

# 可視化プラットフォームの実現に向けた グラフ描画セレクトの基礎検討

Toward a Graph Presentation Selector for Common Visualization Platform

田中 和広  
Kazuhiro Tanaka

松下 光範  
Mitsunori Matsushita

関西大学総合情報学部  
Faculty of Informatics, Kansai University

## 1. はじめに

近年、ネットワーク上には様々な情報が大量に蓄積されるようになってきている。これらの情報を適切に利用することができれば、ユーザにとって様々な場面で有益なものとなるはずである。しかし、これらの蓄積された情報は、粒度が異なっている、断片化されている、情報間に重複や矛盾がある、などの特徴を持つため、これらの情報を適切に利用するには情報の取捨選択に関する様々なスキルが要求される。それゆえ、情報の利用スキルを十分に持たない人々が不利益を被るといった情報格差を招きかねない。そこで、そのような状況の改善を目指して情報編纂という枠組みが提唱されている [加藤 06]。これは多量雑多な情報を知的に編纂し、情報の理解やアクセスを支援するための枠組みである。

ネットワーク上に存在する雑多な情報には、テキスト等の言語情報や数値データや視覚情報等の非言語情報など、異なるモードの情報が含まれている。そのため、これらを利用者が自らの意図に沿って適切に利用できるように支援するには、言語情報を処理する技術や非言語情報を処理する個々の技術だけでなく、モードが異なる情報を纏め上げるための技術や、適切な形式に変換するための技術などそれらを統合的に扱う技術が必要になる。そのため、様々な研究者・研究機関が互いの技術を相補的かつ協調的に利用することでこのような課題に取り組むことが効果的である。情報編纂では、このような連携を円滑に進めるための基盤として、様々な要素技術を組み合わせるためのフレームワークの実現を目指している [松下 09]。我々はこのようなフレームワークのひとつとして可視化プラットフォームの実現を目指している。その足がかりとして複数の可視化インタフェースを切り替えて利用することのできる機構の実現を試みている。

本稿では、新聞記事テキストから取り出した時系列の統計情報を対象とし、それらを異なる可視化インタフェースで描き分けるためのグラフ描画セレクトについて述べる。まず、公開されているグラフ描画ライブラリ等の可視化ツールの調査を通してグラフ描画セレクトの実現に必要な機能や制約を検討し、それに基づいて実装案を提案する。

## 2. 情報編纂促進のためのフレームワーク

### 2.1 情報編纂の概要

情報編纂では、情報への理解やアクセスの支援を目的として、そのシステムや技術の構築を目指している。その取り組みの一例として、ガソリン価格や携帯電話加入者数等のいわゆる統計量（以下では時系列情報と呼ぶ）をテキストである新聞記事の集合から抽出し、その時系列情報を入力として用いて可視化するシステムの構築が考えられている [加藤 07]。この例に挙げたような一連の流れを実現するためには、少なくとも情報要約技術、情報抽出技術、情報可視化技術といった技術を用いることが必要となる。

情報要約技術 [奥村 04] については、大量のテキストからユーザにとって必要と思われるテキストを検索するテキスト検索技術、大量の情報から重要な情報のみを選択して提供し、要点の迅速な把握を支援するテキスト自動要約技術、ユーザの目的を反映した質問文に対して的確な答えを返す質問応答技術など、様々な観点から研究が進められている。

テキスト要約技術については、要約を行うにあたって考慮すべき要因に応じ、いくつかの分類がなされている [難波 02]。例えば、要約の対象が単一か複数かによって、単一文書要約と複数文書要約とに分けられる。情報編纂では、テキスト情報を含む多様かつ雑多な情報を対象としているため、複数文書要約の技術が必要になる。難波らは動向情報の抽出を一種の複数文書要約であると考え、あるトピックに関する文書集合について文書文間

関係解析を行い、複数の文書から動向情報を自動的に抽出し、グラフ化する手法を提案している [難波 05]。

情報抽出技術は、上述した情報要約技術の一つとしても捉えることもあり、要約と抽出を同様にして扱われる場合もあるが、ここでは文献 [関根 04] の「(情報抽出技術は) 特定のテーマの情報を非構造的な文書の中から抽出し、構造化されたかたちで提示する技術である」という定義に則り、情報要約技術とは切り分けて考える。情報編纂における情報抽出技術としては、複数の動向情報を編纂するために必要となる統計量をテキストから抽出するための研究が該当する。例えば、森らは動向情報の要約と可視化を背景に、新聞記事からの統計量の抽出を目的とし、統計の調査方法と統計量名を定義することで、統計量名の抽出を検討している [森 08]。

情報可視化技術は、情報要約技術や情報抽出技術によって取捨選択されまとめられた情報を分かりやすく提示することによって、ユーザに理解を促す技術である。これまで情報編纂を目的として、統計データを元にした可視化手法がいくつか提案されている。松下らは、時系列的にグラフによる可視化を行うシステムとして、Elucignage を提案している [松下 08]。Elucignage は、描画されたグラフからその元となる新聞記事情報へのアクセスを簡便にすることを目的に、時系列数値情報を折れ線グラフと言語情報の内容に応じたアイコンを用いて可視化している。時系列的に情報を提示するだけでなく、空間的に情報を提示するシステムとして、高間らは、地震に関する動向情報の可視化を行うインタラクティブ地震動向情報システムを提案している [高間 06]。このシステムでは、折れ線グラフと日本地図を併用することで、時空間的に地震情報を可視化することを試みている。また、太田らは、統計量に関する数値情報を用いて、その数値的特徴を元に適切なグラフタイプ（棒グラフ、折れ線グラフなど）を作成する手法について提案している [太田 07]。

これらの異なる技術の連携を図るためには、各技術を協調的・競争的に用いるためのフレームワークが必要になる。また同じ技術の範疇であっても、各々の研究者が独自の環境や手法でシステムを構築している場合は、複数のシステム間の比較が難しくなる。例えば、同じデータを用いて同様の挙動を行わせようとした場合、システムの固有の仕様に依じた調整が必要になるため、比較の手間が増大することが懸念される。また、情報可視化システムの研究においては、その視覚的な特性により、同じデータを扱う場合であっても、見た目や操作性がその判断に影響する懸念があることが指摘されている [Pillat 05]。このような観点からも、フレームワークが必要になる。

我々はこのような問題に対処するため、可視化プラットフォームの実現に向けての取り組みを行っている。可視化プラットフォームが構築されることにより、(1) 見

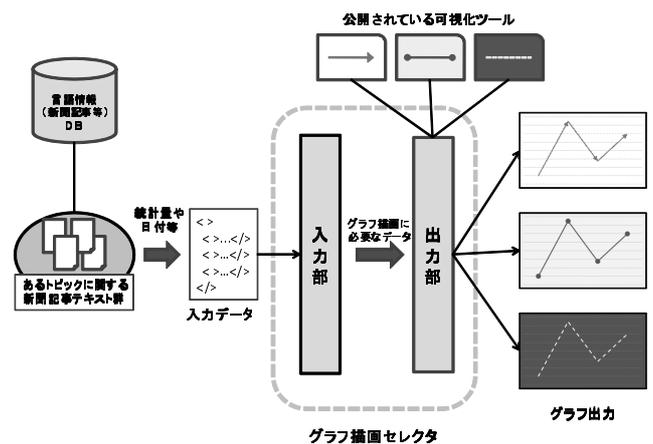


図 1: グラフ描画セレクタの概念図

た目や操作に関する統一性を図り、異なるシステム同士の比較が容易になる、(2) 個々の研究者が各々の環境や独自のプラットフォームで実装したシステムを移植する際の負担が軽減したり、自らの研究の位置付けや役割を認識したりする、という二つの効果が期待できる。

## 2.2 可視化プラットフォーム

2.1 節で述べたように、可視化プラットフォームの主眼は、複数の研究機関で開発された様々な可視化システムを評価するための比較の容易化と情報抽出や要約といった要素技術の連携を可能にすることである。

可視化プラットフォームの実現を考えていく上で、データを可視化するための手法には様々なものがあるが、本研究では情報編纂の現在の主な取り組み [加藤 10] を鑑みて時系列情報の可視化に焦点を絞り、複数の可視化インタフェースに切り替えて利用するためのプラットフォームの実現を目指している。その中でも特に複数のグラフ可視化ツールに切り替えて利用するための機構として、グラフ描画セレクタについて検討している。

グラフ描画セレクタの概念図を図 1 に示す。グラフ描画セレクタでは、あるトピックに基づいた複数の新聞記事テキストから統計量や日付などの情報を抽出し、統計グラフの描画に必要な情報を一つのデータとして纏めた入力データを用いる。そして、それを元に、様々なグラフ可視化ツールを用いて、それらを切り替えて描画できるようにする。このグラフ描画セレクタが統計データを元にしたグラフ描画のフレームワークとして機能することにより、可視化のために必要となる統計量の抽出やその元となる新聞記事などのテキストの要約において、統一的なデータセットの準備が可能となることを期待している。また、可視化技術においては、用意された統計量データの形式に基づいて、可視化を行う方法について考えられる環境づくりを目指す。このように各々の研究者・研究機関が一つのフレームワークに則って研究を進めていくことにより、異なる可視化ツールを用いるこ

とが可能となり、異なる研究機関が相補的・協調的に連携することが期待される。

### 3. グラフ描画セレクタ

#### 3.1 グラフ可視化ツール

今回、様々な異なるグラフ可視化ツールの機能や制約について検討を行うため、WEB上に公開されているグラフ描画ライブラリなどの可視化ツールの調査を行った。代表的なシステムの概要を表1に示す。

これらの可視化ツールは、使用する言語や技術、グラフを描画するための仕様、グラフの外観などが各々で異なっている。例えば、グラフを描画するための技術としては、JavaScript(JS)、PHP、Silverlight、ActionScript(AS)、Ruby、JAVA、WebAPI等が用いられている。また、入力形式においても、コードに直接数値を埋め込むもの、XMLファイルやEXCELファイルに数値を記述し、それを読み込ませるもの等があり、出力形式に関しても、各々の技術を用いるためにプラグインやプレーヤーが必要となるもの、画像として描きだすものがある。グラフの操作性や外観についても、ズーム機能を有するものや、グラフの時間軸を年月日に合わせて自動調整してくれるもの等がある。

#### 3.2 実装

我々は、統計データを元に様々なグラフに描き分けるためのシステムとして、グラフ描画セレクタのプロトタイプとなるシステムの実装を行った。実装を行うに際し、グラフを描くための統計量のデータと可視化ツールが必要となった。可視化ツールに関しては、先ほど述べたWEB上で公開されているものの中から、amCharts、gvChart、jqPlot、Visifire、XML/SWF Chartsを用いた。これらのツールの選択基準は、異なる技術を用いるツールを入れること、入力が異なる形式を用いる物を使うこと、グラフの外観が特徴的なものを入れること、である。これらの入力となる統計データに関しては、

統計量を抽出する元となる情報として、今回は「動向情報の要約と可視化に関するワークショップ (MuST)」[加藤 04]によって配布されている研究用データセットであるタグ付きコーパス (以降、MuST コーパスと呼ぶ) をもとに、人手で値を抽出してXML化したものを用いた。MuST コーパスは、1998年から2001年の毎日新聞の記事を元に、可視化に必要な要素に対して人手でタグを付与したものである。このMuST コーパスのタグ内容を元にして、日付や統計量の情報を読み取り、必要な情報の抽出を行い、入力データとして纏めた。また、統計量を入力としてグラフを描画させる際に、各々の可視化ツールのコードを書き換える必要があった。そのため、入力データから日付と値を対応付けて抜き出し、各ツールの数値を読み込ませる部分の書き換えも人手で行った。

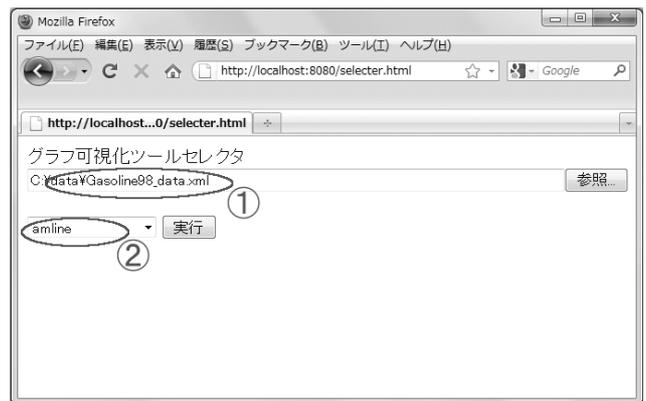


図 2: グラフ描画セレクタの試作インタフェース

実装したシステムのスクリーンショットを図2に示す。図中①で入力データを選択し、図中②で可視化ツールを選択する。本システムを用いたグラフ描画の例を図3に示す。図3中の(A)と(C)はjqPlot、(B)と(D)はAmChartsを用いて描画したものである。図3中の(A)と(B)はMuSTコーパスの中から「レギュラーガソリンの全国平均店頭価格」を、(B)と(D)は「携帯電話とPHSの合計加入台数」を各々描画したものである。今回の実装は、Rubyを用いて行った。また、グラフの描画に関してはブラウザでの表示を想定したツールが多かったため、ブラウザベースで実装している。

#### 3.3 考察

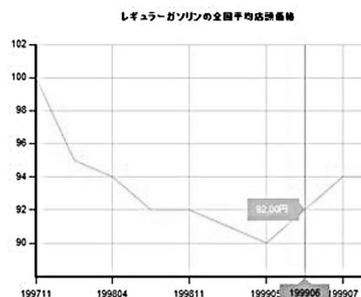
本節では、グラフ可視化ツールの調査とグラフ描画セレクタのプロトタイプシステムの実装を通じて、フレームワークとしての役割と今後の実装案について考察していく。

実装で用いたMuSTコーパスにおける内容へのタグ付けは、プレーンテキストに対して、そこが統計量に関する言及であることを付与する意味処理や文脈処理を行った結果に相当する。また、入力データを作成する際に、複数のMuSTコーパスから必要となる情報を抜き出し、それを用いて一つのデータとして纏めたものである。本実装におけるMuSTコーパスを元にした入力データの作成は、複数文書要約と情報抽出を行った結果であると位置付けられる。また、複数のグラフ可視化ツールは、情報可視化技術の研究の結果として見做すことができる。すなわち、本稿で述べたグラフ描画セレクタのプロトタイプシステムは、新聞記事の統計データを元にしたグラフ描画における、情報要約技術、情報抽出技術、情報可視化技術のフレームワークとなるシステムに相当すると考えられる。

まず、グラフ出力部について考える。今回用いた可視化ツールでは、抽出した統計量をコードに直接埋め込む形やXMLファイルで読み込む形というように各々で入力形式が異なっていた。そのため、各々の可視化ツ



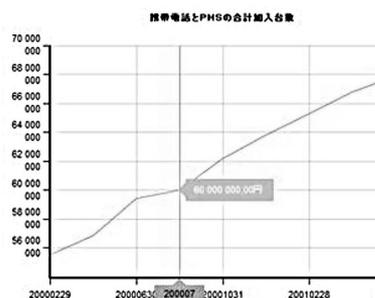
(A)



(B)



(C)



(D)

図 3: グラフの描画例

ルで入力データを用いる際に、入力データがその可視化ツールの入力形式に合致するようにコードの中身の書き換えを行わなければならないという問題が生じた。そのため、複数の可視化ツールで一様に入力データを扱えるようにするには、各々の可視化ツールの統計量の読み込みの部分で、統一した入力形式を用いることが望ましい。また、可視化ツールに用いている技術や操作方法、外観に関しても各々の可視化ツールで異なっていた。使用している技術については、それに必要なプラグインをインストールしている必要があるなど、環境を整えなければ使用できないという場合もあるが、これに関しては各可視化ツールの見せ方の問題なので、特に統一を強制せずに独自で開発を進めてもらえれば良いと考えている。ただしグラフの外観については、グラフ同士の比較が困難になるため、ある程度の基準を設けることが望ましい。

今回用いたグラフ描画ツールでは、jqPlot のみが時間間隔に合わせて、時間軸を自動調整する機能を持っていた。そのため、時間間隔に関係なく横軸に値を描くツールとでは、見た目が大幅に異なっていたため、グラフ同士の比較評価がしづらくなっていた。そこで、可視化ツールに時間間隔に合わせて自動調整する機能を導入するか、プラットフォーム側でその処理を行うことが求められる。各々の可視化ツールでその機能に対処するのであれば問題は無いが、その機能を持たないシステムであっても対

応できるように可視化プラットフォームに時間軸処理の機能を持たせる方が適当であろう。

入力データについて考える。入力データに相当する出力を生成するモジュールは、どの可視化ツールでも用いることができるような形式で作成する必要がある。そのため、標準化したタグ付け規則を整備し、それに基づいて入力データの作成を研究していくことが望ましい。また、グラフを描画するのに必要な情報を抽出する際に、単なる統計量のみを抽出するだけに限らず、定性表現を抜き出すことが求められる場合もある [松下 06] など、各々の研究機関によって、必要であると判断される情報が異なることが想定される。そのため、タグ付けのフォーマットは統一するが、抽出すべき情報の判断に関しては個々の研究機関が決めればよいと考えている。

これらの点を踏まえた上で、今後は実際に各研究者や研究機関に利用してもらったり、我々の今後の研究を通じて検討していくことで、さらなる検討を進めていきたい。

#### 4. おわりに

本稿では、可視化システムの評価比較の容易化や、個々の研究者たちが自分たちの研究の位置付けを認識したり、各々の実装を移植したりすることを容易にするためのフレームワークである可視化プラットフォームの実現に向

け、その一環として複数のグラフ描画ツールで描き分けることのできるグラフ描画セレクタを提案し、それについての検討を行った。

今後は、今回の基礎検討を踏まえ、グラフ描画セレクタに必要な要件について精査していき、可視化プラットフォームの実現を目指していく。

## 謝辞

本研究の遂行にあたり、文部科学省科学研究費（課題番号：22300048）の助成を受けた。記して謝意を表す。

## 参考文献

- [加藤 04] 加藤 恒昭, 松下 光範, 平尾 努: 動向情報の要約と可視化に関するワークショップの提案, 情報処理学会研究報告, Vol. 2004, No. 108, pp. 89–94 (2004)
- [加藤 06] 加藤 恒昭, 松下 光範: 情報編纂 (Information Compilation) の基盤技術, 2006 年度人工知能学会論文集, pp. 1D3–02 (2006)
- [加藤 07] 加藤 恒昭, 松下 光範: 時系列情報の抽出と可視化に基づく情報アクセスのためのマルチモーダルインタフェース –情報編纂の基盤技術に向けて–, 人工知能学会論文誌, Vol. 22, No. 5, pp. 553–562 (2007)
- [加藤 10] 加藤 恒昭, 松下 光範, 神門 典子: 時系列情報の値と変化に関する言語表現コーパスの構築: 動向情報の情報編纂に向けて, 人工知能学会論文誌, Vol. 25, No. 5, pp. 637–650 (2010)
- [松下 06] 松下 光範, 加藤 恒昭: 数値情報の補填とグラフ概形の示唆による複数文書からの統計グラフ生成, 知能と情報, Vol. 18, No. 5, pp. 721–734 (2006)
- [松下 08] 松下 光範, 加藤 恒昭: 言語情報と数値情報の相補的利用を目指した可視化手法, 人工知能学会全国大会論文集, Vol. 21, pp. 3H8–3 (2008)
- [松下 09] 松下 光範, 加藤 恒昭: 情報編纂研究促進のための試み, 人工知能学会論文誌, Vol. 24, No. 2, pp. 272–283 (2009)
- [森 08] 森 辰則, 藤岡 篤史, 村田 一郎: 動向情報編纂のためのテキストからの統計量表現の自動抽出, 人工知能学会論文誌, Vol. 23, No. 5, pp. 310–318 (2008)
- [難波 02] 難波 英嗣, 奥村 学: ここまで来たテキスト自動要約 (特集: テキスト自動要約: 知的活動支援の基本技術として), 情報処理, Vol. 43, No. 12, pp. 1287–1294 (2002)
- [難波 05] 難波 英嗣, 国政 美伸, 福島 志穂, 相沢 輝昭, 奥村 学: 文書横断文間関係を考慮した動向情報の抽出と可視化, 情報処理学会研究報告, Vol. 2005, No. 73, pp. 67–74 (2005)
- [太田 07] 太田 彰, 福本 淳一: 文書中の数値的特徴を用いた情報可視化, 動向情報の要約と可視化に関するワークショップ第二回成果進捗報告会予稿集, pp. 13–16 (2007)
- [奥村 04] 奥村 学: テキスト自動要約 (特集: 自然言語による情報アクセス技術), 情報処理, Vol. 45, No. 6, pp. 574–579 (2004)
- [Pillat 05] Pillat, R. M., Valiati, E. R. A., and Fraitas, C. M. D. S.: Experimental study on evaluation of multidimensional information visualization techniques, in *Proc. CLIHC2005*, pp. 20–30 (2005)
- [関根 04] 関根 聡: 情報抽出: 情報を整理して提示する (特集: 自然言語による情報アクセス技術), 情報処理, Vol. 45, No. 6, pp. 563–568 (2004)
- [高間 06] 高間 康史, 山田 隆志: タグ付きコーパスを用いた地震記事からの動向情報抽出・可視化システム, 知能と情報, Vol. 18, No. 5, pp. 711–720 (2006)

表 1: 主なグラフ描画ツールの概要

名前	使用技術	出力	入力形式
AmCharts ( <a href="http://www.amcharts.com/">http://www.amcharts.com/</a> )	JS+AS	Flash Player	XML or CSV
Artichow ( <a href="http://www.artichow.org/">http://www.artichow.org/</a> )	PHP	画像	コード埋め込み
Bluff ( <a href="http://bluff.jcoglan.com/">http://bluff.jcoglan.com/</a> )	JS	HTML5, JS	コード埋め込み
Fly Chart ( <a href="http://flycharts.net/">http://flycharts.net/</a> )	JS+AS	Flash Player	XML
FusionCharts Free ( <a href="http://www.fusioncharts.com/free/">http://www.fusioncharts.com/free/</a> )	JS+AS	Flash Player	XML
Google Chart API ( <a href="http://code.google.com/intl/ja/apis/chart/">http://code.google.com/intl/ja/apis/chart/</a> )	API	画像	コード埋め込み
Gruff Graphs for Ruby ( <a href="http://nubyonrails.com/pages/gruff">http://nubyonrails.com/pages/gruff</a> )	ruby	画像	コード埋め込み
gvCharts ( <a href="http://www.ivellios.toron.pl/technikalia/2010/06/22/gvchart-plugin-jquery-with-google-charts/">http://www.ivellios.toron.pl/technikalia/2010/06/22/gvchart-plugin-jquery-with-google-charts/</a> )	JS	HTML5, JS	コード埋め込み
HTML5.jp ( <a href="http://www.html5.jp/">http://www.html5.jp/</a> )	JS	HTML5, JS	コード埋め込み
Jfree Chart ( <a href="http://www.jfree.org/jfreechart/">http://www.jfree.org/jfreechart/</a> )	JAVA	画像	コード埋め込み
jqPlot ( <a href="http://www.jqplot.com/">http://www.jqplot.com/</a> )	JS	HTML5, JS	コード埋め込み
JS Chart ( <a href="http://www.jschart.jp/">http://www.jschart.jp/</a> )	API	画像	コード埋め込み
JS Charts ( <a href="http://www.jscharts.com/">http://www.jscharts.com/</a> )	JS	HTML5, JS	XML
Libchart ( <a href="http://naku.dohcrew.com/libchart/pages/introduction/">http://naku.dohcrew.com/libchart/pages/introduction/</a> )	PHP	画像	コード埋め込み
pChart ( <a href="http://pchart.sourceforge.net/">http://pchart.sourceforge.net/</a> )	PHP	画像	CSV
PURE CSS LINE GRAPH ( <a href="http://cssglobe.com/post/4175/pure-css-line-graph">http://cssglobe.com/post/4175/pure-css-line-graph</a> )	CSS	HTML, JS, CSS	コード埋め込み
Raphaël ( <a href="http://dmitrybaranovskiy.github.io/raphael/">http://dmitrybaranovskiy.github.io/raphael/</a> )	JS	HTML, JS	コード埋め込み
Visifire ( <a href="http://www.visifire.com/">http://www.visifire.com/</a> )	Silverlight	Silverlight	コード埋め込み
XML/SWF Charts ( <a href="http://www.maani.us/xml_charts/">http://www.maani.us/xml_charts/</a> )	API	Flash Player	XML