

[解説]

非常時データ提供の問題点

奥村 晴彦

1 はじめに

東日本大震災で、情報システムに大混乱が生じ、現在のシステムの問題点がいろいろ明らかになった。こうした問題点、特に放射線量などの非常時データの集約作業を通じて明らかになったことがらについて、著者は2011年5月21日の情報システム学会シンポジウムで話させていただいた。本稿はその内容に加筆したものである。

2 震災での情報システム

大災害があると、まず電話が輻輳して大規模な通話規制がかけられ、ほとんど使えなくなるのが通例である。東日本大震災でも、固定電話の主要3社で最大80~90%、携帯電話の主要3社で最大70~95%の通話規制がかけられた[1]。PHSやIP電話は問題なく使えたようである。一方、パケット通信はNTTドコモが一時的に30%の規制をかけた程度で、ほぼ問題なく使えた。被災地の携帯電話基地局も、故障やケーブル断がない限り、非常電源に切り替わって数時間使えた。携帯メールは、輻輳での遅延以外に、プッシュ機能の障害で届かない場合があったが、その場合は「センター問い合わせ」で通常のパソコンのメール同様に取得することができた。

Twitterのような軽量のパケット通信が最も頑強であった。震災の日にはTwitterサービス開始以来最多の1.77億ツイートが書き込まれ、翌日にはこれも過去最多の57.2万アカウント

が新規作成された[2]。

首都圏でも、交通機関の運行状況、帰宅困難者のための避難場所の情報の入手にTwitterが役立った。避難場所の情報は東京都のホームページにも掲載されたが過重なアクセスに耐えず、猪瀬直樹副知事(@inosenaoki)自らTwitterに情報を流した。気仙沼の施設の子どもたちが公民館に取り残されているというツイートを受けて、猪瀬副知事はヘリでの救助を手配し、12日に全員救助されたという話は有名である。また、気仙沼市危機管理課(@bosai_kesenuma)はTwitterで津波警報や避難指示、被災状況を流し続けた。被災地で実際にTwitterを活用できた人がどれくらいいたかは不明であるが、気仙沼のツイートはテレビなどでも紹介され、被災状況を多くの人に伝えることができた。

気仙沼市はこのような災害時にNTTドコモの緊急速報「エリアメール」を使う予定[3]であったが、エリアメールの情報発信用端末は停電・通信回線寸断により使えなかった。

エリアメールも含め、セキュリティ上の理由で特定の端末からしか更新できないようにした情報発信システムは多いが、結果的に非常時に情報発信ができないことになる。Twitterは避難しながら携帯端末からも情報発信できるので便利であったが、他の経路による情報発信も携帯端末から行えるように設計することは十分可能なはずである。

3月11日の事態を予測できず現実離れの危機管理体制しかとっていなかった事例はほかにも多数見られた。例えば原子力安全委員会は、緊急事態応急対策調査委員40人を携帯電話のメールで招集しようとしたが、ほとんどの委員に連絡がつかず、連絡がついた委員も交通機関が止まっていたことから、当日参集できたのは歩いてきた数人だけであった。連絡方法は携帯

Haruhiko Okumura

三重大学 教授

[解説]

2011年8月31日受付

© 情報システム学会

メール以外にも複数用意すべきであり、招集できない事態に備えて TV 会議や Skype 等の利用など複数の手段を準備しておくべきであろう。

情報発信サイトが過重なアクセスに耐えず、利用不能になるというケースも非常に多かった。例えば三重県は「防災みえ」[4]というサイトを用意しているが、過去にも台風の際にアクセス過重で使えなかったという経験が生かされず、今回も使いものにならなかった。緊急時にリソースを増やすことのできるクラウド型 (PaaS, IaaS) システムが期待される。

今回はアクセス過多に陥ったシステムにはミラーサイトが多数作られた。例えば文科省サイトは Yahoo!, さくらインターネット, WIDE, Goo, Microsoft にミラーを置いた。県ごとの放射線推移グラフなどは Windows Azure のクラウドに置いた。

ただし、単純に元サイトからミラーサイトへのリンクを張るだけでは、元サイトがアクセスできなくなればミラーサイトにも来てもらえない。

この問題を解決する方法はいろいろ考えられる。例えば検索サイトがミラーサイトに導いてくれるような仕組みも有効であろう。もっと簡単な仕組みで相互ミラー・負荷分散をしている国内の例として、Ring Server Project [5]や Debian JP Project [6]などがある。これらは、各地にミラーサイトを置くだけでなく、代表 URL (Ring では <http://www.ring.gr.jp/>, Debian JP では <http://ftp.jp.debian.org/>) にアクセスすると、DNS ラウンドロビンでランダムなミラーに振り分けられる。Ring ではよりネットワーク的に近いミラーを選ぶ仕組みも実装されている。

データのバックアップ体制も問題になった。自治体の戸籍データは、その地方の法務局に写しがあるだけで、法務局ともに被災した 4 市町の一部データが滅失した。住基データについては、基本 4 情報は住基ネット側にも存在する。それ以外は、委託業者が保管していたデータからほぼ復元できたようである (それはそれで問題があるかもしれない)。

通話規制で注目されたのが災害用伝言ダイヤル (171)、災害用伝言板である。特に NTT 東日本・NTT 西日本の災害用ブロードバンド伝言板「Web171」は、キャリアを問わずパソコンからも使えるサービスとして NHK ニュースなどでも紹介され、アナウンサーが「<https://www.web171.jp/top.php>」という URL を読み上げるのに苦労していたが、top.php のような不要なファイル名まで含み、しかも http から https へのリダイレクトの仕組みがないことから、NHK のニュースを聞いて手で URL を打ち込むのは困難であったであろう (実際は top.php は省略可であった。top.php や index.html のような本来不要なファイル名は URL に含めるべきでない)。しかも、このサイトは、数千文字からなる利用規約を読んで同意しなければ次に進めない。さらに、利用環境として Windows 2000 SP4 と Windows XP SP2, IE 5.5~6.0 を挙げているのは信じがたい。

一方で、地震の 2 時間後に開始された Google パーソンファインダー (短縮 URL は <http://goo.gl/sagas>) は、NHK・朝日新聞社・警察庁・福島県・岩手県・毎日新聞社希望新聞のデータも取り込み、避難所の名簿の写真をボランティアで文字入力したものも加え、国内最大規模の安否情報サイトとなった。

3 放射線などの数値情報提供の問題点

震災直後の福島第一原子力発電所の事故により、放射線の線量や放射性物質の濃度の情報が大量に提供されることになった。しかし、その提供方法には種々の問題があった。

福島第一原子力発電所は敷地境界の 8 個のモニタリングポスト (MP) での放射線を実時間でネット公開していた。サイトの URL は <http://www.tepco.co.jp/fukushima1-mp/monitoring/index2.html> であるが、フレームを使っているため、この URL だけではアクセスできず、さらに左下の「モニタリングポスト測定値」アイコンをクリックしなければならない。しかし、そこで公開されている値は、本稿執筆時点 (2011 年 8 月 21 日) においても、本震直前の

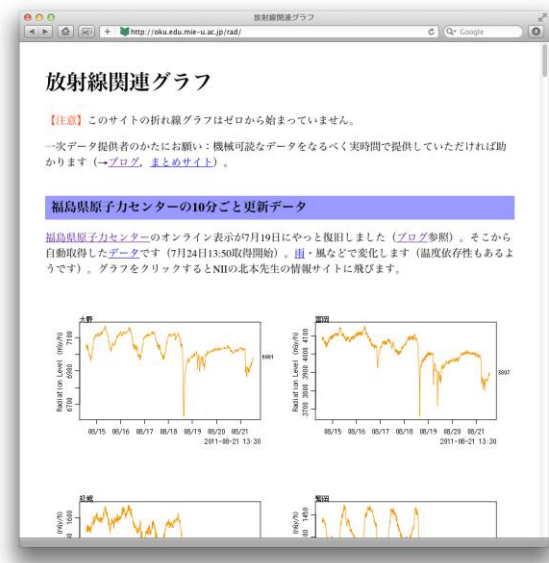
「3月11日14時40分現在」のままである。モニタリングポスト自体は動作しているので、ネット公開がいつこうに復旧せず、マスコミも追及していない。現在、測定値は1日4回 <http://www.tepco.co.jp/nu/fukushima-np/fl/index-j.html> で公開されている（ただし午後11時50分のは翌朝午前9時のものと同時に公開されるので実質1日3回の更新である）。リアルタイムシステムがすぐ復旧していれば原子炉建屋の水素爆発の影響の規模をいち早く知ることができたので残念である。

福島原子力センター <http://www.atom-moc.pref.fukushima.jp/> の環境放射能監視テレメータシステムも3月12日12:00時点で止まっていた。6月10日に修理されるとニュース報道された後もリアルタイム公開はすぐには再開せず、6月18日から福島県のサイトでPDF形式で1日1回公開されていたが、批判（例えば[7]）を受け、7月19日に復旧した。実時間グラフは私のサイト[8]でも公開している。

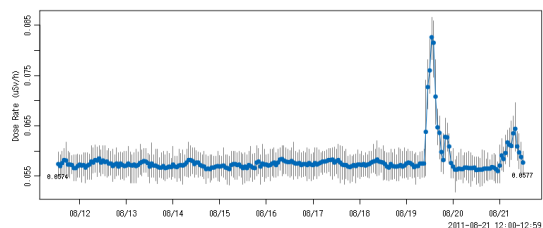
最近では多数の関係者の努力のおかげで改善されてきたが、震災後はほとんどの公的な情報がPDF形式でしか公開されないという状況が続いた。なかには紙資料をスキャンしただけのPDFファイルもあり、ファイルサイズも必要以上に大きく、古い携帯端末では表示されず、スマートフォンでも読みにくい。

そこで、財団法人地方自治情報センター（LASDEC）は3月18日付の「国民へ発信する重要情報のファイル形式について」[9]で、PDF・Excel形式ではなくHTML・CSV形式で情報発信するよう呼びかけた。LASDECの挙げる理由は「容量が大きく、サーバー・回線リソースを圧迫」であり、アクセシビリティや（後述の）情報の再利用については明確には書かれていなかったが、問題の本質は容量ではなくアクセシビリティと再利用の容易性である。

3月13日から著者は東京電力のPDFファイルのCSV化を始めた[10]。3月17日からは文科省の都道府県別データもCSV化を始めた。こういった動きはTwitterを通じて大勢のボランティアに広まり、まとめサイト[11]ができ



著者の放射線関連グラフサイト
<http://oku.edu.mie-u.ac.jp/rad/>



東京都の放射線の変化（著者のサイトの自動グラフの一つ）

た。

しかし、苦勞してPDFをCSVに変換するのは無駄である。特に画像PDFはOCRを利用しても間違いが入りやすい。3月20日、著者はブログで「データは自動処理可能な形で提供してほしい」と訴えた[12]。

東電は、3月23日には電力の使用状況データをCSVでも提供を始める一方、3月21日から放射線データについてはPDFファイルにデータ抽出禁止の保護をかけた。これに対して経産省は3月30日に「東北地方太平洋沖地震等に係る情報提供のデータ形式について」[13]という連絡で機械可読データの公開を呼びかけ、ニュースサイト「ガジェット通信」[14]も東電への質問と回答を掲載してその態度を批判した。

こうした批判を受けて、ついに東電の放射線

データが CSV 化されたのは 4 月 12 日である。

最初は機械可読形式の意義を東電がよく理解しておらず、単位の異なるデータを同じ欄に並べたり、形式を随時に変更したりした。特に 7 月 1 日には電力データの CSV 版の形式を突然変更したため、これを利用するアプリやサイトで表示が乱れた[15]。

文科省の都道府県別データ[16]も 8 月 8 日から主たる配布形式が PDF から HTML に変更された。一般に HTML のほうが PDF より軽量かつデータ抽出が容易であり、この変更はありがたいが、よく見ると妙な「セルの結合」をしたデータを HTML に変換したもので、データの再利用の容易さという点からはいま一步である。

東電の PDF ファイルも、きれいに罫線を引いた美しい作表であるが、よく見ると数値に全角と半角が混在しており、小数点がコンマになったり、単位「 μ Sv/h」を全部の数値に付けていてときどき「 μ 」を落としたり、酷い例では数字「0」が「お」と誤入力していたりで、数値データとして利用することはまったく考えていなかったようである。

4 どのようなデータが望ましいか

ここでは簡単に適用できる私案をまとめておく。

公開のためのフォーマットとしては、CSV、TSV、JSON、XML などのファイル、あるいは API による提供が考えられるが、現時点で一つ選ぶとすれば CSV ファイルがわかりやすいであろう。文字コードがシフト JIS、行末が CRLF、拡張子が csv であれば、ダブルクリックするだけで Excel で開くことができる。細かい注意点として、例えば文科省の都道府県別データは 9 時から 10 時までを「9-10」としているが、Excel で開くと「9 月 10 日」に変換されてしまう。日時データは"2011-03-11 14:46:18"あるいは"2011/3/11 14:46"といった文字列で表せば Excel でも R などのデータ解析ソフトでも正しく解釈される。

位置情報を伴う場合は、緯度・経度データを付けてあれば、Google マップに容易に重ね書

きできる。放射線量の分布を可視化するためには緯度・経度が不可欠であるが、残念ながら地名しか書かれていない公表データが多い。

入力ツールとしては Excel でもかまわないが、「セルを結合」すると正しい CSV 出力が得られなくなる。

よくある失敗として、例えば正門と西門で放射線量を測定するつもりで

日時,正門,西門

2011-05-18 08:40,34.5,16.4

2011-05-18 08:50,,16.5

のようなデータを作り続けて、途中で計測場所が変わったので項目を変えたり増やしたりしてしまうことがある。場所が変わりうるなら

日時,場所,値

2011-05-18 08:40,正門,34.5

2011-05-18 08:40,西門,16.4

2011-05-18 08:50,西門,16.5

のようにする。

5 おわりに

大学関係では、シラバスデータを含め学術情報データは RDBMS で管理し、必要に応じて (X)HTML 形式に変換して Web ページとしたり PDF 形式で帳票出力したりするといったことが広く行われている。企業でもそうであろう。ところが今回の震災で驚いたことは、きれいに罫線を引いた紙ベースの出力のための PDF ファイルがまずあって、それを視覚化・データ解析するために PDF ファイルから数値を抽出してデータを作るという逆の手順を踏まなければならなかったことである。考え方を逆転して、まずは機械可読なデータから出発し、それを用途ごとに変換して使うことで、仕事の効率は著しく向上し、間違いも減るであろう。即座にグラフ化・視覚化することでの的確な判断がしやすくなる。パソコン・携帯端末といったデバイスごとに最適化された出力を用意することも簡単であり、公開された情報のアクセシビリティも高まる。

参考文献

- [1] 総務省『情報通信白書』平成 23 年度版
<http://www.soumu.go.jp/johotsusintokei/whitepaper/> (2011 年 8 月)
- [2] Twitter Blog: #numbers,
<http://blog.twitter.com/2011/03/numbers.html> (2011 年 3 月 14 日)
- [3] NTT ドコモ「宮城県初！緊急速報「エリアメール」を気仙沼市が採用」
http://www.nttdocomo.co.jp/info/notice/tohoku/page/100413_01.html (2010 年 4 月 13 日)
- [4] 防災みえ <http://www.bosaimie.jp/>
- [5] Ring Server Project,
<http://www.ring.gr.jp/>
- [6] Debian JP Project,
<http://ftp.jp.debian.org/>
- [7] 奥村晴彦「福島県・東電の実時間放射線モニタを早く再開してほしい」
<http://oku.edu.mie-u.ac.jp/~okumura/blog/node/2579> (2011 年 6 月 19 日)
- [8] 奥村晴彦「放射線関連グラフ」
<http://oku.edu.mie-u.ac.jp/rad/> (自動更新)
- [9] 財団法人地方自治情報センター (LASDEC)「国民へ発信する重要情報のファイル形式について」
<https://www.lasdec.or.jp/cms/12,22060,84.html> (2011 年 3 月 18 日)
- [10] 奥村晴彦,
<http://oku.edu.mie-u.ac.jp/~okumura/stat/data/>
- [11] radmonitor311: 放射線量モニターデータまとめページ
<http://sites.google.com/site/radmonitor311/>
- [12] 奥村晴彦「データは自動処理可能な形で提供してほしい」
<http://oku.edu.mie-u.ac.jp/~okumura/blog/node/2578> (2011 年 3 月 20 日)
- [13] 経済産業省「東北地方太平洋沖地震等に係る情報提供のデータ形式について」
http://www.meti.go.jp/policy/mono_info_service/joho/other/2011/0330.html (2011 年 3 月 30 日)
- [14] ガジェット通信「東京電力が放射線量測定データを「使いづらく」している理由」
<http://getnews.jp/archives/109133> (2011 年 4 月 6 日)
- [15] 奥村晴彦「東電の CSV 形式が突然変更された」
<http://oku.edu.mie-u.ac.jp/~okumura/blog/node/2580> (2011 年 7 月 2 日)
- [16] 文部科学省「都道府県別環境放射能水準調査結果」
http://radioactivity.mext.go.jp/ja/monitoring_by_prefecture/

著者略歴

1951 年生まれ、名古屋大学大学院理学研究科修士課程修了 (物理学)、高校教諭・松阪大学助教授・松阪大学教授を経て、2004 年度から三重大学教育学部教授 (情報教育)。博士 (学術)。2011 年度は三重大学大学院地域イノベーション学研究科教授・三重大学学長補佐 (情報担当)・CIO 補佐・三重大学高等教育創造開発センター教授・三重大学総合情報処理センター教授を兼務。