

平成 28 年度 修士論文

インフォグラフィックス制作支援システムの
開発と評価

Designing and Evaluating
an Infographics Authoring Support System

提出日 平成 29 年 2 月 3 日
指導教官 矢谷浩司 准教授
所属 東京大学大学院
工学系研究科電気系工学専攻
修士二年
学籍番号 37-156438
岩下知美

要旨

大量の情報源の中から意味のある情報を素早く取り出すために、情報可視化が重要視されている。情報可視化の一例として、絵や文章で情報を表すインフォグラフィックスがある。しかしインフォグラフィックスの制作にはデザインやツールに関する知識が必要であり初心者が容易に行えるものではない。既存のインフォグラフィックスを調査した結果、視覚的な要素としてピクトグラムが重要であることがわかった。この発見に基づき、本研究ではまずインフォグラフィックスの分類分析を行い、どのような視覚的要素が利用されているかを調査した。この結果を踏まえ、インフォグラフィックスの制作を支援するインタラクティブなシステムを開発した。開発したシステム *grafie* はインターネット上の画像からピクトグラムを機械学習によって抽出する。さらに、*grafie* は選択されたピクトグラムを使ったグラフの作成も可能としている。*grafie* と既存のインフォグラフィックス制作システムとの比較を行うユーザテストを行った。結果、パフォーマンスやユーザビリティに関しては数値上の違いは存在しなかった。しかし、実験参加者は自身が制作したインフォグラフィックスのうち既存システムよりも *grafie* で制作したものを好む傾向があった。

Information visualization becomes important to extract useful information out of data resources. One example of common information visualization is infographics, a method to present information through visual graphics with textual descriptions. However, authorizing infographics requires the knowledge of design and tools, which is a large burden for novice users. My analysis on existing infographics revealed that pictograms is an important visual component. Based on this finding, I develop an interactive system for supporting infographics authoring. My system, called *grafie*, extracts pictograms from online graphics with machine learning. In addition, *grafie* also supports creating graphs using the selected pictograms. I conducted a comparative user study against an existing infographics authoring system. The results did not find significant differences in quantitative performance and usability metrics. However, more participants tended to express better user experience in infographics authoring with *grafie* than the reference system.

目次

第 1 章	序論	1
1.1	背景	1
1.2	貢献	3
第 2 章	先行研究と既存システム	4
2.1	先行研究	4
2.2	制作ツール	10
第 3 章	システムの実装	14
3.1	ピクトグラムを取得しグラフを作成する機能の詳細	14
3.2	その他の機能	16
3.3	実装	17
3.4	使用例	17
第 4 章	システムの製作	21
4.1	infographics の調査	21
4.2	機械学習を用いたピクトグラムの判定の予備実験	24
4.3	ピクトグラム判定手法の本実験	28
第 5 章	ユーザ評価実験	33
5.1	実験参加者	33
5.2	比較対象	33
5.3	実験の流れ	34
5.4	結果	35
5.5	考察	40
第 6 章	おわりに	44
	謝辞	45
	参考文献	46
	研究業績一覧	49

目次

2.1	ストーリーテリングのプロセス	10
2.2	ピクトグラムを用いたグラフの例	13
3.1	実装した機能のプロセス図	15
3.2	本システムで作成することの出来るグラフの例	16
3.3	grafie 使用例 1	18
3.4	grafie 使用例 2	19
3.5	grafie 使用例 3	20
4.1	infographics の 7 つの分類	22
4.2	学習データとして用いる画像に行った処理を表した図	25
4.3	正規化された画素の標準偏差の分布	26
4.4	正規化されたコーナー数の分布	26
4.5	コーナー数と画素の標準偏差との散布図	27
4.6	ピクトグラムである画像の例	31
5.1	ユーザに例として提示した作例	35
5.2	各システムの SUS の平均と標準誤差	37
5.3	各システムのそれぞれのタスクにおけるクリック回数の平均と標準誤差	37
5.4	各システムのそれぞれのタスクにおける所要時間の平均と標準誤差	38
5.5	piktochart での満足度が高かった P6 の作例	40
5.6	grafie での満足度が高かった P7 の作例	41
5.7	P4 のタスク 2 の piktochart で制作された infographics	41
5.8	画像を複数枚使用した infographic の例	42
5.9	タスク 2 におけるクリック回数の差による作例の違い	43

表目次

4.1	infographics の分類結果	23
4.2	infographics の各分類のピクトグラムの有無	23
4.3	ピクトグラム認識テストの混同行列	27
4.4	infographics で頻出している単語	29
4.5	頻出単語を用いて検索した結果得られた画像の枚数	29
4.6	AMT による分類の結果	31
4.7	Linear SVC の C 値のグリッドサーチの結果	31
4.8	10-fold cross validation を行った際の混同行列	32
5.1	それぞれのシステムにおいての各タスクの所要時間 (秒)	38

第 1 章

序論

1.1 背景

ICT 技術の進歩により様々なデータを大量に収集することが容易となった。それにより、収集したデータをどうやって活用するかが情報産業以外の幅広い業界でも重要な問題となってきている。しかし、そのようなデータはサイズが非常に大きくなることもありそのままでは理解し活用することが難しいが、流動的である側面もあるため理解しやすい形式に加工するのに時間がかかると意味の無いデータになってしまう。そこで、素早く適切にビッグデータを活用できる手段として情報可視化が重要視されている [1, 2, 3]。情報可視化によってデータに隠された様々な関係性を明らかにすることによってデータ活用をより適切に行うことができるのである。

データ可視化の表現方法として近年注目を集めているのが **infographics** である [4]。**infographics** とはテキストだけではなくグラフや図などを用いてデータや知識をわかりやすく提示するメディアであり、情報可視化においてよく用いられる表現方法である。**infographics** 自体の歴史は古く、棒グラフやヒストグラムを用いてデータを表すという現在用いられている **infographics** に近いものは 18 世紀に **William Playfair** が使い始めたと言われている。近年ではコンピュータの普及、および画像を加工するソフトウェアの発達によりデータに基づいた画像を描画することが容易になったおかげで **infographics** はより一般的なものとなってきている*1。非言語の要素が多いため、適切に表現されていれば直感的に内容を理解することができる一方、文章も含まれているので内容を詳しく説明することも出来るという利点があり、題材にあまり詳しくない人にも興味を持ってもらいやすく、かつ理解しやすくデータを伝えることができる。日本でも平成 23 年に **infographics** を用いて専門家や国が持つデータとクリエイターの「伝える」力を結びつけ

*1 <https://visual.ly/m/history-of-infographics/> (閲覧日:2017 年 1 月 29 日)

るというコンセプトのもと「ツタグラ」プロジェクトの運用を経済産業省が開始する^{*2}など、infographics は今後さらに身近なものになっていくことが予想される。

しかし、infographics は企業の広報や新聞社が統計を元にしたニュースを報道する際の一手法として使用されることはあるものの、一般的なユーザが infographics を制作し活用できている例はまだ少ない。その理由の一つとして、一般的なユーザにとっては infographics は容易に制作することができないという問題がある。infographics は画像を中心としたデザインが特徴的であるが、データをわかりやすく理解できるようなデザインを作るには画像の配置方法や色の使い方、またソフトの扱い方などの知識が必要となり、専門的な教育や経験がない個人が作成するのは難しい。そういった問題を解決するために infographics の制作が出来るインターネット上のサイトもいくつか存在しており、いずれも用意されているテンプレートに沿う限りは上質なデザインの infographics を作成できるが、一方でオリジナルのデザインを一から制作する上ではやはりデザインに関する知識を必要とする。もう一つの理由として、infographics の作成には手間と時間がかかることが挙げられる。infographics の制作プロセスを表現した infographic^{*3}によると、infographics の作成にはデザイナーと編集者とデータ解析者が以下の作業を行う必要がある。

1. トピックの選定
2. 調査
3. データ収集および集計
4. データの解析
5. データからのストーリーの発見
6. 完成形のアイディアのスケッチ
7. 編集作業
8. デザイン
9. 検証

特にデザイナーはほぼすべての過程に携わっており、そのデザインに関する知見は理解しやすい infographics を制作する上で必要とされているが、一般のユーザが infographics を制作する際に同程度の知識を用いることはまず不可能である。

そこで、本研究では infographics の制作を容易とするシステム「grafie」の開発を行った。本稿では、第2章にて先行研究や既存の infographics 作成サービスについて述べた後、既存サービスがどのような問題点を抱えているのかを述べる。次に、第4章にて制作した提案システムの機能紹介および実際の使用方法を説明する。続く第3章にて既存の

^{*2} http://www.meti.go.jp/policy/mono_info_service/mono/creative/111031_release_infographics.pdf (閲覧日:2016年10月1日)

^{*3} <http://visual.ly/infographics-process> (閲覧日:2016年11月20日)

infographicsにはどのような種類があるのか分類を行った上で infographics ではピクトグラムと呼ばれる画像様式が重要な要素であることを示し、インターネット上の画像の検索結果からピクトグラムである画像だけを抽出する機械学習の実装について説明する。第5章にて開発したシステムへのユーザテストを通じて評価を行った後、第6章で課題点や今後の展望を述べる。

1.2 貢献

- 既存の infographics を形状によって分類し、可視化するデータの種類によって用いられる形状に差があることを発見した。
- 形状それぞれのピクトグラムの使用率を調査したところ、平均で6割程度の使用率であることに對し最も多い形状では7割を超える使用率であることが示された。
- インターネット上の画像からピクトグラムのみを抽出する機械学習を実装し、それを組み込んだ infographics 制作支援システムを構築した。
- 構築したシステムのユーザ評価実験では既存システムよりもより制作者自身の満足度が高い infographics の制作が行えることが示された。

第 2 章

先行研究と既存システム

2.1 先行研究

2.1.1 情報可視化

infographics は情報可視化の一手段である。情報可視化の分野においては多次元グラフや色分け、幾何的な変形など様々な手段で可視化を行ってきており [5]、またそれがデータの整理や理解において有用であることが研究により示されてきた。

情報可視化の分類

Chi らは Data State Model を用いた情報可視化の分類に挑戦した [6]。それによると、加工前のデータから情報可視化を作成する際に行われる変換工程は以下の 5 つである。

1. Data Transformation
2. Analytical Abstraction
3. Visualization Transformation
4. Visualization Abstraction
5. Visual Mapping Transformation

情報可視化を行うシステムでは、以上の 5 つの工程をデータの種類や最終的な可視化の様態に合わせて適宜適切な方法で行っている。その際、すでに次の工程で使うことの出来るフォーマットになっているために飛ばされる工程もある。また、この研究では可視化に関する研究は 10 種類に分けられている。

- 科学的な事柄に関する可視化
- 地理的情報に基づいた可視化
- 二次元での可視化

- 多次元プロットによる可視化
- 情報を空間的に表示する可視化
- 木構造での可視化
- ネットワーク構造での可視化
- 文章による可視化
- ウェブサイトの可視化
- スプレッドシートの可視化

螺旋によるデータ可視化

時間のみを可視化する場合は直線を用いることが多いが、時間だけでなく週や月、年をも同時に示したい場合は直線による可視化ではわかりやすく表すことが出来ない。このように連続的かつ周期的な軸を持つデータの可視化に対して Carlis らが提案したのが螺旋を用いたデータ可視化である [7]。 $r = \alpha\theta$ で示されるアルキメデスの螺旋を用いて 2π ごとに一周を表すグラフを棒グラフなどと併用することにより、周期的に変化するデータの可視化に用いる提案がなされた。

スプレッドシートによる可視化

スプレッドシート自体が情報可視化の一種であるが、その拡張を行ったのが Chi ら [8] である。従来のスプレッドシートの各セルに含まれるのは文字列や数値などの単純なデータのみであるが、Chi らはセル内に 3D で描画されたグラフを入れるソフトウェアを作成した。これにより、二つのデータセットを比べる場合のように二つ以上の可視化に対して演算を行うことが容易になり、また複数種の可視化を並べてどの可視化にどのような特徴があるのかをデータを変えつつ比較することができる。

Dynamic Queries

Ahlberg らによって提唱された Dynamic Queries[9] は、データベースの提示方法とクエリの入力画面双方を図によって可視化することによってデータベースの検索性を高めるべく提案されたシステムである。本論文の実験では、元素周期表の中から指定された条件に合致する元素を検索するというタスクをデータベースのみを可視化したシステムとクエリの入力画面のみを可視化したシステムと Dynamic Queries で行った結果、Dynamic Queries を用いた検索が所要時間の短さとエラー率の低さともに他二つよりも優れているという結果が示されていた。

Table Lens

巨大なスプレッドシートや表の中から必要な情報を取り出しやすくする可視化を行ったのが Rao らによる Table Lens[10] である。Table Lens では、フォーカスしている行の数値は数字として、その他の行の数値は線などとして表示し全体としてはグラフのように見せることによってフォーカスしている行の数値が全体としてはどの位置にあるのかということとどれくらいの大きさであるのかということの両方を一目で見ることが出来る。複数行をフォーカスすることができ、個々の値同士の比較と全体との比較を同時に行うことが出来るのも特徴である。

FOCUS

先に挙げた Dynamic Queries や Table Lens の流れをくむ表に対する可視化としては Spanke らによる FOCUS[11] も挙げられる。FOCUS では cases-by-attribute の表から特定の要素だけを抽出して表示する機能や、特定の条件を満たす事例のみを抽出することができる機能を実装しており、商品カタログなど膨大な選択肢の中から比較検討を行う際に有用な可視化を提供している。

prefuse

情報可視化を行うアプリケーションのためのツールキットとして Heer らが作成したのが prefuse というツールキットである [12]。このツールキットは、データから可視化に必要な情報を取り出すアクションを数多く用意することにより、ユーザが自分の望む機能を持った情報可視化アプリケーションを作成することを容易にしている。特にツリーマップを作成することに特化しており、数行の命令でツリーマップを表示させるアプリケーションの作成が可能である。

DataTone

Gao らによって制作された DataTone[13] は、行と列で表されるデータセットから自然言語を用いて必要なデータを取り出し可視化を行うことを可能にしたシステムである。このシステムでは、まずユーザにクエリを命令文の形式で入力させる。例えば "Show me medals for hockey and skating by country." と言った具合である。システムは入力された文章からデータベースからどのようにどんなデータを取り出すのか、可視化にどのようなグラフを用いるかをパーサーなどを使用して計算し、グラフを自動的に作成する。これにより、ユーザが望む可視化をより容易に、よりの確に行えるようになる。

SketchInsight

Lee らによる SketchInsight[14, 15] は大型タッチスクリーンを拡張するシステムであり、簡単なジェスチャーでデータから任意の形状のグラフを作成するだけでなく、ユーザが描いた絵を用いたグラフに適用させることができる。これによりオリジナリティのあるグラフが簡単に作成でき、よりユーザの表現意図に沿った絵をプレゼンテーションに用いることができる。

TurningPoint

TurningPoint[16] は PowerPoint のプラグインとして制作された、ストーリーの組み立てを支援するシステムである。このシステムはタイムライン状の部分とフリースペースから構成されている。ユーザはフリースペースに画像やメモを自由に貼り付けていくことによってブレインストーミングやアイデアをまとめることが出来、ある程度まとまったらそれらをタイムラインに貼り付けていくことによって全体のストーリー構成を考えながらプレゼンテーションの作成を行うことが出来る。

GeoTime Stories

Eccles らによる GeoTime Stories[17] も TurningPoint と同様にストーリーを組み立てる部分を支援するシステムである。これはストーリーをノードの繋がりのみなし、図上に可視化されたノードにユーザがテキストを付与していくことによってストーリーがどのような繋がりになっているのかを視覚的に把握しながらストーリーの組み立てを行うことが出来る。

課題リストに用いられる可視化

Spry は、情報可視化の実用的な用途として病院での使用を提案した [18]。病院においては、患者が抱える問題点を医療従事者全員が効率的に把握し適切な処置を行うことが治療のために不可欠であるが、問題点は複雑多岐に渡り簡単には共有できない。この研究においては必要な情報を適宜タブで区切って可視性を上げ、縦軸を治療のステップ、横軸に日付を持つガントチャートでどのような治療をいつ行ったのかを簡単に把握することができるシステムの開発を行い、実際の病院で運用して情報伝達にかかる時間の短縮やコストの削減に繋がることを示した。

ユーザからのフィードバックを与える支援

情報可視化のデザインを制作している途中にもユーザからのフィードバックが得られる支援も存在している。sense.us[19] という Web サービスでは、可視化されたデータの側に

議論用のスペースが設けられ、サービスのユーザが自由に議論を行うことが出来る。図 2.1 では右端のセグメントに対する支援であり、自分の制作している可視化が意図した情報が伝わるものかどうか読み手から直接反応を得ることができ、正式に発行する前にユーザからのフィードバックを反映させることができる。同様のシステムとして、Tableau*¹上でユーザが可視化されたデータに対してフィードバックを行うことが出来るシステム [20] がある。

可視化に用いられる色に関する研究

Healey らは可視化の際に効果的な色を選択する方法を研究している [21]。それによると、複数の色を使用していても指定された特定の色を素早く正確に見つけることが出来るような色を選択するには以下の3つの要素を調整することが重要である。

- 色の距離……色空間内で色同士のユークリッド距離
- 線形分離……色空間内でターゲット色と非ターゲット色を線形分離できる可能性
- 色のカテゴリ……ターゲット色と非ターゲット色がそれぞれ占める名前の付いた色の領域

これらを最低限保つことによって、可視化に用いる色を最大限多くしつつ指定された色を素早く正確に見つけることが可能となる。

2.1.2 infographics に関する調査

infographics の制作手法

Abilock らは「Recipe for an Infographic」 [22] において infographics の性質や制作方法を詳細に紹介している。それによると、infographics はポスターとは異なるものであり、広告ではなく調査の成果であり、説得ではなく議論を呼び起こすものだとしている。この条件を満たす infographics を制作するにあたっては、制作者は次の5つの要素を意識しなければならないと指摘している。

- infographics のテーマに関心を持つ閲覧者が誰であるか
- 閲覧者がどんな問題や課題を気にかけるか
- 閲覧者が決断するためにどういう選択肢やトレードオフを考慮する必要があるか
- 制作者が集めた情報をまとめるのにどういった考え方が必要か
- 利用できる知識としてどういったコンテンツを学んだことがあるか

*¹ <http://www.tableau.com/>

著者らはこれら全ての要素に気を配る方法として、各要素を行とした行列を用いてこれらを全て俯瞰しながら infographics のテーマを決めることを推奨している。

infographics を活用した教育

Kiber ら [23] は視覚判断能力を養うための教材として infographics を使うことを提案している。大学生を対象として実際に調査を行ったところ、学生の理解度を示す指標である rubric evaluation は「可視化」など複雑な要素では低くなったが、「色」や「フォント」、それに「情報をまとめる」と言った要素では高くなる傾向があった。

報道における infographics の使用例

Dick によると [24], すでにイギリスのオンラインニュースではユーザがインタラクティブに触れることが出来る infographics が日々制作され公開されており、ジャーナリストへの聞き取りやニュース発行社の観察からはある種のデータは他の手段で伝えるよりも伝えやすく、ニュースの価値を上げることが示されている。しかし、従来用いられているグラフなど一部のフォーマットを用いた infographics はユーザが忌避しやすいという指摘がなされている。

メディアにおける infographics の課題

Medina らは、メディアにおける infographics は情報の価値創造に役立っており、その作成作業は企業のワークフローに影響を与えるほど普及しているものだとしている [25]。一方で、infographics の制作には視聴覚を用いた報道やよく検討されたビジネスモデルに精通したプロフェッショナルが必要なため、全ての企業でその真価を発揮しきっているわけではないとも指摘している。

ビジュアルコミュニケーションデザインの教育における infographics の重要性

Dur による調査 [26] では科学的技術的進歩により我々がより多くの情報に接しなければならないことを指摘し、それによる整理されていない情報の氾濫が情報デザインの重要性を高めているとしている。そのため、infographics やデータ可視化などの情報デザインはビジュアルコミュニケーションのデザインの教育を行う際により体系的に教育されるべきものであり、その教育は今日の世界の需要に見合うデザイナーを育てるのに重要な役割を果たすであろうと結論づけている。

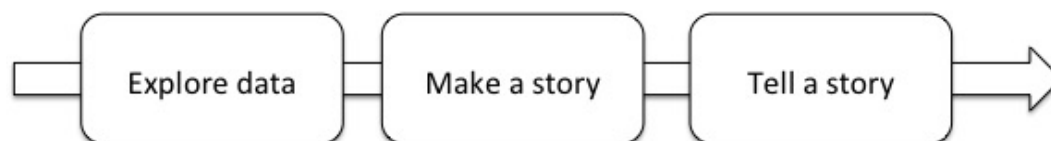


図 2.1: ストーリーテリングのプロセス:ストーリーはデータを分析し、どういったストーリーにするかを決め、ストーリーを伝えるのに必要な材料を集め、発表したストーリーにフィードバックを得ることによって作られる (Lee et al.[27] の論文より再作成).

2.1.3 ストーリーテリング

データ可視化において重要となるのがストーリーテリングである。データからどのようなストーリーを紡ぎだすかによって、同じデータを可視化したものでもそれから受ける印象は異なったものになる [27]。図 2.1 はストーリーテリングのプロセスを図式化した図である。ストーリーテリングはデータの探索、ストーリーの組み立て、そしてストーリーを伝えるという3つの領域に大別できる。データの探索は **infographics** 制作においては制作に入る前に行っておくべきことのため、本研究では特にストーリーの組み立ておよびストーリーを伝える領域に着目する。

2.2 制作ツール

infographics を制作することの出来るツールやサービスは数多く開発されている。以下にその一部を示す。

2.2.1 Adobe Illustrator

デザインのプロが **infographics** を制作する際に頻繁に用いられているのが **Adobe Illustrator***²である。このソフトは **infographics** だけではなく様々なデザインに使用することが出来るほど非常に多機能であり、商用に用いるイラストやデザインを作成するのに使用しているプロのイラストレーターも多い。しかしそれ故に初心者が使用する際には自分の行いたい作業を実現するにはどの機能を使えば良いのかを探る難易度が高く、容易に **infographics** を制作することは出来ないという問題点がある。また、**Illustrator** に備わっている表データを扱う機能は非常に限定的であるため、**infographics** を制作する上ではデータを解釈しどういった可視化を決定する作業は人間が行う必要がある。

*² <http://www.adobe.com/jp/products/illustrator.html>

2.2.2 特定の形式に特化した infographics 制作サービス

iChart^{*3}や Eegraph^{*4}はデータからデザイン性の高いグラフを作成することの出来るサービスである。いずれもデータ入力の手動による入力だけではなく Excel 形式の読み込みにも対応しており、グラフをメインに据えた infographics の制作には有用なサービスである。類似のサービスとして Tableau が挙げられる。

Visually^{*5}は特定の infographics, 例えばユーザの Facebook に登録している情報を可視化した infographics などの作成に特化したサービスである。使用することが出来るデータもデザインも自由度が低く、オリジナリティの高い infographics を作成するのは困難であるが、データを連携させるだけであるので非常に簡単に、素早く infographics を作成することができる。

2.2.3 infographics 制作サービス

easel.ly^{*6}, infogr.am^{*7}, piktochart^{*8}, Venngage^{*9}といったサービスでは, Microsoft Powerpoint でスライドを作成する際と同様の操作で infographics の制作を行うことが出来る。デザインのテンプレートも何種類か用意されており, あらかじめ指定されている場所にグラフや文章を挿入していくことによって簡単にデザイン性の高い infographics を制作することが出来る。自由に使うことのできるピクトグラムや写真, 画像素材が用意されているので, 使用する画像を用意する必要が無いということも infographics 制作を容易にしている。

2.2.4 既存システムの問題点

上に挙げたような既存のシステムにはいくつか問題点がある。まず, 使用するピクトグラムを収集するのに手間がかかるという点である。ピクトグラムのみを集めた画像素材集を利用することもできるが, その素材集に収録されていないオブジェクトのピクトグラムを利用したい場合, 自分で検索して保存し, それをシステム上で使えるようにアップロードあるいはインポートするという手間がかかる。piktochart 等ではシステム内に予めピク

*3 <http://www.icharts.net/>

*4 <https://eegraph.com/>

*5 <http://visual.ly/>

*6 <http://www.easel.ly/>

*7 <http://infogr.am/>

*8 <http://piktochart.com/>

*9 <https://infograph.venngage.com/>

トグラムが用意されているが、上に挙げた問題は依然残っており解決にはならない。

他にも、デザインに関する問題が残っている。既存のシステムの一部には配置や配色を決められているテンプレートが多く用意されているが、オリジナルのデザインを作成しようとする際にはあまり手助けが無く、制作者にデザインに関する知識やセンスが求められる。独自性の高い **infographics** の制作には非常に高い壁となる問題点である。

また、**infographics** では画像やピクトグラムと図形を組み合わせて用いることがある。例えば全ての画像やピクトグラムを円の中に配置して、**infographics** 内でのデザインを統一させるケースがある。全ての画像に対しその処理を施すことは、既存のシステム内では手動で行うこととなり、非常に労力のかかる作業である。既存のピクトグラムですでにそういった処理を施されているものを検索して見つけるという方法もあるが、検索の手間がかかるだけではなく **infographics** を制作しているソフトウェアと検索を行うブラウザとを何度も行き来する必要があり煩わしい作業となり得る。

さらに、**infographics** では図 2.2 のようにピクトグラムを用いて作成されたグラフを使用することがあるが、そのようなグラフを作成するにはピクトグラムの大きさや配置場所、個数を自らデータに合わせて決定しそれに合わせてピクトグラムを加工しなくてはならず、手間がかかる。例えば「30%」という情報を表している図 2.2 の場合であれば、

1. 使用する画像を用意する
2. グラフの形状（今回のケースであれば同一の画像 10 枚のうち 3 枚を塗りつぶす）を決定する
3. 完成したグラフの大きさを決め、それに合わせて画像一枚あたりの大きさを決定し、画像を適宜拡大縮小する
4. 画像を必要な枚数だけ複製する
5. 等間隔に整列させる
6. 必要な枚数に対して塗りつぶしを実行する
7. 最終的に配置する場所に配置する

という手順を踏むことが必要になる。この際、画像全てを塗りつぶさない場合、例えば 37% を表そうとして塗りつぶした画像 3 枚と 7 割を塗りつぶした画像 1 枚が必要な場合は、7 割塗りつぶすためには画像下部から何ピクセルまで塗りつぶすのかを画像の大きさを元に自分で計算する必要があり、さらに時間を要する。

本研究では、上に挙げた問題点のうち特にピクトグラム収集にかかる手間とピクトグラムを用いたグラフの作成に伴う煩わしさの問題を解決する。この二つは [27] で述べられているストーリーテリングのセグメントの中でもストーリーを伝えるセグメントのうちストーリーに用いる素材を探すプロセスに関する問題であり、ここを支援することによって編集作業にかかる手間を省くことが出来ると考えられる。

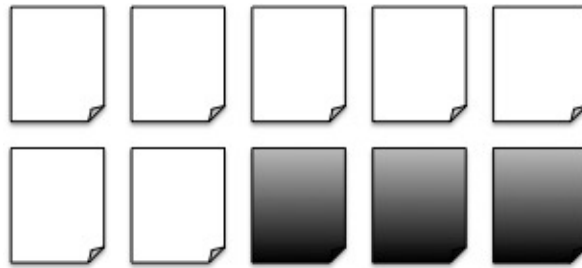


図 2.2: ピクトグラムを用いたグラフの例:「30%」という情報を提示している.

そこで、提案するシステムではインターネットから既存の画像検索サービスを用いて infographics に適切なピクトグラムを検索・入手し、その画像を用いて infographics で用いるグラフを作成する機能を実装した。画像がピクトグラムであるかどうかの判定には画像検索エンジンの検索条件を工夫することだけではなく、機械学習に基づいた判定を用いることによってより適切な画像の入手を試みた。

第 3 章

システムの実装

3.1 ピクトグラムを取得しグラフを作成する機能の詳細

図 3.1 は開発したシステム「grafie」に実装したインターネットからピクトグラムである画像のみを取得し、その画像を用いたグラフを作成する機能の処理を模した図である。以下にフローの詳細を詳しく述べていく。なお、本システムは Web ブラウザ「Google Chrome*1」上での操作を想定している。

1. 最初にユーザは文章を入力し、図として表したい単語を選択する。grafie は選択された単語に、「icon png」という文字列を加えたものをクエリとして Bing Search API にて画像検索を行う。結果は画像への URL が含まれた XML ファイルとして返ってくるので、XML ファイルから URL を取得、その URL を用いて画像のダウンロードを行う。なお、API で提供されている Adult オプションによって性的な画像を排除し、ImageFilters オプションによってピクトグラムの可能性が低い写真などの画像を除いた結果を入手するようにした。
2. 取得した画像を 4.2.2 で述べる手順によって加工し、計算された特徴量に対して正規化を行う。その結果を識別器に入力しピクトグラムかどうか判定する。識別機は 4.3 で詳細を述べる。ピクトグラムである場合は表示させる画像として保存し、そうでない場合は XML ファイルから新たな URL を取得し同じ手順を踏む。
3. もし一番最初にピクトグラムであると判定された画像であるならばキャンバス上に表示し、そうでなければ表示しない。
4. また、選択された単語の類義語を WordNet*2を使って検索する。もし類義語が存在していた場合、各々に対して 1.~3. と同様の手順で画像を収集する。検索に時間が

*1 <https://www.google.co.jp/chrome/browser/desktop/>

*2 <https://wordnet.princeton.edu/>

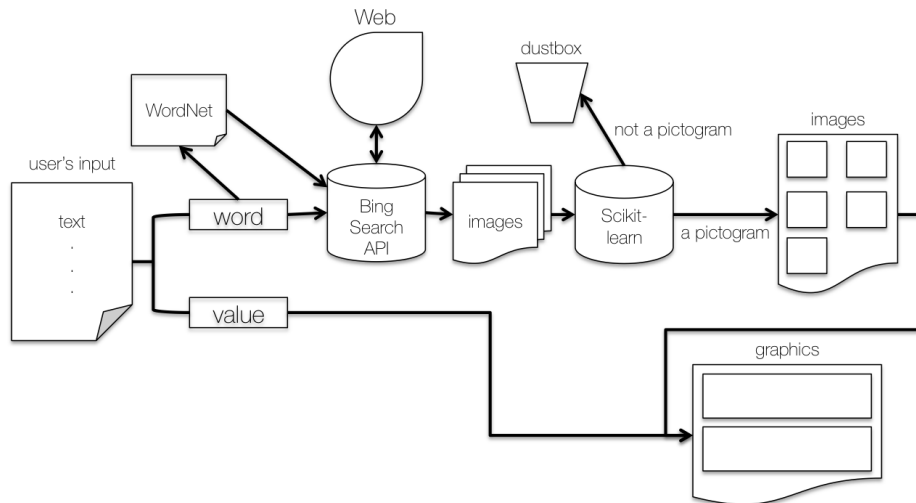


図 3.1: 実装した機能のプロセス図：入力されたテキストから選択された単語を用いて画像検索を行い、ピクトグラムかどうか判定する。ピクトグラムであると判定された画像の中からユーザが選択した画像と入力されたテキスト中の数値を用いてグラフを作成する。

かかりすぎることを防ぐため、類義語は最大5つまでしか取得しない。

5. ユーザは用いられている画像が適当でないと判断した場合、ピクトグラムであると判断された他の画像一覧を閲覧でき、そのうちのひとつをクリックすることによって画像の差し替えを行うことができる。類義語がある場合はその場合の検索結果も閲覧し、用いる画像として選択することができる。
6. ユーザは画像を用いたグラフの作成も行う事が出来る。グラフを作成する画面では次の5つのグラフが候補として表示される。
 - 10個の小さい画像を表示し、文章中の数字を百分率としたものを使用して10個のうち数個に着色する(図 3.2(a))
 - 文章中の数字を百分率としたものを使用して画像の下からその比率分を着色する(図 3.2(b))
 - 画像で文章中の数字上二桁を表す(図 3.2(c))
 - 文章中の数字を百分率としたものを使用して画像の左からその比率分を着色する(図 3.2(d))
 - 画像をそのまま用いる(図 3.2(e))

なお、用いる数字は文章中の数字を抽出したものであり、数字が複数ある場合はどの数字を用いるかユーザが選択することができる。

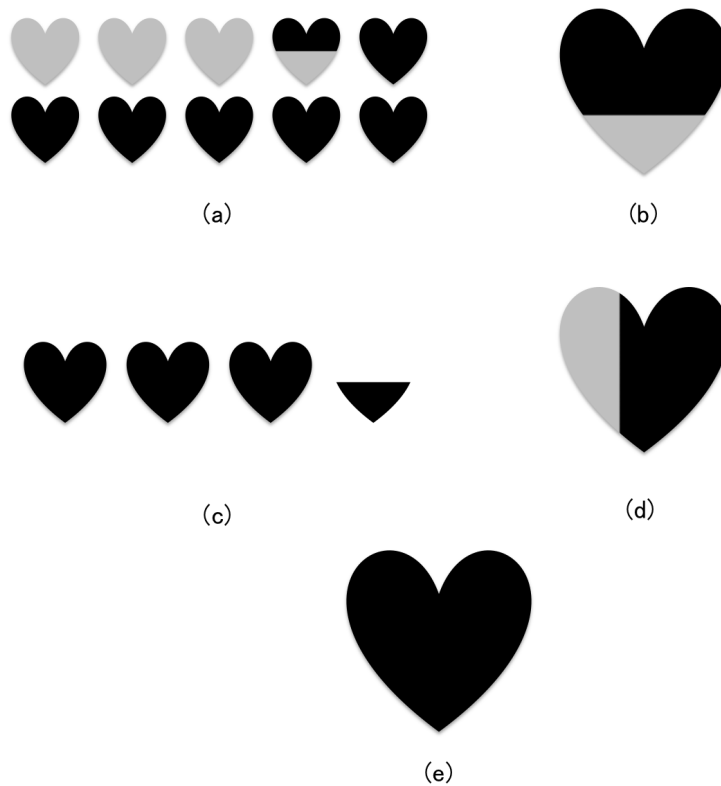


図 3.2: 本システムで作成することの出来るグラフの例：数値として 35 を使用している。

3.2 その他の機能

grafie には infographics を作成するための機能として、以下のような機能も実装されている。

キャンバスの設定 キャンバス（描画域）のサイズとキャンバスの色の設定機能。サイズはブラウザの画面上でのサイズに自動でフィッティングし、色は任意の色を選択することが可能である。

直線の描画 キャンバスが許す限りの任意の長さ、1px から 100px までの任意の太さ、任意の色を持つ直線の描画機能。描画した後も線そのものの位置と始点と終点の位置を変更することが可能である。

図形の描画 円、正方形、角丸正方形、星型を直径 1px から 200px までの任意の大きさで描画する機能。縁の色と塗りつぶしの色をそれぞれ任意に設定することが出来る。描画した後も大きさは変更可能であり、正方形と角丸正方形は長方形にすること

も可能である。

文章の入力 文章を入力する際にはフォントの大きさと色を任意に変更することが可能である。

これらの機能は、対応するモードを選択してからキャンバスをクリックすることによって利用できる。色を変更する際には、どのモードでも16進数でのRGB表記を直接打ち込む方式と、カラーバーによりHSVを選択する方式とどちらの方式でも変更が行えるようになっている。

3.3 実装

実装はブラウザサイドに Javascript, サーバーサイドの CGI として Python2.7 と PHP を使用した。また, Python にて機械学習を行うにあたって scikit-learn^{*3}, 画像処理を行うにあたって OpenCV^{*4}を用いた。画像検索には Microsoft 社が提供している Bing Search API^{*5}を使用した。

3.4 使用例

図 3.3(a) に示すのが grafie が Google Chrome 上で動作している様子である。右側にメニューボタンが配置されており、対応するボタンを押すと描画モードが変わり、ボタンの下にオプションを選択するスライダーやカラーバーが表示される。描画モードは上から「select (選択)」「line (直線を描画)」「shape (図形を描画)」「upload (画像のアップロード)」「background (背景色変更)」「text (文章の入力)」となっており、図 3.3(a) では「text」モードが選択されている。図 3.3(b) は文章を入力したところである。文章を入力するモードを選択してからキャンバス上の任意の位置をクリックするとテキストボックスが現れ、そこに任意の文章を入力することが出来る。

図 3.3(c) では文章の一部を選択している状態で右クリックし、コンテキストメニューを表示させている状態の画像である。一番下の「Search」をクリックすると選択されているテキストを用いて Web 上の画像検索を行い、結果を識別器にかけピクトグラムであると判断したものだけを候補として保存する。なお、誤動作を防止するために検索を行っている間はシステムの画面を暗くし、操作不能にしている。ピクトグラムと判断された画像が一枚でもあると操作不能状態は解除され、一番最初にピクトグラムであると判断された画

^{*3} <http://scikit-learn.org/stable/#>

^{*4} <http://opencv.org/>

^{*5} <https://datamarket.azure.com/dataset/bing/search>

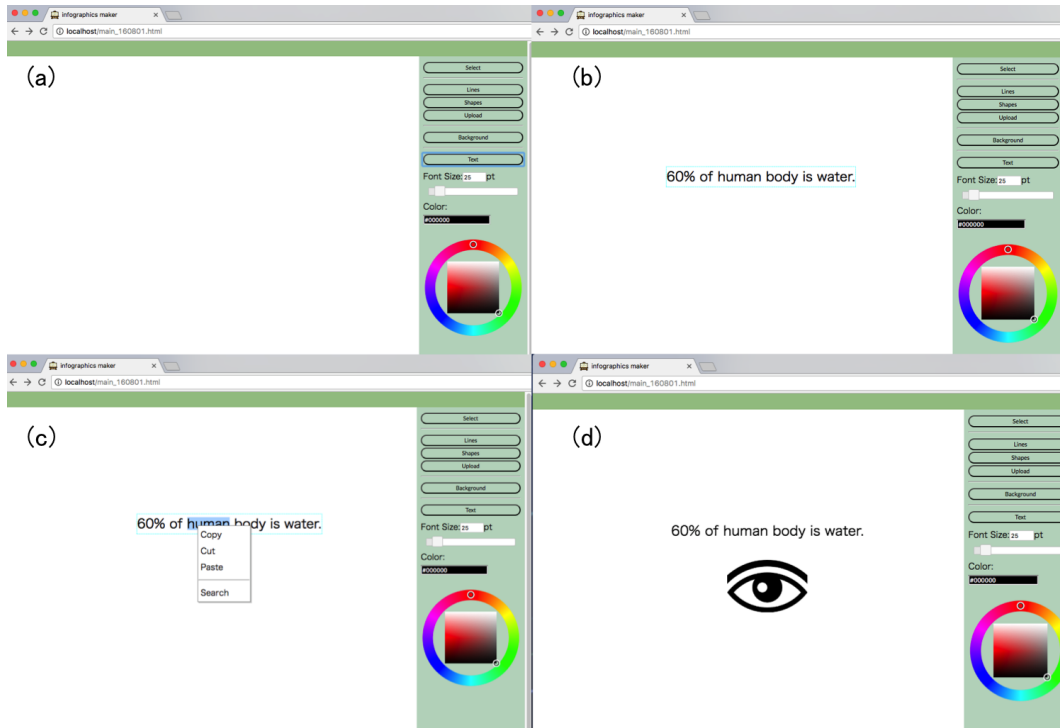


図 3.3: grafie 使用例 1 : (a)grafie を起動した直後の状態。「Text」モードが選択されている。(b) 文章の入力をした状態。現在編集中的の文章は水色の点線枠で囲われる。(c) 文章中の単語（この場合 people）を選択しメニューを出した状態。一番下の「Search」をクリックすると画像検索が開始される。(d) 最初の検索結果が出現した状態。画像が表示された後もバックグラウンドでは検索が引き続き行われている。

像がキャンバス上に表示されて図 3.3(d) の状態になる。この際、バックグラウンドでは引き続き画像の検索、ダウンロードおよび識別が行われている。また、表示された画像は選択された状態になっている。画像は文章に紐付けられており、一つの文章に対して 1 個または 0 個の画像という対応関係になっている。画像をクリックして選択すると青い枠が表示される。その状態で画像上で右クリックすると図 3.4(e) のようにコンテキストメニューが表示される。上から順に「change base color (ピクトグラムの色を変更)」「graph (ピクトグラムからグラフを作成)」「change pic (ピクトグラムの変更)」「delete (ピクトグラムの消去)」である。図 3.4(f) は「ピクトグラムの変更」をクリックして選択した状態である。新たにウィンドウが開き、その時点までにピクトグラムであると識別された画像の一覧が表示される。ウィンドウ上部には、図 3.3(c) で選択している語句の類義語が存在している場合にはそれらが表示されている。これらの類義語をクリックすると、その類義語での検索結果に基づいた画像が表示される。

現在の画像の代わりに使用したい画像をクリックすると（この場合左から二列目最上行）、すでに表示されている画像の代わりにクリックした画像がキャンバスに表示される（図 3.4(g)）。画像の変更は何回でも行うことが可能である。

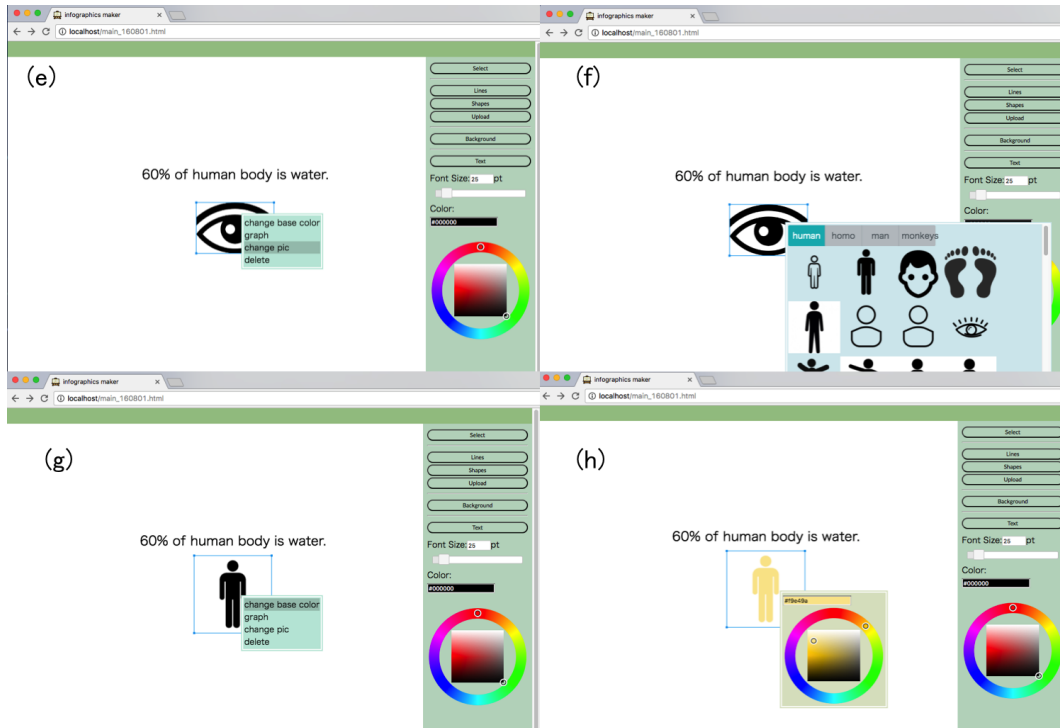


図 3.4: grafie 使用例 2 : (e) 画像をクリックすると選択状態になり青枠が付く。この状態で画像を右クリックすると画像に対する処理のメニューが展開される。現在実装している機能は 4 つである。(f)(e) で「change pic」を選択した状態。上部に **people** の類語が表示され、下部には **people** で画像を検索した結果のうちピクトグラムと判定されたものが並ぶ。ピクトグラム判定が終わった画像から随時追加されていく。(g) 画像を図 (f) の最上段左から二番目に設定した後に再度画像メニューを開いた状態。(h) 「change base color」を選択してカラーサークルを表示させた状態。カラーサークル上の任意の位置をクリックすると都度キャンパス上のピクトグラムの色を変更される。

コンテキストメニューの一番上の「change base color」をクリックして色変更ウィンドウを呼び出したのが図 3.4(h) である。ウィンドウ内に表示されているカラーサークルで任意の色を選択すると、ピクトグラムをその色に着色することができる。

再び画像のコンテキストメニューを表示し (図 3.5(i)) , 一番上の「graph」をクリックしてグラフ作成ウィンドウを呼び出したのが図 3.5(j) である。表示されているグラフの候補は文章中の数値とその文章に紐付いている画像から生成されている。文章中に複数の数値が含まれている場合は、上部タブから用いる数値を選択することが出来る。また、16 進数 RGB が表示されているテキストボックスをクリックするとカラーサークルが表示され、それを用いてグラフに用いる色を選択することが出来る。

図 3.5(k) は完成した infographic である。「人体の 60 % は水である」という文章が下部 60 % が水色に塗りつぶされた人体のピクトグラムにより表現されている。ピクトグラムの色をオレンジ寄りの黄色にすることで人体を表した。文章や画像の位置は「select」モードが選択されている時にドラッグによりキャンバス上の任意の位置へと移動させることが

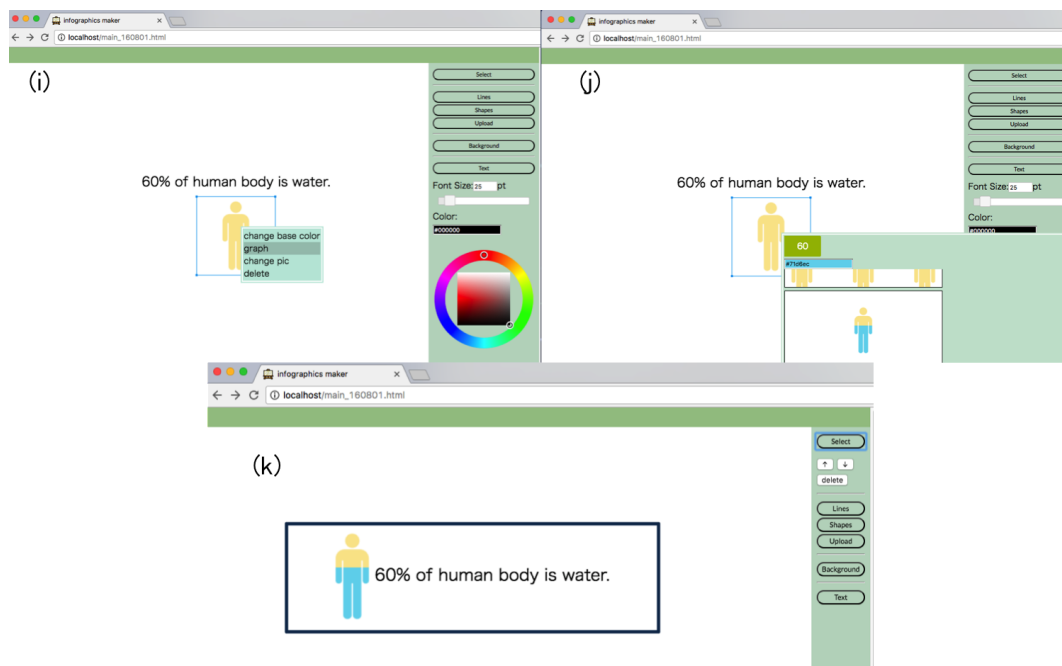


図 3.5: grafie 使用例 3 : (i) 色を変更した後に再度画像メニューを開いた状態. (j) 「graph」を選択してグラフを作成するウィンドウを開いた状態. 上部にはテキスト中に出現する数値が並び, 任意の数値を用いてグラフの作成が行える. カラーコードが入力されているテキストをクリックするとカラーサークルが表示され, グラフに用いる色を選択することができる. 下部には様々なグラフの種類が並び, クリックすることで任意のグラフをキャンバスに挿入できる (k) 作成された infographic. 「Shape」モードで角丸正方形を挿入し, 「Select」モードで変形及び並び順の変更を行って枠線を作成した.

出来る. また, 「shape」モードを選択し, 縁の色を藍色, 塗りつぶしを白色, 形状を角丸正方形に設定して図形をキャンバスに挿入し, 「select」モードで大きさや並び順, 位置を設定して枠線を作成した.

出来上がったグラフが不適當であると感じた場合は, 再び画像を選択して画像メニューから処理を選択することによって何度でもグラフを作り直すことができる.

第 4 章

システムの製作

4.1 infographics の調査

4.1.1 infographics の分類

まず、既存の infographics のデザインにはどのようなスタイルが存在していてどのような素材が用いられているのかを調査するため、Google 画像検索を用いて infographics を収集した。集めた 457 枚のうち重複しているものや極端にサイズが小さいものを除外した 362 枚の infographics を形状によって 7 種類にカテゴリ化した。7 種類の分類を図 4.1 に示す。

1. 箱型

提示する情報一つを一つのブロックとし、ブロックをレイアウトすることによって infographics とするデザイン。ブロックの位置には特に意味の無いことがほとんどであるが、ブロックの大きさによって情報の重要度を示すこともある。ブロックごとの重要度は色や文字のサイズなどで示される事が多い。

2. 線型

情報を縦一列、あるいは横一列に配置したデザインであり、上記した箱型の亜種。読み手が読む順番はある程度決まっているものの、順番そのものには重要性が無い。Web サイト上において縦あるいは横にスクロールさせて読まれることを想定し、それに特化した形状と考えられる。

3. 道型

時系列や手順に沿って情報を提示してゆく infographics。読み手の視線を線に沿って誘導するデザインになっている。

4. 比較型

二つの事柄について、共通している点と異なっている点を比較・強調する

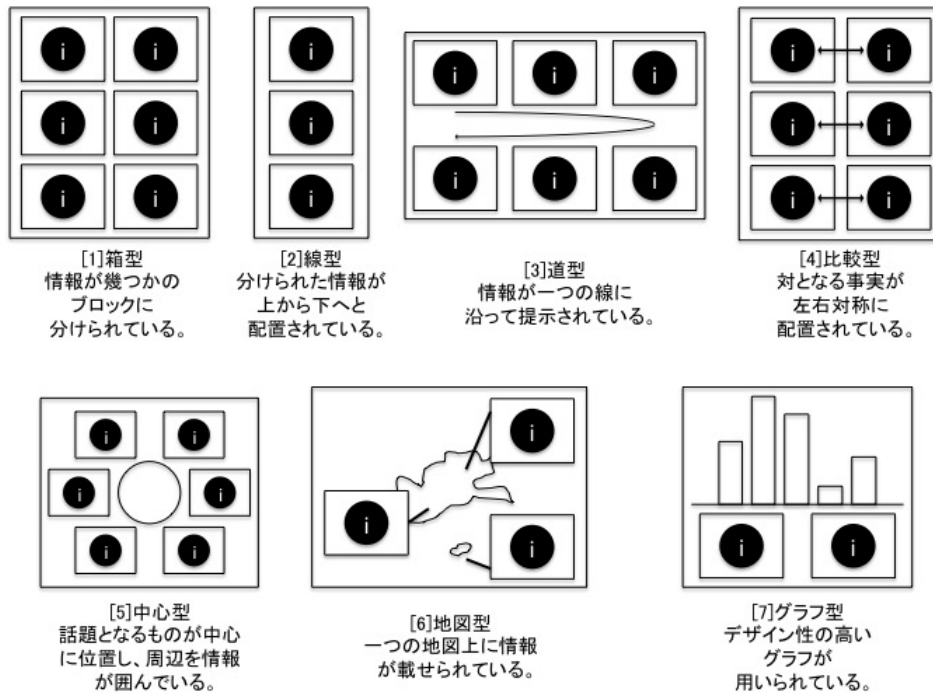


図 4.1: infographics の 7 つの分類

infographics に用いられている。左右対称や上下対称のデザインとなっていることが多い。

5. 中心型

中心に題材を配置し、周辺にその題材についての情報を列挙していくデザイン。一つの事柄のみに関する情報を扱う infographics に用いられる。

6. 地図型

地図を中心としたデザインであり、地理的な情報に基づいた比較を行うデータを扱う infographics に用いられている。

7. グラフ型

一つあるいは複数のグラフからなる infographics。提示する情報の数が少ない時に用いられており、配色や形状に工夫が凝らされている。

なお、7種それぞれの枚数の内訳は表 4.1 の通りである。

4.1.2 ピクトグラムの有無の調査

次に、362枚の infographics それぞれに対して、Amazon Mechanical Turk を用いてピクトグラムが用いられているかどうかの調査を行った。調査は同じ画像を5人のワーカーに見てもらい、ピクトグラムが用いられているかどうかをそれぞれ判断してもらい、その結

表 4.1: インターネットから収集した infographics を 4.1 にもとづいて分類した結果

種類	枚数 (枚)
箱型	145
線型	78
道型	32
比較型	15
中心型	59
地図型	13
グラフ型	20
計	362

果で多数決を行い用いられているかどうかを決定した。その結果を表 4.2 に示す。なお、報酬として infographic 一枚につき 0.04 米ドルをワーカーに支払った。

表 4.2: infographics の各形状ごとにピクトグラムの有無を AMT を用いて分類した結果

種類	ピクトグラムの有無		計
	有	無	
箱型	102(70.3%)	43(29.7%)	145
線型	52(66.7%)	26(33.3%)	78
道型	23(71.9%)	9(28.1%)	32
比較型	8(53.3%)	7(46.7%)	15
中心型	29(49.2%)	30(50.8%)	59
地図型	8(61.5%)	5(38.5%)	13
グラフ型	12(60.0%)	8(40.0%)	20
計	234(64.6%)	128(35.4%)	362

以上の結果により、ピクトグラムは 6 割以上の infographics で用いられていることがわかった。形状によって用いられる割合は異なるものの、地図型と中心型を除く全ての形状でピクトグラムの使用率は 5 割を超えており、特に最も枚数が多い箱型では 7 割を超える使用率となった。このことからピクトグラムは infographics のデザインにおいてはピクトグラムが主要な要素であり、これを適切に選択することが infographics の製作においては重要であると言える。しかし、デザインに関しては個人の好みが多岐にわたる存在すること、また最適解が一つではなく複数存在しうるなどから、システム側はあくまでピク

トグラムの候補を示すこととし、最終的に用いるピクトグラムを選択するのはユーザの意志に委ねることにした。

4.2 機械学習を用いたピクトグラムの判定の予備実験

4.2.1 機械学習の必要性

ピクトグラムを用意することに手間がかかる問題点を解決するために本研究ではオンラインの画像検索 API を利用している。しかし検索クエリや検索条件を設定するだけではピクトグラムだけを選択的に入手することが出来ないことが事前の調査で判明した。一例として、検索クエリを「(キーワード) icon png」として検索オプションで白黒かつイラストレーションである画像を指定しても、ピクトグラムやアイコンだけではなく複数のピクトグラムを一覧にした画像や、白い背景でモノトーンの物体を撮影した写真が検索結果に出現する。そこで、画像検索 API を実行して入手した画像を識別器にかけ、ピクトグラムかどうかを識別することにした。

4.2.2 学習データの準備

学習データの収集には Google 画像検索^{*1}を用いた。まず、収集した infographics ではどのような話題を扱っているかやどのようなピクトグラムが用いられているかを調査し、用いられているピクトグラムから 17 個を選択した。それらを表す単語をクエリとして画像検索を実行した。収集した画像のうち、透かしの文字が入っている画像、複数個のピクトグラムを集めた画像、他の画像と同一である画像などを取り除いた。結果 3277 個の画像を得た。その中から 134 個をランダムに抽出し、それぞれに対してピクトグラム（アイコン）かそうでないかのラベル付けを手動で行った。また、特徴量の抽出を同条件で行えるように、134 個すべての画像に対し、

1. 画像の輝度を用いてグレースケール画像に変換する。画素の RGB 値を $[r, g, b]$ とすると、輝度を求める計算式は以下で示される。この値をグレースケール画像における画素の値として用いる。

$$\text{luminance} = 0.299 \times r + 0.587 \times g + 0.114 \times b$$

2. 画像の余白をトリミングする。画像のサイズを横 w ピクセル、高さ h ピクセルとした時、画像は全ての要素が $[r, g, b, \alpha]$ という 1×4 行列である $h \times w$ 行列と見なすことが出来る。ここで r, g, b はそれぞれ画素の RGB 値を表し、 α は画素の不透明度を表す。 1×4 行列の要素は全て 0 以上 255 以下の整数値を取る。この行列を

^{*1} <https://www.google.co.jp/imghp>

上の行から走査して一番最初に $\alpha \neq 0$ となる要素が存在する行のインデックスを h_{upper} ，下からの走査において同様の行のインデックスを h_{lower} ，左右からの走査において同様の列のインデックスをそれぞれ w_{left} ， w_{right} とすると，画像から左上の座標が (w_{left}, h_{upper}) ，右下の座標が (w_{right}, h_{lower}) となる領域を切り出せば余分な透明部分の余白を除いた画像が得られる。

3. 透明度の情報を無くすために，画像と同サイズで全ての画素が $[r, g, b] = [255, 255, 255]$ である画像，つまり白色で塗りつぶされた画像を画像の下に合成する。
4. 画像が 100 ピクセル四方の白色で塗りつぶされた正方形に収まるように拡大縮小，および画像がその正方形の中心になるように配置する。

という処理を行った。図 4.2 に処理例を示す。グレースケール画像にする際に明度ではなく輝度を用いたのは，色による明るさの違いを考慮するためである。色の違いを処理を行った 134 個の画像のうち，画像のファイル名を降順で並べた際の最初の 20 個を検証用データ，残り 114 個を学習用データとした。

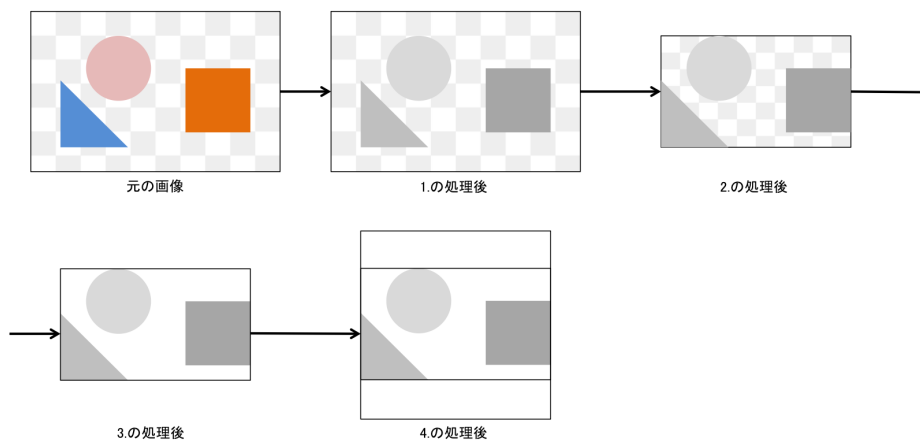


図 4.2: 学習データとして用いる画像に行った処理を表した図：画像を輝度を元に白黒に変換し，トリミングを行い，背景に白い画像を合成し，100px 四方の正方形に収まるように拡大縮小しその正方形の中心になるように配置する。便宜的に元の画像の枠線を表示している。薄い市松模様で示される部分は透明であることを示す。

4.2.3 特徴量

画像のピクトグラムらしさを定量的に表す指標として，機械学習に用いる特徴量には画像に含まれる画素の標準偏差と画像のコーナー数を採用した。画像のコーナー数を計測するにあたっては，Harris のコーナー検出法 [28] を使用した。二つの特徴量をそれぞれ平

均0, 分散が1となるよう正規化し, ヒストグラムを作成したところそれぞれ図4.3, 図4.4のようになった. 図で色が濃くなっているのは2つのヒストグラムが重なっている部分である. おおよそ正規分布になっており, 特徴量としては妥当であると言える. また, 散布図は図4.5のようになった.

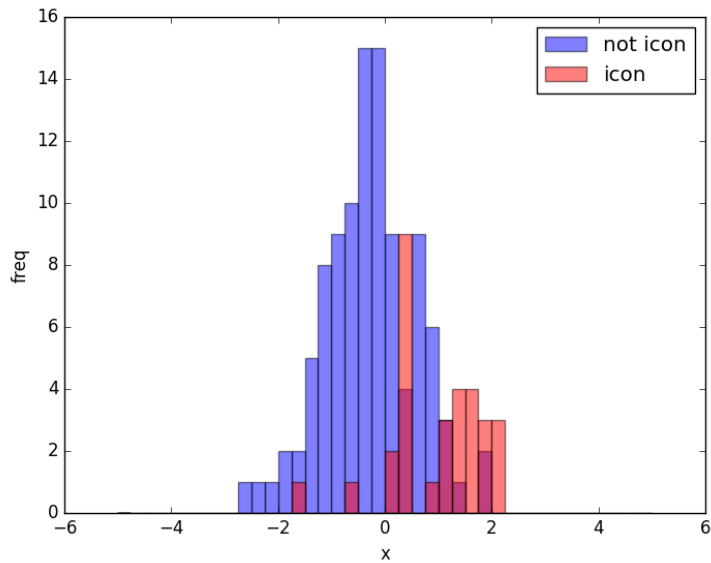


図 4.3: 正規化された画素の標準偏差の分布: 紫色部分は両クラス同数の数であることを示す.

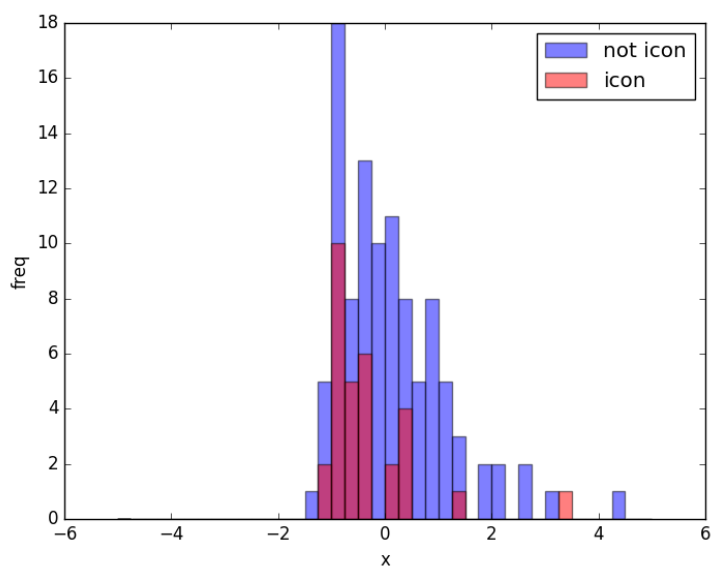


図 4.4: 正規化されたコーナー数の分布: 紫色部分は両クラス同数の数であることを示す.

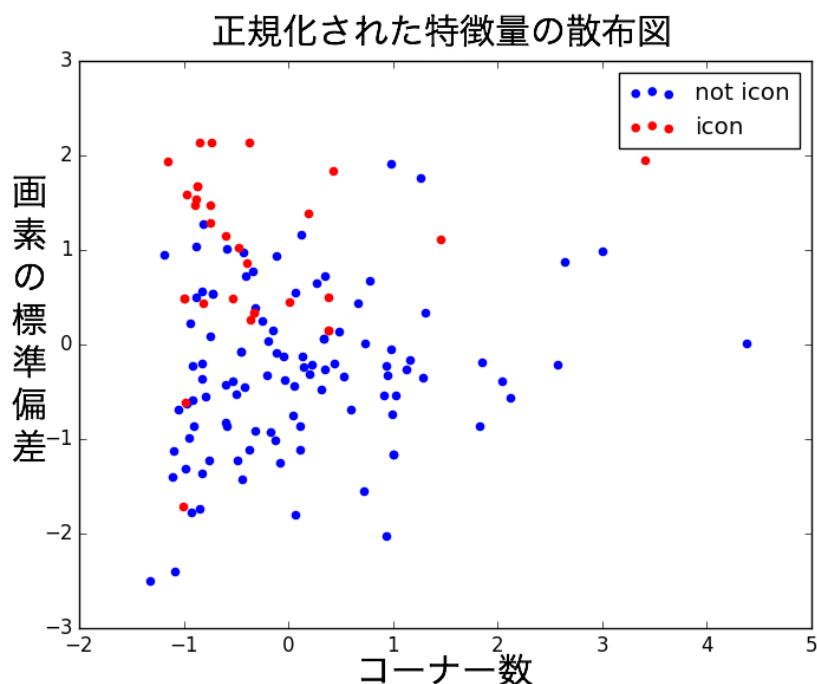


図 4.5: コーナー数と画素の標準偏差との散布図：ピクトグラム (icon) は左上に分布している。

4.2.4 学習

学習には Linear SVC を使用した。Grid search を行った結果、C 値は 0.03125 となった。先に作成した検証用データを用いてホールドアウト法で検証を行ったところ、識別率は約 86% であった。混同行列は表 4.3 の通りである。表では「予測値」が学習した識別器で予測した結果、「正解」が予めラベル付しておいたデータを示し、「陰性」はピクトグラムでないこと、「陽性」はピクトグラムであることを示している。誤判定はピクトグラムでないことラベル付けされている画像をピクトグラムであると予測したケースに集中していた。

表 4.3: ピクトグラム認識テストの混同行列

		予測値	
		陽性	陰性
正解	陽性	12	0
	陰性	6	2

4.3 ピクトグラム判定手法の本実験

以上でピクトグラムかそうでないかの判定に機械学習が有用であることを示すことが出来た。しかし、前節の機械学習では学習データに用いた画像の枚数が少ないため、より実用的な識別器を作成するためにはより多くの画像を学習データとして用いなければならない。また、infographics で用いられるピクトグラムを識別する識別器を作成するためには学習データには infographics で用いられる頻度が高い単語を表した画像が使用されることが望ましい。そこで、先に収集した infographics から出現頻度が高い単語を抽出し、その単語を用いて画像を取得、処理を施した後にピクトグラムかそうでないかのラベル付を行ったものを学習データとして使用した。

4.3.1 infographics で用いられているピクトグラム

4.1.2 の実験の際、ピクトグラムが用いられている infographics に対しては回答と同時に何を表したピクトグラムが使用されているかも任意で列挙してもらった。この設問の回答は任意であることに加え、複数のピクトグラムが使用されている場合は複数回答可である。得られた結果を cat や dog であれば pets に、vegetables や eggs であれば food に分類するように上位概念に手動で分類した結果、5枚以上の infographic に現れたピクトグラムは表 4.4 のような結果になった。

4.3.2 画像の取得

学習データは表 4.4 に出現した語に「icon png」を付与して Google 画像検索で検索を行って得られた png 形式の画像によって作成した。ただし、検索結果にロゴマークが多く出現したために social media は除外した。上から順に画像の取得を行ったところ、digital devices までで 5672 枚という十分な枚数の画像が得られたので後の処理の手間などを考え画像の取得はここまでとした。内訳は表 4.5 の通りである。

4.3.3 画像の前処理

前節で得られた 5672 枚の画像に対して、

1. 300 ピクセル四方に収まるようにリサイズ
2. 300 ピクセル四方の白い画像を背景に合成
3. 輝度によるグレースケール変換

表 4.4: infographics で頻出している単語

単語	出現枚数
people	27
money	23
health	15
animal	11
social media	10
nature	10
food	9
learning	7
vehicle	6
digital devices	5
business	5
building	5

表 4.5: infographics で頻出している単語に「icon png」を付与して得られた png 形式の画像の枚数

単語	枚数
people	380
money	385
health	377
animal	529
nature	999
food	1229
learning	622
vehicle	543
digital devices	608
計	5672

という一連の処理を行った。この際、見かけ上 png 形式であるもののファイルが壊れている等の理由で処理を行うことが出来なかった 168 枚を除いた結果、5504 枚の画像が残った。

4.3.4 AMT による分類

5504 枚の画像のうち、目視でピクトグラムであると判断した画像（図 4.6）を 2 枚、ピクトグラムではないと判断した画像を 2 枚、ピクトグラムが複数枚含まれている画像を 1 枚取り除いた。残りの 5499 枚を 9 枚 1 グループとして 611 グループを作成し、1 グループに先にピクトグラムであると判断した画像、ピクトグラムではないと判断した画像、ピクトグラムが複数枚含まれていると判断した画像を 1 枚ずつ加え、1 グループにつき 12 枚の画像が含まれているようにした。3 枚の画像は予め回答が用意されている画像として不適切な回答者を弾くのに用いた。この 611 グループに対し、各グループにつき二人にそれぞれの画像が

- ピクトグラムである
(It is a pictogram.)
- ピクトグラムでない
(It is NOT a pictogram. (illustration, photo, etc.))
- 複数のピクトグラムが含まれている
(2 or more pictograms are included in this image.)
- 透かしが入っている
(It has a watermark in the image.)
- 何も表示されていない
(There is no image.)

のうちどれに属するかを Amazon Mechanical Turk を通して判断してもらった。回答者であるワーカーには、説明文を読んでもらう際にピクトグラムがどのようなものであるのかを提示した。ただし、報酬目当てで適当な回答を行う回答者を弾くため、予め回答が用意されている 3 枚の回答が間違っているものはリジェクトし、他のワーカーに回答してもらった。報酬は 1 グループにつき 0.03 米ドルとした。

4.3.5 分類結果

5499 枚のうち 3986 枚に対しては二人のワーカーの回答が一致していたが、残り 1513 枚に対しては一致していなかった。回答が一致している場合はその結果を採用し、回答が「ピクトグラムである」の場合はピクトグラムである、「ピクトグラムでない」の場合はピクトグラムでない、それ以外の回答である場合はその他と分類した。異なっている場合は



図 4.6: ピクトグラムである画像の例

目測によってどちらの回答がより相応しいかを判断し分類した。分類した結果は表 4.6 に示すとおりである。

表 4.6: AMT による分類の結果

ピクトグラムである	3555 枚
ピクトグラムでない	1288 枚
その他	661 枚
計	5504 枚

4.3.6 学習と検証

学習は、予備実験と同じく特徴量に画素の標準偏差とコーナー数、識別器に Linear SVC を用いた。まず 5504 枚全てに対して特徴量の計算とラベル付けを行い、次にピクトグラムであるとラベル付けしたものとピクトグラムでないとラベル付けしたのものから 1200 枚ずつランダムに抽出した。特徴量を平均 0、分散が 1 になるように正規化を行ってから 1200 枚のうち 200 枚ずつをランダムに抽出し、合計 400 枚でパラメータのグリッドサーチを行った。結果は表 4.7 の通りである。

表 4.7: Linear SVC の C 値のグリッドサーチの結果

C 値	81.642
識別率	71.8%

次に、得られたパラメータと残りの 2000 枚の画像を用いて学習を行った。学習は、2000 枚をランダムに並び替えた上で一グループ 200 枚とした 10-fold cross validation にて行った。結果、得られた識別率は $71 \pm 7\%$ であり、各回の結果の平均値と標準誤差で表

した混合行列は表 4.8 の通りであった。表の見方は 4.2.4 節と同様である。ピクトグラムであるものを正しく識別出来たケースが最も多く、ピクトグラムである画像をピクトグラムでないと誤判定したケースが最も少ないという結果になった。grafie への実装は、正規化の際に用いたスケーラと作成された識別器を出力し、grafie から呼び出すことで行った。

表 4.8: ピクトグラムである画像 1000 枚, そうでない画像 1000 枚の計 2000 枚の画像を使って 10-fold cross validation を行った際の混合行列。それぞれの要素について 10 回の平均値と標準誤差を示す。

		予測値	
		陽性	陰性
正解	陽性	36.2±2.7%	13.9±3.0%
	陰性	15.9±2.9%	34.2±3.6%

第 5 章

ユーザ評価実験

5.1 実験参加者

grafie のユーザビリティを調査するべくユーザテストを実施した。参加者は男性 5 人女性 2 人の合計 7 人であり、年齢層は 20 歳～25 歳（平均 22.3 歳）であった。参加者を募集する際には、マウスのクリックやキーボードからのタイピングなどの PC の基本操作が行うことが出来る人物という条件を設けた。以下、実験参加者を参加した順に P1～P7 と表記する。

5.2 比較対象

grafie との比較対象として、2.2.3 節で言及した infographics 作成サービスである piktochart を選定した。このサービスを比較対象とした理由として、基本図形やテキストの挿入、図形同士のアライメント支援など infographics 作成に必要な機能を十分備えていること、このサービスを利用したことがないユーザでも直感的に操作を行うことが可能な UI を備えていること、マウスやトラックパッドでも操作がしやすいことなどが挙げられる。piktochart には 14 種類のグラフを生成する機能が備わっており、棒グラフのようなシンプルなグラフからピクトグラムを用いたグラフまで作成を行うことが出来る。piktochart におけるグラフの作成手順は以下の通りである。

1. スプレッドシートにデータを記入する。
2. プレビューを見ながらグラフの種類を選択する。
3. プレビューを見ながら用いる色やピクトグラムの変更を行う。

グラフのプレビューは常にスプレッドシートを参照しており、変更を行った項目があると即時プレビューに反映される。

5.3 実験の流れ

実験の手順を以下に示す。

1. 同意書などの必要な書類の説明を行う。
2. infographics がどのようなものであるかについて説明を行う。
3. 既存の infographic 一枚を見せ、同様の infographic を piktochart と grafie でそれぞれ制作してもらう。(タスク 1)
4. 文章を見せ、その文章を元にして最適だと思うグラフや画像を使った infographic を piktochart と grafie でそれぞれ制作してもらう。(タスク 2)
5. piktochart と grafie それぞれについてシステムユーザビリティスケール (System Usability Scale, SUS)[29] に回答してもらう。
6. 参加者にインタビューを行う。以下にインタビュー項目を示す。
 - 自主学習や絵画塾など、義務教育以外でデザインを学んだ経験の有無、ある場合はどの程度か。
 - デザインを行うソフトウェア (Adobe Photoshop や Adobe Illustrator など) の使用経験の有無、ある場合はどの程度のスキルか。
 - タスク 1 で「使いやすい」と思ったシステムはどちらか、またその理由。
 - タスク 2 で「使いやすい」と思ったシステムはどちらか、またその理由。
 - タスク 1 でより自身の満足度の高い infographic が作成出来たと思うシステムはどちらか、またその理由。
 - タスク 2 でより自身の満足度の高い infographic が作成出来たと思うシステムはどちらか、またその理由。
 - 本実験に関して気になった箇所や質問したいことの有無、あればその内容。

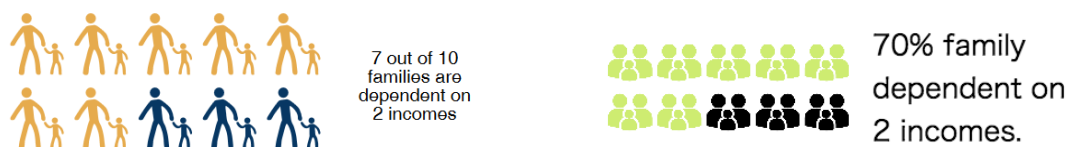
タスク 1 で参加者に見せた infographic は「アメリカの家庭のうち 70% は共働きで生計を立てている」といったデータを表すものである。このタスクにおいては参加者に既存の infographic だけではなく予め piktochart と grafie それぞれで作成した画像も見せている。

これは、どの程度まで似せれば良いのかの一つの指標とするための画像であるが、参加者には必ずしも同じものを制作しなくてもよいと伝えた。また、タスク 2 で使用した文章は The Guardian 紙が提供する Datablog 内の一文を参考に*2、文意を変えない程度に改変した以下の文章である。

*1 <http://www.designinfographics.com/health-infographics/life-insurance-breakdown> (2016 年 11 月 17 日閲覧)

*2 <https://www.theguardian.com/society/datablog/2016/dec/13/europeans-massively-overestimate-muslim-population-poll-shows>

(2017 年 1 月 18 日閲覧)



piktochart で作成した画像

grafie で作成した画像

図 5.1: ユーザに例として提示した作例. 既存の infographic *¹の一部を再作成した.

Australian people's average estimate was that 12.5% of the country was Muslim.

(オーストラリア人に「国民の何 % がイスラム教徒だと思うか」と尋ねたところ、その平均は 12.5% であった.)

英語を得意とす参加者と得意としない参加者との間に文章に対する理解度の差が出てしまうことを防ぐため、文章を見せる際にカッコ内に示した日本語訳も同時に口頭で伝え、意味を理解してもらうようにした。

実験を行う際、piktochart と grafie を使用する順番はタスク 1 とタスク 2 とで異なるようにし、かつ奇数人目の参加者と偶数人目の参加者とでも異なるようにした。実験では MacBook Air 2015 年初期モデル（メモリ 4GB）を使用し、grafie は Google Chrome 上、piktochart は Firefox 上でそれぞれ動作させた。参加者には MacBook Air に付属しているトラックパッドで操作を行ってもらった。

5.4 結果

5.4.1 システムユーザビリティスケール (SUS)

SUS について

SUS(System Usability Scale) とは、参加者に以下に示す 10 個の質問に 1（全くそう思わない）から 5（強くそう思う）の尺度で回答してもらうことによってシステムがどれほど使いやすいのかを定量化する指標である。

1. I think that I would like to use this system frequently.
(このシステムを頻繁に使いたいと思う.)
2. I found the system unnecessarily complex.
(このシステムは無駄に複雑だと思った.)
3. I thought the system was easy to use.
(このシステムは簡単に使えると思った.)

4. I think that I would need the support of a technical person to be able to use this system.
(このシステムを使えるようになるのに専門家のサポートが必要だと思った.)
5. I found the various functions in this system were well integrated.
(このシステムの様々な機能はよく統一されていると思った.)
6. I thought there was too much inconsistency in this system.
(このシステムにはあまりにも多くの矛盾があると思った.)
7. I would imagine that most people would learn to use this system very quickly.
(殆どの人がこのシステムの使い方をすぐに覚えるだろうと思う.)
8. I found the system very cumbersome to use.
(このシステムはとても扱いにくいと思った.)
9. I felt very confident using the system.
(このシステムを使う自信があったと感じた.)
10. I needed to learn a lot of things before I could get going with this system.
(このシステムを使う前に多くのことを学ばなければならなかった.)

これらの質問のうち奇数番号の質問では回答から 1 を引き、偶数番号の質問では 5 から回答を引いた数字を全て足しあわせ、得られた数値に 2.5 をかけた結果がシステムユーザビリティの値である [29]。値は 0 から 100 の間を取り、高い値ほどユーザビリティが高いことを示す。

SUS の比較

結果は図 5.2 に示すように、piktochart が平均 65.0 ± 3.9 点、grafie が 66.8 ± 4.9 点であった。対応のある t 検定を行った結果、piktochart と grafie の SUS の間には差は見られなかった ($t(7) = 0.34, p = 0.74, \text{Cohen's } d = 0.12$)。

5.4.2 マウス操作

実験中にはクリックの回数をトラッキングし、それぞれのタスクにおいてどの程度マウスが操作されているかを調査した。平均値はタスク 1 での piktochart では 92.6 ± 11.7 回、grafie では 82.9 ± 14.3 回であり、タスク 2 での piktochart では 170.4 ± 26.8 回、grafie では 148.9 ± 19.0 回となった。対応のある t 検定を行った結果、タスク 1 ($t(7) = 0.77, p = 0.47, \text{Cohen's } d = 0.29$) においてもタスク 2 ($t(7) = 0.69, p = 0.51, \text{Cohen's } d = 0.26$) においても piktochart と grafie のクリック回数の間には差が見られなかった。

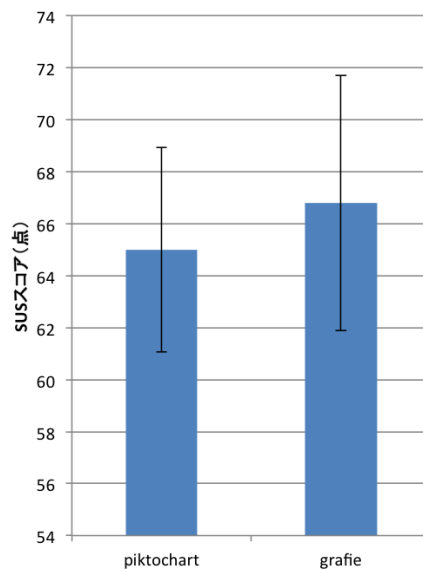


図 5.2: 各システムの SUS の平均と標準誤差

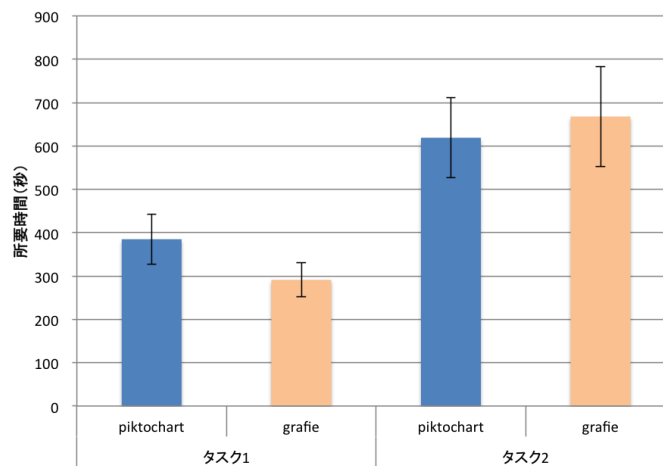


図 5.3: 各システムのそれぞれのタスクにおけるクリック回数の平均と標準誤差

5.4.3 所要時間

各タスクの所要時間を計測した結果を以下の表 5.1, 平均値と標準誤差のグラフを図 5.4 に示す。

対応のある t 検定を行った結果, タスク 1 ($t(7) = 1.86, p = 0.11, Cohen's d = 0.70$) においてもタスク 2 ($t(7) = 0.64, p = 0.55, Cohen's d = 0.24$) においても piktochart と grafie の各タスクに対する所要時間の間には差が見られなかった。

表 5.1: それぞれのシステムにおける各タスクの所要時間 (秒)

参加者		タスク 1	タスク 2
P1	piktochart	470	499
	grafie	253	502
P2	piktochart	226	573
	grafie	172	353
P3	piktochart	337	1080
	grafie	306	1055
P4	piktochart	641	818
	grafie	352	1099
P5	piktochart	374	506
	grafie	206	436
P6	piktochart	195	460
	grafie	262	475
P7	piktochart	451	397
	grafie	487	756

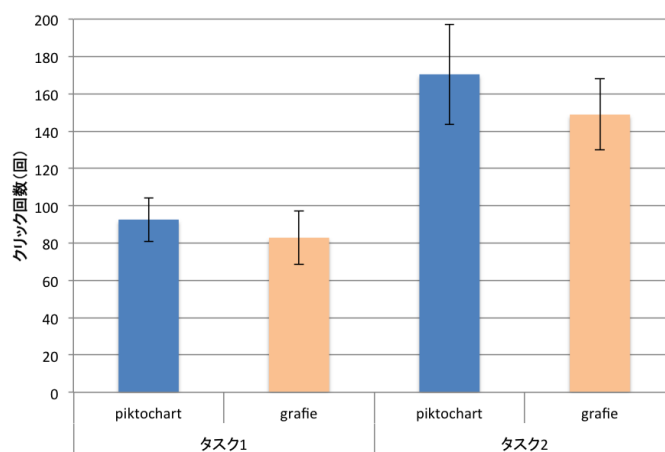


図 5.4: 各システムのそれぞれのタスクにおける所要時間の平均と標準誤差

5.4.4 インタビュー

デザインを学んだ経験

参加者のうち P3, P4, P6 の 3 人はいずれも独学でデザインを学んだ経験があり, P1, P2, P5, P7 の 4 人は学んだ経験が全く無いとのことであった。

ソフトウェアの使用経験

Photoshop の利用経験があると回答したのは P1, P2, P3, P4, P6 で、いずれも簡単な写真加工が行える程度のスキルを持っているとのことであった。Illustrator の利用経験があると回答したのは P2, P3, P4 で、P2 は簡単な図形の作成が行える程度であるが、P3 はイラストを描くためのツールとして、P4 は CAD システムとしてそれぞれ Illustrator を使用していると回答した。その他に使用した経験のあるツールとしては Mac 標準のプレビュー (P1), Inkscape (P3), Microsoft Paint (P7) があった。デザインを行うソフトウェアの使用経験が全く無いと答えたのは P5 のみであった。

最初のタスクで「使いやすい」と思ったシステム

piktochart と回答したのが 2 名 (P2, P6), grafie と回答したのが 4 名 (P3, P4, P5, P7), どちらも同程度と回答したのが 1 名 (P1) であった。piktochart のほうが使いやすいと回答した理由としては「piktochart の機能でデータを変更するとすぐにグラフに反映されるのがわかりやすかった」(P2) や「図形を移動させる際に補助線が出たので (アライメントに) 便利だった」(P6) などの意見が挙げられた。一方, grafie のほうが使いやすいという理由としては「(入力) 文章を打つだけなので楽でいい」(P3), 「グラフに使う画像や数字を文の中から拾えるのがいい」(P4), 「システムが単純なので仕様が把握しやすい」(P5) などが挙げられた。使いやすさは同程度だったという意見の理由は「図の探し方は grafie のほうが良いものの、(検索にかかる) 明確な待ち時間があるのは嫌だった」(P1) といったものだった。

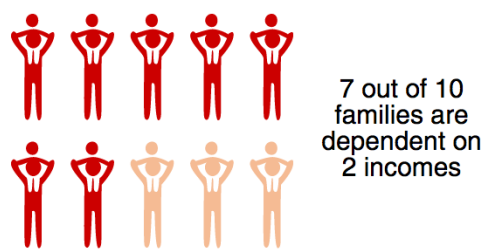
二番目のタスクで「使いやすい」と思ったシステム

piktochart と回答したのが 4 人 (P2, P3, P4, P5), grafie と回答したのが 3 人 (P1, P6, P7) であった。piktochart が使いやすいとした理由としては「グラフの種類が多い」(P2), 「grafie よりも図形の拡大縮小といった機能が多い」(P4, P5, P3) といった機能の豊富さに関連する意見が挙げられた。grafie が使いやすいとした理由には「画像を探す手間が省けているから楽」(P1) 「画像の検索が出来るので良かった」(P7) といったピクトグラム of 検索機能に関する意見や「piktochart は項目が多く複雑なのでシンプルな grafie のほうが楽」(P6) といった意見があった。

最初のタスクでより満足度の高い infographic が制作できたシステム

piktochart と回答したのが 1 人 (P6), grafie と回答したのが 4 人 (P1, P3, P4, P5), どちらも同程度、あるいはどちらにも満足していないと回答したのが 2 人 (P2, P7) であった。grafie で制作した infographic が良いと回答した理由については「piktochart では家族

を表す絵が見つけれなかった」(P1), 「見栄えが良いのが grafie で作成したほうだった」(P3, P4, P5) といった理由が挙げられた. piktochart で制作した infographic が良いと回答した理由には「grafie では色をカラーサークルから選択するが piktochart では何色か予めいい感じの色がパレットに用意されていて, その中から選択したほうが自分の好みの色になった」(P6) といったものがあった. その他の回答の理由としては「どちらのシステムも一つのグラフにつき種類の画像しか使えないという制約があって想像していたものが制作できなかった」(P7) というグラフに対する制約に関する意見があった.



P6 のタスク 2 における
piktochart での作例



P6 のタスク 2 における grafie での作例

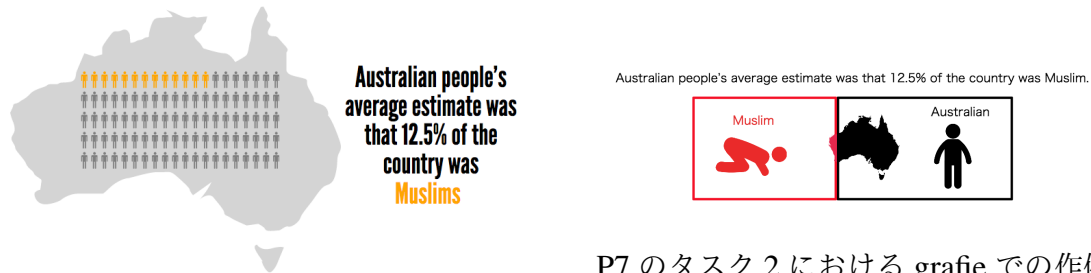
図 5.5: piktochart での満足度が高かった P6 の作例: piktochart での配色は予め定められている 98 色の中から選択するか, RGB カラーコードで直接してはいるかを選択できる. P6 は 98 色の中から選択を行っていた.

二番目のタスクでより満足度の高い infographic が制作できたシステム

piktochart と回答したのが 1 人 (P5), grafie と回答したのが 3 人 (P2, P6, P7), どちらも同程度と回答したのが 3 人 (P1, P3, P4) であった. piktochart と回答した理由は「画像の大きさが変えられないのが不便」(P5) と grafie の機能の少なさを指摘するものがあった一方, grafie と回答した理由には「piktochart では小数点以下を表すグラフが作成出来なかった」(P2, P7), 「piktochart では画像の途中までを塗りつぶすグラフが出来なかった」(P7), 「piktochart では画像の大きさを変更すると画像の中の線の太さまで変わってしまった」(P6) など piktochart の機能では手が届かない点を指摘するものがあった. どちらも同程度と回答した参加者の意見としては「両方ともそれぞれ機能に関してはマイナスの点がある」(P4) とやはりシステムの機能不備を指摘する声があった.

5.5 考察

SUS の値には有意差が見られなかったものの, grafie の SUS を下げている要因としては取得したピクトグラム of 拡大縮小, アラインメント, アンドゥ・リドゥなど補助的な機



P7 のタスク 2 における grafie での作例

P7 のタスク 2 における piktochart での作例

図 5.6: grafie での満足度が高かった P7 の作例:「オーストラリア国民の 12.5%」という情報を表す際に、piktochart では画像の 12.5% だけを塗りつぶすグラフが作成出来ないのが不満点として挙げられた。

能の不足があることがインタビューを通して見えてきた。これらを実装することによって SUS が上がる可能性はある一方、単純に画像を拡大縮小するのでは、piktochart も抱えている図 5.7 のように「画像の大きさを変更すると画像内の線の太さまで変わってしまう」問題をそのまま抱えてしまうことになる。これはラスター形式の画像特有の問題であるた



図 5.7: 大きさの変更により線の太さが変わってしまった例 (P4 のタスク 2 の piktochart で制作されたもの): 吹き出しの左側にある円の枠線の大きさが著しく変化している。

め、現在は png 形式のみを収集している画像の取得を svg 形式などのベクターで表現された画像の取得に切り替えるという解決方法が提案できるが、軽量かつ多くのソフトウェアでの閲覧加工に対応している png 形式の画像と比較すると、svg 形式で収集できる画像の枚数が少ないことが想定される。インターネット上で svg 形式で提供されているピクトグラムの種類及び枚数を調査する必要がある。

各タスクの所要時間にも有意差は見られなかった。grafie では画像の取得に数十秒かかっていることを考慮に入れると grafie での実作業時間はより少ないと考えられる。特にタスク 2 においては画像を複数枚使用する参加者もあり (図 5.8 参照)、画像の検索の度

に数十秒のロードが挟まることを考えるとこれを改善することによってタスクへの所要時間を大きく減らすことが出来ると想定される。現在の grafie の検索システムでは、最も時

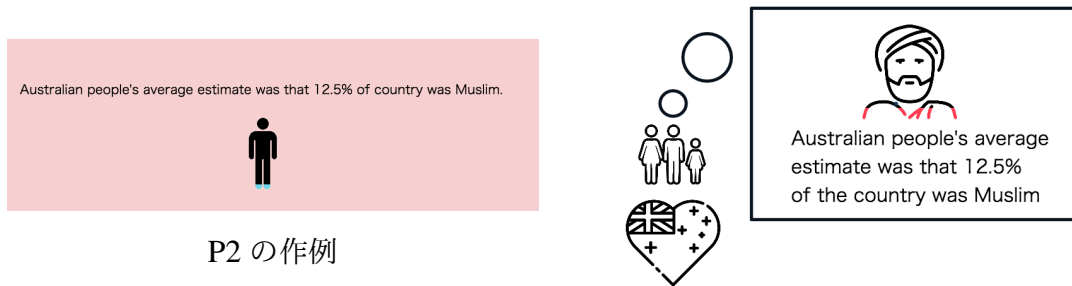


図 5.8: 画像を複数枚使用した infographic の例 (タスク 2 における P4 の grafie での作例): 「Australia」, 「people」, 「balloon」, 「mosque」で計 4 回検索を行っている。

間を要するのは画像の特徴量の計算である。コーナー検出や画素の標準偏差を求めるアルゴリズムの工夫か、あるいは全く別の計算量の軽い特徴量を用いることによって所要時間の短縮に努めたい。この待ち時間はインタビューにおいても grafie の欠点として指摘されており、SUS への影響も考えられる。よく検索される単語を表すピクトグラムは予め grafie 内に保持しておき、ユーザがその単語を検索しようとした時には保持されているピクトグラムを提示し、その中に満足できるものが無かった時だけ検索を行うなどの工夫で時間の短縮を行いたい。

クリック回数には有意差が見られなかったが、参加者によってかなりばらつきが見られた。例えばタスク 2 の grafie でのクリック回数は最小の参加者で 85 回、最大の参加者では 232 回でありおよそ 2.7 倍もの差がある。図形や絵を多く使った凝ったデザインの infographics を制作しようとするとは必然的に作業量が多くなりクリック回数が増えると考えられる (図 5.9 参照)。

興味深い点として、piktochart がより使いやすいと感じた点に関して P6 が「予め定められている色から選択したほうが自分の好みの色になった」と発言していることが挙げられる。これは RGB 形式で扱うことのできる 1677 万色の中から自分で色を選択するという自由度の高さが必ずしもシステムの使いやすさには繋がらないことを示唆している。原因として、調和性の高い色の組み合わせにはある程度の法則があるが、デザインを学んだことのない初心者ではその法則を知らないため、色を適当に選択すると調和性の低い組み合わせになってしまうということが考えられる。システムを発展させていく上では使用する色にあえて制限をかけることも視野に入れていきたい。例えば、ユーザが予め設定したカラーテーマやすでに使われている色、あるいは文章中の単語から連想される色といった情報から調和性の高い色を選択しユーザに提示する機能の実装によってよりユーザビリ



P2 の作例

P4 の作例

図 5.9: タスク 2 においてクリック回数が最小だった参加者 (P2) と最大だった参加者 (P4) の作例. P2 は用いている画像が一枚だけであるが, P4 は画像を複数枚使用した他に図形の挿入も行っている. P2 のクリック回数は 85 回, P4 のクリック回数は 232 回であった.

ティの高いシステムになる可能性がある.

今回のユーザテストでは infographics を制作する過程にのみ注目しているが, infographics はデータを伝えるためのものであるので制作された infographics がどれくらいわかりやすくデータを伝えているかということも今後は調査していきたい. より複雑な infographics を制作してもらい, そこから正しくデータが読み取れるかどうかという実験の設計について考える必要がある.

第 6 章

おわりに

本稿では現代社会でのデータ可視化における infographics の重要性および課題点を述べ、既存の infographics 作成ツールが抱える問題点をインターネット上の画像からピクトグラムのみを抽出する機械学習と文章中に出現する数値とピクトグラムからグラフを生成する機能の実装により解決するシステムの構築について述べた。システムのユーザ評価実験では数値上の有意差は見えなかったが、既存システムよりも製作者自身が満足できる infographics が制作できることが示され、さらなる機能の追加によりユーザビリティのより高いシステムになりうることが示唆された。

現在のシステムでは予め文章化したデータから infographics を制作することしか実現できておらず、infographics 制作過程においては編集作業およびデザインの一部を支援することしか出来ていない。今後は入力された文章の関係性からどのような infographics の形状が良いのかをシステム側で判断してレイアウトの候補を作成して提示したり、ユーザが infographics 内で特に強調したい部分に視線を誘導するような配色を提案する機能など infographics のデザインに関する支援を強化していくことに取り組みたい。また、大量のデータの中から infographics に使用できそうな関連性のあるデータのみを取り出して文章化して提示し、ストーリーの組み立てを支援する機能など infographics 制作プロセスの全般に渡る支援にも挑戦してゆきたい。

本研究では infographics をどうやって制作するかという点に重点を置いていたが、今後の研究課題としてはどのような infographics を制作すべきかという点に着目することが挙げられる。infographics とはデータを伝えるものである以上、読み手が正しくデータを受け取ることが簡単に出来なければならない。読み手が情報を受け取りやすい infographics がどのようなものであるのかについて研究していくことが必要である。

謝辞

本研究を進めるにあたり，終始熱心なご指導を頂いた指導教員の矢谷浩司准教授に深く感謝いたします。また，日々の議論を通じて多くの知識や刺激を頂き，実験の実施の際にも惜しみなく協力をくださった矢谷研究室の皆様，及び相田研究室の皆様に感謝します。さらに，実験の際に参加を快諾してくださった皆様およびに参加者を探す手伝いをしてくださった皆様にこの場を借りてお礼を申し上げます。そして，研究を進めていくにあたり私生活の様々な場面で支えてくださった家族，そして友人の皆様に深謝いたします。

参考文献

- [1] Daniel Keim, Huamin Qu, and Kwan-Liu Ma. Big-data visualization. *IEEE Computer Graphics and Applications*, Vol. 33, No. 4, pp. 20–21, 2013.
- [2] Philip Russom, et al. Big data analytics. *TDWI Best Practices Report, Fourth Quarter*, pp. 1–35, 2011.
- [3] Seref Sagiroglu and Duygu Sinanc. Big data: A review. In *Collaboration Technologies and Systems (CTS), 2013 International Conference on*, pp. 42–47. IEEE, 2013.
- [4] Vitaly Friedman. Data visualization and infographics in: Graphics. *Monday inspiration*, 2008.
- [5] Daniel A Keim. Visual exploration of large data sets. *Communications of the ACM*, Vol. 44, No. 8, pp. 38–44, 2001.
- [6] Ed H. Chi. A taxonomy of visualization techniques using the data state reference model. *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings*, Vol. 94301, No. Table 2, pp. 69–75, 2000.
- [7] Interactive Visualization of Serial Periodic Data.
- [8] Ed Huai-hsin Chi, Phillip Barry, John Riedl, and Joseph Konstan. A Spreadsheet Approach to Information Visualization.
- [9] Christopher Ahlberg, Christopher Williamson, and Ben Shneiderman. Dynamic Queries for Information Exploration: An Implementation and Evaluation. 1992.
- [10] The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus+Context Visualization for Tabular Information.
- [11] Michael Spence, Christian Beilken, and Thomas Berlage. FOCUS: The Interactive Table for Product Comparison and Selection.
- [12] Jeffrey Heer, Stuart K Card, and James A Landay. prefuse: A Toolkit for Interactive Information Visualization.
- [13] Tong Gao, Mira Dontcheva, Eytan Adar, Zhicheng Liu, and Karrie G Karahalios. Datatone: Managing ambiguity in natural language interfaces for data visualization.

- In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pp. 489–500. ACM, 2015.
- [14] Bongshin Lee, Rubaiat Habib Kazi, and Greg Smith. Sketchstory: Telling more engaging stories with data through freeform sketching. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 12, pp. 2416–2425, 2013.
- [15] Bongshin Lee, Greg Smith, Nathalie Henry Riche, Amy Karlson, and Sheelagh Carpendale. Sketchinsight: Natural data exploration on interactive whiteboards leveraging pen and touch interaction. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 199–206. IEEE, 2015.
- [16] Larissa Pshetz, Koji Yatani, and Darren Edge. Turningpoint: Narrative-driven presentation planning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1591–1594, 2014.
- [17] Ryan Eccles, Thomas Kapler, Robert Harper, and William Wright. Stories in geotime. *Information Visualization*, Vol. 7, No. 1, pp. 3–17, 2008.
- [18] Kenneth C Spry. An infographical approach to designing the problem list. *Proceedings of the 2nd ACM SIGHIT symposium on International health informatics - IHI '12*, p. 791, 2012.
- [19] Jeffrey Heer, Fernanda B Viégas, and Martin Wattenberg. Voyagers and voyeurs: supporting asynchronous collaborative information visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 1029–1038. ACM, 2007.
- [20] Jeffrey Heer, Jock D. Mackinlay, Chris Stolte, and Maneesh Agrawala. Graphical histories for visualization: Supporting analysis, communication, and evaluation. In *IEEE Transactions on Visualization & Computer Graphics Volume:14, Issue:6*, pp. 1189–1196, 2008.
- [21] C.G. Healey. Choosing effective colours for data visualization. *Proceedings of Seventh Annual IEEE Visualization '96*, pp. 263–270., 1996.
- [22] Debbie Abilock and Connie Williams. Recipe for an infographic. *Knowledge Quest*, Vol. 43, No. 2, pp. 46–55, 2014.
- [23] Pınar Nuhuğlu Kibar and Buket Akkoyunlu. A new approach to equip students with visual literacy skills: Use of infographics in education. In *European Conference on Information Literacy*, pp. 456–465. Springer, 2014.
- [24] Murray Dick. Interactive infographics and news values. *Digital Journalism*, Vol. 0811, No. January 2015, pp. 1–17, 2013.
- [25] Media Technologies. Information Graphics Design Challenges and Workflow

- Management Marco Giardina, University of Neuchâtel, Switzerland, Pablo Medina, Sensiel Research, Switzerland. *Online Journal of Communication and Media Technologies*, Vol. 3, No. 1, pp. 108–125, 2013.
- [26] B.I.U. Dur. Data Visualization and Infographics In Visual Communication Design Education at The Age of Information. *Journal of Arts and Humanities*, Vol. 3, No. 5, p. 39, 2014.
- [27] Bongshin Lee and Nathalie Henry Riche. More than telling a story: Transforming data into visually shared stories. In *IEEE Computer Graphics and Applications Volume:35, Issue:5*, pp. 84–90, 2015.
- [28] Chris Harris and Mike Stephens. A combined corner and edge detector. In *Alvey vision conference*, Vol. 15, p. 50. Citeseer, 1988.
- [29] John Brooke, et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, Vol. 189, No. 194, pp. 4–7, 1996.

研究業績一覧

国内研究会

- (1) 岩下知美, 矢谷浩司. 「インフォグラフィックスの作成をインタラクティブに支援するシステム」 情報処理学会第 170 回ヒューマンコンピュータインタラクション研究会, pp. 1-7, 2016 年 10 月.