

Title	会議情報の自動編集システムの作成
Author(s)	新井, 誠亮
Citation	
Issue Date	1999-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1218
Rights	
Description	Supervisor:佐藤 理史, 情報科学研究科, 修士

会議情報の自動編集システムの作成

新井 誠亮

北陸先端科学技術大学院大学 情報科学研究科

1999年2月15日

キーワード： 重要情報抽出, 情報検索, 情報統合, 情報表現, 情報の内部構造把握

現在、インターネットの利用は一般的に広く普及してきており、インターネットを通して利用できる情報は膨大になりつつある。インターネット上の情報の利点は情報の更新が素早く、リアルタイムにやりとりを行なうことすら可能で、他のメディアよりも早く最新の情報の取得が可能である。また、他のメディアはプロの手によって一度編集された情報をユーザが利用するという形態であったが、インターネットでは個人が容易に情報を公開することが可能であり、情報の多様性という面では、従来のメディアにはない可能性を持っていると考えられる。

このようにインターネット上の情報は従来のメディアにはない新しい可能性を持ったものであるが、以下のような問題点を持つ。

- 必要な情報の取得が困難
- 情報の形式が統一されていない

1つ目の問題点はインターネット上での情報の配置がばらばらであり、例えばある車に関する情報がどこにあるかという事について考えてみると、Web上でメーカーのホームページで紹介されている場合、個人のホームページで紹介されている場合、特定のニュースグループでその話題が出てくる場合などさまざまである。2つ目の問題点としてはインターネット上の情報は個人が発信したものが多く、従来のメディアのようにプロの手によって十分に編集されたものではないため、形式が統一されたものとなっていない。このため、従来のメディアより利用しづらいメディアとなっている。

インターネット上の情報の有効利用に関する研究が盛んに行なわれている。本研究ではインターネット上の会議情報という一つのカテゴリーに絞って研究を行なった。従来、ネットニュース fj.meetings 上の会議情報のダイジェスト自動生成に関する研究が行なわれ

ていた。電子ニュースを通じて多くの情報が流通し、多くの人々がその情報を利用している。しかし、電子ニュースを読むためのソフトウェアである現在のニュースリーダーは、その特徴や利用者の要求に合致した、適切な機能を提供しないため、読者にとっては必ずしも利用しやすい情報メディアであるとは言えない。そこで、従来研究では電子ニュースからの情報を利用しやすくするために、ニュース記事のダイジェストを作成することが試みられた。ダイジェスト作成のための中心的技術は、サマリーの自動抽出技術であり、会告記事に見られるスタイル上の特徴、言語表現パターンを利用する事により、実用に十分な精度でのサマリーの抽出を実現している。

上記の研究のようにダイジェストを作成する事により、統一された表記方法を取られていない電子ニュースの記事が、一様な形式に編集された。このことにより、電子ニュースの情報はユーザにとっては非常に利用し易く編集されたといえる。そこで本研究では、さらにニュース記事から抽出した情報を、ユーザの要求に応じて適切な形式で提供するためのシステムの作成を行なった。

電子ニュースに掲載される会告記事は、ダイジェスト作成に必要な全ての情報を含んでいない場合がある。このような場合、不足情報を他の情報源から得ることによって補完することが求められる。本研究では、これらの不足情報を、WWWのサーチエンジンを利用して見つけだし、それを補完する機能を実現する。本システムは以下に示すモジュールから構成される。

重要情報抽出モジュール ニュースグループ `fj.meetings` 上の会告記事を入力として受けとり、そこから会議タイトル、場所、日時、締切といった重要情報を抽出する。

内部構造把握モジュール 入力として抽出された会議の重要情報を受けとり、その情報の内部構造を把握する。

インターフェイス ユーザからの検索要求に対して、該当する会議情報を適切な形式に編集して提示する。

会議情報補完モジュール 会議情報データベース中の会議情報において住所情報が不完全である場合に、WWWを探索し、不足情報を補完する。

重要情報抽出モジュールでは、従来研究のものをベースに、よりテキストのスタイル情報を考慮して、改良を行なった。内部構造把握モジュールでは、抽出された会議情報に対して、情報の詳細度をもとに、その大きいレベルから小さいレベルへ段階的に分割を行ない、後述するインタフェース部分で利用しやすい内部構造を表現した形に変換を行なう。インターフェイスモジュールでは、ユーザの検索条件、および一般的な常識レベルから、ユーザにとって利用しやすいダイジェスト情報の形式を判断して、検索要求に該当する会議情報を適切な形式に編集して提示する。会議情報補完モジュールでは、会議情報データベース中の会議情報において住所情報が不完全である場合に、WWWを探索し不足情報を補完する。

これらの機能により、より品質の高い会議情報ダイジェストを生成することが実現できた。また、インターネット上の単一情報源からの情報が不十分な場合、WWW 等の他のリソースの情報の探索を行なうことによって、情報の補完が可能であることが示された。