

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

FUKUDA, Kazuyo<sup>1\*</sup>, Ruri Funakoshi<sup>1</sup>, Shin Miyagi<sup>2</sup>, Hideaki Saito<sup>1</sup>, Makoto Nakamura<sup>2</sup>, Akira Sonoda<sup>1</sup>

<sup>1</sup>DrC/JAMSTEC, <sup>2</sup>Marine Works Japan Ltd.

In order to disseminate research and development achievements of the Japan Agency for Marine-Earth Science and Technology (JAMSTEC), JAMSTEC has provided back issues of various publications in PDF archived at the Global Oceanographic Data Center<sup>1)</sup> via the website "JAMSTEC Document Database" since April 2002. This site has been succeeded by the new website "JAMSTEC Document Catalog<sup>2)</sup>" since November 2011. The documents in each publication cover great deal of fields such as deep-sea living, climate change, and solid earth. As of February 2012, seven different kinds of 390 public relations and four different kinds of 1975 research papers were provided in this website. Recently, we at the Data Research Center for Marine-Earth Sciences (DrC) have analyzed traffics for over 10 websites operated by DrC as part of user needs assessment. The two websites described above have been also analytical objects. In this presentation, we will report the change of user access to JAMSTEC document data focusing on user behavior before and after the 2011 off the Pacific coast of Tohoku Earthquake.

For the traffic analysis, we used the server logs from the "JAMSTEC document database", except access from JAMSTEC domain, from April 2010 to November 2011. The logs were processed by using a web analytics reporting tool AWStats<sup>3)</sup> to extract traffic parameters such as number of visits and search phrases. Furthermore, text mining applied to search phrases used at search engines. Term frequency and collocation about search phrases were analyzed using by a Japanese morphological analyzer MeCab 0.98<sup>4)</sup>, a statistical software R 2.13.1<sup>5)</sup> and a package software RMeCab 0.94<sup>6)</sup>.

Number of daily visits has increased suddenly on March 12th, 2011, the day after the earthquake. The six-month average number of both visits and unique visitors after the earthquake was 1.4 times larger than before the earthquake. After the earthquake, the six-month average number of entries to the site via direct access or bookmark, external pages and search engines was 1.2, 1.6 and 1.3 times larger than before the earthquake, respectively. This indicates that the number of both new and repeat users has increased after the earthquake. According to an analysis of search phrases, the frequency of term "jishin", which means "earthquake" in Japanese, has been ranked in the top 5 during the whole period. It shows that earthquake has been one of the user's major concerns in daily life. The term frequency of "jishin" from March to April 2011 was seven times higher than before the earthquake, and the frequency after May remained three times higher than before the earthquake. Frequently collocated terms with "jishin" in each phrase from March to April 2011 were "mekanizumu", "kyodai" and "hassei", which mean "mechanism", "mega" and "occurrence" in Japanese, respectively. The rapid rise of frequency of terms related earthquake seems to influence by newspaper and television coverage. The number of access to the "Blue Earth", which is a marine-earth science information magazine aimed to above high school student, has markedly increased after the earthquake. The titles of six documents include the term "jishin" and the bodies of the other five documents include terms related to earthquake among top twenty most accessed documents within the "Blue Earth". These results suggest that providing documents online is essential for outreach activity.

### References

- 1) Sonoda, A., M. Nakamura, H. Miyagi and F. Shidara, 2005, In the case of information processing and management at Global Oceanographic Data Center (GODAC)/JAMSTEC, *Journal of Information Processing and Management*, 48(4), 214-219.
- 2) [http://www.godac.jamstec.go.jp/catalog/doc\\_catalog/index.en.html](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>

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