

An Access Analysis Tool Based on an Access Analysis Program Originally Developed for Food-Insect Site

メタデータ	言語: jpn 出版者: 公開日: 2019-12-20 キーワード (Ja): キーワード (En): Access Analysis, Food-Insect Site, Radiation-Influence Site, Access Behavior, Website Operation 作成者: 曲山, 幸生, 七里, 与子, 宮ノ下, 明大, 今村, 太郎, 古井, 聡, 和田, 有史, 増田, 知尋 メールアドレス: 所属:
URL	https://doi.org/10.24514/00002947

技術報告

食品害虫サイト用に開発したアクセス解析プログラムとそのツール化

曲山幸生*, 七里与子, 宮ノ下明大, 今村太郎, 古井聡, 和田有史, 増田知尋

独立行政法人農業・食品産業技術総合研究機構 食品総合研究所

キーワード：アクセス解析, 食品害虫サイト, 放射線影響サイト, アクセス行動, ウェブサイト運営

**An Access Analysis Tool Based on an Access Analysis Program Originally
Developed for Food-Insect Site**

Yukio Magariyama*, Kumiko Shichiri, Akihiro Miyanoshita, Taro Imamura, Satoshi Furui,
Yuji Wada, and Tomohiro Masuda

National Food Research Institute, National Agriculture and Food Research Organization
2-1-12 Kannondai, Tsukuba, Ibaraki 305-862, Japan

Abstract

We had originally developed a computer program to analyze Food-Insect Site more than 4 years ago, and have been using and improving the program ever since. By separating the part specified for Food-Insect Site from the computer program to form a configuration file, we made an access analysis tool that could be used to analyze the other websites more easily. The tool has the advantage of giving information about a visitor's movement among webpages in addition to the number of views of a webpage. To show the value of the tool, we did an access analysis of Radiation-Influence Site. The results showed that the proportion of visitors who read several pages during one visit of Food-Insect Site was larger than that of Radiation-Influence Site. The documents about the design of the tool are already available.

Key words: Access Analysis, Food-Insect Site, Radiation-Influence Site, Access Behavior, Website Operation

ウェブは現在ほとんどの組織が利用する重要なメディアであると同時に、新聞、テレビ、ラジオ等に比べると手軽に情報発信できるために、情報の内容や見やすさ、使いやすさ等の質に関して玉石混淆である。しかし、ウェブサーバに記録された閲覧記録を解析する（アクセス解析）ことによって、運営しているウェブサイトを客観的に評価して改善することができる。

著者ら（食品害虫サイト運営グループ）は、2007年より食品害虫サイト^{a,1)}を運営しているが、早期にアクセス解析を導入しサイトの改善等に利用してきた。2010年から独自にアクセス解析ツールを開発し、これまでに食品害虫サイトの運営に必要な機能を追加するなどの改良を重ねた。この独自開発アクセス解析プログラムは他のサイトの運営においても利用できるのではないかと考え、プログラムの中からウェブサイト固有の情報を独立させて簡単に設定変更できるようにした。これにより食品害虫サイト以外にも適用できるアクセス解析ツールが完成した。

本技術報告では、食品害虫サイトにおけるアクセス解析が果たした役割を紹介した後、今回作成したアクセス解析ツールの利用方法について、放射線影響サイト^{1,10)}への適用例を交えて、述べる。

アクセス解析ツールの利用準備

1. アクセスログに含まれる情報

ここで報告するアクセス解析ツールで採用している方法は「サーバログ型」である³⁾。他に、「パケットキャプチャ型」や「ウェブビーコン型」等があり、それぞれに特徴がある⁴⁾。サーバログ型を採用した最大の理由は、導入時に大きな作業や費用が発生しないこと（パケットキャプチャ型ではウェブサーバに出入りする情報を入手するためのハードウェアを追加する、ウェブビーコン型ではすべてのウェブページのソースコードにデータ収集サーバに情報を送信するタグを埋め込むことが必要である）と、過去にさかのぼって解析できることの2点である。これらは試験的にアクセス解析ツールを導入する場合に重要な要件である。

食品害虫サイトと放射線影響サイトのウェブサーバはウェブサーバソフトウェアとしてApacheを使用して

いる。そのアクセスログの書式はCombined Log Formatで、サーバの1作業ごとに次の情報セットが記録されている¹¹⁾。

①%h	ユーザのIPアドレス
②%l	(使用しない)
③%u	ユーザID
④%t	処理時刻
⑤¥"%r¥"	リクエスト
⑥%>s	HTTPステータスコード
⑦%b	転送容量(バイト)
⑧¥"%{Referer}i¥"	参照元情報
⑨¥"%{User-Agent}i¥"	ユーザエージェント

すべての解析結果はこの情報から導いたものになるので、ブラウザの戻るボタンの操作などサーバが関知しない情報を直接得ることはできない。

2. アクセスログの入手

食品害虫サイトは食品総合研究所ウェブサイト（農研機構ウェブサイトの一部）に含まれているので、農研機構本部広報部門を通じ農研機構ウェブサーバアクセスログを入手した。2014年8月時点で放射線影響サイトは食品総合研究所ウェブサイトと農林水産研究情報総合センターのバーチャルラボシステムに分かれて配置されている。著者(曲山)は放射線影響サイト(バーチャルラボシステム部分)のサーバ管理者の一人でもあるので、バーチャルラボシステムのウェブサーバアクセスログは直接サーバから読み出した。

農研機構ウェブサーバログファイルは毎月1回gzip¹²⁾で圧縮して授受することにしており、解凍には圧縮解凍ソフト7-zip¹³⁾を用いた。

3. 解析プログラム

巨大容量ファイルを取り扱うことができるエディタEmEditor¹⁴⁾を用いて、アクセスログファイルの月の境界の微調整と解析プログラムに入力できる形式への変換を行った。解析プログラムはExcel Macro (VBA)を用いて作成し、結果もExcelシートに出力するようにした。

a 食品害虫サイト：貯穀害虫・天敵図鑑を中心に、害虫の防除情報やニュース・コラム等、多様な情報を提供するウェブサイト
b 放射線影響サイト：放射性物質の食品への影響に関する情報を提供するために、東日本大震災に伴う原子力発電所事故後に食品総合研究所が開設したウェブサイト¹⁾。関係する情報へのリンク集と関連論文集²⁾にもまとめられている）から構成される。

食品害虫サイトで利用されたアクセス解析

2007年開設の食品害虫サイト¹⁾は、現在では食品総合研究所ウェブサイトの中で閲覧数の多いページを多く含み、多様な情報を提供する人気コンテンツの一つになっている。この発展に貢献した要因はいくつかあるが、アクセス解析もその一つである。以下に食品害虫サイトの取組みの歴史の中で、アクセス解析が果たした役割を追ってみたい(図1)。

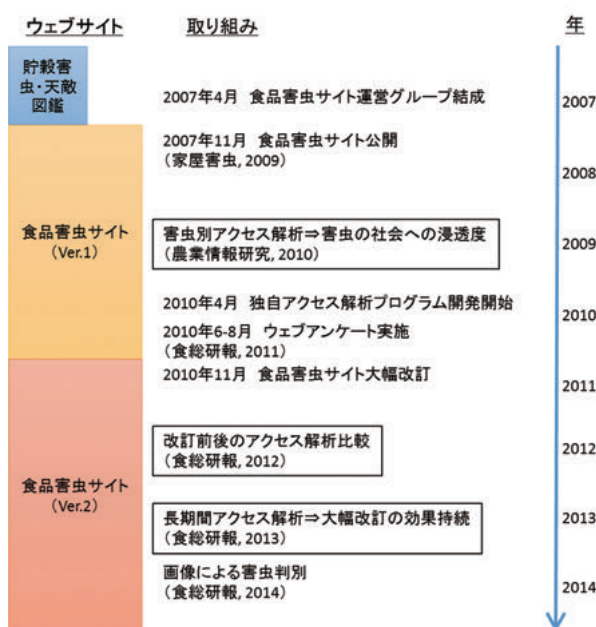


図1. 食品害虫サイトの取組みの歴史

1. アクセス解析で示された害虫の社会への浸透度

食品害虫サイトは2007年11月にインターネット図鑑「貯穀害虫・天敵図鑑」を拡張する形で開設された⁵⁾。開設の初期に図鑑の各ページのアクセス解析結果を比較することによって、下記の新たな発見があった。貯穀害虫・天敵図鑑に掲載されている食品害虫の各ページについて、閲覧数(対数)と検索サイト経由の比率のグラフにプロットすると、右肩上がりの直線の周辺に分布した。ただし、コクゾウムシとノシメマダラメイガは検索サイト経由の比率が低く、この分布から大きくはずれていた⁶⁾。つまり、この2種はすでに多くの人に知られていて、貯穀害虫・天敵図鑑にリンクを張っているウェブページが多く存在しているなど、社会に十分浸透していると推測された。ある食品害虫に注目した場合、図鑑ページの閲覧数はその時点における社会の注目度を反映していると考えられるが、それ以外に社会への浸透度(一過性の注目か長い年月にわたって

の注目か)までアクセス解析でわかることが示された。

2. 長期間アクセス行動に影響するウェブサイトの設計

食品害虫サイトの目的は、貯穀害虫・天敵図鑑を利用しやすくすることに加えて、食品害虫に関する食品総合研究所における研究成果とその周辺情報を提供することである。食品害虫サイトの訪問者を対象に2010年6月から8月に実施したウェブアンケートによれば、これらの目的は概ね達成していたが、貯穀害虫・天敵図鑑とその他のコンテンツの間の関連性を強くすることで、図鑑の訪問者を他のコンテンツに誘導する確率をさらに高くできると考えられた⁷⁾。そこで、図鑑と他のコンテンツの間の関連性の強化を目指し、2010年11月に食品害虫サイトの大幅改訂を実施した。改訂の効果をアクセス解析により解析したところ、図鑑から他のコンテンツへの移動、その逆の他のコンテンツから図鑑への移動がともに増加し、大幅改訂が食品害虫サイト訪問者のアクセス行動に変化をもたらしたことが明らかになった⁸⁾。この変化したアクセス行動は大幅改訂後1年以上経過しても継続していることが、アクセス解析結果によって示されている³⁾。これらの結果は、ウェブサイトのコンテンツ設計の変更により訪問者のアクセス行動が影響を受けることと、その影響が長期間継続することを示している。

3. 食品害虫サイトにおけるアクセス解析の役割

上述した食品害虫サイトの歴史の中で、アクセス解析はコンテンツの定期的な更新(ほぼ毎月)とともに重要な役割を果たした。作成した食品害虫サイトがユーザにどのように受け取られているのか、客観的なデータを獲得するために利用できる技術は少ない。例えば、ウェブアンケートの場合、ある程度詳細にユーザの意識を調査することができるが、アンケート回答者はモチベーションの高い訪問者に限られていること等、サンプルに偏りがあると考えられ、調査結果が食品害虫サイト訪問者全体の傾向を反映しているとは言えない。一方、アクセス解析の場合、獲得できる情報の種類は限られているものの、食品害虫サイト訪問者全員のアクセス行動情報を入手できるというメリットがある。食品害虫サイトにおいては、アクセス解析による閲覧パターン解析から、サイトの問題点を抽出し改善の方向を決定してきた。

食品害虫サイトでは独自に開発したアクセス解析プログラムを利用している。当初は農研機構が提供していたアクセス解析ツールを利用していた。しかし、訪

問者のサイト内の移動経路等の情報が得られなかったことや、ツール自体が変更されたことがあった。そこで、必要な情報を安定して得るために2010年より独自開発アクセス解析プログラムを利用し、開発後もプログラムに改良を加えてきた。2011年から食品害虫サイトのコンテンツとしてアクセス解析結果（人気害虫、人気ページランキング等）も掲載している^{viii)}。

食品害虫サイト用に独自に開発したプログラムのツール化

食品害虫サイトの客観的評価のために2010年より独自に開発してきたアクセス解析プログラムは現在まで日常的に利用しており、この2年ほどは大きなトラブルもなく動作している状況である。今回、このプログラム中のソースコードから食品害虫サイト専用に記述された部分を外部設定シートとして分離し、プログラムを汎用的に利用できるように修正した。ここでは作成したアクセス解析ツールの機能と具体的な使用方法について説明する。

1. アクセス解析ツールの解析項目

アクセス解析プログラムは次の流れ（図2）に沿って解析を進めていく。

- ① 読み込み：アクセスログファイルから食品害虫サイトに関する行だけを読み込み、情報の項目別に分解して、解析結果ファイルのシートlogに書き込む。この工程で、ユーザのIPアドレスとユーザエージェントの情報から各ユーザを区別し、固有の名前を付けてホストとした。データは時刻順に並んでいる。
- ② 並べ替え：シートlogの同一日内のデータをもとにホストをキーにして昇順に並べ替え、結果をシートsortに書き込む。
- ③ 訪問解析：同一ホストの30分以内のアクセスを1訪問と解釈し、結果をシートpath0に書き込む。次に、直前の閲覧ページと参照元が異なる場合はユーザがアクセスログに記録されないブラウザ操作を行ったと仮定して、その間を矛盾なく連結する仮想アクセスを追加する。この結果はシートpathに書き込む。
- ④ 参照元解析：シートpathのデータを使って、参照元情報から1訪問ごとの参照元分類（農研機構サイト、検索サイト、その他の外部サイト、不明）、参照元

- URL、検索サイトの場合キーワードの3つのデータを抽出し、結果をシートrefererに書き込む。
- ⑤ トレンド解析：サイト内のすべてのページについて1日の閲覧数をカウントし、結果をシートtrendに書き込む。
- ⑥ ページ間結合解析：シートpathのデータを使って、ページ間移動（参照元ページ→閲覧ページ）の回数をカウントし、結果をシートconnectionに書き込む。
- ⑦ 順位解析：次の項目について回数の多いものから順位付けし、結果をシートrankに書き込む。解析する項目：参照元ページ、閲覧ページ、ページ間移動、訪問の参照元ドメイン、入口ページ、出口ページ。
- ⑧ 報告書作成：人気ページ、参照元情報、1訪問閲覧ページ数、入口ページと出口ページのランキングをグラフ化したものを、A4サイズ1ページで表示できるようにまとめたものをシートReportに書き込む。

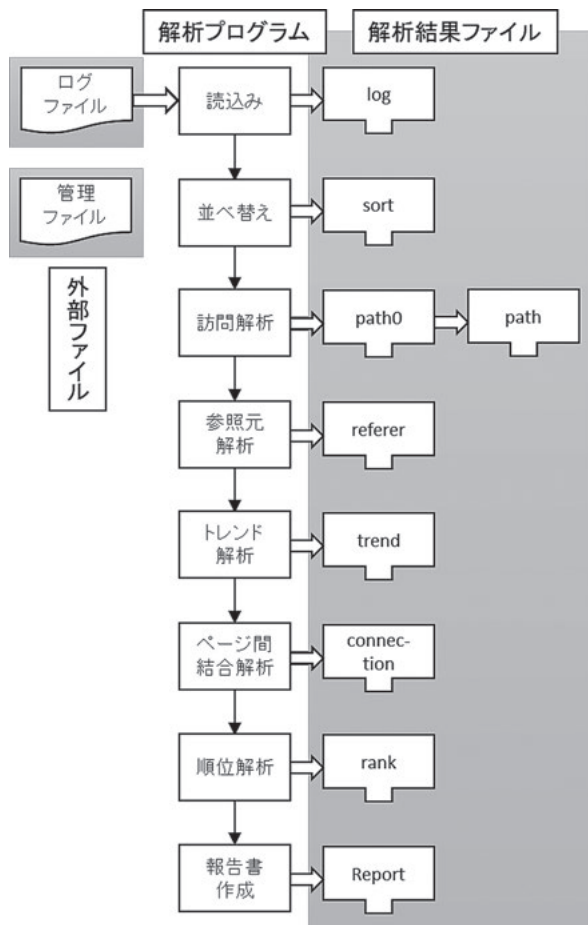


図2. 解析プログラムの動作の流れ

2. アクセス解析手順

食品害虫サイトのアクセス解析手順（図3）で関係するファイルは下記のとおりである。

- ① Lists.xlsm: サイト管理情報等を格納した管理ファイル。
- ② access_combined_log.yyyyymm.gz: ウェブサーバから出力されたままのyyyy年mm月のアクセスログファイル（圧縮版）。
- ③ access_combined_log_yyyyymm_naro.log: アクセス解析プログラムの解析対象となる修正済みアクセスログファイル。
- ④ macro_Foodinsect.xlsm: アクセス解析プログラム（本体）。
- ⑤ Analysis_Food_yyyyymm.xlsx: yyyy年mm月の解析結果ファイル。

次に、具体的な解析手順を簡単に説明する。

(ア) 管理ファイルLists.xlsmの更新

最初に、サイト管理情報等が格納された管理ファイルLists.xlsmを最新情報に更新する。初めてアクセス解析するウェブサイトはもちろん、毎月解析をしているウェブサイトでも、新しいページが追加された場合等にはこの作業が必要になる。

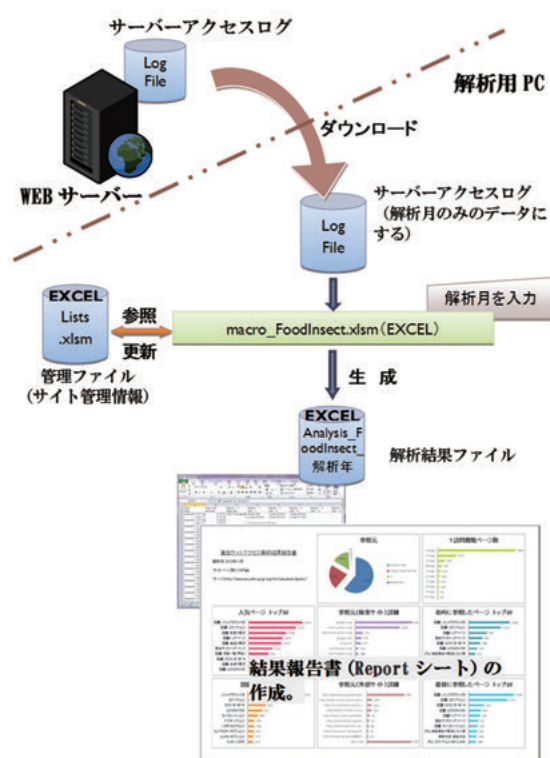


図3. 食品害虫サイトのアクセス解析手順の概略

このファイルには次の種類の情報が格納されている。

- ① Server: サーバ情報
サイト格納サーバ, サイト格納フォルダ, 開始日, 終了日, 有効, サイト説明, 備考。
- ② Pages: サイトのすべてのページ一覧
HTMLファイル名, ページ名称, 開始日, 終了日, 有効, HTMLファイル名 (パス付), 備考。
- ③ Insects: 貯穀害虫・天敵図鑑の害虫ページ一覧
害虫名, (未使用), 開始日, 終了日, 有効, 害虫名, HTMLファイル名, 図鑑番号。
- ④ Robots: ロボット (クローラ) 一覧
識別用文字列, (未使用), ロボット名。
- ⑤ SearchSites: 検索サイト一覧
識別用文字列, キーワードのタグ, (未使用), 検索サイト名。
- ⑥ Queries: 漢字変換一覧
ヒット件数, 文字コード, エンコードされた文字列, デコードされた文字列 (漢字), 区分 {15文字一致/新規/新規(二重)}, 備考。
- ⑦ Replace: サイト内置換えページ一覧
置換え前ページ, (→), 置換え後ページ。
- ⑧ Classification: サイト区分一覧
区分ID, 種別, サイトドメイン, (未使用), 注意事項1, 備考1。
- ⑨ TEMP: Classificationの説明

(イ) アクセスログファイルの準備

利用するアクセスログはウェブサーバソフトウェア ApacheのCombined Log Formatで出力されたものを前提としている。したがって、この書式のアクセスログなら、農研機構ウェブサーバやバーチャルラボシステムウェブサーバに限らず、特別な変更もなく以下に説明する作業を実行してアクセス解析することができる。

アクセス解析ツールはWindows上で動作するExcel macro (VBA) で作成されているので、アクセスログファイルもWindows用に変換する必要がある。すなわち、文字コードをEUCからShift-JISへ、改行コードをLFからCR+LFへ変更する。また、Apacheが出力したアクセスログの範囲が厳密に月初めから月末ではなく、数行ずれていることが多いので、それも修正して入力用アクセスログファイルを作成する。ファイル名に関する特別な決まりはないが、食品害虫サイトでは慣例でaccess_combined_log_yyyyymm_naro.log (yyyyymmの部分は201406等) としている。

(ウ) アクセス解析ツールmacro_FoodInsect.xlsmの実行次の手順に従って解析を実行する。なお、この手順はmacro_FoodInsect.xlsmのシート「解析手順詳細」に記載されている。すべての操作終了後は、数値結果を記載したシートだけでなく、グラフ中心のまとめのシートも作成される(図4)。

- ① 解析月を指定する。(解析したい月が2013年9月の場合、2013/09/01と入力。)
- ② 解析結果ファイル名指定 (Analysis_FoodInsect_yyyymm.xlsx) して「保存(S)」を押す。
- ③ アクセスログファイルを指定 (【対象R抽出】の出力F: access_combined_log_yyyymm_naro.log) して「開く(O)」を押す。
- ④ "クローラ?"のコメントに答える。
ロボットらしき物には「はい」、そうでなさそうな物には「キャンセル」。
USER_AGENT: 情報にゴミが入っている時や何も

ない (none) の時にも「はい」を選択する。※"はい"にすると解析対象外レコードとなる。

- ⑤ 検索サイト判別のための文字列。
表示されたアドレスが検索サイトなら「はい」、そうでないなら「キャンセル」。
※検索サイトURLとキーワードの識別コードを設定する。
- ⑥ "日本語変換"のコメントに答える。
変換正常なものが出たら「はい(Y)」で、別変換が必要な時「いいえ(N)」で、変換しないとき「キャンセル」。

(エ) 管理ファイルList.xlsmの保守

シートRobots, SearchSites, Queriesに新たにデータが追加されているので、それらをチェックし、問題があれば修正して保存する。

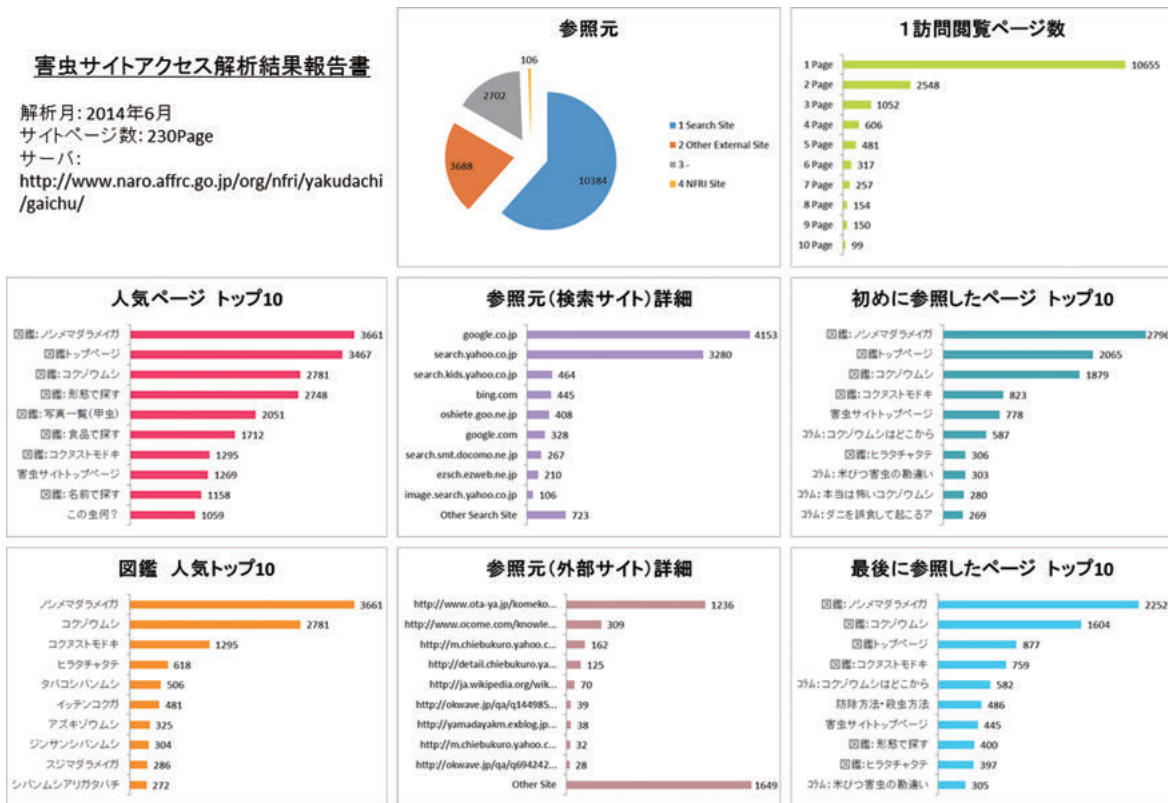


図 4. アクセス解析結果のまとめ (食品害虫サイト)

c ロボット (クローラ): 検索エンジンロボット. 検索サイトがウェブページの情報を収集するために動作させているプログラム. ウェブサーバの視点からは, 人間がブラウザを利用してウェブページを閲覧するのも, ロボットがウェブページにアクセスするのも違いはない.

アクセス解析ツールの放射線影響サイトへの適用

実際に他のウェブサイトにもアクセス解析ツールが使えることを示すために、以下に本アクセス解析ツールを用いて、放射線影響サイトの解析をおこなった例を示す。

1. アクセス解析手順

アクセス解析は次の手順で実施した。

(ア) Lists.xlsm: Serverの設定

放射線影響サイトは農研機構ウェブサーバとバーチャルラボシステムウェブサーバに分かれて配置されているので、シートServerには表1のようにその二つの情報を入力した。

(イ) Lists.xlsm: Pagesの設定

2014年5月時点の放射線影響サイトは、PDFファイルを含めて、17ページで構成されていた。表2のように、これをすべてシートPagesに登録した。

(ウ) Lists.xlsm: Replacesの設定

放射線影響サイトの改訂により、同じ内容のページが別の名称になったり、配置が変わったりしたものが4つあった。表3のように、これをシートReplaceに登録した。

(エ) アクセスログファイルの準備

(ア)で述べたように、放射線影響サイトは農研機構ウェブサーバとバーチャルラボシステムウェブサーバに分かれて存在するため、それぞれからアクセスログファイルを入手し、アクセス解析用入力ファイルとして、yyyyymm_naro.logとyyyyymm_rc0311.logを作成した。

(オ) アクセス解析ツールmacro_FoodInsect.xlsmの実行

この過程は、食品害虫サイトであろうが他のウェブサイトであろうが、大きな違いはない。2(ウ)のうち、入力ファイル名、出力ファイル名を適切なものに変更するだけでよい。放射線影響サイトの解析結果のまとめは図5のようになった。

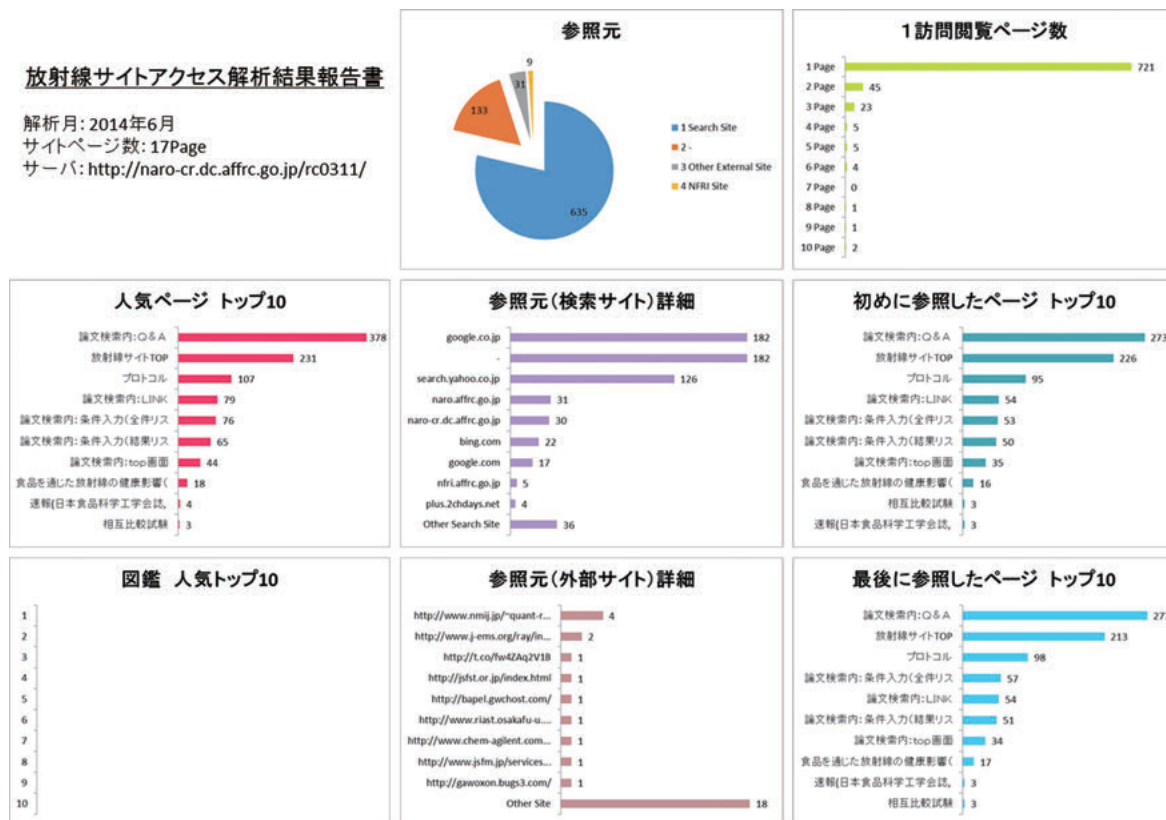


図5. アクセス解析結果のまとめ(放射線影響サイト)

表1. 放射線影響サイト用の設定 (シートServer)

サーバ	メインの階層	開始日	終了日	有効	サイト名称	備考
農研機構	http://www.naro.affrc.go.jp /org/nfri/topics/	2011/3/11		Yes	放射線サイト (メイン)	
バーチャルラボ	http://www.naro.affrc.go.jp /rc0311/	2011/3/11		Yes	放射線サイト (論文検索)	

表2. 放射線影響サイト用の設定 (シートPages)

ファイル名	ページ名称	開始日	終了日	有効	サイトID	備考
/R_C.html	放射線サイト TOP	2011/3/11		Yes	/R_C.html	
/protocol_kakunin25.pdf	プロトコル	2011/3/11		Yes	/protocol_kakunin25.pdf	
/kakunin_FAQ.pdf	よくある問い合わせ	2011/3/11		Yes	/kakunin_FAQ.pdf	
/kakunin_131023.pdf	相互比較試験	2011/3/11		Yes	/kakunin_131023.pdf	
/GE-2011-06-20.pdf	ゲルマニウム半導体検出装置	2011/3/11		Yes	/pdf/GE-2011-06-20.pdf	
/20110921-NaI.pdf	速報 (日本食品科学工学会誌, 58(9), 464-469(2011)) として掲載	2011/3/11		Yes	/pdf/20110921-NaI.pdf	
/CRinfo-end.pdf	放射性物質の食品影響と今後の対応	2011/3/11		Yes	/pdf/CRinfo-end.pdf	
/sympo1.pdf	放射性物質の基礎を学ぶ (小林泰彦先生)	2011/3/11		Yes	/pdf/sympo1.pdf	
/sympo2.pdf	食品を通じた放射線の健康影響 (滝澤行雄先生)	2011/3/11		Yes	/pdf/sympo2.pdf	
/Questionnaire.pdf	緊急シンポジウム アンケート集計結果	2011/3/11		Yes	/pdf/Questionnaire.pdf	
/ronbun_index.html	論文検索内: top 画面	2011/3/11		Yes	/rc0311/ronbun_index.html	
/link.html	論文検索内: LINK	2011/3/11		Yes	/rc0311/link.html	
/sympsiun.html	論文検索内: 緊急シンポジウム	2011/3/11		Yes	/rc0311/sympsiun.html	
/faq_01.jsp	論文検索内: Q & A	2011/3/11		Yes	/rc0311/faq_01.jsp	
/ronbun_01.jsp	論文検索内: 条件入力 (全件リスト付き)	2011/3/11		Yes	/rc0311/ronbun_01.jsp	
/ronbun_02.jsp	論文検索内: 条件入力 (結果リスト付き)	2011/3/11		Yes	/rc0311/ronbun_02.jsp	
/ronbun_03.jsp	論文検索内: 用語対訳リスト	2011/3/11		Yes	/rc0311/ronbun_03.jsp	

表3. 放射線影響サイト用の設定 (シートReplace)

置換え前	置換え後
rc0311/index.html	rc0311/ronbun_index.html
rc0311/symposium_files/sympo1.pdf	pdf/sympo1.pdf
rc0311/symposium_files/sympo2.pdf	pdf/sympo2.pdf
rc0311/symposium_files/Questionnaire.pdf	pdf/Questionnaire.pdf

(カ) 管理ファイルLists.xlsmの保守

この過程も食品害虫サイトの場合と同じである。

2. 放射線影響サイトと食品害虫サイトの比較

放射線影響サイトと食品害虫サイトのアクセス解析結果の比較を以下に記す(図4と図5)。

2014年6月時点で放射線影響サイトは17ページ、食品害虫サイトは230ページから構成されていた。

人気ページトップ10のグラフから、人気ページの1カ月の閲覧数を知ることができる。食品害虫サイトの

ページ閲覧数は放射線影響サイトのおよそ10倍であることから、放射線影響サイトの訪問者は専門家の割合が多く、食品害虫サイトは専門家以外の人も含まれていると思われた。

参照元の円グラフから、放射線影響サイトでは80%ほどが検索サイトからの訪問者であるのに対して、食品害虫サイトは他の外部サイトや「-」の割合が多くなっていった。「-」には、①訪問者のブラウザが参照元情報を返さない場合、②お気に入り(ブックマーク)から直接そのページを閲覧する場合が含まれている。

検索サイトを介さずに食品害虫サイトを訪問する割合が多いことから、すでに食品害虫サイトを認知しているユーザの割合が、放射線影響サイトよりも多いと言える。参照元（外部サイト）詳細のグラフからわかるように、食品害虫サイトのほうが外部サイト経由の訪問者が桁違いに多かったことも、これを裏付けている。

1訪問閲覧ページ数は、訪問者が1回の訪問で何ページ閲覧したかというデータである。放射線影響サイトでは1回の訪問で1ページ、つまり他のページを閲覧せずにそのまま出て行ってしまう場合がほとんどである。それに対して、食品害虫サイトでは複数ページを閲覧する割合が多くなっていた。この違いの原因は、サイトに含まれるページ数の違いもあるが、食品害虫サイトでは他のページに訪問者を誘導する工夫を施したページ作りをしている⁵⁾こともあるだろう。

始めに参照したページトップ10と最後に参照したページトップ10のデータから、食品害虫サイトで特徴的な傾向が見られた。防除方法・殺虫方法というページはほとんど1訪問の最初の閲覧ページ（入口）にはならないが、最後の閲覧ページ（出口）としては多かった。家庭で発生した害虫を調べるために図鑑ページにたどり着いた訪問者が、最後に防除方法を調べて満足したと考えられる。これは食品害虫サイトのページ作りの意図と合致している。一方、放射線影響サイトではこのようなことは見られなかった。

以上、開発したアクセス解析ツールが出力する結果のまとめを分析することで知ることができる訪問者のアクセス行動を簡単に紹介した。アクセス解析をしなければこのようなこともわからないため、ウェブサイトの評価も主観的になってしまうだろう。現在では食品害虫サイトの運営にアクセス解析は不可欠になっている。

本アクセス解析ツールの利用について

1. 本アクセス解析ツールを利用したい方へ

市販されているアクセス解析ツールは、本アクセス解析ツールよりも高速・高機能である。日常の解析項目が確定しており、見栄えのよい報告書を素早く作成したいという用途には、市販ツールの利用をおすすめする。

一方、ウェブサーバApacheからCombined Log Formatで出力されたアクセスログファイル、Windows PC、巨大ファイルを扱えるエディタ、Excelを用意すれば、ここで紹介したアクセス解析ツールを利用することが

できる。本報告のほかに、次節で述べる文書も用意しており、アクセス解析ツールそのものをさらにカスタマイズすることも可能である。食品害虫サイトの運営に利用してきた経験から、これからアクセス解析を導入するかどうか検討してみたい人や、小さなウェブサイトの運営に利用したい人にとって、本アクセス解析ツールは適していると考えている。本アクセス解析ツールを利用したい方は、食品総合研究所にお問い合わせいただきたい。

2. 関連文書

本アクセス解析ツールを他のウェブサイトに適用するときの参考にするために、下記の文書を整備している。

(ア) 解析ツール概要書.docx

解析内容など、概略を説明している。

(イ) 解析ツール設計書（全体仕様）.docx

アクセス解析ツールの仕様について全体像を示している。環境、処理の流れ、サーバアクセスログ形式、管理ファイルLists.xlsxmのシートの説明、解析結果ファイルAnalysis_FoodInsect_yyymm.xlsxのシートの説明に関して記述している。

(ウ) 解析ツール詳細設計書（Procedure）.docx

アクセス解析ツールの下記のプロシージャを解説している。

- ① Mod_main(): 全体の制御
- ② Sub1_Input(): 関係行だけを抽出
- ③ Sub2_Sort(): 並べ替え
- ④ Sub3_Path0(): 訪問経路解析
- ⑤ Sub4_Path1(): 訪問経路修正
- ⑥ Sub5_Connection(): ページ間結合解析
- ⑦ Sub6_Ranking(): ランキング
- ⑧ Sub8_Report(): 報告書作成

(エ) 解析ツール詳細設計書（Function）.docx

アクセス解析ツールで使われる下記のオリジナル関数を解説している。

- ① isCheckExtract(): 出力対象チェック
- ② strRevisedHost(): Host編集
- ③ strRevisedReferer(): Referer編集処理
- ④ strRevisedRequest(): Request編集処理
- ⑤ iPageID(): Connectionシート編集ライン算出処理
- ⑥ isCheckSearchSite(): サーチサイトチェック処理

- ⑦ strKeyword(): キーワード編集処理
- ⑧ strKeyword0(引数1): キーワード編集処理

(オ) 添付資料ファイル設計書 (Lists).xlsm

アクセス解析ツールで使われる変更の可能性があるパラメータをまとめたファイルLists.xlsmを説明している。内容は2.(1)で記したとおりである。

要約

著者らは4年以上前に食品害虫サイト用にアクセス解析プログラムを独自に開発し、これまで改良しながら使い続けてきている。今回、食品害虫サイト専用に記述された部分をプログラム本体から外部設定シートに分離することにより、簡単な作業で食品害虫サイト以外にも適用できるアクセス解析ツールを作成した。このアクセス解析ツールは、サイト内の各ページの閲覧数だけでなく、ページ間の移動に関する情報を解析できることに特徴がある。有用性を示すために、このアクセス解析ツールを放射線影響サイトに適用した。その解析結果を食品害虫サイトと比較したところ、食品害虫サイトのほうが1訪問中に複数ページを閲覧する割合が多いことがわかった。このアクセス解析ツールの利用希望者のために、設計資料等の文書も準備できている。

農研機構ウェブサーバのアクセスログの入手において新田宜史氏と大川清弘氏にお世話になった。また、ウェブサイトの開設と運営について石井馨氏にアドバイスを受けた。

文献

- 1) 濱松潮香, 放射性物質の食品への影響研究について, 食糧, **52**, 5-26 (2014).
 - 2) 食品総合研究所, 「食糧, 第50号」(食品総合研究所), (2011).
 - 3) 曲山幸生, 七里与子, 宮ノ下明大, 今村太郎, 古井聡, 和田有史, 増田知尋, 食品害虫サイトの長期間アクセス解析, 食品総合研究所研究報告, **77**, 51-61 (2013).
 - 4) 衣袋宏美, 「PROFESSIONAL アクセス解析」(技術評論社, 東京), (2011).
 - 5) 曲山幸生, 七里与子, 宮ノ下明大, 今村太郎, 食品害虫サイトの開設とそのアクセス解析, 家屋害虫, **31**, 93-99 (2009).
 - 6) 曲山幸生, 七里与子, 宮ノ下明大, 今村太郎, アクセス解析から推定した食品害虫の注目度と浸透度, 農業情報研究, **19**, 1-9 (2010).
 - 7) 曲山幸生, 七里与子, 宮ノ下明大, 今村太郎, 和田有史, 増田知尋, 木村敦, ウェブアンケートによる食品害虫サイト利用状況調査, 食品総合研究所研究報告, **75**, 55-61 (2011).
 - 8) 曲山幸生, 七里与子, 宮ノ下明大, 今村太郎, 古井聡, 和田有史, 増田知尋, 食品害虫サイトの大幅改訂による訪問者のアクセス行動の変化, 食品総合研究所研究報告, **76**, 59-66 (2012).
- (引用URL)
- i) <http://www.naro.affrc.go.jp/org/nfri/yakudachi/gaichu/index.html> (2014. 10. 23)
 - ii) http://www.naro.affrc.go.jp/org/nfri/topics/R_C.html (2014. 10. 23)
 - iii) <http://httpd.apache.org/> (2014. 10. 23)
 - iv) <http://www.gzip.org/> (2014. 10. 23)
 - v) <http://sevensip.sourceforge.jp/> (2014. 10. 23)
 - vi) <https://jp.emeditor.com/> (2014. 10. 23)
 - vii) http://www.naro.affrc.go.jp/org/nfri/yakudachi/gaichu/ranking_00.html (2014. 10. 23)
- 1) 濱松潮香, 放射性物質の食品への影響研究について, 食糧, **52**, 5-26 (2014).