

イベント情報のXML化 -NewsMLによるイベント情報配信-

井上 明^{*1}、猪狩 淳一^{*2}、小野寺 尚希^{*2}
藤原 隆弘^{*3}、永井 智子^{*4}、金田 重郎^{*4}

^{*1} 甲南大学 情報教育研究センター、^{*2} 毎日新聞社

^{*3} イースト株式会社、^{*4} 同志社大学 工学部

概 要

本研究では、ニュース情報の配信管理フォーマットである NewsML を利用し、イベント情報の受信・発信・管理のための XML フォーマットを規定する。NewsML は、テキストデータはもとより、画像、動画、音声といったさまざまな素材を管理することができ、新聞紙面はもとより、Web、携帯端末、組版などの各種メディアへ配信する機能を有する。これらの NewsML の機能を活用した以下のシステムを試作した。1) NewsML 化したイベント情報の入出力をおこなうエディタ、2) イベント情報の携帯端末への出力システム、3) 地図情報とリンクしたイベント情報発信システム。これらのシステム構築を通じてイベント情報の共通化・受発信の新たな仕組みの可能性を検討する。

Event information using NewsML

*AkiraINOUE^{*1}, JunichiIGARI^{*2}, NaokiONODERA^{*2}
TakahiroFUJIWARA^{*3}, TomokoNAGAI^{*4}, ShigeoKANEDA^{*4}*

^{*1} Konan University, ^{*2} The Mainichi Newspapers Co., ^{*3} EAST Co.,Ltd.

^{*4} Doshisha University

E-mail:inoue-a@center.konan-u.ac.jp

Abstract

This paper demonstrates a practical event information inputting and publishing system by using NewsML. NewsML is the news management format of the next generation. NewsML manages data, such as not only text data but an animation, a sound, etc. Moreover, It can offer information to media, such as a newspaper, television, and radio and Web. We built the following system. 1)Event information Editing system. 2)The output system to the Cellular phone of event information 3)Map information and the linked event information dispatch system

1 はじめに

観光・祭り・展覧会など様々なイベント情報が、新聞社や行政などの Web ページや紙媒体を通じて広く公開されている。しかし、そのデータフォーマットは統一化されておらず、同一のイベント情報が様々

な形式で配信・管理されている。また、これらの情報の流過程はほとんど電子化されていない。本来、イベント情報に限らず、情報は容易に再利用可能な形で提供されるべきである。

ある地方自治体でのイベント情報の作成から、新聞社への情報提供までの作業を調査した。

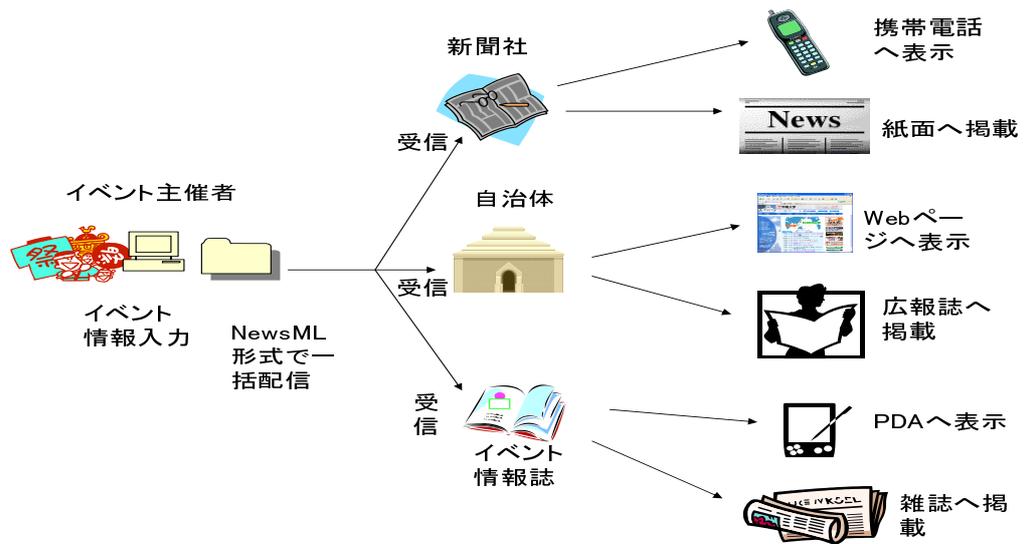


図 1: イベント情報のワンストップサービス・イメージ図

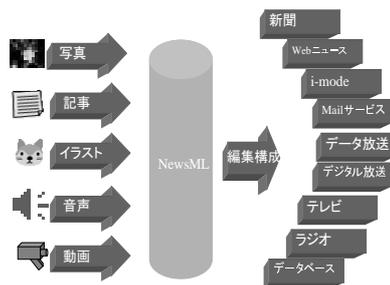


図 2: NewsML の機能

行政各部署はイベント開催者からの情報提供や取材により情報を集める。次に、担当部局にてワープロ入力し印刷する。次に、印刷した紙を広報担当部署へ渡す。広報部署では、原稿をもとに広報誌用組版データを作成する。また別途、Web ページ用 HTML、そしてプレスリリースを作成する。

新聞社はこのプレスリリースを FAX などを受け取る。そして、プレスリリースからのデータ投入がおこなわれ、その後、紙面用データ作成、Web ページ用データ作成などが必要になる。このように人的作業やデータの再入力が多に多い。

現時点でも、イベント情報を収集し、ホームページ等で公開するシステムはすでに各地で部分的には

開発が行なわれている。しかし、以下のコンセプトに基づいていることが多い。

1. サーバを設置して、情報発生元（例えば、自治体）から遠隔地より Web ブラウザ経由で入力してもらう
2. 種々の組織から入力された情報を、ホームページに公開する

このようなアプローチは、幾つかの問題がある。最大の問題は、ワンストップサービスになっていないところである。例えば、自治体が、ある新聞社の公開するこの種のシステムにイベント情報を入力したとする。この情報は確かに、当該システムにより新聞社のホームページに掲載することは可能である。

しかし、他の新聞社・放送局は、この情報をどうやって利用すればよいのであろうか。何度もホームページを見て、必要であれば、紙で打ち出して、再度、パソコン投入するか、しかない。また、同類のシステムがいくつもの新聞社から公開されていた場合、自治体は同じ作業を繰り返さなければならない。これでは、前述の自治体の中の情報流通と同様に、大きな無駄を生むのみである。

本研究では、社会全体でのイベント情報のワンストップサービス (図 1) を実現するために、1) イベント情報のデータ形式の標準化、2) イベント情報の入

出力・蓄積が容易にできるツールの提供、3) 実際の活用事例の試作、を行い、イベント情報公開・流通システムの構築を目指す。

2 NewsML

本研究では、イベント情報の管理・配信フォーマットへ「NewsML」を適用する。まず、NewsML について説明する。

NewsML は、XML(Extensible Markup Language) をベースにしたニュース管理標準フォーマットである。テキスト、画像、動画、音声といったさまざまなニュース素材を管理し、新聞紙面はもとより、Web、携帯端末、テレビ、データ放送などの各メディアへ配信する機能を有する。

例えば、あるニュースについて「いつ」「どこで」「どのようなことが起こったのか」を決められたフォーマットにしたがって管理できる。さらに、ニュース記事の中に日本語、英語など異なる言語の文章を混在させたり、動画コンテンツや音声データを使用したマルチメディアニュースの作成、配信を可能にする。このようなニュース構造の共通化や扱うデータフォーマットの多様化という機能だけでなく、ある事象がどのように変化していったかというニュースの履歴管理や、新聞社や執筆者など情報の提供元に対する著作権管理機能なども組み込まれている。

ニュース配信元はデータ形式を NewsML にしておくことで、新聞やテレビ、PDA や携帯電話といった配信先の利用形態を意識することなく NewsML という共通フォーマットで効率よく情報を配信できる。一方、情報の受け側は、文章、画像、動画といった各種データフォーマットの中から必要な情報のみを使用することが可能になる(図2)。NewsML では、わが国だけでなく、世界的にニュースフォーマットの構造を標準化することを目的としており、その普及と発展を目指し様々な組織が NewsML の採用を進めている。NewsML の普及により、データフォーマット差異や言語の壁を気にすることなく、ニュース情報の共有化や蓄積に関して新たな発展が期待されている。

3 イベント情報の NewsML 化

同志社大学と毎日新聞社は、2001年4月、イベント情報を XML 化し、ひとつのイベントデータを新聞紙面と Web ページへ提供するシステムを共同開発した。このシステムでは、イベントデータは XML 化されているが、NewsML 化されていない。今回、本システムを元にイベント情報の NewsML 化を行った。

3.1 毎日新聞社版イベント XML の再構成

まず、既存の、毎日新聞社で使用していた XML 化したイベントデータを再構成した。XML インスタンスは独自に規定していたために、NewsML 化するために再度、全てのタグ構成を見直した。図3は、既存 XML を NewsML 化するための検討図である。

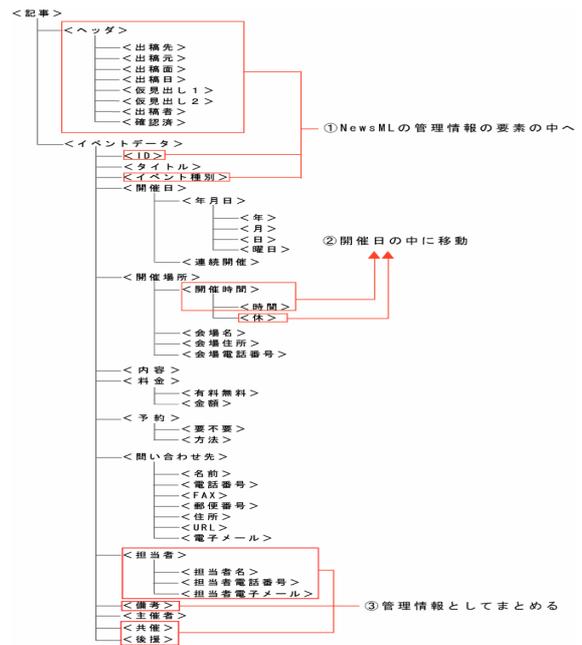


図3: 毎日新聞社イベントシステムの XML を NewsML へ再構成

3.2 NewsML 化の流れ

次に、NewsML との統合化をおこなった。NewsML はその構造、使用するタグが詳細に決められている。例えば、表示用に用意されている記事の作成者

メタデータ	意味	NewsML 上の格納場所
出稿先	イベント情報の送信先	NewsEnvelope/SentTo
出稿元	イベント情報の単に送信する元	NewsEnvelope/SentFrom
出稿面	新聞紙の何面に載るかを示す	NewsComponent/Metadata
出稿日	新聞紙面に載る日	NewsComponent/NewsLines/NewsLine NewsComponent/Metadata
仮見出し 1	ファイル名の一部	NewsComponent/AdministrativeMetadata/FileName
仮見出し 2	イベント種別	DiscriptiveMetadata/SubjectCode NewsComponent/Metadata
出稿者	記者名	NewsComponent/NewsLines/ByLine
ID	ファイル名	NewsItem/NewsIdentifier/NewsItemId

表 1: メタデータの NewsML での格納場所

の記述は、”NewsComponent/NewsLines/ByLine” という位置に格納される。また、ユーザが独自に拡張できる部分もある。

NewsML 化にあたり、イベントデータの構造を、1) イベント管理情報、2) イベント詳細情報、に分類した。イベント情報には、イベントの名称や開催場所といった「イベント詳細情報」の他に、いつ誰がそのイベント記事を書いたのか、といった「イベント管理情報」がある。

そこで、「管理情報である記事作成者は NewsML の ByLine を使う」「イベント詳細は NewsML の”記事”本文中に入れる」といったルールを規定しながら、作業を進めていった。次項以降にその詳細を述べる。

3.3 イベント管理情報の NewsML 化

イベント管理情報として扱うメタデータを、要素名、その意味、NewsML での格納場所を表にしたものが表 1 である。

注意点として、以下の 4 つある。

1. 出稿日は、新聞紙に載る日として『組日』と呼ばれる用語があるため、より意味が分かるように変更
2. 「イベント種別」「組日」は 2 つの格納場所がある
3. 「イベント種別」「出稿者」「組日」は NewsML 上にない管理情報を追加して管理
4. 「組日」は具体的に新聞紙面に載る日が分かる YYYYMMDD 表記を用いて Metadata 中で管理し、「何日組何刊」などのテキスト表記を NewsLine 中に管理

次に、上記の仕様が決まっただけでは、NewsML の管理は完全には成立しないことを示すため、FormalName 属性について説明する。

FormalName 属性は NewsML を読むポイントである。なぜなら、NewsML のほとんど全ての要素に存在するからである。例えば、出稿先を例に挙げると、< Party FormalName=”京都支局” /> と記述するが、これは次のように読む。

『Party の値はコントロールド・ボキャブラリ¹の京都支局である。京都支局の値は TopicSet ファイルに記述されている』

TopicSet は、送信者・受信者間で管理したい値を定義する。また、TopicSet で定義したボキャブラリの位置を Catalog で指定する。よって、新しく定義する FormalName 属性の全ての値に対して、既存の TopicSet ファイルを拡張し、参照先である Catalog を記述しなければならないことが分かる。具体的には、以下の 3 つがあげられる。

1. 出稿先、出稿元のために、「Party」の TopicSet ファイルを拡張
2. 組日のために「NewsLineType」の TopicSet ファイルを拡張
3. イベント種別、出稿面、組日のために、「MetadataType」「Property」の TopicSet ファイルを拡張

また、上記の他にも専用の TopicSet ファイルを作らなければならない箇所がある。それは、Metadata 中の、「ハイキング」、「二京都面」、「20040115」などイベント種別、出稿面、組日 3 つに対する各々の Value 属性である。

次に、イベントの名称や開催場所などのイベント詳細情報を NewsML 化した。

3.4 イベント詳細情報の NewsML 化

イベント情報本体は、NewsComponent/ContentItem/DataContent 中に全て一括して格納できるため分かり易い。DataContent は通

¹NewsML 上の FormalName 属性に記述する値は、任意の文字列は許されない。送信者、受信者間でデータ交換する場合、決められた値しか入れられない。このように値そのものが、コントロールされたボキャブラリであるという意味で、「コントロールド・ボキャブラリ」と呼ばれる

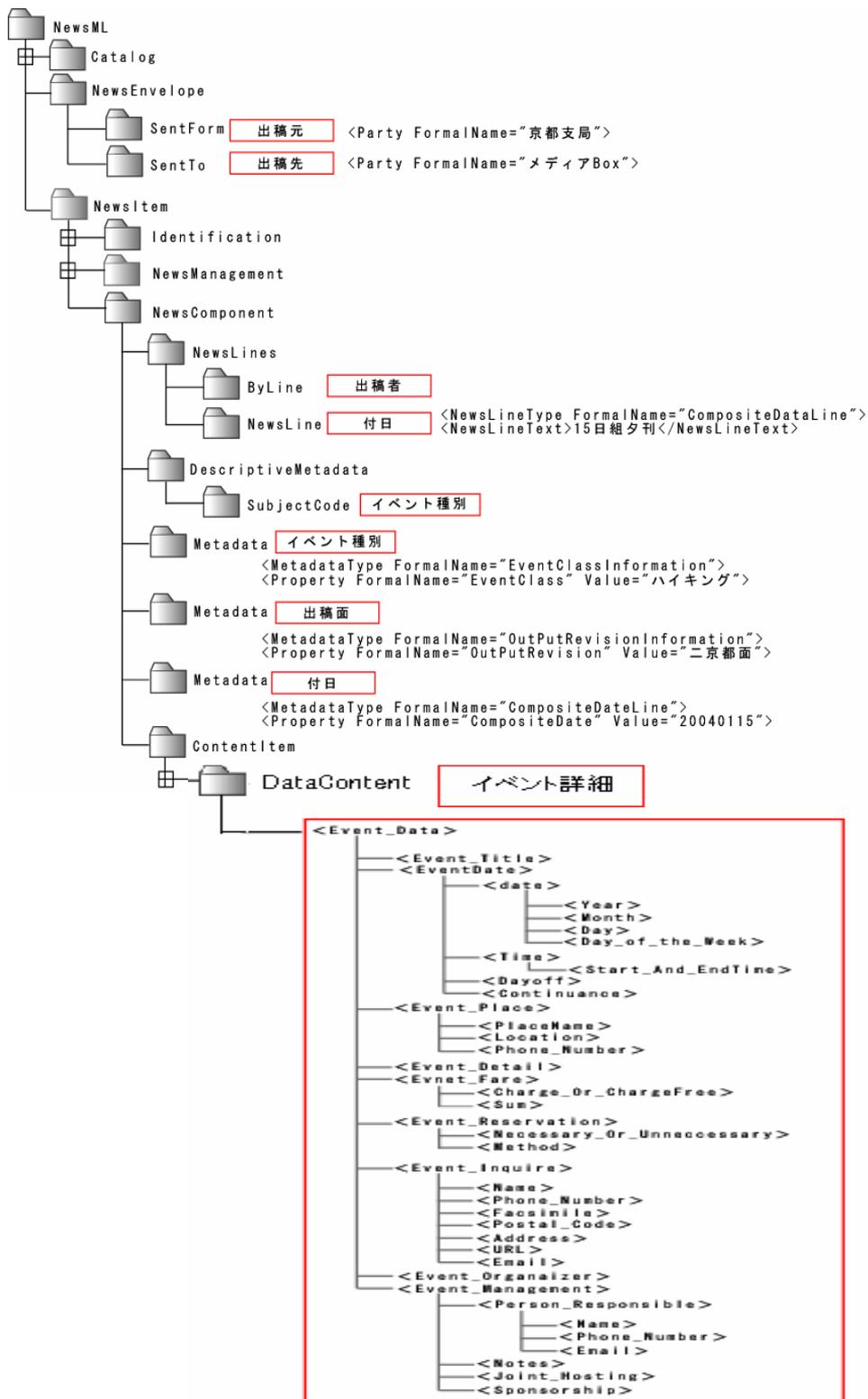


図 4: NewsM 化したイベント情報 (概略)

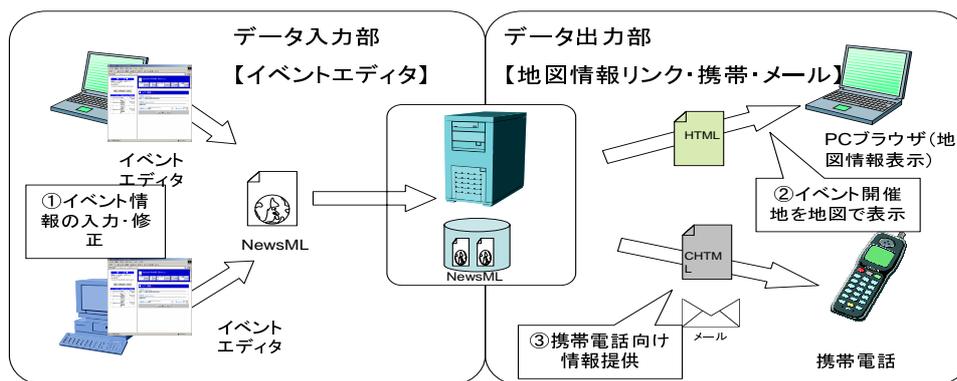


図5: イベントエディタと地図情報リンク・携帯電話への表示システム全体概要図

常、新聞記事でいうと記事そのもの（本文）が格納される場所である。

DataContent 内に、イベント名、開催日、場所、連絡先などの詳細イベントデータを定義した。ここで使用しているタグの要素名、構造は、ENTITY 宣言で定義している。

管理情報、イベント詳細全てを NewsML 化した結果が、図4である²。フォーマットの概略を図式した。

4 イベントエディタの構築

イベント情報の NewsML 化により、イベント情報を管理する共通フォーマットが規定できた。次に、実際に NewsML 化したイベント情報を入出力するツールの構築をおこなった。

自治体を含む多くの組織は、つぎつぎとソフトウェアを購入できるほど裕福ではない。システム開発費が高価であれば、たとえ、優れた XML 仕様を定めても使用されない。そこで、イベント情報を XML 形式で打ち込むためのツール、エディタを開発し、これをオープンソースとして、広く無料公開する予定である。

イベントエディタとは、WWW ブラウザ上からイベントデータを入出力できるエディタである。ユーザは、フォーム画面から、イベント名、開催日、開

² イベント情報の共通フォーマット化の世界的な状況としては、現在、IPTC(International Press Telecommunications Council :国際新聞電気通信評議会)にて「EventsML」としてフォーマットの規定作業が進行している。ただ、現段階での EventsML のフォーマットを見ると、例えば連続開催の日付処理や、何月何日まで開催しているがある特定の曜日のみ休み、後援者や共催、といったいった詳細規定がまだ確定していない。今後、EventsML との統合や我々が規定したフォーマットの提案なども視野にいれ研究を進める

催場所などを記入し、登録すれば自動的に NewsML 化されたイベントデータが出力される(図5)。

先の例のように、NewsML は非常に複雑な構造をしており、一般利用者がこのようなデータを手作業で記述することは困難である。また、いくつもの組織が同じようなシステムを構築することは無駄である。そこで、イベントエディタを、誰もが容易に NewsML 形式でのイベントデータを作成可能にするツールとして提供する。

エディタは、インターフェース部分を HTML と JavaScript で実装した。サーバ側でのファイル出力処理や、データの生成、修正には BayServer と Xi(eXtended It!以下 Xi) を使用している。Bayserver、Xi とは、オープンソースコミュニティの「横浜ベイキット」が公開している、XML アプリケーションサーバと開発言語である。Apache+Tomcat+Java+DOM といった環境よりも容易に XML アプリケーションの構築が可能である。

4.1 イベントエディタによるデータ入出力

ユーザは、まずイベントの管理情報を入力する。プロトタイプでは管理情報を「ヘッダ情報」としてしている。出稿者(記事作成者)、イベント種別、データの送付先などを入力する。「出稿面」「組日」とは、新聞紙面向けの管理情報である。ステータスは、「データは入力したが公開は待機する」「即時に公開する」といったイベント情報の公開・非公開管理情報である(図6)。

次に、イベント詳細情報を入力する。イベントのタイトル、開催日、会場名、会場住所、連絡先、内

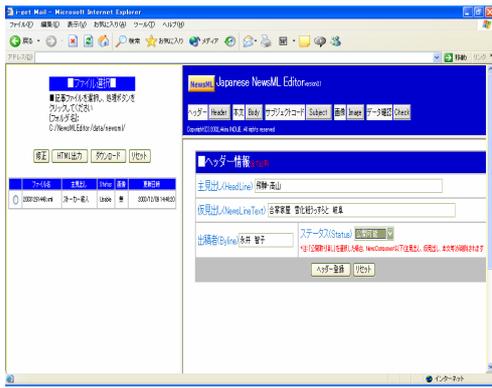


図 6: イベントエディタ: 管理情報入力画面

容、有料/無料などの情報を入力する。開催日付はカレンダー形式で表示され、開催日をチェックする。通常は 3 か月分表示されており、展覧会など年間を通じて開催されるイベントの日付は、「年間カレンダー」をクリックすると 1 年分カレンダーが表示される (図 7)。

すべての情報を入力し、確認画面を経由しデータを登録する。登録されたイベント情報はトップ画面左に一覧として表示される。

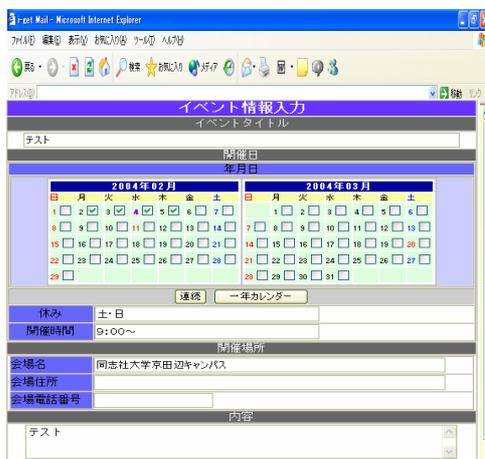


図 7: イベントエディタ: イベント詳細情報入力画面

出力された NewsML 形式のイベントデータは、ユーザの要望に応じて、Web 用 HTML や組版用データなどへコンバートすることができる。また、ある組織で作成したイベントデータを直接他の組織へ転送できる。NewsML として共通化されているために、情報の受け取り側では、必要な箇所の情報を取り出すだけで情報の再利用が行える。

5 携帯端末への配信と地図情報とのリンク

NewsML 化の効果を検証するために、2つのシステムを試作した (図 5 右側)。一つ目は、イベント情報の携帯端末への配信である。現在でも、インターネット接続機能を有する携帯電話向けのイベント情報は、各所から配信されている。ただ、その多くは、携帯端末用に、もとあるイベントデータとは別のデータを手作業または自動で生成している。

NewsML 化することで、ひとつのデータをスタイルシートで表示形式のみ変更すれば、PC 用、携帯電話用、組版用へ出力といったワンソース・マルチユース (One Source, Multi Use) が可能である。例えば、ある行政機関が作成したイベント情報を、自動的にある新聞社の携帯電話サイトへ公開する、といったことが実現可能である。

今回試作した携帯電話への配信は、イベント一覧や本日行われるイベント別に表示したり、日付やジャンルでイベントを検索する機能を持つ。また、メールアドレスを登録すれば、イベント情報がメールでプッシュされ、ユーザの利便性を高める (図 8)。

二つ目が、地図情報とのリンクである。開催場所を地図情報として視覚的に表示することで、直感的に位置を認識できる。また、周辺で開催されている他のイベントとの位置関係も把握しやすい。位置情報の登録には、1) イベント情報をマッピングする全範囲を表す広域地図を用意、2) 広域地図を適当に分割し、各矩形領域に値する狭域地図を用意、3) 分割された地図に識別番号を振る、4) 登録する会場・施設を含む地図を選択、地図上の座標を得る、5) データベースに地図番号、X 座標、Y 座標を登録する、という手順でおこなっている (図 9)(図 10)。

尚、今回構築したエディタ・携帯端末への出力・地図情報リンクはプロトタイプであり、本稿執筆時点ではすべてが稼動しているものではない。

6 まとめ

本研究では、ニュース情報の配信管理フォーマットである NewsML を利用し、イベント情報受発信・管理の共通フォーマットを規定した。さらに、NewsML 化したイベント情報を入出力するエディタを構築した。また、活用事例としてイベント情報の携帯端末

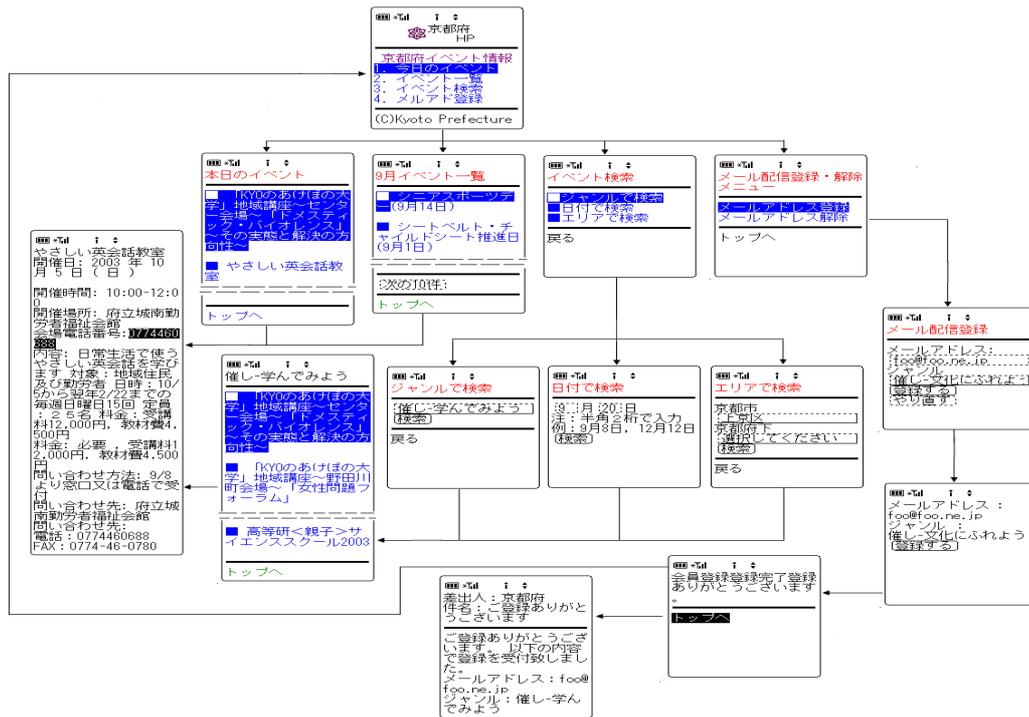


図 8: イベント情報の携帯電話への表示 (画面フロー図)

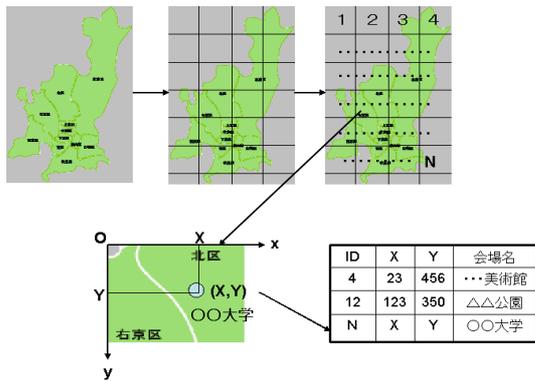


図 9: 座標の指定方法

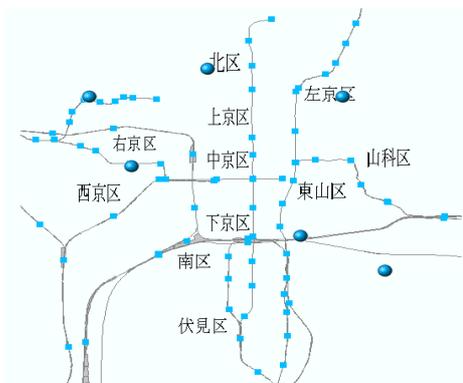


図 10: マッピングされたイベント情報

への出力、地図情報とのリンクをおこなうシステムを試作した。

NewsML は画像や動画、音声を管理する仕組みを有することから、将来的には、例えば、イベント開催地の画像や動画をリンクさせることも可能であろう。また、一般の市民からパソコンや携帯電話経由で情報を受け、直接、新聞や市の広報誌へ掲載するといった、情報の入力から出力までの流れを変える可能性が考えられる。

今後、京都地区で自治体・新聞社との実証実験を予定している。京都は日本の中でも上位に入る文化・観光都市としての特徴を持つ。各種の祭り、神社仏閣の一般公開等、京都在住者にとって価値があるだけでなく、京都を訪れるであろう旅行者にも参考になるイベント情報が非常に多い。そこでの実験を通じてシステム導入の効果を検証したい。

参考文献

[NewsMLv1.0] 社団法人日本新聞協会, 日本新聞協会 NewsML レベル 1 解説書, http://www.pressnet.or.jp/newsml/files/newsml1_0_2.pdf

[Mainichi] 毎日新聞社, 毎日インタラクティブ, <http://www.mainichi.co.jp/digital/newsml/ibm/index.html>