Journal of Information Processing Vol.22 No.4 566–573 (Oct. 2014)

[DOI: 10.2197/ipsjjip.22.566]

Invited Paper

The 3.11 Disaster and Data

Haruhiko Okumura[1,a)]

Received: September 13, 2013, Accepted: May 26, 2014

Abstract: "3.11"—the worst disaster in postwar Japanese history, consisting of the Great East Japan Earthquake (March 11, 2011), the subsequent tsunami and the nuclear accident at the Fukushima Daiichi power plant—taught us many valuable lessons. This paper reviews the disaster from a computer scientist's perspective, paying special attention to the problem of presenting data to the public, and discusses what we could do and can still do.

Keywords: Japan, Fukushima, earthquake, disaster, data

1. Introduction and Overview

The Great East Japan Earthquake of magnitude $M_w = 9.0$, the worst in Japanese history, occurred at 14:46 on March 11, 2011 [*1]. Thanks to Japan's advanced early warning system, alarms rang seconds before the main shock. *Shinkansen* (Bullet Trains) slowed down immediately without trouble. The first tsunami warning was issued on 14:49. It took about half an hour to one hour for the tsunami to arrive at the coast, but the initial forecast of their heights—point estimates rather than interval estimates—turned out to be too small. Many people did not evacuate soon enough. Moreover, it became clear that the nuclear power plants were not designed to withstand a tsunami of an "unexpected" size. The Fukushima Daiichi power plant lost control, resulting in the worst nuclear accident since Chernobyl. We shall call the whole disaster "3.11".

After the earthquake, telephones were mostly unusable. People used Twitter and other digital media to communicate and find nearby buildings that could accommodate stranded commuters. Even in the worst-affected area, movies of the tsunami were taken with smartphones and uploaded to YouTube. However, the central and local governments and TEPCO, the company that runs the nuclear power plant, had trouble communicating and presenting data.

This paper reviews the disaster from a computer scientist's perspective, paying special attention to the problem of presenting data, and discusses what we could do and can still do.

Parts of the content of this paper were published in Japanese elsewhere [1], [2], [3], and presented in symposium talks at the Information Processing Society of Japan, the Physical Society of Japan, the Information Systems Society of Japan, and Google's *Project 311* workshop, among others.

The *2011* and *2012 White Papers on Information and Communications in Japan* [4], [5] are valuable sources of information. Official reports on the nuclear accident are available in English from the Government [6], the National Diet [7], and an independent organization [8]. Numerous books on 3.11 have been published in Japanese, by authors including the then Prime Minister, top officials and scientists. However, the contents were quite diverse and sometimes contradictory.

Some open-access journal issues and proceedings are devoted to 3.11, including *Information Processing* Vol.52, No.9 (2011), https://ipsj.ixsq.nii.ac.jp, and *Journal of Natural Language Processing* Vol.20, No.3 (2013), https://www.jstage.jst.go.jp/browse/jnlp/20/3/_contents.

2. Failure of Information Systems

After the earthquake, voice calls became overloaded and mostly unusable in a wide area including Tokyo [4]. However, public telephones were often usable, and citizens were advised to turn off anonymous-call rejection setting, so as not to reject calls from public telephones.

Internet-based systems, including IP phones and Skype, were usable. Cell sites that were not washed out by the tsunami functioned for several hours, until the emergency batteries died.

Emails on PCs were usable, but texting on cell phones did not work as expected, presumably because the "push" mechanism was in trouble. Manual "pull" operations often seemed to work.

Many disaster-related websites were too overloaded to be usable. Some websites set up mirror sites, but if the links to the mirror sites were on the original sites, it was difficult to discover the mirrors. Inoue et al. [9] started a URL shortening service that replaces the original URL with a "coralized" URL, i.e., one with `.nyud.net` appended to the domain name, which makes CoralCDN content distribution network automatically mirror the web content [10]. However, Inoue's system was not widely used. CoralCDN has its own drawbacks in that its DNS packets exceed 512-byte UDP limit.

NHK, the national broadcaster, remained a reliable source of information, but many people were unable to reach TV sets. At 15:03, a 14-year-old boy began to re-broadcast NHK TV via Ustream using his iPhone. He knew that his deed infringed the

1 Faculty of Education, Mie University, Tsu, Mie 514–8507, Japan
a) okumura@okumuralab.org

*1 Time is shown in 24-hour HH:MM format, Japan Standard Time (JST).

rights of NHK and he might be punished, but he was determined to help people in dire need of information. His re-broadcasting became known through Twitter. Even `@NHK_PR`, the popular NHK public-relations Twitter account, advertized the link to the re-broadcasting, saying that it would take the consequences. The boy and the person behind `@NHK_PR` were not punished. Instead, NHK officially began broadcasting its TV content via Ustream later that day. Other TV stations followed suit. Fukuda et al. [11] noted that these multimedia traffic had a non-negligible effect on the Internet, although the effect of the increased amount of text communications was insignificant.

Because Japan has experienced many earthquakes and tsunami, local governments had set up contingency plans for such disasters. For example, Kesennuma, a coastal city with a population of 73,000 (as of 2010), had planned to use Area Mail, a congestion-free broadcast-type texting service for mobile phones [12]. The city had also set up an emergency Twitter account, `@bosai_kesennuma`, just in case [13]. The city office had a dedicated computer terminal for sending Area Mail—the single point of failure in hindsight. After the earthquake, the city workers were unable to get the terminal to work, so they had to rely on the siren and Twitter to warn about the tsunami. Their first tweet was at 14:55, nine minutes after the main quake. The tsunami arrived at the city office some time after 15:30. They escaped to the rooftop, where they continued to tweet using a laptop. After the laptop battery died, they turned to a mobile phone to keep tweeting. Their last tweet on the day was at 22:37, after which, the cell site for the mobile phone ceased to work because the batteries died [14]. In Kesennuma, more than 1,400 citizens were lost or missing.

Some city and town offices were washed out by the tsunami. Paper documents and data that had not been backed up were lost.

People rushed to send donations, causing an abnormal termination of Mizuho Bank's batch job on the night of March 14, resulting in the megabank's large-scale trouble that lasted until March 24 [15].

Google set up its renowned Person Finder by 16:32 on March 11. Volunteers converted many photos of handwritten lists of evacuated people into text data for Person Finder. Google's Crisis Response website included many other services such as a map of passable roads based on location data from Honda cars [16], [17].

After the failure of the Fukushima Daiichi nuclear reactors, some twenty members of the then Nuclear Safety Commission (NSC) were summoned using priority texting, but the message did not get through to the members. A few members voluntarily walked to the NSC office, in the midst of complete transportation paralysis in the Tokyo area [18, p.32]. Were there no other means of communication? Moreover, it was revealed later that many of the official emergency meetings failed to keep records of what was discussed. A voice-recorder app on a smartphone or even photos of the whiteboard could have been used for keeping records.

3. Twitter, Chain Emails, and Hoaxes

TV experts, who assured anxious viewers that Japan's nuclear reactors were safe, lost their confidence when explosions occurred at the Fukushima Daiichi plant. On Twitter, few experts spoke up, but those who did won many followers. For example, Ryugo Hayano, `@hayano`, Professor of Physics at the University of Tokyo, had only 2,272 followers on March 10, but within ten days the number had jumped to over 150,000, as shown in **Fig. 1**. Even this author, `@h_okumura`, experienced a significant increase of followers.

To study how people used Twitter during the first week after 3.11, the author analyzed all of the 179,286,297 tweets in Japanese (including retweets) for the period of 7×24 hours starting at 09:00 on March 11. The data were kindly provided by Twitter Japan, Inc., through the coordination of Google Japan [19].

During the period, 3,691,599 different accounts, amounting to 2.9% of Japan's population, tweeted or retweeted. Among the tweets and retweets, 21.4% were retweets, 2.4% were tweets that were retweeted during the period, and the remaining 76.2% were "lone" tweets which were not retweeted [20].

These tweets give an idea of when and to what extent people were interested in specific subjects. For example, the word SPEEDI (see Section 7) appeared only 201 times during this week, of which 115 were retweets.

Twitter was also a vehicle for hoaxes and false rumors. An example of false advice was to drink *Isodine* (brand name for the popular gargle solution containing povidone-iodine) to prevent radioactive iodine intake; see **Fig. 2**.

Chain emails were also rampant, often including stereotypical phrases such as "According to a friend of mine (or brother-in-law, or whatever) who is working for the relevant company (or government office, or whatever)," and "Please forward this message to as many people as possible."

One of the most popular chain emails in the Tokyo bay area on March 11, after an explosion at an oil plant, was: "Due to a Cosmo Oil plant explosion, the rainfall contains toxic substances. Carry an unbrella or a rain jacket to prevent contamination." Cosmo Oil announced that the explosion did not involve

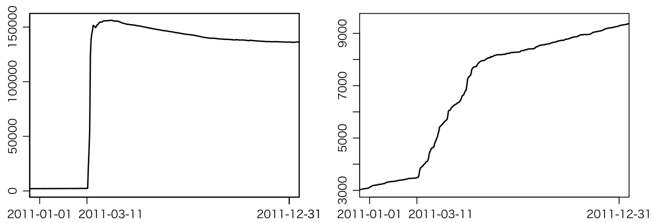


Fig. 1 Numbers of Twitter followers in 2011, for @hayano (left) and @h_okumura (right).

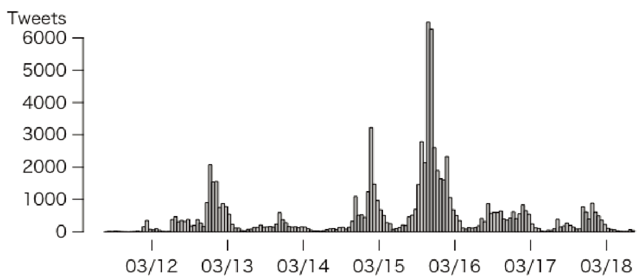


Fig. 2 Hourly occurrences of the word "うがい薬" (gargle solution) or "イソジン" (*Isodine*, a popular brand name) on Twitter.

toxic matters.

In the western part of Japan, people received many emails like this on March 12: “The eastern part of Japan will run out of electricity at 6 p.m. today. Electricity will be transmitted from the western part of Japan, so save electricity!” In fact, power lines in the eastern and western parts of Japan use different frequencies, and cannot be exchanged easily.

After the accident at the Fukushima Daiichi nuclear plant, a variety of hoaxes on radioactivity, such as the following, began to spread through Twitter and other social media:

- Radioactivity is coming to Tokyo. Drink much water, eat *kelp*, or drink diluted *Isodine* solution. [Kelp contains iodine. *Isodine* is the aforementioned gargle solution containing some iodine. Drinking such solutions is dangerous.]
- The Japan Medical Association recommended doctors not to enter within 50 km of Fukushima Daiichi. [The Medical Association announced that this was untrue.]
- Even in Tokyo, many people are suffering from nosebleeds due to radioactivity. [Absurd. Even decommission workers at the Fukushima Daiichi do not suffer from any radiation-related symptoms such as nosebleeds.]

On March 13, the Ministry of Internal Affairs and Communications (MIC) asked the public not to forward chain emails [21], and on April 6, asked ISPs to take appropriate measures with regard to hoaxes on the Internet [22].

Even today, however, hoaxes such as “the government and mass media are concealing the facts about radioactivity” are abundant. Such hoaxes have led some people to escape from safe places like Tokyo to western parts of Japan, or even overseas.

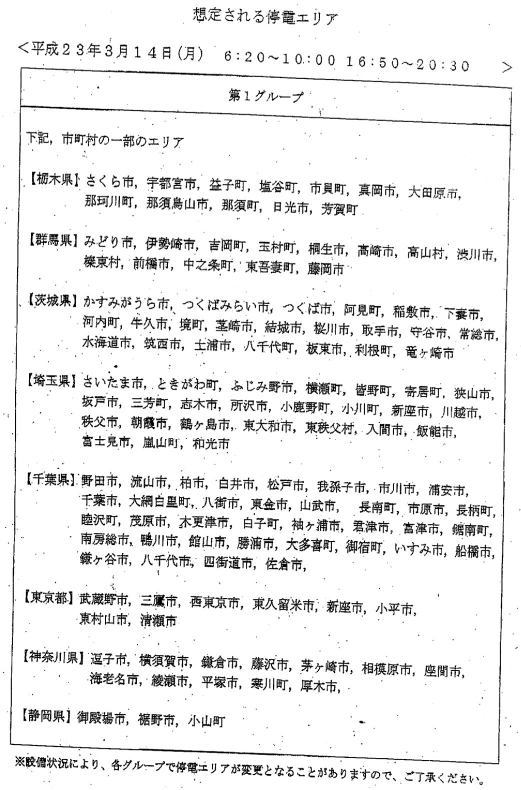
4. TEPCO’s Electricity Data

Because of the failures of its Fukushima Daiichi (No.1) and Daini (No.2) nuclear power plants due to the tsunami, the Tokyo Electric Power Company (TEPCO) carried out scheduled blackouts in a wide area including Tokyo, from March 14 onward. TEPCO’s first announcement of the blackouts was an 8-page PDF file apparently scanned from a fax output, as shown in **Fig. 3**. As people rushed to download the file, the TEPCO server became overloaded. People began to mirror it on other sites and converted it to readable/searchable formats. TEPCO’s announcements were improved after that, but their practice of overwriting earlier files when new announcements were made was confusing.

Image PDF files like that shown in Fig. 3 were widely used for presenting data after 3.11. Many were scanned from faxed outputs. The fact that faxes are still the preferred communication tool in Japan has led some researchers to develop a fax-based data transfer system [23].

TEPCO provided a graphic chart showing the current electricity demand on its website [24], but failed to give raw numbers. This was inconvenient for application developers. A volunteer developed an API that extracted approximate values from the image. Later, TEPCO began to provide raw numbers in CSV format. The data were used by various smartphone apps and portal sites such as Yahoo! Japan.

However, on July 1, TEPCO changed the format of the CSV file so that “yesterday’s demand” became “today’s forecast.”


想定される停電エリア

<平成23年3月14日(月) 6:20~10:00 16:50~20:30 >

第1グループ

下記，市町村の一部のエリア

【栃木県】さくら市，宇都宮市，益子町，塩谷町，市貝町，真岡市，大田原市，那珂川町，那須烏山市，那須町，日光市，芳賀町

【群馬県】みどり市，伊勢崎市，吉岡町，玉村町，桐生市，高崎市，高山村，渋川市，榛東村，前橋市，中之条町，東吾妻町，藤岡市

【茨城県】かすみがうら市，つくばみらい市，つくば市，阿見町，稲敷市，下妻市，河内町，牛久市，境町，茎崎市，結城市，桜川市，取手市，守谷市，常総市，水海道市，筑西市，土浦市，八千代町，板東市，利根町，竜ヶ崎市

【埼玉県】さいたま市，ときがわ町，ふじみ野市，横瀬町，皆野町，寄居町，狭山市，坂戸市，三芳町，志木市，所沢市，小鹿野町，小川町，新座市，川越市，秩父市，朝霞市，鶴ヶ島市，東大和市，東秩父村，入間市，飯能市，富士見市，嵐山町，和光市

【千葉県】野田市，流山市，柏市，白井市，松戸市，我孫子市，市川市，浦安市，千葉市，大網白里町，八街市，東金市，山武市， 長南町，市原市，長柄町，睦沢町，茂原市，木更津市，白子町，袖ヶ浦市，君津市，富津市，鋸南町，南房総市，鴨川市，館山市，勝浦市，大多喜町，御宿町，いすみ市，船橋市，鎌ヶ谷市，八千代市，四街道市，佐倉市，

【東京都】武蔵野市，三鷹市，西東京市，東久留米市，新座市，小平市，東村山市，清瀬市

【神奈川県】逗子市，横須賀市，鎌倉市，藤沢市，茅ヶ崎市，相模原市，座間市，海老名市，綾瀬市，平塚市，寒川町，厚木市，

【静岡県】御殿場市，裾野市，小山町

※設備状況により、各グループで停電エリアが変更となることがありますので、ご了承ください。

Fig. 3 Part of TEPCO’s plan of blackouts for March 14, a PDF file apparently scanned from fax output. At least two versions of the same filename `teiden20110313.pdf` were released.

The Ministry of Economy, Trade and Industry (METI) noticed TEPCO’s announcement on the night of June 30, and warned the developers about the change at 22:01 from the `@openmeti` Twitter account [25]. Yahoo! Japan fixed their API early the next day (by 2:34), but smartphone apps that directly accessed TEPCO’s CSV files displayed meaningless results. `@openmeti` tweeted that METI was disappointed by TEPCO’s unexpected change of format. This was one occasion when TEPCO and other information providers did not understand the significance of machine-readable data.

5. TEPCO’s Radiation Data

Before 3.11, TEPCO had a Web-based realtime monitoring system, which showed radiation levels every ten minutes for eight monitoring posts in the Fukushima Daiichi site. But the system was stopped because of the earthquake.

The reactor crisis progressed rapidly, but instead of releasing radiation data as soon as they were obtained, TEPCO took the trouble of tabulating them on neatly bordered PDF files, as in **Fig. 4**, and released them only once in a while. For example, the file shown in Fig. 4 has `/ModDate` (PDF modification date) 14:00 and timestamp 14:20 of March 15; i.e., it was released more than five hours after the record 11,930.0 μSv/h was observed at 09:00 near the main gate of the Daiichi site.

TEPCO apparently put in the unnecessary units by hand; on at least one measurement they mistakenly entered Sv/h in place of μSv/h. Other typos included commas “,” in place of decimal points “.”, confusion between 午前 (a.m.) and 午後 (p.m.), and a queer occurrence of “お” (the Japanese letter “O”) in place of “0”.

計測日	計測時間	計測場所	γ線	中性子線	風向	風速(m/s)
	午前9時00分	正門	11930．0μSv/h	0.01μSv/h未満	北北東	1.5
	午前9時15分	MP−4付近	58．0μSv/h	−	−	−
	午前9時20分	MP−4付近	50．0μSv/h	−	−	−
	午前9時35分	正門	7241．0μSv/h	0.01μSv/h未満	北東	5.3
	午前10時15分	正門	8837．0μSv/h	−	−	−
	午前11時40分	西門	253．8μSv/h	0.01μSv/h未満	−	−
	午前11時45分	西門	162．4μSv/h	−	−	−
	午後0時05分	西門	2431．0μSv/h	0.01μSv/h未満	南東	1.2
	午後0時15分	西門	2434．0μSv/h	0.01μSv/h未満	東	1.3
	午後0時25分	正門	1407．0μSv/h	0.01μSv/h未満	東南東	3.4
	午後0時35分	正門	1325．0μSv/h	0.01μSv/h未満	南東	1.3
	午後0時45分	正門	1267．0μSv/h	0.01μSv/h未満	南	1.4
	午後0時55分	正門	1216．0μSv/h	0.01μSv/h未満	南	1.8
	午後1時00分	正門	1191．0μSv/h	0.01μSv/h未満	南	1.3
	午後1時10分	正門	1148．0μSv/h	0.01μSv/h未満	南	1.3
	午後1時20分	正門	1100．0μSv/h	0.01μSv/h未満	南南東	1.4
	午後1時30分	正門	1068．0μSv/h	0.01μSv/h未満	南	1.0

Fig. 4 The last page of the PDF file http://www.tepco.co.jp/cc/press/betu11_j/images/110315d.pdf TEPCO released on March 15, 2011. The file was linked from http://www.tepco.co.jp/cc/press/11031509-j.html, but was later retracted and replaced by a new file.

Moreover, most of the numbers were in *zenkaku* (fullwidth) characters, presumably for aesthetic reasons. To create a time-series plot of the radiation level, as shown in **Fig. 8**, it was necessary to convert the PDF files into a machine-readable format.

TEPCO's files were typical examples of "bad data" [26]: data intended for human consumption, not machine consumption.

For data to be machine-friendly, cells should not be merged, and numbers should be in ASCII numerals without units. Date-time values should be in a format that data-analysis software such as Excel, Google Spreadsheets and R can interpret, as in the example below (in CSV format [27]):

```
DateTime,Place,Gamma,Neutron,WindDir,WindSpeed
2011/03/15 06:00,Front Gate,73.2,,N,0.8
2011/03/15 08:20,Front Gate,807.7,,NE,1.5
2011/03/15 08:31,Front Gate,8217.0,,NE,1.5
2011/03/15 08:40,Front Gate,1726.0,,N,1.6
2011/03/15 08:50,Front Gate,2208.0,,N,1.8
2011/03/15 09:00,Front Gate,11930.0,,NNE,1.5
2011/03/15 09:15,Near MP-4,58.0,,,,
```

The gamma-ray and wind direction/speed data were very important at that time. They should have been published as soon as they were obtained. They could have shared the data using an online tool such as Google Spreadsheets (Google Drive). That way, they did not even have to save the file and upload it to the server.

During the months after 3.11, the author converted much of such "bad data" into machine-readable CSV files at http://oku.edu.mie-u.ac.jp/~okumura/stat/data/, and posted plots similar to those shown in Fig. 8 on Twitter. At that time, many scientists and IT specialists felt that it was their responsibility to convert obfuscated PDF files into machine-readable data, visualize them, and give scientific explanations. A summary website [28] links to many such activities.

On March 30, the Ministry of Economy, Trade and Industry (METI) requested [29] the *Keidanren* (Japan Business Federation) member organizations (including TEPCO) to present the disaster data in HTML, CSV, or other machine-readable formats, rather than PDF; or provide APIs for machine processing. But TEPCO's way of releasing data did not change.

Moreover, TEPCO "secured" (copy-protected) their PDF files, as shown in **Fig. 5**.

Citing the author's blog article [30], an online media company

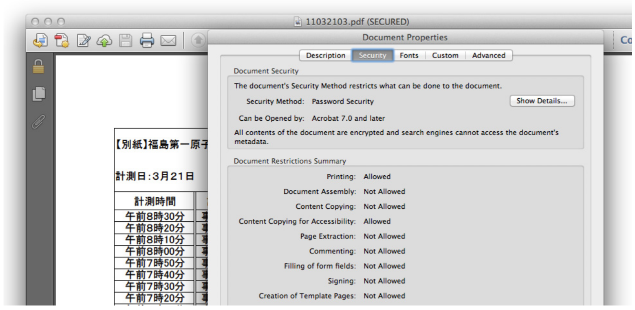


Fig. 5 TEPCO "secured" their PDF files so that document assembly, content copying, and page extraction were not allowed.

asked TEPCO why they secured their PDF files [31]. TEPCO answered on April 4: "As for the publication of the monitoring data, we have been securing the PDF files to prevent diverting or copying all or part of the information provided by TEPCO verbatim or with modification," which did not make sense.

METI apparently contacted TEPCO directly. On April 12, TEPCO finally began to release CSV files in addition to PDF files. A METI twitter account `@openmeti` revealed on the next day, "For CSV of TEPCO's radiation data, the arrangement with TEPCO was done by [METI's] Information Economy Division" (https://twitter.com/openmeti/status/58168947572355073).

On October 17, 2012, TEPCO resumed the Web-based real-time monitoring system, after 19 months of discontinuation, at http://www.tepco.co.jp/nu/fukushima-np/f1/. It shows radiation readings every 10 minutes, although with a curious delay of about 20 minutes.

6. MHLW and Food Contamination Data

After 3.11, the Japanese Ministry of Health, Labour and Welfare (MHLW) has been publishing radioactivity measurements of food every weekday (until April 2014, when the schedule was changed to once per week). Measurements are mostly done by local governments, which email the results to MHLW in Excel form. MHLW copies them into one or more Excel files, converts them to PDF, and publishes them on the MHLW website.

In MHLW's old format, the cells for ^{134}Cs and ^{137}Cs had to be merged when only the sum of these two cesium isotopes were known. With Excel, it is easy to merge cells and format tables visually, but it is impossible with CSV and RDBMS tables. On March 30, 2012, MHLW revised the format. The new format has three separate columns for ^{134}Cs, ^{137}Cs and the total. If only the total is known, the cells for ^{134}Cs and ^{137}Cs are filled with "–" (for "not available"). This small improvement made computer processing much easier.

MHLW's files are not named systematically. The new results are linked from the "What's New" section of the MHLW home page. This makes automatic discovery of new data difficult. They should at least name CSV files systematically. A better approach is to create a database system into which the local governments can enter measurements directly.

Sugiyama and Uda [32] set out to make such a database system. They had to convert MHLW's PDF files to machine-readable data "by hand," which were double-checked by the voluntary effort of the Organization of Food-marketing Structure Improvement (OFSI). The search interface, http://yasaikensa.cloudapp.net, built with the app engine provided by Microsoft, was released on April 11, 2011. The website came to be known as the *yasaikensa* (vegetable test) site after its URL, although it was not limited to vegetables.

For one year, the people behind the *yasaikensa* site continued updating the data every time MHLW published new PDF files, but they finally decided to discontinue updating the data at the end of March, 2012, and close the entire website at the end of April. They released their code and data on GitHub (https://github.com/udawtr/yasaikensa), and asked volunteers to continue with the effort.

But it was not easy to deploy their code on a Linux server. On April 1, 2012, the author of this paper created a simple search interface from scratch [33]. To convert MHLW's PDF files to CSV, he used the *pdftotext* command-line utility, a simple Ruby script, and some human intervention. The daily CSV files and a merged one have been maintained at http://oku.edu.mie-u.ac.jp/~okumura/stat/data/mhlw/. Volunteers developed several websites and an iPhone app that used the data.

In May 2012, MHLW started to release monthly Excel files, but daily files were still in PDF. On November 1, 2013, a new website, http://www.radioactivity-db.info, that succeeded the *yasaikensa* site, was opened by the National Institute of Public Health under contract with MHLW. The advent of this site finally obviated the one-man effort of this author. On March 11, 2014, the author discontinued updating the data.

MHLW does not report measurement errors. The measured value (in Bq/kg) is reported only when it exceeds the detection limit (3σ limit); otherwise, it is written as < (less-than) followed by the detection limit, e.g., "< 25". In the past, insignificant values were denoted as "ND" (not detected). People often could not understand the difference between "< 25" and "25", or so pretended in order to exaggerate the results [*2].

7. Measurement and Simulation

After 3.11, people rushed to buy Geiger counters and other radiation detectors, some of which were of poor quality. These tools were generally calibrated, if at all, against the gamma ray from some specific radioisotope, but many were also sensitive to beta rays that have very short ranges, so amateur measurements on soil surfaces often gave values that were far too large. Some even thought that statistical (Poisson-distributed) fluctuations of radiation counts were caused by blowing winds that carry radioactive substances. This misconception was reinforced by newspapers that reported "air" dose rates. In fact, virtually all radiation comes from cesium-contaminated soil surface.

Fluctuating values must be averaged, and this is especially true with radiation whose effect on humans is assumed to be proportional to its time integral, but people often reported the maximum value, which can be made arbitrarily large by decreasing measurement intervals. For example, **Fig. 6** shows the radiation at Shinjuku, where Tokyo's official monitoring post resided. The first plot shows hourly mean values with error bars showing fluctuations. The gray belt represents the range of a natural radiation level, 0.028–0.079 μGy/h. The mean value for 10:00–10:59 of March 15 was 0.496 μGy/h, with a fluctuation of 0.160–0.809. The mean value for the whole day was 0.111 μGy/h. The second plot is by NHK, the Japanese national broadcaster. Unfortunately, they used the maximum value of the fluctuation as the representative value of each day.

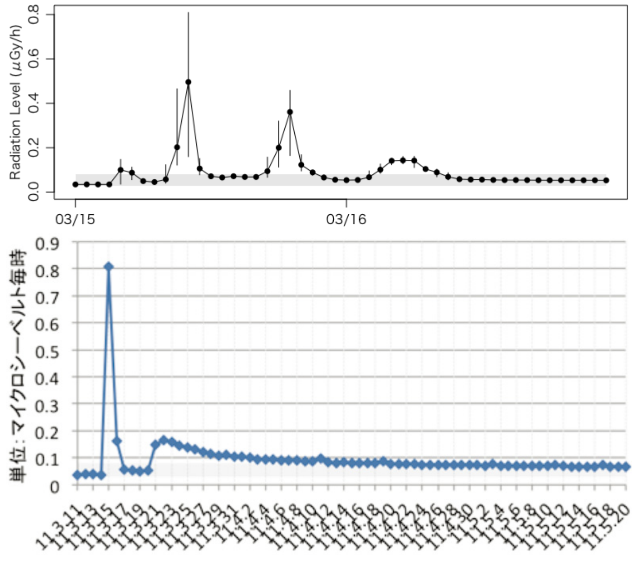


Fig. 6 (Top) Radiation in Tokyo after 3.11, based on the data http://monitoring.tokyo-eiken.go.jp/mp_shinjuku_air_week_list.html, (Bottom) the same data plotted by NHK.

The Japan Meteorological Agency kept silent about the radioactive plume, but institutions overseas began publishing various simulation results. Examples were *Spiegel Online* [34], Australian Central Institution for Meteorology and Geodynamics http://www.zamg.ac.at, and German Meteorological Service http://www.dwd.de/. Most of these simulations merely showed wind trajectories, for at that time there was no reliable information as to when and how much radioactive substances were emitted. Google Japan, which developed the renowned Person Finder and many other useful services for the disaster, also considered services related to radioactivity as early as March 13, but the idea was abandoned because of the uncertainty of interpretations [17].

Academic societies kept silent. On March 18, the Meteorological Society of Japan published an unusual statement [35] on its website, asking the members of the society not to speak out, because the Ministry of Education, Culture, Sports, Science and Technology (MEXT) would provide official information. Such a "single voice" (so called in Japan) attitude was much criticized later.

Indeed, a system called SPEEDI (System for Prediction of Environmental Emergency Dose Information) [36], [37] was being operated by an organization affiliated with MEXT. At first, because no information of radioactive emission was known, only "unit emission" simulations, which supposed an emission of 1 Bq/h, were carried out. The results were not disclosed to the public, but were sent to the Fukushima prefectural office and other relevant parties by email or fax. SPEEDI calculates the distributions of emitted particles, as in **Fig. 7**, from realtime wind

*2 For general food, allowable limit of radioactivity is now set to 100 Bq/kg, which is more stringent than in most other countries. Virtually no food with detectable radioactivity is on the market.

Journal of Information Processing Vol.22 No.4 566–573 (Oct. 2014)

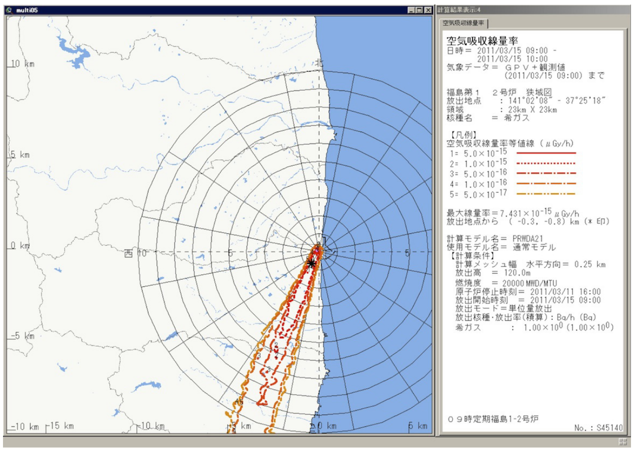


Fig. 7 A SPEEDI simulation result emailed to Fukushima prefecture on March 15, 09:32, just after the crucial radioactive emission occurred.

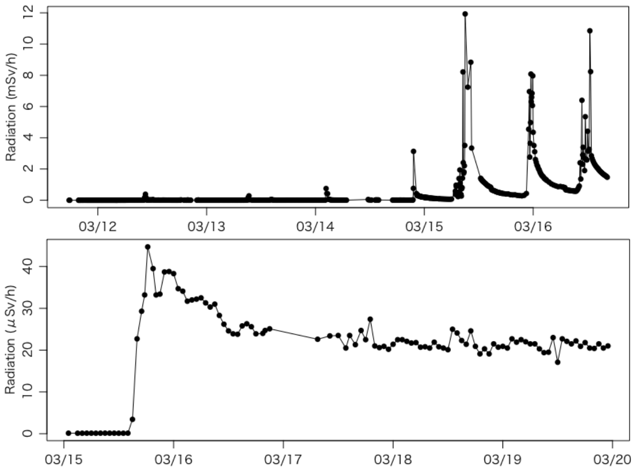


Fig. 8 (Top) Radiation at the main gate of Fukushima Daiichi, based on TEPCO data. The biggest emissions occurred during March 15–16. (Bottom) Radiation at Iitate village, northwest of Fukushima Daiichi.

data, based on the advection-diffusion model.

People at the Fukushima prefectural office were unable to make use of such simulations. They even deleted many such emails, because their mailbox was limited to a scant 25 MB [38].

All of the 86 SPEEDI results that Fukushima prefecture received by email and fax were identified and disclosed on the Web [39]. As always, they were in PDF. Volunteers created animated renditions of them (e.g., Ref. [40]), which showed that the wind directions were continually changing; without emission information, it was difficult to tell in which direction people should evacuate. It was reasonable for them to just go farther from the nuclear plant, provided that the road was not blocked by tsunami debris or fallen trees.

The crucial damage to one of the reactors occurred in the morning of March 15, making a surge in radiation levels, as shown in Fig. 8. At that time the wind was blowing southwest, as in Fig. 7, but during the course of the day the wind gradually changed to northwest. Rain in the evening, which later turned to sleet, made airborne radioactive substances drop and stay on the soil in the northwest direction. This can also be seen in Fig. 8 for Iitate village, 39 km northwest of the plant. During 13:00–14:00 radiation began to rise, and at 18:20 it recorded 44.70 μSv/h. Even the city of Fukushima, the capital of Fukushima prefecture, 63 km northwest from the plant, recorded 24.24 μSv/h at 18:40. These values were tabulated and published on the Fukushima prefecture website in PDF.

Fig. 9 Internal radiation dose of March 12–23, a “prediction” backward in time, released on March 23, 2011.

On March 15, a car from MEXT headed for the Fukushima Daiichi plant via Fukushima prefectural office. Later that day, at 20:40–20:50, they found an area of quite strong radiation, 200–300 μSv/h, 20 km northwest of the plant. A two-page PDF document [41, the last PDF file] showing the result and the map was published on the MEXT website early next day at 01:05, and `@mextjapan` tweeted the URL at 05:12. This was the first evidence of strong contamination in the nortwest direction.

More extensive on-vehicle measurements by MEXT on March 16–17 confirmed skewed contamination. The results were promptly released on the MEXT website. However, some newspapers accused MEXT of hiding the details.

On March 17–20, the US DOE/NNSA conducted aerial measurements [42]. The results clearly showed a skew toward the northwest direction. They were provided to Japan on March 18–22, but were not disclosed to the public [43].

On March 16, the office in charge of SPEEDI was transferred from MEXT to NSC [18].

On March 23, NSC released a map, **Fig. 9**, showing the first realistic estimation of exposure, for the period between 06:00 March 12 and 00:00 March 24, using the observed contamination as input, and running SPEEDI in the inverse simulation mode [44]. This “prediction” backward in time confused some media, which accused the government of concealing the computer prediction and exposing people to radiation [45], [46], [47] *3.

8. Changing URLs

According to Tim Berners-Lee, “cool URIs don’t change” [49], but many 3.11-related URLs have changed. An example of a strangely short-lived URL is the government’s recovery support website, http://hirogetai.go.jp, which was established on September 17, 2011 and vanished on June 30, 2012.

Radiation-related data from local governments were some-

*3 Indeed, some of the many early simulations “predicted” the northwest contamination. An example is the third PDF file of the second set of calculations linked from the NISA SPEEDI page [48].

times released on the Web in such a way that previous data were overwritten by new ones or silently deleted.

Just after 3.11, MEXT's website http://www.mext.go.jp was the best source of information on radioactivity from the government. MEXT soon established a dedicated website http://radioactivity.mext.go.jp. In hindsight, such a special-purpose website should not have been under the MEXT domain; it should have been independent of the administrative hierarchy of the government.

In April, 2012, the website was completely reorganized, and the old site was moved to http://radioactivity.mext.go.jp/old/. At this point the comprehensive realtime radiation map http://radioactivity.mext.go.jp/map/ was introduced. Fukushima prefecture also created a similar map at http://fukushima-radioactivity.jp. Most of the radiation telemeters within the prefecture are listed on both sites, but some are listed on only one of them, presumably for administrative reasons.

On September 19, 2012, the Nuclear and Industrial Safety Agency (NISA) and NSC were abolished and replaced by the new Nuclear Regulation Authority (NRA, http://www.nsr.go.jp; the domain name suggests that the organization name was changed at the last minute). The NISA and NSC websites were moved to http://www.nsr.go.jp/archive/nisa/ and http://www.nsr.go.jp/archive/nsc/, respectively. The radioactivity-related taskforce of MEXT was transferred to NRA, and the website http://radioactivity.mext.go.jp was changed to http://radioactivity.nsr.go.jp. The renowned radiation map http://radioactivity.mext.go.jp/map/ was also moved to http://radioactivity.nsr.go.jp/map/. At this point, the "old" site http://radioactivity.mext.go.jp/old/ was lost.

As another example, SPEEDI simulation results were initially published on at least three websites: MEXT [50], NSC [51], and NISA [48]. All of these URLs have changed since then. On March 2014, the Fukushima Prefecture website has undergone drastic restructuring, which changed every URL except the home page.

To solve the problem of ever-changing and vanishing Web contents, Japanese National Diet Library (NDL) has been conducting the Web Archiving Project (WARP) [52] since 2002. Compared to the renowned Internet Archive http://archive.org, the scope of WARP is limited because of the stronger copyright law of Japan. WARP archives some of the above-mentioned websites, as well as the National Diet of Japan Fukushima Nuclear Accident Independent Investigation Commission website, http://naiic.go.jp, which immediately went offline after its duty was accomplished. On the other hand, the Government's Investigation Committee (ICANPS) site http://icanps.go.jp/ is now part of the Cabinet Secretariat site, http://www.cas.go.jp/jp/seisaku/icanps/.

On November 5, 2012, NDL released the experimental version of its Great East Japan Earthquake Archive. The completed version was released on March 7, 2013 [53]. In accordance with Japanese copyright law, web archiving requires the consent of the copyright holder. Many organizations, including the central and local governments, opted to refuse many of their archived pages to be viewed on the Internet. Many pages are only allowed to be browsed inside the NDL building.

The Ministry of Internal Affairs and Communications (MIC) published a 354-page guideline for constructing and operating such digital archives [54]. One of their recommendations is to use durable formats such as PDF/A. However, archiving numerical data needs a separate consideration.

Recently, the Physical Society of Japan and the Japan Society for Archival Science jointly announced their effort to preserve and archive radiation data of the Daiichi accident [55], although the fruit is yet to be borne.

9. Conclusion

In the aftermath of 3.11, Twitter and other services on the net were valuable, but in-house information systems and systems with single points of failure often turned out to be unusable. The central and local governments relied on phones, emails and faxes when online groupware and videoconferencing were becoming the norm.

After the nuclear accident, TEPCO and the governments had trouble presenting data. They tabulated data neatly into paper-friendly PDF files, sacrificing promptness, accessibility, and machine-readability. Further, they were unaware of the importance of data archiving when they overwrote old files, changed URLs, and reorganized websites.

Media literacy was important when speculations and false information were rampant on the net and sometimes on the conventional media.

The 3.11 data and memories are quickly dissipating. It is our responsibility to preserve them and keep learning lessons from them.

References

[1] Okumura, H., Tatsumi, T. and Tohma, M.: Earthquake and ICT Education, *Proc. IPSJ Computer Education Symposium SSS2011*, pp.25–32 (2011). (in Japanese).

[2] Okumura, H.: Earthquake Disaster and Social Networks, *Information Processing*, Vol.52, No.9, pp.1072–1073 (2011). (in Japanese).

[3] Okumura, H.: Hilarious uses of Excel, *Proc. IPSJ Computer Education Symposium SSS2013*, pp.93–98 (2013). (in Japanese).

[4] 2011 White Paper on Information and Communications in Japan, Ministry of Internal Affairs and Communications, available from ⟨http://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2011/2011-index.html⟩ (accessed 2014-05-10). [especially, Part 1: Information and Communications in the Aftermath of the Great East Japan Earthquake].

[5] 2012 White Paper on Information and Communications in Japan, Ministry of Internal Affairs and Communications, available from ⟨http://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2012/2012-index.html⟩ (accessed 2014-05-10). [especially, Chap. 3: Lessons of the Great Earthquake and the Role of ICT].

[6] Secretariat of the Investigation Committee on the accident at the Fukushima Nuclear Power Station, available from ⟨http://www.cas.go.jp/jp/seisaku/icanps/eng/⟩ (accessed 2014-05-10).

[7] The National Diet of Japan, available from ⟨http://warp.da.ndl.go.jp/info:ndljp/pid/3856371/naiic.go.jp/en/⟩ (accessed 2014-05-10).

[8] The Independent Investigation on the Fukushima Nuclear Accident: *The Fukushima Daiichi Nuclear Power Station Disaster: Investigating the Myth and Reality*, Routledge (2014).

[9] Inoue, T., Toriumi, F., Shirai, Y. and Minato, S.: Great east Japan earthquake viewed from a URL shortener, *Proc. Special Workshop on Internet and Disasters*, DOI: 10.1145/2079360.2079368 (2011).

[10] The Coral Content Distribution Network, CoralCDN, available from ⟨http://coralcdn.org⟩ (accessed 2014-05-10).

[11] Fukuda, K., Aoki, M., Abe, S., Ji, Y., Koibuchi, M., Nakamura, M., Yamada, S. and Urushidani, S.: Impact of Tohoku Earthquake on R&E Network in Japan, *Proc. Special Workshop on Internet and Disasters: SWID '11*, pp.1:1–1:6, New York, NY, USA, ACM, DOI: 10.1145/2079360.2079361 (2011).

[12] NTT DOCOMO: Area Mail, available from ⟨https://www.nttdocomo.co.jp/service/safety/areamail/⟩ (accessed 2014-05-10). (in Japanese).

[13] Area Mail for disaster mitigation and Kesennuma Crisis Management Division's Twitter account, *LASDEC*, No.12, pp.41–46 (2010). (in Japanese).

[14] Calling for evacuation on Twitter and expectations for Area Mail, *LASDEC*, No.9, pp.5–8 (2011). (in Japanese).

[15] Mizuho Financial Group, available from ⟨http://www.mizuho-fg.co.jp/csr/mizuhocsr/report/list/2011/highlight_system/plan.html⟩ (accessed 2014-05-10). (in Japanese).

[16] Miyadai, S. et al.: *Earthquake and Nuclear Disasters in the IT Era*, Impress Japan (2011). (in Japanese).

[17] Hayashi, N. and Yamaji, T.: *Google's 72 hours: Great East Japan Earthquake, Information, and the Internet*, Kadokawa Shoten (2013). (in Japanese).

[18] Okamoto, K.: *Testimony: Haruki Madarame*, Shinchosha (2012). (in Japanese).

[19] Google Japan: Project 311, available from ⟨http://googledevjp.blogspot.jp/2012/09/project-311.html⟩ (accessed 2014-05-10).

[20] Okumura, H.: Project 311 — Twitter data, available from ⟨http://oku.edu.mie-u.ac.jp/~okumura/prj311/tw/⟩ (accessed 2014-05-10). (in Japanese).

[21] Ministry of Internal Affairs and Communications, available from ⟨http://www.soumu.go.jp/menu_kyotsuu/important/kinkyu01_000096.html⟩ (accessed 2014-05-10). [released May 13, revised May 15].

[22] Ministry of Internal Affairs and Communications, available from ⟨http://www.soumu.go.jp/menu_news/s-news/01kiban08_01000023.html⟩ (accessed 2014-05-10). (in Japanese).

[23] Shinsai FaxOCR, available from ⟨https://code.google.com/p/faxocr/⟩ (accessed 2014-05-10). (in Japanese).

[24] TEPCO: Electricity Forecast, available from ⟨http://www.tepco.co.jp/en/forecast/html/⟩ (accessed 2014-05-10).

[25] Ministry of Economy, Trade and Industry, Twitter, available from ⟨https://twitter.com/openmeti/status/86419160002859008⟩ (accessed 2014-05-10).

[26] McCallum, Q.E. et al.: *Bad Data Handbook*, O'Reilly (2012). especially Chapter 3, Data Intended for Human Consumption, Not Machine Consumption, by Paul Murrel.

[27] RFC 4180 – Common Format and MIME Type for Comma-Separated Values (CSV) Files, available from ⟨http://tools.ietf.org/html/rfc4180⟩ (accessed 2014-05-10).

[28] A summary page of radiation monitoring data and graphs related to the nuclear power plant accident in Fukushima, Japan, radmonitor311 collaboration, available from ⟨https://sites.google.com/site/radmonitor311/⟩ (accessed 2014-05-10). (mostly in Japanese).

[29] Ministry of Economy, Trade and Industry, available from ⟨http://www.meti.go.jp/policy/mono_info_service/joho/other/2011/0330.html⟩ (accessed 2014-05-10). (in Japanese).

[30] Okumura, H.: Provide data in machine-readable format, Okumura's Blog, available from ⟨http://oku.edu.mie-u.ac.jp/~okumura/blog/node/2578⟩ (accessed 2014-05-10). (in Japanese).

[31] Gadget Tsuushin, available from ⟨http://getnews.jp/archives/109133⟩ (accessed 2014-05-10). (in Japanese).

[32] Sugiyama, J. and Uda, W.: Information processing of food for safety and security, *Information Processing*, Vol.52, No.11, pp.1370–1375 (2011). (in Japanese).

[33] Okumura, H.: Food Radioactivity Data Search, available from ⟨http://oku.edu.mie-u.ac.jp/food/⟩ (accessed 2014-05-10). (in Japanese).

[34] Becker, M.: Risiko Wetter: Wind bläst radioaktive Wolke nach Tokio, Spiegel, available from ⟨http://www.spiegel.de/wissenschaft/natur/bild-750835-191816.html⟩ (accessed 2014-05-10).

[35] The Meteorological Society of Japan, available from ⟨http://www.metsoc.or.jp/others/News/message_110318.pdf⟩ (accessed 2014-05-10).

[36] Kitamoto, A.: Radioactive material diffusion simulation by SPEEDI: Problems revealed from between the ideal and reality, *Suugaku Seminar*, Vol.50, No.12, pp.18–23 (2011). (in Japanese).

[37] Sato, Y.: *SPEEDI: Why it wasn't used*, Toyo Shoten (2013). (in Japanese).

[38] Fukushima Minpo, available from ⟨http://www.minpo.jp/pub/topics/jishin2011/2012/04/post_3757.html⟩ (accessed 2014-05-10). (in Japanese).

[39] Fukushima Prefecture, available from ⟨http://www.pref.fukushima.lg.jp/sec/16025c/genan13.html⟩ (accessed 2014-05-10). (in Japanese).

[40] Sumiya, T.: Hiroshima University, available from ⟨http://sumi.riise.hiroshima-u.ac.jp/speedi/⟩ (accessed 2014-05-10). (in Japanese).

[41] MEXT, available from ⟨http://radioactivity.nsr.go.jp/ja/list/207/list-201103.html⟩ (accessed 2014-05-10). (in Japanese. Original URL: http://www.mext.go.jp/a_menu/saigaijohou/syousai/1303726.htm, archived at https://web.archive.org/web/20110320023130/http://www.mext.go.jp/a_menu/saigaijohou/syousai/1303726.htm. The current page with corrections is at http://radioactivity.nsr.go.jp/ja/list/207/list-201103.html).

[42] US DOE/NNSA Response to 2011 Fukushima Incident—Data and Documentation, U.S. Department of Energy, available from ⟨http://energy.gov/downloads/us-doennsa-response-2011-fukushima-incident-data-and-documentation⟩ (accessed 2014-05-10). Also available at http://www.data.gov/.

[43] Upper House proceedings information, 180th National Assembly, Budget Committee, No.22, available from ⟨http://kokkai.ndl.go.jp/SENTAKU/sangiin/180/0014/18007100014022a.html⟩ (accessed 2014-05-10).

[44] Nuclear Safety Commission: Press Release, March 23, 2011, Nuclear Safety Commission, available from ⟨http://www.nsr.go.jp/archive/nsc/info/110323_top_siryo.pdf⟩ (accessed 2014-05-10). [URL as of release: http://www.nsc.go.jp/info/110323_top_siryo.pdf].

[45] Onishi, N. and Fackler, M.: Anger in Japan Over Withheld Radiation Forecasts, New York Times, available from ⟨http://www.nytimes.com/2011/08/09/world/asia/09japan.html⟩ (accessed 2014-05-10).

[46] Birmingham, L. and McNeill, D.: *Strong in the Rain: Surviving Japan's Earthquake, Tsunami, and Fukushima Nuclear Disaster*, Palgrave Macmillan (2012).

[47] Samuels, R.J.: *3.11: Disaster and Change in Japan*, Cornell University Press (2013).

[48] Nuclear and Industrial Safety Agency, available from ⟨http://www.nsr.go.jp/archive/nisa/earthquake/speedi/speedi_index.html⟩ (accessed 2014-05-10). (in Japanese. Originally: http://www.nisa.meti.go.jp/earthquake/speedi/speedi_index.html).

[49] Berners-Lee, T.: Hypertext Style: Cool URIs don't change, W3C, available from ⟨http://www.w3.org/Provider/Style/URI.html⟩ (accessed 2014-05-10).

[50] MEXT, available from ⟨http://radioactivity.nsr.go.jp/ja/list/201/list-1.html⟩ (accessed 2014-05-10). (in Japanese. Originally: http://radioactivity.mext.go.jp/ja/distribution_map_SPEEDI/).

[51] Nuclear Regulation Authority, available from ⟨http://www.nsr.go.jp/archive/nsc/mext_speedi/⟩ (accessed 2014-05-10). (in Japanese. Originally: http://www.nsc.go.jp/mext_speedi/).

[52] Web Archiving Project, National Diet Library, available from ⟨http://warp.da.ndl.go.jp⟩ (accessed 2014-05-10). (in Japanese).

[53] NDL Great East Japan Earthquake Archive, National Diet Library, available from ⟨http://kn.ndl.go.jp⟩ (accessed 2014-05-10). (in Japanese).

[54] Ministry of Internal Affairs and Communications, available from ⟨http://www.soumu.go.jp/menu_seisaku/ictseisaku/ictriyou/02ryutsu02_03000114.html⟩ (accessed 2014-05-10). (in Japanese).

[55] The Physical Society of Japan, available from ⟨http://www.jps.or.jp/information/2013/11/post-11.html⟩ (accessed 2014-05-10). (in Japanese).

Haruhiko Okumura was born in 1951. He received his M.S. degree in Physics from Nagoya University in 1978, and his doctorate from the Graduate University for Advanced Studies in 1999. His current research interest is computer education. He is a member of IPSJ, ACM, and JPS (the Physical Society of Japan).