

## 外国人居住者向け Web サイトのアクセスログ解析による

## 主要な行政サービス情報の特定

筑波大学／つくば市役所 \*有澤 瑠利子 ARISAWA Ruriko  
0120677 筑波大学 繁野 麻衣子 SHIGENO Maiko

## 1. はじめに

多様化する「個」に応じた行政サービスの質の維持・向上には、行政の保有する情報を活用した証拠に基づく政策立案（EBPM）やDX等の取組みによる業務の改善・効率化が不可欠である。デジタル庁が推進するオンライン化の流れを受け、市民一人ひとりに合わせた情報を配信するポータルアプリの提供など、必要な情報を必要な人に提供できるように行政サービスのあり方も見直されている。そのために、市民個々が必要とする情報を把握することが求められる。

本稿では、外国人居住者向けの情報発信の業務改善を目的に、つくば市の外国人居住者向け Web サイトのアクセスログに基づいて、閲覧者の関心の高い情報を特定する。そして、住民基本台帳データに基づく住民の特徴を分析し、関心の集まった背景について考察する。自治体が保有するデータを住民サービスや政策に還元する行政のDX化の取組の事例である。

地方自治体のインターネットを利用した情報発信に関しては、アンケートで評価するものがあるが[1]、アンケートだと情報ニーズの変化は観測しがたい。また、アクセスログから閲覧者の関心を明らかにする研究では、多くは対象分野を定めて分析をしている[2,3]。本稿は、分野は特定せずに対象者を外国人居住者に限定して分析することが特徴である。

## 2. 分析データ

つくば市が管理・運営する外国人居住者向けWebサイト「Welcome to the city of Tsukuba」(図1)から取得した2015年1月1日から2020年12月31日までの6年間のアクセスログを分析対象とする。本サイトの提供言語は英語、中国語、韓国語、やさしい日本語の4言語である。韓国語のページは、2018年5月1日より公開のためデータ取得期間が短い。

まず、Webサイト内の各ページに記載されている内容からページを58のページブロック(PB)に分類した。次に、アクセスログはIPアドレスとユーザーエージェントが一致する場合は同一の閲覧者のログとし、同一閲覧者の連続するアクセス時間が45分を超えないログを集めて一つのセッションとした。セッション



図1 外国人居住者向けWebサイト

ンを決める45分は、この値を変化させたときのセッション数の変化率から定めた。そして、閲覧ページの提供言語が主に単一言語に閉じた閲覧に焦点をあてるため、セッション内の80%以上が同一言語のページであるもののみを抽出し、それらセッションを該当提供言語のグループに分類した。なお、編集者によるアクセスは、対象から除外した。

## 3. 関心の高い情報の特定

最初に、提供言語ごとにPBを分類した。当該PBを含むセッション数、ページあたりの平均滞留時間を算出し、各々降順に並べて累積をとり、累積構成比70%を境として70%以内を上位として4分類して以下のようにラベル付けした。

**高関心**：セッション数、滞留時間ともに上位のPB

**頻出**：セッション数のみ上位のPB

**深掘**：滞留時間のみ上位のPB

**小事**：セッション数、滞留時間ともに上位にないPB

いずれの提供言語でも**小事**ラベルでない9種類のPBを関心の高い内容として抽出した。各言語でのラベルを表1に示す。「病院」は時間外診療や救急要請に関わる情報であり、どの言語でも**高関心**ラベルが必要とされている情報であることがわかる。英語サイトにおいては**高関心**ラベルのPBが「病院」のみで、突出してよく閲覧されていることがわかる。図2に英語サイトでの「病院」と「ニュース」のPBを含むセッション数の推移を示す。新型コロナウイルス感

表1 関心の高いページブロックのラベル (高: 高関心, 頻: 頻出, 深: 深掘, -: 小事)

閲覧者集合	全閲覧者				PB数, 閲覧時間ともに上位				PB数のみ上位				閲覧時間のみに上位				PB数, 閲覧時間ともに上位でない			
	英	中	韓	日	英	中	韓	日	英	中	韓	日	英	中	韓	日	英	中	韓	日
PB\提供言語	英	中	韓	日	英	中	韓	日	英	中	韓	日	英	中	韓	日	英	中	韓	日
病院	高	高	高	深	高	深	頻	-	頻	頻	深	頻	高	高	高	-	高	高	高	深
ニュース	頻	高	高	高	高	高	高	高	高	高	高	高	-	頻	高	頻	頻	頻	高	高
暮らし	頻	高	高	高	頻	高	高	高	頻	高	高	高	-	高	高	頻	頻	高	頻	高
私のつくば市	頻	高	頻	高	高	高	-	高	頻	高	頻	高	-	頻	-	深	-	頻	頻	-
スポーツ施設	深	深	深	深	-	深	深	深	-	-	-	深	-	-	深	-	-	深	-	-
日本語を学ぶ	頻	高	深	頻	-	高	深	-	頻	頻	-	頻	-	-	-	深	頻	頻	-	深
英語広報紙	深	頻	高	高	高	高	高	高	深	深	高	高	-	深	深	深	-	深	頻	深
葡語広報紙	深	深	高	高	高	高	高	高	深	深	高	高	-	頻	頻	-	-	深	頻	-
中国語広報紙	深	頻	頻	高	深	高	高	高	深	深	頻	高	-	-	深	深	-	深	頻	-

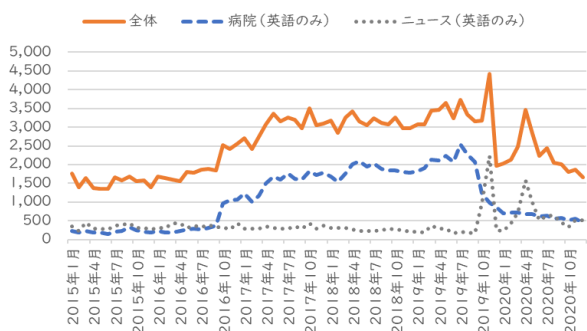


図2 英語サイトのセッション数推移

感染症の拡大に伴い「病院」を含むセッション数は減少し、「ニュース」を含むセッション数が増えている。「病院」への関心は感染症の影響でないものといえる。中国語サイトでは「日本語で学ぶ」は高関心ラベルで、言語障壁に対する支援サービスに関心があるといえる。これは、中国籍の4割の在留資格が留学であり、その多くが2年以内で転出しているため、来日の際に確認している可能性が考えられる。合わせて転入や転出に関するPBも高関心ラベルであった。外国語広報紙は8言語で提供しているが、提供言語サイトでその言語の広報紙を中心に閲覧しているということではなく、図3に示すようにすべての言語広報紙にアクセスが分散している、想定外の使い方がなされている可能性がうかがえる。

次に、提供言語でグループ化したセッションに基づき、閲覧者ごとに閲覧したPB数と累計閲覧時間を算出し、累積構成比70%を境として70%以内を上位として4分類した。この4分類ごとに再度PBにラベル付をして閲覧者の特性別の関心の高いPBを観察した(表1)。PB数、閲覧時間ともに上位の積極的閲覧者内では、高関心ラベルのPBが多く幅広い情報を求めていることがわかる。PB数のみ上位の閲覧者では「ニュース」が高関心ラベルで、卒読の傾向のある閲覧者であることがわかる。これら以外の閲覧者では「病院」への関心が顕著に現れた。

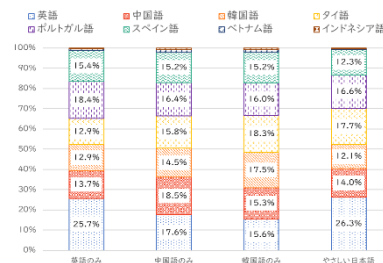


図3 言語別外国語広報紙のアクセス内訳

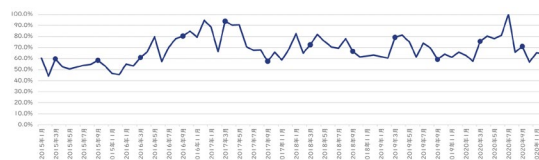


図4 「暮らし」の月次セッション数推移

「暮らし」は、ごみの分別方法が中心のPBである。転入・転出が集中する3月や9月のアクセス数が増加することが想定されるが、図4に示す月次セッション数の最大値比の推移では、印のある3月、9月に限らず、安定して推移していることが分かる。このことから、当該PBの情報は、必要に応じて都度閲覧されている可能性が高いことがうかがえる。

#### 4. おわりに

本稿で紹介した以外にも、PB間の遷移も併せて観察することで、関心の高い情報について、その閲覧の傾向や情報の必要なタイミングが明らかになり、今後、行政サービスの改善に役立つと期待される。

#### 参考文献

- [1] 馬場, 福田:外国人支援から見た地方自治体のWebサイト-多文化共生とICT-. 日本社会情報学会会誌, 21-1 (2009), 5-11.
- [2] 瀧本, 橋本:小さな自治体向け災害時支援システムの開発~情報の整理, 共有, 発信に関して~. 地域安全学会論文集, 4(2002), 335-344.
- [3] 後藤:福祉における情報の壁-自治体の事例調査に基づく考察-. 社会政策, 9-2(2017), 135-146.