

NewsML 記事配信のためのメタ記事記述言語 KDML について

大園 忠親[‡] 新谷 虎松[‡] 青崎 保好[†][‡]名古屋工業大学大学院情報工学専攻 [†]共同通信社システム局

1 はじめに

膨大なニュース記事（以降、記事と記す）の配信には、記事そのものだけでなく、ニュースの価値などのメタ情報を適切に配信する必要がある。メタ情報は、記事の受け手にとって、記事をどのように扱うのかを決定するための指標となり、記事選択のためのコストを抑えつつ、重要なニュースを逃さないことを可能にする。

本論文におけるメタ記事とは、記事に関わるメタ情報を記述した文章である。本論文では、ニュースは事象を表し、ニュースを文章などで表現したものを記事と呼ぶ。

メタ記事を体系化し、より効果的にメタ記事を作成・利用するための研究が必要である。例えば、メタ記事を記述するための標準フォーマットが必要である。標準フォーマットを決めるためには、単に記述方法を定めるだけでは不十分であり、メタ記事とは何かを深く追求する必要がある。

本論文では、メタ記事について議論し、ニュース記事を効果的に扱うためのメタ記事を記述するための記述言語としてKDMLを提案する。

2 NewsMLによる記事配信

大まかに言って記事の配信は、ニュースを集め記事を執筆する通信社、通信社から記事を受け取り適切なメディアに変換し読者に届ける新聞社、そして読者の3者から構成される。

日本を代表する通信社として共同通信社を例とすると、通信社から新聞社への記事の配信には、NewsML[1,2]が用いられている。NewsMLは、2000年に国際新聞電気通信評議会（The International Press Telecommunications Council）が策定した、ニュースの制作、交換、ユーザでの利用のすべての過程で活用できるXMLベースの標準フォーマットである。

共同通信社は、1日に2万件のNewsML形式の記事を配信している。これら2万件の中には、記事だけでなく、記事の更新情報などニュース

とは、直接は無関係な情報も含まれている。

通信社から配信される記事のうち約5%程度が読者に届く。新聞社は人力のフィルタリングシステムであり、記事の価値を見定めている。記事の中でも、ニュースの価値を伝えるための特殊な記事として、メモと呼ばれる記事が存在する。新聞社は、メモも参考にして情報をフィルタリングしている。

メモは記事のメタ情報の一種である。現状では、NewsMLを用いてメモを配信しているが、NewsMLを離れてメタ情報について考察する。

3 メタ記事配信

メタ情報の種類を慎重に選択する必要がある。1日に何千件も発生するニュースに対して、適切にメタ情報を付与するのは相当なコストがかかる。とはいえ、メタ情報に情報がなすぎ過ぎて役に立たないというのも問題である。

3.1 ニュースの価値

ニュースの価値は個人の主観に強く依存するが、ニュースの相場観が鍛えられたプロの記者の判断した価値を配信することにも意味はある。記事の配信だけでなく、記事のアーカイブまで考えると、記者が与えた価値だけでなく、新聞社や読者が与えた価値や、計算機が計算した価値も与えることが可能である。

ニュースの価値は、誰にとってどのような価値があるのかを含む必要がある。ニュースのカテゴリもニュースの価値情報として考える。

ニュースの価値に関しては、社会的な影響などを考慮する必要もあるため人間のエキスパートですら、判断は困難である。ニュースの価値を絶対的に評価するのは困難であり、相対的な評価にならざるを得ない。現在は、1日の内での重要度の相対的な評価が行われている。結果として、ニュースの価値情報の再利用には、相対評価された価値をどのように扱うかという問題が残る。

ニュースの絶対的な価値基準を定めることは不可能であるため、様々な価値を記述できるような形式にするのが好ましい。多種類の価値を付けられるようにするために、価値の記述には、価値付けの手法を記述する必要がある。また、価値の情報を再利用するために、価値付けに対する自信などが表現できると好ましい。

On Meta Article Markup Language KDML for NewsML Delivery

[‡]OZONO Tadachika, Graduate School of Engineering, Nagoya Institute of Technology, ozono@nitech.ac.jp

[‡]SHINTANI Toramatsu, Graduate School of Engineering, Nagoya Institute of Technology, tora@nitech.ac.jp

[†]AOSAKI Yasuyoshi, KYODO NEWS, Systems Department, aosaki.yasuyoshi@kyodonews.jp

3.2 記事間の関係

ニュースは複数の記事によって読者に伝えられることが多い。例えば、ニュースの速報や経過を伝える記事、写真やイラストや音声などのマルチメディアデータ、関連する用語や状況を説明するための記事、事実だけでなく関係者の感想などを伝える記事などが挙げられる。さらに、訂正記事や関連するメモなども挙げられる。これらの記事間の関係を適切に表現することが重要であるが、全ての記事が揃ってから配信されるわけではなく、その都度配信するため、新聞社側で記事間の関係をトレースすることが困難になっている。

記事間の関係は、記事をノードとし、記事間の関係をリンクとするグラフ表現で表すことができる。

記事の速報性と記事配信の柔軟性を両立するために、将来作成される記事との関係を示すためのリンクを作成するという運用が行われている。複数の記者が将来の記事へのリンクをそれぞれ個別に作ってしまうような場合に問題が生じる。

3.3 記事のリビジョン管理

ニュースは突然発生し、その内容も時間の経過に連れて詳細化されていく。記事もニュースの全貌が明らかになるに従って更新される。また、記事の誤字や誤記などの訂正によっても記事が更新される。

新聞作成においては、最新の記事を送るだけでなく、更新部分を適切に連絡することが重要である。

4. メタ記事記述言語KDML

KDMLはXMLに基づくメタ記事記述言語である。KDMLでは、多様な価値観に基づく価値を内包するメタ記事の配信を目標とする。

4.1 ニュースの価値

KDMLにおけるニュースの価値は、次のように記述される。

```
<newsValue>
  <category></category>
  <metrics>評価手法</metrics>
  <valuator>評価者</valuator>
  <value>価値</value>
  <confidence>信頼度</confidence>
</newsValue>
```

newsValueタグは、categoryタグで示されたカテゴリにおけるニュースの価値を表す。metricsタグは評価手法とvalueタグ内の価値の記述方法を指定

する。valuatorタグは評価者を指定する。confidenceタグ内は、評価者の評価に対する自信を100%表記で記す。評価者が計算機システムの場合は、精度などを記述する。必要に応じて、評価理由や想定する読者についても記述すると良い。newsValueは、1記事に対して複数付加できる。これにより、特定の評価手法にこだわらず、様々な評価手法や評価者による価値を記述することが可能になる。

4.2 記事間の関係とリビジョン管理

記事間の関係で管理が困難なのは、執筆予定の記事との関係を管理することである。あるニュースに対して記者がそれぞれ独立して執筆を進めると、存在しない記事との関係を管理するのが困難になる。

ニュースの内容は時々刻々と変換し、しかもその変化は突然であるため、記事間の関係も柔軟に変更できることが要求される。KDMLでは、リビジョン管理も記事間の関係として取り入れる。これらの記事間の関係をグラフ表現で表し、そこでの最新のリンクの一覧をメタ情報として適切に配信することで、リンクのトレースが容易になる。

4.3 KDMLを利用した記事編集システム

KDMLを用いることによってニュースの価値情報や記事の現在の状態を効果的に把握可能になる。記事編集システムでKDMLを有効利用することで、記事の配信だけでなく、記事の作成段階からのニュースの動向の把握のための情報配信に利用できる。例えば、どの記者がどのようなニュースを扱っているのかがわかる。

また、様々な価値観を内包した記事アーカイブは、情報の再利用という観点から、新たな用途を生み出す可能性がある。例えば、紙の新聞にQRコードなどでデジタル化された価値情報を印刷しておくことで、記事内容だけでなく価値情報の利用も容易になる。

5. おわりに

効率的な記事配信のために必要なメタ記事について議論し、メタ記事の記述のための言語としてKDMLを提案した。今後の課題は、KDMLの作成と再利用のワークフローを考慮した記事作成支援システムの実現である。

参考文献

[1]The International Press Telecommunications Council (IPTC), "News Markup Language : NewsML", <http://www.newsml.org/pages/index.php>, Accessed on 2007/Jan/03

[2](社)日本新聞協会, "NewsML.jp", <http://www.newsml.jp/>, Accessed on 2007/Jan/03