

東北地方太平洋沖地震前後における JAMSTEC 文書データに対するユーザアクセスの変化

Change of user access to JAMSTEC Document data before and after the 2011 off the Pacific coast of Tohoku Earthquake

福田 和代^{1*}, 船越 留里¹, 宮城 伸², 齋藤 秀亮¹, 仲村 亮², 園田 朗¹

FUKUDA, Kazuyo^{1*}, Ruri Funakoshi¹, Shin Miyagi², Hideaki Saito¹, Makoto Nakamura², Akira Sonoda¹

¹ 海洋研究開発機構地球情報研究センター, ² マリン・ワーク・ジャパン

¹DrC/JAMSTEC, ²Marine Works Japan Ltd.

海洋研究開発機構では研究開発成果を広く一般に紹介するため、国際海洋環境情報センターにおいて PDF ファイルとしてアーカイブされた JAMSTEC のさまざまな刊行物のバックナンバーを、2002 年 4 月からウェブサイト「JAMSTEC 文書データベース」を通じて公開してきた¹⁾。このウェブサイトは 2011 年 11 月末に新しいウェブサイト「JAMSTEC 文書カタログ²⁾」に移行した。各刊行物の文書は JAMSTEC の研究対象である、深海生物、気候変動、固体地球などの幅広い分野をカバーしており、2012 年 2 月現在、7 種類の広報誌 390 件と 4 種類の学術誌 1975 件が公開されている。近年、地球情報研究センターでは、ユーザニーズ調査の一環として、同センターが運用する 10 以上のウェブサイトについてアクセス解析を行っており、上述した 2 つのサイトも解析対象である。本発表では、JAMSTEC 文書データに対する 2011 年 3 月 11 日東北地方太平洋沖地震前後のユーザ行動に着目し、アクセス数の変化や検索エンジンにおける検索語の変化について調査した結果を報告する。

解析に用いたデータは JAMSTEC ドメインのアクセスを除外した 2010 年 4 月から 2011 年 11 月までのアクセスログである。訪問数や検索フレーズなどの有用な情報を抽出するため、ウェブ解析レポートツール AWStats³⁾ を用いてログファイルを処理した。さらに、検索エンジン用いられた検索フレーズの特徴を把握するため、テキストマイニングの手法を適用し、入力された語の頻度や共起について、日本語形態素解析器 MeCab 0.98⁴⁾、統計ソフトウェア R 2.13.1⁵⁾ およびパッケージソフトウェア RMeCab 0.94⁶⁾ を用いて解析した。

地震の影響を受け、日次の訪問数は 2011 年 3 月 12 日に急増していた。地震後の 6 カ月平均の訪問数および訪問者数は地震前と比較して 1.4 倍に増加した。参照元別の 6 カ月平均の変化を調べると、地震後の直接、外部ページ、検索エンジンからのアクセス数は地震前のそれぞれ、約 1.2 倍、約 1.6 倍、約 1.3 倍に増加しており、新規ユーザ、リピータともアクセスが増加していたことが推察される。

検索フレーズについて調べると、検索語「地震」の頻度は全期間を通して各月で 5 位以内にランクしており、普段からユーザの関心事であることがうかがえる。2011 年 3 月から 4 月にかけて、「地震」の出現頻度は地震前の約 7 倍となり、5 月以降も地震前の 3 倍を維持していた。地震と共起する語を抽出すると、2011 年 3 月から 4 月にかけて「メカニズム」、「巨大」、「発生」の語の急増がみられた。また、地震に関連するいくつかの語の頻度の急増は、新聞やテレビ報道の影響を受けていることが推測された。文書データに対するアクセスパターンを調べると、高校生以上を対象とした海洋地球科学情報誌「Blue Earth」へのアクセスが地震後に増加していた。全期間における「Blue Earth」のアクセス上位 20 位の記事のうち、6 つの記事はタイトルに「地震」が含まれており、それ以外の 5 つの記事の本文中には地震に関連する語が含まれていた。これらの結果は、文書データをオンライン公開することがアウトリーチ活動にとって極めて重要であることを示唆する。

参考文献・参考 URL

1) 園田朗、仲村亮、宮城博、設楽文朗、2005、国際海洋環境情報センター (GODAC) における情報管理について、情報管理、Vol.48、No.4、214-219.

2) http://www.godac.jamstec.go.jp/catalog/doc_catalog/index_en.html

3) <http://awstats.sourceforge.net>

4) <http://mecab.sourceforge.net>

5) <http://www.R-project.org>

6) <http://rmecab.jp/wiki/index.php?RMeCab>

キーワード: ウェブ解析, テキストマイニング, 東北地方太平洋沖地震, アウトリーチ

Keywords: Web analytics, Text mining, Tohoku Earthquake, Outreach