を占める。これら 90 の主題では、ひとつの主題あたり平均 323.1 件の質問が存在した。つまり、特定の主題に質問数が集中する傾向が認められるが、他の主題が皆無ということではなく、多様な主題の質問も一定数行われていると言える。

#### 4.2.2. Q&A サイト

Yahoo!知恵袋データに収録されているすべての質問 16,257,413 件のカテゴリを集計した結果、全カテゴリ件数は 16,098,580 件となった。中カテゴリでは 108 種、小カテゴリでは 446 種の主題が存在した。

中カテゴリ集計の結果、Yahoo!知恵袋では「恋愛相談、人間関係の悩み」に関する質問が約 120 万件と、2 番目に多い「スポーツ」に約 50 万件の差をつけて、圧倒的に多いことが分かった。また、全体的に「音楽」、「ゲーム」、「芸能人」といったエンターテインメントに関する質問が上位を占めており、「音楽」、「ゲーム」に関しては事実タイプの質問が多くみられ、「芸能人」では意見タイプの質問が多くみられた。一方で「健康、病気、病院」といった回答に専門知識を要する可能性の高い質問も人気であり、そのなかでは事実タイプの質問が比較的多く行われていることが分かった。

また、人気の主題 1~10 位までで 6,148,155 件であるが、これは全体の 38.1% となり、約 4 割を占めるこ

とが明らかになった。残りの質問では 98 の主題が含まれており、これらは全 108 主題のうち 91% を占める。また、これら 98 の主題ではひとつの主題あたり平均 101,534.9 件の質問が存在した。このことから、Yahoo! 知恵袋においても、図書館でのレファレンスサービスと同様に、人気の主題に質問がある程度集中するが、その他の主題も幅広く存在し、それら主題のなかで質問も一定数行われていることが明らかになった。

## 5. 考察

質問タイプの分析結果より、レファレンスサービスでも Q&A サイトでも、正解があるようなファクトイド型の質問が半数を占める一方、残りの半数はノン・ファクトイド型の比較的多様性のある回答が要求される質問であるという構造に大きな違いはないといえる。レファレンスサービスにおいては、上記のようなノン・ファクトイド型質問が半数を占めるが、そのなかでは「提案」タイプの質問が大多数を占め、回答者の考えや思い付きを求めるような質問はほとんどみられない。一方、Q&A サイトでもノン・ファクトイド型質問が半数を占めるが、レファレンスサービスとは反対に、回答者の意見や経験を回答に求めるような「経験」「意見」タイプの質問が多いということが明らかになった。以上より、両者の特徴として、レファレンスサービスでは記述された知識を求める傾向があるのに対し、Q&A サイトでは回答者の考えや思い付きを求めるような質問が多く行われるということが明らかになった。

主題分析の結果より、レファレンスサービスと Q&A サイトでは最も人気な主題に大きな違いが存在することが明らかになった。レファレンスサービスでは「日本史」、Q&A サイトでは「恋愛相談、人間関係の悩み」に関する質問が圧倒的に多く行われていた。さらに、レファレンスサービスでは全体的に人文学系の質問、Q&A サイトでは全体的にエンターテインメントに関する質問が多く行われていた。主題ごとの質問件数の偏りから、両者とも特定の人気主題に質問件数が集中することが明らかになった。一方で、その他の質問も少数ながら幅広く行われており、それらの質問も全体のなかでは大きな割合を占めるということも明らかになった。

レファレンスサービスと Q&A サイトの両者で行われている質問の主題及びその質問タイプの特徴の比較から、レファレンスサービスと Q&A サイトの両者とも「健康・医学」に関する質問など上位で似たような主題の質問が多くあり、それらの質問では質問タイプも似ているものが比較的多くあった。一方で、最もよく行われている質問の主題とその質問タイプには違い

があり、図書館で最も行われている「日本史」に関する質問については「事実」タイプの質問が多くを占める一方で、Q&A サイトで最も多く行われている「恋愛相談、人間関係の悩み」に関する質問については、そのなかで「意見」タイプの割合が多くを占めていたことが明らかになった。

## 6. 結論

本研究ではレファレンスサービスと Q&A サイトの差異を明らかにすることを目的として、質問タイプと主題の 2 つの軸を用いて質問の分析を行った。レファレンスサービスと Q&A サイトは、ファクトイド型／ノン・ファクトイド型の質問で二分される構造や、特定の人気主題に質問が集中するなど、似ている部分も存在することが明らかになった。一方で、両者の間には、回答者自身の考えや思い付きを求めるような質問が多くなされているかどうか、多くの質問が集まる主題、それらの主題での質問タイプの傾向などに明確な違いが存在する。

両者の質問傾向として、レファレンスサービスでは資料を求めるような質問が行われており、歴史に関する主題が人気である一方、Q&A サイトでは個人の考えや思い付きを回答に期待するような質問が行われており、恋愛や人間関係に関する主題が人気である、といったことが明らかになった。これらの傾向の違いから両者のユーザーが別種の期待をレファレンスサービス、Q&A サイトそれぞれに抱いているということが考えられる。以上より、人々はレファレンスサービス及び Q&A サイトをうまく使い分けているのではないかと考えられる。

今回の研究では質問のみの分析を行ったが、今後は質問に解答も含めたより複雑な状態でレファレンスサービスと Q&A サイトの比較分析を行いたい。図書館でのレファレンス質問に対する解答はプロフェッショナルによるひとつの解答であり、Q&A サイトに寄せられる解答は不特定多数の人々による複数の解答である。このような大きな性質の違いがあるため、質問と解答の文言そのものを比較した際、質問のみでは見えてこなかった具体的な解答の傾向や差異が明らかになると考えられる。

## 謝辞

本研究の実施にあたっては、ヤフー株式会社が国立情報学研究所に提供した「Yahoo! 知恵袋データ（第 2 版）」、国立国会図書館による「レファレンス共同データベース」を利用しました。

## 文 献

- [1] De Rosa, Cathy, Lorcan Dempsey, and Alane Wilson. The 2003 OCLC Environmental Scan: Pattern Recognition : a Report to the OCLC Membership. Dublin, Ohio: OCLC Online Computer Library Center, 2004.
- [2] 陳力. 視点: Google と図書館. 情報管理, 2005, 48(5) p.292-294.
- [3] BROPHY, Jan.; BAWDEN, David. Is Google enough? Comparison of an internet search engine with academic library resources. Aslib Proceedings. 2005, p. 498-512.
- [4] JOINT, Nicholas. The web 2.0 challenge to libraries. Library Review. 2009, 58(3), p. 167-175.
- [5] 日本図書館情報学会用語辞典編集委員会. 図書館情報学会用語辞典. 第3版, 丸善株式会社, 2009, 286p.
- [6] Tenopir, C. Virtual reference services in a real world. Libr J. 2001, 126(12), p. 38-40.
- [7] 国際基督教大学図書館. "図書館の崩壊". ICU 国際基督教大学. 2003-10-06. [http://www-lib.icu.ac.jp/en/library\\_blog/date/2003/10/](http://www-lib.icu.ac.jp/en/library_blog/date/2003/10/), (参照 2014-06-09).
- [8] 齋藤泰則. デジタル環境の進展による図書館と利用者との関係の変容 : レファレンスサービスの仲介的機能の展開を中心に. 情報の科学と技術, 2007, 57(9), p. 429-433.
- [9] 辻慶太, 榎原衣恵, 木川田朱美, 原淳之. Q&A サイトと公共図書館レファレンスサービスの正答率比較. 図書館界, 2010, 61(6), p. 594-608.
- [10] 齋藤泰則. "第1章 レファレンスサービスの動向". 利用者志向のレファレンスサービス : その原理と方法. 初版, 勉誠出版. 2009, p. 9.
- [11] 渡邊齊志. 短報 公立図書館におけるレファレンスサービスの意義の再検討. Library and information science. 2011, 66, p. 153-165.
- [12] 長澤雅男. "レファレンスの種類". レファレンスサービス. 丸善株式会社 1995, p.132-133.
- [13] Katz, W. A. "QUESTIONS AND SEARCHES". Introduction to reference work. 8th ed., McGraw-Hill, 2002. p.167-17.
- [14] Sears, J. Chat reference service: An analysis of one semester's data. Issues in science and technology librarianship. 2001, 32, p. 200-206.
- [15] Warner, D. G. A new classification for reference statistics. Reference & User Services Quarterly. 2001, 41(1), p. 51-55.
- [16] Numminen, P.; Vakkari, P. Question types in public libraries' digital reference service in Finland: Comparing 1999 and 2006. Journal of the American Society for Information Science and Technology. 2009, 60(6), p. 1249-1257.
- [17] Tsuji, K.; Arai, S.; Suga, R.; Ikeuchi, A.; Yoshikane, F. Analysis of the Questions Asked through Digital and Face-to-Face Reference Services. US-China Education Review A. 2013, 3(1), p. 51-58.
- [18] Shah, C.; Kitzie, V. Social Q&A and virtual reference? comparing apples and oranges with the help of experts and users. Journal of the American Society for Information Science and Technology. 2012, 63(10), p. 2020-2036.
- [19] Zhang, Y.; Deng, S. Social Q&A vs. library virtual reference: User choices and comparisons. Proceedings of the American Society for Information Science and Technology. 2013, 50(1), p. 1-4.
- [20] Radford, M. L., Connaway, L. S., & Shah, C. Convergence and synergy: Social Q&A meets virtual reference services. Proceedings of the American Society for Information Science and Technology. 2012, 49(1), p. 1-8.
- [21] KIM, Soojung; OH, Jung Sun; OH, Sanghee. Best - answer selection criteria in a social Q&A site from the user - oriented relevance perspective. Proceedings of the American Society for Information Science and Technology. 2007, 44(1), p. 1-15.
- [22] Ignatova, K., Toprak, C., Bernhard, D., Gurevych, I.. Annotating question types in social Q&A sites. In: Tagungsband des GSCL Symposiums 'Sprachtechnologie und eHumanities. 2009. p. 44-49.
- [23] 渡邊直人, 島田諭, 関洋平, 神門典子, 佐藤哲司. "QA コミュニティにおける質問者の期待に基づく質問分類に関する一検討". 第3回データ工学とマネジメントに関するフォーラム. 静岡県伊豆市. 2011-02-27/03-01, 電子情報通信学会, 2011, B5-1.
- [24] 国立国会図書館. "レファレンス協同データベース". レファレンス協同データベース. <http://crd.ndl.go.jp/reference/>, (参照 2014-12-25).
- [25] 国立国会図書館. "レファレンス協同データベース事業累積統計 (平成 26 年 11 月末現在)" レファレンス協同データベース. [http://crd.ndl.go.jp/jp/library/documents/stats\\_201411.pdf](http://crd.ndl.go.jp/jp/library/documents/stats_201411.pdf), (参照 2014-12-25).
- [26] 国立情報学研究所. "「Yahoo! 知恵袋データ (第2版)」の提供について". 国立情報学研究所. [http://www.nii.ac.jp/cscenter/idr/yahoo/chiebr2/Y\\_chiebukuro.html](http://www.nii.ac.jp/cscenter/idr/yahoo/chiebr2/Y_chiebukuro.html), (参照 2014-12-25).
- [27] DIMITRIADOU, "Package e1071". CRAN. <http://cran.r-project.org/web/packages/e1071/index.html>, (参照 2014-12-25).