

博士論文 2023 年度

データビジュアライゼーションのための
創造的なデザインプロセスモデルの研究

慶應義塾大学大学院政策・メディア研究科

山邊真幸

論文要旨

データビジュアライゼーションのための 創造的なデザインプロセスモデルの研究

近年、AI技術やデータ解析は様々な分野で活用されている。特に、政府が提唱するデータ駆動型社会は、膨大なデータと機械学習や統計手法を駆使し、様々なシステムの効率化・自動化を推進するものである。一方で、ビッグデータから導き出される分析結果やモデルを使用するAIの精度は、確率やデータの偏りなど、さまざまな要因に左右される。そのため、AIシステムのユーザーはモデルの限界や不確実性を理解する機会が欠かせない。また、大規模なデータセットは様々な変数を含むため、データからどのようなストーリーを導き出すかは、分析者の関心や目的に依存する。そのため分析者の主張だけでなく、膨大なデータの背後にある多様なストーリーを理解することも重要である。

複雑さや多様性を情緒的に表現するデータビジュアライゼーションは、鑑賞者に膨大なデータの背後にあるストーリーを感じさせることができる。こうした作品は、卓越したビジュアルデザイナーやアーティストによって開発されることが多い。しかし、こうした作品のデザインプロセスモデルは明らかになっていない。それだけでなく、熟練者の経験知を学習者に適切に共有する方法は存在していない。

本研究では、学習者や実践者を支援するために、データビジュアライゼーションプロジェクトにおける一般的なデザインプロセスモデルと実践知のパターンを提案した。具体的には、データの専門家とプログラミングスキルを持つビジュアルデザイナーが共同で行ったデータビジュアライゼーションデザインプロジェクトの観察から、デザインプロセスモデルとパターン・ランゲージを作成した。モデルはプロジェクトの概要を示し、各ステップの役割や関係性を理解するのに役立つ。また、パターンは生じやすい問題を記述するための共通のボキャブラリーとして機能する。さらに、実践者向けにモデルやパターンを解説した冊子「データビジュアライゼーション・デザインパターン」を作成した。

デザインやアートを学ぶ学生を対象として専門家と一緒にデータビジュアライゼーションを開発するワークショップと、実務でデータビジュアライゼーションに携わるデザイナーへのインタビューを行い、プロセスモデルやパターンの影響を調査した。その結果、参加者はデザインプロセスの進め方を理解し、パターンを参照することでプロジェクトにおける問題を自らの工夫によって解決できることが示された。

キーワード：データビジュアライゼーション, 情報デザイン, コミュニケーション, デザインプロセス, コンピュータショナルデザイン

慶應義塾大学大学院 政策・メディア研究科
山邊真幸

The research on a model of the creative design process for data visualization

Abstract

In recent years, AI technology and data analysis have been employed in various fields. Specifically, the government's vision of a data-driven society promotes using vast amounts of data, machine learning, and statistical methods to increase efficiency and automation in various systems. However, the accuracy of AI systems that rely on big data-derived analytical results and models depends on various factors, including probability and data bias. Therefore, users of AI systems must have the opportunity to understand the limitations and uncertainties of the models. Additionally, in large data sets containing a variety of variables, what story is derived from them depends on the analyst's interests and objectives. Therefore, it is important to understand the analyst's views and the diverse stories behind the vast data.

Data visualizations that emotionally express complexity and diversity can make viewers feel the stories behind vast amounts of data. Some exceptional designers or artists develop such remarkable works. However, no general design process model has been proposed. Nor is there a way to appropriately share the experiential knowledge of skilled workers with learners.

To support a learner or a practitioner, this study proposed the general design process model and patterns of practical knowledge on data visualization projects. Specifically, we developed the model and pattern languages based on observations of actual data visualization design projects carried out by a data expert and a visual designer with programming skills. The model provides an overview of the project and helps understand each step's role and relationships. And the patterns are functional as a common vocabulary to describe recurring conflicts. In addition, we produced a booklet entitled "Data Visualization Design Patterns" that explains the models and patterns for practitioners.

We conducted a workshop for design and art students to develop data visualizations with data experts and an interview for an experienced designer to investigate the impact of process models and patterns. The results showed that participants understood how to proceed with the design process and were able to solve problems in their projects through their devising by referring to the patterns.

Keywords : Data Visualization, Information Design, Communication, Design Process, Computational Design

Graduate School of Media and Governance, Keio University
Masaki Yamabe

目次

第 1 章	序論	1
1.1	はじめに	1
1.2	本研究の背景	2
1.3	情報可視化とデータビジュライゼーション	3
1.4	可視化の感情的な価値	6
1.5	社会で情報を共有するデータビジュライゼーション	7
1.6	データビジュライゼーションとアート	9
1.7	情報可視化とグラフィックデザイン	10
1.7.1	科学革命と図の利用	11
1.7.2	統計学の興隆とデータ	13
1.7.3	統計グラフとプレゼンテーション	14
1.7.4	交通や物流の発展に伴う地図とデータの融合	16
1.7.5	市場経済の成熟と反動－図像統計とアイソタイプの誕生	19
1.7.6	グローバリズム隆興とメディアの発展	20
1.7.7	情報可視化の公共化・大衆化－ビジュアルコミュニケーションとしての情報可視化	23
1.7.8	動的な情報可視化の登場－対話的インターフェイスの研究	25
1.7.9	データから人や社会の動きを映し出すアート、デザインの登場	26
1.8	本研究の目的	28
第 2 章	関連研究	29
2.1	はじめに	29
2.2	情報可視化の基礎的研究	30
2.3	参照モデル	31
2.4	システムとしての情報可視化の設計	33
2.5	コミュニケーションメディアとしての情報可視化の設計	36
2.6	情報可視化システムと美観に関する研究	38
2.7	情報可視化と芸術性に関する研究	39
2.8	情報可視化とグラフィックデザインに関する研究	41
2.9	美観を伴う情報可視化デザイン手法の研究	44
2.9.1	Computational Information Design	44

2.9.2	CID の事例と作品の特徴	45
2.9.3	CID のプロセス	53
2.9.4	CID の影響と現在の状況	57
2.10	本研究の位置付け	58
第 3 章	本研究の手法	60
3.1	本研究手法の概略	60
3.2	データビジュアライゼーションプロジェクトの調査	61
3.3	プロセスの分析	62
3.4	形式知化	62
3.4.1	プロセスのモデル化	62
3.4.2	パターン・ランゲージによる形式知化	62
3.4.3	データビジュアライゼーション・デザインパターンの作成	63
3.5	運用と検証	63
第 4 章	プロセスの分析	64
4.1	分析の目的	64
4.2	対象とするプロジェクト	65
4.2.1	新型コロナウイルスゲノム系統樹の 3 次元可視化の概要	65
4.2.2	プロジェクトの推進方法及びコミュニケーション手段について	67
4.3	分析手法	67
4.3.1	分析対象とするデータ	67
4.3.2	分析手順	67
4.3.3	新たに定義したカテゴリーについて	68
4.4	プロジェクトの観察	69
4.4.1	プロジェクトの全体像および前提	69
4.4.2	フェーズ 1：第 1 回公開までの記録	70
4.4.3	フェーズ 2：第 2 回公開までの記録	78
4.4.4	フェーズ 3：第 3 回公開までの記録	85
4.4.5	フェーズ 4：第 4 回公開までの記録	93
4.4.6	フェーズ 5：第 5 回公開までの記録	100
4.5	考察	119
4.5.1	各カテゴリーの相互関係に見られる特徴と役割	119
4.5.2	視覚造形の試行錯誤と探索用インタラクション	120
4.5.3	視覚造形の探索的な進展例	122
4.5.4	プロセスのモデル化	126
4.5.5	モデルによる事例の分析	128
4.6	まとめ	130
第 5 章	パターン・ランゲージの作成	131
5.1	パターン・ランゲージについて	131

5.2	パターン・ランゲージの制作プロセス	133
5.3	本研究におけるパターン・ランゲージの制作プロセス	135
5.3.1	パターン・マイニングにおけるパターン・ダイアログについて	135
5.3.2	パターン・マイニングにおけるクラスタリングについて	136
5.3.3	パターン・マイニングにおけるシード・メイキングについて	137
5.3.4	パターンの再構造化	140
5.3.5	パターン・ライティングについて	141
5.3.6	パターン・シンボライジングについて	142
5.4	開発したパターン・ランゲージ	143
第 6 章	データビジュアライゼーション・デザインパターン	145
6.1	制作の目的	145
6.2	想定読者	146
6.3	構成	147
第 7 章	ワークショップでの運用と考察	148
7.1	はじめに	148
7.2	ワークショップの概要	148
7.2.1	参加者の概要	148
7.2.2	ワークショップのスケジュール	149
7.3	ワークショップの実施計画	149
7.3.1	目的の提示	150
7.3.2	テーマ設定	150
7.3.3	制作の方向づけ	150
7.3.4	レビューの方法	151
7.3.5	データ	151
7.3.6	参考資料の提供	152
7.3.7	レビューに専門家を招聘	152
7.3.8	参加者のコミュニケーションツール	153
7.4	記録方法	153
7.4.1	事前アンケート	153
7.4.2	事後アンケート	154
7.4.3	対話の記録	154
7.5	制作事例	154
7.5.1	事例 1: 参加者 Y の作品	156
7.5.2	事例 2: 参加者 U の作品	158
7.5.3	事例 3: 参加者 M の作品	160
7.5.4	事例 4: 参加者 K の作品	162
7.5.5	事例 5: 参加者 N の作品	164
7.5.6	事例 6: 参加者 O の作品	166
7.5.7	事例 7: 参加者 T の作品	168

7.6	アンケート結果と考察	171
7.6.1	事前・事後回答の比較	171
7.6.2	事前・事後の回答から見た変化	172
7.6.3	各パターンについての参加者の感想	174
7.6.4	制作中にパターンを参照する機会が少なかったケース	175
7.6.5	パターンに記述した状況ではない状況でパターンを活用したケース	176
7.6.6	専門家を交えたレビューについての言及	177
7.6.7	ワークショップの構成についての言及	178
7.7	まとめ	178
第 8 章	実務での運用と考察	180
8.1	はじめに	180
8.2	前提と手法	180
8.3	制作したデータビジュアライゼーション	181
8.4	アンケート結果と考察	184
8.4.1	事前・事後回答の結果	184
8.4.2	事前・事後回答の比較と考察	185
8.4.3	自由記述での回答	186
8.4.4	自由記述での回答の考察	187
8.5	作業の記録と考察	187
8.5.1	パターンを活用したエピソード	187
8.5.2	エピソードの考察	189
8.6	まとめ	191
第 9 章	結論	192
9.1	本研究のまとめ	192
9.2	今後の課題と展望	194
9.2.1	実践の機会創出と認知の向上	194
9.2.2	創造的なプロセスモデルはどうあるべきか	194
9.2.3	データビジュアライゼーションの良い質とは何か	196
9.2.4	データビジュアライゼーションを社会に位置づけるために	199
9.3	おわりに	203
	参考文献	207
	付録 A 付録	214

図目次

1.1	DIKW モデルの概念図 (Shedroff のモデルを参考に筆者作図)	4
1.2	検索の傾向から見た「Data Visualization」と「Information Visualization」の人気の変化	6
1.3	データ・ジャーナリズムの例、朝日新聞社「見えない交差点」 [7]	7
1.4	インフォグラフィクスとストーリーテリングの例、National Geographic「These are the best and worst states for women」 [8]	8
1.5	ダッシュボードの例、内閣府「V-RESAS」 [9]	9
1.6	バビロニア地方出土の世界図 (大英博物館蔵)	11
1.7	ロバート・プロットによる 1684 年の気圧計のプロット (左) と日毎の風と天候の記録 (右)	13
1.8	The Commercial and Political Atlas: (右) 1700 年から 1780 年までのイギリスにおける対デンマーク・ノルウェーの貿易収支 (左) 1780 年のスコットランドにおける対外貿易収支。黒の矩形が輸出額を斜線の矩形が輸入額を示す	15
1.9	Statistical Breviary: ヨーロッパ各国の人口 (左の棒の高さ)、税収額 (右の棒の高さ)、国土の面積 (円の面積) の可視化を試みている。この図では円を用いて現代のパイチャートのような割合表現、直径による総量の表現、ベン図のような表現などのバリエーションが見られる	15
1.10	Diagram of the causes of mortality in the army in the East と題されたダイアグラム。報告書『Notes on Matters Affecting the Health, Efficiency, and Hospital Administration of the British Army』の中で使用された。	16
1.11	ミナールによるディジョン-ミュルーズ間の流動地図、濃い線幅は通過客の数を表す (1845 年)	17
1.12	ミナールによる徒歩でパリに出荷された食肉用家畜の統計地図 (1858 年)	18
1.13	『社会と経済』 (左上) 世界の人口及び民族の構成 (左下) ユーラシア大陸の森林面積 (右上) 世界の栽培地面積 (右下) 人口 25 人あたり大都市圏に住む人数の比較	20
1.14	「Dimaxion World Map」: (上) 1952 年 Shoji Sadao と Fuller による正 20 面体版。平均気温が着色されている (左下) 1944 年図法特許申請のための図解 (右下) 1953 年 Herbert Bayer により「World Geo-Graphic Atlas」に使用された図。正方形と正三角形の区画からなるオリジナル版に Fuller が提唱した独自のエネルギー生産指標が可視化されている (図版引用 [22])	22

1.15	「Understanding USA」、それぞれのページのテーマが問いの形式で右上に記されている：(左上) 人口 (右上) 社会保障 (左下) 選挙制度 (右下) 出版と同時に公開されたインタラクティブなウェブ版の「歳出と歳入」	24
1.16	(左上) 木構造 (ファイルシステムなど) を 2 次元空間の入れ子の矩形で表す「TreeMap」 [24] (右上) 動的な映画検索システム「Film Finder」 [25] (左下) 複雑なネットワーク構造を 3 次元平面にマッピングした「H3」 [26] (右下) プログラムのソースコードなど長いテキストファイルを抽象化して可視化する「SeeSoft」 [27]	25
1.17	(左) 株式市況を伝えるウェブサイトで (finviz.com) 「Treemap」 を使用している例 (右) ソースコードエディタ (VSCode) と併置してソースコード全体を表示させる「SeeSoft」 のようなマップ	26
1.18	(左上) 「Web Hopper」 [32] (右上) 「SHISEIDO 30-YEAR ANNIVERSARY POSTER」 (左下) 「History Flow」 (右下) 「Flight Patterns」	27
2.1	Reference Model のダイアグラム ([42, p.17] より引用)	32
2.2	情報可視化の用途。伊藤貴之. 意思決定を助ける 情報可視化技術, p. 6. コロナ社, 2018 より引用 [44]	33
2.3	Nested Model の入れ子構造。Tamara Munzner. Visualization analysis and design. p66. CRC press, 2014.[48] より引用	35
2.4	Kirk の「Data Visualization Design Process」 [49, p.31] より引用	36
2.5	Domain model for information aesthetics[11] より引用	40
2.6	Lau らによる Categories within the model of information aesthetics [11] より引用	41
2.7	「Information Landscape」 のイントロダクション [54]。キネティックな 3D タイポグラフィとズームリングによるシームレスな表現	43
2.8	VLW に所属した Earl Rennison による「Browsing The Net」 [54]	43
2.9	VLW に所属した Lisa Strausfeld による「Financial Viewpoints」 [55]	44
2.10	CID による Fry の作品群 [56, pp.133–161] : (A).Handheld Genome Browser Prototype, (B).Genome Valence, (C).Chromosome 21, (D).Strippy	46
2.11	CID による Fry の作品群 [56, pp.133–161] : (E).Axonometric Introns and Exons, (F).Haplotype Lines	47
2.12	CID による Fry の作品群 [56, pp.133–161] : (G).Revisionist, (H).Dismap	48
2.13	CID による Fry の作品群 [56, pp.133–161] : (I).Redesign of gff2ps, (J).isometricblocks	49
2.14	CID による Fry の作品群 [56, pp.133–161] : (K).tendrill, (L).shopactivity, (M).Deconstructulator, (N).radialsize	50
2.15	Lau らの分析による作品 B,C,J の分布 [11] より引用 (一部筆者加筆)	52
2.16	CID の 7 つのステップと研究分野 [56, p.13] より引用	54
2.17	ステップのフィードバック [56, p.30] より引用	55
2.18	ゲノムデータプロジェクトで開発されたソフトウェアのビュー [56, pp.82–85] より引用	56

2.19	ゲノムデータプロジェクトでのフィードバック [56, p.86] より引用	56
3.1	本研究の流れ	61
4.1	新型コロナウイルスゲノム系統樹の3次元可視化	65
4.2	nextstrain.orgでのゲノム系統樹可視化状況(2020年7月)	70
4.3	筆者による初期スケッチ	71
4.4	フェーズ1のタスク遷移図	73
4.5	フェーズ1の放送時の映像①	76
4.6	フェーズ1の放送時の映像②	77
4.7	フェーズ1の放送時の映像③	77
4.8	フェーズ1のタスクネットワーク図	78
4.9	フェーズ2のタスク遷移図	80
4.10	フェーズ2のセミナーの様子①	83
4.11	フェーズ2のセミナーの様子②	83
4.12	フェーズ2のセミナーの様子③	84
4.13	フェーズ2のタスクネットワーク図	85
4.14	フェーズ3のタスク遷移図	88
4.15	フェーズ3の放送時の映像	92
4.16	フェーズ3のタスクネットワーク図	93
4.17	フェーズ4のタスク遷移図	96
4.18	フェーズ4の展示の様子(上)スクリーンショット(下)設営された展示	99
4.19	フェーズ4のタスクネットワーク図	100
4.20	フェーズ5(VisE)のタスク遷移図	106
4.21	フェーズ5(VisE)のタスクネットワーク図	110
4.22	フェーズ5(VisI)のタスク遷移図	111
4.23	フェーズ5の放送時の映像(VisI)	113
4.24	フェーズ5の放送時の映像(いくつかの流行株の後でデルタ株(インド株)が出現する様子)	114
4.25	フェーズ5(VisI)のタスクネットワーク図	115
4.26	フェーズ5(VisQ)のタスク遷移図	116
4.27	フェーズ5の放送時の映像(VisQ、UAEからの渡航者検疫ウイルスと同じ系統の国内ウイルスが検出される様子)	118
4.28	フェーズ5(VisQ)のタスクネットワーク図	119
4.29	最終的に実装された、すべての探索用インターフェイス	121
4.30	探索用インタラクションにより、同じデータから豊富なバリエーションをデザインできる	121
4.31	最初期の時間と空間の表現	122
4.32	日本へ到達する系統だけの表示に切り替え比較する	123
4.33	スライダーの操作に応じて時間軸を移動する地図面	123

4.34	排除された系統の透明度を調整するスライダー	124
4.35	放送時の映像に反映された表現	124
4.36	開発したアルゴリズムの概略	125
4.37	バンドリングの効果	125
4.38	DEXAS の 3 パート	127
4.39	DEXAS における各パートとタスク	127
4.40	DEXAS における各パートの進行	128
4.41	フェーズ 1 のプロセスを DEXAS で整理した	129
4.42	プロジェクト全体の DEXAS モデル	130
5.1	パターン同士の意味的構造、Takashi Iba and Taichi Isaku. A pattern language for creating pattern languages: 364 patterns for pattern mining, writing, and symbolizing.[73] fig.7 より引用	133
5.2	カードの記入例、破線部がカテゴリーを表す	136
5.3	クラスタリング後のグループ	137
5.4	「チームの立ち上げ」グループ	138
5.5	「再現性の確保」グループ	138
5.6	シードメイキングにおけるボトムアップの構造化	138
5.7	シードメイキングにおけるトップダウンの構造化とボトムアップの構造化の重ね合わせ	140
5.8	再構造化されたパターン	141
5.9	パターンの記述例	142
5.10	パターンの体系	143
7.1	Y の作品（蔓延防止等重点措置発令下の都市とそうでない都市）	157
7.2	Y の作品（地方都市から大都市をのぞむ構図）	157
7.3	U の作品（2D 版）	159
7.4	U の作品（3D 版）	159
7.5	M の作品（関東）	161
7.6	M の作品（近畿）	161
7.7	K の作品（全国を俯瞰）	163
7.8	K の作品（細部の状態）	163
7.9	N の作品（ワクチン接種率が米国などで高まる時期）	165
7.10	N の作品（ワクチン接種率が高まるも再び感染者数が増える時期）	165
7.11	O の作品（全景表示）	167
7.12	O の作品（細部の表示）	167
7.13	T の作品（2021 年 1 月頃を表示）	169
7.14	T の作品（2022 年 1 月頃を表示）	169
7.15	T の作品のドットにマウスオーバーした状態（上：2020 年 1 月頃の変異ウイルス、中：同 2021 年 8 月、下：2021 年 9 月）	170

7.16	設問 1 について事前・事後の回答と比較	172
8.1	N のプロトタイプ	182
8.2	緯度、経度、時間の範囲を指定して該当部分のみ表示した	183
8.3	特定の記事をクリックして絞り込んだ状態	183
8.4	設問 1 について事前・事後の回答と比較	185
8.5	N が制作した初期スケッチ	189
8.6	プロットの色に多様な変数を割り当てている（ピンク：記事作成日、シアン：施設お気に入り登録日、黄色：記事お気に入り登録日、白：閲覧日）プロットにマウスを当てると詳細情報が表示される。	190
8.7	1 週間ごとに重ねて表示し曜日ごとの変化を観察している	190
9.1	論文用の作図とデータビジュアライゼーションの比較。上段が専門家による作図。画像出典：左：Abundance of non-conservative microplastics in the upper ocean from 1957 to 2066[84] / 中：Characterizing the Anti-Vaxxers' Reply Behavior on Social Media[85] / 右：COVID-19 時空間発生マップ [86]	202
9.2	データビジュアライゼーションの社会での位置づけ	203

表目次

4.1	公開日と公開形態	69
4.2	フェーズ 1 のタスク一覧	74
4.3	フェーズ 2 のタスク一覧	81
4.4	フェーズ 3 のタスク一覧	89
4.5	フェーズ 4 のタスク一覧	97
4.6	展示されたムービーモードのテロップ	99
4.7	フェーズ 5 (VisE) のタスク一覧	107
4.8	フェーズ 5 (VisI) のタスク一覧	112
4.9	フェーズ 5 (VisQ) のタスク一覧	117
5.1	パターン一覧	144
9.1	映像表現について回答人数とパーセンテージ	198
9.2	視覚要素についての回答人数とパーセンテージ	198
9.3	印象に残った内容について回答人数とパーセンテージ	198

第1章

序論

1.1 はじめに

近年、情報技術の進展によって大規模なデータの収集や分析が可能となった。大規模なデータは高度な分析や AI、機械学習等で活用されており社会に広く浸透しつつある。情報可視化もまたデータを分析する過程や、得られた結果の周知において広く利用されている。中でデータビジュアライゼーションは、大規模で複雑なデータを可視化し、映し出される対象の複雑さや多様な解釈のあり方を提示するものであり、データの大規模化に歩みを合わせるように広く応用が進んでいる。データビジュアライゼーションは専門家が高度なデータ分析を行う際のツールとしてだけでなく、アーティストやクリエイターによって大規模なデータを介して表現するアートやグラフィックとしても世界的にも多くの作品が作られている。一方、その制作手法は作家の属人的スキルと考えられており、そのプロセスは形式化・体系化されてこなかった。本研究の主題はこのような作品としてのデータビジュアライゼーションの制作プロセスを明らかにし、クリエイターが利用できる形式知化を行うことである。

序論では本研究の背景、主題、対象を明確にする。まず、今日の我々を取り巻くビッグデータとそれを利用するデータ駆動型社会のあり方について、公平性の観点から本研究の背景を述べ、データの活用に際して行われる情報可視化とデータビジュアライゼーションの関係を整理する。次に、本研究で注目するデータビジュアライゼーションの情緒的な価値について述べた上で、社会の中での情報伝達やコミュニケーション、アートとしてのデータビジュアライゼーションについて概観する。さらに、近代以降今日に至るまでの情報可視化とグラフィックデザインの文脈を社会的背景の変化を踏まえて概観し、最後に本研究の目的を述べる。

1.2 本研究の背景

情報技術の進展によって、現代ではかつてない大規模なデータの蓄積と分析が可能となった。機械学習やデータサイエンスなどを通じてデータの利用が促進され、私たちの日常はデータによって便利さを増している。

収集された大規模なデータ、いわゆるビッグデータを分析することで、複雑な課題を解いたり最適化を目指す行為はこれまでのように科学の諸分野にとどまらず、行政や民間など実用の分野へも広く浸透しつつある。一例として、我が国の科学技術基本政策においては、IoT とデータサイエンスを組み合わせ、データ駆動型社会を実現し持続可能な発展と人間中心的な社会の実現を「Society 5.0」の基本コンセプトに掲げる。

一方で、複雑で膨大なデータを分析して推論したり、モデルを導くことができるのは、高度に専門的な知識を持つ人々に限られる。そもそも大規模なデータの全容は人間の知覚を通して理解することが難しく、統計的知識や緻密な分析作業、あるいは深層学習などコンピュータの計算力を通して理解できる。そのため、分析結果や AI などのモデルに準拠したサービスを使う一般の人々が、そのプロセスを理解することは困難であるといえる。

日常生活に関係する物事の予測や分類にさまざまな AI が普及し始めているなか、それらの精度が高く感じられたとしても、信頼に足るのか、正当な評価をしているのかという利用現場における不安を払拭する確かな方法が提供されることは極めて稀である。また、いくらモデルを説明されたとしても、元のデータの複雑な有り様や多様な解釈の幅を俯瞰的にみて、利用現場でその妥当性を判断することは困難である。従って、一般のサービス利用者の立場としてはモデルや分析主体を信用するしか方法がない。

このような不安はデータ独占への批判にもつながっている。各種 IT サービスを提供するテクノロジー企業は、膨大なデータを集め分析する高い能力を持つ一方、消費者は個人の属性や行動パターンをデータとして提供するという、情報の非対称が拡大していることが背景にある。その問題を緩和しようと、データ利用の法的整備が国際的に進んでいる。企業が個人の同意を得ずに情報収集することに歯止めをかけることが目的である。こうした動きは、企業間競争の公正を保ち、個人が意図しないデータ提供を防ぐという意味で人々の不信感を解消する一定の効果は期待できる。しかし、企業や組織のデータ収集と蓄積が何を目的にしているのか、それによって個人あるいは社会にどんな利益をもたらそうとしているのかを理解可能にする効果はあまり期待できない。前述のデータ駆動型社会の問題点を唐沢らは、AI を含む膨大なデータ分析の結果をどのような納得して受け入れるか心理的なメカニズムについての考察や、データ化される自己と実空間の自己のつながりの

中で主体性をどのように担保するのかという課題についての考察が不十分であると指摘する [1]。

これはすなわち、データ駆動型社会を実現するにあたって、データを安全に集めることや、公正に利用することに努めるだけでは不十分であることを示唆している。それよりも本質的な問題は、複雑で大規模なデータについて一般市民とデータ分析を行う者の間にあ

る社会的なコミュニケーションの不足にあると言えよう。それらが不安や不信感を抱かせ、やがて利用者はただモデルに従うだけで、真実を見通せない不可視な社会を生きることを余儀なくされうる。

このように、社会における情報の総量と処理能力が著しく向上したにもかかわらず、我々が接する日常の体験においては、必ずしも見通しがよくなったとは言い難い。むしろ情報が非対称となり、不可視な部分が拡大する世界を生きていると言える。

1.3 情報可視化とデータビジュライゼーション

本研究が対象とするデータ・ビジュライゼーションについて述べる。またデータ・ビジュライゼーションに近い言葉として情報可視化 (Information Visualization) も広く普及している。ここで、このふたつについて説明する。

言葉としてのデータビジュライゼーション (以下 Data-Viz) と情報可視化 (以下 Info-Viz) は、ほとんど同じ意味で使用される場合が多い。確かに形式的に見れば、Data-Viz、Info-Viz のいずれも、記録された数値を視覚的な要素に変換して表現する点で同じである。視覚的な表現は、人間の視覚認知機能の働きを利用する。この効果により、数値の羅列で表現される場合に比べて、視覚的に表現されたものは数値同士の比較やパターンの把握を効率化できる。両者の性質の違いは、データから意味のある情報を取り出して新たな知見を導く流れの中で見ると明確になる。ここでは Nathan Shedroff が提唱し、データを普遍的な知恵へと導く過程を抽象化した「DIKW モデル」 [2] からその違いを確認する。



図 1.1 DIKW モデルの概念図 (Shedroff のモデルを参考に筆者作図)

図 1.1 は DIKW モデルの概念図である、データが知恵に変わる過程において最も原始的な状態であるデータは、ただ集められた数値の集合に過ぎずそれ自体は特に意味を持たない。しかし、分析者が自身の関心に従ってデータから特定の数値を整理したり取り出した時に初めて意味を持つ情報に変わる。そしてその情報を分析によって精錬し知識へと昇華させ、最終的に普遍的に利活用できる知恵を得る。

このように、データと情報の違いには分析者の関心の所在が関連している。分析者の関心が異なればデータから取り出される情報もまた異なる。データには多様な意味が内在しており、分析者の関心に依存して取り出される情報は異なるため、データと情報は一意には一致しない。データの規模や変数が多いほどその傾向は強い。

Data-Viz はこのプロセスにおける初期段階、すなわちデータを情報化する段階で行われる。分析者は何か関心を持ちながらも、まずデータの中からいくつかの変数に絞って可視化したり、その組み合わせを変えるなどして、さまざまな方向からデータに内在する情報を探索する。この際行われるのが Data-Viz である。一方、Info-Viz はプロセスの中盤以降すなわち、情報化以後の段階で行われる。分析者は自身の関心に沿って必要なデータが整理され情報として取り出された後で、その意味を読み解き、分析した結果を踏まえた考察を第三者へ周知する。この際行われるのが Info-Viz である。

Data-Viz と Info-Viz の性質の違いを端的に述べれば、分析者が持っている関心に基づいた論理性や主張とその表現がどれだけ密接に結びついているかの度合いに違いがあると言えるだろう。Data-Viz の場合、まだ分析者にとってデータの素性が掴めず、異常値の発見、分布の把握、変数間の大雑把な相関を見るような状況である。このような状況での可視化を含む一連のデータ分析は「探索的データ分析 (Exploratory Data Analysis:

EDA)」と呼ばれる。EDA を通じて、分析者はデータ全体の中から自身の関心を満たす特徴を持つ部分を選択的に対象とすることができる。

一方、Info-Viz の場合、表現される数値は既に分析者の関心を軸としたものに絞られており、その表現方法は分析者の関心や主張を裏付けるものとして、読者を明瞭に了解させる内容である必要がある。Data-Viz はデータの全体像を俯瞰し、内在する多様な情報を吟味することに主眼を置くのに対して、Info-Viz は分析者の目的に沿って明瞭闊達に特定の情報や意味を浮き彫りにすることに主眼を置く。このような違いから、特に大量のデータをインタラクティブな操作によって表示を切り替え、多面的な視覚表現を行うものを Data-Viz と呼ぶ傾向が見られる。ただし、用語としての Data-Viz と Info-Viz の使い分けは一般に定着しているとはいえず、データ分析や統計学の分野では EDA も含めて Info-Viz と呼ぶことがある。

近年、オープンデータ化の普及や、ビッグデータ利活用の普及と合わせ、変数や行数が大きいデータセットの普及と流通が進み、大規模な生データの素性或信頼性を確認する上でも Data-Viz の重要性は増している。また、データを分析した結果よりも、データそのものを可視化して意外性や多様性を表出させることで話題性を作り出し、読者への問題提起や関心の惹起を図る表現として作成される Data-Viz も見られる。これらは、Info-Viz 的な役割を持つ Data-Viz と考えられる。図 1.2 は Google trends^{*1}における過去 10 年間の「Data Visualization」と「Information Visualization」の検索人気度の推移をグラフ化したものである^{*2}。Info-Viz が横ばいであるのに対して Data-Viz が年々上昇している。これらから、Info-Viz と Data-Viz は重複する部分があるものの、Data-Viz の汎用的な応用例が広がり、人々の関心が高まっていると考えられる。

*1 <https://trends.google.co.jp/>

*2 Google 社によれば人気度とはグラフ上の最高値を基準として相対的に表すものである。100 はそのキーワードの人気度が指定された期間内で最も高い状態を示している。このグラフでは地域を指定せず比較した。

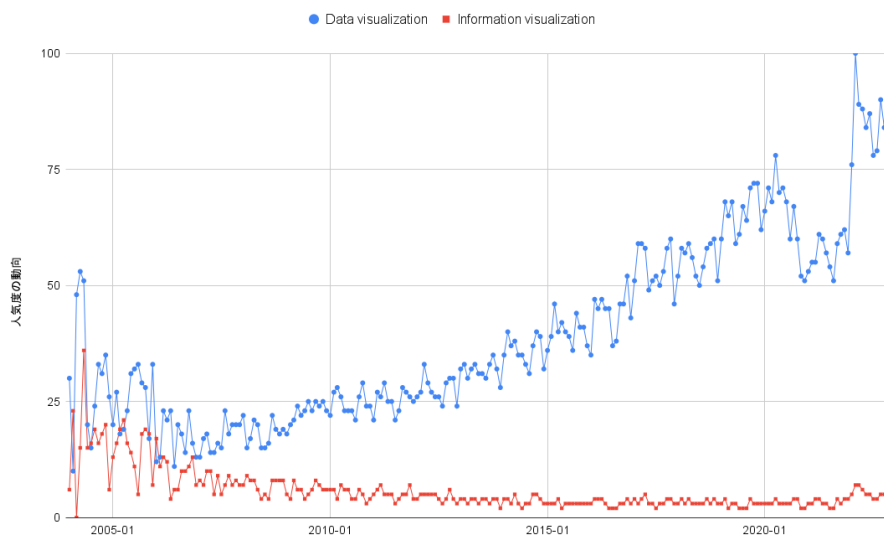


図 1.2 検索の傾向から見た「Data Visualization」と「Information Visualization」の人気の変化

1.4 可視化の感情的な価値

Data-Viz や Info-Viz を評価する際、受け手にとって正確に理解できか否か、あるいは効率的に理解できるかといった機能性に着目する方法がある。しかし、そうした機能的価値だけでなく可視化表現の感情的な作用も注目されている。Yung らは、可視化研究者の間では、機能面での分析的価値が最も確立された価値指標である一方、人間は分析的な存在ではなく、感情的な存在であるという認識が生まれ、他の研究者は異なる解釈を提供し始めていると述べている [3]。Cawthon らは、利用者テストの結果から、情報可視化から知覚される美的感覚と検索タスクの効率には相関があることを示している [4]。可視化研究者のトップカンファレンスの一つである IEEE VIS では毎回 Arts Program を実施している。同プログラムは、IEEE VIS の正式な関連イベントとしては最大のものであり、可視化に関するインタラクティブアート、デザインプロジェクト、新規性のある可視化手法等についての論文や展示を募集し、可視化研究者、メディア・アーティスト、デザイナーらが新しい可視化の価値について議論する場となっている [5]。全世界で約 2 万人以上の会員を有し可視化の普及や制作者間の情報共有に貢献している Data Visualization Society が開催する IIB Awards では、審査指標として、正確性や利便性の他に、感情に働きかける審美性や話題性、創造的なセンスを掲げている。逆に、過剰に分析的なものや、

可視化されるデータに変数が少なく 1,2 次元程度の意味しか持たないものについては、評価を下げることを警告している [6]。

1.5 社会で情報を共有するデータビジュアライゼーション

パーソナルコンピュータ、スマートフォン、デジタルサイネージなどデジタルな情報表示メディアの偏在化が進んだ現在では、日常生活の中でデータビジュアライゼーションを利用する機会は増している。特に、データ分析やパーソナルユースを目的としたものではなく、広く社会にメッセージを伝え情報を共有する目的でのデータビジュアライゼーションを使用する例は以下のものが挙げられる。

データ・ジャーナリズム

ジャーナリストが一般に公開されているデータを根拠に取材を進める手法。またそのような根拠となるデータを合わせて開示することで読者の理解を促進する（図 1.3）。

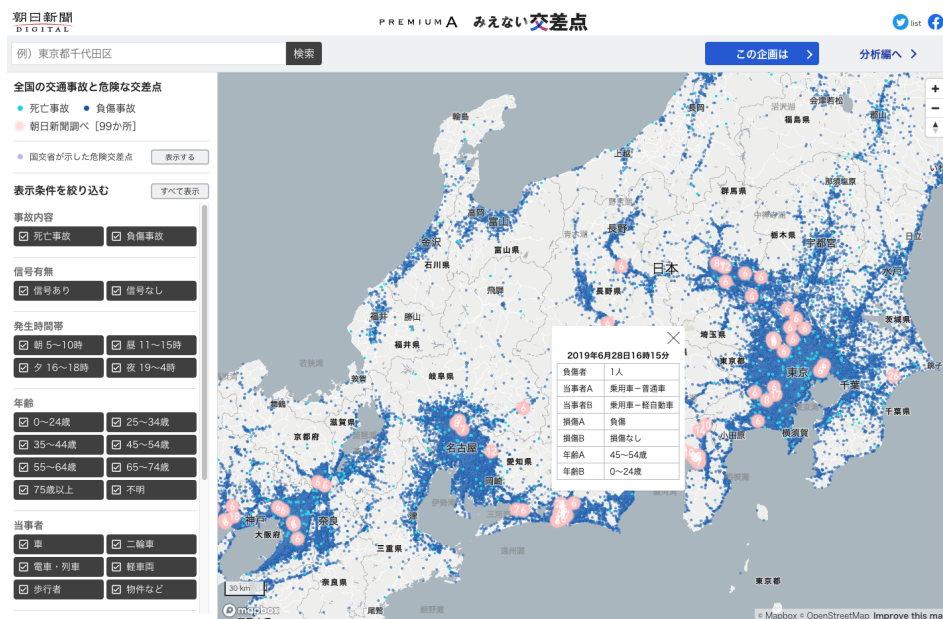


図 1.3 データ・ジャーナリズムの例、朝日新聞社「見えない交差点」 [7]

インフォグラフィクスとストーリーテリング

特定のストーリーについて広く読者に伝える役割を持つ。その際、データをグラフィックとして表現することで理解を促進する。2010 年代以降では、ブラウザのスクロール操

作に合わせてグラフィックやコンテンツが可変する「スクローリーティング」と呼ばれる手法が注目を集めている（図 1.4）。

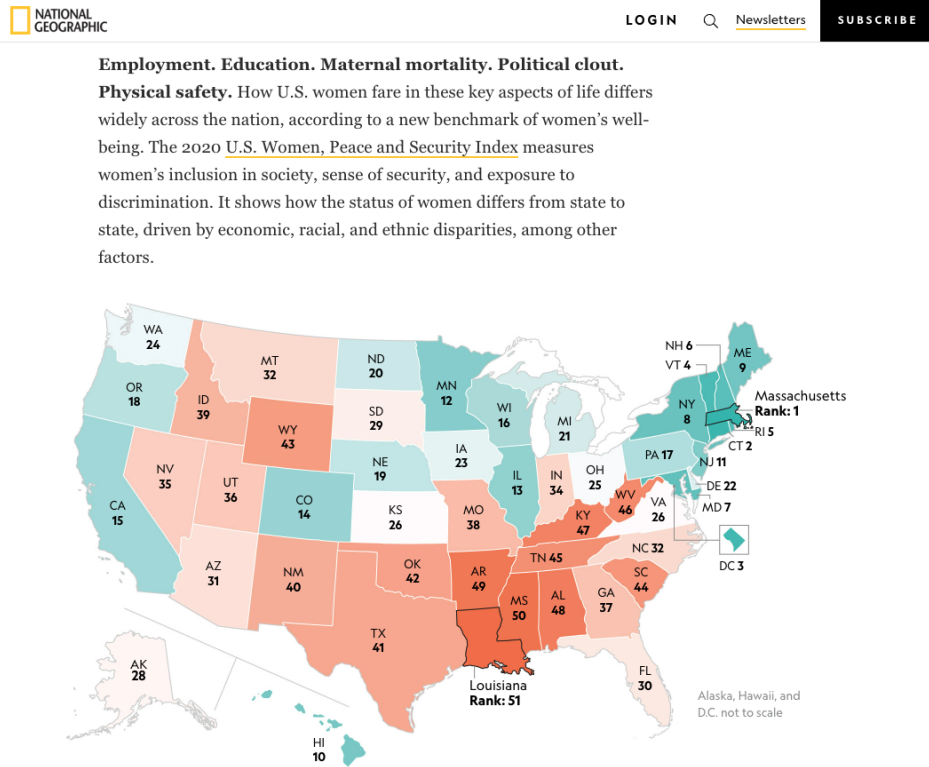


図 1.4 インフォグラフィクスとストーリーテリングの例、National Geographic 「These are the best and worst states for women」 [8]

ダッシュボード

閲覧者の意思決定に必要な情報を一覧性よくコンパクトに提示する手法（図 1.5。多くの情報を総合して素早く決断する必要がある場合や、物事が変化していく様子を即時的に映し出す場合に利用される）。

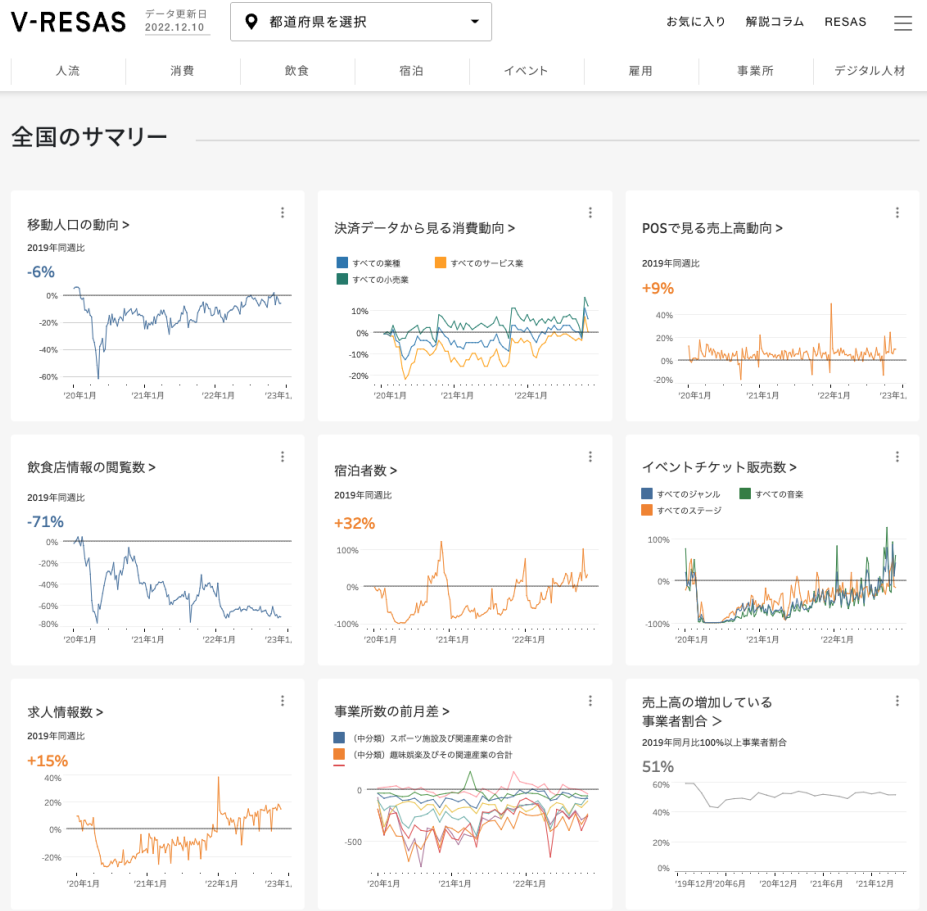


図 1.5 ダッシュボードの例、内閣府「V-RESAS」 [9]

1.6 データビジュアライゼーションとアート

データビジュアライゼーションの情動的な価値の探求は主に物理現象を対象とした可視化研究者や Human-Computer Interaction (HCI) の研究者らによってなされてきた。一方、データビジュアライゼーションの美学的な価値や芸術的価値はメディアアーティストや、インタラクションデザイナーらによって探求されてきた。Gaviria は機能的な情報可視化は、利用者がすぐに理解できるメタファーによって、メッセージを伝えたり、表現されたデータに隠されたパターンを明確にしたりすることを目的としているのに対し、美的な情報可視化は、利用者から直感的または感情的な反応を引き出すことによって、データセットの主観的な印象を提示することに重点を置いていると指摘している [10]。Lau らは、デザイナーやアーティストによる、このようなビジュアライゼーション・アート (ま

たはデータ・アート)の目的は、芸術的な動機付けを持ちながらデータ駆動型の視覚的形態によって、情報社会の主観的な経験を表現することであると指摘する [11]。

また、Lau らは前述の報告の中で、アートやポップカルチャーとしてのビジュアルライゼーションが台頭してきた要因として、ソフトウェアの可用性の向上、データセットの利用可能性の拡大、インターネットの速度と配信の高速化、学際的でクリエイティブなスキルに対する関心、クリエイティブな創作者相互の刺激と研鑽による進化をあげている。特にソフトウェアの可用性向上には、クリエイティブな個人を対象にデザインされたソフトウェア開発環境の普及が貢献したことを指摘している。そのため、デザイナーは複雑な技術的問題を完全に理解することなく高度な情報表現技術を用いながら、直接的かつ反復的な方法でアイデアを実現することができる。

データビジュアルライゼーションの進化において HCI 研究者らの中心的な関心は、複雑な情報構造を表現するための革新的なグラフィックおよびインタラクティブな技術を見つけることである。そのため HCI 分野での情報可視化研究が一巡したと言われる 2000 年代以降、その研究領域は学際性を帯びデザインやアートとも近接するようになった。HCI 研究者の増井は 2000 年代前半時点の情報可視化の研究を総覧した上で、より手軽な開発ツールが普及してきた点に触れつつ「技術者とデザイナー、アーティストの協力により画期的で便利な情報可視化システムが出現してほしいものである」[12]と期待を寄せている。同じく研究者の Judelman は、インスピレーションを得るため情報可視化研究者らと創造的なデザイナーやアーティストとのコラボレーションを提案している [13]。Judelman は提案の中で示唆的な作品群を取り上げ、それら作品の属する分野に蓄積されている情報可視化との接点を次のように紹介している。グラフィックデザインは視覚的なコミュニケーションを最適化するために文字、形、色の使用に関する知識と経験を蓄積している。建築は、空間の設計と概念化を支援するために 3D ビジュアルライゼーションを使用している。インタラクティブデザインや、メディアアートは、デジタル領域でのアニメーション、インタラクティブティ、3D、データビジュアルライゼーションの実験を行なっている。

1.7 情報可視化とグラフィックデザイン

大規模なデータを扱う情報可視化研究が計算機科学の中で活発化したのは、およそ 1980 年以降である。一方、それ以前の時代において情報の図化や表現はグラフィックデザインの領域で探求されてきた。数値を視覚的に表現する技術の歴史は古くその変遷をたどると、特に近代以降、生身の身体では捉えることができない自然や社会の構造や変化を理解しようと努めた人類の足跡を見ることができる。20 世紀以降ではメディアの発達とともに

に、グラフィックデザイナーらは社会に新たな世界観を提示したり強いメッセージを含む表現の一つとして情報可視化を多く開発し研鑽してきた。本研究の対象であるデータビジュアライゼーションもまたこうしたメッセージを持つ視覚表現の一つとして捉えることができる。本節では、近代以降の情報可視化の変遷を概観するとともに、グラフィックデザインと情報可視化の接点について述べる。

1.7.1 科学革命と図の利用

図の起源は相当に古く、例えば、紀元前 600 年頃にバビロニアで作られたと思われる粘土板に描かれた現存する最古の世界図がある（図 1.6）。壁画の類も図に含めればその起源は約 3 万年前まで遡る。



図 1.6 バビロニア地方出土の世界図（大英博物館蔵）

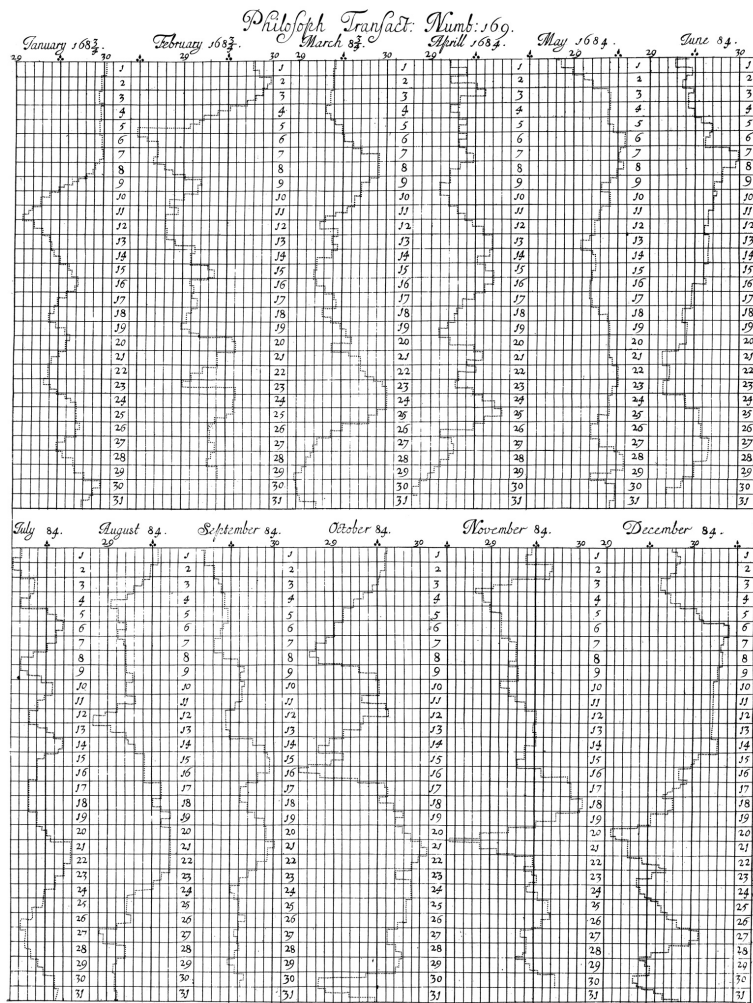
具体物の形態や位置関係をそのまま図化あるいはシンボル化するもの（例えば、洞窟壁画、古代地図、ヒエログリフ等）ではなく、観測された数値と図を結びつけるものを情報可視化とすれば、17 世紀ごろの科学革命を境にその端緒といえる図が多数見られるよう

になる。

この背景にあるのは、おおよそ 16 世紀まで支配的だったアリストテレスの目的論的自然観に基づいた物語的な世界観が科学革命によって大きく転回する過程である。代わりに立ち現れた科学的世界観は、デカルト (René Descartes)、ベーコン (Francis Bacon)、ガリレオ (Galileo Galilei)、ケプラー (Johannes Kepler)、ニュートン (Isaac Newton) らによって牽引され、現代へ続く自然科学の基盤が形づくられた。この転回の中で、図も重要な役割を果たしている。

デカルトは座標系によって数学と図を結びつけ、代数学と幾何学を連携させる解析幾何学を確立した。一方、ベーコンら、観念的な理論よりも実験や観測で得た証拠を元に自然法則を導こうとした経験主義者たちは、得られたデータを記録し図にすることで視覚的に自然の法則に迫った。統計学研究者の Friendly と Wainer は、1684 年にロバート・プロット (Robert Plot) が制作した「History of The Weather」[14] (図 1.7) と、プロットがイギリス王立学会フェロー、マーチン・リスター (Martin Lister) に送付した手紙の内容 (後述) から、グラフによって視覚的洞察が得られることをプロットらが認識していたことを指摘している [15]。プロットは手紙に「このような観測者を離れた国外の多くの場所に置けば同時刻にそれぞれの地点でどのような風が吹いていたか知ることができ、風そのものの境界や幅や流れ方だけでなく、風が呼び起こす天候について根拠を持って調べることができる。ひいては寒暖、飢餓、疫病についても予見可能になるだろう (筆者訳)」と記している。

このように、図と数式を往復して思索を巡らすことや、観測データの可視化から予測につながるパターンを発見することは、この頃から自然科学における定石のひとつとなった。



[935]

A Scheme of the Wind, and Weather at Oxford, April, 1684.

Day.	Wind.	Weather.
1	N & W	fair fun-lhine weather.
2	W & by N	rain i'th morning, fair evening.
3	S & by W	fair morning, wind and rain at night.
4	S & by W	clofe moift weather.
5	S W	rainy weather.
6	due W	fair morning, wet at night.
7	due W	fair weather.
8	S W	clofe weather.
9	N W & N	clofe weather but fair.
10	N E	clofe fair weather.
11	due E	rain and wind moderate
12	N & W	clofe cold weather
13	N & by E	fair fun-lhine weather.
14	E N E	fair morning, clofe evening.
15	NE & by E	clofe cold weather.
16	E & by N	clofe cold weather.
17	E & by N	clofe cold weather.
18	S & W	variable dry weather.
19	due S	moift morning, wet night.
20	SE & SW	clofe warm weather.
21	S E	fair fun-lhine weather.
22	S. N. E	fair morn. inclinable to rain at night.
23	S W	rainy morning, fair evening.
24	S W	fair fun-lhine weather.
25	S W	fair fun-lhine weather.
26	due S	clofe dry weather.
27	S W	fair fun-lhine weather.
28	S W	clofe dry weather.
29	S W	clofe dry weather.
30	due S	rain morning, fair evening

A Seebme

図 1.7 ロバート・プロットによる 1684 年の気圧計のプロット (左) と日毎の風と天候の記録 (右)

1.7.2 統計学の興隆とデータ

科学革命以後、計数と分析が普及する契機としては統計学の発展によるところが大きい。古代ローマやエジプトの時代から、権力者にとって人口や土地を測り国力を把握することは、外部からの侵略をしのぎ繁栄をもたらすためになくはならないものであった。17世紀以降、ヨーロッパでは国家間の競争が激しくなると、競争を有利に進めるために、国家の合理的なマネジメントが必要となってきた。ドイツを中心に国勢学が発展し、人口や土地面積、歳出歳入などを国家として記録するようになる。イギリスでは教会に残る出

生数や死亡数などの大量の記録から人口を予測する試みが行われ、国家の将来予測のために活用が模索される。それらの手法を体系化した『政治算術 (Political Arithmetic)』が1690年にウィリアム・ペティ (William Petty) によって出版された。フランスではデールブルゲームの勝ち負けの議論から確率論が発展し、これらを統合した統計学が確立した。統計 (statistics) の語源が国 (state) であるように、国の状態を把握し運営を支えるためにさまざまな定量化が行われたことが統計の起源である。

また、18世紀から19世紀にかけては印刷技術や航海術の発展も相まって、経済活動が大陸を越えて拡大し情報も多くの人に届くようになった。こうした状況から、集計したデータや統計的に導かれる結果をわかりやすく伝える手段の必要性が高まった。

1.7.3 統計グラフとプレゼンテーション

■The Commercial and Political Atlas 1786年イギリスでウィリアム・プレイフェア (William Playfair) による『The Commercial and Political Atlas』と題された一冊の本が刊行された。政治算術から約100年、外洋航海技術も発達し、ヨーロッパ、アフリカ、アメリカ大陸を巡航するいわゆる三角貿易も活発となり、国家統治だけでなく商業や物流においても国家間競争は激しくなった時代である。膨大なデータを効率的に正確に理解したい、そんな社会のニーズに合致したのがこの本である (図 1.8)。

続いてプレイフェアは『Statistical Breviary』(1801年)を出版している。棒グラフ、折れ線グラフ、円グラフなど、現代でも日常的に目にするダイアグラムの試作を繰り返している様子がうかがえる (図 1.9)。その用例を多数提示し、数表に変わるダイアグラムの有効性を示した。

プレイフェアはデカルト座標系など自然科学の世界で見出された情報可視化の知見を実用の世界に転用し、国家経済という抽象的現象を理解するための視覚的な発明で後世に大きな足跡を残したとされる。

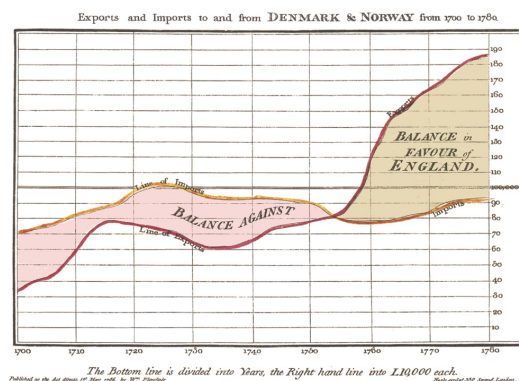
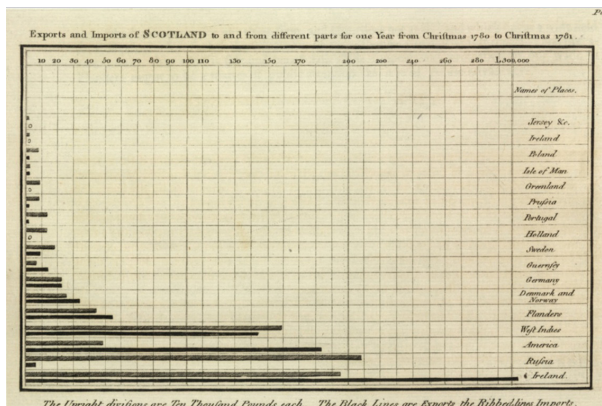


図 1.8 The Commercial and Political Atlas: (右) 1700 年から 1780 年までのイギリスにおける対デンマーク・ノルウェーの貿易収支 (左) 1780 年のスコットランドにおける対外貿易収支。黒の矩形が輸出額を斜線の矩形が輸入額を示す

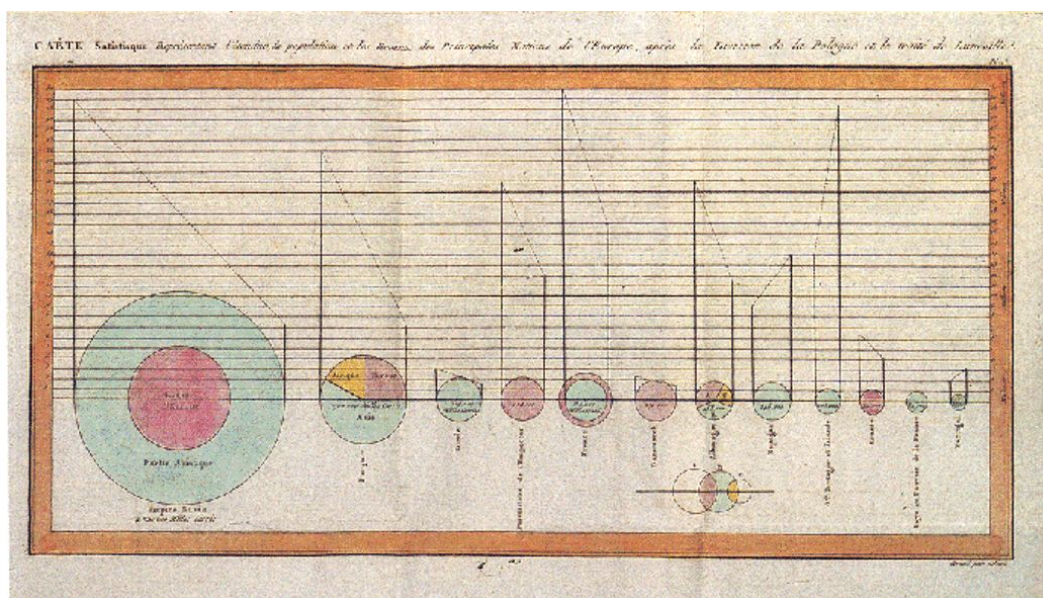


図 1.9 Statistiacal Breviary: ヨーロッパ各国の人口 (左の棒の高さ)、税収額 (右の棒の高さ)、国土の面積 (円の面積) の可視化を試みている。この図では円を用いて現代のパイチャートのような割合表現、直径による総量の表現、ベン図のような表現などのバリエーションが見られる

■ローズダイアグラム 分析以外の情報可視化の主要な役割として、他者への説明すなわちプレゼンテーションがある。プレゼンテーションとして情報可視化を利用する事例もこの頃より確認できる。フローレンス・ナイチンゲール (Florence Nightingale) はクリミア

戦争（1854年-1856年）に従軍して帰国すると、統計学者ウィリアム・ファー（William Farr）、衛生学者のジョン・サザーランド（John Sutherland）とともに、医療状況の分析を行いさまざまな統計資料を製作、提言を行った。図 1.10 はクリミア戦争で死亡した兵士の数と死亡原因を時系列で可視化したいわゆるローズダイアグラムである。これは、戦死者のうち不衛生による病によって死亡した兵士の人数が、負傷によって死亡した人数よりも多く、病院の衛生管理を徹底した 1855 年 4 月を境に激減したことを図化したものである（図中の青い部分は病死、赤が負傷、黒がその他を示す）。

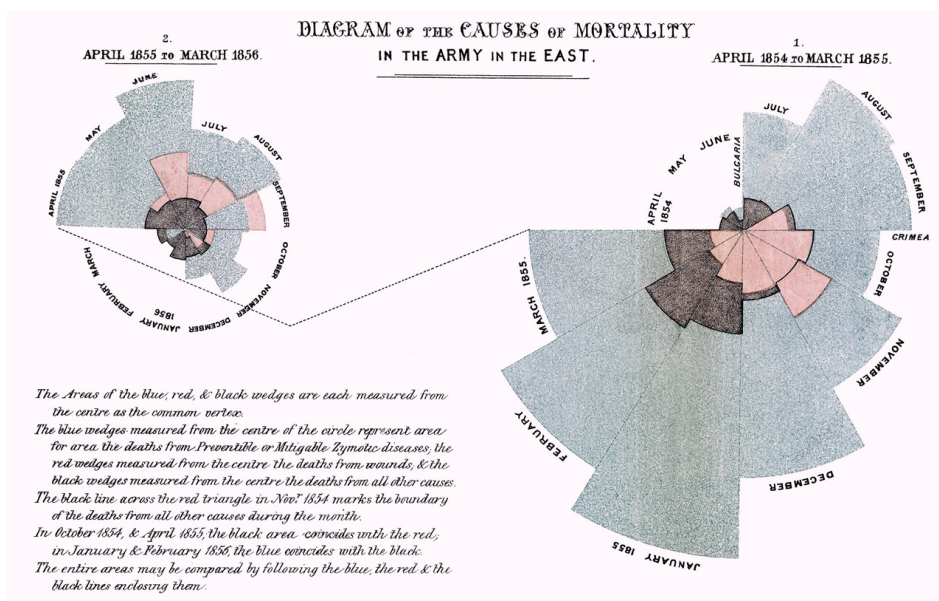


図 1.10 Diagram of the causes of mortality in the army in the East と題されたダイアグラム。報告書『Notes on Matters Affecting the Health, Efficiency, and Hospital Administration of the British Army』の中で使用された。

1.7.4 交通や物流の発展に伴う地図とデータの融合

産業革命を迎えて鉄道が登場すると、物流や交通の発展に伴い生産や消費の形態も一変し、政策上の理由から地図の上にデータを表す必要が増した。19 世紀初頭、イギリスで蒸気機関の鉄道が実用化されると大量の物資を離れた都市まで輸送できるようになり生産と消費の拡大を支え、産業革命の重要な牽引役となった。各国ではそれまでの水路や道路での輸送に鉄道が加わり、大量輸送のための交通網が整備されていく。鉄道や、鉄道専用の橋梁などの大掛かりなインフラストラクチャは一度完成してしまうと安易に変更できないため、敷設の長期的な費用対効果をよく検討しコンセンサスを得ることは避けて通れな

い。そうした中で、国土の地図に各地の生産量、消費量、交通量を重ね合わせた統計地図や流動地図が考案され、広く活用されることになる。こうした需要に応えた代表的な情報可視化はフランスのシャルル・ジョゼフ・ミナール（Charles Joseph Minard）に見ることができる。

■ディジョン-ミュルーズ間旅客量の流動地図 フランスの技術者ミナールは、1781年に生まれ土木技師として多くの重要な交通を担う運河、港湾、橋梁などの設計に携わり、1838年からは土木学校の教員の職に就いている。輸送や物流について見識を備えた、図面設計や地図表現のプロフェッショナルである人物である。

ミナールは鉄道先進国であるイギリスに渡り鉄道に関する知識を吸収すると、母国フランスに戻り統計と地図を融合させたダイアグラムを数多く手掛けた。ディジョン-ミュルーズ間の流動地図（図 1.11）は、鉄道の計画にあたり各都市間の乗合馬車の旅客量を調べ、道幅の太さで表したものである。河川など地理的条件や既存の輸送量を可視化したこの図をたたき台として議論がなされ、線幅が最も太いドーウー溪谷の道沿いに線路を敷設することが決定された [16, p. 110]。

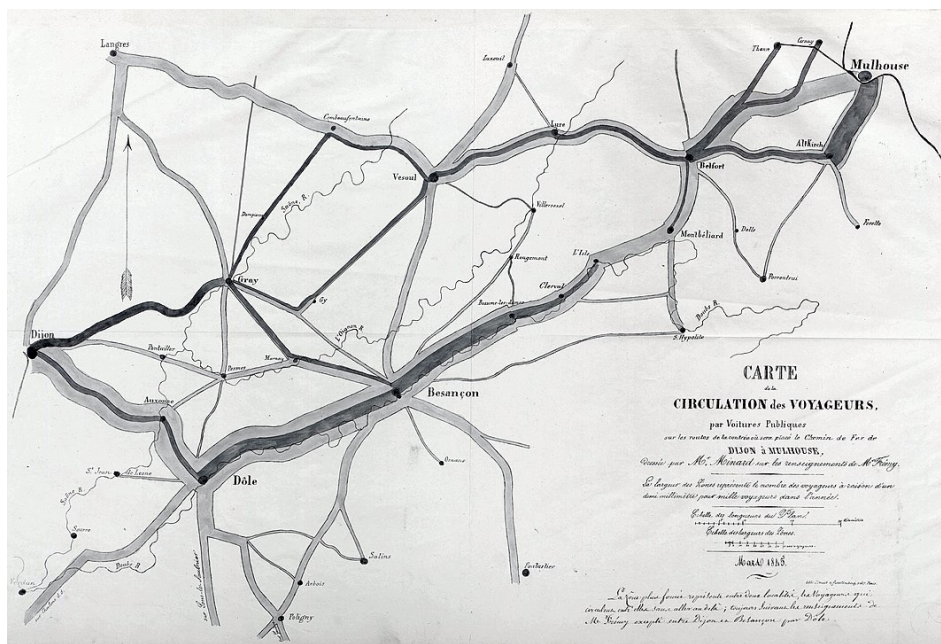


図 1.11 ミナールによるディジョン-ミュルーズ間の流動地図、濃い線幅は通過客の数を表す（1845年）

■**地図と統計** ミナールは生産物統計量と地図の融合においても優れた可視化を実践した。図 1.12 中の円グラフは各県からパリに徒歩で出荷された食肉用家畜の量を円の面積で表している（出荷実績のない県はグレーで塗られている）。家畜の種類による輸送量の内訳を成牛（黒）仔牛（赤）羊（緑）で示している。図から特定の県が多くの成牛を出荷していることや、仔牛の出荷はパリ隣接の県に限られること*3が読み取れる。この図は分円図と地図を融合して統計を表現し、多面的な読み取りを可能とした先駆的な例とされている [16, p. 129]。

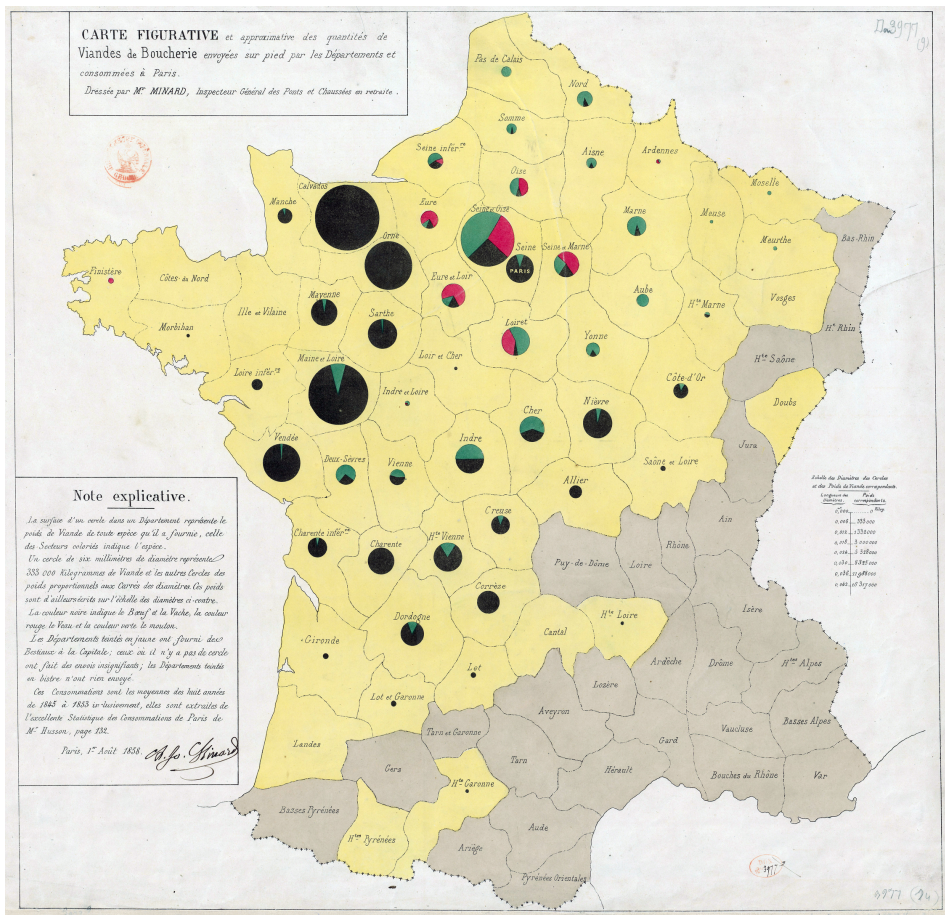


図 1.12 ミナールによる徒歩でパリに出荷された食肉用家畜の統計地図（1858 年）

*3 長距離移動に耐え得る仔牛の体力の限界とされる [16, p. 129]

1.7.5 市場経済の成熟と反動 – 図像統計とアイソタイプの誕生

産業革命を経て自由市場経済が成熟すると、資本家に富が集中する一方で、不衛生や貧困にあえぐ労働者が大量に生まれ、資本家対労働者という階級構造が固定化すると次第に労働者の不満が蓄積していく。1867年以降、マルクスとエンゲルスは『資本論』を出版し資本主義を批判。経済を衰退させずに資源や富の配分をどのように公平に行うべきか、この問いをめぐって経済学上の論争が巻き起こる。

オットー・ノイラート (Otto Neurath) は、1882年ウィーンに生まれ戦時経済、計画経済研究に取り組んだ哲学者、社会経済学者である*4。ノイラートは、版画家のゲルト・アルンツ (Gerd Arntz)、数学と美術に精通したマリー・ライデマイスター (Marie Reidemeister) らと協働し、図像的シンボルを体系的に用いて意味を表現する「アイソタイプ*5」を考案し、その普及に尽力した。その端緒は、目に見えない、富、労働、人口、資源からなる複雑な経済の連関を文字や数字によらず図像のシステムによって示す図像統計 (pictorial statistics) にある。「ウィーン・メソッド」と呼ばれたこの図像体系を駆使して、ノイラートらは視覚教育を通して経済の仕組みを啓蒙する活動を展開した。その舞台となったのはノイラートが館長として関わったウィーン社会経済博物館 (Gesellschafts- und Wirtschaftsmuseum in Wien) での展示である。

博物館時代のノイラートらの重要な事業に、1930年『社会と経済：図的統計の基礎教材』の出版がある (図 1.13)。『社会と経済』は、多色刷りされた100枚のダイアグラムと30枚の解説からなり、綴じられることなく函に収められている。そのため、読み手が自由に図版を組み合わせたたり、解説にしたがって読み解くことで人口、資源、労働、農業、商工業、物流、戦争、民族、宗教、政治体制などさまざまな切り口で、先史時代から近代までの時間的比較や、国家や地域、地政学的分布政に基づく空間的比較を可能にしている。

*4 ノイラートの活動は多面的でその影響は経済学、哲学、グラフィックデザインなどに及び、その活動や思想を巡ってさまざまな視点からの研究がある [17, 18, 19, 16, 20]

*5 「アイソタイプ (ISOTYPE)」という呼称は著書「国際図像言語」(1936年)の中で初めて使用された

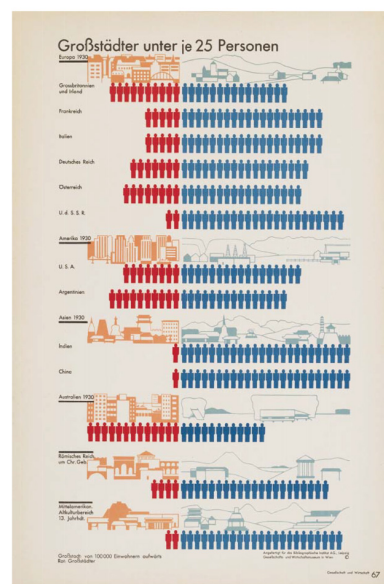
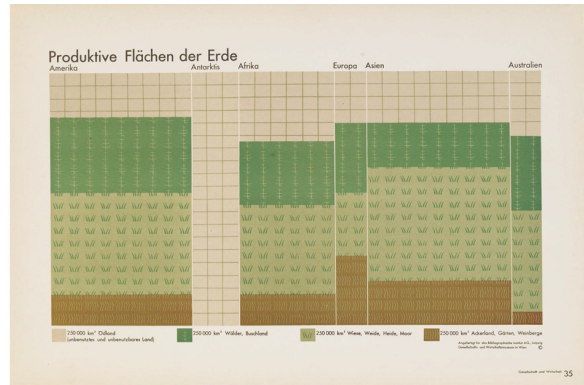
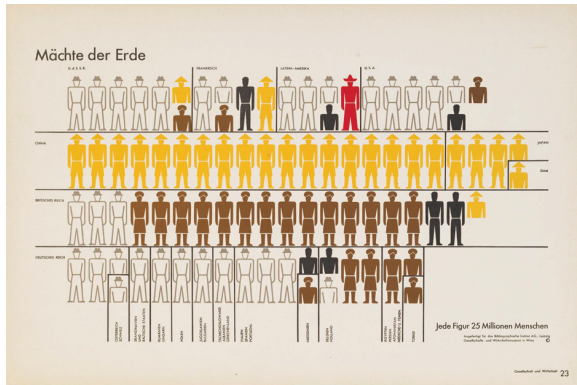


図 1.13 『社会と経済』(左上) 世界の人口及び民族の構成 (左下) ユーラシア大陸の森林面積 (右上) 世界の栽培地面積 (右下) 人口 25 人あたり大都市圏に住む人数の比較

1.7.6 グローバリズム隆興とメディアの発展

大戦の甚大な被害とその後の混乱を経て、世界は再び成長拡大へと向かう。一方、無軌道な拡大が環境を破壊し資源を枯渇させる危機への警鐘から、グローバリズムやエコロジーが芽生える。

哲学、建築、アートなど学際的に活動したアメリカの思想家リチャード・バックミンスター・フラー (R. Buckminster Fuller) は、戦中から戦後にその思想と実践を通じて広く問題を提起した人物の一人である。この活動の中で象徴的な情報可視化が行われている。

■Dymaxion World Map Fuller は世界的な資源分布の統計調査を始めて以降、地球規模の課題を正しく理解するためには地図の刷新が必要であると考えた。航海のための正確な方位表示を目的に作られたメルカトル図法では、高緯度ほど地形が激しく歪むため、極付近の面積や、極を挟んだ2点間の距離を正しく理解することはできない。16世紀以降の航海術の地図で20世紀の世界の関係を可視化しても正しい地理的考察はできないと考えたからである。また、航空機の時代に入り東西イデオロギーの対立の深まり始めていたことから海よりも大陸同士の関係に関心に移る時代でもあった。Fuller が考案した Dymaxion World Map は地球を14の区画に分け、球面を平面に投影する際に生じる歪みが最小限になるように、縮尺、方位、形に気を配りながら、全体に均等に歪みを配分した新しい地図投影法である。地図と情報で描く主題に応じてその区画は並べ替えることができる。

一方、Dymaxion World Map にも課題がある。面積比と形状の正確さを優先しているため、展開すると不連続な領域が生じ矩形平面を充填できない。これは切れ目のないはずの世界と認識のずれを生む。この問題に着目した鳴川は、連続的で並べ替えによる多中心的な表現も可能な投影法による地図「Autha Graph」を考案した [21]。このように世界認識の基準となる地図投影法は、認識の時代変化やずれがリ・デザインの契機となる場合がある。

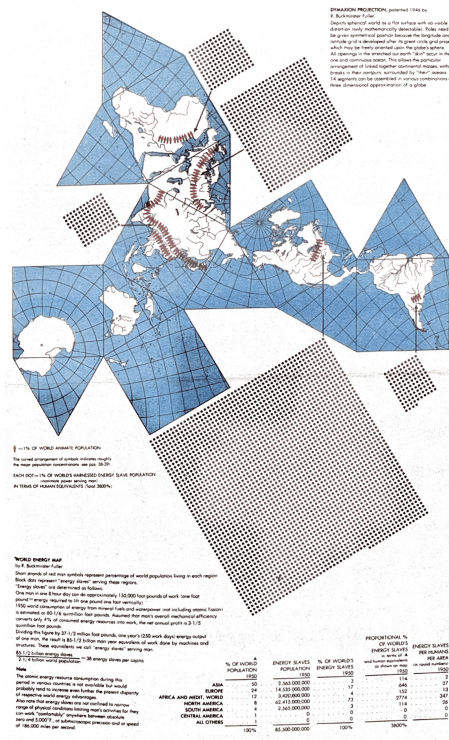
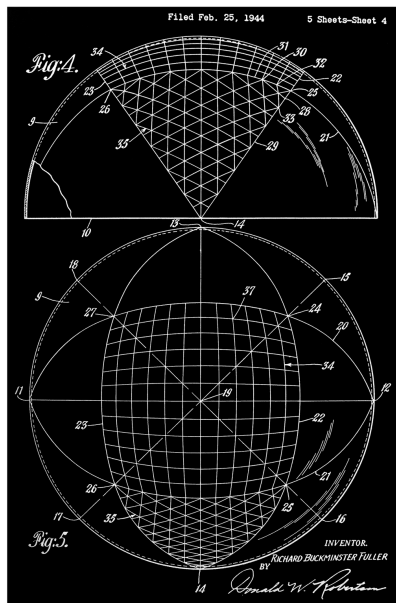
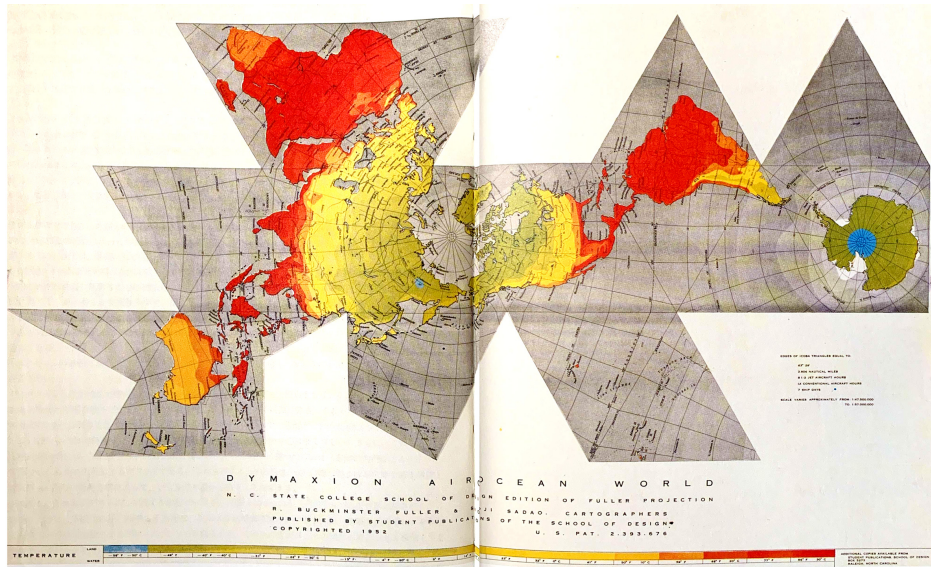


図 1.14 「Dimaxion World Map」：(上) 1952 年 Shoji Sadao と Fuller による正 20 面体版。平均気温が着色されている (左下) 1944 年図法特許申請のための図解 (右下) 1953 年 Herbert Bayer により「World Geo-Graphic Atlas」に使用された図。正方形と正三角形の区画からなるオリジナル版に Fuller が提唱した独自のエネルギー生産指標が可視化されている (図版引用 [22])

1.7.7 情報可視化の公共化・大衆化ービジュアルコミュニケーションとしての情報可視化

20 世紀後半はオフセット印刷や写真製版の技術が高度化し品質の良い印刷物の大量生産・流通が可能になった。加えて、テレビ放送も始まり視覚情報が急速に普及した時期である。商品のスタイルやブランドイメージが消費を刺激することはすでに定着していたが、高品質になったメディアを通じてそれらはさらに拡大し、報道、広告、娯楽など日常のさまざまなシーンに洗練された視覚体験が求められるようになった。

■The Visual Display of Quantitative Information 新聞、雑誌を通じてビジュアルコミュニケーションとしての情報可視化は急速に大衆化し、データを可視化し洞察へと導く技術は統計家だけが必要とするものではなくなった。しかし、ニーズが高まると同時に、誤認を誘発するもの、理解を阻害するもの、詐欺的なものなど粗雑な情報可視化も多く作られた。1983 年、アメリカの政治学者・統計学者のエドワード・タフティ (Edward R. Tufte) が出版した『The Visual Display of Quantitative Information』は、統計と情報可視化の専門家がもつ熟達したデザインの技を体系化し、非専門家に向けて門戸を開いた初めての本格的な書籍である。タフティはこれ以降一貫したテーマで、1990 年に『Envisioning Information』、1997 年に『Visual Explanations』、2006 年に『Beautiful Evidence』を出版し、いずれも今日まで、情報可視化を洗練されたビジュアルコミュニケーションツールとして定着させる知識の普及に大きく貢献した。

■Understanding USA リチャード・ソール・ワーマン (Richard Saul Wurman) は、増大していく一方の情報と人間の理解を自然に結びつける方法について「インフォメーション・アーキテクチャ」を提唱した。特に、20 世紀後半から研究と実践に意欲を注ぎ情報デザインを牽引した。2000 年、ワーマンのディレクションにより制作された『Understanding USA』[23] (図 1.15) は、「理解」をその名に冠し、膨大な統計をもとに、情報デザインとインフォグラフィクスによってさまざまな視点からアメリカを読み解こうとするプロジェクトである。

ワーマンが本書の序文に掲げたメッセージは「Public information should be made public.」である。永原はこの「Public」に「公共」と「大衆」の二つの意味の「公（おおやけ）」を読み解いている [20, p.112]。これは、大量の公開データだけでは国民は実態を理解できないことに対する政府への批判であると言えよう。本書の出版当時、すでにインターネットの利用者数も拡大をみせ情報爆発の足音が聞こえ始めていた。情報があること

が当たり前の時代に入らる中で、大量のデータを大衆の理解に結びつけることは容易でないことを示している。この問題についてワーマンが重視したものは、問うことである。優れた問いがなければ膨大な情報は理解へとつながらないとする主張である。それは次のように序文に表れる。

アメリカ人が心に抱いている基本的で単純な疑問は、簡単に、即座に、明快に答えが得られる。質問をして答えを得ることは、私たちに与えられた権利だ。本書は、理解へ通じる質問と答えの祭典であり、質問と答えが理解に通じることを視覚的に実証するものだ。本書では理解することに内在するパワーを、そして、理解はパワーだという考えを実証している。(筆者訳) [23]

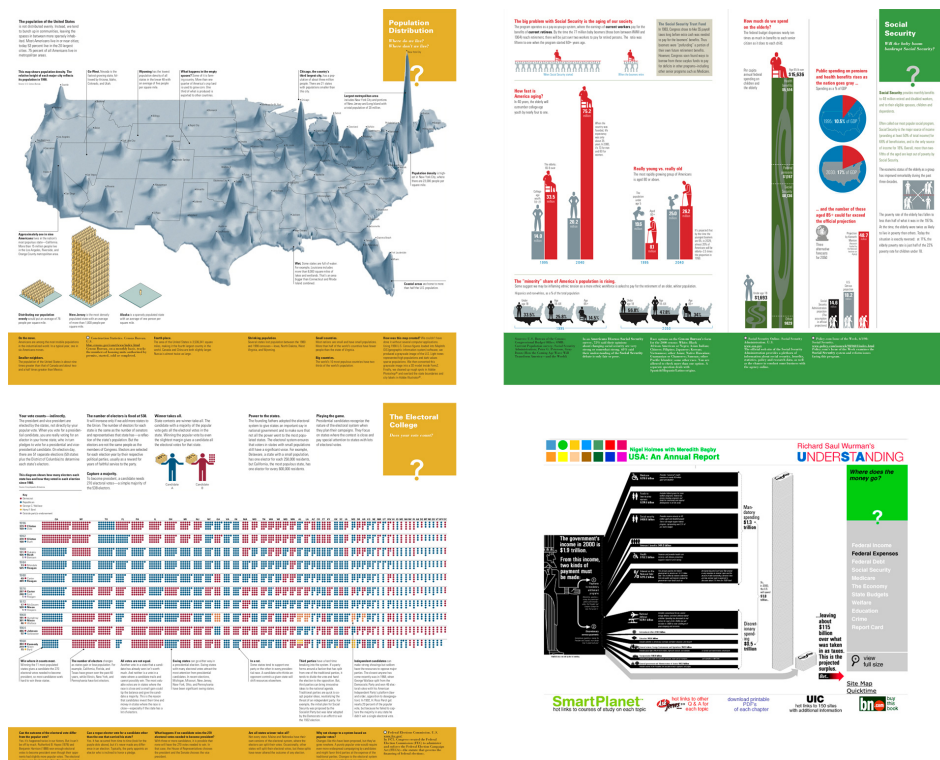


図 1.15 「Understanding USA」、それぞれのページのテーマが問いの形式で右上に記載されている：(左上) 人口 (右上) 社会保障 (左下) 選挙制度 (右下) 出版と同時に公開されたインタラクティブなウェブ版の「歳出と歳入」

1.7.8 動的な情報可視化の登場—対話的インターフェイスの研究

コンピュータが高額で大型な装置から、安価で個人でも所有できる装置に変わり始めると、大量の情報の中から目的の情報を容易に見つけるための検索システムやインターフェイスが研究のモチベーションとなった。情報可視化は、その中であって、人の要求（ときに曖昧な要求）に対しても満足に行く結果を表現する役割を期待される。情報可視化に必要なグラフィカルなシステムを効率的に構築できる環境が充実したことも研究開発の多様化につながった。このような背景から、独創的な情報可視化システムが数多く提案された（図 1.16）。また、この時期にグラフィックデザインと動的な情報表現の研究も行われている（2.5 節「Information Landscape」を参照）。現在普及している情報可視化の表現手法には、この時期に提案され定着した手法もある（図 1.17）。

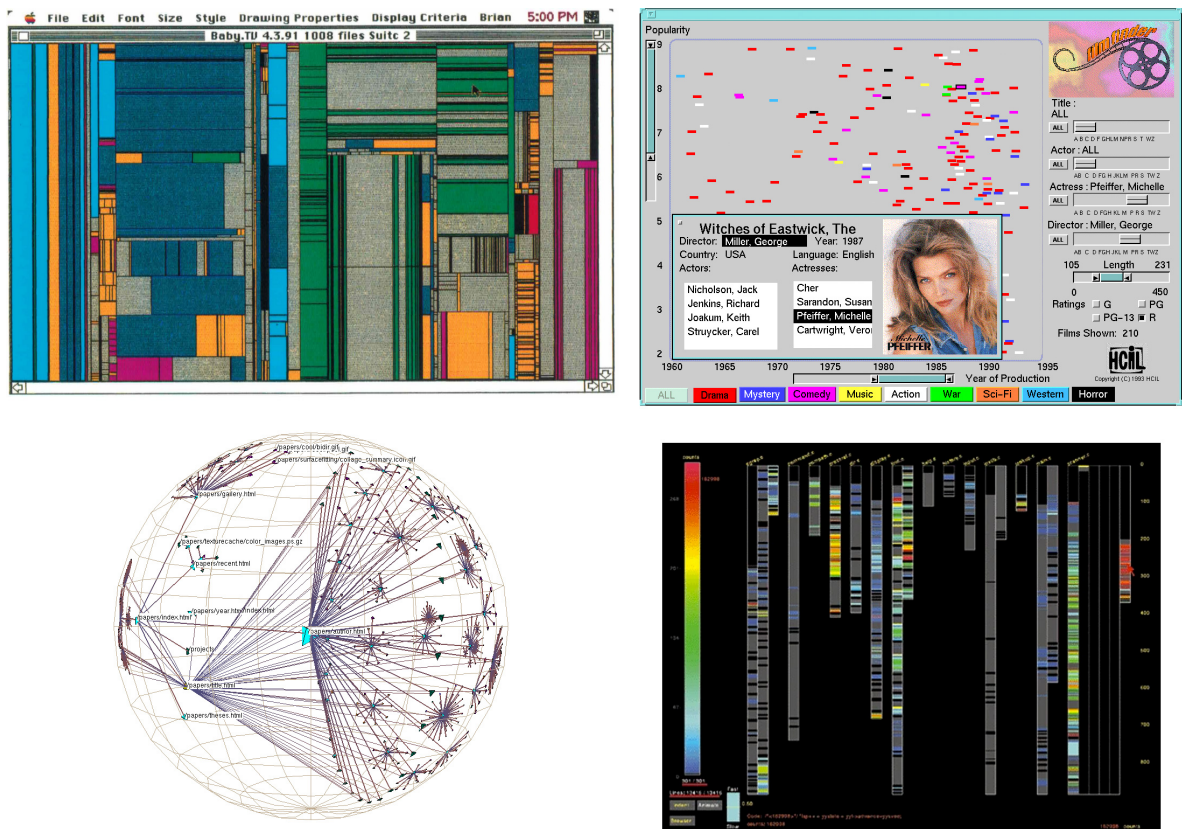


図 1.16 (左上) 木構造 (ファイルシステムなど) を 2 次元空間の入れ子の矩形で表す「TreeMap」[24] (右上) 動的な映画検索システム「Film Finder」[25] (左下) 複雑なネットワーク構造を 3 次元平面にマッピングした「H3」[26] (右下) プログラムのソースコードなど長いテキストファイルを抽象化して可視化する「SeeSoft」[27]

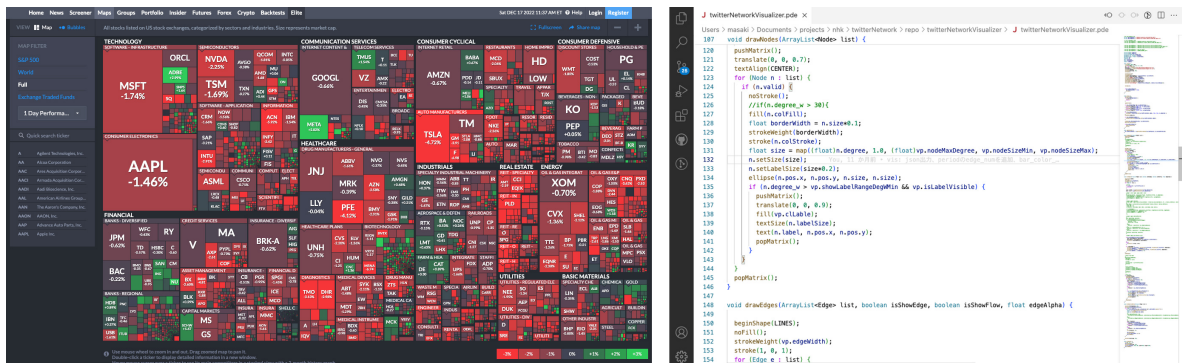


図 1.17 (左) 株式市況を伝えるウェブサイトで (finviz.com) 「Treemap」を使用している例 (右) ソースコードエディタ (VSCode) と併置してソースコード全体を表示させる「SeeSoft」のようなマップ

1.7.9 データから人や社会の動きを映し出すアート、デザインの登場

インターネット商用化が始まった 90 年代後半以降、人々の生活はデータによって支えられ、またデータによって人や社会の活動自体を可視化することが可能になってきた。この頃から研究者だけでなくアーティストやデザイナー自身がプログラミングによってデータの表現を探求し、巨大なデータやリアルタイムデータを可視化した作品が多く生み出された (図 1.18)。そうした作品はデータビジュアライゼーションによる効率的な理解促進や情報伝達ではなく、写真、映像のように感性を刺激しデータを通じて社会や人間活動を描き出すことで注目されるようになった。

■Sensorium(Web Hopper) Sensorium は竹村真一を中心に渡辺保史、西村佳哲、上田壮一、東泉一郎、小崎哲哉、他多数のメンバーからなるプロジェクトである。インターネット黎明期から情報空間での新たな知覚 (Sense) をテーマに作品を制作した。「Web Hopper」は江渡浩一郎が開発を担当しインターネット上を流れる自分や他人の HTTP パケットの位置情報をリアルタイムに可視化する作品である (図 1.18 左上)。97 年 Sensorium は Ars Electronica Prix ネット部門で金賞 (Golden Nica) を受賞した [28]。

■SHISEIDO 30-YEAR ANNIVERSARY POSTER 1995 年に企業のテレビ CM 放送 30 周年を記念してジョン・マエダ (John Maeda) によりデザインされたポスター [29] (図 1.18 右上)。30 年間に制作された商業フィルムのフレームを極小サイズで配置したイメージと、タイトルのアスキーコードやサウンドデータをもとにアルゴリズムにより生成された有機的なパターンで構成されている。(ニューヨーク近代美術館蔵)

■History Flow 2003年、フェルナンダ・ビエガス (Fernanda B. Viégas) とマーティン・ワッテンバーグ (Martin Wattenberg) による作品 [30] (図 1.18 左下)。ネット上で人々が自由に編集できる Wikipedia の記事がどのように編纂され、攻撃的な編集に対してどのように復元していくのかという問いから生じた作品。同図は、abortion (中絶) を説明するページの編集履歴を可視化したもの。黒い溝は攻撃により削除された期間を表す。(ニューヨーク近代美術館蔵)

■Flight Patterns 2005年、アロン・コブリン (Aaron Koblin) とバルディン・クランプ (Valdean Klump) による、24 時間の間に北米上空を飛行した 14 万機以上の航空機のログデータを可視化したインタラクティブアニメーション作品 (図 1.18 右下)。飛行時刻、位置、高度、機種などのデータから人口の分布や時間帯ごとの航空機の分布など巨大な輸送システムの上で動く人の流れを可視化している [31]。

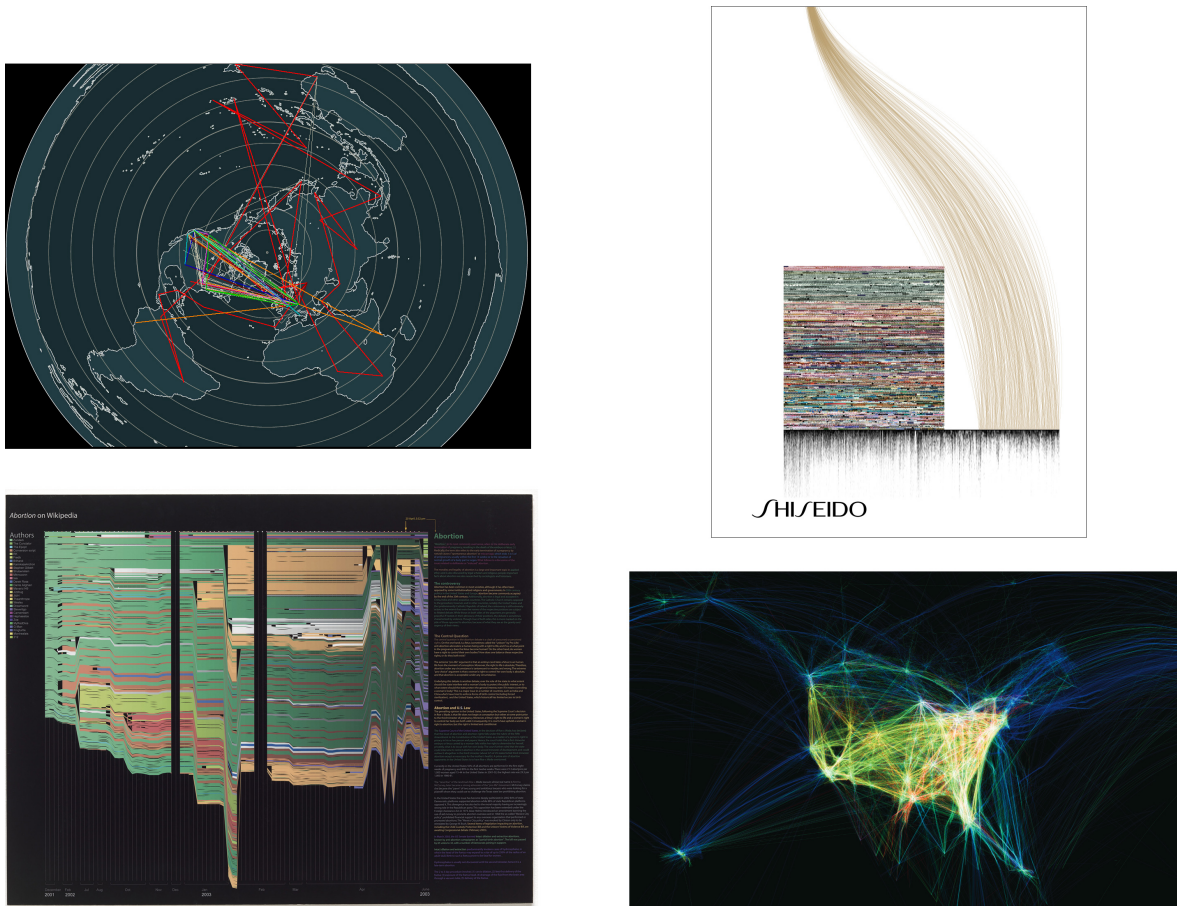


図 1.18 (左上)「Web Hopper」[32] (右上)「SHISEIDO 30-YEAR ANNIVERSARY POSTER」 (左下)「History Flow」 (右下)「Flight Patterns」

1.8 本研究の目的

序論ではまず、かつてない膨大なデータの出現を背景として、機械学習技術（いわゆるAI）や高度なデータ解析にもとづいた新たな価値の創造や社会の変革を推し進める「データ駆動型社会」のビジョンに触れた。また、これら技術主導で進む先に残される一般市民の受容についての問題提起にも触れた。

続いて、情報可視化はデータ分析を支援し、結果をわかりやすく伝達する側面に注目が集まっているが、それだけでなく感情的な価値や複雑で多様な解釈を共有する価値を持つことについても論じた。現在我々が目にするチャートやグラフは見慣れたものかもしれないが、歴史を振り返れば、情報可視化は科学革命と共に端緒を開き、科学技術の進展や社会の構造変革の中で自然や社会に対する新たな見方を創造し獲得してきたことに触れた。このことから情報可視化は社会の有り様の変化に応じて探求され開発されることが望まれるはずである。

現在、大規模なデータの解析的な見方だけでなく、複雑さや多様さを共有する見方の価値は高まりつつあり、そのような探求は情報可視化研究者だけでなく、ソフトウェアエンジニアリングやプログラミングの知識を持つ情報デザイナーやクリエイターによって多く提示されてきている。しかし、そのような取り組みはどのようにして行われるのか、あるいは、卓越した実践者の経験則や実践知はどのように記述されるべきなのかという問いには未だ答えはない。

このことから、本研究ではデザイナー、アーティストなど視覚表現を行うクリエイターによる大規模なデータを扱う創造的な情報可視化のデザインを支援することを前提として、データビジュアライゼーションのデザインプロセスの経験則を形式知化する手法の提案を目的とする。

第2章

関連研究

2.1 はじめに

本研究は大規模で複雑なデータや、専門的なデータを可視化し、一般の鑑賞者にとって現象を多様な視点から眺め、洞察へと導くことができるデータビジュアライゼーションのデザインプロセスを観察し、創造的表現を探求するデザイナーのために、プロジェクト推進の枠組みを提供するものである。すなわち、データの理解に始まり具体的なメディアに可視化を定着するまでを包括するプロジェクト全体の概観、プロジェクトを構成する各ステップの意味やステップ間の相互関係、プロジェクトにおける専門家と非専門家のコミュニケーション、プロジェクトが抱える本質的な困難さやそれを解消するための道筋を、実際にプロジェクトで得られた経験や観察を通して記述することである。

情報可視化のデザインは幅広く探求されている。コンピュータで大規模なデータを処理する情報科学、統計学、気象学、流体力学、社会科学、医学などにとどまらず、近年では人文学や歴史学などでもその応用的手法が見られる。

本研究では、大規模なデータを処理し動的な表示や操作を含むシステムとしてのデータビジュアライゼーションをデザインすることから、Human-Computer Interaction (HCI) の領域で議論されてきた問題解決型の情報可視化システム開発手法や、領域専門家と可視化専門家の協働による可視化開発のためのフレームワークと共通点は多い。

また、本研究では、大規模なデータの情報可視化を通じて広く一般に向けた作品を作ることに焦点を当てている。このようなメディアとしての性質を持つ情報可視化は、データ・ジャーナリズムやインフォグラフィクスなど、一般に向けたコミュニケーションを目的とした可視化開発フレームワークとも共通点を持つ。

以下では、こうした分野での既往研究を述べつつ本研究との差異を論じ、本研究の位置

付けを明確化する。

2.2 情報可視化の基礎的研究

近代以降、専門分野から一般領域に至るまで広く利用されるグラフやチャートの設計において、基盤となる原則を示したのは Edward R. Tufte である。Tufte は統計学の専門家でありながら優れた情報可視化設計のパイオニアである。Tufte の研究をまとめた書籍は扱うデータの領域を限定せず、また、読者層も広く想定して書かれ広く普及した [33, 34, 35]。Tufte の著書の最も強調されている原則は「情報可視化の価値は複雑な対象を、明瞭、正確、効率的に伝えること」である [33, p. 13]。中でも「Data-Ink Ratio」と名付けられた情報可視化の評価尺度は、あるグラフを描くのに必要な総インク量に対する、データそのものを表現するインク量の比である [33, p. 93]。この比率が小さければ、本来不要な視覚要素や装飾を排除が少なく、効率的な情報可視化につながると主張した。

不必要な装飾的要素を排除することは、理解を妨げる要因を減らすことにつながる。Tufte の主張は現在でも有効であるが、一方で、データを表す以外の要素が少ないグラフは、その背後にある意味や物語を想起させる視覚要素を付加したグラフに比べて、記憶への定着が弱いことが指摘されている [36]。また、Tufte の論考は静的な印刷物における情報可視化に焦点を当てているため、コンピュータとディスプレイによる動的な情報可視化やインタラクティビティについてはほとんど記述がないため、本研究の中心的な議論からは外れる。

認知心理学では人間が感覚器を通じて外部刺激を知覚し記憶として定着するプロセスを明らかにしている。感覚器官から得られる 500m 秒程度の瞬間的な記憶である感覚記憶の一部が短期記憶（数分程度の保持）に転送され、さらにそれらを定期的に思い起こし復習することで神経細胞のシナプス構造が強化され、数十年という長期記憶に定着する。視覚からの感覚記憶には「前注意処理」と呼ばれる意識の集中を要せず瞬間的に知覚するメカニズムがあり、これを利用することで特定の視覚刺激を与える属性を持つ要素だけが周囲のものから浮かび上がり、差異として瞬時に知覚される（例：Pop-Out 効果）。

Ware は、認知心理学の研究から前注意処理を喚起し情報可視化にするための視覚属性（色、形、位置、動き）をカテゴライズした [37]。データセットが持つ複数の変数を視覚属性に割り当てる（エンコードする）場合、人間が一度に知覚できる視覚属性の数には限界がある。また、視覚属性を定量的な表現に割り当てた場合の鋭敏に知覚する（デコードする）度合いについても視覚属性間に優劣がある。Cleveland と McGill はこれらを実験的に検証している [38]。また、Few は実験から数量データ、序列データ、分類データと視覚

属性の適合性を評価している [39]。

情報可視化がコンピュータディスプレイの上で実現できるようになると、印刷媒体や手作業による作図では実現できなかった大量の情報を少ないスペースで動的に可視化するズームングやスクロール、関心に沿ってデータを選択するフィルタリングなどの対話的操作（インタラクション）が利用できるようになった [25]。また、CG、物理演算アルゴリズムなどを応用した視覚的な計算処理を施した可視化が可能となった [40]。

情報可視化と Graphical User Interface (GUI) を組み合わせて、利用者の持つニーズ（時に、明確ではないニーズ）を満たすには、情報可視化システムが利用者の要求に従ってディスプレイされる可視化を動的に変更し、その結果から、さらに利用者が目的とする情報へとナビゲートされる（または、ひとつ前に立ち戻って要求をやり直す）ことによって利用者のタスク実現を支援する設計が必要となる。Shneiderman はこのような情報空間の全体像の提示と利用者要求に従って情報を絞り込む対話的で反復的なインターフェイスの設計思想を「Visual Information seeking Mantra」 [41] と称して提唱している。

Card らの『Readings in Information Visualization』 [42] は 90 年代までに研究された HCI 分野における情報可視化や検索システムに関する主要な研究を総覧する。特に第 1 章では、各研究成果を踏まえつつ、複雑なデータを処理してグラフィカルに情報を表示し対話的操作を通じて効率的にタスクを処理するシステムを実現する上での総合的な知見を提供している。

このように、情報可視化の基礎となる分野は広範囲に及んでいる。Mazza の『Introduction to Information Visualization』 [43] は、情報可視化分野の成り立ち踏まえつつ、それらを理解する上での参照先を網羅的に提供している。

2.3 参照モデル

Card らはインタラクティブな情報可視化システムの典型的な処理手順を示す参照モデル (Reference Model) を提案した [42, p. 17]。

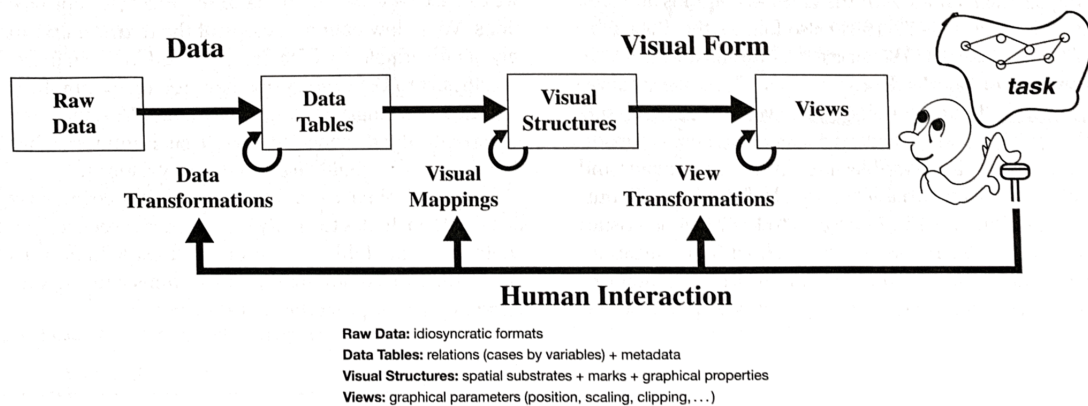


図 2.1 Reference Model のダイアグラム ([42, p. 17] より引用)

図 2.1 では、左側にある Raw Data からあるタスクを持った利用者と接する View に向けてステップに別れた矢印が描かれる。これらは各ステップ間の変換の連鎖を表す。利用者から各変換ステップに向けた矢印は、それらの変換を調整するための利用者自身の操作を示す。Raw Data はスプレッドシート、電気信号、小説の文字列などあらゆる形式として存在しており、目的とする表現に直ちに交換できるとは限らない。Data Transformations はこれを論理的に構造化し可視化システムに適した形式に変換する。このように変換されたデータ構造は視覚構造へと変換される。Visual Mappings は視覚構造とデータ構造を対応づけることである。視覚構造とは、前節で示した認知心理学的研究から得られた知見を基盤として定義されたものであり、次の 3 つの要素を指す。

1. 空間定義
2. 視覚要素
3. 視覚属性

空間定義とは、2次元のグラフでは縦軸と横軸、3次元ではさらに奥行きをどのようなデータでマッピングするかであり、視覚要素は表現されるオブジェクトの本質的な形状のことである。視覚属性は形を持ったオブジェクトに付与される色、テクスチャ、大きさ、向きなどである。データ構造と視覚構造の対応関係が適切でなければ、利用者の求めるタスクを支援できないだけでなく、正確に情報を読み取ることすら困難になる。決定された視覚構造は、View Transformation を通じて可視化アルゴリズムによって具体的にディスプレイに表現される。表現された可視化は、ズーム、スクロール、パン、フィルタなどの GUI を通じた操作によって動的にレイアウトの変更に対応する。このモデルが示すフローによって、前節で示した Shneiderman の Mantra に象徴されるように、利用者の持

つコンテキストに応じた適切な表示に変換され、情報を取り出すことができる。

2.4 システムとしての情報可視化の設計

ここからは、情報可視化の設計プロセスについての関連研究を述べる。

情報可視化の設計・実装・評価の手法は、情報可視化の用途によって異なる。伊藤は情報可視化の用途を以下の4つに分類する [44]。

1. 概観 (Overview)
2. 解明 (Clarification)
3. 操作 (Handling)
4. 報告 (Announcement)

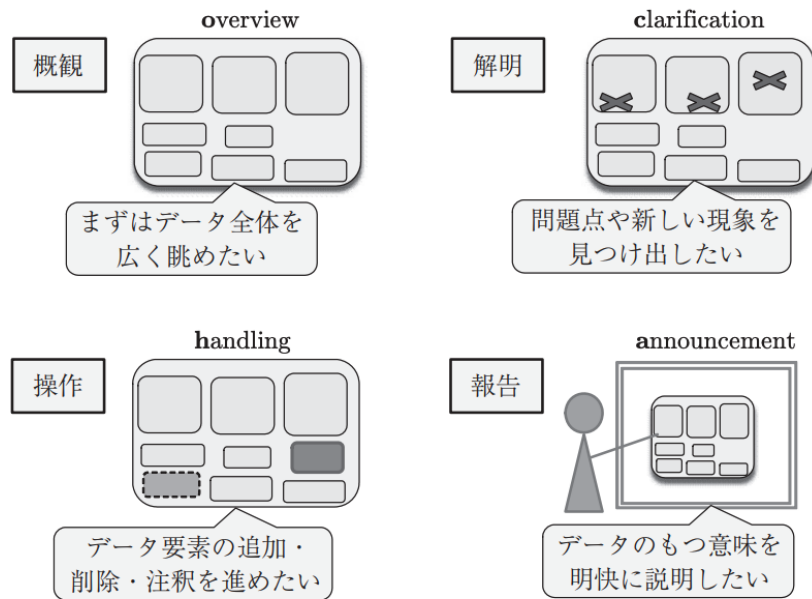


図 2.2 情報可視化の用途。伊藤貴之. 意思決定を助ける 情報可視化技術, p. 6. コロナ社, 2018 より引用 [44]

4つの用途のうち、「概観」「解明」「操作」は利用者自身が何らかの課題や問題の解決に向けて積極的に関わるものである。一方、「報告」は他者に向けたコミュニケーションに主眼を置くものであり、他の3つの用途とは異なっている。

ここでは、「概観」「解明」「操作」の3つをまとめて、利用者が抱える課題を支援するこ

とを目的とした「システムとしての情報可視化」に、「報告」を広く一般に向けたコミュニケーションを図ることを目的とした「コミュニケーションメディアとしての情報可視化」に大別する。

まずは、前者についての研究動向を述べる。システムとしての情報可視化は、利用者自身が目的を持って一次情報 (Row Data) を探索し、そこで成し遂げたい課題解決や理解を支援することを目的として開発される。つまり、利用者自身が一次情報の領域について知識を持ち、探索可能性空間を自身の関心において独立して探索できることが前提となる。

近年、情報可視化を伴う分析 (Visual Analytics) を必要とする研究領域は非常に広範囲におよび、個々の専門性も高いために、情報可視化設計者が各専門分野 (ドメイン) のニーズを理解し抽象化することは (重要でありながらも) 困難であると認識されている。

加えて、あるドメイン領域における利用者が抱える課題解決を支援する情報可視化システムを構想する際、そのタスクを抽象化したり、システムが提供する機能の定義、機能性を評価する指針も必要となる。こうした問題解決型の情報可視化システムの研究では、ドメインニーズの的確な把握と正しい理解のために、エスノグラフィなど Human-centered design (HCD) アプローチが取り入れられている [45, 46]。このアプローチを「Design Study」という場合もある。

また、視覚表現手法やインタラクションの選択、実装のフェーズにおいてもさまざまなトレードオフに満ちており、最適な解を選択するためには膨大な選択肢の中から最適解を検討する必要がある。そのため、多くの検証と試作を繰り返しながら具体的なソフトウェアとして完成していく。

Munzner はこのような情報可視化設計プロジェクトの観察的研究から、プロセスの方法論として「Nested Model」を提案している [47, 48]。Nested Model の 4 階層の入れ子構造を図 2.3 に示す。

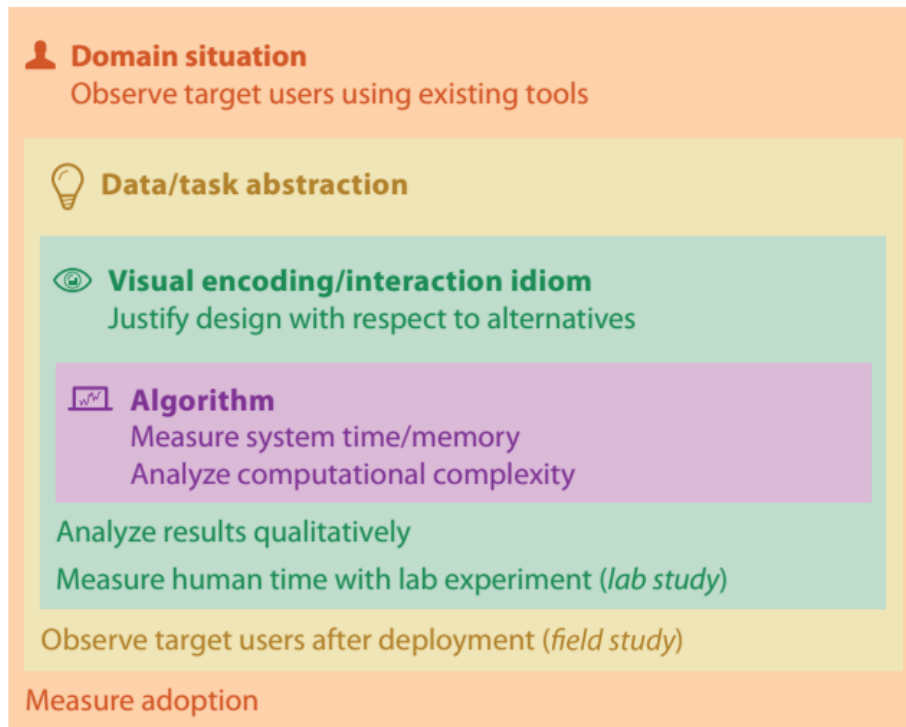


図 2.3 Nested Model の入れ子構造。Tamara Munzner. Visualization analysis and design. p66. CRC press, 2014.[48] より引用

上層から「Domain Situation」、「Data/Task Abstraction」、「Visual Encoding/Interaction Idiom」、「Algorithm」の4階層からなる。

Domain Situation では、ターゲットとする利用者のグループの関心やそのドメインの状況を利用者と密接に関わり合いながらデザインスタディを通して調査し記述する。

Data/Task Abstraction では、上層で得られたそのドメイン特有の状況をドメインに依存しないレベルに抽象化する。対象となるデータやタスクがそのドメイン特有の語彙から抽象化され、情報可視化システムでの語彙のレベルに翻訳されることで、対象となるドメインとは異なる状況や分野で得られたタスクやデータ構造との比較やアナロジーを利用することができる（抽象化されたタスクの例は、要約する、比較する、ブラウジングするなどである）。

Visual Encoding / Interaction Idiom は、抽象化されたタスクやデータ構造をもとに、視覚的なエンコードやインタラクションを決定する。イディオムとは視覚的エンコードとインタラクションを組み合わせた戦略のことであり、複雑なタスクになればいくつものイディオムが選択の候補となり得る。

決定されたイディオムは最後に最下層の Algorithm で実装が検討される。この階層では計算リソースと現実的な応答速度という制約の中でイディオムを実現する最適な方法が決定される。

これらの4つの入れ子のブロックは各ブロックの最後にその妥当性を検証することで、最終的に具体化される情報可視化システムが、本来の潜在的な課題を解決するにあたり機能的に作用することを目指して検証性を確保する狙いがある。

2.5 コミュニケーションメディアとしての情報可視化の設計

次に、コミュニケーションメディアとしての情報可視化について述べる。この分野での可視化の利用例として、インフォグラフィクスを使用したジャーナリズムにおけるストーリーテリング、インフォメーション・ダッシュボードなどが挙げられる。これらの共通点は、可視化制作者の関心に沿って一次データが整理され、閲覧者へ提示される点である。

提示によって期待する効果は、閲覧者の理解の促進、意思決定の支援など、閲覧者に対して（主には大衆に向けて）何らかの働きかけを伴う情報の発信である。そのため、この場合の設計モデルは、必然的に編集的視点やデザインの思考を含むものとなる。

コミュニケーションメディアとしての情報可視化の実務家であり研究者でもある Andy Kirk はこのプロセスを次のモデルで説明している（図 2.4）[49]。



図 2.4 Kirk の「Data Visualization Design Process」[49, p. 31] より引用

プロセスは次の4つのステージからなる（和訳は筆者による）。

1. Formulating your brief（ブリーフの作成）
2. Working with data（データを整える）
3. Establishing your editorial thinking（編集的思考で吟味する）
4. Developing the design solution（デザイン手法の開発）

ブリーフの作成の目的は、プロジェクトの文脈を定義し、ビジョンを明確にすることで

ある。文脈とは、聴衆は誰なのか、制約条件（期間・コスト・技術的課題）は何か、成果物の特性（媒体や更新性）は何かを特定することである。ビジョンとは、その情報可視化によって閲覧者に何がもたらされるか（説明、理解、探索的な経験等）を特定することである。

データを整えることの目的は、必要なデータ（一次データ）を取得し求める情報可視化に利用できる形式へと変換することである。

編集的思考で吟味することについて Kirk は次のように述べる。

Editorial thinking is concerned with making informed judgements about the content you intend to include in your visualisation. In my view this is one of the most defining activities that separates the best visualisers from the rest, possibly even more so than any technical talent or design flair. Before we move on to making design choices, you need to consider: given all the things you could show, what will you show?[49, p. 119]

編集的思考とは、情報可視化に含める内容について入念な調査と検討に基づいて判断を下すことである。私の考えでは、これは良いビジュアライザーであるか否かを左右する決定的な活動の1つであり、おそらく技術的な才能やデザインのセンスよりも重要なものである。デザインの選択に移る前に、見せられるものすべての中で何を見せるかを考える必要がある。（筆者訳）

また、編集の必要性について次のように述べている。

Editing is about making selections: choosing what clips you leave in a movie, what contents you leave in a book, how you arrange music into coherent whole. In visualisation we need to make selections about what analysis we are going to portray to our audience in order to satisfy our articulated curiosity.[49, p. 119]

編集とは、映画の中にどのシーンを残すか、本にどの内容を残すか、音楽をどのようにアレンジして全体として一貫性を持たせるか、といった選択をすることである。情報可視化において私たちの明確な関心を満たすためには、どのような分析を閲覧者に見せるかについて選択する必要がある。（筆者訳）

編集的思考は取得したデータを表現しうる全ての可視化の選択肢の中から制作者自身の責任と判断において選択することである。この選択的プロセスはエキスパートではない一般の閲覧者に対して行う情報可視化に特有なものと言え、前節で述べたシステムとしての情報可視化設計と対比すると特徴的である。Kirk はこの取捨選択の指針を、情報可視化

によって取り上げる主題について、閲覧者に生じる経験と、制作者が取りうる選択の適合性 (relevance) と説明している。つまり、閲覧者が主題を理解することにどのような意義があるのかという観点と、制作者がこの主題についてデータから何を可視化することができるかという観点がうまく噛み合っていることが重要であると主張する。また、Kirk はデータを整え、編集的思考で吟味することは反復的であるとしている。これは、データが整えられた結果として自ずと適合性が見出されるわけではなく、整えられたデータを題材として編集的思考によって吟味した結果がデータの整え方に影響を及ぼすことがありうることを意味している。

デザイン手法の開発は具体物としての情報可視化の設計行為自体を指す。つまりは、前節で触れた Munzner の Nested Model における第 3 階層「Encoding / Interaction Idiom」に相当する。もし、インタラクティブティや動的な表現を取り込んだ情報可視化を開発するのであれば、第 4 階層の「Algorithm」も含まれるだろう。Munzner のモデルとの違いは、エンコーディングやインタラクションだけでなく、レイアウトや余白の活用といったグラフィックデザインにおける視覚的コントロールについても重要性を指摘している点である。そのことが端的に現れている箇所を引用する。

Elegant design: Like colour, composition decisions are always relative: an object's place and the space it occupies within a display create a relationship with everything else in the display. Unity in composition provides a similar sense of harmony and balance between all objects. The flow of content should feel logical and meaningful.[49, p. 292]

エレガントなデザインには、色彩だけでなく構成も常に関係する。ディスプレイの中のオブジェクトの位置と空間は、ディスプレイ内の他のあらゆるものとの関係を作り出す。構成に統一性を持たせることで、すべてのオブジェクトの間に調和とバランスをもたらす。コンテンツの流れは、論理的で意味のあるものでなければならない。(筆者訳)

これらは、一般の読者の興味や関心を高めると同時に見た目心地よだけでなく、論理性と感覚を融合させた優れた視覚的体験をもたらすことを意図したものである。

2.6 情報可視化システムと美観に関する研究

2.2 節では、初期の情報可視化がコンピュータサイエンス分野の HCI とソフトウェア工学の文脈で広く研究されたことに触れた。しかし、この研究潮流には、コミュニケーション

ンメディアとしての情報可視化において重視されていたようなグラフィックデザインのコントロールはほとんど取り入れられてこなかった。この理由を明確にすることは難しく本研究の趣旨と異なるが、考えうるいくつかの背景は以下の通りである。

初期の HCI 分野の情報可視化研究の多くは課題解決型、技術志向型研究である。大規模なデータをインタラクティブに可視化してコミュニケーション用途に使うことはほとんど普及しておらず、グラフィックデザインを取り入れる大きなニーズが生まれていなかった。そのようなコミュニケーションとしての情報可視化は静止的なグラフィックデザインが主流であった (1.7.6 の『Understanding USA』を参照)。また、情報可視化研究の多くが計算機科学など技術的な教育を受けた人々によって行われていたため、芸術やデザインにおける視覚表現との接点は限定的であったことが考えられるだろう。

本研究は大規模で複雑かつ専門的なデータを可視化し一般の鑑賞者にとって現象を多様な視点から眺め洞察へと導く情報可視化に焦点を当てることから、美観 (aesthetics) は重要なキーワードである。

Cawthon と Vande Moere は一般の人々を対象とした実験結果から、情報可視化に対する美観の感覚とそれによるタスクのパフォーマンスには相関があることを示した [4]。Cawthon らはこの結果から、Norman の主張

It is only through our emotions do we unravel problems, as the human emotional system is intertwined with our cognitive abilities.[50] 人間の感情システムは認識能力と絡み合っているため、問題を解くのは感情を通じてのみである (筆者訳)

は支持されるとし、情報可視化における美観の重要性を示した。

2.7 情報可視化と芸術性に関する研究

Lau と Vande Moere は、工学的文脈と芸術分野にまたがる情報可視化の研究を「Information Aesthetics」と称し、それぞれの研究分野の特徴を位置付ける見取り図的なモデルを提案した (図 2.5) [11]。このモデルでは「抽象的なデータを表現する」「インタラクティブなインターフェースを提供する」「視覚的な魅力でユーザーを引きつける」の 3 つの課題に焦点を当て各研究領域をマッピングする。

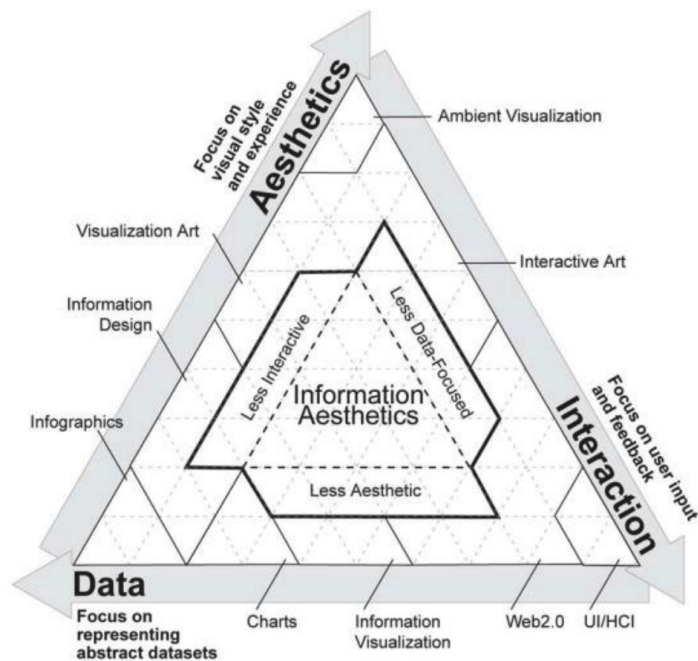


図 2.5 Domain model for information aesthetics[11] より引用

Lau らはさらに、この複合領域にある 47 の情報可視化アプリケーションを定性的に分類した (図 2.6)。ここで用いられる 2 つの評価軸は、ひとつは制作者が選択したデータの表現手法 (Mapping Technique) が直接的か、解釈的か、もうひとつは利用者が可視化を通じてデータから引き出すことができる意味 (Data Focus) が、内在的か、外在的かである。表現手法が直接的な作品ほど、2.2 節で述べた一般的な情報可視化のガイドラインに則った手法が取られており、より直接的にデータの値を知覚することができる。一方、解釈的な表現は一般的な情報可視化ガイドラインにとらわれず作者自身の表現が反映されるものであり、表現から個々のデータの値を正確に読み取ることは困難である。

内在的な作品ほど、外れ値、傾向、クラスタなどデータそのものが持つパターンを知覚しデータに対する洞察に目を向けさせるものである。一方、外在的な作品はデータセットの根底やその背後にある意味の発見を促すものであり、利用者の個人の考察を呼び起こすことを主眼に置く。

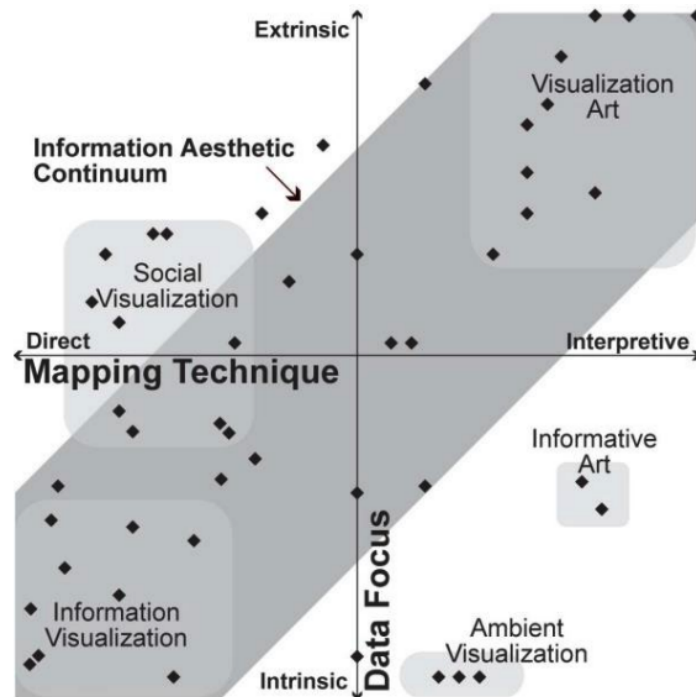


図 2.6 Lau らによる Categories within the model of information aesthetics [11] より引用

Lau らはこの考察から、芸術的な情報可視化はパターンの提示よりもデータの根底にある意味の表現を促進し、開かれた解釈を誘発することを目的としていることを指摘した。また Information Aesthetics の領域は従来の情報可視化と芸術的な情報可視化の連続した分布であることを指摘した。技術面的側面では、芸術的な情報可視化は従来の情報可視化から取り入れた既存のデータマッピングの手法を再文脈化し、その科学的信頼性に疑問を投げかけ、あるいは、信頼性に依拠する人間の意見や態度を揺さぶることを意図する傾向があることを指摘している。

2.8 情報可視化とグラフィックデザインに関する研究

この節では、コンピュータをコミュニケーションメディアと捉え、ディスプレイ上のインタラクティブな情報表現にグラフィックデザインの観点を取り入れる研究を述べる。

コンピュータのグラフィカルなアルゴリズムやインターフェイス、インタラクティブティとグラフィックデザインの知見を接続するアプローチの端緒においては、MIT Media Lab の Muriel Cooper らによる「Visible Language Workshop (VLW)」の活動と研究がある [51]。

Cooper はマサチューセッツ芸術大学でデザインの学士号と教育の学士号を取得。パウ

ハウスのデザイン原則を習得し、ニューヨークを中心に広告や出版などのグラフィックデザイン分野でその業績が高く評価されたデザイナーである。1952年、MIT Pressに入社し、1967年、MIT Pressの初代アートディレクターに就任。1974年にはMITで「Messages and Means」というコースを教え始め、タイポグラフィ、グラフィックデザイン、テクノロジーの教育に取り組む。1977年には、グラフィック・デザイナーとして初めて建築学科の教員となり、VLWを主宰し指導した（のちにVLWはMIT Media Labのリサーチグループとして移設された）。

VLWでの研究の多くは、メディアとしてのコンピュータには、従来のメディアとは異なる能力、異なる制約があることに着目している。自動レイアウト、プログラムによるデザイン生成、キネティックな情報表示などのテーマが研究された[52]。Cooperによれば、コンピュータにおけるグラフィックデザインの探求は、情報がどのように画面に表示されるかだけでなく、情報を機械のアーキテクチャにおいてデザインするための新たな視覚言語を探索することである。

Cooperらが発表した「Information Landscape」[53]はこのようなデジタルメディアの特性を具現化した情報可視化のショーケース的作品である。同作品は、Richard Saul Wurmanが主催する1994年のTED Conferenceで公開され話題となった。ディスプレイに表示された空間を連続的に視点を移動させながら、フライトデータ、リアルタイム通信ネットワーク、株式レポートなどを題材としたビジュアライゼーションを「Work in progress」として紹介した（図2.7、2.8、2.9）。これらの3次元空間におけるレイアウトやタイポグラフィを駆使した表現は、当時としては高度なCG技術とグラフィックデザインを融合させた情報可視化を実現したものである。



図 2.7 「Information Landscape」のイントロダクション [54]。キネティックな 3D タイポグラフィとズームングによるシームレスな表現

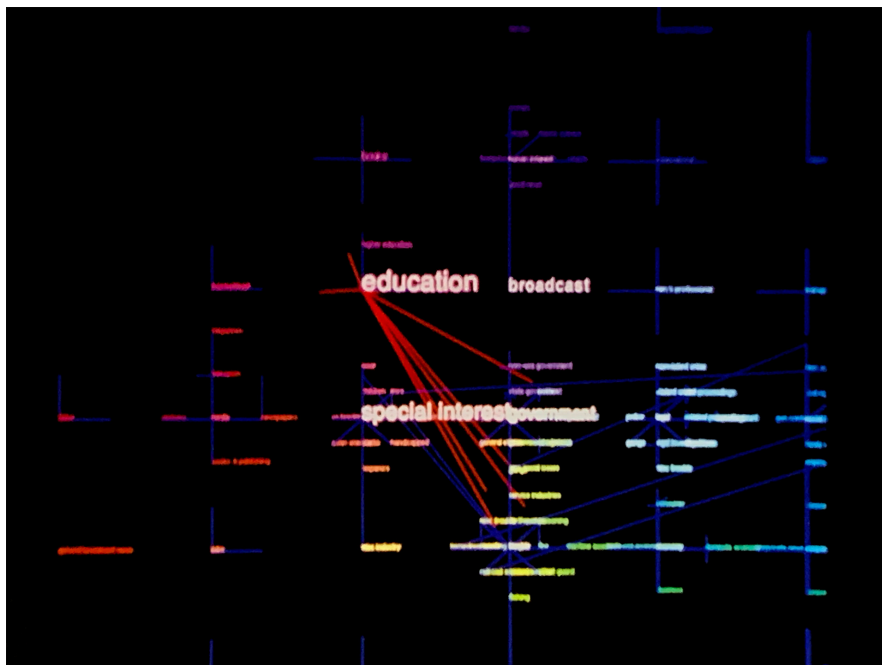


図 2.8 VLW に所属した Earl Rennison による「Browsing The Net」 [54]

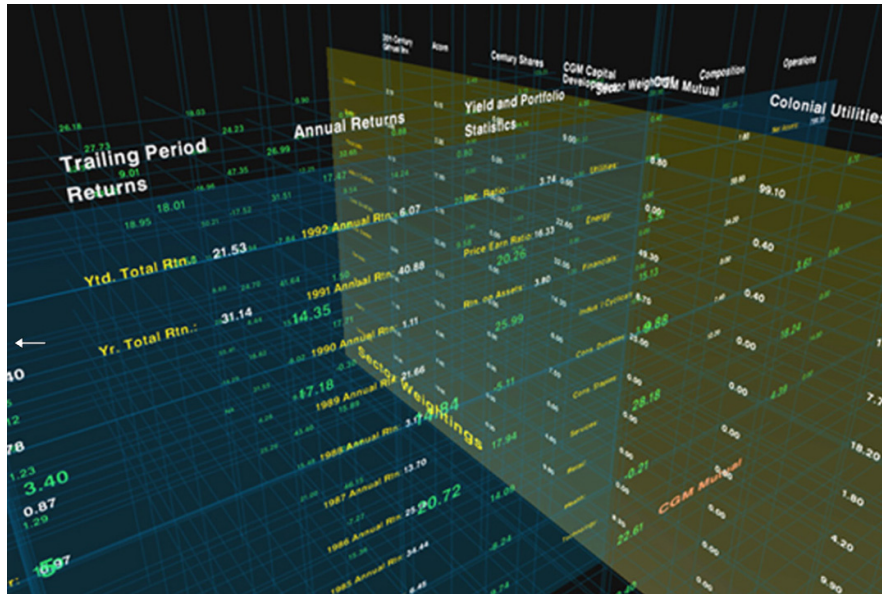


図 2.9 VLW に所属した Lisa Strausfeld による「Financial Viewpoints」 [55]

2.9 美観を伴う情報可視化デザイン手法の研究

本節では、グラフィックデザインの知識を持つ開発者が美観に配慮された情報可視化を開発するプロセスについての研究を述べる。

2.9.1 Computational Information Design

「Computational Information Design (CID)」は Ben Fry が提案した情報可視化のための設計プロセスモデルである [56]。Fry は計算機科学とグラフィックデザインのバックグラウンドを持つ情報デザイナーでありソフトウェアエンジニアである。Fry は先述した VLW の研究を引き継いだ John Maeda の指導のもとで研究を進め、CID を博士論文にまとめている。CID について、Fry は次のように述べる。

The thesis describes the computational designer as more of a choreographer, whose task is to order and move the elements of a continually changing space.[56, p. 50]

本論文では、コンピューショナルデザイナーを、絶えず変化する空間の要素を秩序立てて動かす振付師のような存在と表現している。(筆者訳)

つまり、CIDの志向は、情報をデザインする上で最終的な段階で静止的なレイアウトや色彩表現を担当するのではなく、データから情報可視化までの全ての段階に関わり、データ、表示、インタラクションなど可変する要因を理解し、それらが調和するようにコントロールするシステムをデザインすることである。

Fryは、大規模なデータを題材にして視覚的に洗練された情報表現やインタラクションを含んだ制作プロセスをモデル化することに焦点を当てている。Fryは全米の全ての住所と郵便番号のデータや、ヒトゲノムの全シーケンスデータなどの大規模なデータを題材として、自ら行った情報可視化システムプロジェクトのプロセスを記録し、その考察からCIDモデルを構築している。観察の対象プロジェクトは、専門家が利用するもの、一般の理解のために提示されるもの、芸術作品として提示されるものがある。プロセスは、グラフィックデザインとソフトウェアエンジニアリングの両方の知識を有するFry個人の視点から記述されており、具体的でボトムアップ的である。この点においてMunznerのモデル(2.2節)やKirkのモデル(2.5)とは異なる。

またFryは、デザイナーやアーティストが視覚的表現の実践やプロトタイプを支援するためのプログラミングライブラリである「Processing」[57]の開発者としても知られる。Fryはこの研究においてProcessingをこのプロセスを支援するツールとして位置付けている[56, p. 123]。

2.9.2 CIDの事例と作品の特徴

ここではFryがCIDによって制作した作品群を概観し、3作品について解説する。図2.10～2.14は、作品群のスクリーンショットである*1。

*1 ここで取り上げた作品は、当該論文に掲載された作品に限ったが、Fryのウェブサイトにはここに含まれない作品が多数紹介されている。<https://benfry.com/projects/>

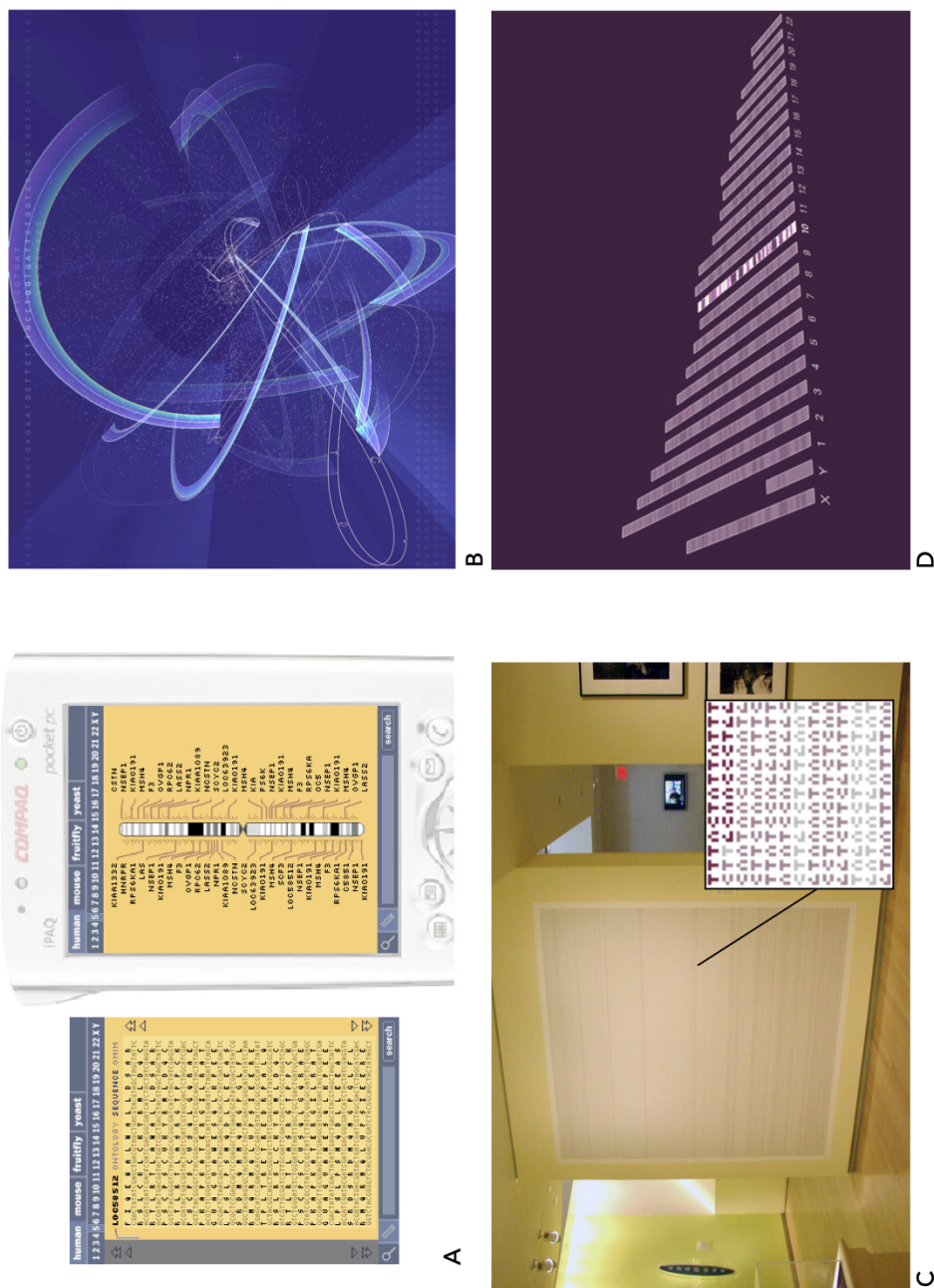
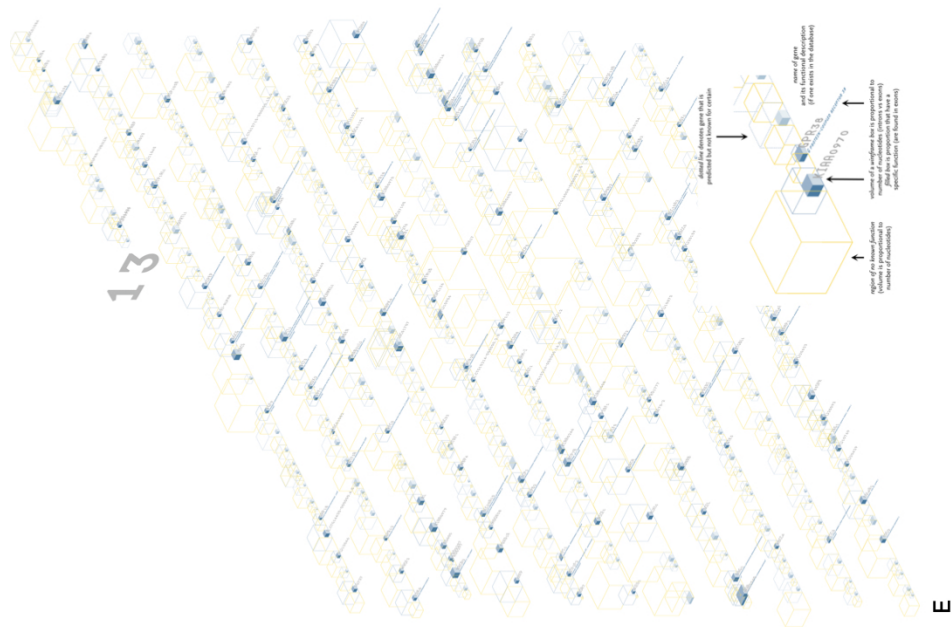


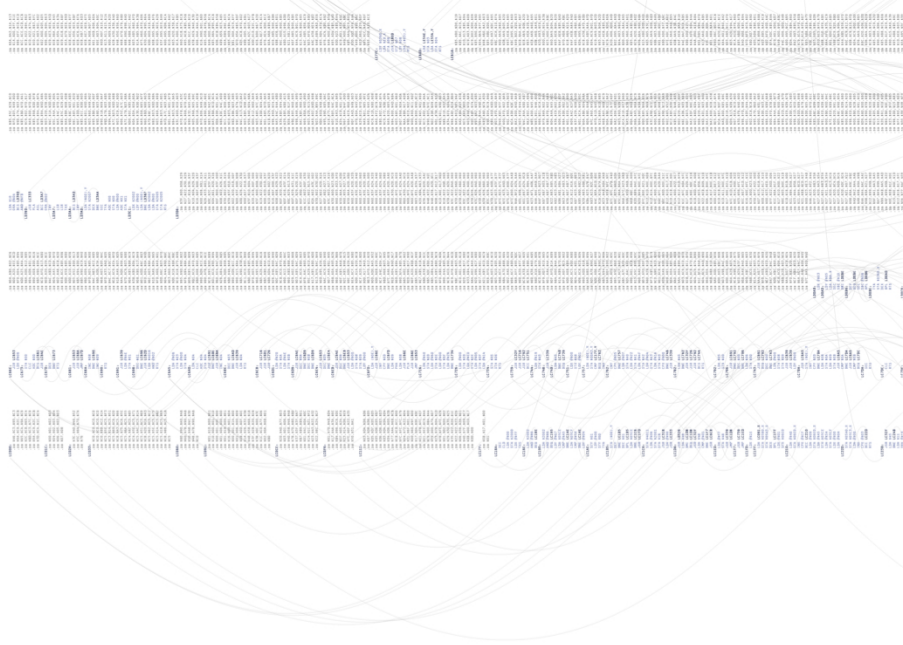
図 2.10 CID による Fry の作品群 [56, pp. 133–161]: (A). Handheld Genome Browser Prototype, (B). Genome Valence, (C). Chromosome 21, (D). Strippy



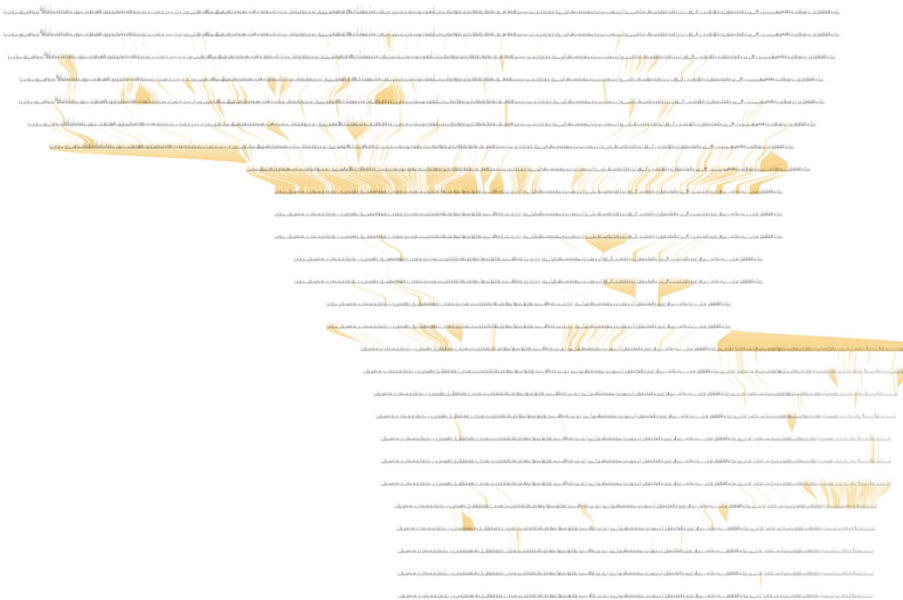
13

F

図 2.11 CID による Fry の作品群 [56, pp. 133–161] : (E).Axonometric Introns and Exons, (F).Haplotype Lines



H



G

図 2.12 CID による Fry の作品群 [56, pp. 133–161] : (G).Revisionist, (H).Dismap

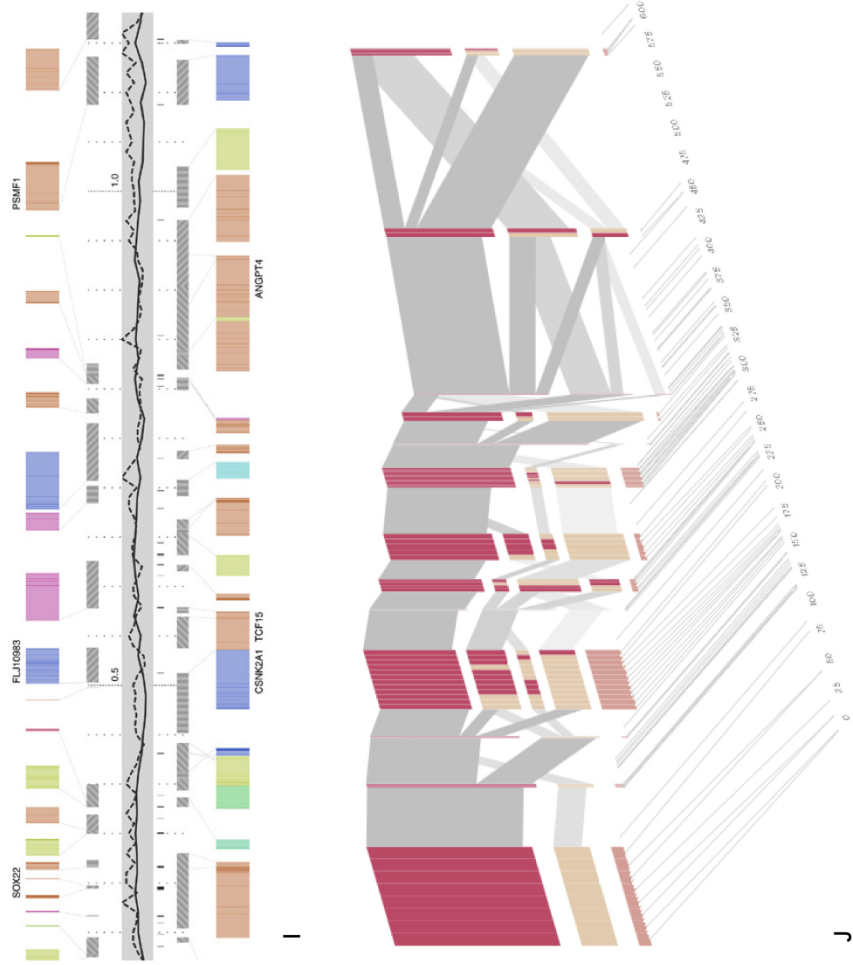


図 2.13 CID による Fry の作品群 [56, pp. 133–161] : (I).Redesign of gff2ps, (J).isometricblocks

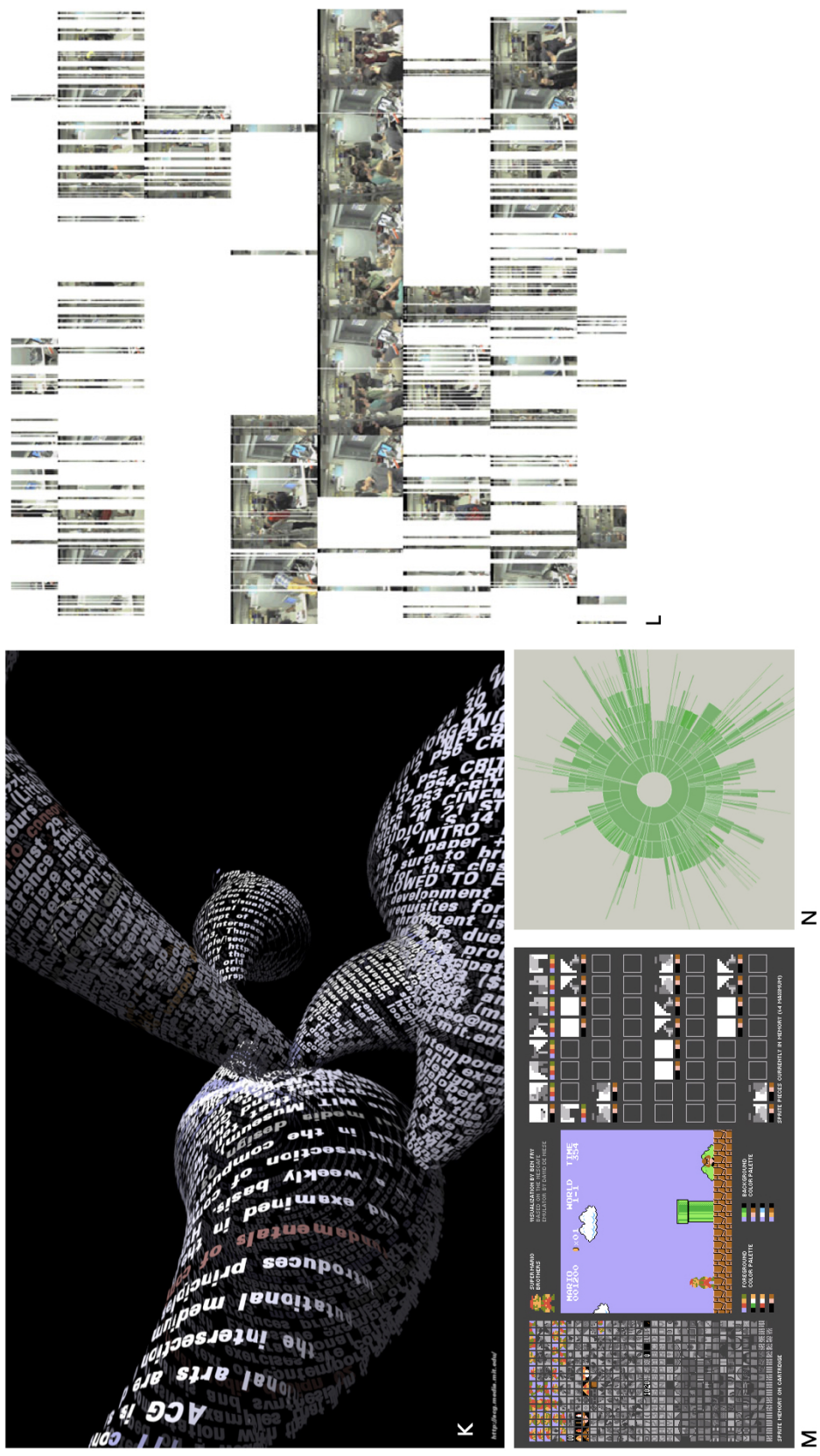


図 2.14 CID による Fry の作品群 [56, pp. 133–161] : (K).tendril, (L).shopactivity, (M).Deconstructulator, (N).radialsize

Genome Valence

『Genome Valence』（図 2.10-B）はゲノム配列解析における標準的なツールである BLAST（Basic Local Alignment Search Tool）アルゴリズムを可視化した作品である。BLAST では、DNA の塩基配列を特定のシーケンス（例えば GACTTAGTA）を異なる生物同士で検索することができる。これによりある生物で特定された遺伝子のシーケンスを他の生物で検索し類似したパターンを調べることができる。この作品では配列を表す a,c,g,t を 9 文字ずつに分割して空間上の点で表し、点は中央から順に並べられ、最も一般的なセットは外側に、一般的でないセットは中央に向かって配置される。上部には、検索したい配列が表示され、その配列の中で見つかった 9 つの文字の集合に対して、空間内のその点と次の 9 つの文字の集合を表す点との間に弧が描かれる。

Chromosome 21

『Chromosome 21』（図 2.10-C）は 21 番染色体の 1300 万文字の遺伝情報を、ヒトゲノムのスケールで表現した 2.5 メートル四方に印刷されたインスタレーションである。タンパク質をコードする遺伝子領域（exons）は濃色で、機能が知られていない領域は灰色で表される。1 平方センチメートル当たり約 220 文字を配置できる専用のフォントを開発して出力された。21 番染色体は、人間の染色体の中でも最も短いものの一つであるが、この画像では、約 5000 万文字で構成される染色体の 4 分の 1 しか表示されていない。ヒトゲノム全体の遺伝情報を描くには、このサイズの画像が 250 枚必要でその文字数は 31 億文字になる。

Isometric Blocks

『Isometric Blocks』（図 2.13-J）は 500 人の集団のゲノムデータ内の 5q31 セクションの SNP とハプロタイプブロックを可視化している。SNP とは個体同士のゲノム DNA を比較した際、1 塩基が異なる現象である。SNP は 1000bp に対して約 1~3 箇所程度の割合で出現する。SNP は全て表現型変化に影響するわけではなく、タンパク質をコードする遺伝子領域（exons）でアミノ酸変異を起こすものなどその可能性が高いものとそうでないものが混在している。ハプロタイプは複数の SNP を組み合わせた領域であり、親から子へと引き継がれる。3 色に塗り分けられた縦長のバーが SNP を表し、バーを横方向に連続させた矩形がハプロタイプブロック（ブロックを定義するパラメータは画面上で任意に変更できる）を表す。

SNP の色は変異のバリエーションを表し、最もメジャーな変異を濃色、あまり見られ

ない変異を淡色、集団の 5% 以下にしか発生しない変異をやや濃い色で表している。各部の縦方向の長さは全集団に対する出現割合を示す。また奥行き方向に LD (Linkage Disequilibrium) 値を割り当てることができる。その他いくつかのモードがあり、このソフトウェアによって、ハプロタイプの特定や集団における遺伝的疾患との関係を探ることができる。

作品の特徴

CID による Fry の作品は創造性に富み魅力的な美観を備えている。また、作品群には専門家の支援を目的とする情報可視化システムと、一般に向けたコミュニケーションを目的とする情報可視化のどちらのタイプの作品も含まれる。前述の 3 作品 (B,C,J) については 2.7 節で述べた Lau らの分析 [11] に含まれている。各作品と比較した結果を図 2.15 に示す。

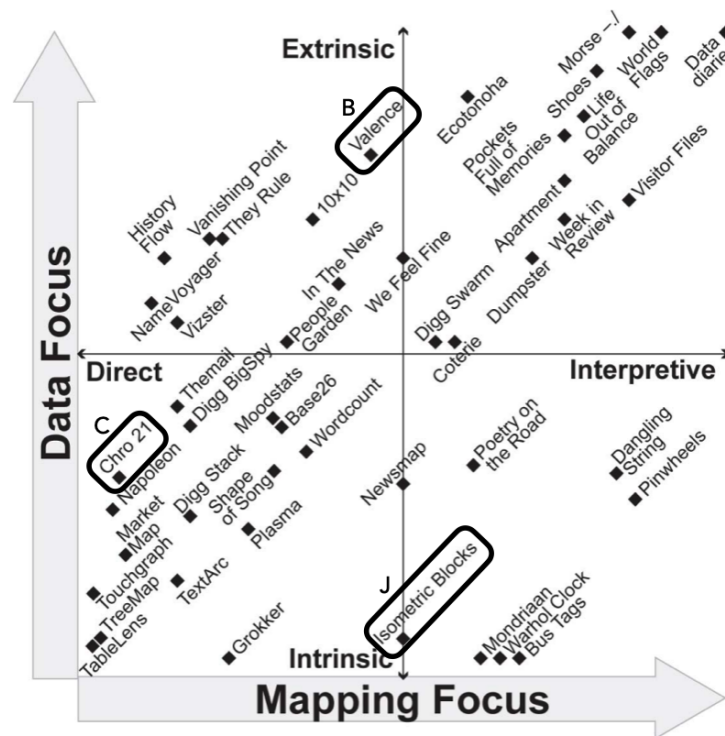


図 2.15 Lau らの分析による作品 B,C,J の分布 [11] より引用 (一部筆者加筆)

Isometric Blocks (J) は他作品に比べて内的度合いが高い。対照的に Genome Valence (C) は外的度合いが高い。これらから、CID はデータ内部の意味を求める情報可視化と、

データの背後にある意味を求める情報可視化の両方に対応できることを示唆している。

2.9.3 CID のプロセス

Fry は「データを理解するプロセスは、一連の数値と、そのデータに関する疑問に答えるという目標から始まる。このプロセスのステップは次のように説明できる（筆者訳）」と述べており [56, p. 13]、データから読み取りたい情報や仮説を設定した上で、行うべきステップを次のように定義している。

1. acquire – the matter of obtaining the data, whether from a file on a disk or from a source over a network.
2. parse – providing some structure around what the data means, ordering it into categories.
3. filter – removing all but the data of interest.
4. mine – the application of methods from statistics or data mining, as a way to discern patterns or place the data in mathematical context.
5. represent – determination of a simple representation, whether the data takes one of many shapes such as a bar graph, list, or tree.
6. refine – improvements to the basic representation to make it clearer and more visually engaging.
7. interact – the addition of methods for manipulating the data or controlling what features are visible.

(以下、筆者訳)

1. 「取得」 ディスク上のファイルやネットワーク上のソースなど、データを取得する
2. 「解析」 データの意味するところを構造化し、カテゴリーに分類する
3. 「フィルタ」 目的のデータ以外のものを取り除く
4. 「マイニング」 統計学やデータマイニングの手法を応用してパターンを見出したり、数学的な文脈において分析する
5. 「表現」 データが棒グラフ、リスト、ツリーなど様々な形をしているかどうかに関わらず、簡潔な表現を特定する
6. 「精緻」 基本的な表現をより明確に、より視覚的に魅力的にするために改良する
7. 「インタラクション」 データを操作したり、どの機能を表示するかを制御する

方法を追加する

Fry は、図 2.16 に示すようにこれらの個々のステップで必要となる知識や技術が、独立した研究分野によって蓄積されてきたことを指摘し、次のように述べる。

Part of the problem with the individual approaches of dealing with data is that the separation of the fields leads to each person solving an isolated part of the problem, and along the path towards a solution, something is lost at each transition—a “telephone game” for context, where each step of the process diminishes aspects of the initial question under consideration. [56, p. 13]

データに対してそれぞれの分野に固有のアプローチだけで対処するならば、各分野の人間がその分野に切り分けられた部分の問題だけを解決することになり、このプロセスを経るごとに伝言ゲームのように何かが抜け落ち、プロセスの各段階で検討すべき本質的な問いの側面が減少する。(筆者訳)



図 2.16 CID の 7つのステップと研究分野 [56, p. 13] より引用

Fry は出版の分野ではコンピュータの技術的革新によって、レイアウト、植字、印刷といった別々の専門知識が統合され効率化しただけでなく、Desktop Publishing(DTP) という柔軟で全く別のデザインプロセスをもたらしたことを引き合いに、CID の目的を次のように述べている。

The intent of Computational Information Design is not to make a interdisciplinary field, but rather that single individuals should be able to address the spectrum of issues presented—the goal is to understand data, so one needs to be familiar with the issues in handling data from start to finish.[56, p. 87]

CID の目的は、学際的な分野を作ることではなく、むしろ一個人が立ち現れる幅広い問題に対処できるようにすることである。目的はデータを理解することであるため、個人が最初から最後までデータを扱う際の問題に精通している必要があるの

だ。(筆者訳)

このような観点から、CID は一人のデザイナーがデータから情報可視化に至る必要な幅広い知識と対処能力を有し、コンピュータをインタラクティブなメディアとしてデザインすることを目的として構成されている。また、その実践を支えるツールとして Processing を位置付けている。

Nested Model (2.4 節) や、Kirk のモデル (2.5 節) では原則として前方のステップは後方のステップを拘束する構造を持っている。特に表現 (エンコーディング) は探求するものではなく特定するものと位置付けている。一方、CID では一般的に表現やインタラクションをなど後方のステップを担当するデザイナーが前方のステップに対してもアプローチできるため、後方ステップの結果を前方のステップにフィードバックすることができる (図 2.17)。

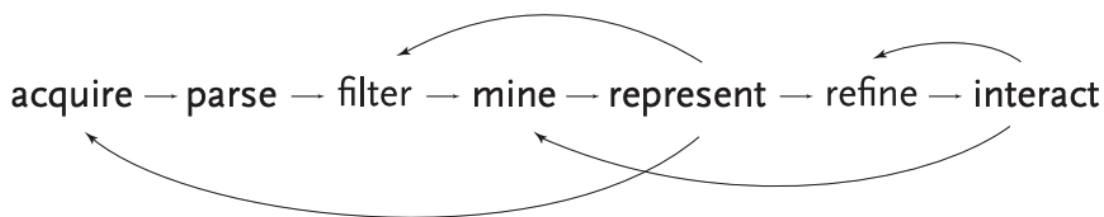


図 2.17 ステップのフィードバック [56, p.30] より引用

このようなフィードバックが何をもたらすのか。Fry は、先述したヒトゲノムプロジェクト内の事例で示している。

図 2.18 は、このプロジェクトで開発されたインタラクティブな情報可視化システムのスクリーンショットである。これは、500 人のゲノムの特定のセクションにおける変異の頻度データを可視化するソフトウェアで、ゲノム分析の専門的な利用者を想定している。5 つのグラフは同じデータを使用しているが、利用者が持つニーズの違いに沿って、2D と 3D、質的・量的の異なる表現で描かれる。これらのビューはインタラクティブかつトランジションを伴って切り替えることができるため、利用者のコンテキストを損なわずにさまざまな視点から情報を読み取ることができる。

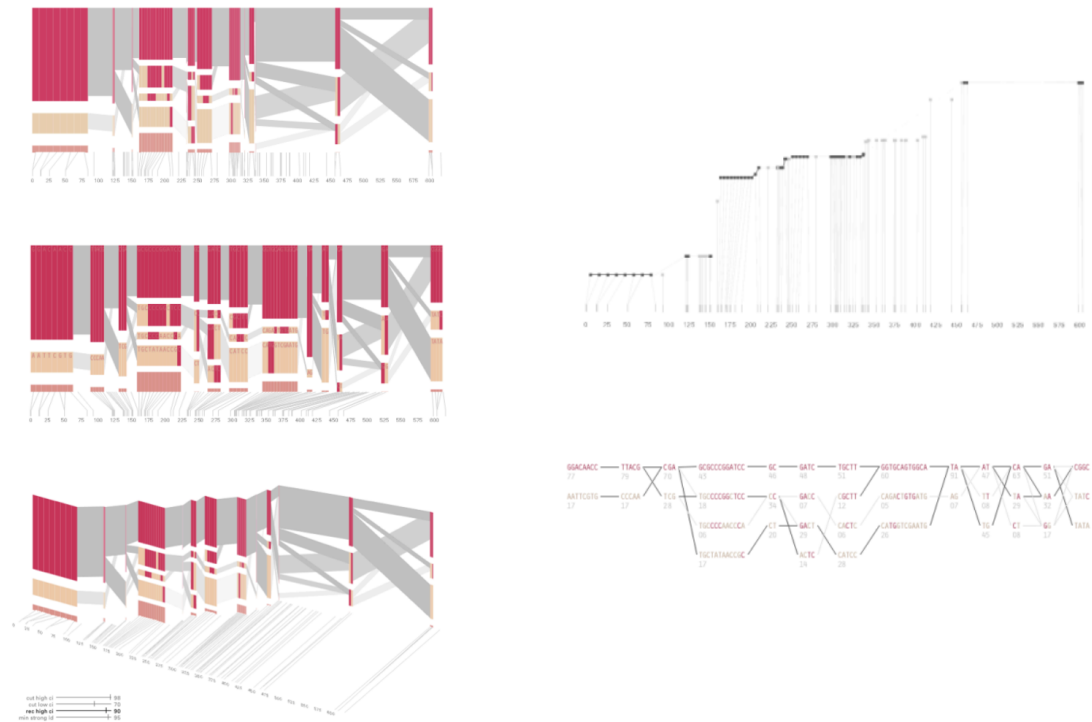


図 2.18 ゲノムデータプロジェクトで開発されたソフトウェアのビュー [56, pp. 82–85] より引用

この制作プロセスでは、図 2.19 に示す通り、ひとつのビューをインタラクションまで含めて実装した後で、別のビューに遷移する際にデータ全体から描画に必要なデータ読み込む「Filter」、描画アルゴリズムの開発のための「Mine」、量的データと質的データで表現を変更する「Represent」、限定された表示領域内に収めるための「Refine」が行われた。

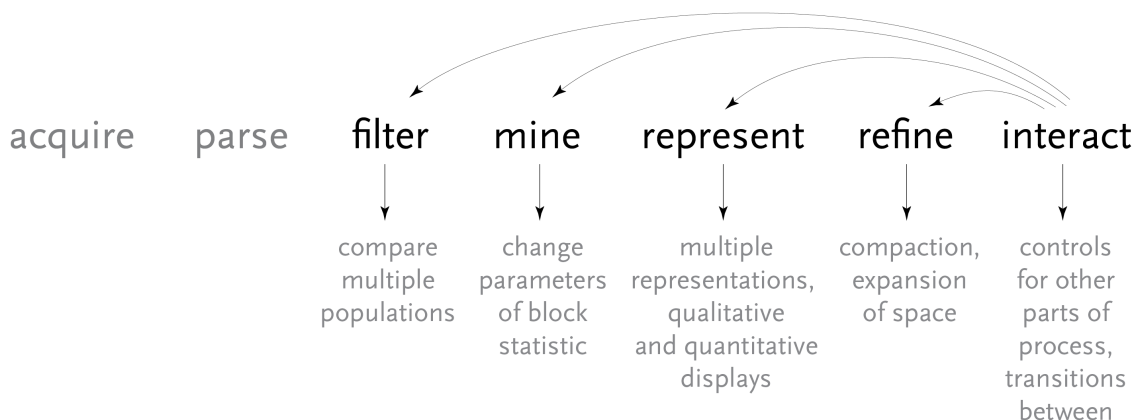


図 2.19 ゲノムデータプロジェクトでのフィードバック [56, p. 86] より引用

Fry は、このように多くのパラメータを含むデータにおいて、それぞれ異なる複数の

ビューを柔軟に切り替えられるシステムの開発について次のように述べている。

The many dimensions in genetic variation data necessitate multiple perspectives for how it is viewed, and an interactive software visualization provides a means to transition between these views in an informative manner, showing how the many views are related, yet at the same time highlight different aspects of the data set.[56, p. 86]

遺伝的変異のデータには多くの次元があるため、それをどのように見るかについて複数の視点が必要である。インタラクティブなビジュアライゼーションは、豊かな情報を保ったまま各視点間を遷移できる手段を提供し、多くの視点がどのように関連しているかを示すと同時に、データセットの異なる側面を引き立てることができる。(筆者訳)

複雑なデータほど多面的な表現方法があり、表現方法によって得られる意味が異なることが多い。このように複数のビューをシームレスに繋ぐ表現は事前に机上で検討することは難しく、実装を試しながら前方のステップを調整する以外に方法はない。また、単独ビューであっても、この例のようにイディオム化されていない全く新しいグラフを考案することは実装と切り離して机上で行うことはできず、検討段階から実装してみる以外に方法はない。

2.9.4 CID の影響と現在の状況

Processing は、クリエイティブコーディング (Creative Coding) やジェネラティブアート (Generative Art) を実践する上での入門的な環境としてその後も開発が続けられているだけでなく、多くの利用者に支持され、さまざまなタイプのデジタルアート作品開発を支援してきた。また、デザイナーやアーティストが Processing を使用して斬新なデータビジュアライゼーションに取り組んだ事例も見られる [58, 59]。

Processing の思想は他の言語や環境にも影響を与えた。C++ 言語のフレームワークである openFrameworks、JavaScript ライブラリの p5.js、オープンソースハードウェアの Arduino など、その思想を受け継ぐツールが登場し、デザイナーやアーティストが直接プログラミングすることで自らアルゴリズムや実装レベルの探求を幅広く行うことを可能にした。これは、前節でみた Cooper らが VLW で実践した、コンピュータとデザインの深いレベルで接続する探求的な試みを誰もがに行えるようにした点で大きな貢献と言えるだろう。一方それとは対照的に、CID が目指した従来の情報可視化プロセスを統合し新しいグ

ラフィックデザインプロセスの地平を開くような取り組みは多いとは言えない。

この要因を特定することは困難だが考えうるいくつかを述べると、CID が発表された 2000 年代前半はいわゆるビッグデータ以前であり、現代に比べて大規模で複雑なデータが少なかったこと、ビッグデータ以後には爆発的にデータが増え、情報可視化よりも統計解析や機械学習の技術革新が進んだことが影響していると考えられる。そのような流れを受けて、情報可視化デザインはデータサイエンスのニーズの高まりから探索的データ分析を支援するグラフィックライブラリ^{*2}や、ストーリーテリングを行うニュースサイトにおけるインタラクティブなインフォグラフィクスを実現するライブラリ^{*3}などに特化した形で発展しているものの、依然として図 2.16 のような、分断した状況となっている。

2.10 本研究の位置付け

序論では、これまで人間が持ち得なかったような大規模で複雑なデータが増え、その多くは一般の人々の目には届かない形で個人や社会の重要な意思決定に影響を与えることが日常化しつつあることを述べた。

このような大量のデータに内在する意味についてさまざまな方向から光を当てるためには、すでに開発されたグラフ形式や視覚的エンコードのパターンを利用して情報可視化を行うだけでなく、新たな表現形式を探求する必要がある。序論では、そのような新たな表現形式は社会の状況の変化と呼応するようにグラフィックデザインの文脈で育まれてきた歴史的経緯を概観した。前述した現代の状況の中で、グラフィックデザインやビジュアル表現を得意とする人々が積極的にデータに対する新しい見方を探求し提案することが望まれる。そのような探求を行う上での足掛かりとなる枠組みは存在するだろうか。存在しないとすればそれをどのように構築することができるだろうか。これが本研究の主題である。

本章では、情報可視化の開発フレームワークについて、主に専門家の問題解決ユースを想定したアプローチである Munzner の Nested Model と、一般に向けたコミュニケーションユースを想定した Kirk のモデル、表現探究型のアプローチで問題解決とコミュニケーションの両方を想定した Fry の CID を参照した。これらはいずれも元データ (Raw Data) が視覚表現へと変換される一連のプロセスの中で、表現およびその実装は最終的な

^{*2} データ解析に用いられる主な可視化パッケージには、R の ggplot2[60]、Python の matplotlib[61] などがある

^{*3} インタラクティブなビジュアライゼーションをブラウザで実装するパッケージには、D3.js[62]、three.js[63]、deck.gl[64] などがある。

ステップに位置付けられていた。

Nested Model は専門分野での特定のタスクを同定し効率化することを動機としたモデルであるため、表現やアルゴリズムの探究はこうした動機に対して支配的な影響を受ける構造を持っている。

Kirk のモデルは、編集的思考を取り入れ、制作者がデータから得られるさまざまな情報の中で、鑑賞者に伝えるべきメッセージを厳選するプロセスが特徴的である。編集的思考が先行するため、表現やアルゴリズムの探究はこの選択されたメッセージの中でのみ行われる。

CID は全てのステップについてデザイナーが熟知することを前提として、表現やインタラクションの実装も前方のステップの判断を問い直すための材料として積極的にフィードバックする点が、前述の二つのモデルとは異なっている。CID は表現の試行探索を繰り返し、データを多面的な視点で掘り起こしていくことから、本研究の主題と合致するフレームワークであると考えられる。

一方で、CID には次のような検証すべき課題が残されている。一つは、具体的なプロジェクトの記述のみに留まっており、他のプロジェクトで利用するには抽象化し応用性を高める必要がある。次に、Fry 本人の主観的な記述が中心となっている点で、フィードバックが起こる契機となるような議論や、プロジェクトに関わるメンバー（例えば、ゲノムの専門家）との応答について記述がない点である。最後に、現代へ適応性である。CID はビッグデータ以前であり、現在とは状況が異なる。ゲノムデータは現在のデータのように大規模かつ複雑であるが、リアルタイムな動的性やデータの偏りなどのビッグデータ特有の事情はない。加えて、巨大なデータを扱う分野は広範囲にわたっており、デザイナーが全てに習熟することは現実的とは言えない。デザイナーとそれぞれのドメイン知識を持つ専門家とのコミュニケーションをモデルの一部として記述する必要がある。

このような観点から、本研究では CID を先行するモデルとしつつも CID が取りこぼしている抽象化や現代への適応性に着目し、デザイナーが探究的に情報可視化を構築するためのモデルの提案を行う。

第3章

本研究の手法

3.1 本研究手法の概略

本研究の主題は、デザイナー、アーティストなど視覚表現を行うクリエイターによる、大規模なデータを扱う情報可視化メディアのデザインを支援することを前提として、データビジュアライゼーションの制作プロセスを形式知化することである。このことから、まず、データビジュアライゼーション制作プロジェクトを観察し、その記録からプロセスの分析を行う。次に、プロセスのモデル化と、パターン・ランゲージの作成を通して形式知化する。得られた形式知を「データビジュアライゼーション・デザインパターン」としてデザイナー、クリエイター向けの小冊子としてまとめる。小冊子を実践の中で運用し、どのような影響があるかを調べるため、データビジュアライゼーションワークショップを企画する。参加者には小冊子を配布しその制作プロセスの観察とアンケートから考察を行う。以上の流れを図 3.1 に示す。

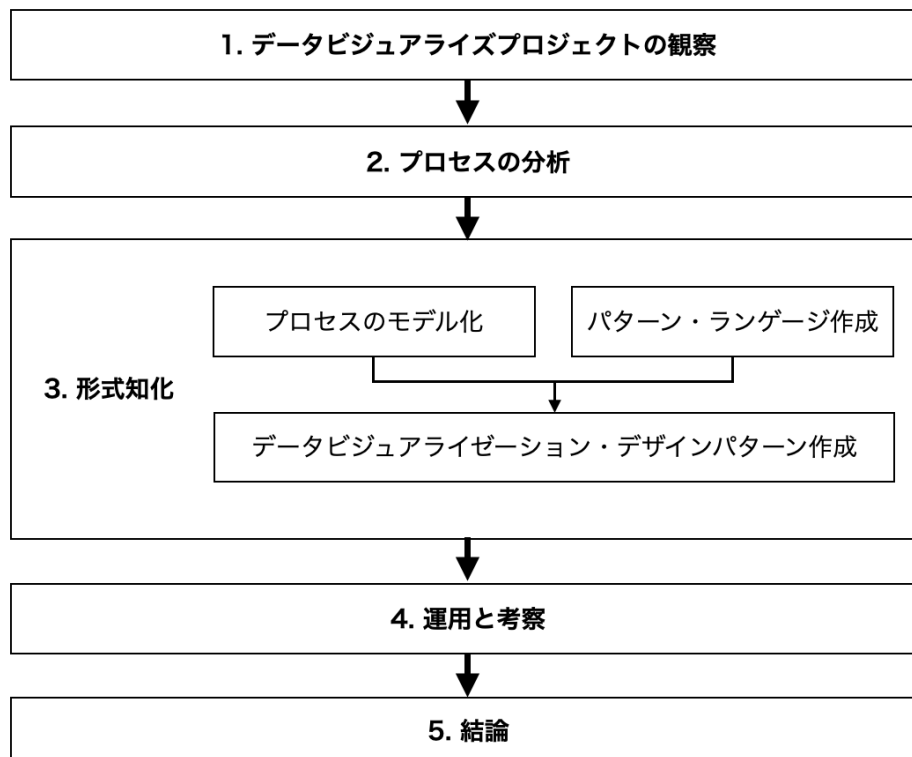


図 3.1 本研究の流れ

3.2 データビジュアライゼーションプロジェクトの調査

本研究で分析対象とするプロジェクトの詳細は4章で述べる。ここではプロジェクトの特徴を述べる。

分析対象としたプロジェクトの特徴は5つである。第1に現場での観察が可能な対象である。対象としたプロジェクトは、著者がデザイナーとしてプロジェクトの工程に参加しそのプロセスを記録したものである。第2に探究的である。つまり、あらかじめ開発する情報可視化やそれによって主張するストーリーを固定せず、さまざまな情報可視化を制作しながら視聴者（鑑賞者）に伝えるべき内容とデザインを検討するものである。第3に、横断的である。プロジェクトは可視化するデータの分野に精通した専門家、データ処理に精通した技術者、情報可視化に精通したグラフィックデザイナー、メディアやコンテンツのディレクター、インタフェースエンジニア等、分野の異なる関係者がプロセス全体に関与し、検討を加えながら進めるものである。第4に、その最終成果物は広く一般に向けて公開されるものである。制作されたデータビジュアライゼーションは報道番組の解説CG

やインタラクティブ展示の形式で一般への公開を前提に制作され、幅広い層の鑑賞者を対象としたものであり、外部からの評価が可能である。第5にビッグデータを扱う。静的で変化のないデータや変数が少なく分析の視点が限定されるデータではなく、動的でさまざまな視点から分析が可能なデータをビジュアライズする。

3.3 プロセスの分析

データビジュアライゼーション制作手法の先行研究として2章で述べた、Benjamin Jotham Fryの「Computational Information Design」(以下、CID)がある[56]。

プロジェクトの観察の記録をもとに行われたタスクを抽出し、そのタスクがCIDにおける7ステップすなわち「取得」「分析」「フィルタリング」「マイニング」「表現」「精緻」「インタラクション」(2.9.3節を参照)のいずれに相当するものであるかという観点から分類する。そして、それらのタスクの順序や反復関係に注目し、CIDで明らかになっていないタスク間の相互関係を明らかにしていく。

3.4 形式知化

3.4.1 プロセスのモデル化

前述のプロセス分析の結果から得られたタスクの相互関係から、それらが最終的なデータビジュアライゼーションにどのように作用し、どのような意味を持つかについて検討し、プロセスモデルとして一般化を試みる。

3.4.2 パターン・ランゲージによる形式知化

プロセスのモデル化はタスク間の相互関係や順序に注目してプロセスの形式知化を試みるものである。しかし、プロセスに従うだけでは、さまざまな実際の現場で起こり得る具体的な問題に対処することは難しい。そこで、観察対象プロジェクトで生じた課題やその対処に注目し、経験則を抽出したパターン・ランゲージによる形式知化を試みる。パターン・ランゲージは、ある課題を生じさせる要因と解消のための工夫、解消後のあるべき状況をまとめるものであり、経験則を形式知として他者と共有することを目的とする。

3.4.3 データビジュアライゼーション・デザインパターンの作成

上記のプロセスモデルとパターン・ランゲージを取りまとめた小冊子「データビジュアライゼーション・デザインパターン」を制作する。本小冊子は、クリエイターを読者と想定し、彼らが本冊子を読んで学習することで、前述のワークフローとパターン・ランゲージを理解し、制作に活用できることを目的とする。

3.5 運用と検証

「データビジュアライゼーション・デザインパターン」の運用と検証は次の方法で行う。デザインやアートを学びビジュアル表現のためのプログラミングを習得した学生を対象としたワークショップを実施し、制作にどのような影響をもたらすかを検証する。また、経験のあるデザイナーがデータビジュアライゼーションを制作する実務においてどのような影響をもたらすかを検証する。

第4章

プロセスの分析

4.1 分析の目的

本研究の目的は、大規模なデータを対象としたデータビジュアライゼーションのデザインプロセスを明らかにすることである。同様なデザインプロセスの先行研究として、2章で論じた Fry による CID がある [56]。CID では、データビジュアライゼーションに必要なプロセスについて、計算機科学、統計学、グラフィックデザイン、インタラクションデザインなど、個別の領域で培われた技術や知識の統合により行われるプロセスであると述べている。具体的には「取得」「分析」「フィルタリング」「マイニング」「表現」「精緻」「インタラクション」7つのステップ*1をデザイナーが熟知した上で、それらのステップを反復的かつ探索的に行うことで、デザインの質を高めていく手法である。しかし、CID の課題は、第1に、反復が発生する状況や条件が明示されていないこと、第2に、対象とするデータがヒトゲノムや全米の郵便番号など静的なデータであり、今日のビッグデータのように動的でないこと、第3に、デザイナー本人が単独でプロセスを進めており、コラボレーションによってプロジェクトを進める場合の記述がないことである。

そこで、本研究ではビッグデータを対象として前述のように専門領域の異なるメンバーからなるチームがデータビジュアライゼーションをデザインするプロセスを観察した記録をもとに、行われた作業やコミュニケーションを CID の観点から解釈する。次に、前述の7つのステップ間の反復や、参加者の相互作用について検討を行う。最後に、そのデザインプロセスのモデル化を行う。

*1 Fry による用語の定義については 2.9.3 節を参照

4.2 対象とするプロジェクト

本研究では、『新型コロナウイルスゲノム系統樹の3次元可視化』*2（図4.1）のデザインプロセスを調査対象とした。プロジェクトの概要を以下に示す。

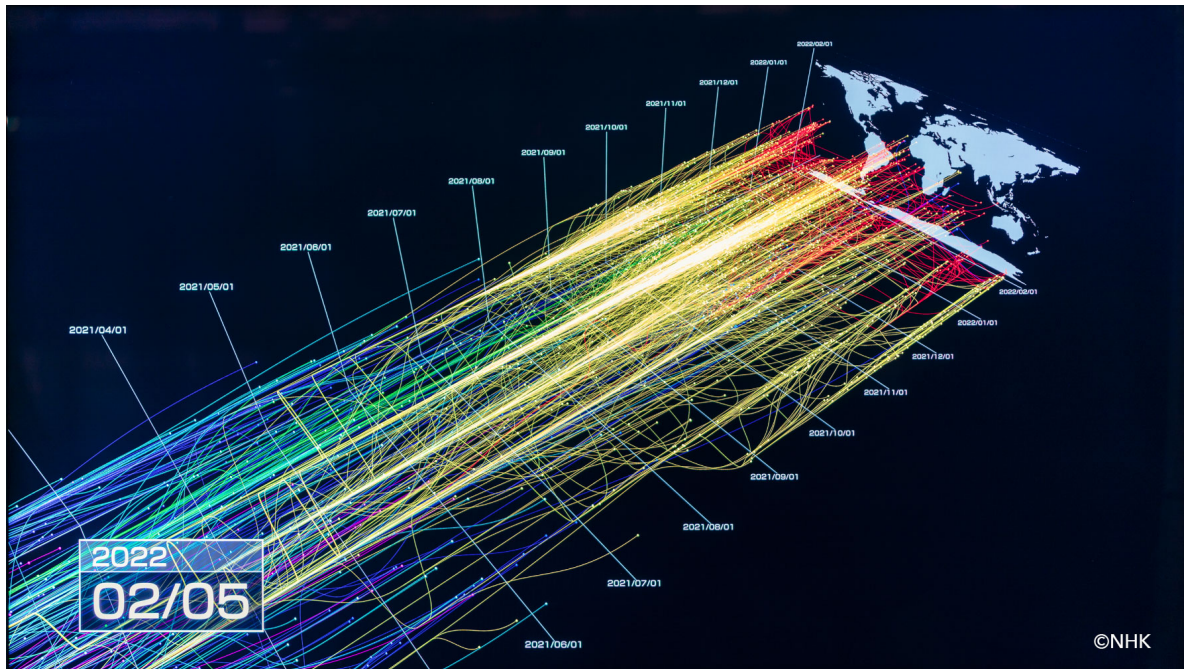


図 4.1 新型コロナウイルスゲノム系統樹の3次元可視化

4.2.1 新型コロナウイルスゲノム系統樹の3次元可視化の概要

同プロジェクトは2020年7月に開始し、新たな変異株の流行の兆しが見られるごとにデータを更新し、表現手法を調整しながらデータビジュアライゼーションを順次制作し一般に公開した。公開の形態については後述するが2022年6月までにインタラクティブ展示を3回、全国放送報道番組での解説用CG映像としての利用を4回、解説員付きのセミナー型展示を1回実施している。

本プロジェクトの主体は科学番組などの放送用映像やインタラクティブ展示を制作するE社である。本プロジェクトの動機は、大量の新型コロナウイルスゲノム解析データから、世界的に広がる変異の様子を可視化し、同ウイルスやCOVID-19感染症について

*2 プロジェクトのウェブサイト <https://timespacephylogeny.xyz/>

の先端的な研究結果から示唆される内容を一般の視聴者に届けることである。その際、研究成果として確認された特定の内容を伝えることを目的とするのではなく、研究が進行過程にある中で研究者が注目している現象や、仮説を含んだ多様な解釈のあり方を感じられることを目的としたものである。データは、世界各地で検出される変異ウイルスのゲノムデータが投稿されるオープンデータベースである GISAID^{*3}から取得した。GISAID のデータは最新の動向に合わせて更新される動的なものである。

本プロジェクトのチーム構成について述べる。番組や展示の企画と構成や取材を担当する E 社のプロデューサー (I)、科学論文の調査や取材を行う E 社のリサーチャ (K)、国内大学医学部准教授で新型コロナウイルスのゲノム解析を専門とし、科学的な見地から助言を行う生命科学研究者 (N)、データ処理や可視化プログラミングを担い、さまざまな可視化のプロトタイプを行うデザイナー (筆者)、番組用 CG の制作や一般に向けた展示用インタラクションを実装する制作会社のエンジニア (O) である。

本プロジェクトの基本方針における特色や、メンバー同士の関係性について示す。E 社と筆者はデータビジュアライゼーションの制作について委託者と受託者の関係にあったが、制作物の具体的な要求や仕様等の事前の取り決めはなかった。代わりに、I が提示する大まかな関心に沿って筆者が探索的に試作を繰り返し、その過程で得られた可視化がもつ意味やストーリーを I、K、N、筆者らが科学的に見極めながら、視聴者に向けて提供する価値があると判断される可視化を採用するという方針があった。そのため、筆者は試作について E 社から指示を受ける関係ではなく、自らの判断でデータビジュアライズ用のソフトウェアを試作し、その提供を受けた E 社が試用を通じて効果を見極め最終的な成果物とする関係にあった。O の役割は E 社から委託され、筆者と E 社の間で採用することが決まった可視化について、E 社や筆者の指示に基づき、別途 O らによって開発された放送用途や展示用途のソフトウェア上で、決定した可視化を忠実に再現することであった。

プロジェクトの外部からの評価として、その成果は 2022 年グッドデザイン賞を受賞した。また、制作過程において筆者らが考案したゲノム系統樹の 3 次元可視化手法は、国際会議での発表 [65] を経て、第 33 期可視化情報学会賞、第 49 回可視化情報シンポジウムアートコンテスト大賞を受賞した。

^{*3} <https://gisaid.org/>

4.2.2 プロジェクトの推進方法及びコミュニケーション手段について

同プロジェクトが開始された時期における社会状況は、COVID-19の第一波を経てソーシャルディスタンスの意識が広がり、リモートワークの習慣が定着していた。そのため、同プロジェクトにおいても、進行管理及びメンバー間のコミュニケーションは主にリモートツールによってなされた。不定期の会議ではMicrosoft社製のTeamsを使用した。会議ではアジェンダが配布され、その内容について討議した後、メンバーのタスクと大まかな期限を協議の上で決定し、それらを簡潔にまとめた議事録が配布された。そして、会議での検討に沿って筆者がビジュアライゼーションの試作を行いメンバーに提供した後、それについて検討する会議を行うというプロセスを繰り返した。筆者が行ったビジュアライゼーションの試作はProcessing^{*4}で開発した。試作はインタラクティブなアプリケーションとしてWindowsまたはMacOS上で動作するものであり、検討要因に沿ったアップデート版のアプリがメンバーに配布された。操作に変更がある場合はマニュアルも配布された。メンバーはマニュアルに沿ってアプリを確認した上で会議に臨んでいた。検討事項がアプリの特定の状況での見え方に限定されている場合は、アプリを配布せず検討を反映したスクリーンショットを配布する場合もあった。

4.3 分析手法

4.3.1 分析対象とするデータ

プロジェクトでは、一連のタスクや関係者間のコミュニケーションが電子的に記録されている。具体的には、会議に向けたアジェンダ、議事録、試作したアプリのバージョン管理ツール(Github)のコミットログ、関係者間のメールに記録されている。そのため本研究ではこれらの一連の資料と筆者の作業メモを分析対象とした。

4.3.2 分析手順

まず対象となる記録から、タスクを抽出した。各タスクは、日時、内容、内容を要約したタイトルによって記述した、また、各タスクの内容を判断しCIDにおける7つのカテゴリーでラベリングした。7つのカテゴリーに分類できないが頻出するタスクについては、

^{*4} <https://processing.org>

次項に記すラベルを付与した。これらタスクを集めたデータを「タスク一覧」と称する。次に、時系列とラベルに注目してタスクを時間軸上に配置した。これを「タスク遷移」と称する。またタスク同士の結びつきの強さを調べるため、連続するタスクのネットワークを有効グラフで可視化した。これを「タスクネットワーク」と称する。タスクネットワークでは、エッジ（矢印）の太さがタスク間の遷移回数、ノードの大きさが遷移先の多様度を表す。

4.3.3 新たに定義したカテゴリーについて

CID の定義に従い各タスクをカテゴリーに分類した結果、どのカテゴリーにも当てはまらないタスクがあった。それらのタスクの共通性を検討して新たなカテゴリー「レビュー」を設けた。また、既存のカテゴリである「精緻」「インタラクション」「取得」については、元の定義を踏まえながらタスクの性質を考慮して解釈を更新した。以下ではその詳細を述べる。

新設したカテゴリー「レビュー」

7つのカテゴリーに分類できないが頻出するタスクとして「レビュー」を設けた。「レビュー」はデザイナーが試作した可視化をメンバーに共有し、それをメンバーらで吟味するタスクである。このタスクでは可視化から読み取れる意味や科学的な確かさ、視聴者への価値を見極める。また、レビューによって課題を発見し、次の試作の指針とすることもある。

「精緻」と「インタラクション」の探索用途と提示用途の区別

「精緻」と「インタラクション」についてはプロジェクトメンバーによる探索の利便のために行う場合と、一般視聴者に公開する際の配慮のために行う場合で性質が異なる。そのため、前者を目的に行う場合の精緻は「探索用精緻」、後者は「提示用精緻」に、インタラクションも同様に「探索用インタラクション」と「提示用インタラクション」に分けた。

「取得」の解釈

「取得」について CID では、プロジェクト開始直後に十分内容や構造が把握できていない未知の大規模なデータをデータソースから取得してくることを指す。しかし、いったん内容や構造が把握できた後で、データが更新されたために「再取得」することを CID では明確に想定していない。本プロジェクトでは再取得は頻繁に行われることから、その場

合については「取得」とせず、未知のデータを取得した場合にのみ「取得」とした。

4.4 プロジェクトの観察

4.4.1 プロジェクトの全体像および前提

全体像と観察時期

プロジェクトは2020年7月に開始し、2021年6月末までに5点のデータビジュアライゼーションを順次制作し一般に公開した。公開日と公開形態は表4.1に示す。第4回と第5回の間は若干の休止期間があったものの、それ以外の期間は連続しており、チームメンバーの変更は無い。本プロジェクト開始時には第1回と第2回の公開予定だけが確定しており、第3回以降はそれまでの成果応じて漸次的に公開機会が決定したものである。プロジェクトの進行中、可視化されるデータソースは常に最新版に入れ替えられ、その折々の感染状況や新たな変異株についての研究動向および視聴者のニーズを考慮して表現を変更している。

表 4.1 公開日と公開形態

	公開日	公開形態
第1回	2020年9月26日	報道番組
第2回	2020年12月5日	セミナー型展示
第3回	2021年1月13日	報道番組
第4回	2021年3月15日	インタラクティブ展示
第5回	2021年6月23日	報道番組

前提

プロジェクトの初回ミーティングで、次のように初期の方向性が確認された。ゲノム解析の国際的なプロジェクトである nextstrain.org^{*5}では GISAID に集められたウイルス遺伝子変異のデータを分析し、変異がいつどのように広がったかを系統樹で可視化していた(図4.2)。これを知ったIは、日本にいつどのように入ってきたかを可視化できるのではないかと動機を提起した。専門家(N)は、野生株(武漢株)が変異し感染力が高まったと推定される変異「D614G」がヨーロッパで拡散したとする論文[66]を紹介した。デザイナー(筆者)は、GISAIDのデータを取得して変数を精査し、どのような可視化が可

*5 <https://nextstrain.org/ncov/gisaid/>

能かを調査することとした。

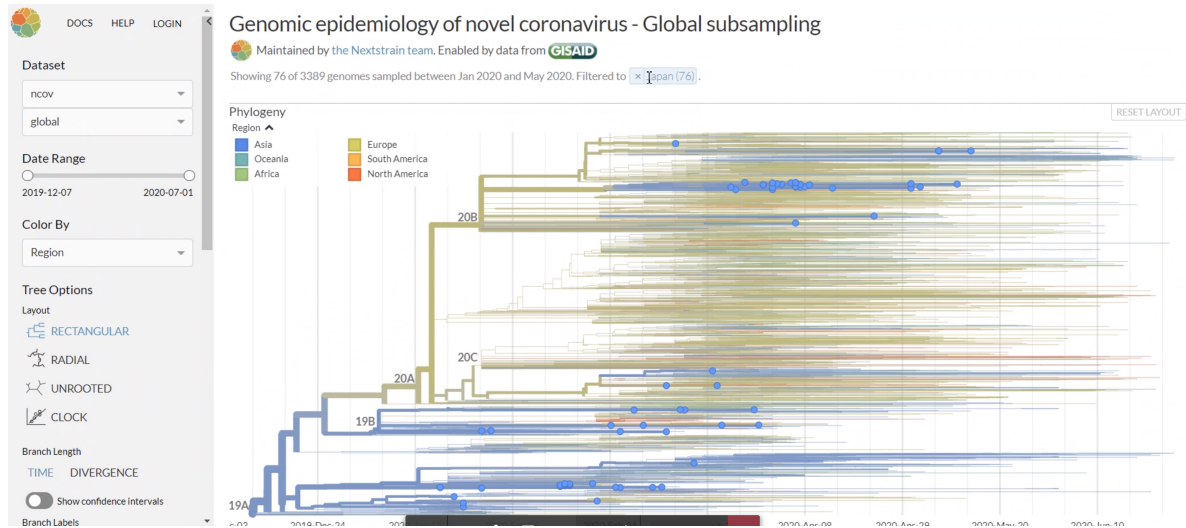


図 4.2 nextstrain.org でのゲノム系統樹可視化状況 (2020 年 7 月)

以下では、第 1 回公開までを「フェーズ 1」と呼称し、第 5 回公開までそれぞれ同様に「フェーズ (公開回)」と呼称する。タスクの実行者は、I、K、N、O などの略号で示す。特に略号がない場合はデザイナー (筆者) である。レビューのタスクについては、タスクだけでなくレビューで出された意見についても記載する。なおこの記載ルールはフェーズ 1 からフェーズ 5 までの記録全てにおいて同様に適用する。

4.4.2 フェーズ 1：第 1 回公開までの記録

時期

2020 年 7 月 20 日から 9 月 8 日

開始時の概況

フェーズ 1 開始時の概況は E 社メンバーおよび O は以前からの仕事仲間であったが、筆者及び専門家 N はこのプロジェクトにおいて初めて面識し、日が浅い状況であった。初回の打ち合わせ (2020 年 7 月 10 日) に、I からプロジェクトの方針が示され、続いて K からこれまで調査した新型コロナ変異株の拡散について資料とともにメンバーに説明された。変異の系統樹を可視化する先行事例として nextstrain.org を閲覧した。その際、専門的すぎてわかりにくい、地域的な広がり認識しにくいといった課題が挙げられた。第 2 回目の打ち合わせ (2020 年 7 月 13 日) において、筆者は変異株の拡散を時間の推移と空

間的な広がりの方を同時に可視化する手法として3次元可視化をスケッチ(図4.3)と共に提案した。筆者の調査により使用するデータソースはGISAIDから入手できる見込みがあったが、データの変数の数や種類など詳細は不明であった。

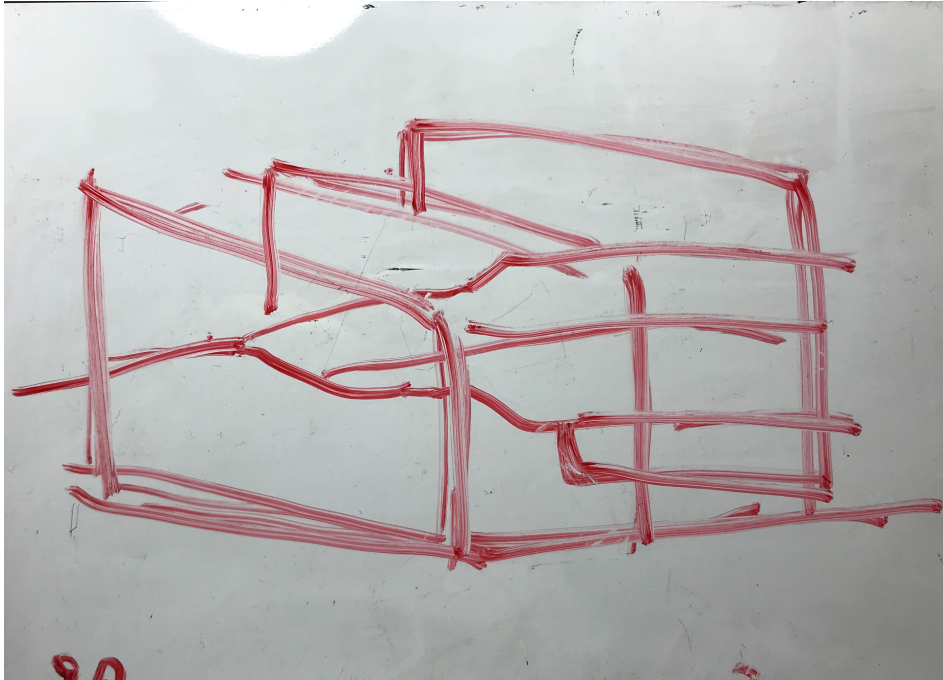


図 4.3 筆者による初期スケッチ

プロセスの概略

フェーズ1のタスク遷移図は図4.4に示す。タスク一覧を表4.2に示す。遷移図内の各タスクのタイトル先頭に付した()はタスク一覧表におけるタスクIDを表す。

フェーズ1は、他のフェーズと異なり、データの取得と解析のタスクに時間を多く割いている。変異の仕組みや系統樹の記述法といった基本知識を論文等で得つつ、スケッチで提案した3次元の系統樹をプロトタイプしていった。試作のレビューは合計4回行っている。

第1レビュー(1-7)では初期のアイデアの通り、時間軸と緯度経度にマッピングされた空間に系統樹を可視化した。日本に流入した系統だけを時空間で可視化することができたため、初期アイデアの有効性に一同は肯定的な反応を示した一方、「検出場所は正確にわかるが、系統が分岐した位置の特定は困難なので表現を変えてみてはどうか」、「D614G変異をみたい」などの課題が提起された。Kが調査していたD614G株の拡散を可視化す

る方針に沿って、変異株データと配下の子孫数の多寡を可視化する方針を立てた。

第2レビュー(1-12)では方針に沿って野生株とD614Gを色で可視化し、子孫数はエッジの太さで可視化することでD614Gの拡散を理解できることを示した。次に、ある国で検出した変異がどのようなルーツを持つのかという問いに答える可視化を試作する方針となり、変異の採取国データで系統をフィルタリングできる機能を実装した。

第3レビューでは、日本国内で検出した変異のルーツだけを可視化した状態を示した。また、課題として、系統の分岐が不自然に直角的であること、分岐位置を採取位置から算出するために筆者が作成したアルゴリズムと実際に分岐したと推測できる位置に乖離が起こる場合があることなどから、それらの改善が挙げられた。分岐を曲線状に表現することに変更した。また分岐位置の3次元的な推定アルゴリズムは存在しなかったことから、筆者とNが協議し、科学的解釈を逸脱しない範囲で自然な位置に配置できるアルゴリズムを開発した。

第4レビュー(1-21)では、拡散状況の実態に近い表現によって、国内に流入した変異株のルーツを表示できた。

表 4.2: フェーズ 1 のタスク一覧

ID	カテゴリー	時期	タイトル	タスク内容
1-1	取得	7/11	可視化に適したデータを探す	nextstrain.org のガイドンスに従い.nwk 形式の系統樹データをダウンロードし解析を試みる。.nwk データに関する論文を調査したが、専門性が高く解析が非常に困難であると気づく。.nwk から json 形式に変換するためのいくつかの方法を nextstrain の開発者向けページ内で発見し調査。この方法が複雑なため、一旦保留し、json 形式の提供がないかを調査。GISAIID に json 形式での入手が可能であることを突き止めた。
1-2	解析	7/12-15	データ構造と変数の意味の理解	json 内の変数をパースし、変異の採取位置情報（採取国代表値）と、時間情報が確認できた。このことから、緯度経度、時間軸の 3 次元的な表現が可能であると気づく。7.13 のミーティングで 3 次元可視化のスケッチを提案した。Nextstrain のプロジェクトのプロジェクトドキュメントからデータスキーマを確認し、可視化のための変数を吟味した。系統樹がどのように記述されているか確認する。系統樹は木構造で表現されており末端のノードに変異ウイルスの情報がある (strain)、末端以外のノードは分岐 (branch) として表現されている。branch は strain のような明確な採取国情報を持たないことを確認した。
1-3	フィルタリング	7/15	初期アイデア用の変数を取り出す	json を processing でパースし、すべてのノードの時間情報と位置情報を格納した。初期のアイデアをプロトタイプするために、json から strain の位置と時間、および branch の時間のみをフィルタした。
1-4	マイニング	7/15	データを可視化に適した形式へ変換	時間データを X 座標に割り当てるため正規化。緯度経度を Y-Z 座標に割り当てるため変換。特定の国に入った系統を分類
1-5	表現	7/16	全体像の把握とアルゴリズムの検討	世界地図を読み込み、Strain のだけをプロットしてみる。表示された位置（時間軸と緯度経度）が正しい確認する。分岐から strain を結ぶエッジを描くためのコードをメインのプログラムとは別に開発してテスト。同じ国でほぼ同時に採取された strain は重なってしまうので少し位置をずらす。特定の国で見えられた strain につながる系統をハイライトさせる
1-6	探索用インタラクション	7/20	全体と細部を観察する	マウスドラッグで視点変更とズームを実装
1-7	レビュー	7/20	初期アイデアの試作を通じた方向性の検討	ブランチの座標はストレインほど正確にはわからないので、表現を変える必要があるかもしれない（波線にするとか）／過ぎ去った時間のものの色を薄くするとかすると見やすくなるか／国以下の位置情報を入れられるか（都道府県のレベル）／ローカルバージョンも作れるか（日本だけのバージョン）／D614G の他の変化／日本に入ってきた複数のストレイン同士の共通の祖先との距離 (divergence) が見られると面白い（色で分ける）／一つのストレインから先祖を迎えると面白い
1-8	フィルタリング	7/21	変異株の種類を同定する変数追加	D614G(欧州株) 変異を持つか持たないかのデータを追加
1-9	マイニング	7/21	系統ごとに検出した子孫数計算	D614G 変異データで系統でカテゴリ化 (S プロテインの 614 変異が G であるものにフラグ)。各エッジの子孫の数を計算。
1-10	表現	7/22 - 24	変異株子孫数の寡多を表現	D614G 変異によって色分けして表現／多くの子孫が検出されたエッジは太くなるように表現
1-11	探索用インタラクション	7/24-25	時間選択スライダー	時間を操作するスライダーによって地図が時間軸上を平行移動し、採取位置と時間を把握しやすくした
1-12	レビュー	7/26	D614G 変異の拡散を表現	ヨーロッパに入った経路を複数みたい／最終地点の国とそこに共通する祖先に関する関心 (Italy/CDG1/2020) ／D と G それぞれの地域への広がり色分けしたい
1-13	フィルタリング	7/27-30	変異の詳細データを追加	ストレイン (変異体を表すノード) をクリックしたときに表示する／ストレイン ID の文字列をデータに含める
1-14	マイニング	7/27-30	変異のルーツ構造解析	ストレイン (変異体を表すノード) に自分のルーツ (親のノードのリスト) を格納

(次ページへ続く)

表 4.2 – (前ページから続き)

ID	カテゴリー	時期	タイトル
1-15	表現	8/1-10	国と株の選択的表示
A. 特定の国で見つかったストレインだけを可視化 (対象国外のノードとエッジは半透明に) / B. 特定の変異株の系統 (D614G) とそれ以外を色分け / A と B を同時に適用すると特定の国にフォーカスして変異系統を可視化できる			
1-16	探索用インタラクション	8/1-10	クリックによるルート表示
ストレインをクリックするとツールチップで ID を表示する / クリックされたストレインからオリジナルまでの遡るルートを表示する / フィルタ対象外の系統の透明度を調整するスライダーを実装			
1-17	レビュー	8/11	注目したい変異株の拡散ルートと時期の可視化
クリック操作で特定の株へのルートが確認できるようになった。注目している変異 (D614G) の経路の様子も探索できる。 / ID を見ることで Nextstrain と比較することができる。 / ストレインの重なりを抑えるため実際の位置からずらしているが、ズレが大きすぎる問題がある。(日本採取のストレインが海上にあるなど) / ノードが分岐する位置の推定をもう少し自然にできないか。分岐した後時間的に近い strain の場所を重視すると自然になるのでは。			
1-18	マイニング	8/12-19	自然な分岐表現の幾何学計算
系統の分岐を自然に表現するため、分岐位置推定方法を「子孫 A と子孫 B の中点」から、「A、B のうち先に採取された子孫の位置」に改める (v0.4)			
1-19	表現	8/12-19	系統分岐表現の改良
エッジを直角に曲がる表現から曲線の表現に変更。(v0.41)			
1-20	探索用精緻	8/12-19	分岐位置推定計算の精緻
実際に観察された感染ルートとてらし、より自然に表現するため、分岐する位置推定方法をさらに変更。(v0.42)			
1-21	レビュー	8/20	分岐表現の改良により拡散状況の実態に近い表現と評価
エッジ自体もより見やすく整理され、初期の移動がより事実に沿って表現されるようになっている。 / 推定に用いた仮定は全く問題ない。今回の系統樹と位置関係のほうが確からしいと思う。(専門家)			
1-22	提示用精緻	9/1	複雑さ緩和のための精緻
同一国で同じ日に採取されたストレインが重なって表示されるのを防ぐため相互にずらして位置決めする方法を見直し採取国から大きくずれないように改善			
1-23	提示用インタラクション	9/1-8	テロップ付記・画角調整用 UI
エンジニア O は、映像制作用の CG カメラワークや、テロップ機能を実装した			

公開までのプロセス

第 4 レビュー (1-21) の終了後、I と K らはビジュアルライズのインパクト、感染拡大実態の説明性、この可視化から発見できた特徴的な事象の 3 点を念頭に番組の台本を作成した。筆者は台本に基づき使用する表現を把握したのち、実際の番組用 CG を制作する O へ、ソフトウェアから出力した 3 次元系統樹の形状データを提供し再現の助言をおこなった。O は K との調整を経て番組 CG 制作用ソフトウェアを完成させた (1-22,1-23)。完成した CG をもとに I らは N への取材および収録を行った。N は可視化された CG を示しながら番組内で解説を行った。

公開されたデータビジュアライゼーション

放映された番組の概要を示す。変異データを分析する国際的な研究が進められていることを説明したのち、ビジュアライズ映像を使用して以下の3点が解説された。①武漢で発祥した野生株から派生した変異株（D614G株）がヨーロッパ各地へ拡散したこと（図4.5）。②日本で拡散したウイルスの多くはヨーロッパ経由のD614G株と見られること（図4.6）。③国境を跨ぐ密集した系統樹が疎に変わり移動制限の効果が視覚的に確認できること（図4.7）。①は第2レビューの成果(1-12)、②は第3レビューの成果(1-17)、③は第4レビューの成果(1-21)に対応している。

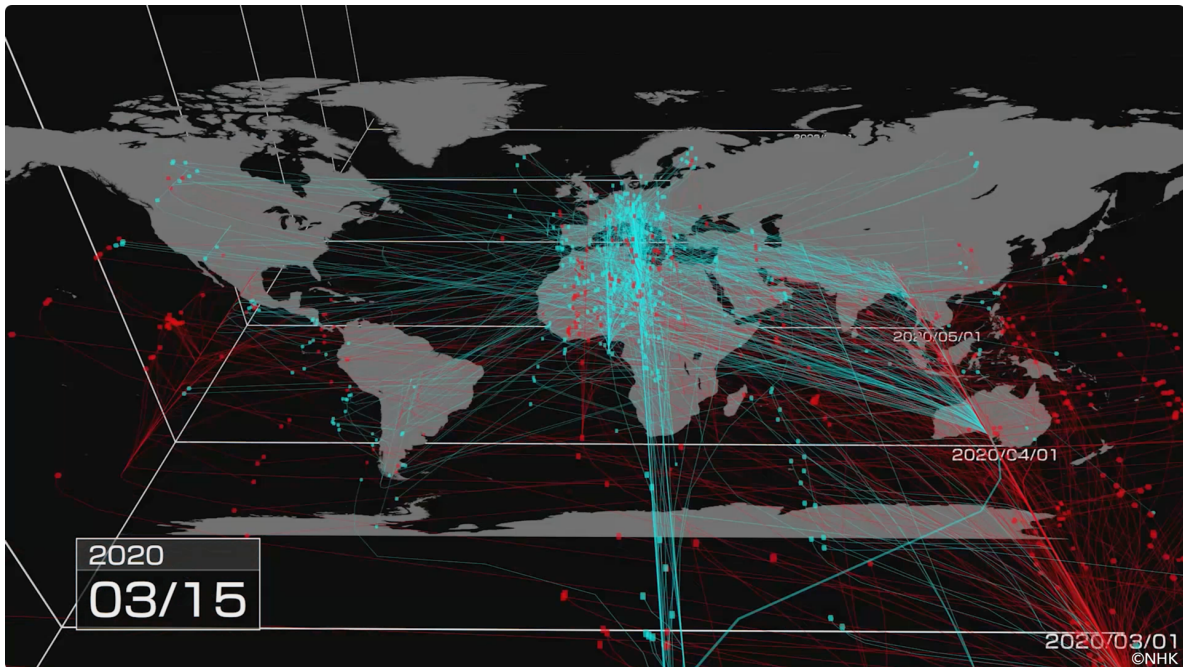


図 4.5 フェーズ 1 の放送時の映像①



図 4.6 フェーズ 1 の放送時の映像②



図 4.7 フェーズ 1 の放送時の映像③

タスクのネットワーク

フェーズ1のタスクネットワークを図4.8に示す。タスクネットワークでは、エッジ(矢印)の太さが遷移回数、ノードの大きさがタスクからの遷移先の多様度(out degree)を表す。プロジェクトの序盤(取得からフィルタリング)および最終段階(レビューから展示用インタラクション)ではタスクの反復は見られなかった。フィルタリングからレビューまでのタスクには4回の循環が見られる。レビューは遷移先の多様度が最も多かった。

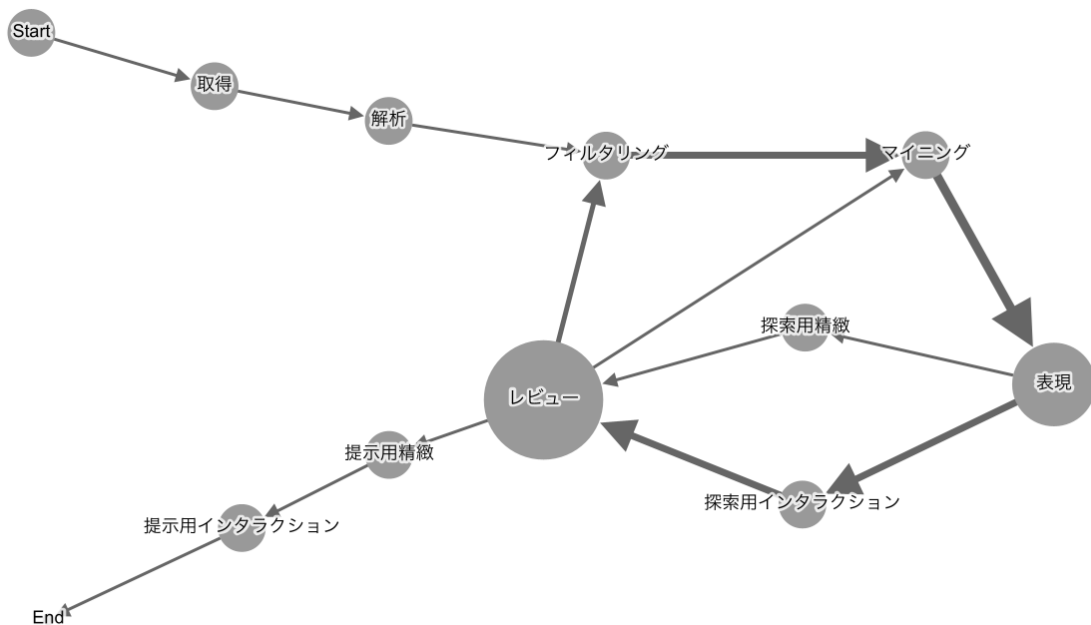


図4.8 フェーズ1のタスクネットワーク図

4.4.3 フェーズ2：第2回公開までの記録

時期

2020年10月16日から11月24日

開始時の概況

E社が主催し、子どもから大人まで幅広い年齢層を対象に科学館で行われるイベントが予定されており、そのうちひとつのブースで本可視化を使ったミニセミナーが企画されて

いた。セミナーでは、当時最新の新型コロナウイルス研究でわかってきたことについて、本可視化のアニメーションを交えながら説明員が解説を行う計画であった。

プロセスの概略

フェーズ2のタスク遷移図は図4.9に示す。タスク一覧を表4.3に示す。フェーズ2では、レビューを3回実施し5種の試作を行った。

第5レビュー(2-5)ではウイルスの変異数によって着色する機能を実装した。変異数による着色のアイデアは、フェーズ1の第1レビュー(1-7)の時にNが提案したものである。変異数(Divergence)は採取されたウイルスの塩基が、系統樹の最古に位置する野生株の塩基と比較してどれだけ変異しているかを示す。変異は一定の確率で生じるため変異数はおおよそ変異した回数と一致し、時間軸で可視化すると変異の速度を示すことができる。Iは結果について遺伝の進化速度が見えることを確認した上で「遺伝の進化が分かりやすく見えるために色の差を強調したい」とコメントした。

第6レビュー(2-15)では3種の試作を提示した。前述のIの指摘の修正版(2-12)。また、Kが調査していた特定地域で優勢になっている変異S477N(フランス)とA222V(スペイン)を可視化した(2-10)。さらに、筆者の提案で6大地域(北米、南米、アジア、オセアニア、ヨーロッパ、アフリカ)ごとに系統を着色する機能を追加した(2-11)。この機能とフェーズ1で開発した検出国までのルートをフィルタする機能を同時に使うと検出国に至るまでどの地域を経由してきたかが明らかになる。K提案の可視化結果(2-10)についてNは「変異が独立に平行に起きている状況が見えてとてもよい」とコメントした。筆者提案の可視化した結果(2-11)についてIは、「各地域(日本含め)どこから流入しているかわかり、価値のある機能になりそうだと感じた」とコメントした。またこの時Nは、追加でN439K変異についても着色して可視化することを提案した。理由は、N439Kは免疫逃避性が確認されたことから感染が拡大することを予測したためである。筆者は、N439K変異のデータを着色できる機能を実装した。

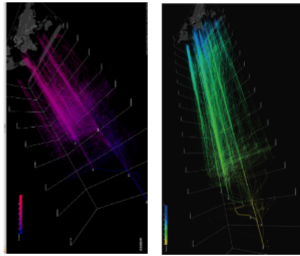
第7レビュー(2-19)ではN439Kを可視化し免疫逃避性のある変異が流行する兆しを表現した。結果についてIは「数は多くないが、この変異による感染が継続していることがわかる」とコメントした。

カテゴリ	10/22 - 28	10/22 - 28	10/22 - 28	10/29	11/06 - 09	11/06 - 09	11/06 - 09	11/10	11/11	11/12	11/12	11/12-13	11/14	11/15	11/15	11/24
取得																
解析																
フィルタリング																
マイニング																
表現																
探索用精緻																
探索用インタ																
レビュー																
表示用精緻																
表示用インタ																

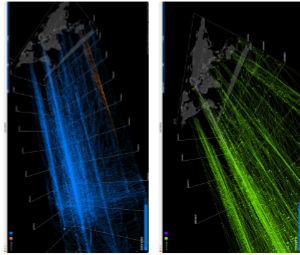
(2-1)変異数のデータを追加 (2-6)局地的な流行株、A222V (スペイン)、S477N (フランス) のデータをフィルタ (2-7)世界6大地域 (region) データをフィルタ (2-8)変異を持つ子孫数計算 (2-9)regionごとに検出した子孫数計算 (2-10)局地 (2-11)6大地域地域別に系統を表現 (2-12)変異数の色の表現を修正 (2-17)免疫逃避が確認された変異を表現 (2-16)免疫逃避が確認された変異 (N439K) のデータを追加

(2-2)変異数の分布解析 (2-3)変異数を色で表現 (2-4)変異数表示のフルダウンメニュー (2-13)変異の表示切替えにメニューを追加 (2-14)地域の表示切替えプルダウンを追加 (2-15)局地的な流行株を表現、6大地域別の色分けを表現 (2-19)免疫逃避が確認された変異の分布をレビュー (2-20)説明用アニメーションの実装

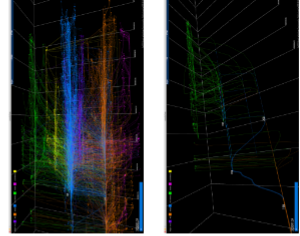
第5レビュー (2-5) 遺伝子の進化速度を可視化 (上：精細画、下：精細後)



第5レビュー (2-5) 遺伝子の進化速度を可視化 (上：精細画、下：精細後)



第6レビュー (2-15) 局地的な流行株を表現 (上：フランスS477N、下：スペインA222V)



第6レビュー (2-15) 6台地域別を表現 (上：全部表示、下：特定の経路の地域を表現)



第7レビュー (2-19) 免疫逃避のある変異を表現

図 4.9 フェーズ2のタスク遷移図

表 4.3: フェーズ 2 のタスク一覧

ID	カテゴリー	時期	タイトル
2-1	フィルタリング	10/22 - 28	変異数のデータを追加 変異の Divergence を追加
2-2	マイニング	10/22 - 28	変異数の分布解析 Divergence の分布を計算、Divergence の値で 7 段階クラスに分け
2-3	表現	10/22 - 28	変異数を色で表現 Divergence のクラスに色を対応させる
2-4	探索用インタラクション	10/29	変異数表示のプルダウンメニュー Divergence をプルダウンに追加
2-5	レビュー	10/29	遺伝の進化速度を表現 (I コメント)「キャプチャをみた限りでは、色のバリエーションがもう少し多い方がいいと感じた。8K では色数が多く、綺麗に出るのでそれを生かす意味と、遺伝的の進化がわかりやすく見えた方がいいと感じた。来週、全員で議論したい。」
2-6	フィルタリング	11/06 -09	局地的な流行株、A222V (スペイン)、S477N (フランス) のデータをフィルタ 現時点で特徴的な変異、A222V (スペインが中心)、S477N (フランスが中心) の地域的な広がりを見る / (T コメント)「欧州を中心にこの夏に広がった S タンパク質の変異として A222V (スペインが中心)、S477N (フランスが中心) の 2 つの報告がある。S は 614G が世界的に major になってしまったので、現在地域特異的に広がっている変異として示す例としてはよいかもしれない。」
2-7	フィルタリング	11/06 -09	世界 6 大地域 (region) データをフィルタ 採取地の Region を追加
2-8	マイニング	11/06 -09	変異を持つ子孫数計算 A222V 変異、S477N 変異データで系統でカテゴリライズ
2-9	マイニング	11/06 -09	region ごとに検出した子孫数計算 Region でノードとエッジをカテゴリライズ
2-10	表現	11/10	局地的な流行株を表現 A222V 変異、S477N 変異を持つ場合にエッジとノードの色を変更
2-11	表現	11/11	6 大地域地域別に系統を表現 Region を選択すると、特定の Region ごとに着色される
2-12	探索用精緻	11/12	変異数の色の表現を修正 Divergence の表現を修正
2-13	探索用インタラクション	11/12	変異の表示切替えにメニューを追加 プルダウンに A222V 変異、S477N 変異を追加
2-14	探索用インタラクション	11/12	地域の表示切替えプルダウンを追加 Region モードをプルダウンに追加
2-15	レビュー	11/12-13	局地的な流行株を表現、6 大地域別の色分けを表現 (T コメント)「A222V と S477 の変異の追加について、変異が独立にパラレルに起きている状況などが見えてとてもよい。こちら展示発表のシナリオ次第かもしれないが、同じく S タンパク質の N439K についても追加することは可能か。先日追加したものと異なり、こちらは免疫から逃れるという機能について明らかになり、また欧州で独立に 2 回生じていて、それぞれ 500 サンプル以上シーケンズされているので、Nextstrain でもそれなりに見られるのではないと思う。」(I コメント)「Region の色分けは 8K で表示することを考えると美しく見えると感じた。また、各地域 (日本含め) にどこから流入しているかがわかり、価値のある機能になりそうと感じた。」
2-16	フィルタリング	11/14	免疫逃避が確認された変異 (N439K) のデータを追加

(次ページへ続く)

表 4.3 – (前ページから続き)

ID	カテゴリー	時期	タイトル
タスク内容			
N439K を追加。免疫から逃れるという機能について明らかになり、また欧州で独立に 2 回生じていて、興味深いため。			
2-17	表現	11/15	免疫逃避が確認された変異を表現
N439K を追加			
2-18	探索用インタラクション	11/15	変異の表示切替えにメニューを追加
N439K プルダウンに追加			
2-19	レビュー	11/15	免疫逃避が確認された変異の分布をレビュー
(I コメント)「S439 の着色について、Nextstrain では数は多くないようだが、感染が継続していることがわかる。他の変異より優位なのかななどと想像した。A222 と S477 は、現状では台湾では触れる時間はなさそうなので、同じ説明内容を S439 でできると考えている。」			
2-20	提示用インタラクション	11/24	説明用アニメーションの実装
エンジニア O は、イベント解説員の説明用に CG カメラワークや表示切り替え機能を実装した			

公開までのプロセス

第 7 レビュー (2-19) の終了後、I はセミナーの所要時間を加味した上で、①進化の速度、② D614G 株の流行、③免疫逃避性のある新しい変異の出現の 3 点を念頭にアニメーションのシナリオ作成した。筆者はシナリオに基づき使用する表現を把握したのち、アニメーションを制作する O ヘデータを提供し再現の助言をおこなった。O はシナリオに沿って CG カメラワークのアニメーションを実装した (2-20)。イベント当日は O が作成したソフトを操作しながら解説員が解説を行った。

公開されたビジュアライゼーション

科学館で行われたミニセミナーの概要を示す。変異データを分析する国際的な研究が進められていることを説明したのち、ビジュアライズアニメーションを使用して以下の 3 点が解説された。①ウイルスが変異し現在までに 30 回以上変異したこと (図 4.10)。②日本で拡散したウイルスの多くはヨーロッパ経由の D614G 株と見られること (図 4.11)。③新たに免疫を逃避する可能性がある新しい変異が出てきていること (「ワクチンの効果に影響が出るかも」と表現) (図 4.12)。①は第 5 レビューの成果 (2-5)、②第 2 レビューの成果 (1-12)、③は第 7 レビューの成果 (2-19) に対応している。

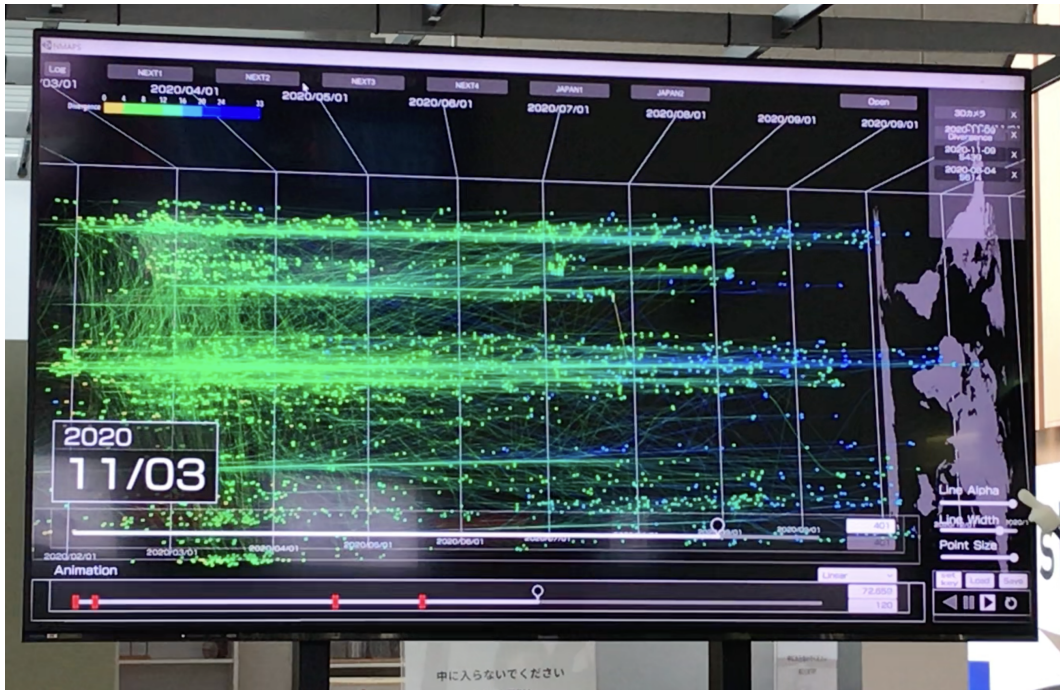


図 4.10 フェーズ 2 のセミナーの様子①



図 4.11 フェーズ 2 のセミナーの様子②



図 4.12 フェーズ 2 のセミナーの様子③

タスクのネットワーク

フェーズ 2 のタスクネットワークを図 4.13 に示す。本フェーズでは、データは最新に更新されたものの構造には変化がないため、取得、解析のタスクはない。また、特定の変異を着色する機能はフェーズ 1 で実装したため、フィルタリング後マイニングを行わずに表現する遷移がある。前回同様、フィルタリングからレビューまでの循環が 3 回確認できる。

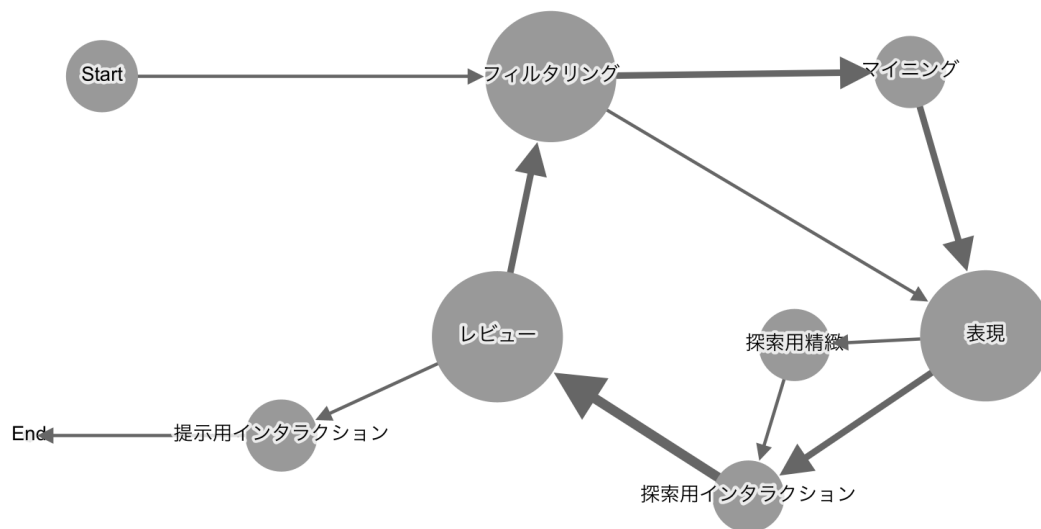


図 4.13 フェーズ 2 のタスクネットワーク図

4.4.4 フェーズ 3：第 3 回公開までの記録

時期

2020 年 12 月 22 日から 2021 年 1 月 4 日

開始時の概況

翌月に全国放送される報道番組にて、ビッグデータを分析して COVID-19 のワクチン開発や感染拡大防止を図ろうとする研究を特集することとなった。それに伴い、これまでに本プロジェクトで制作したビジュアライゼーションを活かして番組用映像を制作することとなった。野生株からこれまでの変異を振り返るとともに、最新の状況を伝えるためにイギリスと南アメリカで拡大しつつあった N501Y 変異に注目することとなった。N501Y 変異とはいわゆる「イギリス株」現在では「アルファ株」と命名されている系統である。当時すでにイギリス国内では感染者が急増していたが、我が国においてはこれから流行が懸念される状況であった。

プロセスの概略

フェーズ 3 のタスク遷移図は図 4.14 に示す。タスク一覧を表 4.4 に示す。N501Y をリサーチしていた K が調査した結果からこれまでの可視化に使用していた GISAID デー

タではなく、neherlab のデータで可視化することとなったため取得 (3-1) を行った。GISAID データとは異なり、N501Y 変異ウイルスに絞ったデータであるため、感染拡大地域に表示が集中してしまう懸念があったが、筆者はこのデータから試作を行った。

第 8 レビュー (3-4) ではフェーズ 2 の第 6 レビューで実装した地域別の可視化機能で N501Y の分布を着色した。結果は、流行地域 (イギリスやヨーロッパ) に短期間で相当数の変異が集中したため国境外へプロットがはみだす「染み出し」が起き、現在のアルゴリズムでは対処できないことがわかった。I は、地域別でなく Clade 別に着色することを提案した。Clade とは国際的に取り決められている変異の系統を示す名称である。つまり、表示が集中した地域でも Clade に差をつけて可視化することで切り分けて見られるのではないかという目論見である。筆者はこれに応じ、Clade を識別するデータを追加し表示の切り替え機能の試作を行った。

第 9 レビュー (3-8) では Clade 別に着色した neherlab データの可視化をレビューした。I は「20B (筆者註: Clade の名称) から N501Y が急速に入れ替わっていることがわかりますね」とコメントしている。これは、イギリスやヨーロッパでの 20B から派生した N501Y/V1 変異が急速に優位になっていることを切り分けて表示できたことを評価している。しかし、I は「現状ではこれを取り上げることは止めることにしました。(ヨーロッパ上でたくさんの UK データが円上に配置されることが説明しにくいと判断しました。)」とコメントしている。これは、特定の箇所にプロットが集中するデータセットは今回の可視化システムでは、内容と表現がうまく結びつかないと判断したと考えられる。I は対応策として、従来通り GISAID のデータに変更して N501Y を可視化する方針を提案した。筆者はこれを了承し、データを差し替えた上で、N501Y と Clade の組み合わせで着色する機能を付加した。すなわち、Clade=20B の N501Y と Clade=20C の N501Y を色分けすることができ、Clade=20B の N501Y はいわゆるイギリス株であるため、この部分に注目すれば拡大しているイギリスの状況を可視化することにつながる。

第 10 レビュー (3-14) では、Clade=20B に属する N501Y (ピンク) と Clade=20C に属する N501Y (水色) を着色して提示した。I は「20B 501Y のデータが更新されて、世界各地に拡散されていることがわかりますね。(GISAID の日本のデータは採用されていませんが、韓国のデータがあるためアジアへのエッジもひかれている)」とコメントしている。これは、イギリス株のアジアへの流入の端緒を捉えていることを評価している。ここで番組のプロデューサー (M) から、武漢株から欧州株 (D614G) の置き替わりを経て、さらにイギリス株 (N501Y) へと置き換わりつつあることを連続で可視化できないかという提案があった。筆者はフェーズ 1 第 2 レビュー (1-12) の機能を改修することで可能と判断した。

第 11 レビュー (3-20) では、メジャー株 (置き換わり株) の変遷を色で可視化した。M は放映時の色の伝わりやすさの観点を指摘したがそれ以外の点ではこれを評価し、これを精緻したバージョン (3-21) を正式に採用することとした。

	12/25	12/25	12/25	12/26	12/26	12/27	1/2	1/2	1/2	1/3	1/3	1/4	1/4	1/4	1/4
カテゴリ取得	(3-1)inheritabのデータ取得	(3-9)ISADのデータ取得	(3-10)N01Yのデータをフィルタ	(3-11)N01YとCladeの掛け合わせで分類	(3-12)N01YとCladeの掛け合わせで表示	(3-13)N01YとCladeの掛け合わせでフィルタリング	(3-14)N01YとCladeの掛け合わせで観察しビュー	(3-15)アプレビ用調整	(3-16)アプレビ用調整	(3-17)メジャー株の世代変化に注目した条件で分類	(3-18)野生株、欧州株、アルファ株(20B,20C) を色分け	(3-19)表示切り替え用フィルタリングを追加	(3-20)野生株から欧州株さらにアルファ株の派生	(3-21)色の調整	(3-22)説明用アニメーションの実装
解析															
フィルタリング															
マイニング															
表示															
探索用機能															
探索用インタフェース															
表示用機能															
探索用インタフェース															

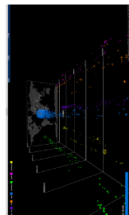
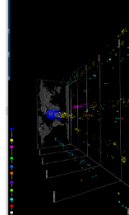


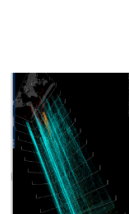
				
<p>第9レビュー (3-4) N01Yに注目したヤンブリングデータを地域別に着色</p>	<p>第9レビュー (3-8) N01Yに注目したヤンブリングデータをClade別に着色</p>	<p>第10レビュー (3-14) Clade-20Bに属するN01Y (ピンク) とClade-20C (黄緑) のN01Y (灰色) を着色。ヒングは初期のアルファ株</p>	<p>第11レビューの調整後 (3-20) 野生株 (黄緑) 欧州株 (青色) アルファ株 (イギリス株) (緑) アルファ株 (南ア) (赤) アルファ株 (南ア) (赤) 次々に派生する株子</p>	<p>第9レビュー (3-4) N01Yに注目したヤンブリングデータを地域別に着色</p>

図 4.14 フェーズ 3 のタスク遷移図

表 4.4: フェーズ 3 のタスク一覧

ID	カテゴリー	時期	タイトル
3-1	取得	12/25	neherlab のデータ取得
<p>N501Y の拡散を見るため、neherlab の系統樹の json データを取得\、(I コメント)「先日ご相談していた N501Y の変異について、Nextstrain の Author が新しい分析を出しているようで、以下の N501Y (VUI 202012/01 という表記する変異) の Nextstrain 画面のデータを可視化してみたい。このデータが 3 次元可視化に合うかどうか、N501Y は地域間の移動がまだあまりないので、イギリスのところに大量の四角が載るといった絵になるのではないかと想像する。」</p>			
3-2	マイニング	12/25	neherlab の region 分布を計算
<p>ストレインが 4312 個。そのうち、ヨーロッパが 2462 (57.1%)、UK が 2219 (51.5%、ヨーロッパの 90%)</p>			
3-3	表現	12/25	neherlab の region 分布を表現
<p>neherlab の region 分布を表現</p>			
3-4	レビュー	12/25	neherlab の region 分布レビュー
<p>(筆者コメント)「画像は全て region で色分けしてあり、なおかつ、07.png だけは UK でフィルタをかけている。07 を見ると分かる通り、UK のストレインの数が突出して多い為これまでの方法でプロットすると、ヨーロッパ全域にしみ出してしまう。06 を見ると分かる通り、10 月中旬から現在までの非常に短い期間に集中して検出されているのでエッジでつなぐと識別がかなり困難になると予想する。」(I コメント):「(ストレインが UK の上に円形に配置されているように見えることについて) 4K8K の際にも最初は円形に配置される形になっていたところを各国の上になるべく置かれるようプログラムを調整したという認識である。今回またそのように見えるは数が多すぎるなどの問題があるからだろうか。(ストレインを region で色分けについて)「これを clade で分けたバージョンを見ることは可能か。Clade でわけると、もう少し話題になっている N501Y.V1 部分だけが切り分けて見えるので、言えることが出てくるように思う。」(筆者コメント)「中心から外側に向かってストレインを埋めていくため、早い時間に現れたストレインが先に円の中心付近を占め、離れた時間で大量にストレインが出たのでドーナツ型に見えるのだと思う。(これを clade で分けたバージョンについて) 趣旨は了解した。やってみる価値はあると思うのでトライするが、clade で色分けする機能はないので、新たに実装すると早くても週明けくらいになると推測する。」</p>			
3-5	フィルタリング	12/26	clade のデータをフィルタ
<p>clade のデータをフィルタ</p>			
3-6	表現	12/26	clade のデータを色で表現
<p>clade のデータを色で表現</p>			
3-7	探索用インタラクション	12/26	clade 表示切替えプルダウンを追加
<p>clade 表示切替えプルダウンを追加</p>			
3-8	レビュー	12/27	neherlab の clade 分布レビュー
<p>(I)「501Y の Clade わけを確認した。こうしてみると、20B から 501Y に急速に入れ替わっていることがわかる。このサンプルセットについて先程、M ディレクターと検討した。その結果、いったん止めようという判断になった。今後、neherlab で、日本や世界各地への広がり大幅にアップされた場合には再検討するが、現状ではこれを取り上げることは止めることにした。(ヨーロッパ上でたくさんの UK データが円上に配置されることが説明しにくいと判断したため。)なお、Nextstrain は、これまでの Global subsampling では取り上げる予定であり、1 月 2 日までのデータで 501Y だけを色分けするセット(json)を作成いただきたいと希望している」(筆者)「諸々承知した。局所的な事象を扱うのに課題があるということがわかりよかった。Global subsampling で可視化する際の確認だが、Global subsampling では、neherlab のように 501Y/V1 と 501Y/V2 を Clade として保持しておらず、代わりに、20B の clade で発生した 501Y、20C で発生した 501Y という色分けで良いか? 前者が UK、後者が南アフリカである。」注: いわゆるイギリス株(アルファ株)が 20B の clade で発生した 501Y</p>			
3-9	取得	1/2	GISAID のデータを取得
<p>GISAID のデータを取得</p>			
3-10	フィルタリング	1/2	N501Y のデータをフィルタ
<p>N501Y のデータをフィルタ</p>			

(次ページへ続く)

表 4.4 - (前ページから続き)

ID	カテゴリー	時期	タイトル
3-11	マイニング	1/2	N501Y と Clade の掛け合わせで分類
			N501Y と Clade の掛け合わせで分類
3-12	表現	1/2	N501Y と Clade の掛け合わせを色で表現
			N501Y と Clade の掛け合わせを色で表現
3-13	探索用インタラクション	1/2	N501Y と Clade の掛け合わせをプルダウンに追加
			N501Y と Clade の掛け合わせをプルダウンに追加
3-14	レビュー	1/3	N501Y と Clade の掛け合わせ表現をレビュー
			(I) :20B 501Y のデータが更新されて、世界各地に拡散されていることがわかる。GISAID の日本のデータは採用されていないが、韓国のデータがあるためアジアへのエッジもひかれている。注目したいのは「20B 501Y」なので、色を交換してみてもどうか。／ M(I 同僚)：中国武漢由来のウイルスから、D614G (欧州型) が派生し、その欧州型から 501 に変異がおきたもの N501Y が発生したことを一連の CG で言えないかと考えている。スパイクタンパク質の中でも 614 と 501 の違う場所での変異なので、N501Y への変化は、欧州株で起こっているの、よいのではないかと議論になった。CG が 2 つに分かれて色が又変わると、一般視聴者がついていけないのではないかと (筆者)「カラーリングの条件式を変えるだけで対応できそう。こういう感じでよろしいか。S614 = D → 武漢型…黄色 (薄め)、S614 = G → 欧州型…水色 (薄め)、欧州型 かつ Clade = 20B かつ S501=Y → VOC-202012/01…赤色、欧州型 かつ Clade = 20C かつ S501=Y → 501Y.V2…オレンジ色」／ (M) はこれを了承。
3-15	提示用精緻	1/3	テレビ用調整
			ピンクと水色を逆転 (アルファ株が主題であるため)
3-16	フィルタリング	1/4	D614G をフィルタリング
			D614G をフィルタリング
3-17	マイニング	1/4	メジャー株の世代変化に注目した条件で分類
			欧州型 かつ Clade = 20B かつ S501=Y → VOC-202012/01…赤色 / 欧州型 かつ Clade = 20C かつ S501=Y → 501Y.V2…オレンジ色
3-18	表現	1/4	野生株、欧州株、アルファ株 (20B,20C) を色分け
			野生株、欧州株、アルファ株 (20B,20C) を色分け
3-19	探索用インタラクション	1/4	表示切り替え用プルダウンに追加
			表示切り替え用プルダウンに追加
3-20	レビュー	1/4	野生株から欧州株さらにアルファ株の派生
			(M)：武漢型の黄色と欧州型の青色がこれまでの CG と比べて発色がとても悪い。
3-21	提示用精緻	1/4	色の微調整
			(筆者)：青の部分 (欧州型) の部分は、以前の欧州型の明るい水色に、また、黄色と指定した武漢型の部分は、Divergence4~9 の部分の蛍光の黄緑色の色にした。くすみはなくして、透明度で N501Y とそれ以外の差をつけている。
3-22	提示用インタラクション	1/4	説明用アニメーションの実装
			エンジニア O は、映像制作用の CG カメラワークや、テロップ機能を実装した

公開までのプロセス

最終形の色微調整 (3-21) の終了後、Iはこのレビューで確定した可視化のアニメーション制作を O へ依頼した。筆者はデータを提供し再現の助言した。O はシナリオに沿って CG カメラワークのアニメーションを実装した (3-22)。I は O のソフトウェアを使って

約 90 秒の CG アニメーション映像を制作した。また M は O のソフトウェアをワクチン研究の専門家への取材ロケに携行し、専門家にソフトウェアを使用してもらいながらワクチンの研究状況をインタビューする映像を制作した。

公開されたビジュアライゼーション

番組では 90 秒のアニメーション CG を使用して、野生株から欧州株へ全世界で置き換わった様子と、現在イギリスを中心に新たな変異株が主流になりつつあることを伝えた (図 4.15)。また、ワクチン研究者への取材ロケ映像では、ソフトウェアを操作しながら、今後の変異株に対するワクチン研究の現状などを伝え「ビッグデータでワクチン開発はどう進むのか」「変異を続けるウイルスにビッグデータで対抗できるか」について研究者の見解を引き出した。研究者はゲノム配列のデータがあることでウイルスそのものがなくてもワクチンを作ることができる意義を説明した。世界中のウイルス変異をビッグデータによって常に監視することで、今後ワクチンに耐性を持つ変異を持つウイルスを察知する可能性が高まると述べた。

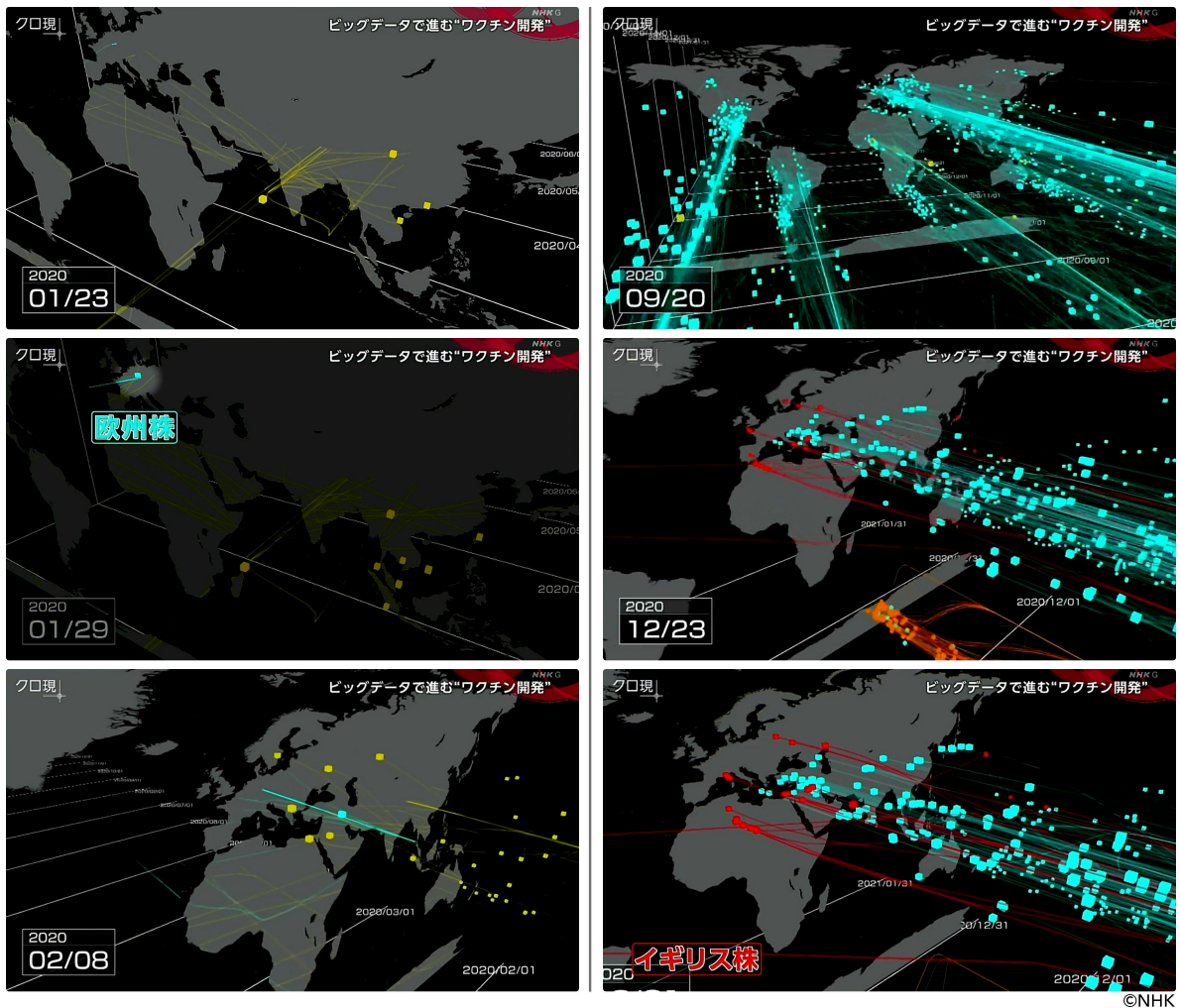


図 4.15 フェーズ 3 の放送時の映像

タスクのネットワーク

フェーズ 3 のタスクネットワークを図 4.16 に示す。本フェーズでは、まず新しいデータソース neherlab からデータを取得した。必要な変数 (N501Y) はデータ上で絞り込まれているためフィルタリングはなかったが、特定地域に極端に変異が偏っている可能性があったため確認のためマイニングを行なって可視化し最初のレビューとなる (図中記号 A-B-C-J の遷移)。その後、データは変えずに別の変数で可視化しレビューした (K-I-D-E)。しかしこのデータを採用しないこととし、GISAID からデータを取得してレビューする (F-G-H-C-D-E)。いったん提示用に精緻したが複数の流行株を可視化する方針に変わったためレビューした (M-L-H-C-D-E)。結果、フィルタリングまたはマイニン

グのどちらか1つを含む循環が4回確認できる。

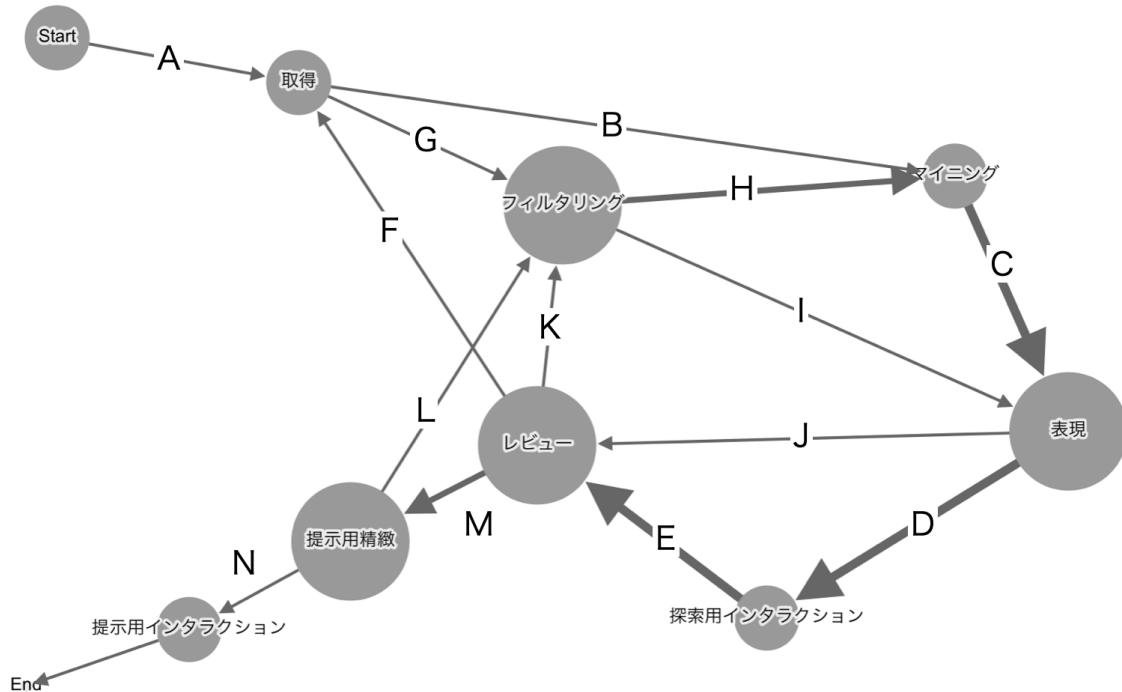


図 4.16 フェーズ 3 のタスクネットワーク図

4.4.5 フェーズ 4：第 4 回公開までの記録

時期

2021 年 2 月 22 日から 3 月 15 日

開始時の概況

フェーズ 1 の完了直後、E 社の関連施設において開催される展示会にて、一般来場者向けにインタラクティブな展示として本ビジュアライゼーションを公表する方針が示された。その展示内容案について I からプロジェクトメンバーへ提案し検討するところが本フェーズの開始状況であった。展示会は E 社の関連する放送局が主催し、8K 解像度の映像コンテンツを特集して一般に公開する趣旨であった。本プロジェクトにおいては 32 インチ型 8K ディスプレイを用いること、インタラクティブであること、会場には展示説明員が待機すること、解説用パネルも併置することが決定していた。本プロジェクトにおいて、不特定多数の一般来場者が直接操作できる形式での展示は、本フェーズが初めての機

会となった。

Iが提示した内容の要旨は次の通りである。鑑賞者がいない場合は、系統樹全体が映る画角で系統樹の周囲 360 度をゆっくり回転する視点からの映像が表示される（「スクリーンセーバーモード」と呼称、以下「SS モード」）。ディスプレイ脇に設置されたトラックボールの特定のボタンを押下すると、系統樹の幹から末端に向かって自動的に視点が移動しながら 30 秒程度で時系列に拡散した変異株を説明するテロップが表示される（「ムービーモード」）。トラックボールや、ボールの周囲にあるホイールを操作すると鑑賞者がカメラの視点や時間を操作し、任意の位置と時間で系統樹を眺めることができる（「操作モード」）。表示するビジュアライゼーションについては、イギリス（N501Y.V1）、南アメリカ（N501Y.V2）、ブラジル（N501Y.V3）を中心に検出が相次いでいた N501Y 変異、いわゆる、イギリス株、ブラジル株など、のちにアルファ株と称される変異に関連する変異が日本国内でも検出され始めたことを契機として、これらの拡散状況及び、従来株である野生型、欧州型（D614G）の変異を表示することが提案された。

プロセスの概略

フェーズ 4 のタスク遷移図は図 4.17 に示す。タスク一覧を表 4.5 に示す。本フェーズでは、開始時点で目指す展示形式や鑑賞者へ提供する操作が具体的に I によって示され、結果的にはその通りの実装がなされた。これは開始から 1 ヶ月未満という短期間で展示を完成させる必要があり、またインタラクティブな展示の計画について一定程度の経験を I が有していたこと、前フェーズでのビジュアライズで提示するストーリーと可視化方法を踏襲しており、ビジュアライズ内容について試作を繰り返す必要がなかったことが背景にある。

まず、I の提案に沿って筆者は、N501Y.V1、N501Y.V2、N501Y.V3 をフィルタリングした (4-1) これは、前フェーズの第 11 レビュー (3-20) と同じ結果になるが、Nextstrain におけるデータの定義が変更となったため必要な作業である。データを 2 月 16 日に再取得した上で表現した (4-2)。(3-20) に比べて日本へ N501Y.V1 が流入していることが確認できた。これに加え、系統樹を表現する曲線の描画方法の改善を行なっている (4-3) (図)。これはこれまで直線を連続させて曲線を表現していた点を改善し滑らかにするものである。改善の意図について筆者は「これまでは映像という形で「動き」がついた状態だったのであまり気にならなかったのですが、展示で任意の場所をじっくり見る場合は際立ってしまう可能性もあり (中略) ベジェ曲線で表現できればと思いますが、いかがでしょうか。」と述べている。8K という高解像度ディスプレイを使用すること、ズーム操作に配慮した意図がうかがえる。

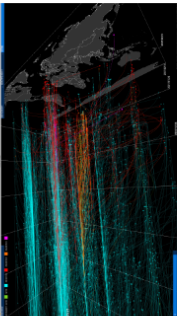
第 12 レビュー (4-4) では I、O ともに同意した。

続く (4-5) では O が提示用のインタラクションを実装し、第 13 レビュー (4-6) を行なった。ここで SS モードのカメラ視点の回転中に、地図平面が系統樹よりも手前に表示される場合、系統樹を遮蔽して位置関係が理解しにくくなる点を I が指摘した。この対処方法について筆者は地図平面の透過率を上げることを提案し、O は改善を行なった (4-7)。

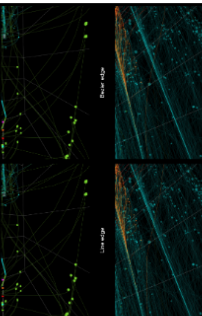
第 14 レビュー (4-8) では、この改善がなされたことを確認した。I は「地図の裏側に回った際の半透明もちょうどよいバランスと感じました。」とコメントしている。また同レビューでは、筆者が用意した 3 月 2 日に再取得したデータに差し替えた。さらに多くの N501Y 変異が日本で検出されていることが確認でき、これを展示するデータとして確定した。

カテゴリ	2/23	2/24	2/24	2/24 - 3/1	3/2	3/3 - 3/4	3/5
取得							
解析							
フィルタリング							
マイニング							
表示用階級							
探索用階級							
探索用インタ							
レビュー							
表示用階級							
表示用インタ							

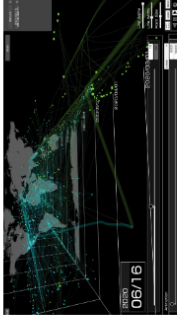
	(4-1)NS01Yの垂種をフィルタリング						
	(4-2)武漢型、欧州型、NS01Y (英、ブラジル、南ア) を表現						
	(4-4)NS01Yの垂種と、曲線の表現更新をレビュー	(4-6)操作、ムービー、SSの各モードをレビュー	(4-8)地図による運搬時の表現をレビュー				
	(4-3)曲線の表現の滑らかさを改善	(4-5)操作モード、ムービーモード、SSモード	(4-7)地図による運搬時の調整				



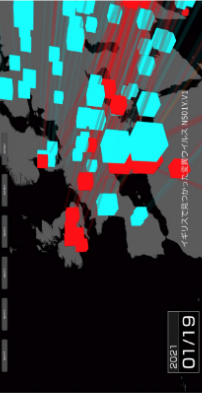
第12レビュー(4-4)
NS01Yの垂種(上) と、曲線の表現更新 (下)



第13レビュー(4-6)
操作モードムービーモードをレビュー



第14レビュー(4-8)
地図による運搬時の表現が改善された状態



第14レビュー(4-8)
地図による運搬時の表現が改善された状態

図 4.17 フェーズ 4 のタスク遷移図

表 4.5: フェーズ 4 のタスク一覧

ID	カテゴリー	時期	タイトル
タスク内容			
4-1	フィルタリング	2/23	N501Y の亜種をフィルタリング イギリスで見つかった変異ウイルス N501Y.V1、南アフリカで見つかった変異ウイルス N501Y.V2、日本で検出されたブラジル由来の変異ウイルス N501Y.V3 を追加
4-2	表現	2/24	武漢型、欧州型、N501Y (英、ブラジル、南ア) を表現 「野生型」「欧州型」「N501Y.V1」「N501Y.V2」「N501Y.V3」の出現を表現
4-3	提示用精緻	2/24	曲線の表現の滑らかさを改善 インタラクティブに操作すること、高解像度で展示することを考慮し、曲線の描き方を直線の連続からベジェ曲線に変更
4-4	レビュー	2/24	N501Y の亜種と、曲線の表現更新をレビュー 筆者コメント:「ストレインを結んでいる線について、現在は添付画像 (fig01) の上のように直線を細かく連結させて曲線を表現しているが、この方法だと直線の集まりなので 8K の高精細な表現にはやや不自然さがある。これまでは映像という形で「動き」がついた状態だったのであまり気にならなかったが、展示で任意の場所をじっくり見る場合は際立ってしまう可能性もあり、できれば fig01 の下にあるように、ベジェ曲線で表現できればと思う」 / O コメント: ベジェ曲線で描画する機能はないため計算して頂点列化して対応する。json が小さくなるのは魅力がある。 / I コメント: O が問題ないのであれば、データも軽くなり、綺麗になるので異存はない。
4-5	提示用インタラクション	2/24 - 3/1	操作モード、ムービーモード、SS モード O は I のディレクションに基づき展示用インタラクションを実装。一定時間操作しないと系統樹が自動的に回転するモード / ボールを回すと系統樹が回転 / ホイールを回すと時間が変化 / 何もせずに 3 秒立つと自動で時間が進みながら回転 / トラックボールの右下のボタンでムービーモード。ムービーモード中に、トラックボールのボールの回転、ホイールの回転を行うと操作モードに遷移
4-6	レビュー	3/2	操作、ムービー、SS の各モードをレビュー I コメント: トラックボールになったことで、操作性の幅が広がった。とてもスムーズに動くのと、ホイールを回すと地図が固定されながら時間軸だけが動くのが良かった。よく見ると系統樹のカーブも綺麗。SS モードで視点が裏側に回り込んだ際に、地図が透過しているため、系統樹と地図の関係性がよくわからない。裏側に回り込んだ際には、地図が系統樹を遮る形になり、地図が透過しないようにできるだろうか。それが難しい場合には、裏側まで回り込まないという動きでもよい。 / 筆者コメント: 「スクリーンセーバーモードの際は、主役である系統樹だけがはっきり見えていて、地図やグリッドは、例えば、明度を 70%、アルファを 50% 程度に控えめに表示すると良い。それであれば、裏側に回り込んだときもさほど違和感ないように思う」
4-7	提示用精緻	3/3 - 3/4	地図による遮蔽時の調整 地図によって系統樹が遮蔽される時は地図に透過度を与えて系統樹を遮蔽しないようにする。
4-8	レビュー	3/5	地図による遮蔽時の表現をレビュー O コメント: 地図データに関しては半透明にした。データ線も半透明なこともあり、あまり薄いとやはり前後関係がわかりにくいので、うっすら見える程度にした / I コメント: 地図の裏側に回った際の半透明もちょうどよいバランスになった。

公開までのプロセス

本フェーズの最終レビュー (4-8) 後、操作モードから SS モードに遷移するまでの無操作待機時間の調整、現地設営状態にて線の太さやドットの大きさを調整した上で一般への公開を開始した。

公開されたビジュアライゼーション

展示では下記の凡例を併置して来場者が自由に操作できるよう設置した。

「変異ウイルスの拡散」

表示されている四角い点は、すべて新しく見つかった変異ウイルスです。

この最示では5つの色に分類して代表的な変異ウイルスを表示しています。

- 黄：中国・武漢で初めて報告された新型コロナウイルス
- 水色：欧州から世界へ広がった変異ウイルス（D614G）
- 赤：イギリスで見つかった変異ウイルス（N501Y.V1）
- オレンジ：南アフリカで見つかった変異ウイルス（N501Y.V2）
- 紫：ブラジルで見つかった変異ウイルス（N501Y.V3）

設営された展示とスクリーンショットを図 4.18 に示す。ムービーモードで表示するテロップを表 4.6 に示す。

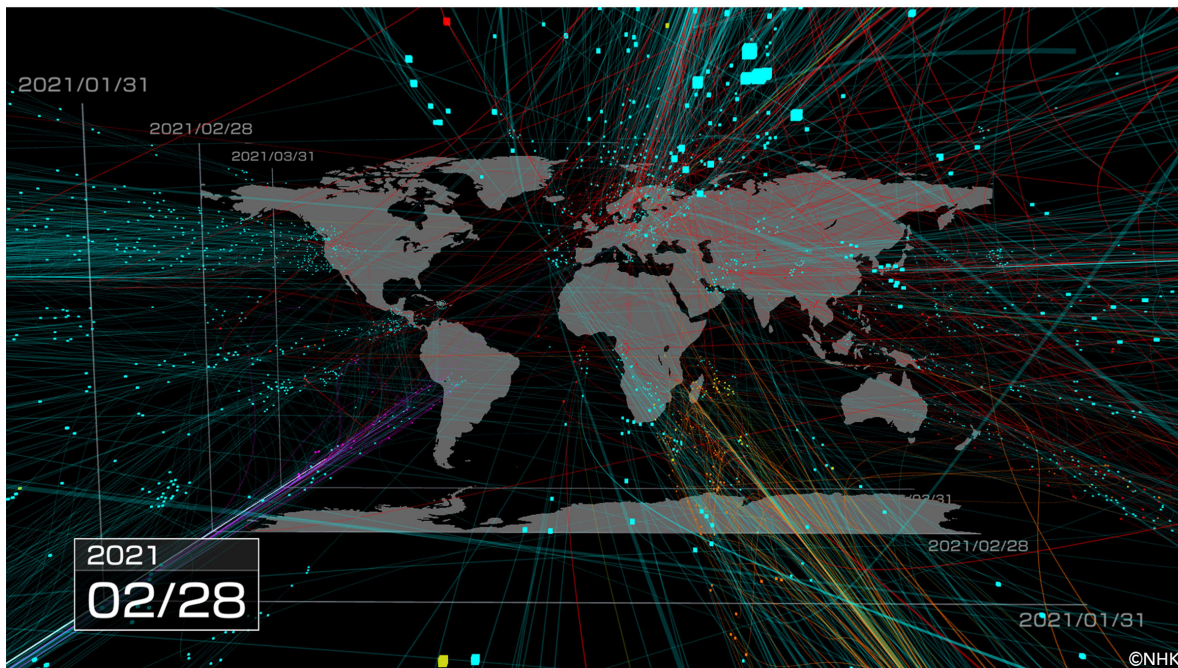


図 4.18 フェーズ 4 の展示の様子 (上) スクリーンショット (下) 設営された展示

表 4.6 展示されたムービーモードのテロップ

秒数	テロップ
0:03-0:07	武漢で初めて報告されたウイルス
0:11-0:15	欧州で広がった変異ウイルス D614G
0:27-0:31	イギリスで見つかった変異ウイルス N501Y.V1
0:34-0:38	南アフリカで見つかった変異ウイルス N501Y.V2
0:45-0:49	日本で検出されたブラジル由来の変異ウイルス N501Y.V3

タスクのネットワーク

フェーズ4のタスクネットワークを図4.19に示す。前述の通り、本フェーズでは開始当初より提示する内容を決定した上で、展示する8Kディスプレイや鑑賞者に提供するインターフェースの精緻化に主眼を置いて短期間で開発が進められた。そのため、表現とレビューの反復は見られず、鑑賞者のインタラクションと8Kならではの高精細に適した表現を高めるための反復が見られる。

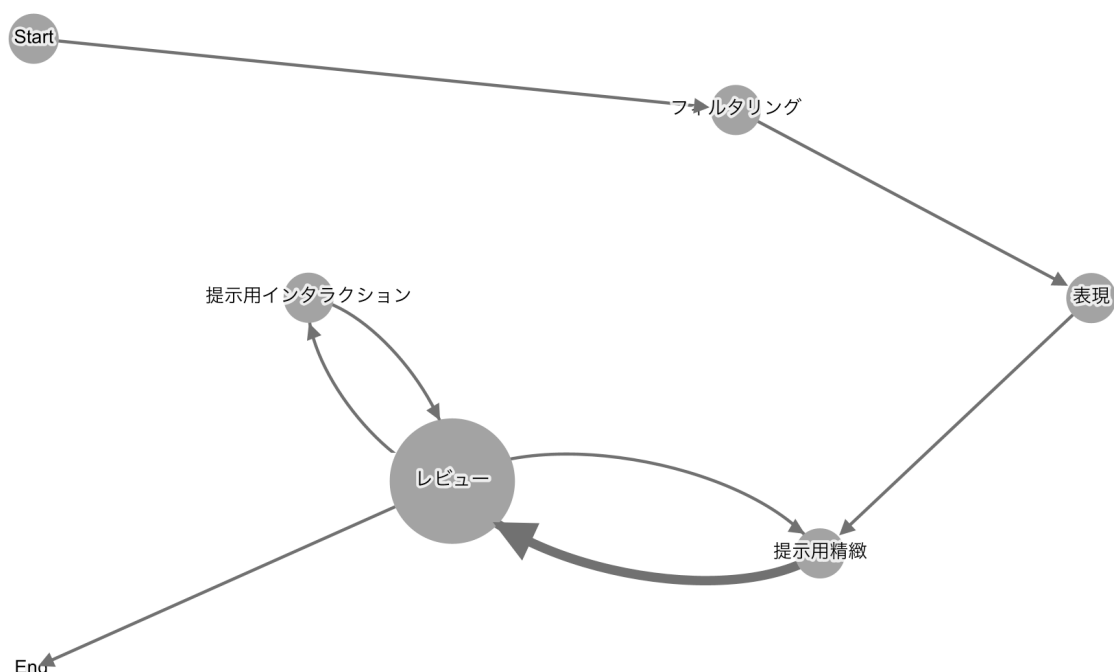


図 4.19 フェーズ4のタスクネットワーク図

4.4.6 フェーズ5：第5回公開までの記録

時期

2021年6月7日から6月22日

開始時の概況

開始時における感染状況は第4波のピークアウトにあたる。全国の新規感染者数は、報告日別で5月中旬以降減少に転じ、直近の1週間では10万人あたり約26人となってい

た [67]。変異株について 6 月 11 日時点の報告によれば、国内のウイルスはイギリスで初めに検出されたアルファ株 (B.1.1.7) へほぼ置き換わった。国外ではアルファ株よりも感染力・伝播性が高いと見られるインドで初めに検出されたデルタ株 (B.1.617.2) がアジア諸国を中心に感染を拡大させていた [68]。I らは、アルファ株が海外で先行して拡大し、その後、国内でも流行するという現象に着目して、変異株の国内への流入問題や、政府による検疫の強化、いわゆる「水際対策」について取り上げる報道番組を企画した。その一部として、本プロジェクトのビジュアライゼーション映像を使用し、ヨーロッパで拡大したアルファ株やアジアで拡大したデルタ株が、さまざまなルートで日本に流入する様子を可視化することを試みた。本フェーズ開始時においてどのようなビジュアライゼーションを作成するかについては未定であったが、放送日は開始から約 2 週間後の 6 月 23 日に決定していた。

本フェーズ開始時において、I らは遺伝学を専門とし新型コロナウイルスゲノム解析を行っていた大学医学部教授である S らへの取材を終えていた。I らは S への取材から、ウイルスの遺伝子解析をもとに扱うテーマについて次の関心を寄せていた。それらは、変異株 (イギリス株とインド株) について、(1) どのような経路で世界から国内へ流入したか、(2) 国内ではどのような経路で拡散したか、(3) 検疫対策は十分だったかである。本フェーズではこれらの関心に沿って、試作を進めることとなった。また本フェーズにおけるデータソースは、S または S の研究補助員である U から提供を受けた。S らのデータと本フェーズ以外のデータの違いは、本フェーズ以外では GISAID に登録されたウイルス変異ゲノムをランダムサンプリングし Nextstrain が提供するデータ (Global Sampling) を使用している。一方、本フェーズでは、GISAID のデータからサンプリングしたものであるが、S や U が注目する地域や変異を指定し、目的の現象を分析することに適したサンプリングを行ったデータを使用することとなった。そのサンプリングの調整は主に U によって随時行われた。

プロセスの概略

VisE のタスク遷移図は図 4.20 に示す。タスク一覧を表 4.7 に示す。

本フェーズでは、次の 3 つの独立した試作プロセスが並行して行われた。第 1 にイギリス株の経路の可視化、第 2 にインド株の経路の可視化、第 3 に検疫での陽性者から検出された変異とほとんど同じ変異が国内で検出されたケースの可視化である。これらはそれぞれ E、I、Q と略号を付して「VisE」「VisI」「VisQ」と表記する。タスク ID にも同様に e、i、q の略号を付与して表記する。

イギリス株流入経路の可視化 (VisE) の概略

Sらは日本で検出された変異株に注目して GISAID からサンプリングしたデータを提供した (5-1e)。このデータはこれまでのフェーズで使用してきた Nextstrain が提供する全世界の検出変異をランダムサンプリングしたデータ (Global Sampling) とは異なり、日本において検出した変異に特化してサンプリングしたデータである。検出位置についてはこれまで国情報 (変数名: Country) をもとに各国の代表地点を可視化に使用していたが、国内での変異の広がり方について関心を寄せていたことから、今回は都道府県情報 (変数名: Admin Division、以下 AD) をもとに各都道府県の中央位置を可視化に使用することを試みた。まず、データから AD を抽出する解析を行った (5-2e)。AD 値の頻度を調べたところ (5-3e) 全体の半数程度は AD 値を保持していなかった。このうち、検疫で検出された変異は Country 値が JapanQuarantine となっていた。これらは便宜上、本州の南方沖の座標を設定した。検疫検出でもなく AD 値を持たないものについては検出地不明として取り扱うこととした。これらは便宜上本州の東方沖の座標を設定した。従来株 (欧州型) を青、イギリス株を赤として表現した (5-4e)。この可視化および、この可視化から検疫検出変異だけをフィルタした可視化をもってレビューを行った (第 15 レビュー、5-5e)。

第 15 レビュー (5-5e) では、I は「イギリス株が徐々に国内で広がって従来株と置き換わるのがわかる」とコメントし、置き換わりの時間的な進行や空間的な広がりやのわかりやすさを評価した。一方、国内の経路については「国内がかなり線が交錯して、あまりにデータの数が多すぎるのかも知れない。海上にある「Japan」にデータが集まる位置は、直感的には理解が難しいと感じる。」とコメントし、国内の経路を示すエッジが密集しすぎ、理解しにくいことを指摘した。また、「Japan データが集まる位置」とは AD 値を持たない変異を東海上に集めたことを指しているが、このような欠損値の表現処理についても疑問を呈している。検疫検出変異については、「房総沖などにある「keneki」の位置と合わせて、うまい整理を考えたい」とコメントし、現状では直感的な理解とうまく結びついていないが、改善する余地が残ることを示した。レビュー後、AD 欠損値が多いことから国内での拡散経路を正確に表現することは難しいため、国内経路についての可視化は番組では使用しないこととなった。また検疫検出変異についてはサンプリングデータが少ないことから、サンプリング数を調整し再度可視化することとなった。

レビューの結果から、国内の遷移については番組では取り扱わないこととし、イギリス株の国内への流入経路や検疫検出変異の到達経路に関心を向けて試作することとなった。Uらは検疫検出変異の数を増やすように調整したサンプリングデータを提供した (5-6e)。

前回の表現では検疫検出変異を洋上の同じ座標に表現したため、国外のどの地域から到達したかを読み取ることが困難であったことから、日本の代表座標を中心とする同心円周上に配置した。円周上のどの位置にプロットされるかは、日本代表座標と流入エッジの始点座標の角度によって定まる。そのため流入元の方法が可視化される (5-7e,5-8e)。この可視化および、この可視化から検疫検出変異だけをフィルタした可視化をもってレビューを行った (第 16 レビュー、5-9e)。

第 16 レビュー (5-9e) で I は「日本が地図の一番右端にあることで、検疫が左側でストップされているように見えると理解した。」とコメントした。これは検疫検出変異が日本と流入元の間にあるため「壁」のように見えることを語っている。また、検疫検出変異だけをフィルタした可視化については「手前側の 3 つの検疫から奥への延びるように線が引かれると予想したが、現在の状態だとすべて中東から入ってきているように見える」と述べている。これはエッジが分岐する位置の特定の方法に問題があり、国内で分岐したと推測できる系統についても、国外で分岐して流入したように見える問題について指摘している。これについて分岐位置の可視化方法を見直すこととなった。I のコメント「国内蔓延している株と、海外の近いものが直接紐付けされて線が引かれるとすると、誤解を防ぐには一旦国内に入った後で系統樹の下流に紐付く株は、海外からの線が引かれないように処理される必要があると感じる。」については、国外から国内に入った系統が再び国外へ展開するケースがあるが、今回については国内への流入経路に注目することから、国内から国外へ引かれるエッジを消去することを提案している。

系統が分岐した位置を空間的に決定する方法に標準的なものはなく、本プロジェクトではフェーズ 1 (1-20) で開発した分岐を束ねるバンドリングを独自の手法を使用してきた。すなわち、共通の幹の下流にある同じ変異数の分岐 (直接の「子」と見做せる分岐) が複数ある場合は、その分岐の中で時間的に最も早く検出された変異の検出位置を分岐位置として決定し分岐の位置を少なく見せ、過剰な複雑さを抑制するものである。この手法では、例えば、幹が国外にあり、同じ変異数を持つ分岐が複数ある中で、最も早く検出された変異が国外にあるケースでは、その後、国内で検出された変異も国外に分岐位置をもつことになり、国外からエッジが描かれる。これが第 16 レビュー (5-9e) での I の指摘の原因である。この点を解消するため、分岐とその「親」にあたる分岐の Country 値を比較して、両者が異なる場合は「親」の Country 値を無視して、「子」の値を採用することとした。これにより、国内で分岐したと推定される分岐とその下流にある変異は、国内の位置同士で結線されることとなり、指摘の解消に効果を発揮する。この精緻 (5-10e) を採用した可視化を試作し (5-11e)、レビューを行った (5-12e)。

第 17 レビュー (5-12e) では、精緻 (5-10e) の結果が議論となった。I は「イメージに

近いと感じた。検疫からではなく中東からまだ数本の線が伸びているのはどういう理由だろうか。」とコメントしている。事前の想定との乖離が少なくなったことを評価しつつ、依然として国外から引かれるエッジについて疑問を呈している。筆者はこれについては元のデータを踏まえて解説し表現の妥当性を示しながらも、Sらの確認と検証を求め、Iもこれに同意した。Uらは問題となっている特定の変異（図中①～⑤）についてフィンガープリントを調べ、④⑤については国内分岐の変異と推定した。筆者はこのような推定に基づいて特定の変異の分岐位置を個別に指定できるプログラムを実装した。

第16レビュー（5-9e）でIが指摘した、国内流入した系統の下流でふたたび国外の変異が表示されることが視認性を下げることにについて、対策を試みた。国内流入系統の下流で国外検出した系統を削除する処理を行い可視化した（5-13e、5-14e）。これをレビューしたものが第18レビュー（5-15e）である。

第18レビュー（5-15e）では、データは日本検出の変異の上流にあたる国外検出変異が少ないサンプリングを使用しているため、国内流入の下流で国外検出される系統を削除（不可視化）すると可視化される系統が減少し、流入元の分布があまり視認できていない点が浮かび上がった。Iは「海外での初期のデータを多めにしたデータセットを作ることによって解決できないか」と提案し、Uは「サンプリングの数などパラメータを調整しつつ解析を試みる」と応じ、具体的なサンプリングパラメータを提案した。

Uが再サンプリングしたデータ（5-16e）を可視化し（5-17e）、第19レビューを実施した（5-18e）。

第19レビュー（5-18e）では、Iは懸念された国内から国外へ流出する系統がほぼ不可視化できたことを評価した。一方、「海外・国内ともにかなりデータ数が多いためか、欧州からのラインが多く引かれる見た目になっている。中東やオセアニアからのラインもあるが、ほとんどが欧州からラインがひかれている。」とコメントした。多様な流入経路があることから、欧州だけが突出することを防ぎたいという観点からフィルタリングのパラメータを再度調整することとなった。

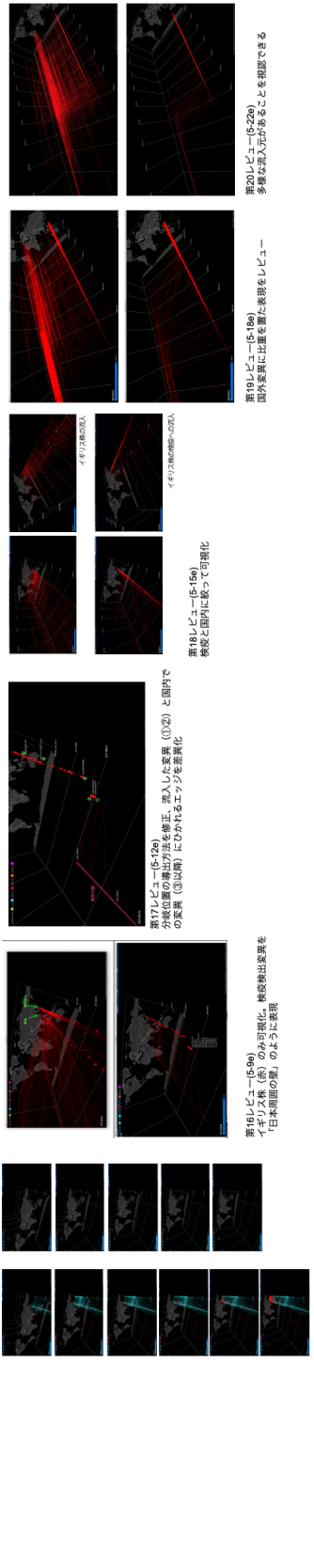
Uが再々サンプリングしたデータ（5-19e）を可視化し（5-20e）、第20レビューを実施した（5-21e）。

第20レビュー（5-21e）では、流入元の系統が欧州に過剰に偏って可視化されることを抑制でき、国内から国外への流出系統も抑制できた。Iは「Sが伝えようとしている点も押さえつつ、データが多すぎて理解できないという点も解消されていると判断した。」とコメントして評価し、この可視化データを番組用に調整することとなった。番組用CG化に向けて透明度の調整を行った（5-22e）。

VisEの可視化は、番組で取り扱う内容に沿って使用が検討されたが、イギリス株の国

内流入について解説するシーンは後述する別の可視化を用いることとなり、Iらの最終的な判断で番組での使用は見送られた。しかし、後述する VisI、VisQ に先行して実施されたため、VisE のプロセスで得た手法が VisI、VisQ の手法に影響を与えた。

カテゴリ	6/4	6/4	6/7	6/8	6/9-12	6/12	6/13	6/13	6/15	6/15	6/16	6/18	6/18	6/18	6/18	6/18	6/18
取得	(5-16)国内検出装置の系群データ取得																
解析	(5-20)検出場所データの解析																
フィルタリング	(5-56)検出装置に注目して系群を削除																
マイニング	(5-130)国内流入系群の下流で国外検出した系群を削除																
表現	(5-196)国内より上流の国外検出装置を動的にサンプリング (5-196)装置のサンプリングのバランスを再変更																
探検用機能	(5-30)国内各都市での検出と検出検出に分類 (5-176)装置を変更したサンプリングデータで可視化 (5-200)再度サンプリングしたデータで表現																
レポート	(5-46)座標検出装置の流入元を視認し易く (5-100)分岐位置の導出手法を修正 (5-176)検出装置の流入元を視認し易く (5-100)分岐位置の導出手法を修正																
探検用機能	(5-59)従来からの履歴検出と検出をレビュー (5-120)流入した装置と国内で生じた装置の検出をレビュー (5-150)検出と国内検出した系群を表現をレビュー (5-210)国外装置に比重を置いた表現をレビュー (5-220)検出装置CG用機能																
探検用機能	(5-59)従来からの履歴検出と検出をレビュー (5-120)流入した装置と国内で生じた装置の検出をレビュー (5-150)検出と国内検出した系群を表現をレビュー (5-210)国外装置に比重を置いた表現をレビュー (5-220)検出装置CG用機能																
探検用機能	(5-59)従来からの履歴検出と検出をレビュー (5-120)流入した装置と国内で生じた装置の検出をレビュー (5-150)検出と国内検出した系群を表現をレビュー (5-210)国外装置に比重を置いた表現をレビュー (5-220)検出装置CG用機能																



第15レビュー(5-59) 従来からの履歴検出(59)への
 従来検出(59)からイナズマ穴(59)への
 流入経路が可視化され、
 右は経路上にプロット、右は検出
 出のみ可視化

第16レビュー(5-90)のみ可視化、検出検出装置を
 「日本国の置」のように表示

第17レビュー(5-120) 分岐位置の導出手法を修正、流入した装置(020)と国内で
 の装置(59)に絞って可視化

第18レビュー(5-150) 検出と国内に絞って可視化

イナズマの穴
 イナズマの穴の穴

第19レビュー(5-180) 国外装置に比重を置いた表現をレビュー
 国外装置に比重を置いた表現をレビュー

第20レビュー(5-220) 多様な流入元があることを確認できる

図 4.20 フェーズ 5 (VisE) のタスク遷移図

表 4.7: フェーズ 5 (VisE) のタスク一覧

ID	カテゴリー	時期	タイトル
5-1e	取得	6/4	国内検出変異の系統データ取得
日本で検出された変異のデータを K から受領			
5-2e	解析	6/4	検出場所データの解析
日本国内の検出場所 (Admin Division) を解析			
5-3e	マイニング	6/4	国内各都市での検出と検疫検出に分類
検疫での検出、国内各都市での検出、国内不明地検出に分離			
5-4e	表現	6/7	従来株からの置き換わりと検疫を表現
欧州株とイギリス株を着色、各県の検出変異は県庁所在地にプロット、検疫の検出変異は本州の南洋上に、国内の検出場所不明変異は本州東洋上にプロット			
5-5e	レビュー	6/7	従来株からの置き換わりと検疫をレビュー
I コメント: イギリス株が徐々に国内で広がって従来株と置き換わるのがわかる。このデータセットだと検疫が少ないので分析にはあまり役立たないが、どんなビジュアルになるかはわかる。国内がかなり線が交錯して、あまりにデータの数が多すぎるのかも知れない。海上にある「Japan」にデータが集まる位置は、直感的には理解が難しいと感じる。房総沖などにある「keneki」の位置と合わせて、うまい整理を考えたい。/協議の結果、国内で場所が特定できない変異が多いことを踏まえ、国内での拡散経路を表現することは難しいと判断した。			
5-6e	フィルタリング	6/8	検疫検出変異に着目して系統をフィルタ
U による検疫で検出されたイギリス株を焦点にサンプリング			
5-7e	探索用精緻	6/9-12	検疫検出変異の流入元を視認し易く
検疫のプロット位置を一定の場所ではなく、流入元の位置と日本の位置の線上に置くことで流入元の方がわかるように修正。国内検出変異は検出地不明が多かったことから、都市ごとのプロットは行わず、日本の代表地点座標周辺にプロットとすることとした。			
5-8e	表現	6/12	検疫検出変異の流入元を視認し易く表現
検疫のプロット位置を修正 (日本の周囲を取り囲むように)			
5-9e	レビュー	6/13	検疫検出と国内流入状況をレビュー
I コメント: 日本が地図の一番右端にあることで、検疫が左側でストップされているように見えると理解した。国内流入していないという直感的な理解とも齟齬がないのでこれでいいように思う。中東から流入したものが、国内で広がったケースと考えると、奥の国内に広がっているものについても中東から線が引かれているのが気になる。手前側の 3 つの検疫から奥への伸びるように線が引かれると予想したが、現在の状態だとすべて中東から入ってきているように見える。国内蔓延している株と、海外の近いものが直接紐付けされて線が引かれるとすると、誤解を防ぐには一旦国内に入った後で系統樹の下流に紐付く株は、海外からの線が引かれられないように処理される必要があると感じる。			
5-10e	探索用精緻	6/13	分岐位置の導出手法を修正
系統樹分岐位置推定のアルゴリズムを変更。フェーズ 1 (1-20) で開発した分岐を束ねるバンドリング手法を改良し、分岐とその「親」にあたる分岐の Country 値を比較して、両者が異なる場合は「親」の Country 値を無視して、「子」の値を採用することとした。			
5-11e	表現	6/13	流入した変異と国内で生じた変異を表現
系統樹分岐位置推定のアルゴリズムを変更し、国内で流行したことが自然に見える表現に変更			
5-12e	レビュー	6/14	流入した変異と国内で生じた変異の表現をレビュー

(次ページへ続く)

表 4.7 – (前ページから続き)

ID	カテゴリー	時期	タイトル
	タスク内容		
	<p>I コメント：P1 と P2 確認し、イメージに近いと感じた。P2 で、検疫からではなく中東からまだ数本の線が伸びているのはどう理由だろうか。系統樹ではそのようにみえないため念の為確認したい。／筆者コメント「これについては、実際、中東から入ってきている変異系統が見えている結果なので、これで正しいかと思われる。念のため系統樹と対応させて番号をふると、まず、系統樹の紫色の線（ブランチ）を見ると、3 次元表示では中東（United Arab Emirates）の位置に描かれている。恣意的にそうしているわけではなくこのブランチの Country が「United Arab Emirates」になっているため。（その対応として、例えば、このブランチの AdminDivision は「JapanQuarantine」なので、検疫の位置に線を表示することも考えたが、そうすると、変異とブランチが同じ位置に描かれるため、どこから来たのかがわかりにくくなる。）中東から直接入ってきているのは⑤で、③や④は 1 つ上流の分岐が日本国内にある。ただ、③、④の上流に国内で見つかった変異はないのでこれらも直接入ってきたとみなせると考えられる。」／I コメント：①②は JapanQuarantine、⑤は中東から直接入った可能性が高い、③④は国内での分岐がある、上流で見つかった変異がない（フィルタリングの過程で落ちている可能性もあるか）について、どのようなデータから言えるのか。／筆者コメント「口頭ではなかなか説明しづらいので、昨日送った系統樹を「Country」で色付けしたものを添付する。ちなみに、昨日送ったものは「Admin Division」（行政区画）で色付けしたものである。データはいずれも U からのもの。添付の図で見ると、左上の凡例に出ていないがグレーは「United Arab Emirates(UAE)」である。この図から、①②＝UAE で分岐して検疫で採取された③④＝UAE で分岐したものがさらに日本で分岐して日本で採取された⑤＝UAE で分岐して日本で採取されたとわかる。「検疫で分岐する」というのは現象としてもあり得ないと思われるため、これに合わせて 3 次元化している状態である。解釈として正しいのかは、S に確認が必要。」／I コメント：理解できた。ただ抜け落ちているデータがある（とくに国内株）という前提で考えると、②は系統的に近い⑤の祖先、①も系統的に近い③の祖先と考えるのも考え方として自然ではないか。一方で④は近い祖先がないため検疫でつかまらずに入り込み、国内で分岐したという解釈も成り立つ。ただ Nextstrain では「Country」でわけていることも考慮して S に確認の必要があると考える。／K らが①～⑤のフィンガープリントを精査し、変異数の観点から中東から横のノードが引かれるのは、①と②だけで良く、④⑤はその下にぶら下がる形として問題ないと判断した。</p>		
5-13e	フィルタリング	6/15	国内流入系統の下流で国外検出した系統を削除 日本で検出された変異を親にもつ海外で検出された変異を削除
5-14e	表現	6/15	検疫と国内検出した系統を表現 日本で検出されたイギリス株のルーツを表現・日本の検疫で検出された変異とその子孫にあたる国内の変異を表現
5-15e	レビュー	6/16	検疫と国内検出した系統を表現をレビュー 日本検出の変異の上流にあたる国外検出変異が少ないサンプリングデータを使用しているため、国内流入の下流で国外検出される系統を不可視化すると可視化される系統が減少し、流入元の分布があまり視認できていない。サンプリングを上流の国外変異を優先してサンプリングしなおすこととした。また、検疫と国内は区別せず全て日本に到達した変異として表現することとした。／I コメント：海外での初期のデータを多めにしたデータセットを作ることで解決できないか／U コメント：サンプリングの数などパラメータを調整しつつ解析を試みる。具体的には 9 月から 11 月の海外株を多めに Subsampling し、合計 5 千件程度の海外株を Priority にする
5-16e	フィルタリング	6/18	国内より上流の国外検出変異を優先的にサンプリング 6/16 の提示条件で U がフィルタリング
5-17e	表現	6/18	変更したサンプリングデータで可視化 筆者コメント：U のイギリス株のデータを可視化した E1 と E2 を提示。受領データで日本から出る系統はほぼ消えている。
5-18e	レビュー	6/18	国外変異に比重を置いた表現をレビュー I コメント：日本から出る矢印についてはほとんど記載されず、わかりやすくなった。海外・国内ともかなりデータ数が多いためか、欧州からのラインが多く引かれる見た目になっている。中東やオセアニアからのラインもあるが、ほとんどが欧州からラインがひかれている。／多様な流入経路があることから、欧州だけが突出して見えることを防ぎたいという観点からフィルタリングのパラメータを調整することとなった。
5-19e	フィルタリング	6/18	変異のサンプリングのバランスを再変更

(次ページへ続く)

表 4.7 – (前ページから続き)

ID	カテゴリー	時期	タイトル
タスク内容			
I コメント：サンプル数は減っても、国内の感染の始まりにあたる株や、それと近い塩基配列の海外の株などのポイントになるものは残っていると理解しているが正しいか？ / U コメント：Probablic Sampling である程度ランダムにしている。日本のものでも残っているという理解で正しい。ただ、それでも除外される恐れがあるため日本の系統 (Strain) で、ある程度似たものうち早い日付という基準で数十個を選び、Global の系統樹に当てはめた。			
5-20e	表現	6/18	再々サンプリングしたデータで表現 再々サンプリングしたデータを表現
5-21e	レビュー	6/19	再々サンプリングした表現のレビュー I コメント：S が伝えようとしていた点も押さえつつ、データが多すぎて理解できないという点も解消されていると判断した。この試作を採用したい。
5-22e	提示用精緻	6/18	番組 CG 用精緻 番組 CG 用に不透明度を変更

VisE のタスクのネットワーク

VisE のタスクネットワークを図 4.21 に示す。本フェーズはこれまでのフェーズとは異なり、日本への変異株流入、防疫、国内での拡散経路という全世界と特定地域間の流入出に焦点を当てた。まず、日本だけを対象としたデータをこれまでのプログラムに流し込み、その可視化された様態から期待する現象が読み取れるかを確認している (5-5e)。そして、国内の経路という局所的な現象については推論できる可視化が得られないことから断念している。検疫検出変異とその後に国内で検出された同系列の変異を可視化し、その関係を可視化から明らかにすることについては、実態をうまく説明するように系統の分岐手法を精緻化し、それだけでなく変異のフィンガープリントを検証して確証を得ている (5-10e)。国外から国内へ向かう経路については、データのサンプリングの微調整を繰り返すことで偏りの少ない経路の可視化を追求している (5-16e~5-21e)。これまでに開発した基本的な可視化プログラムを目的に応じて改修するための探索的精緻と、データを調整するフィルタリングを介して表現とレビューの反復を繰り返している。

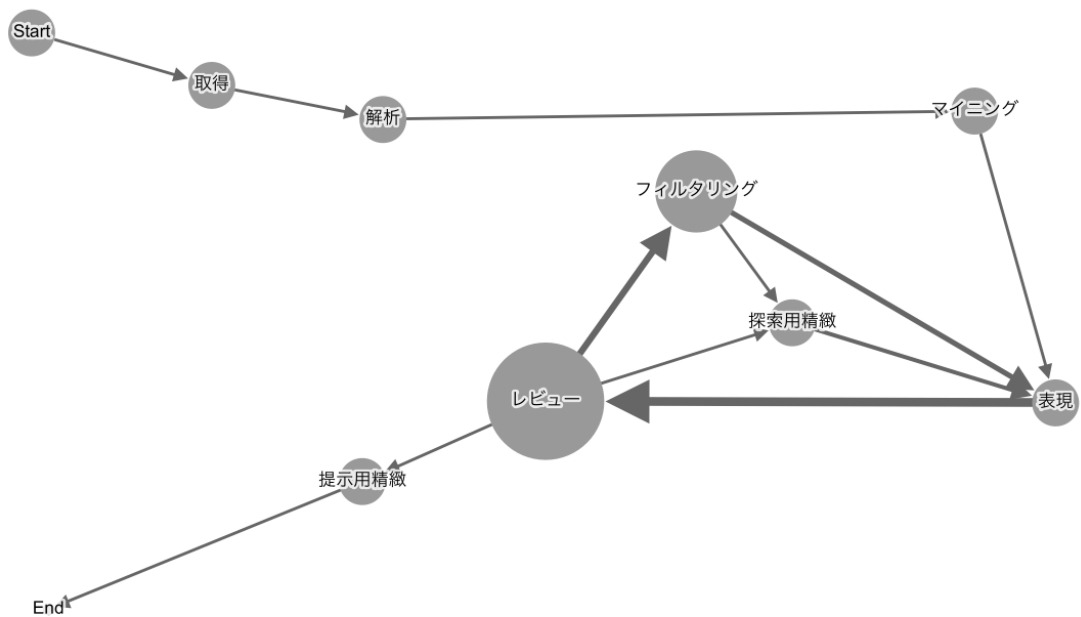


図 4.21 フェーズ 5 (VisE) のタスクネットワーク図

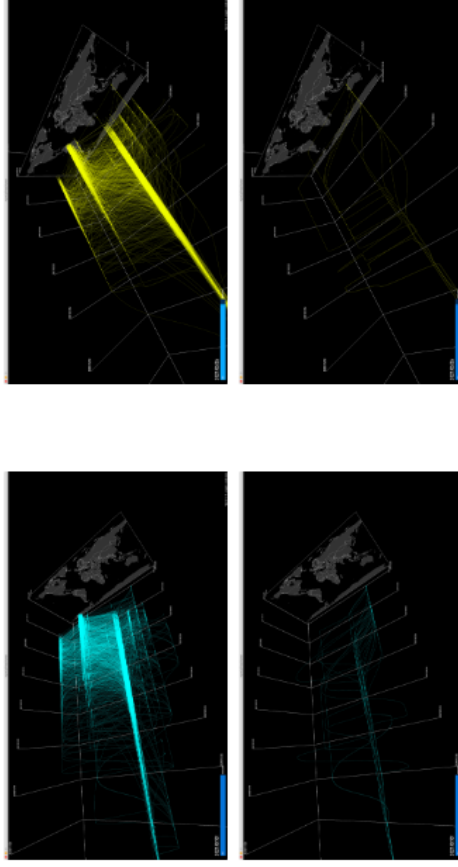
インド株流入経路の可視化 (VisI) の概略

VisI のタスク遷移図は図 4.22 に示す。タスク一覧を表 4.8 に示す。

インド株の日本への到達経路を把握したいという関心から、S、U らは GISAID から日本検出のインド株変異を優先的にサンプリングした系統データを作成した (5-1i)。検疫検出変異、国内検出変異、国外検出変異の 3 パターンのうち検疫と国内を統合し、系統全体を表示する可視化と、日本で検出された変異を含む系統だけを可視化する 2 種類の表現を作成した (5-2i、5-3i)。

第 21 レビュー (5-4i) では、I は「現状の内容で問題ない」とコメントした。これについて S や U から異論はなかった。その後、I は視聴者の識別のしやすさを加味して、番組内で紹介する場合に各変異株の色を統一する意向を示し、筆者は、着色を青から黄色に変更した。また VisE で実装した系統分岐位置算出アルゴリズムの改良 (5-10e) をインド株にも適用し、国内流入する経路の視認性を改善した (5-5i)。

カテゴリー	6/10	6/11-12	6/13-14	6/16	6/18
取得					
解析					
フィルタリング	(5-1)インド株の国内流入サンプリングデータ作成				
マイニング	(5-2)国内検出と国外検出に分類				
表現	(5-3)系統全体の表現と国内流入系統のみの表現				
探索用精緻					
探索用インタ					
レビュー	(5-4)系統全体の表現と国内流入系統のみをレビュー				
提示用精緻	(5-5)系統分岐位置算出アルゴリズム調整				
提示用インタ					



第21レビュー (5-4)
世界に拡散するインド株 (上) と日本に流入したインド株 (下) を表現

提示用精緻 (5-5) を施した後の系統図

図 4.22 フェーズ 5 (VisI) のタスク遷移図

表 4.8: フェーズ 5 (VisI) のタスク一覧

ID	カテゴリー	時期	タイトル
	タスク内容		
5-1i	フィルタリング	6/10	インド株の国内流入サンプリングデータ作成 U がインド株の流入動向を見るため、検疫検出データに注目してサンプリングした系統データを作成。「日本からヨーロッパへの戻りもありません」とコメント。
5-2i	マイニング	6/11-12	国内検出と国外検出に分類 国内検出変異と検疫検出変異を統合、国外検出と分離
5-3i	表現	6/13-14	系統全体の表現と国内流入系統のみの表現 日本に流入したインド株の系統と全ての系統を表現
5-4i	レビュー	6/16	系統全体の表現と国内流入系統のみをレビュー I コメント：現状の内容で問題ない。見せ方としてはすべての系統を薄く表示して、日本へ入ってくる系統と変異のみ濃く表示させたい。／S や U から異論や指摘はなかった。
5-5i	提示用精緻	6/18	系統分岐位置算出アルゴリズム調整 番組 CG 用に色、不透明度を変更した。また、改良版の系統分岐位置算出アルゴリズム (5-10e で開発したもの) を適用し、流入の視認性を高めた

VisI の公開までのプロセス

筆者は VisI の最終的な CG 用データを O らに提供した。I らはこのデータをもとにインド株の国内への流入の様子を表す約 1 分の CG を作成した。VisI の系統全体を投影した後、日本に流入した系統のみを映し出した。インド株の流行地であるアジアではなく、アメリカやイギリスからも流入があることを伝えるナレーションを付した。CG シーン最後に、S へのインタビュー映像を組み込んだ。インタビューで S は「比較的早い段階から英国や米国を介して日本に入ってきていることがデータからわかったことは驚きを感じた」と述べた。

VisI の公開されたビジュアライゼーション

図 4.23 は、放送された VisI の映像である。I らは VisI の CG カットの前に、これまでに欧州株、イギリス株、インド株の順に世界的に流行する株が出現したことを説明する CG カットを挿入した (図 4.24)。これはフェーズ 4 で開発したビジュアライゼーションのデータを更新して、インド株が出現する様子を納めた CG 映像である。

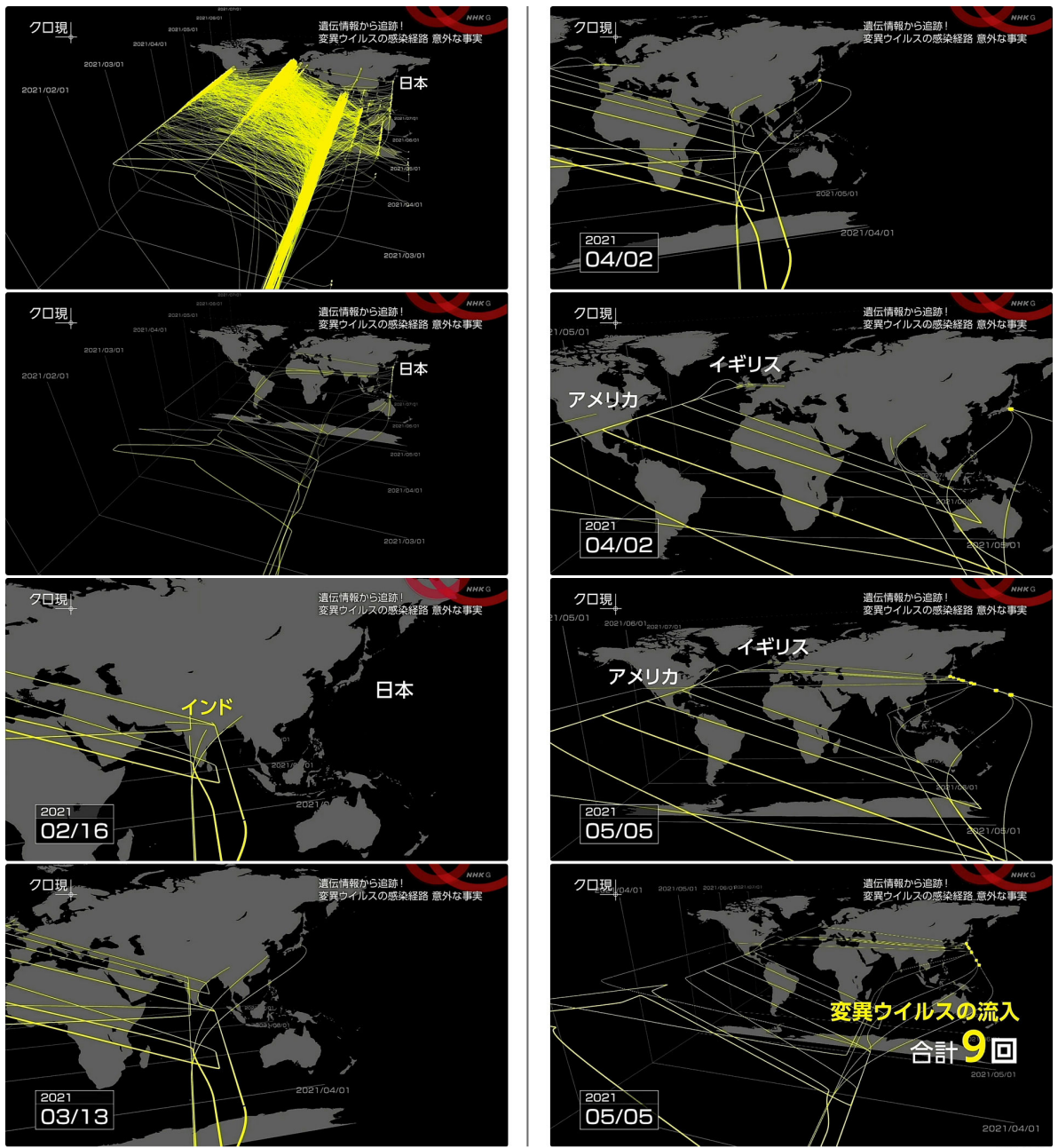


図 4.23 フェーズ 5 の放送時の映像 (VisI)

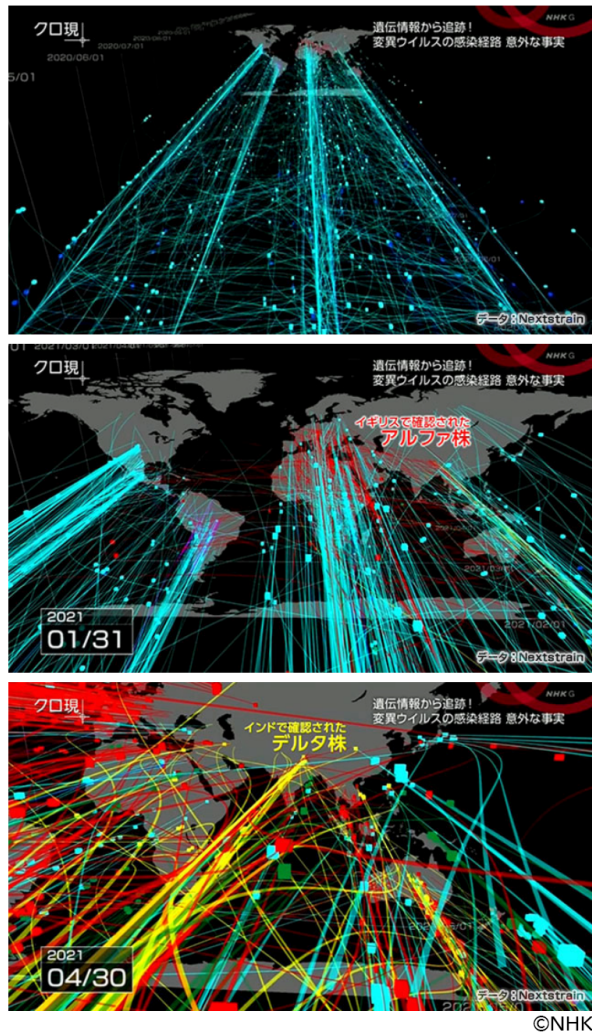


図 4.24 フェーズ 5 の放送時の映像 (いくつかの流行株の後でデルタ株 (インド株) が出現する様子)

VisI のタスクのネットワーク

VisI のタスクネットワークを図 4.25 に示す。

VisI では、開始時における関心を満たす表現ができたことが 1 度目のレビューで確認された。そのためタスク間の反復は起こらずに終了となった。VisE で得られた探索用精緻 (5-10e) がここでは提示用精緻 (5-5i) として用いられている。

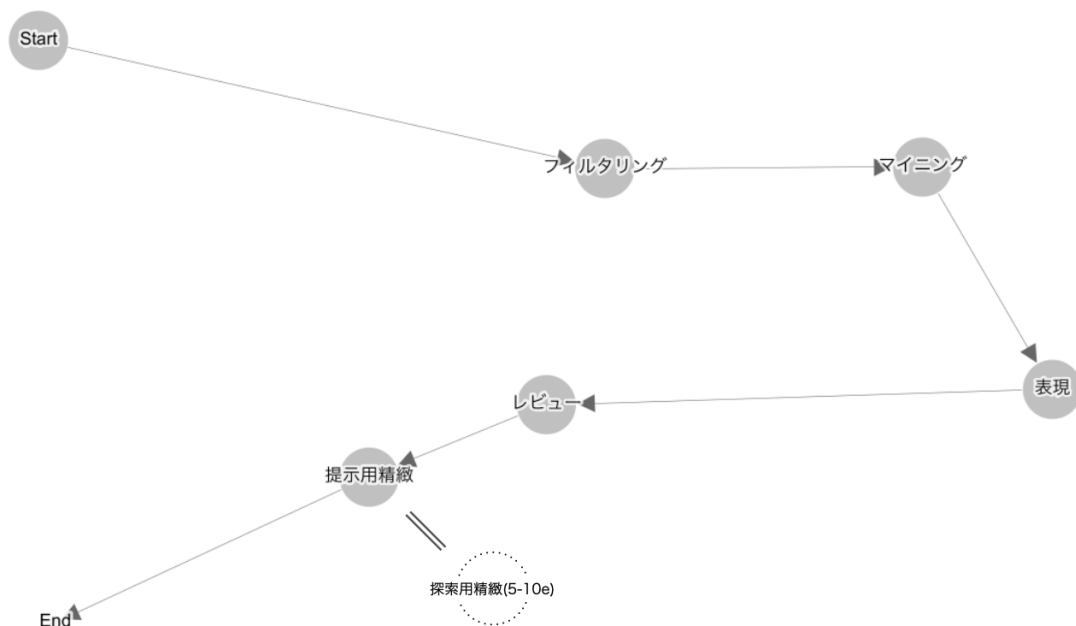


図 4.25 フェーズ 5 (VisI) のタスクネットワーク図

検疫からの流入の可視化 (VisQ) の概略

VisQ のタスク遷移図は図 4.26 に示す。タスク一覧を表 4.9 に示す。

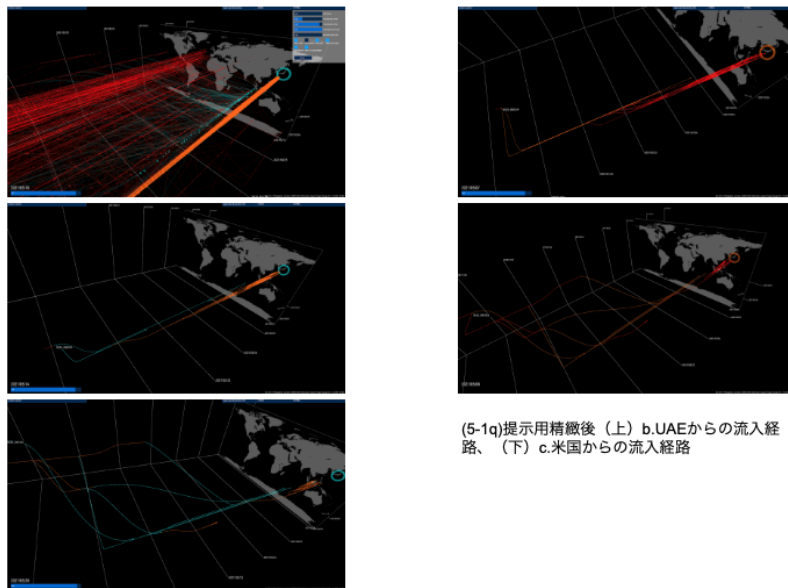
S らは、本フェーズ以前に、検疫陽性者から検出されたウイルスのゲノムと、国内で渡航歴のない患者の検出されたウイルスのゲノムの比較から、検疫のウイルス同じ由来を持つウイルスが国内で感染を広げていた可能性を示唆していた。VisQ ではその現象について番組内で可視化して解説したいという関心から試作を行った。I は S らとの議論の上で、次の 3 タイプの可視化の試作を提案した。(a) 国外で蔓延している系統、日本の検疫で検出された系統、日本の国内で検出された系統を色分けする。次に、現象を説明する特定の事象として、(b)a のうち、検疫で検出された特定の UAE 由来の系統にある変異と、その下流にある国内検出変異を色分けする。(c)b と同様に米国由来の変異とその下流の国内検出変異を色分けする。なお、いずれもイギリス株を対象としている。

U は、a,b,c 用のデータを GISAID からサンプリングした (5-1q)。筆者はこれらのデータを可視化した (5-2q)。この時、検疫検出変異の表示位置は、VisE で開発した検疫を壁のように見立てる手法 (5-7e) を適用している。また、分岐の表示位置については、VisE で開発した流入元の視認性を高める手法 (5-10e) を適用している。なお、地図上で日本を取り囲むように描かれる円周の形状は検疫を意味するものである。

第 21 レビュー (5-3q) で、この試作について、肯定や否定の意見は出なかった。I の元で番組用映像を制作する Y は、番組で使用する際の色の統一を考慮し検疫検出変異の色をオレンジ、国内検出変異の色を国外と同じ赤で表現するよう依頼した。

b,c について、筆者は、検疫検出よりも上流にある分岐のうち国内分岐ではないにもかかわらず国内の位置で分岐しているシステムを手動で調整して精緻化した (5-4q)。これについて I は「かなり分かりやすくなった」とコメントした。

カテゴリー	6/15	6/19-20	6/20	6/21
取得				
解析				
フィルタリング	(5-1q)3パターンデータをサンプリング			
マイニング				
表現		(5-2q)全体拡散状況とUAEからの流入、USAからの流入経路を表現		
探索用精緻				
探索用インタ				
レビュー			(5-3q)全体拡散状況とUAEからの流入、USAからの流入経路をレビュー	
提示用精緻				(5-4q)検疫検出より上流のエッジの位置を調整
提示用インタ				



(5-1q)提示用精緻後 (上) b.UAEからの流入経路、(下) c.米国からの流入経路

第22レビュー (5-3q)
 (上) a.世界に拡散するイギリス株 (中) b.UAEからの流入経路 (下) c.米国からの流入経路

図 4.26 フェーズ 5 (VisQ) のタスク遷移図

表 4.9: フェーズ 5 (VisQ) のタスク一覧

ID	カテゴリー	時期	タイトル
5-1q	フィルタリング	6/15	3 パターンのデータをサンプリング U がイギリス株について検疫検出変異のサンプリングデータを作成した。a. 全世界での蔓延、b.UAE 由来の検疫検出変異、c.USA 由来の検疫検出変異
5-2q	表現	6/19-20	全体拡散状況と UAE からの流入、USA からの流入経路を表現 拡散状況を表現、UAE 由来の国内流入と USA 由来の国内流入を表現
5-3q	レビュー	6/20	全体拡散状況と UAE からの流入、USA からの流入経路をレビュー a,b,c を表現、国外の系統と検疫検出、国内検出を着色し、レビュー / Y コメント：(b,c について) 検疫の色合いについて、日本検疫 (水色) →オレンジ、日本 (オレンジ) →赤 (アルファ株と同色) に変更して欲しい。
5-4q	提示用精緻	6/21	検疫検出より上流のエッジの位置を調整 番組内の別の CG で使用する配色を考慮して、国内流入前と後の色を変更。さらに、筆者は b,c の流入状況を視認しやすくするため、検疫検出変異より上流にあるエッジを手動で検疫の位置に調整した。筆者コメント：「検疫検出後のブランチを国内から検疫側に寄せたのでだいぶ見やすくなったと思う。それに伴い、国内から検疫に出ているように見えるストレインもなくなっている」。I コメント：かなり分かりやすくなった。

VisQ の公開までのプロセス

筆者は VisQ の最終的な CG 用データを O らに提供した、I らは、このデータのうち b (UAE からの流入) を使用した CG を用いてイギリス株が検疫をすり抜けて国内に感染者を広げた可能性を伝える内容約 4 分半の映像を制作した。映像には、検疫で陽性と判定されたウイルスとほとんど同じ遺伝子を持つウイルスが 150 件以上国内で確認されたことを伝える字幕とナレーションが付された。また、映像の最後に、S へのインタビュー映像を組み込んだ。S は「検疫で陽性となった方と同行や接触された方が国内で発症し感染源となった可能性がある」と述べ、「入国後に十分な観察期間をおくことで発症を未然に防ぐことが引き続き重要である」との見解を述べた。

VisQ の公開されたビジュアライゼーション

図 (4.27) は、イギリス株が検疫をすり抜けて国内に感染者を広げた可能性を伝える内容約 4 分半の映像である。

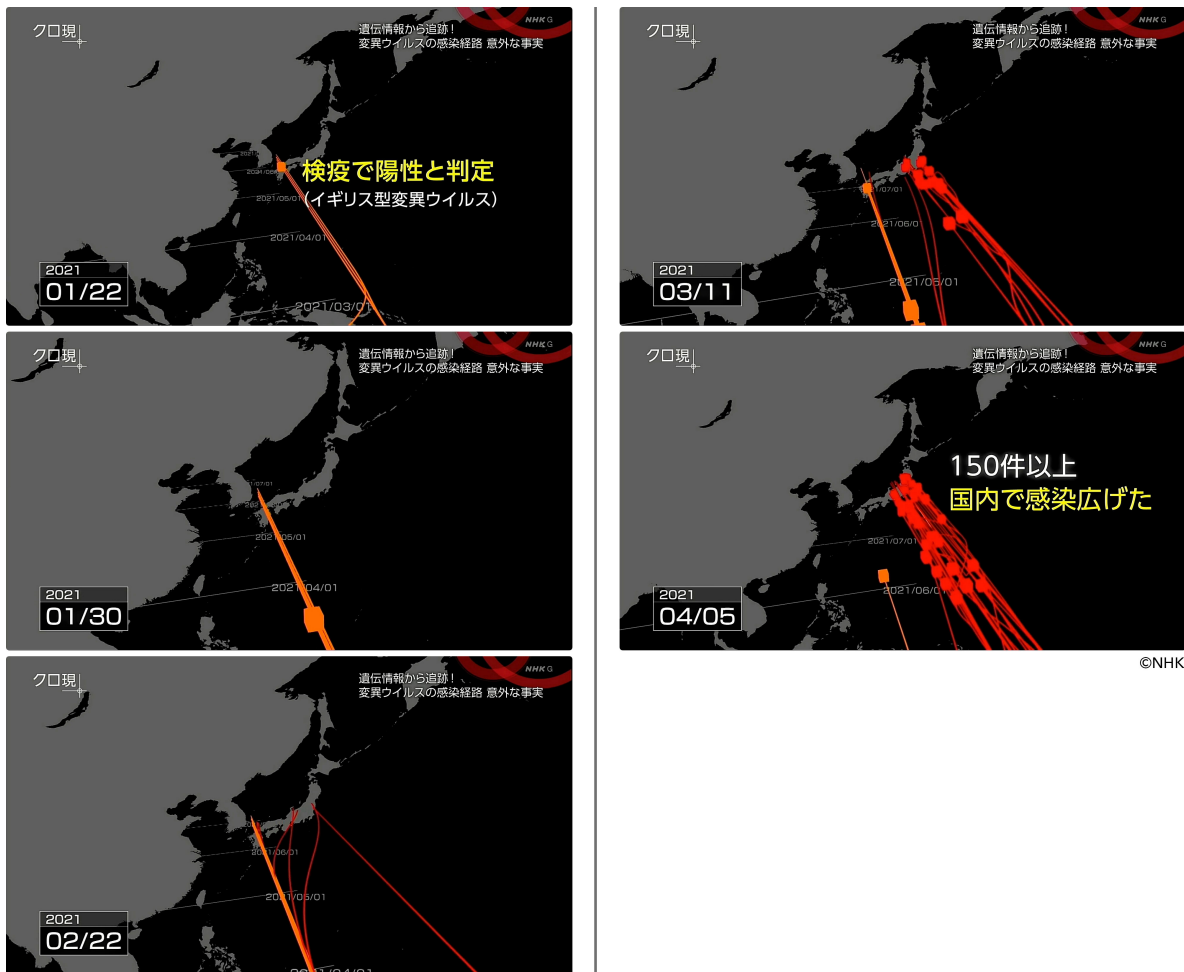


図 4.27 フェーズ 5 の放送時の映像 (VisQ、UAE からの渡航者検疫ウイルスと同じ系統の国内ウイルスが検出される様子)

VisQ のタスクのネットワーク

VisQ のタスクネットワークを図 4.28 に示す。

VisQ では、開始時における関心を満たす表現ができたことが 1 度目のレビューで確認された。そのためタスク間の反復は起らずに終了となった。VisE で開発した探索用精緻 (5-7e、5-10e) が表現に適用されている。

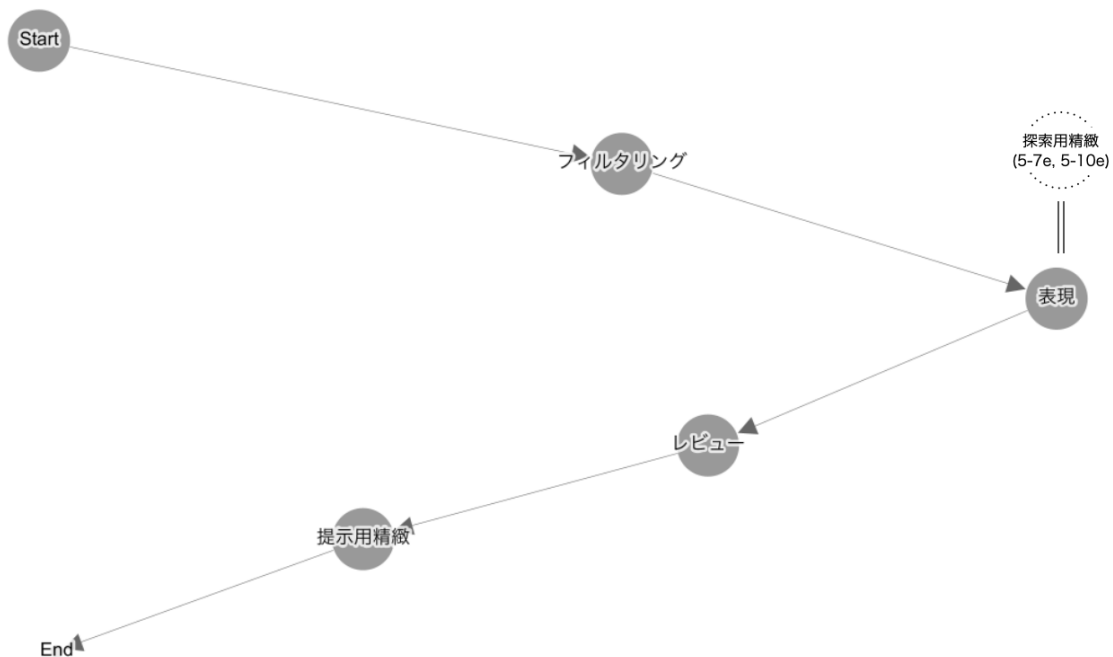


図 4.28 フェーズ 5 (VisQ) のタスクネットワーク図

4.5 考察

4.5.1 各カテゴリーの相互関係に見られる特徴と役割

前節までの分析をもとに、各カテゴリーの相互関係と役割を考察する。

「取得」「解析」はプロジェクトの初期フェーズに行われ、デザイナーにとって未知のデータの構造を読み解き、ビジュアライズの初期プロトタイプを作る足がかりとなっている。

「フィルタリング」「マイニング」「表現」「レビュー」を繰り返す循環が全てのフェーズで見られる。「レビュー」は全てのカテゴリーの中で最も遷移先の種類が多く、レビューからの遷移先の上位は「フィルタリング」「マイニング」「取得」「提示用精緻」となった。「レビュー」ではメンバーらによって可視化から読み取れる意味や提供する価値が吟味される。その結果から、新たな問いや課題が見出され、可視化をアップデートしていく過程がこの循環を形成している。データから別の変数を可視化する場合は「フィルタリング」へ、表現方法を変更する場合は「マイニング」へ、別の構造のデータを取得する場合は「取得」へと遷移している。

「レビュー」の結果、鑑賞者へ提示する価値があると判断された可視化は、最終的な成果物へと組み込まれる「候補」となっている。候補となった可視化は、公開されるメディアの特性に応じて提示方法が検討され、「提示用精緻」「提示用インタラクション」を経て鑑賞者に適した形で公開される。その際、複数の候補が統合され一つのメディアに集約される場合もある。

4.5.2 視覚造形の試行錯誤と探索用インタラクション

レビューは専門家とデザイナーが対話を進展させ、可視化の科学的な解釈を導き出していくプロセスである。しかしそれだけでなく、最終的に作品を見る一般鑑賞者によって知覚される視覚的な造形をデザイナーが模索するプロセスでもある。前者だけに力点を置けば、可視化は単に専門家の分析を明快に説明するものになろう。ここで求められるのは学術論文に掲載されるような専門的で明晰な可視化ではなく、専門知識や関心を持たない鑑賞者であってもその可視化が意味するところを解釈しようと前向きに意識を向けるように働きかける魅力とインパクトを持った可視化である。そのために専門家は科学的知見を、デザイナーは情報可視化の原則と視覚造形の知識を動員して、論理性と感性がせめぎ合う造形的試行を繰り返す。

このような議論の場面では、探索用インタラクションを利用して数値と視覚属性をパラメトリックに変更しながら、言語化が難しい造形の印象変化を観察し議論を進めている。パラメトリックに変更する要素はスライダーなどのインターフェイスとして画面上に実装され、レビューを重ねるごとに追加され多様なコントロールが可能となっていた。図 4.29 は、最終的に画面上に実装されたすべてのインターフェイスである。上部右側には変異属性に応じたカラースキームを変更するドロップダウン、変異の検出地域を指定してフィルタリングするドロップダウンが並び、最右部は、エッジのアルファを変更するスライダー、フォーカスしたエッジをハイライトするスライダー、カーブの形状を計算するアルゴリズムを切り替えるスイッチ類が並び、即座に造形や配色を切り替えられる。また上部左側には、カメラアングルやトランジションを設定しアニメーションを作成するツールが集約されており、動きを伴った可視化の印象を検討できる。検討したパラメータはプリセットとして保存する機能も有している。

このような探索用インタラクションを実装することで、図 4.30 のように同じデータからいくつもの視覚的なバリエーションをその場で作成でき、議論を前進させることで理想的な視覚造形を探ることができる。

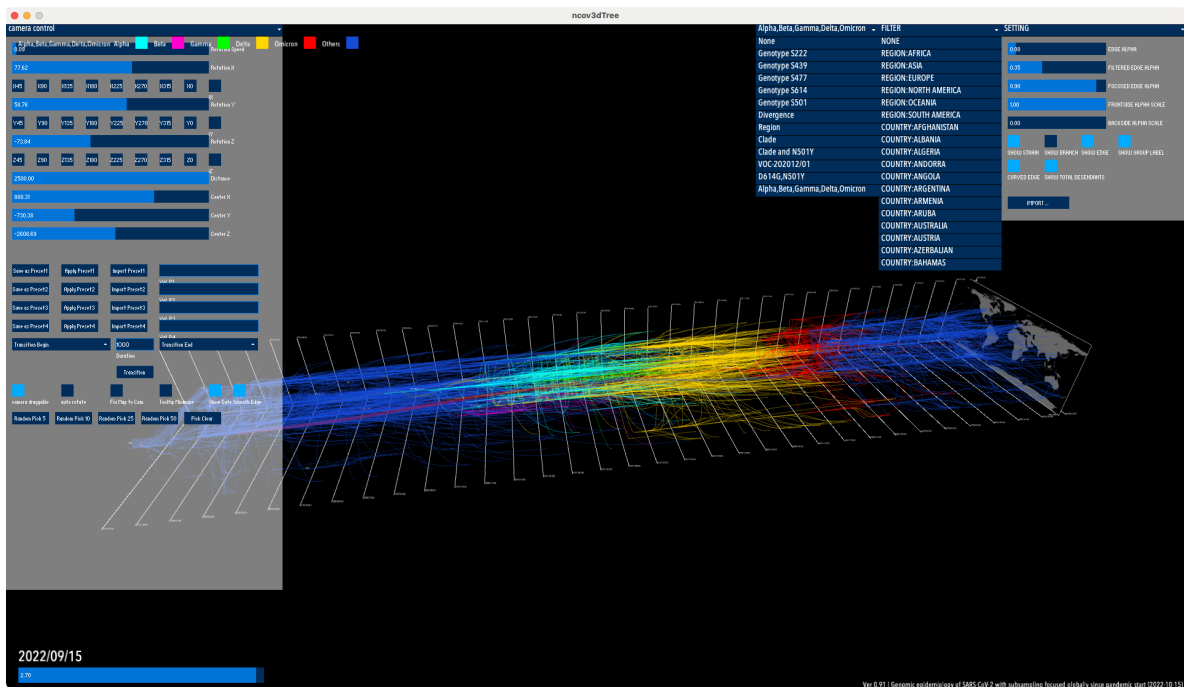


図 4.29 最終的に実装された、すべての探索用インターフェイス

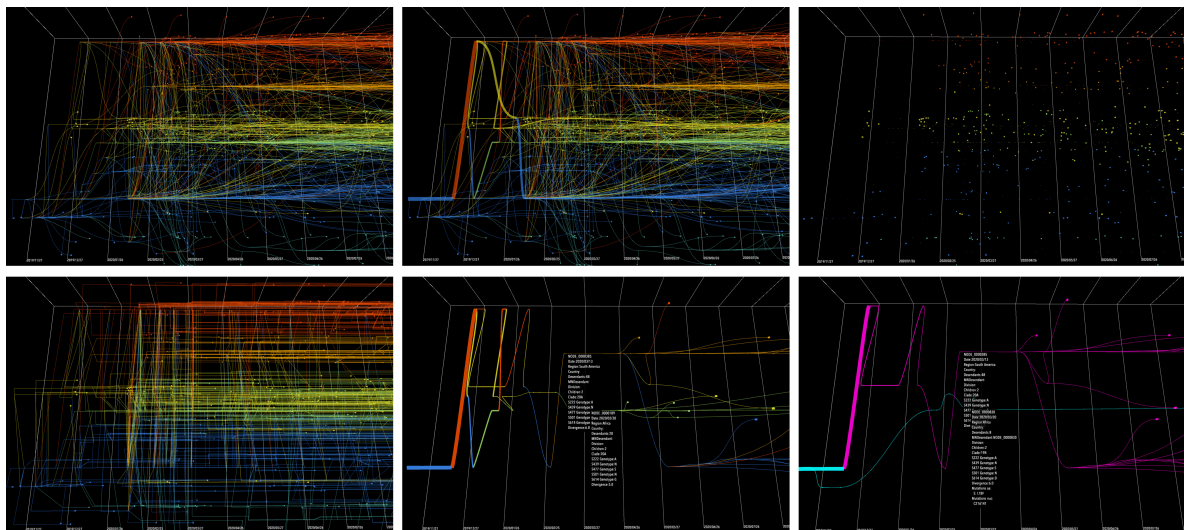


図 4.30 探索用インタラクションにより、同じデータから豊富なバリエーションをデザインできる

4.5.3 視覚造形の探索的な進展例

前項では可視化の造形が探索的に進展していく様子を述べた。本項ではその具体的な場面を示す。

時間と空間

図 4.31 の左はプロジェクトの初回打ち合わせで筆者が提示したアイデアスケッチである。右はその直後に開発したプロトタイプである。スケッチに従い、時間軸を画面奥の方向に取り、世界地図を複数並べることで時間の経過と位置情報を同時に表そうと試みている。

図 4.32 では、日本で検出された変異とつながる系統だけを目立たせるために赤に着色しその他の系統を薄く表示する機能を実装し、キーボードの操作で最初の状態と表示を切り替えて比較し易くしている。

第 1 回目のレビュー (フェーズ 1、1-7) でプロトタイプを眺めて専門家らと議論したところ、スケッチに非常に忠実で経路の概要は把握できるものの、並べた地図が系統樹を遮蔽し、また、地図と系統樹の交点を読み取りづらいために、詳細な動向を把握するには不向きであるとの意見が出た。筆者はこのレビュー後に、時間軸に沿って並行可動する地図を実装し、スライダーで任意の時間を指定できるインターフェイスを用意した (図 4.33)。これにより、検出された変異の時間と位置を 1 日単位で詳しく観察できるようになった。

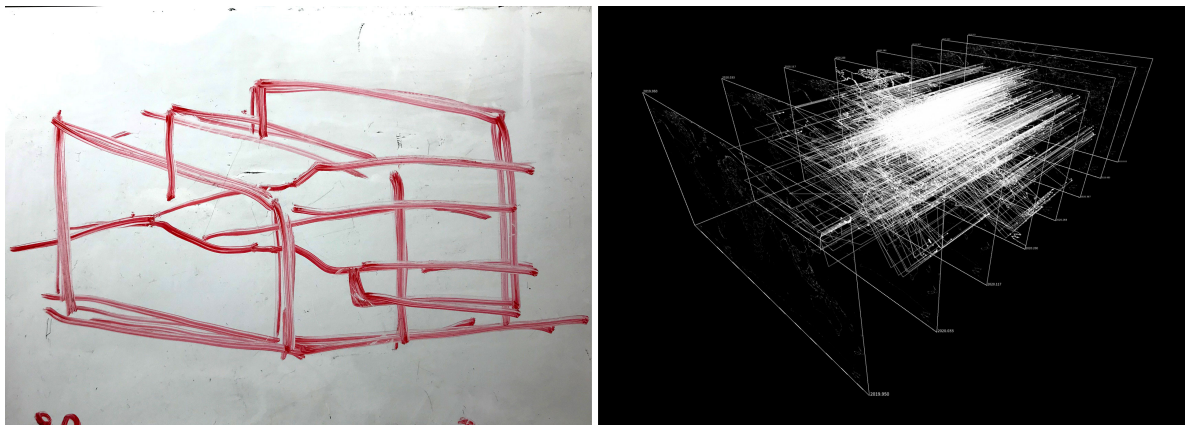


図 4.31 最初期の時間と空間の表現

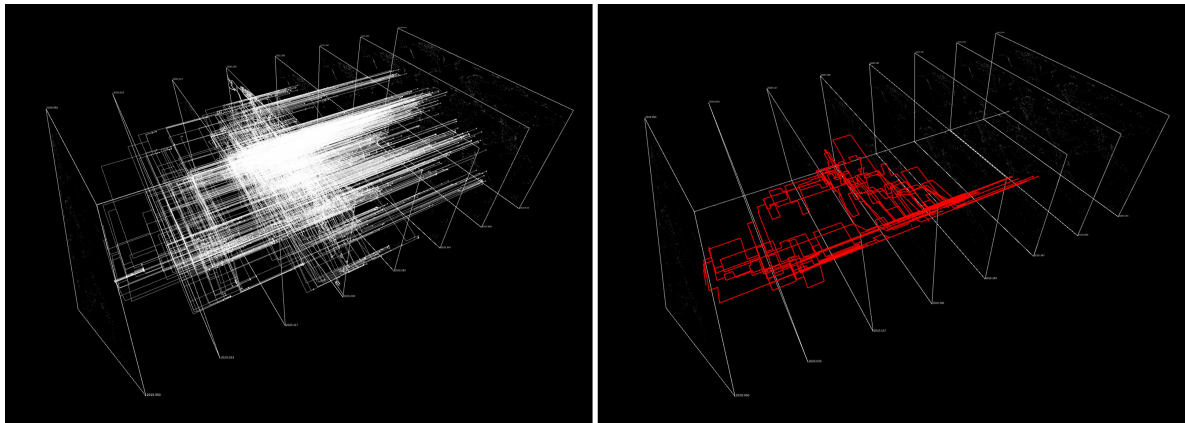


図 4.32 日本へ到達する系統だけの表示に切り替え比較する



図 4.33 スライダーの操作に応じて時間軸を移動する地図画面

あえて表示を残す

第2回のレビュー（フェーズ1、1-12）では議論を経て、ヨーロッパに入った経路や特定の採取地点の国と共通する先祖（ルーツ）について探索を試みるようになった。そこで筆者は、画面上で国名を選択するドロップダウン型のインターフェイスを実装した。国を選択することで、その国で検出された変異につながる系統だけを排他的に表示する機能である。

ここで筆者はもう一つのインターフェイスを用意した（図4.34）。それは表示から排除された系統の透明度を可変するスライダーである。表示から排除された系統はここでは関心の対象外であるから、分析的な観点では非表示にすれば事足りる。しかし、完全に非表示にすれば全世界への拡散という現象の全体像を失い視覚的な印象は大きく変わる。そこでノイズにならない程度の半透明な状態で残すことを想定して、その適度な具合を議論すべくスライダーを実装し、第3レビューの場面で議論した。議論を受けて調整した透明

度は最終的な番組用の CG にも反映された (図 4.35)。

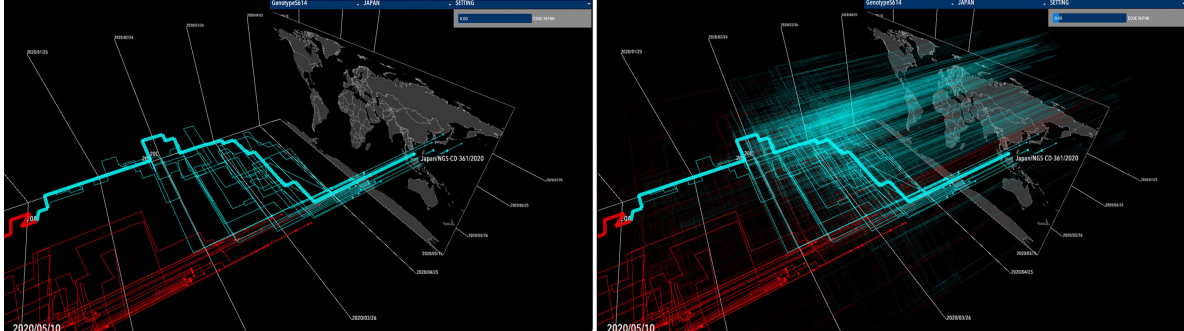


図 4.34 排除された系統の透明度を調整するスライダー



図 4.35 放送時の映像に反映された表現

形のアルゴリズム

第 3 回レビュー (フェーズ 1、1-17) では、系統が分岐する位置の表現が議論となった。一般的に系統学で用いられる系統樹は空間の次元を持たない。今回使用したデータも離れた地点 A と地点 B で検出された 2 つの変異がどこで分岐したかを特定する座標情報は持っていない。そのため、単純に A と B の中点の座標に分岐点を表現していたが、この

手法では大量に分岐点が現れ、系統を示す線が複雑に交錯してしまう。また、海上や陸上の境界で分岐が生じるなど意味の上でも不自然なものとなっていた。そこで、筆者はエッジバンドリング [69] と呼ばれるネットワーク図表現における視覚的な整理手法を応用して複雑さの緩和を試みた。しかし、ここでも分岐位置をどこに置くかで印象が大きく変わる。筆者と専門家は試作を行い、変異数 (divergence) が同じ分岐をグループ化し、グループの中で最も古い時間に変異が検出された地点の座標を分岐位置とみなして表現した (図 4.36)。この手法では、分岐位置と変異の検出位置が一致する*6。そのため視覚的に整理され意味としても不自然さが少ない表現に至ることができた (図 4.37)。

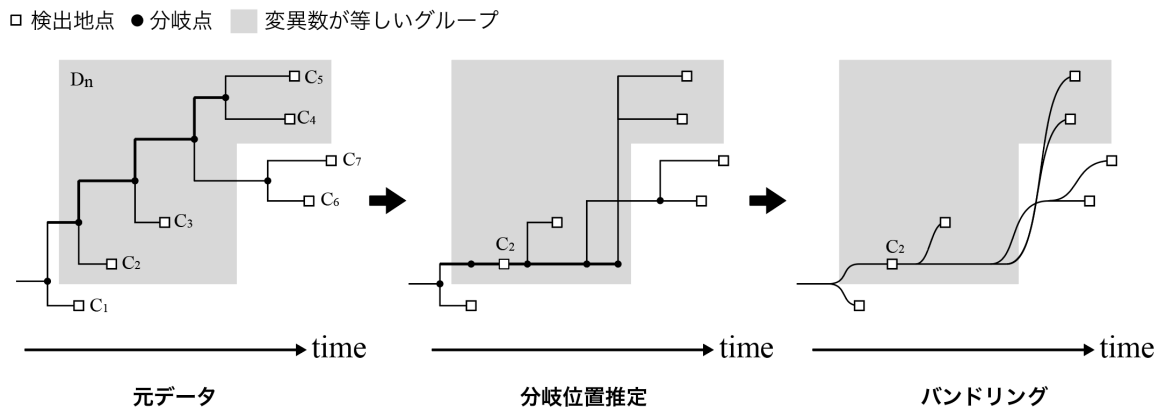


図 4.36 開発したアルゴリズムの概略

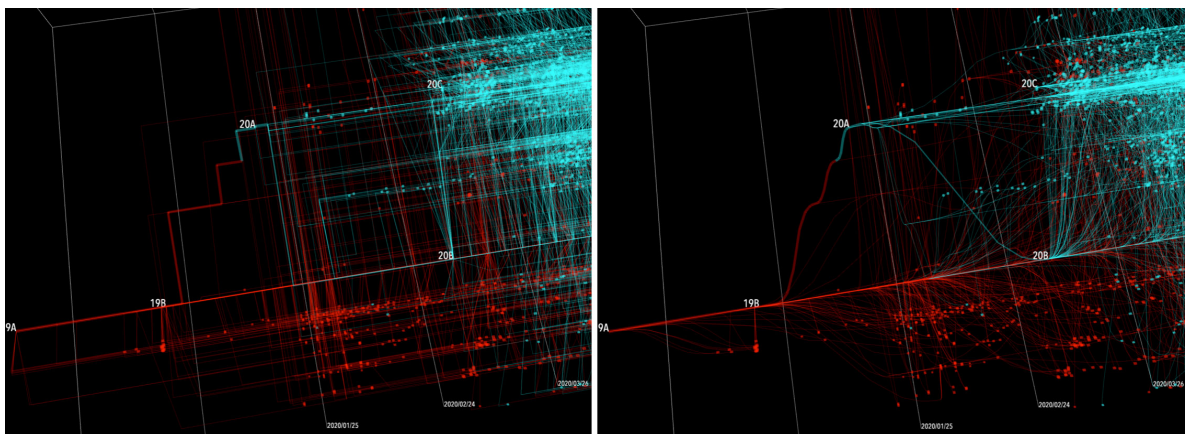


図 4.37 バンドリングの効果

*6 この可視化手法は、VIZBI 2021 で報告し採録された [65]。 <https://vizbi.org/Posters/2021/vB06>

4.5.4 プロセスのモデル化

以上の考察を踏まえ、プロセスのモデル化を試みる。考察からプロセスは次の3つのパートによって構成されていると考えられる。

パート1：情報収集の段階

「取得」「解析」によって、未知のデータの構造を深く理解し、探求への足がかりを築く

パート2：可視化して意味を読み解き探求する段階

「フィルタリング」「マイニング」「表現」「探索用インタラクション」「探索用精緻」「レビュー」の循環によって、多様な視点で可視化し探求する

パート3：可視化をメディアとして統合する段階

探求で得られた意味と、公開するメディア形式を考慮し「提示用精緻」「提示用インタラクション」を通じて作品として組立てる

各パートの名称と「DEXAS」について

各ステップの意味を考慮して、パート1を「Diving」、パート2を「Exploring」、パート3を「Assembling」と称する。またこの3パートからなるプロセス全体のモデルについては、その頭文字をとって「DEXAS」と称する（図4.38）。各パートにおけるタスクをDEXASモデルに従って整理すると図4.39となる。



図 4.38 DEXAS の 3 パート

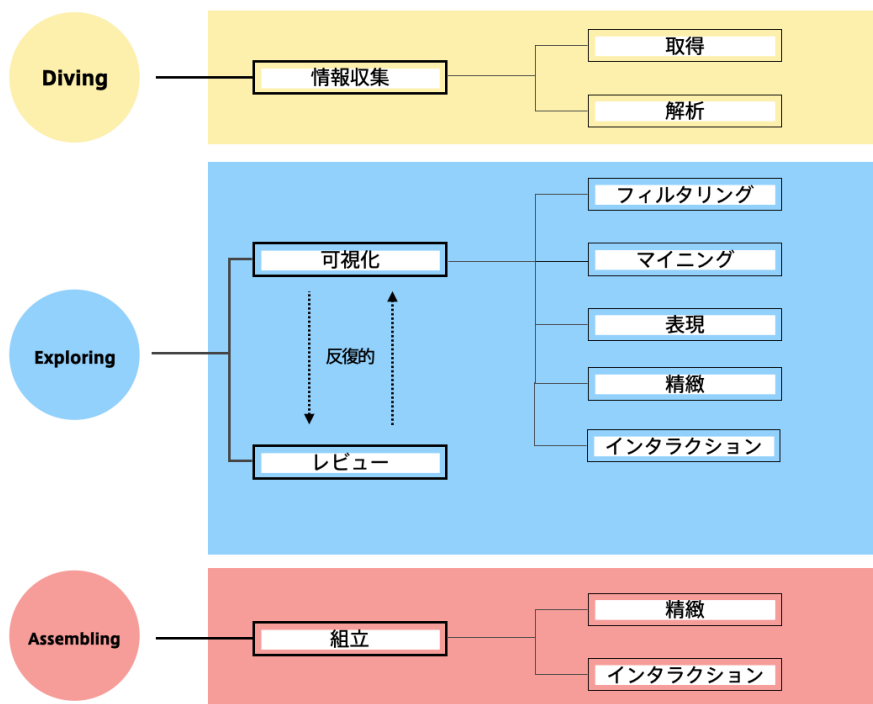


図 4.39 DEXAS における各パートとタスク

次に、DEXAS の全体像を各パートの関係と順序に従って整理すると図 4.40 となる。Diving を終えた後、Exploring は循環的に何度も試行される発散的なプロセスとなり、

Exploring を終えた後、Assembling は多様な可視化の意味をひとつのメディアに統合していく収束的なプロセスとなる。

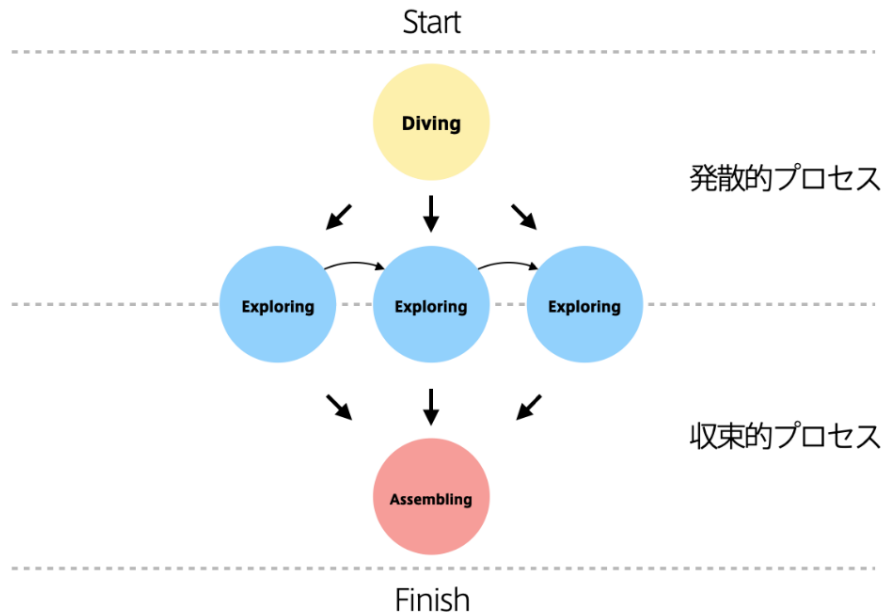


図 4.40 DEXAS における各パートの進行

4.5.5 モデルによる事例の分析

DEXAS モデルからプロジェクトを分析し検討を加える。ここではフェーズ 1 を取り上げる。

図 4.41 は、フェーズ 1 のタスク遷移図（図 4.4）を DEXAS の観点から分析した図である。フェーズ 1 では Diving の後、4 回の試作とレビュー（タスク ID：1-7,1-12,1-17,1-21）を行なっている。このひとつひとつのレビューが Exploring の成果に対応しており、開発された可視化表現とそこから読み取ることができる意味のペアがその Exploring の成果になっている。それぞれの成果を考慮しながら番組として伝えるメッセージを吟味したうえで精緻化を行い放送映像として Assembling を行なっている。

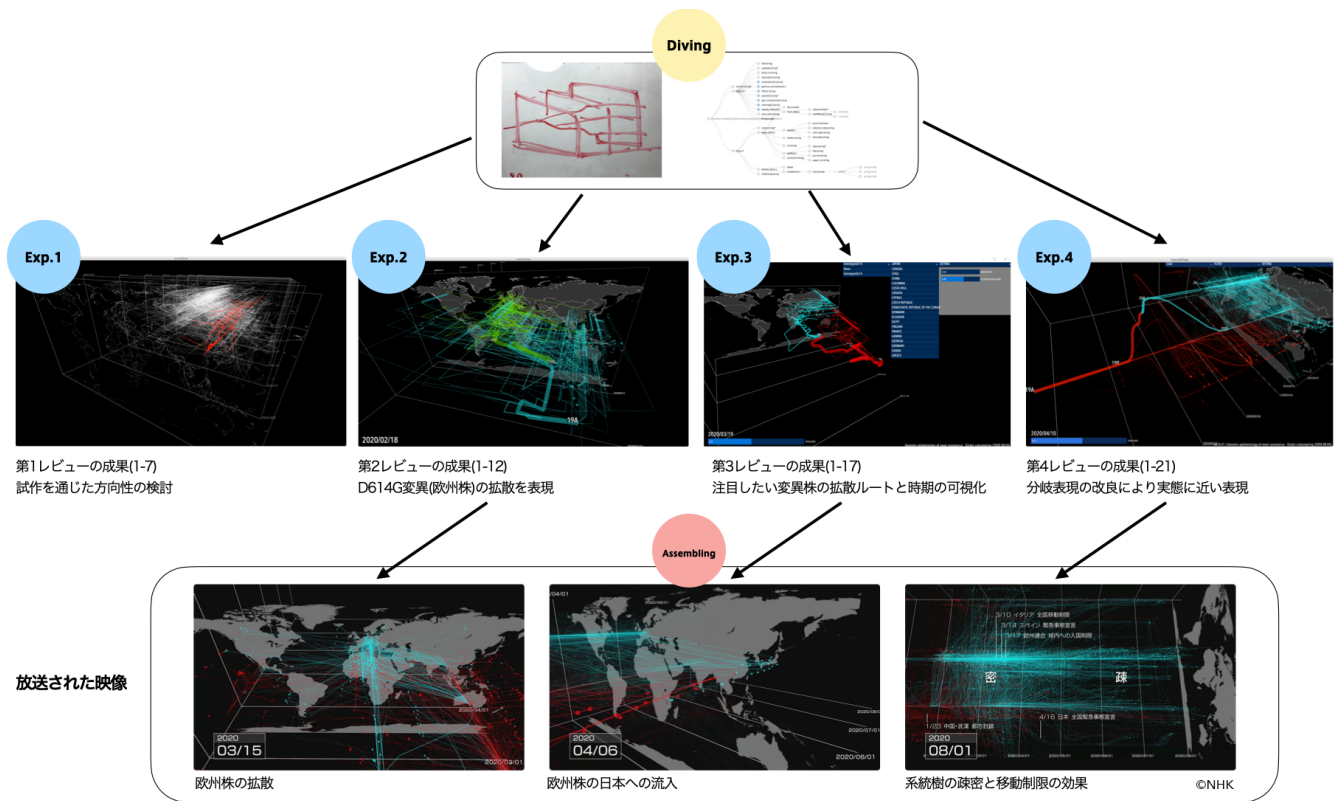


図 4.41 フェーズ 1 のプロセスを DEXAS で整理した

同様に、フェーズ 1 からフェーズ 5 までを DEXAS モデルで示したものが図 4.42 である。

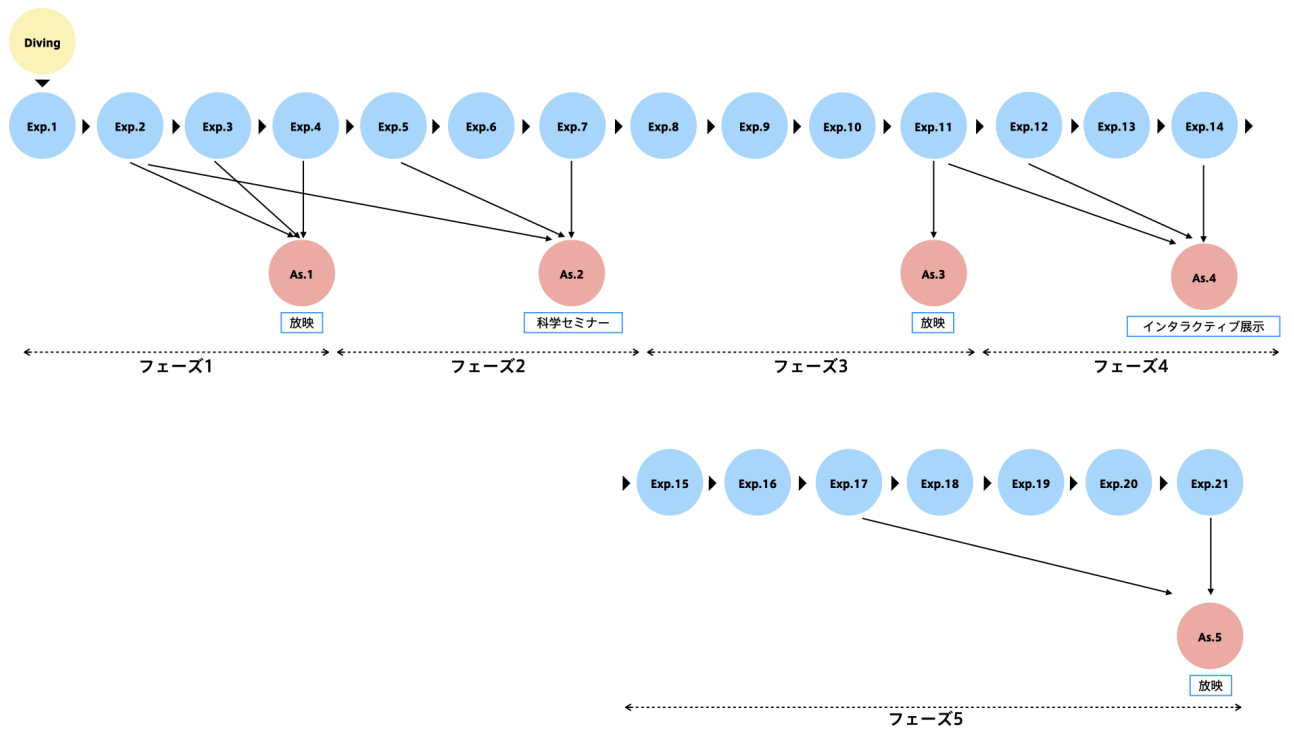


図 4.42 プロジェクト全体の DEXAS モデル

4.6 まとめ

大規模なデータを対象としたデータビジュアライゼーションプロジェクトである「新型コロナウイルスゲノム系統樹の3次元可視化」を対象として、デザイナー・エンジニアとした関わった筆者の観察に基づき CID の観点からそのプロセスを記録した。プロジェクトを5つのフェーズに分け、その各々で CID が提唱する7つのステップがどのように結びついているかを検討した。その結果から各ステップは「情報収集の段階」「可視化して意味を読み解き探求する段階」「可視化をメディアとして統合する段階」の3つにグルーピングできることを示した。それらをそれぞれ「Diving」「Exploring」「Assembling」と名づけ、これらを反復しながらデータの多様な意味を可視化しメディアとして組み立てるプロセス「DEXAS モデル」を提案した。

第5章

パターン・ランゲージの作成

5.1 パターン・ランゲージについて

本章では、プロジェクトの観察から得られたデータビジュアライゼーションの経験則を形式知として記述することを試みる。まず、形式知の記述形式としてパターン・ランゲージを選択した理由を述べる。

パターン・ランゲージ (pattern language) の起源は都市や建築に生き生きとしたよい質をもたらす設計の共通パターンを記述法として建築家 C.Alexander が提唱したものである。現在、その活用範囲は拡張され人工物の質に対してだけでなく、研究、学習、企画などさまざまな人間の活動の質に対して活用が進んでいる。井庭によればパターン・ランゲージは、ある領域において良い質を生む設計 (デザイン)・やり方に潜む型 (パターン) を抽出して概念・言葉にし、体系化して、共通言語 (ランゲージ) として用いることができるようにまとめたものである。[70, 71, 72]

ここで言うパターンとは、実践上でのコツや経験則であるが、特定の事例における個別具体的なものに着目するのではなく、いくつかの事例において繰り返し見られる共通点に着目して抽出されるものである。

また、パターンは具体的な指示ではなく、かといって抽象的な理念でもない、その中間的な抽象度で記述される。これにより読者が具体的な実践のイメージを掴みやすく、またそれぞれの現場の実情に合わせ柔軟な実現の仕方を想像することを支援できる。

パターン・ランゲージは、数十からなるパターン同士を相互に関係させ、全体として一つの体系を成す (図 5.1)。パターンは階層的に組織され、大きなパターンの実現を支える下層のパターンがあり、これらを実践することで、全体の質を高めることができる。例えば、大きな問題に対処する際、それを構成する小さな問題について読者が把握しその対処

を支援することができる。

こうした構造から、パターン・ランゲージによる形式知化は読者に以下の効果をもたらすことが期待できる。

1. パターン・ランゲージを解釈し、自己の状況に当てはめて具体的な行動を考えられる
2. パターン・ランゲージを媒介にして、他者と経験を共有したり学び合うことができる
3. パターン化された概念をもとに問題の本質を見極め、不足を補い、改善につなげることができる

これらは本研究の目的に合致しており、形式知の記述形式として選択した。

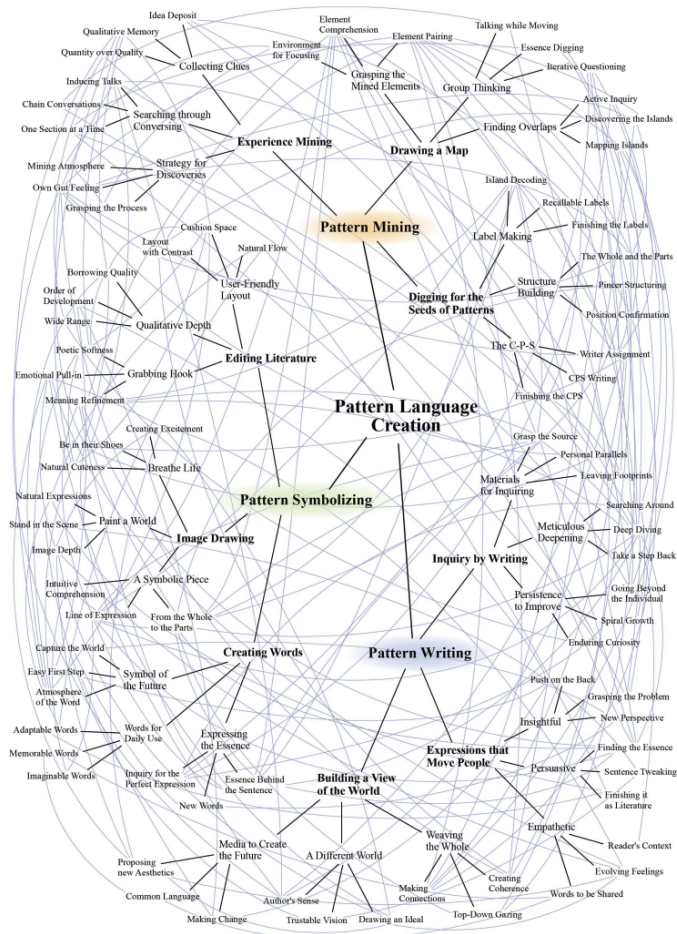


図 5.1 パターン同士の意味的構造、Takashi Iba and Taichi Isaku. A pattern language for creating pattern languages: 364 patterns for pattern mining, writing, and symbolizing.[73] fig.7 より引用

5.2 パターン・ランゲージの制作プロセス

パターン・ランゲージの作成プロセスは井庭らの参考文献 [73, 74, 75] に詳しい。ここではまずそのプロセスについて概説し、次節で本研究でのプロセスを述べる。

一般的なパターン・ランゲージの作成プロセスは「パターン・マイニング」「パターン・ライティング」「パターン・シンボライジング」の3フェーズからなる。

「パターン・マイニング」は、実践知を持つ経験者の談話や記録などを分析してパターンの元になる情報を収集するフェーズである。パターン・マイニングは「マイニング・ダイアログ」、「クラスタリング」、「シード・メイキング」3つのステップで行う。マイニ

ング・ダイアログではパターンを制作したい対象の活動を熟知した経験者から、経験談や実践知を収集し情報を網羅的に書き出す。その際個別の経験がどのような状況で起こりうる、どのような問題に対して、どうやって解決へと導いたのかを掘り下げる。

次に集めた情報をもとにクラスタリングを行う。クラスタリングは膨大に集められた情報をグループ化していく作業である。この際、先入観や既存の概念構造にとらわれたり、上位に位置する大きな類似性から個別な類似性へとトップダウン的に分類するのではなく、情報のひとつひとつを一つ一つで比較してその意味の類似性に注目して検討を行う作業を繰り返し、ボトムアップ的にグループを構成することが肝要となる。こうすることで一つのグループに集められた情報には何らかの共通のパターンがあるとみなすことができ、以降このグループを対象にそれが何を意味しているのかを深く掘りさげパターンを特定する作業へと接続していく。具体的な手法としては、個別の情報を付箋などに要約し KJ 法で行う手法が推奨されている（KJ 法ではグループを「島」と表現する）。

パターン・マイニングの最後のステップは、シード・メイキングである。シード・メイキングはクラスタリングで得られた各グループごとに、内容を検討しながら「状況 (Context)」「問題 (Problem)」「解決 (Solution)」(以下、CPS) を記述する。これらは「パターンの種」と呼ばれる。パターンの種が記述できたところで、パターンの構造化を行う。パターン・ランゲージは独立した個々の問題を解決する知識の集積ではなく、パターン同士が連携して全体的な質の向上を目指すものであるため、質を生み出すようなパターン同士の関係性や共通性に注目して構造化することが肝要である。また、構造の中でのパターンそれぞれの位置付けが読者にうまく伝わるように吟味しなければならない。構造化の作業を通じて、全体に抜け漏れがないか、ひとつひとつのパターンの粒度は揃っているか、活用の際にニュアンスが伝わるかといった視点で点検する必要も生じる。構造化の作業はクラスタリングと同様に、まずはボトムアップで進め、構造の全体像が見えてきたところで、前述の視点に留意しつつトップダウンで構造をつくっていくことが推奨される。また、次のパターン・ライティングのプロセスの間にもパターンへの理解が深まるため、構造の中でのパターンの位置付けが修正され再構造化される場合もある。

「パターン・ライティング」では、パターンの種で記述した CPS をもとに、パターンを深く考察して記述を整えていく。この作業における重要な視点は、パターンが単なる解決のマニュアルではなく、ひとつの顕在化した問題は、諸要因である「フォース」の不協和状態「コンフリクト」であることを読み手に認識させ、その不協和の解消に向けて主体的なアクションへと導くことである。そのため、目指すべき理想の記述は、個別のエピソードのような具体性と教訓のような抽象性の中間、なおかつ、特別な状況下でしか起こり得ない特殊性と、何にでも当てはまる一般性の中間である。井庭はこれを「中空の言

葉」[76]と表現する。ここではCPSの記述を洗練していくことに加え、やや具体的な指針を示す「アクション(Action)」や、課題が解消した後の状況を想起させるための「結果(Consequence)」の記述を加えることもある。それらすべての作業は、読み手の認識を考慮に入れながら何度もリライトを重ねる。

「パターン・シンボライジング」は、パターンライティングで記述したCPSをさらに洗練させつつ、それらを象徴的に表す要素を加えるフェーズである。パターンを象徴するのは、名前やイラストレーションである。象徴化はパターンを活用する際、それらを指し示す言葉として機能し、パターンを通じて現場で起きていることを認識したり他者と議論するために重要な要素である。最後に、パターン・ランゲージとして活用するためにどのようなメディアに定着させるかを検討し、具体的なデザインの作業を行う。以上がパターン・ランゲージ開発の一般的なプロセスである。

5.3 本研究におけるパターン・ランゲージの制作プロセス

次に、本研究におけるパターン・ランゲージ作成プロセスについて述べる。本研究におけるパターン・ランゲージ開発においても、前節で示したプロセスに沿っているものの、一部で本研究の特性に合わせて変更を行っている。その際、パターン・ランゲージの持つ効果が損なわれないように注意を払った。次節以降で、具体的な変更点とその理由を挙げる。

5.3.1 パターン・マイニングにおけるパターン・ダイアログについて

本来、分析対象とする領域について複数の経験者へのインタビューデータからパターンになりえる経験を抽出し、属人性を排除すべきであるが、本研究では、CIDを熟知して実践した経験者が非常に少ないことから、4章で示したものと同一事例の筆者の経験を抽出した。この時、単なる筆者の回想ではなく4章で示した活動のログデータから、具体的なエピソードが読み取れる部分を抽出した上で、その部分の前後における筆者と他者との対話やプロセスの文脈から、背景にある意識や思考を掘り起こして記述することで手法の妥当性を確保した。またその際、対象としている部分がCIDのどのタスクを行っている状況であるのかによってラベリングした。これを「カテゴリー」と呼ぶことにする。データから実践知の表出と思われる箇所を抽出し、50のカードに記述を行なった(図5.2)。

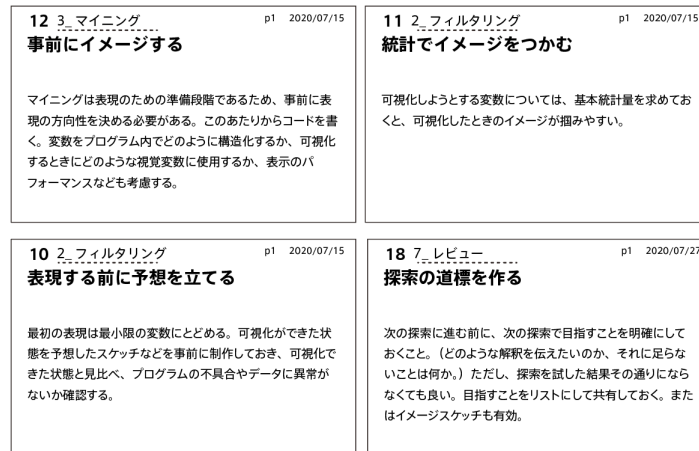


図 5.2 カードの記入例、破線部がカテゴリーを表す

5.3.2 パターン・マイニングにおけるクラスタリングについて

本研究では、ダイアログの記述をクラスタリングした結果、17 のグループに分類できた (図 5.3)。この時、カテゴリーが異なっても同じグループになるものが観察された。ここでは特徴的な 2 例について述べる。なお、カード番号は< >、カテゴリーは [] で表記する。

「チームの立ち上げ」というグループ (図 5.4) には、< 8 >理解のレベルを揃える [解析]、< 40 >達人を模倣する [マイニング]、< 37 >初めは手堅く素早く [表現] というカテゴリーが異なる 3 つのカードが属している。これらに共通するのは、いずれもプロジェクト開始初期に観察されたという点である。これらのカードは、カテゴリーの分類で見れば異なっても、まだプロジェクトが開始されて間も無く、立ち上がったばかりのチームで専門が異なるメンバー同士での認識を合わせたり、関係をスムーズにして方向性を共有するという一貫した意図を持って関連しており、その意識が [解析] → [マイニング] → [表現] と推移しても持続していると判断できる。

「再現性の確保」というグループ (図 5.5) には、< 6 >出所は厳重に管理する [取得]、< 20 >レシピを丁寧に残す [フィルタリング]、< 48 >良さのレシピ [表現] というカテゴリーが異なる 3 つのカードが属している。これらに共通するのは、データビジュアライゼーションプロジェクトに特有の課題、すなわち、データが更新される可能性があること、データが大規模で注目する変数の組み合わせが無数にあること、表現を頻繁に変更し試行錯誤することなどから、得られる結果が移ろいやすいという問題への対処である。これらのカードの共通性からどんなタスクにおいても記録の管理を怠ってしまうと再現でき

なくなってしまうという意識において一貫していると判断できる。



図 5.3 クラスタリング後のグループ

5.3.3 パターン・マイニングにおけるシード・メイキングについて

シードメイキングにおいて、パターンの種を構造化する際、まずは個別のパターン同士の間を重視しボトムアップ的に構造化した。その際、17のグループ（小グループ）はさらに大きな5つのグループ（中グループ）に編成できた（図 5.6）。以下では中グループ名を《 》、小グループ名〈 〉で示す。

中グループは《関係の構築》《表現行為》《表現の解釈》《表現を味わう》《味わいの体験化》である。小グループのうち〈データの取得〉はどの中グループにも属さなかった。ま

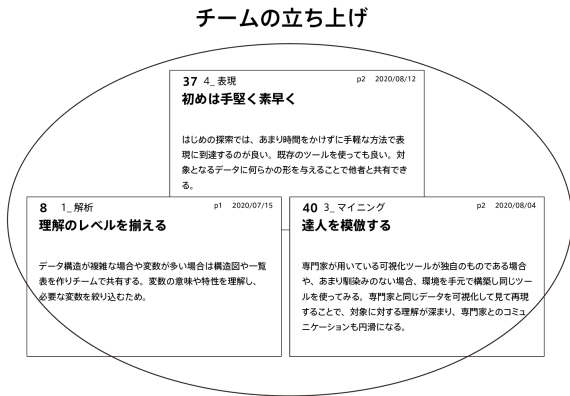


図 5.4 「チームの立ち上げ」グループ

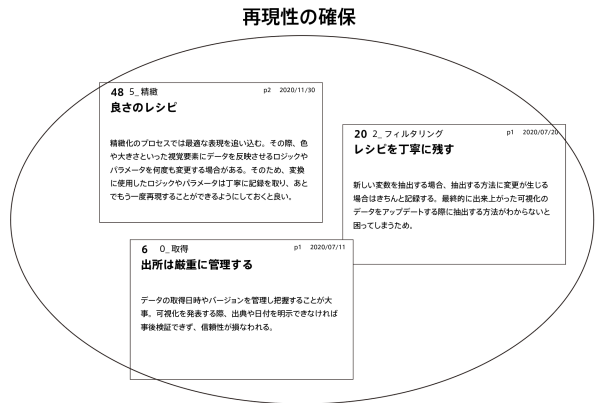


図 5.5 「再現性の確保」グループ

た小グループのうち2つ〈テクニク〉〈生き生き感〉のふたつは、それぞれ《表現行為》と《表現の解釈》、《表現を味わう》と《味わいの体験化》に重複して属した。〈チームの立ち上げ〉と〈裏切られる予想〉は《関係の構築》と《表現行為》に重複して属した。

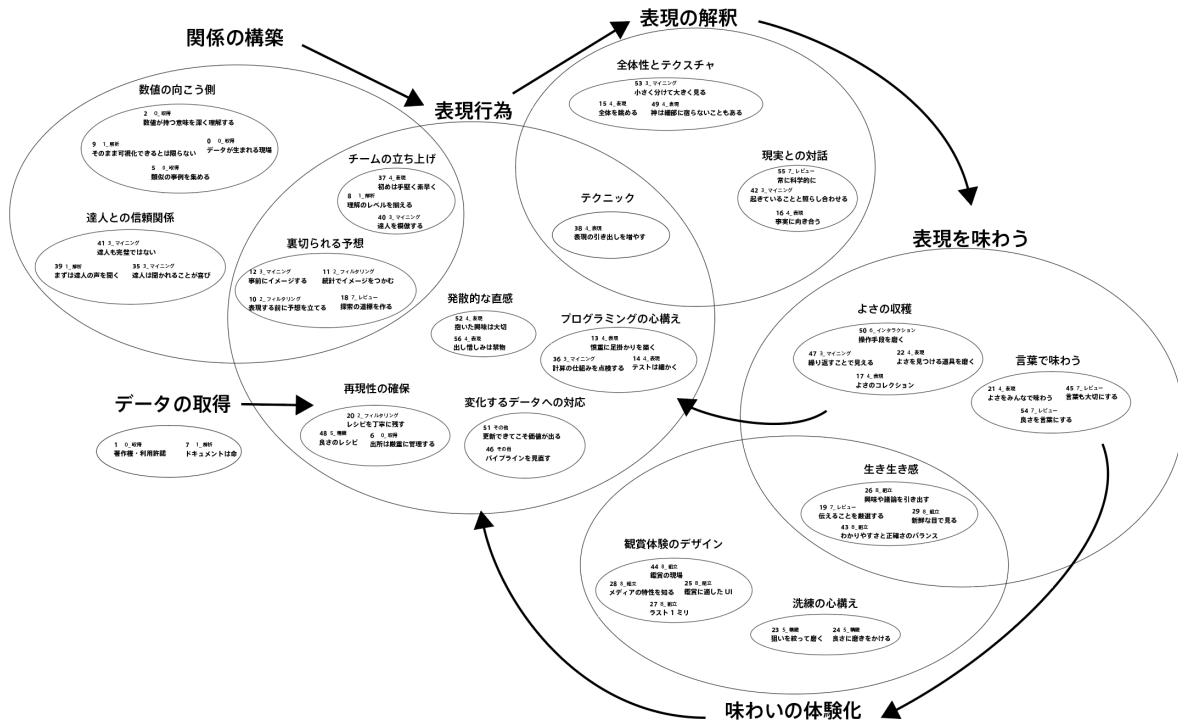


図 5.6 シードメイキングにおけるボトムアップの構造化

次に、トップダウンの構造化にあたっては DEXAS との対応関係から検討した。DEXAS は出来事や作業の順序に注目した分析であるのに対して、パターンは行為と意

味の関係に注目した分析である。そのため DEXAS の各パートの構造とパターン・ランゲージの構造を厳に一致させることは難しく、パートを越境するようなパターンが存在する可能性がある。しかし、そうすることで DEXAS の構造とパターン・ランゲージの構造の一貫性が崩れ、読者の混乱を招く可能性がある。そのため、ここでは両者の構造の意味を尊重し、読者が DEXAS モデルとパターン・ランゲージを同時に学ぶ際の認識のしやすさに配慮して構造化した (図 5.7)。

具体的には、次のようになる。《関係の構築》と〈データの取得〉を Diving パートに、《表現の解釈》のうち表現の工夫に関するものと解釈に関するものに分け、前者と《表現行為》を合わせて Exploring パートの「可視化」に、後者と《表現を味わう》を合わせて Exploring パートの「レビュー」に、《味わいの体験化》を Assembling パートに対応させた。〈生き生き感〉については「組立」で収集した情報が大半を占めることから「組立」フェーズに対応させた。〈チームの立ち上げ〉は《関係の構築》と関連は深いがデータを可視化する行為を伴っていることから Exploring パートの「可視化」に含めた。〈テクニック〉は意味としては《表現行為》に属すると考えられるが、可視化表現の習得のための学習や日頃のトレーニングについて述べているため、プロセスの途中で問題を解消する手立てとはならず、DEXAS の中で位置付けることはできなかった。そのため、パターン・ランゲージとしては扱わないこととした。

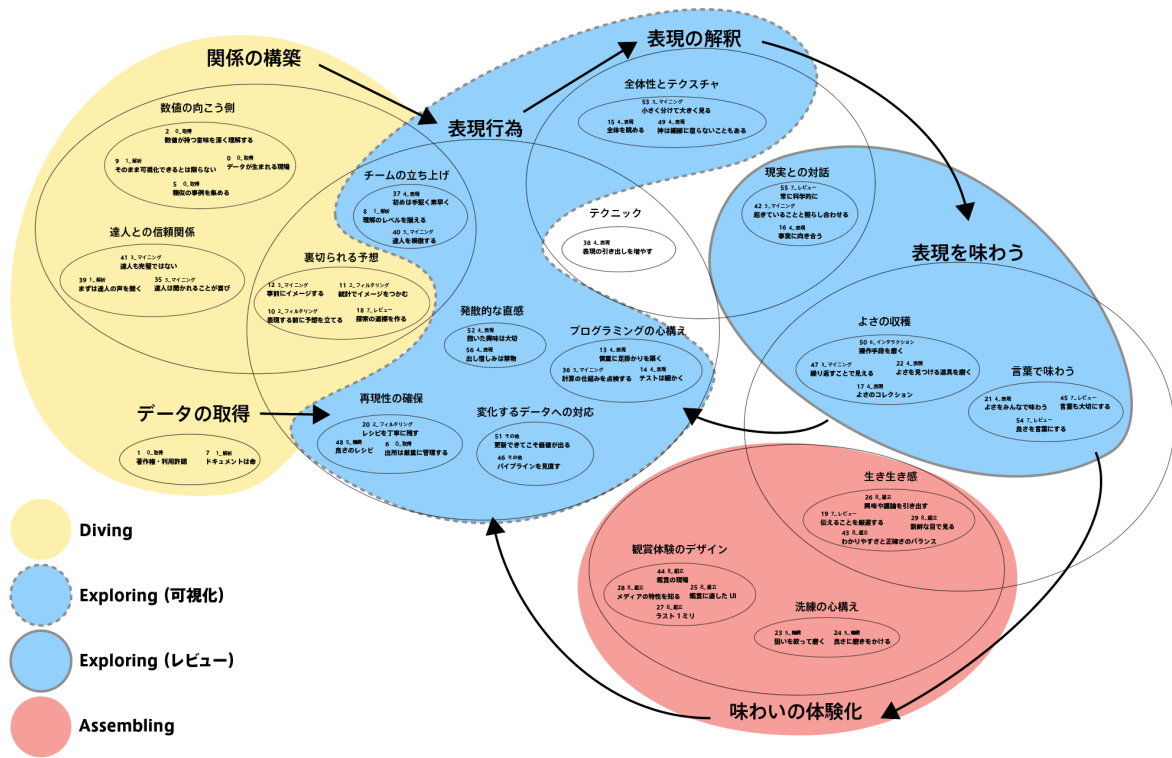


図 5.7 シードメイキングにおけるトップダウンの構造化とボトムアップの構造化の重ね合わせ

5.3.4 パターンの再構造化

前項で示した小グループをシードとして再構造化を行った。その際、前項の図 5.7 で分類したトップダウン構造の 4 つのパート内でパターンの数や粒度が揃うようにシードの組み合わせ方を調整した。次に、4 つのパートそれぞれを代表するパターン (2,9,19,27) と、さらにその 4 つを束ねるパターン (1) を作成した。再構造化されたパターンを図 5.8 に示す。

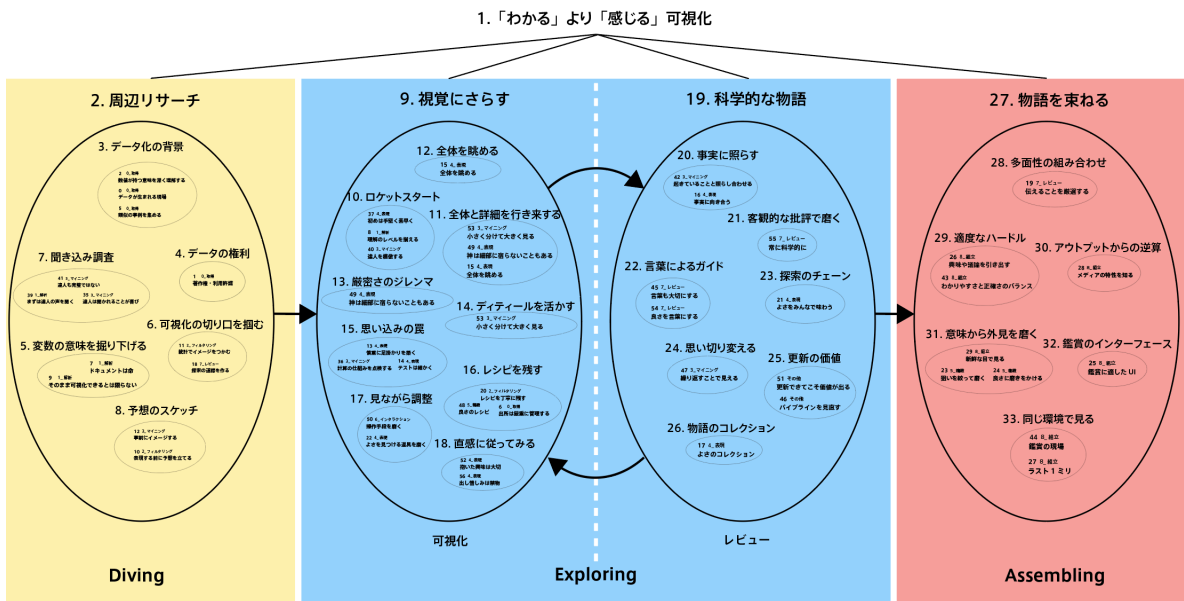


図 5.8 再構造化されたパターン

5.3.5 パターン・ライティングについて

ライティングでは、CPS の洗練に加えて、「結果 (Consequence)」の記述を加えた (図 5.9)。また、「アクション (Action)」は独立して記述はしなかった。その理由として、結果の記述は、不協和が解消された後の状況について、制作物状態や、実践者本人の心理、チームメンバーとの人間関係などさまざまな視点で記述でき、状況の記述と併せて読むことで、読者にとってリアルにその場面を想起しやすいと判断したためである。アクションを記述しない理由は、アクションを独立して記述することでパターンの具体性が増し、特に経験が浅い読者にとって創造的な選択肢を考え出す余地を狭める可能性があると判断したためである。一方、やや表現が抽象的で、読者がイメージを得られにくいと感じたパターンについては、解決の記述の最後に「例えば～」のような書き出しでアクションの要素を付加したものもある。

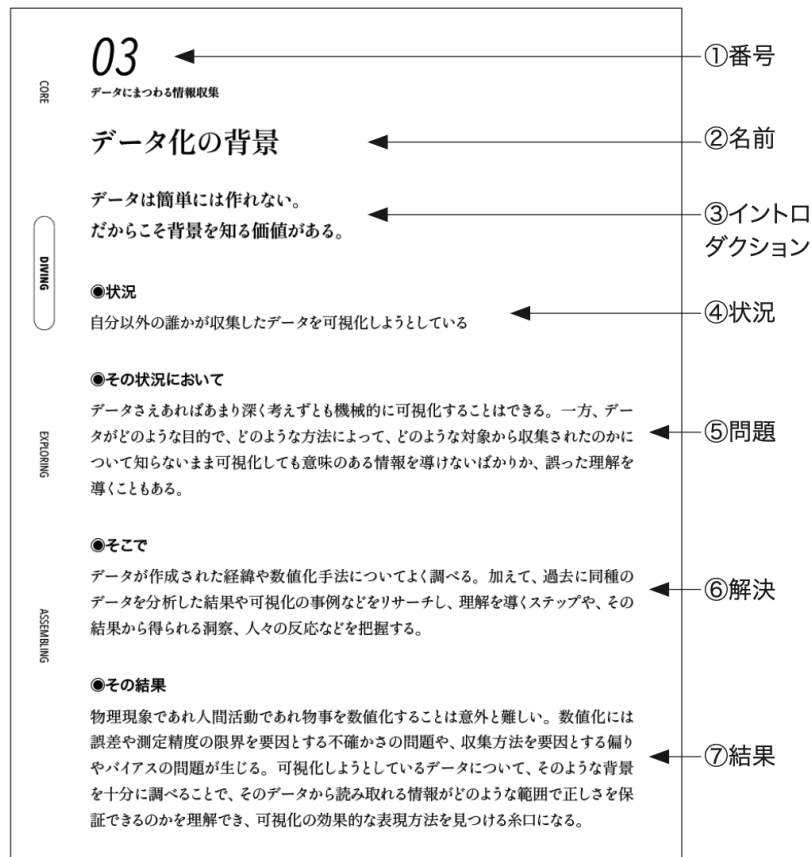


図 5.9 パターンの記述例

5.3.6 パターン・シンボライジングについて

シンボライジングでは、パターンの名称のほか、イントロダクションを記述した。また、イラストは記述しなかった。イントロダクションは1行説明と呼ばれ内容を想起しやすくする役割がある。イラストも同様の目的で描かれる。パターンは名前、イントロダクション、イラストの3点で内容の7~8割が伝わるのが目安とされる。イラストを記述しなかった理由は、主には時間的コストの要因による。また、本研究では、パターン・ランゲージ単独ではなく、プロセスモデル説明も同時に読者に提供することを想定しているため、プロセスの流れを理解することで、イラストがなくても名前とイントロダクションで内容を想起できると判断したためである。

また、DEXAS との対応を想起しやすくするために5つのカテゴリーを設けた。カテゴリー「コア」は全体を束ねるパターンに付与した。カテゴリー「データにまつわる情

報収集」は Diving に属するパターンに付与した。カテゴリー「データを可視化する」は Exploring（可視化）に属するパターンに付与した。カテゴリー「可視化をレビューする」は Exploring（レビュー）に属するパターンに付与した。カテゴリー「作品を組み立てる」は Assembling に属するパターンに付与した。

5.4 開発したパターン・ランゲージ

本研究で開発したパターン・ランゲージのパターン一覧を表 5.1 に示す。パターンの体系を図 5.10 に示す。パターンの記述詳細については付録「データビジュアライゼーション・デザインパターン」を参照されたい。

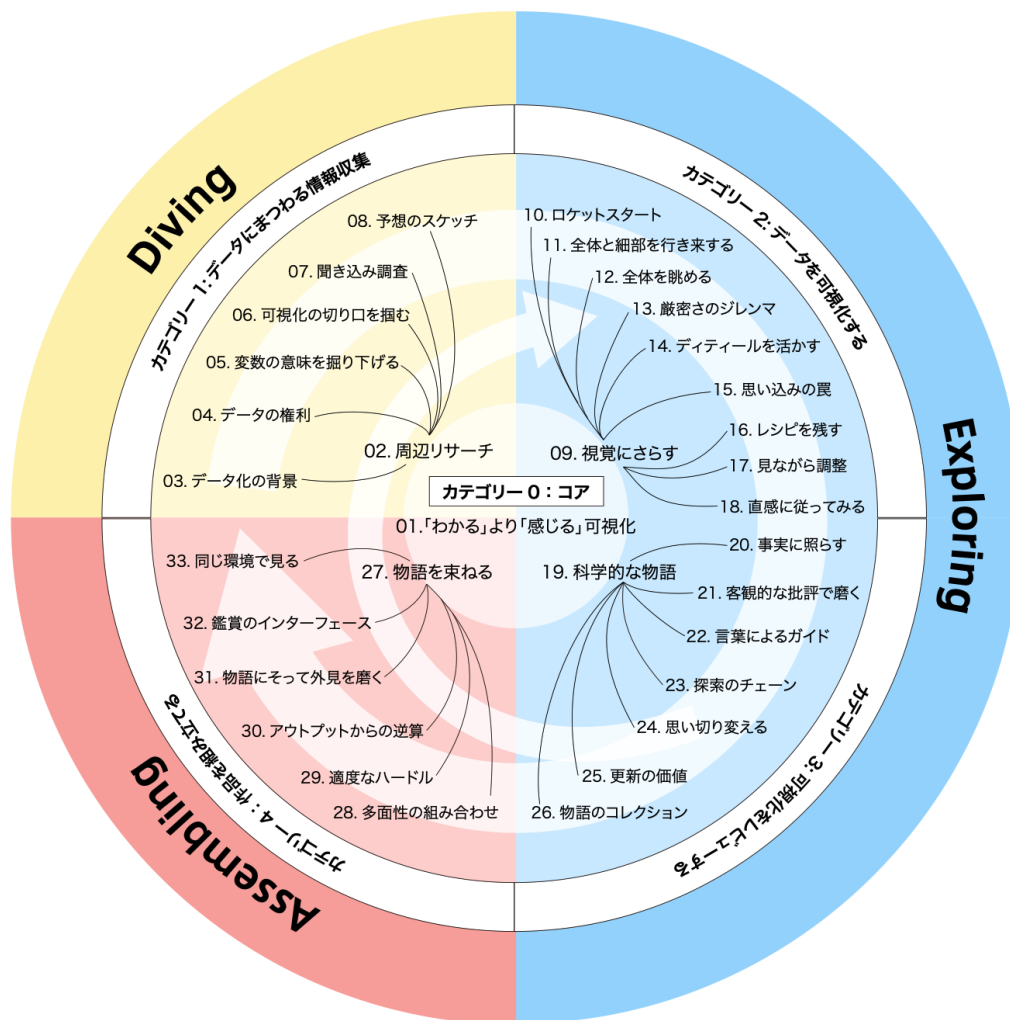


図 5.10 パターンの体系

表 5.1 パターン一覧

No	カテゴリー	タイトル
1	コア	「わかる」より「感じる」可視化
2	データにまつわる情報収集	周辺リサーチ
3	データにまつわる情報収集	データ化の背景
4	データにまつわる情報収集	データの権利
5	データにまつわる情報収集	変数の意味を掘り下げる
6	データにまつわる情報収集	可視化の切り口を掴む
7	データにまつわる情報収集	聞き込み調査
8	データにまつわる情報収集	予想のスケッチ
9	データを可視化する	視覚にさらす
10	データを可視化する	ロケットスタート
11	データを可視化する	全体と細部を行き来する
12	データを可視化する	全体を眺める
13	データを可視化する	厳密さのジレンマ
14	データを可視化する	ディティールを活かす
15	データを可視化する	思い込みの罠
16	データを可視化する	レシピを残す
17	データを可視化する	見ながら調整
18	データを可視化する	直感に従ってみる
19	可視化をレビューする	科学的な物語
20	可視化をレビューする	事実に照らす
21	可視化をレビューする	客観的な批評で磨く
22	可視化をレビューする	言葉によるガイド
23	可視化をレビューする	探索のチェーン
24	可視化をレビューする	思い切り変える
25	可視化をレビューする	更新の価値
26	可視化をレビューする	物語のコレクション
27	作品を組み立てる	物語を束ねる
28	作品を組み立てる	多面性の組み合わせ
29	作品を組み立てる	適度なハードル
30	作品を組み立てる	アウトプットからの逆算
31	作品を組み立てる	意味から外見を磨く
32	作品を組み立てる	鑑賞のインターフェース
33	作品を組み立てる	同じ環境で見る

第6章

データビジュアライゼーション・デザインパターン

6.1 制作の目的

4章でDEXASモデルについて、また5章でパターン・ランゲージについて述べた。本章では、この二つを読者が活用できるようにまとめた小冊子「データビジュアライゼーション・デザインパターン」の制作について報告する。

本研究の目的は、データビジュアライゼーションの制作プロセスを明らかにするだけでなく、実務者や学習者が制作プロセスを理解した上で、各々の状況に合わせて創造性を働かせて制作することを支援することである。ここでの制作の支援とは次に掲げる3点であり、データビジュアライゼーション・デザインパターンはこれを目的に記述する。

1. 制作経験がない場合の学びの支援

制作を経験していない読者に向けて、データビジュアライゼーションの概要、目的、過去に作られた参考作品を紹介しその意義を理解することを助ける。制作プロセスの全体像を示し、その要点や必要となる知識を整理し、自身に不足している関連知識を自覚させ学習を促す。

2. 制作過程における支援

制作を開始するにあたって、必要なメンバーを構成し、チームにおける自身の役割を意識させる。加えて、プロジェクトをリードするにあたってその工程の全体像がイメージできる。制作の各過程で生じる問題とパターンを関連づけて考察し、解決に向けた方策を選

択できる。

3. 制作を経験した後の学びの支援

制作の中で経験したこと、いかに問題を解決したかななどを、新たに形式知として記述し蓄えることができる。

6.2 想定読者

データビジュアライゼーション・デザインパターンは主としてデザイナーを読者と想定して記述する。その理由をいかに示す。

前述の通り、データビジュアライゼーションプロジェクトはデザイナーと専門家の協働で行う。しかし、DEXAS はデザイナーの視点で整理する。これはプロジェクトそのものを終始デザイナーがリードして進めていくことを意味する。一見、データ分析や可視化に詳しい専門家がリードして、デザイナーに指示した方が都合がよさそうに思える。なぜデザイナーがリードするのか、その理由は2つである。

ひとつは、大規模なデータといえども、専門家が見るポイントはかなり絞られているためである。なぜなら、専門家は自らの問題意識から生まれた研究の問い、すなわち関心事に対してその答えや洞察を得るべくデータを分析する。言い換えれば、自らの関心から外れる部分や関連が薄い部分についてはあまり探索を行わないし、関心のある部分であっても必要以上に複雑な可視化を行うことはない。データがあるからといって闇雲に覗き込んで調べたりしないのは、研究者としてごく自然な振る舞いであるからである。

一方、データビジュアライゼーションでは一つの視点からではなく、データのより多面的な見方を探る必要がある。これはひとつの問いを深く研究する専門家の使命とずれており、プロジェクトの目的とは合致しない。そのため、目的に到達するためにはデザイナーが専門家の視点を借りるようにして情報を聞き出し、リサーチしながらその外側へと探索していく必要がある。

もうひとつの理由は、専門家とデザイナーが常に密接にコラボレーションできるとは限らないためである。すでに書いた通り、さまざまな視点でデータビジュアライゼーションを探求するのはデザイナーの仕事であり、専門家の仕事とは言えない。そのため、専門家はデザイナーからリサーチされる対象として、またデザイナーが探索した可視化から読み取れる現象の科学的な解釈や、リサーチを深めるための助言をする立場として、要所で力を発揮する存在として位置付ける必要がある。しかしながら、デザイナーにとって都合の良いところだけ専門家の力を借りるようでは信頼関係を構築できず、目的の達成を阻害す

る。そのため、専門家との信頼関係や環境づくりも結果を左右する要因として心がけるべきである。

6.3 構成

データビジュアライゼーション・デザインパターンは次の要素で構成される。

はじめに

[I] データビジュアライゼーションとは何か

データビジュアライゼーションの目的

作品紹介 1

作品紹介 2

作品紹介 3

[II] データビジュアライゼーションをつくる

データビジュアライゼーションの難しさ

DEXAS でデータビジュアライズを共創する

DEXAS は誰の視点で作られている？

DEXAS の詳細

DEXAS のチーム構成

[III] DEXAS パターン・ランゲージ

DEXAS パターン・ランゲージとは

パターン・ランゲージの使い方

パターン・ランゲージの全体像

パターンの読み方

パターン一覧

第7章

ワークショップでの運用と考察

7.1 はじめに

6章ではデータビジュアライゼーション・デザインパターン（以下、「Data-Viz デザインパターン」と称す）の制作目的として次の3点を挙げた。

1. 制作経験がない場合の学びの支援
2. 制作過程における支援
3. 制作を経験した後の学びの支援

上記3点のうち、1と2に掲げた目的が達成できたかどうかを確認するために、学生7名を対象に、1ヶ月間のデータビジュアライズワークショップを実施した。本章ではその結果と考察を述べる。

7.2 ワークショップの概要

7.2.1 参加者の概要

ワークショップには、メディアアートやデザインを学ぶ学生7名（学部4年5名・修士課程1年2名）が参加した。参加者の平均プログラミング経験年数は4年（最大値7年、最小値2年、中央値4年）であった。参加者が「使用経験がある」と回答したプログラミング言語はProcessing（6名）、Python（6名）、JavaScript（4名）、C++（3名）、R（2名）、Typescript（1名）、Rust（1名）、Swift（1名）、Ruby（1名）、GLSL（1名）であった。参加者一人当たりの平均言語使用数は3.7（中央値3）であった。プログラミングによるデータビジュアライゼーションの経験は、2名が「全くない」、1名が「半年程度」、

3名が「1年程度」、1名が「2年程度」と回答した。データ分析の経験は、3名が「全くない・ほとんどない」、2名が「1年程度」、2名が「2年程度」と回答した。全ての参加者がプログラミングによる視覚的表現を含んだ作品制作を少なくとも1回は経験していた。

参加者の募集方法については、「専門的なビッグデータを使ってビジュアライゼーションを制作するプロセスを経験する」というワークショップの趣旨を示し参加者を募った上でプログラミング経験の有無を確認した。その結果、初級以上の実装スキルがあると認められた希望者が参加した。

7.2.2 ワークショップのスケジュール

ワークショップは以下の日程で実施した。

- 第1回 2022年10月18日：ガイダンス
- 第2回 2022年10月25日：構想発表
- 第3回 2022年11月16日：中間レビュー
- 第4回 2022年11月29日：最終レビュー

このワークショップでは作業時間を指定せず、参加者自身が日常の中で作業時間を確保し制作を進める形式をとった。

ガイダンスではワークショップの進め方と、次節に述べる目的、テーマ、制作のポイント、使用データ、参考資料について説明した。また、Data-Viz デザインパターンのPDFを配布し次回の構想発表までに（制作に入る前に）全体を読むように依頼した。

構想発表、中間レビュー、最終レビューでは各自3分程度の発表（レビューにおいては作品のデモ）とディスカッションを行った。最終レビュー以外はオンラインミーティングツール（Zoom）を使用した。ただし、事情により来場が困難な2名のみ最終レビューもオンラインで参加した。

上記4回の全体会合以外に毎週末に参加者と15分程度の面談をオンラインで実施した。面談ではその週の作業時間と制作状況をヒアリングした。

7.3 ワークショップの実施計画

ワークショップの計画とその意図について述べる。

7.3.1 目的の提示

ワークショップの目的を「複雑なデータを魅力的な外観でビジュアライズし、鑑賞者の興味をかき立てつつ、その意味を探究させるビジュアライゼーションを作るプロセスを体験する」とした。この理由は、本研究の主題である、複雑なデータを題材にしてデータの背後にある意味を鑑賞者自身が考察を促されるような魅力的な作品を作ることと合致している。

7.3.2 テーマ設定

制作テーマを「パンデミックをビジュアライズする」とした。また、その趣旨説明を「COVID-19 は、初めて大規模なデータによって記録されたパンデミック（世界的な疫病の蔓延）である。2020 年から現在までの状況をデータで振り返り可視化する。」とした。このテーマ設定の理由は、参加者自身が COVID-19 についての個人的な経験や社会的な出来事を記憶している一方、公的機関によるさまざまな公開データによって客観的・科学的に振り返ることも可能なことから、主観と客観を行き来しながら、多様な視点をとらえやすいと想定したことによる。

7.3.3 制作の方向づけ

ワークショップの目的をより具体化しつつ制作の方向性を示すために制作のポイントを下記のように示した。

ビジュアライゼーションの作り方は自由。ただし3つのポイントを重視して制作してください。

1. 多様なあり方が感じられるような表現で可視化する（シンプルな棒グラフなど、一つの意味しか見出せない可視化は NG）
2. 多面的な視点でデータを眺めることができる（インタラクションできる等）
3. 興味を喚起したり、問題について幅広く議論する場を生み出す（興味を起こさせるビジュアル＋ストーリーを読み取れるものを目指す）

この指示の意図は、2.5 節で示したような編集的視点によって参加者が意図的に表現の選択肢を狭め、特定の事象を意味する表現のみを洗練することを避けるためである。

7.3.4 レビューの方法

中間レビューと最終レビューでは、制作物のデモをするだけでなく以下の点を依頼した。

1. 作品テーマの説明
2. データソース（データの取得先）の説明
3. 「可視化からこういうことが言えるのではないか」という自分の意見を投げかける

この理由は、デモ後のディスカッションにおいて作品の視覚的特徴の部分だけでなく意味や手法についても話題を広げることを意図したためである。

7.3.5 データ

このワークショップでは題材とするデータを指定した。データはゲノムウイルス変異の国際的なオープンデータベースである GISAID からランダムサンプリングした約 3000 の変異ウイルスのリストである（期間は 2019 年末～2022 年 10 月）。このリストには一つの変異ウイルスにつき、検出国、検出場所の緯度経度、検出日時、系統の国際分類指標である「Pango lineage」、オミクロンなどのいわゆる変異株の名称を含む「Clade」、野生株と比較したゲノムの変異量を示す「Divergence」など 17 の変数を含む。これらをプログラミングで扱いやすい CSV (Comma Separated Value) 形式に筆者が変換したものである。変数の意味を詳細に解説することは避け概略の説明にとどめ、後述する参考資料を足がかりとしてリサーチするように依頼した。

データを指定した理由は、筆者自身の過去の経験から参加者に自由にデータを探索させるとデータソースのリサーチやプログラミングしやすい形式に加工することに時間を費やしてしまい、本質的な作業に必要な時間が十分に確保できない傾向があるためである。また、参加者にとってあまり馴染みのない変数を持つデータを提示することで変数の意味や対象をリサーチする必要性を感じ取ることも本研究の趣旨と合致すると考えたためである。

一方、データを指定することは参加者のモチベーションや創造的な表現の幅を減じることに繋がるとも考えられたため、このデータに関連する意味を持つデータを参加者が独自に追加することを許容した。

しかし、実際にワークショップを進めるにつれて一部の参加者は提供したデータの使用が必須かどうかの確認を求めた。その主な理由としてデータが世界全体を対象としたラン

ダムサンプリングデータであるため、日本だけに焦点を当てた可視化を行いたい場合に使用できるデータ量が非常に少なくなることが挙げられた。参加者らは代替となるデータを既に見つけリサーチや形式の変換を進めており、ワークショップの趣旨も理解していることから上記に挙げた支障には当たらないと判断し、最終的にこのデータの使用は必須としないこととした。

7.3.6 参考資料の提供

前述のデータに含まれる変数の意味を理解したり、変異と感染拡大の関係を理解するためには生命科学の基礎的な知識が必要となる。データをもとにリサーチを始める足がかりとなる参考資料として、『生物の科学 遺伝』Vol.75 特集「新型コロナ・パンデミック その衝撃と科学世界への影響」（株式会社エヌ・ティー・エス）の一部を参加者へ提示した。提示した特集記事のタイトルは次のとおりである。

1. 新型コロナパンデミックその衝撃と科学世界への影響
2. 新型コロナウイルスの遺伝子と変異
3. 新型コロナウイルス顕微鏡写真
4. 新型コロナウイルスとこの一年（2020年版）

7.3.7 レビューに専門家を招聘

7.2.2 項で示したとおり、このワークショップでは中間レビューと最終レビューを設けた。この際、感染症や生命科学の見地から見解を示したり助言を行う専門家を招聘した。招聘の目的は、参加者が自らリサーチと表現する過程で見出した特徴やパターン、データの背後にあると考えられる現象についてプレゼンテーションした後、ディスカッションで専門家からその妥当性の評価を引き出したり、専門家ならではの着目点を引き出すためである。

専門家にワークショップへの参加を依頼する際、本研究の内容や Data-Viz デザインパターンについては明かさず、前述のワークショップの目的、テーマ、データ、参加者の属性のみを説明し参加の承諾を得た。

ワークショップへ参加した専門家は、ワークショップ実施時点で国内の大学医学部准教授の職位にあり、ウイルス学やゲノム生物学を専門とする博士号（理学）を有する研究者である。COVID-19 の世界的な感染状況に着目した研究に従事し、新型コロナウイルスの

ゲノム解析に関して複数の原著論文、総説の掲載などの業績がある。

7.3.8 参加者のコミュニケーションツール

ワークショップの期間中、参加者（筆者を含む）同士のコミュニケーションツールとして、「Discord」*¹を使用した。ワークショップ専用の Discord サーバを設置し、参加者と筆者を登録した。「サーバ」内に「連絡用チャンネル」と「雑談用チャンネル」を設置した。雑談用チャンネルでは参加者らが試作した可視化のスクリーンショットを添付したり、データの入手先や処理の仕方を相談する会話が行われた。先述の専門家は Discord には含まず、メールにてコミュニケーションが取れることを参加者に説明し、メールアドレスを開示した。専門家にメールを送付する際は筆者を CC に入れるように依頼した。なお、期間中に専門家とメールでコンタクトした参加者はいなかった。

7.4 記録方法

参加者の制作プロセスと Data-Viz デザインパターンについての内省を次の方法で記録した。

7.4.1 事前アンケート

事前アンケートは制作に入る前の段階（第 2 回：構想発表の直後）に行った。回答時、回答者が第 1 回ガイダンスで配布した Data-Viz デザインパターンを読了したことを確認した上で行った。

設問 1 は次のとおりである。配布した PDF を手元に準備した状況で「各パターンの“状況”と“その状況において”を読んだ上で、その問題について当てはまるものをお答えください。」という指示があり、回答者はパターン 1 からパターン 33 について以下の 4 つの選択肢から一つを選ぶ。

選択肢 A このような問題が起こる場면을想像できない

選択肢 B このような問題が起こることを知っている

選択肢 C このような問題を実際に認識したことがある

選択肢 D このような問題を解決したことがある

*¹ Discord 社が提供するコミュニケーションサービス。<https://discord.com>

この設問の作成にあたっては、大学生を対象とした Project Based Learning の現場でパターン・ランゲージを活用し、学生が持つパターンへの認識の変化を調査した長田の研究手法 [77] を参考とした。

設問 2 は自由記入形式で「その他、パターン・ランゲージについて感じたことがあれば記入してください。」とした。

7.4.2 事後アンケート

事後アンケートはワークショップ終了時に行なった。

設問 1 は事前アンケートの設問 1 と同じである。設問 2、設問 3 はいずれも自由記入形式で次のとおりである。

設問 2 今回のワークショップでは、パターン・ランゲージを含む「データビジュアライゼーション・デザインパターン」を読んだ上でデータビジュアライズ作品の制作に取り組み、専門家の方も同席したレビューを 2 回行いました。この方法についての感想を聞かせてください。

設問 3 「データビジュアライゼーション・デザインパターン」を読んだことでどのような効果があったと思いますか？ できるだけ具体的な利用シーンを思い出して記入してください。(パターン・ランゲージ以外の部分でも、読んで参考になった部分があれば記入してください)

7.4.3 対話の記録

第 1 回ガイダンスを除く 3 回の会合と、参加者との毎週末に行った面談については映像と音声記録した。

7.5 制作事例

本節ではワークショップ参加者が最終レビューで提示した作品について述べる。ここでは便宜上、作品と呼称するがこの時点での制作物は展示したり鑑賞者が自由に体験できるように完成させる必要はなく、当初の目論見通りできたことや、そうでなかったことをデモを通じて説明できるものを求めた結果であることを書き添える。このように求めた理由は、筆者の経験からデータビジュアライゼーションを展示に耐え得る作品にするためにはさらに長い時間をかけた検証、ソフトウェアの調整、鑑賞体験のデザインが必要となる。

そのためワークショップの期間と参加者の負担を加味したためである。

7.5.1 事例 1: 参加者 Y の作品

Y は報道で「全国的に感染者が増えている」と伝える緊迫した感覚と、身近にはあまり感染者がいないという実感と報道のギャップを感じるが多かったことを思い出し、住んでいる場所（地方都市）と大都市の状況に差があることを可視化することに興味を持った。そこで、感染者数の推移を地図上に立体的に表し、カメラの視点を特定の都市に合わせつつその背後に全国の様子が見え、鑑賞者にとって身近な都市と全国の状況の対比が可視化をされる作品を制作した（図 7.1,7.2）。感染者数を都市の座標を中心とする球体の直径に割り当て、時系列アニメーションを実装した。鑑賞者はマウス操作で視点を変更できる。特に何も操作しなくても時間が進行し（auto play）感染者の変動を見ることが出来る。その時々々の生活状況を想起しやすくするために、政府による緊急事態宣言下の地域を赤で、蔓延防止等重点措置発令下の地域を黄色で着色した。感染者数は厚生労働省オープンデータ*2から、各都市の新規陽性者数（日別）を取得した。

*2 <https://www.mhlw.go.jp/stf/covid-19/open-data.html>

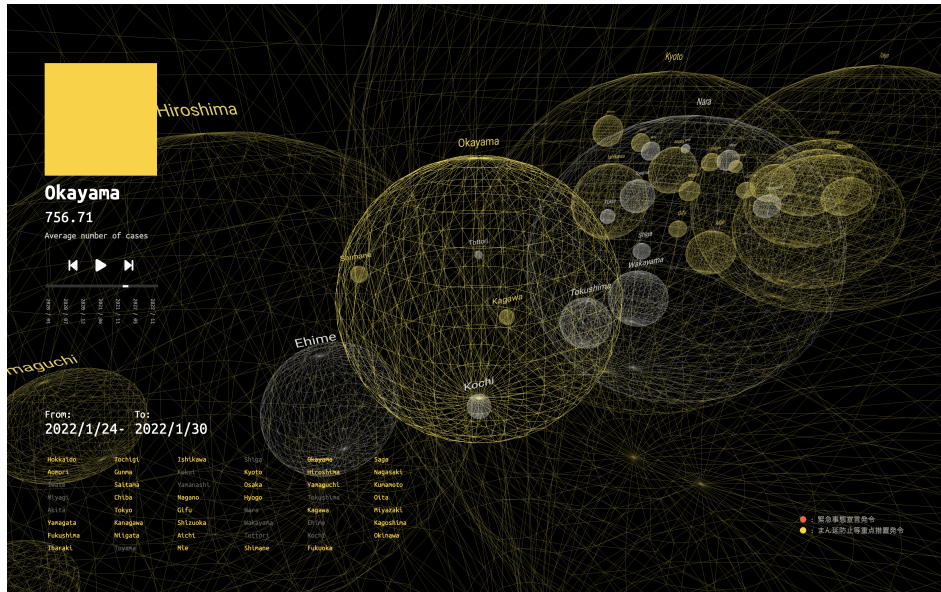


図 7.1 Y の作品（蔓延防止等重点措置発令下の都市とそうでない都市）

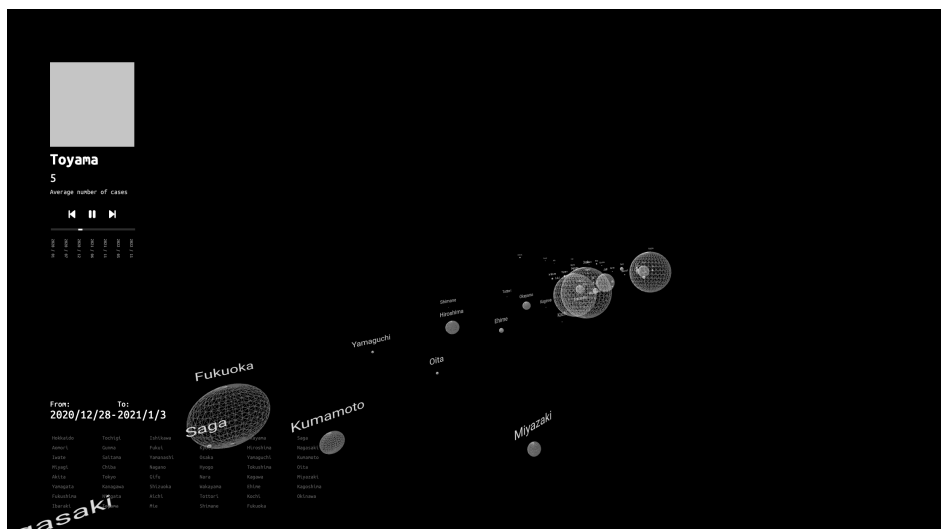


図 7.2 Y の作品（地方都市から大都市をのぞむ構図）

7.5.2 事例 2: 参加者 U の作品

U は日本国内の感染者数の推移と流行した変異株の関係に注目し、さらに各都市ごとに流行した変異株の種類や時間に差があるかという関心を抱いた。U は国立感染症研究所がウェブサイトで発する報告書「新型コロナウイルス ゲノムサーベイランスによる系統別検出状況」の 2022 年 9 月 22 日報告版^{*3}をもとに、2020 年 1 月から 2022 年 9 月までの期間の 1 週間ごとに各都市で最も多く検出された変異株 (PANGO lineage) を調べた。また感染者数については、厚生労働省オープンデータ^{*4}から、各都市の新規陽性者数 (日別) を取得した。Processing を使用してこれらのデータを 2D (図 7.3) と 3D (図 7.4) でビジュアライズした。2D 版では併置された各都市の棒グラフで感染者数の推移を表し、色で流行した変異株を表した。3D 版では時間軸と緯度経度を 3 次元空間に割り当て、各都市の位置に配置した矩形で感染者数の推移を表した。流行株は 2D と同様に色で表現している。3D 版ではマウス操作で回転とズームイン・アウトを、スライダの操作で時間を選択すると選択した時間軸状に地図が移動するインタラクションを実現している。

^{*3} https://www.niid.go.jp/niid/images/cepr/covid-19/20220928_genome_surveillance.pdf

^{*4} <https://www.mhlw.go.jp/stf/covid-19/open-data.html>

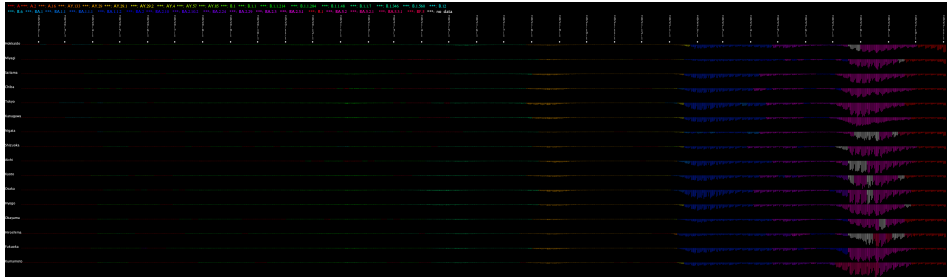


図 7.3 U の作品 (2D 版)

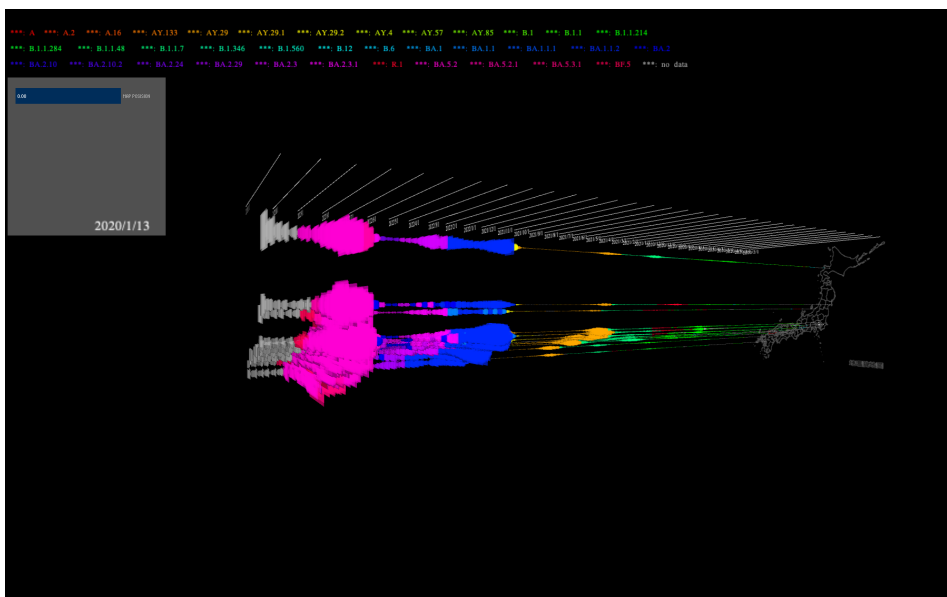


図 7.4 U の作品 (3D 版)

7.5.3 事例 3: 参加者 M の作品

M は、全国の病床使用率、新規感染者数、政府の感染拡大防止対策の 3 つの関連に着目した。各都市の日別新規感染者の推移を地図と垂直にとった時間軸上に円の直径で表している（図 7.5,7.6）。画面左上には時間を表すスライダーがあり、赤いドットがスライダーを移動することで参照している日付を表す。スライダーの下には、参照している日付に最も近い未来に政府がとった対策が文字で表示される。新たな対策が表示されるとサウンドが再生される。参照する日付と連動して各都市の円の中心にある赤い小さな球が垂直に移動し日付を示す。新規感染者数を表す円の半分は病床使用率に応じて透明度が変化し使用率が高まるほど不透明になる。閲覧者は右側のボタンで地域を選択できる。またマウス操作で自由に視点を移動できる。日付の進行は Auto Play である。対策のデータは、内閣官房のウェブサイト*⁵から、新規感染者数のデータと病床使用率のデータは、厚生労働省のオープンデータ*⁶からそれぞれ取得した。

*⁵ 新型コロナウイルス感染症緊急事態宣言の概要、内閣官房新型コロナウイルス等感染症対策推進室
https://corona.go.jp/news/news_20200421_70.html

*⁶ 厚生労働省、データからわかる - 新型コロナウイルス感染症情報 - <https://covid19.mhlw.go.jp/extensions/public/index2.html>

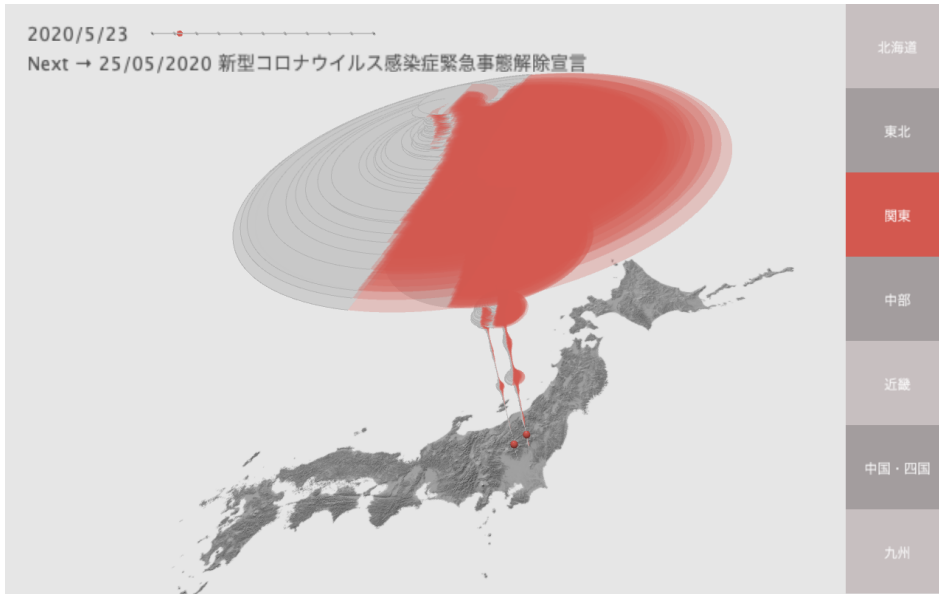


図 7.5 M の作品（関東）

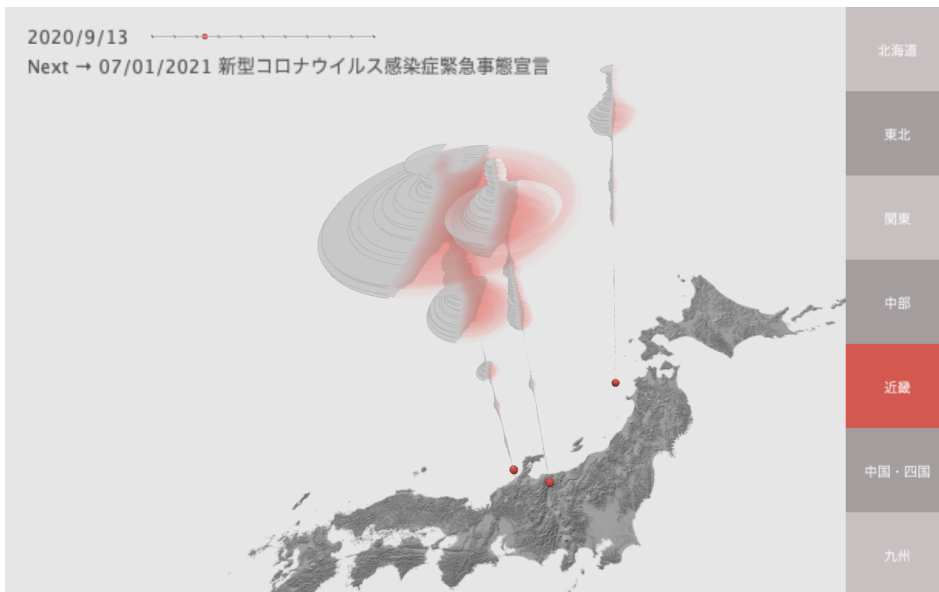


図 7.6 M の作品（近畿）

7.5.4 事例 4: 参加者 K の作品

K は人の移動の激しさと感染の拡大に関心を持った。そこで、1km グリッドの人の滞在者数データを 20km グリッドに結合したデータを使い、地図上のバーの高さでそのエリアの滞在者数を可視化した（図 7.7,7.8）。各都市の新規感染者数日別数と緯度経度から、地図面に対して垂直な時間軸をもつ空間に小さな立方体を配置し、その色で感染者数を可視化した。立方体の透明度と彩度が人数を表す。人数が多いほど濃紫色に、少ないほど透明で淡い紫色となる。閲覧者は視点と時間を操作できる。U の作品と同様に、時間の選択に応じて地図が時間軸状を並行に移動する。感染者数のデータは、厚生労働省オープンデータ^{*7}から、滞在人数のデータは、G 空間情報センターのオープンデータ^{*8}から取得した。

^{*7} <https://www.mhlw.go.jp/stf/covid-19/open-data.html>

^{*8} 全国の人流オープンデータ（1km グリッド、市区町村単位発地別） - データセット <https://www.geospatial.jp/ckan/dataset/mlit-1km-fromto>

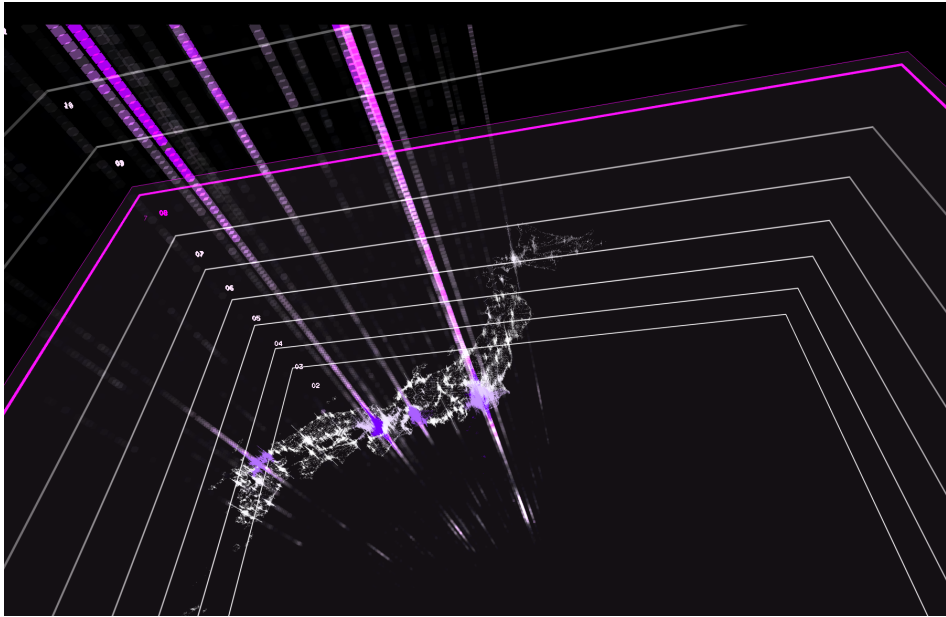


図 7.7 K の作品（全国を俯瞰）

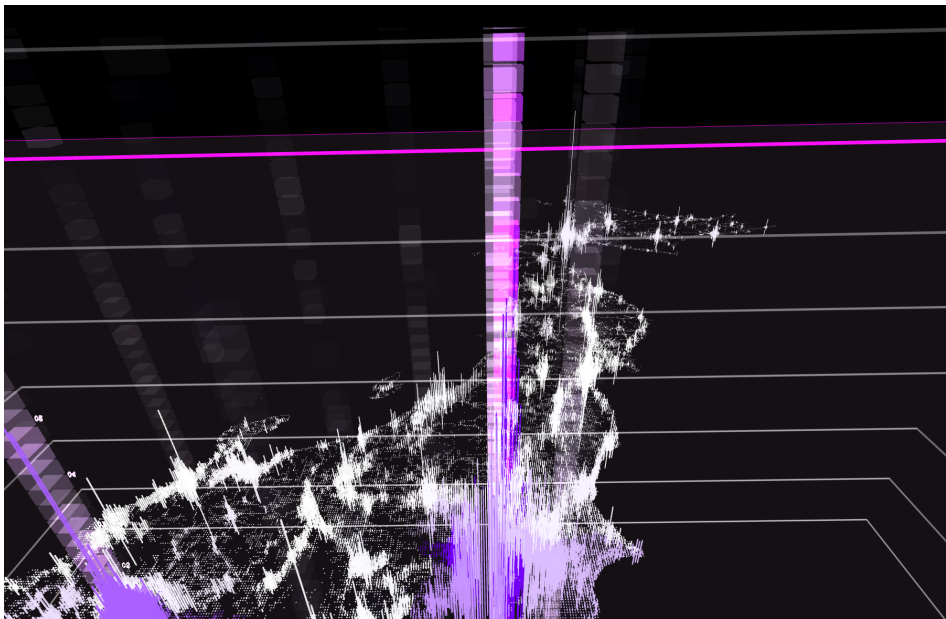


図 7.8 K の作品（細部の状態）

7.5.5 事例 5: 参加者 N の作品

N はアフリカなど発展途上国の状況が日本では特に報道されにくいなど、世界全体に感染が蔓延しているにもかかわらず、先進国など一部の国々の状況にしか触れていないことに関心を持った。各国の感染者の推移とワクチン接種率の推移から世界的な感染の拡大と収束の様子を描き出そうとした（図 7.9,7.10）。各国のワクチン月別接種率をコロプレス図のように各国の色相に割り当てて塗り分けた世界地図を表現した（接種率 0% は青、100% は赤、50% は緑）。月別新規感染者数（対人口比）を各国の大きさに割り当て、変化に応じて形が伸縮するように表現した。左下に時間が表示されている。時間の進行は自動である。各国同士の国境部分では接続を維持しておらず切り離されているが、これは N 本人は接続を維持したまま変形したいと考えていたが実装にかかる時間が足りなかったためであり意図的な表現ではない。新規感染者数、ワクチン接種率とも Our World in Data のオープンデータを取得した*⁹

*⁹ Our World in Data <https://ourworldindata.org/coronavirus>

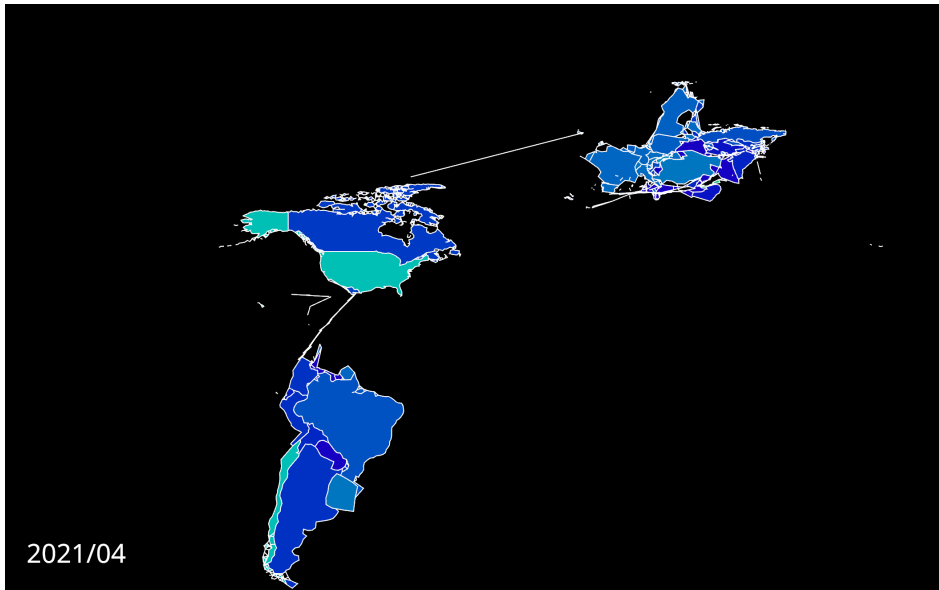


図 7.9 N の作品（ワクチン接種率が米国などで高まる時期）

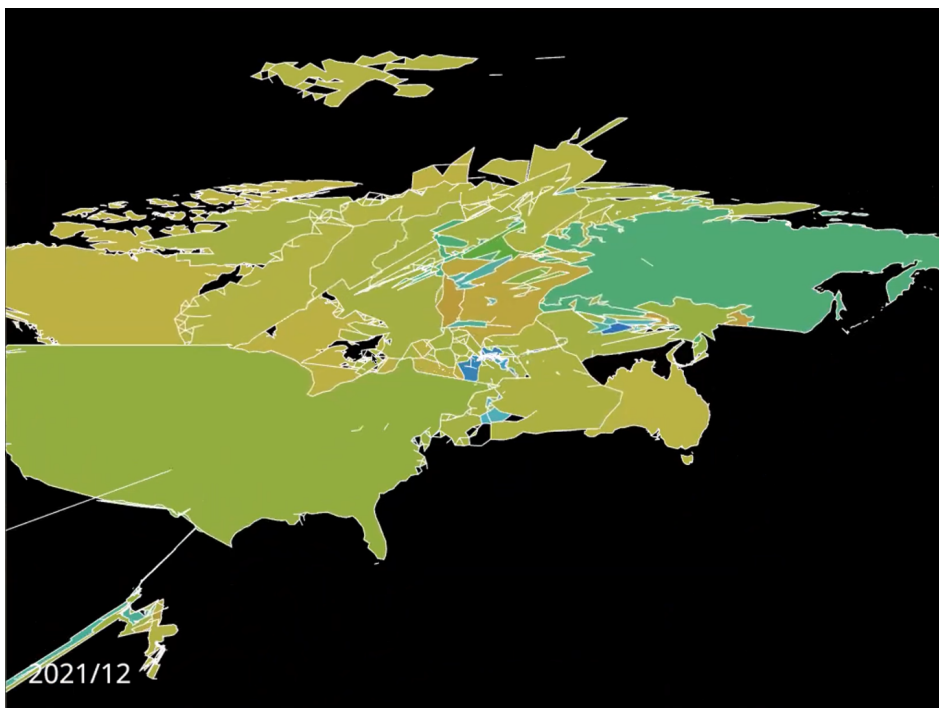


図 7.10 N の作品（ワクチン接種率が高まるも再び感染者数が増える時期）

7.5.6 事例 6: 参加者 O の作品

O は地球儀のように地球表面全体にデータをマッピングする表現に関心を持った。ワークショップ用に筆者が準備した変異ウイルスのデータセットを使用して、変異ウイルスの検出位置、時間、Divergence を可視化した (図 7.11,7.12)。Divergence とは変異したゲノムの塩基配列が当初の配列と比較して何箇所変異しているかを示す値である。変異が発生する頻度はおおむね一定であるため Divergence と検出された時期には正の相関がある。O は検出位置に三角柱形状のバーをプロットし、Divergence をバーの長さで、検出時期を色で可視化した。閲覧者は自由に視点を移動できる。また、右上に検出時期によってフィルタを行うためのテキストフィールドを設け、開始時期と終了時期を入力することで表示を限定できる。マウスを柱に合わせることで変異の検出地と Divergence が表示される。O は制作過程で可視化の形態と顕微鏡写真でみたウイルスの形状とが似ていることに興味を持ち、対比させる意味でその画像も可視化とともに表示している。

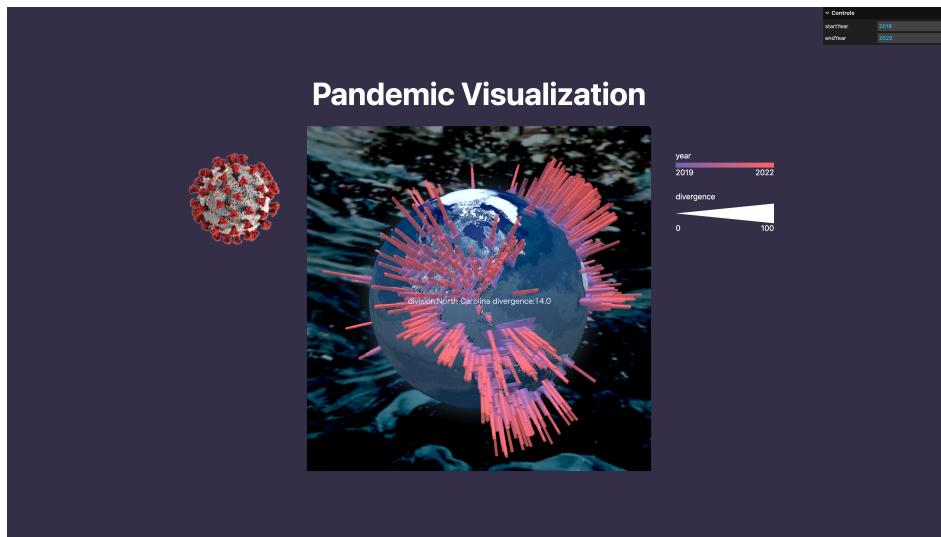


図 7.11 O の作品（全景表示）

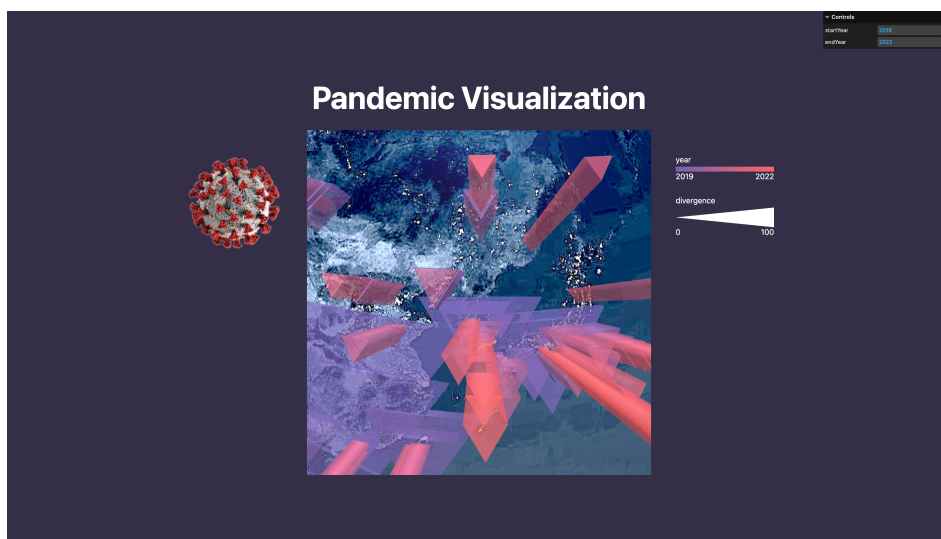


図 7.12 O の作品（細部の表示）

7.5.7 事例 7: 参加者 T の作品

T は時間経過とともに増える変異の多様性を可視化することに関心を持った。また COVID-19 以降、特に変異株の拡散以降において日常的に使用されるようになった「株」について視覚的なイメージ得たいと考えた。T はワークショップ用に筆者が準備した変異ウイルスのデータセットを使用し、横軸に Divergence^{*10}を縦軸に時間を割り当て（下が正の方向）、ドットプロットで変異ウイルスの検出時間と多様性を表現した。また変異ウイルスの Clade の違いをドットの色で表現した（図 7.13,7.14）。Clade とは名義尺度であり系統に付与される名称である^{*11}。この可視化は下向きスクロールすることによって全体を見る。またドットにマウスポインタを当てると、画面の右半分が大きく円が表示されその中に変異の検出地など詳細情報が表示される（図 7.15）。この時、円の大きさは感染力に影響するとされるウイルスのスパイク構造（突起の部分）の遺伝子の多様度を表す「S Mutation」によって可変し、多様であるほど（変異が蓄積しているほど）円が大きくなる。

*10 Divergence については O の作品の注釈を参照のこと

*11 たとえば、WHO が指定した VOC (Variants of Concern) では「Omicron」に分類される変異であっても Clade では「21K (Omicron)」や「21L (Omicron)」などさらに細分化される。

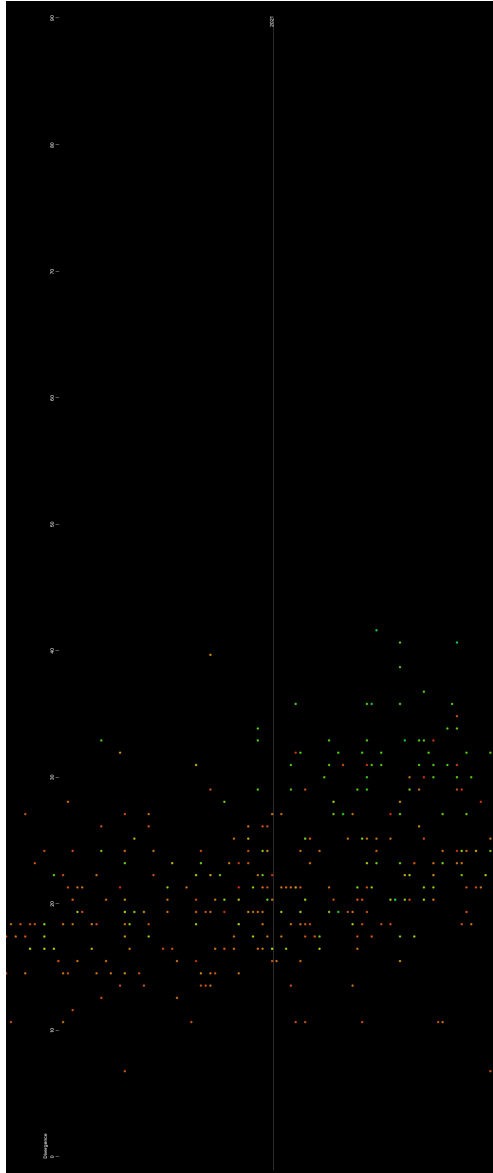


図 7.13 T の作品 (2021 年 1 月頃を表示)

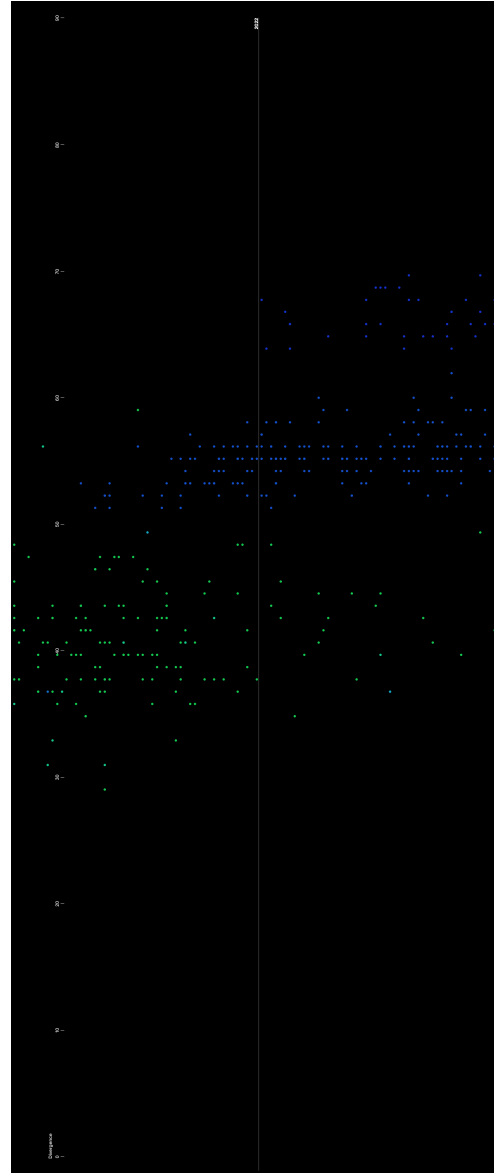


図 7.14 T の作品 (2022 年 1 月頃を表示)

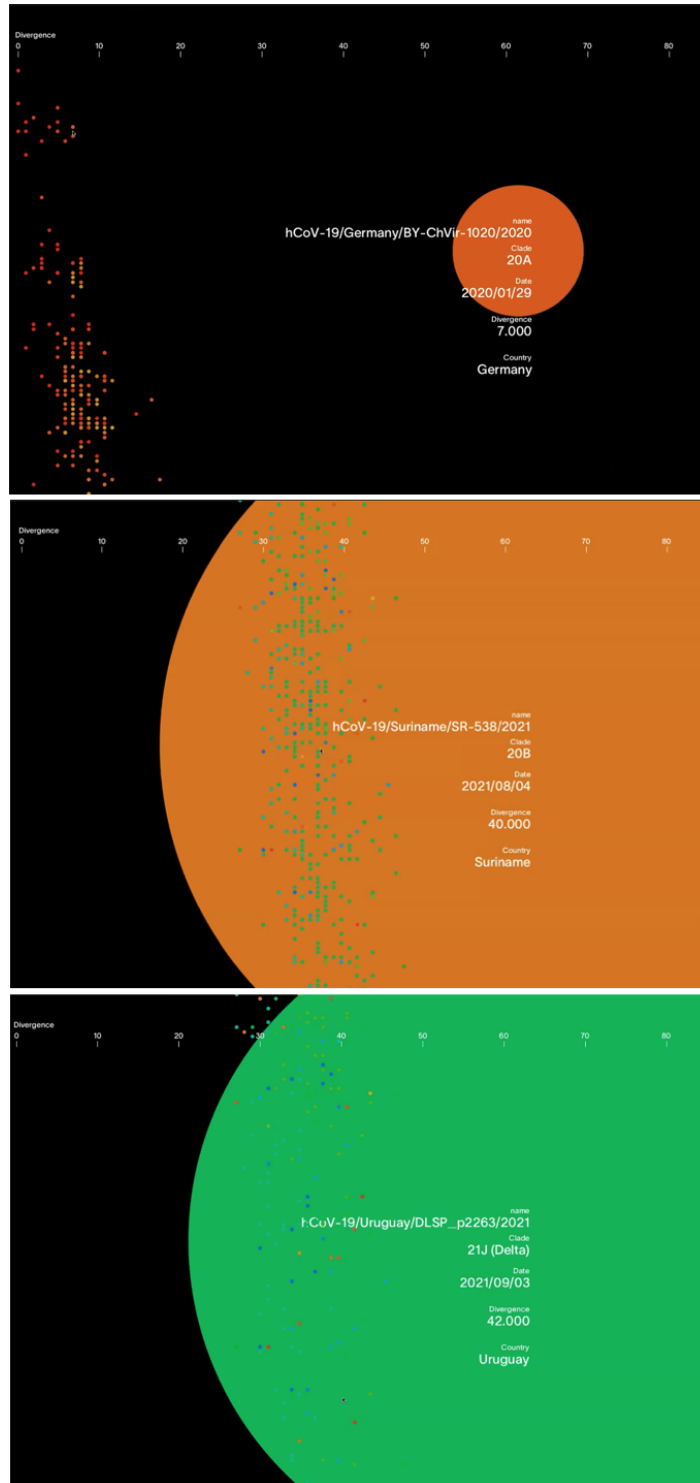


図 7.15 T の作品のドットにマウスオーバーした状態（上：2020 年 1 月頃の変異ウイルス、中：同 2021 年 8 月、下：2021 年 9 月）

7.6 アンケート結果と考察

本節ではアンケート結果を示す。アンケートは事前、事後ともに参加者全員が回答した。

7.6.1 事前・事後回答の比較

事前と事後の設問 1 は、パターン・ランゲージの各パターンについての認識を問うものであることは先述した通りである（7.4 節参照）。

各パターンについての回答結果を事前と事後で比較し、図 7.16 に示す。図では、各パターンについて A から D の選択した人数を示した。人数の分布をスパークライン（表内の小さなチャート）で示し、最も回答が多かった選択肢を緑で示した。平均点については、A、B、C、D をそれぞれ 1 点、2 点、3 点、4 点とし選択した人数と点数を掛けた合計を参加者人数で割り求めた。事後の平均点から事前の平均点を引いた差異を示した。差異では差の大きさ順に上位 10% を赤で示した。また上位 25% は下線で示した。

パターン名	事前アンケート結果					事後アンケート結果					事前・事後の差異											
	A	B	C	D	平均	A	B	C	D	平均	カテゴリ平均	平均の差異	カテゴリ平均の差異									
カテゴリ0	1	2	4	1	—	—	—	—	2.86	2.86	0	1	2	4	—	—	—	—	3.43	3.43	0.57	0.57
1. 「わかる」より「感じる」可視化																						
2. 周辺リサーチ	1	3	2	1	—	—	—	—	2.43	2.51	0	0	3	4	—	—	—	—	3.57	3.37	1.14	0.86
3. データ化の背景	0	4	3	0	—	—	—	—	2.43		0	2	2	3	—	—	—	—	3.14		0.71	
4. データの権利	0	3	3	1	—	—	—	—	2.71		0	2	1	4	—	—	—	—	3.29		0.57	
5. 変数の意味を掘り下げる	2	0	5	0	—	—	—	—	2.43		0	0	3	4	—	—	—	—	3.57		1.14	
6. 可視化の切り口を掴む	0	4	2	1	—	—	—	—	2.57	0	2	3	2	—	—	—	—	3.00	0.43			
7. 聞き込み調査	0	6	0	1	—	—	—	—	2.29	0	0	2	5	—	—	—	—	3.71	1.43			
8. 予想のスケッチ	0	3	3	1	—	—	—	—	2.71	0	1	3	3	—	—	—	—	3.29	0.57			
9. 視覚にさらす	0	5	0	2	—	—	—	—	2.57	2.57	0	1	3	3	—	—	—	—	3.29	3.37	0.71	0.80
10. ロケットスタート	0	1	4	2	—	—	—	—	3.14		0	2	1	4	—	—	—	—	3.29		0.14	
11. 全体と細部を行き来する	0	4	3	0	—	—	—	—	2.43		0	2	2	3	—	—	—	—	3.14		0.71	
12. 全体を眺める	1	4	0	2	—	—	—	—	2.43		0	1	1	5	—	—	—	—	3.57		1.14	
13. 厳密さのジレンマ	1	3	2	1	—	—	—	—	2.43	0	1	4	2	—	—	—	—	3.14	0.71			
14. ディティールを活かす	1	4	2	0	—	—	—	—	2.14	0	2	2	3	—	—	—	—	3.14	1.00			
15. 思い込みの罫	0	5	0	2	—	—	—	—	2.57	0	1	2	4	—	—	—	—	3.43	0.86			
16. レンビを残す	0	5	1	1	—	—	—	—	2.43	0	0	1	6	—	—	—	—	3.86	1.43			
17. 見ながら調整	0	4	1	2	—	—	—	—	2.71	0	1	1	5	—	—	—	—	3.57	0.86			
18. 直感に従ってみる	0	2	4	1	—	—	—	—	2.86	0	1	3	3	—	—	—	—	3.29	0.43			
19. 科学的な物語	0	5	2	0	—	—	—	—	2.29	2.21	0	1	5	1	—	—	—	—	3.00	3.16	0.71	0.95
20. 事実に照らす	1	6	0	0	—	—	—	—	1.86		0	2	3	2	—	—	—	—	3.00		1.14	
21. 客観的な批評で磨く	1	3	2	1	—	—	—	—	2.43		0	0	4	3	—	—	—	—	3.43		1.00	
22. 言葉によるガイド	1	3	3	0	—	—	—	—	2.29		0	2	2	3	—	—	—	—	3.14		0.86	
23. 探索のチェーン	1	4	1	1	—	—	—	—	2.29	0	1	1	5	—	—	—	—	3.57	1.29			
24. 思い切り変える	0	6	0	1	—	—	—	—	2.29	0	1	4	2	—	—	—	—	3.14	0.86			
25. 更新の価値	1	5	0	1	—	—	—	—	2.14	0	3	3	1	—	—	—	—	2.71	0.57			
26. 物語のコレクション	2	3	1	1	—	—	—	—	2.14	0	2	1	4	—	—	—	—	3.29	1.14			
27. 物語を束ねる	0	7	0	0	—	—	—	—	2.00	2.31	0	3	4	0	—	—	—	—	2.57	2.84	0.57	0.53
28. 多面性の組み合わせ	1	5	1	0	—	—	—	—	2.00		0	2	3	2	—	—	—	—	3.00		1.00	
29. 適度なハードル	1	4	2	0	—	—	—	—	2.14		0	1	5	1	—	—	—	—	3.00		0.86	
30. アウトプットからの逆算	0	4	2	1	—	—	—	—	2.57		0	3	3	1	—	—	—	—	2.71		0.14	
31. 物語にそって外見を磨く	0	4	2	1	—	—	—	—	2.57	0	2	2	3	—	—	—	—	3.14	0.57			
32. 鑑賞のインターフェース	1	2	4	0	—	—	—	—	2.43	0	4	2	1	—	—	—	—	2.57	0.14			
33. 同じ環境で見る	0	5	1	1	—	—	—	—	2.43	0	3	2	2	—	—	—	—	2.86	0.43			
									2.42									3.21		0.78		

回答選択肢 A:このような問題が起こる場面を想像できない 平均点算出方法 A、B、C、Dをそれぞれ1点、2点、3点、4点とし選択した人数と点数を掛けた合計を参加者人数で割り平均を求めた
 B:このような問題が起こることを知っている
 C:このような問題を実際に認識したことがある
 D:このような問題を解決したことがある

図 7.16 設問 1 について事前・事後の回答と比較

7.6.2 事前・事後の回答から見た変化

状況への意識の変化

パターンの「状況」への意識の変化について考察する。事前の回答では参加者が「A. どのような問題が起こる場面を想像できない」と回答するパターンが 14 件あった（2、5、12、13、14、20、21、22、23、25、26、28、29、32）。しかし、事後では、それを否定する回答（B,C,D）に変化した。このことから、パターンが示す状況や問題は参加者にとって想像することができるか、できないとしてもワークショップを通じて想像できるようになると考えられる。

問題の認識から解決に向けた実践への変化

事前には「B:このような問題が起こることを知っている」の回答が最も多いパターンが28件あり全体の85%であった。また「D:このような問題を解決したことがある」の回答が最も多いパターンはなかった。一方、事後では、Dの回答が最も多いパターンは20件で全体の61%となり他の回答に比べて最も高かった。次いでCが同42%であった。Dが最優位であるパターンが同一のカテゴリ内に占める割合を見ると（ただし、カテゴリ0を除く）、カテゴリ1（86%）、カテゴリ2（90%）、カテゴリ3（38%）、カテゴリ4（14%）となった。カテゴリ4については、他のカテゴリ同様に、BからCまたはDへの傾向が見られたが、事後にD（解決できた）の回答は全カテゴリの中で最も少なく、カテゴリ平均の差異を見ると他と比較して（0.53）と変化が小さいことが顕著に表れた。

このことから、カテゴリ1（データにまつわる情報収集）とカテゴリ2（データを可視化する）については、事前に想像できた問題を実践を通して実際に認識し、自分なりに解決できたと感じた参加者が多いことがわかった。一方で、カテゴリ3（可視化をレビューする）、カテゴリ4（作品を組み立てる）については、事前に想像できた問題について実践を通じて認識することができたが、解決には至らなかったと感じた参加者が多いことがわかった。

以上の理由を考察する。カテゴリ1とカテゴリ2のパターンは、自分自身の作業の中から生じてくる状況が多く、参加者は記述された状況を認識し実践する機会を多く得たと考えられる。一方、カテゴリ3のパターンは、自分以外（チームメンバーや専門家など共同作業）との関わりから生じる状況を記述したパターン（21）や、レビュー以降の反復の中で現れる状況を記述したパターン（19,20,24,25）である。ここでワークショップを振り返ると、このワークショップでは途中でレビューの機会が一度しかなく最終レビューまでの期間も短い。また、多くの作業を参加者がそれぞれの時間に単独で進めてきた。そのため、参加者にとってカテゴリ3のパターンに記述された問題の多くは、その問題が生じなかったか、生じても実践して解決できると言える手応えを感じる事ができなかったと考えられる。しかしその中でも、レビューに臨むための準備に関するパターン（22,26）やレビュー直後の方策に関するパターン（23）など、前述の状況においても生じた問題については解決できたと感じる参加者が多かったと考えられる。

カテゴリ4はExploringの結果をまとめ上げる段階であるが、前述の通りレビューの機会が少なく期間も短いためこの段階を十分に経験できず解決に至らない参加者が多かったと考えられる。しかし、制作のポイント（7.3.3項を参照）で挙げた「多様なあり方が感じられる表現」や「多面的な視点でデータを眺めることができる」を志向した問題を扱

うパターン（27,28,29,31）については、事後の平均の差異が他のパターンよりも向上していることから、作業の過程で少なくとも問題を認識できたと考えられる。一方、鑑賞者へ提示するための作品準備の問題（30,32,33）は特に変化が少なく、実際に鑑賞者を想定して作品としてまとめる状況に至らない参加者が多かったと考えられる。

7.6.3 各パターンについての参加者の感想

事前アンケートの設問 2 に対して 2 名の回答があった。これを以下に示す。この設問は回答を必須としなかった。末尾の括弧内の記号は回答者を示す。

- 全体的に様々なシーンが想定されて助言があり、自分自身でブラッシュアップするのに重宝すると感じた。29 で記されていた“適度なハードル”について具体的な例が思い浮かばず十分に理解できなかった。(M)
- 鑑賞者のスタンドアロンレベルに応じて Assemble 段階の思考範囲はだいぶ変わるように感じたので、そこの記述もぜひリファレンスとして読みたいと思った。(T)

M の回答から、パターンと自分の経験を照らし合わせ「自分自身でブラッシュアップする」ことが期待できると予想し、一方で、一部のパターンはイメージできないと感じている様子がうかがえる。T の回答から、「鑑賞者のスタンドアロンレベル」（筆者の解釈では、鑑賞者が独力で意味を読み解く力）と「Assemble 段階の思考範囲」（筆者の解釈では、Assemble 時の制作者が考慮すべき事柄）に関連があることをイメージし、参考になる知識を得たいと感じている様子がうかがえる。

次に、事後アンケートの設問 2 と 3 の回答から各パターンについて参加者が言及した箇所を抜き出した。先頭の括弧内の数字は言及されたパターンのナンバーである。末尾の括弧内の記号は回答者を示す。なおカテゴリー 0 とカテゴリー 4 のパターンに言及はなかった。

カテゴリー 1 について

- (8) 事前に読んでおくことで陥りそうな罠に気をつけることができ、時間ロスの削減になった (U)

カテゴリー 2 について

- (10,16) パターンを読む前から自分が心がけていたが、これが載っていると自信に繋がった (O)
- (14) できるだけ細かく表示できるように、描画プログラムについては本当に工夫して効率化した。結果的に高解像度の可視化を行うことができた (K)
- (16) 事前に読んでおくことで陥りそうな罠に気をつけることができ、時間ロスの削減になった (U)
- (16) 可視化のスクリーンショットは全て残しておいた。今までの可視化を見比べることで、どこがどう良いのかを客観的に考察することができた (K)
- (16,17) 開発初期段階から意識することで、スムーズな開発ができた (N)
- (17) 可視化制作の際、パラメーターは全てインタラクティブに調整できるようにして制作を行なった。それのおかげで、スイートスポットを見つけることができた (K)

カテゴリー 3 について

- (20) アフリカのデータがうまく取得できていない現状を把握できた (N)
- (23) 行き詰まりを感じた時に状況に応じて、読むことでやるべきことが明確になった (O)
- (24) 可視化がいったん完成した際、自分が思っていたよりも良い可視化ではない時があった。その時にこのデザインパターンを思い出し方向性を思い切り変えることができた (K)
- (25) 事前に読んでおくことで陥りそうな罠に気をつけることができ、時間ロスの削減になった (U)
- (25) 開発初期段階から意識することで、スムーズな開発ができた。前処理で json に書き出すことで、違うデータを同一プログラムで見比べられるようにしたことで、データ間の比較が楽になった (N)

7.6.4 制作中にパターンを参照する機会が少なかったケース

Y は事後アンケートの設問 2、3 について、制作中にパターンを参照する機会が少なかったと回答した。Y の回答を引用する。

- 事前にデザインパターンに触れたことが作品制作に影響していると考えられるが、制作中に意識的に参照する機会はあまり無かった。(設問 2 より)
- デザインパターンは問題が起こったときに引くつもりでいたが、テーマが決まってからはあまり幅広く他の可能性を検討することもなかったので、スムーズに制作が進んだ印象。自分のアイデアが全て実装し終わった時点で、よりよくするためのインスピレーションを得るのには使えるかもしれない。(設問 3 より)

Y は、事前にパターンを読んだことは役に立ったと考えているが、事後のアンケートでは全てのパターンについて B または C を選択し、事前、事後ともに D を選択したパターンはなかった。つまり、振り返ってみれば認識できた状況や問題があったにもかかわらず、制作中はあまりパターンにとらわれずに制作していたことがうかがえる。Y は構想発表の段階で、COVID-19 についてデータよりも自分の感覚について語っている。それは、Y が COVID-19 の期間中の多くを地方都市で過ごした経験と報道に見るパンデミックのギャップについての語りであり、そのような自分の主観から立ち上がるパンデミックについての問いや関心を中心として制作したいという動機の表明であった。そして Y が経験した、感染が広がっていない地方都市と逼迫した状況となった大都市の対比がテーマとなり、テーマにマッチするデータと表現を模索していった結果が作品となった。このように自分の経験に根ざした動機を持ってデータをリサーチし可視化した参加者は Y 以外にはいなかった。このようなアプローチは 2.5 節で述べた「コミュニケーションメディアとしての情報可視化」であると言え、Kirk のモデルのように編集的思考が働き、テーマに沿った表現の選択が行われたと考えられる。そのことは「テーマが決まってからはあまり幅広く他の可能性を検討することもなかった」という Y の言葉にも表れている。表現を広く探求する必要が薄れたことにより、Y がプロセスを進める上ではパターンを参照する機会が少なかったと考えられる。

7.6.5 パターンに記述した状況ではない状況でパターンを活用したケース

M はパターンに記述した状況以外でパターンを活用したことを述べた。M の回答を引用する。

1. パターンは個人で黙々と作業するとき、自分自身でアイデアのブラッシュアップをするときによりデザイナーの助けになると感じた。
2. 実際に問題に直面した際に辞書のように使うだけでなく、それ以外のパターンを読み込んでアプローチの幅を広げるのに役立った

3. 今回の可視化に取り組む前では具体的なシーンを思い浮かばないパターンも、実際に取り組んでみてより問題の理解が進み、デザインの指針として非常に助けられた

1の回答は、特に問題となる状況ではなく作業が進行している状態でもパターンから良い影響が得られることを示唆している。

2の回答は、問題に直面したときに、その問題について記述されたパターン以外のパターンからも良い影響が得られることを示唆している。これはパターン同士が関連し、構造を成していることが効果を生んでいると考えられる。

3の回答は、パターンを事前に読むことで未経験のパターンについても意識でき、指針に良い影響を与えることを示唆している。このことは、パターン・ランゲージの制作目的の一つである制作経験がない場合の学びの支援（6.1節参照）に結びつく可能性がある。

7.6.6 専門家を交えたレビューについての言及

専門家を交えたレビューについて参加者が言及した箇所を抜き出した。

- 専門家の方の意見によって、可視化の方向性を大きく変える決心がついた。また、自分では知らなかったような新しい知見をたくさん得ることができ、可視化制作にとっても役立ったと感じた (K)
- 専門家と複数回ディスカッションができたので、改善点やその他のお話を聞いてからもう一度考え直すことができ、より有意義な作品を作ることができたと思う。(U)
- 専門家のフィードバックは、このまま可視化を進めていくことについて問題がないのかどうか、といった不安要素を取り除く上で大いに助けられたと思う。また、専門家の方が非常に協力的で、可視化の上での代替案や更なるアイデアを提示してくださり、とても参考になった。(Y)
- 専門家の方を交えたレビューではそれまでなかった新しい見解を知ることができたため、より質の高いブラッシュアップができたと感じましたし、データをデザインする上で、可視化のアプローチを大きく左右する“データへの理解度”の重要性について改めて学びました。(M)
- 調べてもわからないようなことをたくさん教えていただいた上、自分がデータから読み取った背景などの答え合わせができたため、その後のブラッシュアップの際により強調できる見せ方を考えることができた。(N)
- 全く知らない情報を得るきっかけとして、自分で情報検索ができるレベルの視座が

得られた。(T) (筆者註：意識していなかった視点を専門家の指摘によって得られ、自分で調べる際の手がかりが得られた。)

- 専門家の方の意見を可視化に反映することができたため良かった (O)

全ての参加者が専門家を交えたレビューは効果的と感じたことがわかった。参加者自らがリサーチしたり表現した中にはない専門家ならではの視点が得られたこと、可視化を通じて自分が得た意味に確証が持てたこと、それらを次のステップに反映させることができたことが、その主な理由であった。ワークショップ全体の期間に比べて専門家のレビューは短時間であり、効果的な指摘を引き出す機会は限定的である。その中でも効果を感じられたことから、参加者自身がディスカッションに耐えうるリサーチとプロトタイプの実装、その言語化に達していたと考えられる。加えて、レビューの方法 (7.3.4 節) の影響も考えられる。また、専門家が協力的であったという回答についての理由には、専門家自身が今回のワークショップの趣旨を的確にとらえ、その意義について深く理解していたことも挙げられるだろう。

7.6.7 ワークショップの構成についての言及

ワークショップの構成について参加者が言及した箇所を抜き出した。

- パターンに含まれている聞き込み調査がワークショップ内に組み込まれてある点が良かった。(O)
- 「データビジュアライゼーション・デザインパターン」があることでワークショップの流れが明確でわかりやすかった。(U)
- レビューが2回あったことで、1度目のレビューで得られた専門家の方の意見を可視化に反映することができた (O)

今回のワークショップのように限定された環境でパターンを運用する場合、どのようにすれば参加者が実際のプロジェクトに起こりうる状況を体験できるかが課題となる。上記の回答は、ワークショップの計画がパターンの活用の効果と強く関連していることを示唆している。

7.7 まとめ

本章では、開発した Data-Viz デザインパターンの小冊子を利用しながらデータビジュアライゼーションを制作するワークショップを実施し、参加者の作品制作プロセスを観察

した。

選択肢式アンケートの結果から、参加者は Data-Viz デザインパターンを読みプロセスの全体像をイメージでき、各々のパターンに描かれた状況の多くを想像できたと考えられる。

自由記述式回答の結果から、参加者はパターンを読み実践することで、未経験の問題を認識でき、解決の糸口を掴むことができたと考えられる。

参加者が単独で進められるフェーズ（Diving と Exploring の可視化）では解決の手応えが強いものの、専門家との対話が必要なフェーズ（Exploring のレビュー）では問題の認識までにとどまるケースが多く見られた。その理由として、専門家のレビューの機会が少なく、また、レビューの内容を反映させ更なる探求を繰り返す期間が短かったためと考えられる。

パターンに記述した状況ではない状況でパターンを活用したケースが確認できた。具体的には、個人が黙々と作業するときにアイデアをブラッシュアップするために活用したケースや、問題に直面した際に、その問題以外のパターンを読み込むことでアプローチの幅を広げたケースがあった。これらは、問題となる状況ではなく、作業が進行している状態でもパターンから良い影響が得られることを示唆している。その問題について記述されたパターン以外のパターンからも良い影響が得られる理由は、パターン同士が関連し構造を成していることが効果を生んでいると考えられる。

ワークショップの設計やレビューに専門家を招聘したことが効果的だったとの回答が多く見られた。このことから、適切なワークショップ設計とともに Data-Viz デザインパターンを運用することで、実践を通した学びを深められると考えられる。

以上の結果から、Data-Viz デザインパターンは制作経験がない場合の学びの支援及び、制作過程における支援に有用であると考えられる。ただし、レビュー回数不足や、展示の機会を設けられなかったことから Assembling のステップの検証は十分とは言えない。これらを考慮してワークショップを計画することで、さらなる学びの機会を得られる可能性がある。

第 8 章

実務での運用と考察

8.1 はじめに

前章では、本研究の目的に沿って計画され制御されたワークショップにおいて、実務経験のない学生を対象に Data-Viz デザインパターンの影響を観察した。対して本章では、制御されない実務環境において、実務経験のあるデザイナーにとって Data-Viz デザインパターンがどのように機能するかを考察する。

8.2 前提と手法

本章では、メーカーで働くデザイナー N に協力を依頼し調査を行った。

N は入社 5 年目のデザイナーである。プログラミングの経験年数は約 4 年 (JavaScript, C#)、データビジュアライゼーションの経験年数は約 3 年である。また、年間の業務のうち 4 割程度は製品やサービスの User Interface (UI) のプロトタイプをデザインし、その一環としてデータを画面上で視覚的に表現するためのコーディングを行なう機会がある。

N が所属する企業では、スマートフォンアプリを介した個人向けの無料情報サービスを運営している。同サービスは、全国の特定の公共施設を中心にその周辺の街を取材した情報を記事として提供する。また、施設の検索機能、ユーザー自身が興味を持っている情報ジャンルを登録する機能等を有している。ユーザーは全国に分布し、施設や記事を閲覧し、お気に入り登録機能などで情報を収集できる。同サービスの運営部門では、施設事業者や観光協会などとの協業により、地域の盛り上げや PR のための記事の提供を行なっている。そのため、記事の閲覧やお気に入り登録などユーザーのリアクションに注目して

いる。

N は、運営部門から提供された施設の位置情報や記事の配信記録、匿名化されたユーザーの利用ログなど、多数の変数を含む時系列データを使用して、運営部門との対話を重ねながら問題発見と美観の両面に主眼をおいたビジュアライゼーションをプロトタイピングするプロジェクトに従事した。その際の言語は JavaScript を使用した。

N および所属企業の上承を得て調査の趣旨を説明した上で、Data-Viz デザインパターンを提供し、制作過程におけるエピソードの記録を依頼した。N にはプロジェクト開始前に Data-Viz デザインパターンを読了するように依頼し、読了時に事前アンケートを実施した。開始から約 3 ヶ月間に制作したプロトタイプと、事後アンケート、インタビューにより、Data-Viz デザインパターンの利用状況やプロセスの実態を調査した。

事前と事後のアンケートでの設問はいずれも前章で使用した内容と同じである。ただし、事後アンケートの設問 2（ワークショップに対する感想）は本調査の趣旨に合わないため省略した。

8.3 制作したデータビジュアライゼーション

本節では、N が制作したデータビジュアライゼーションのプロトタイプについて述べる。これは、あくまでも本研究の調査を終えた段階の成果物であり、展示や鑑賞を目的として完成させた作品ではないことを書き添える。

図 8.1 は地図平面に対して鉛直方向に時間軸（地表側が過去）を割り当て、ユーザーが記事を閲覧した時刻とその記事に関連する施設の緯度経度から、全ての閲覧記録を 3 次元空間にマッピングし、その密度分布をボリュームレンダリング（濃い青色の雲状部分）で表現している。可視化する密度の閾値はスライダーで変更できる。また、空間に点在するシアンのドットは、記事の作成時刻とその施設の緯度経度を元に空間内にプロットしたものである。画面左側には記事のタイトルとその記事の閲覧数をバーチャートで表現している。

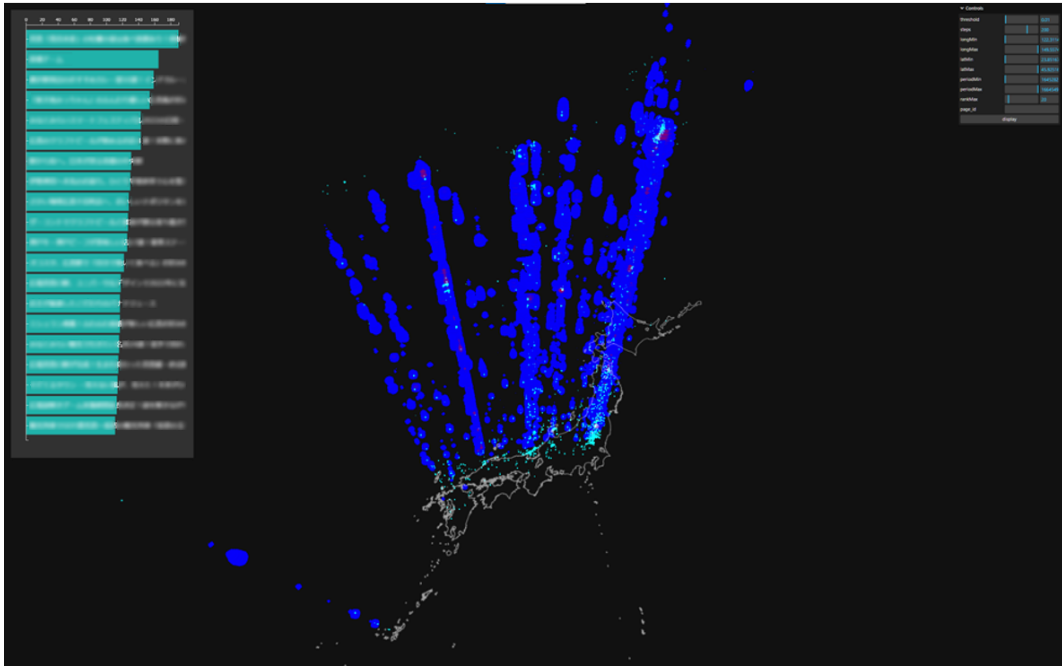


図 8.1 N のプロトタイプ

右側のスライダーで、経度、緯度、時間の範囲を指定することで、雲状部の一部だけを可視化し、また左側のチャートも該当する記事のみに絞り込まれる（図 8.2）。また、記事を示すプロットをクリックすることで、その記事に対する閲覧行動だけを可視化することができる（図 8.3）。

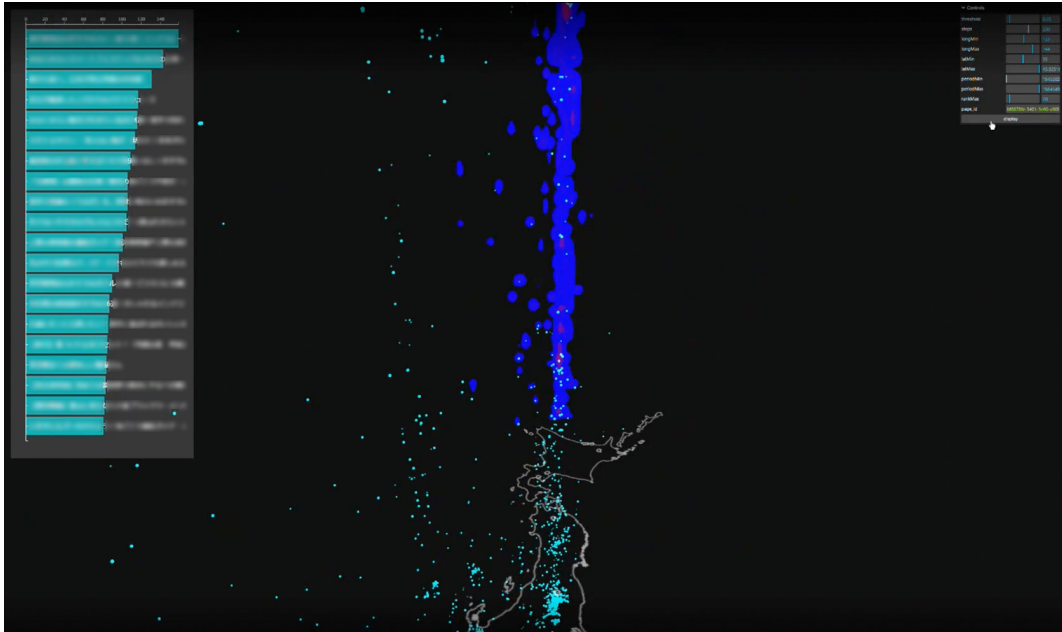


図 8.2 緯度、経度、時間の範囲を指定して該当部分のみ表示した



図 8.3 特定の記事をクリックして絞り込んだ状態

8.4 アンケート結果と考察

8.4.1 事前・事後回答の結果

事前と事後の設問1は、パターン・ランゲージの各パターンについての認識を問うものであることは先述した通りである（7.4節参照）。

各パターンについての回答結果を事前と事後で比較し、図8.4に示す。図では、各パターンについて、A（このような問題が起こる場면을想像できない）、B（このような問題が起こることを知っている）、C（このような問題を実際に認識したことがある）、D（このような問題を解決したことがある）の回答の分布を示した。前章と同様に、A、B、C、Dをそれぞれ1点、2点、3点、4点とし事前と事後の各パターンの得点の差異で変化を表した。

パターン名	事前アンケート結果				事前アンケート結果 カテゴリ平均	事後アンケート結果				事後アンケート結果 カテゴリ平均	事前・事後の差異		
	A	B	C	D		点数	A	B	C		D	点数	点数の差異
カテゴリ0	1. 「わかる」より「感じる」可視化		■		3	3.00		■		2	2.00	-1.00	-1.00
カテゴリ1	2. 周辺リサーチ		■		2	2.86		■		2	2.86	0.00	0.00
	3. データ化の背景		■		2			■		3		1.00	
	4. データの権利			■	4			■		2		-2.00	
	5. 変数の意味を掘り下げる			■	4			■		2		-2.00	
	6. 可視化の切り口を掴む		■		2				■	3		1.00	
	7. 聞き込み調査		■		2				■	4		2.00	
	8. 予想のスケッチ			■	4				■	4		0.00	
	カテゴリ2	9. 視覚にさらす			■		4	3.40		■			
10. ロケットスタート				■	4				■	4	0.00		
11. 全体と細部を行き来する				■	4				■	3	-1.00		
12. 全体を眺める				■	4		■			2	-2.00		
13. 厳密さのジレンマ		■			1		■			2	1.00		
14. ディテールを活かす				■	4				■	4	0.00		
15. 思い込みの罨				■	3				■	4	1.00		
16. レシピを残す				■	3		■			2	-1.00		
17. 見ながら調整				■	4				■	4	0.00		
18. 直感に従ってみる				■	3				■	4	1.00		
カテゴリ3	19. 科学的な物語		■		2	2.25		■		2	2.13	0.00	-0.13
	20. 事実に照らす		■		2			■		2		0.00	
	21. 客観的な批評で磨く		■		2			■		2		0.00	
	22. 言葉によるガイド		■		2				■	4		2.00	
	23. 探索のチェーン		■		2		■			1		-1.00	
	24. 思い切り変える			■	4			■		2		-2.00	
	25. 更新の価値		■		2			■		2		0.00	
	26. 物語のコレクション		■		2			■		2		0.00	
カテゴリ4	27. 物語を束ねる		■		2	2.29		■		2	2.29	0.00	0.00
	28. 多面性の組み合わせ		■		2			■		2		0.00	
	29. 適度なハードル		■		2			■		2		0.00	
	30. アウトプットからの逆算		■		2			■		2		0.00	
	31. 物語にそって外見を磨く			■	4			■		2		-2.00	
	32. 鑑賞のインターフェース		■		2			■		2		0.00	
	33. 同じ環境で見る		■		2				■	4		2.00	
					2.76						2.61	-0.15	

図 8.4 設問 1 について事前・事後の回答と比較

回答の分布を比較すると、事前・事後ともに回答が多い順に B（事前 52%、事後 58%）、D（33%、27%）、C（12%、9%）、A（3%、3%）となった。また、カテゴリ毎の点数の平均を事前と事後で比較すると、ほとんど差異が認められなかった。また、事前には D を回答したパターンが、事後では B や C になるなど、差異が負となるケースが複数認められた。

8.4.2 事前・事後回答の比較と考察

差異が負になるケースについて、事後のインタビューで次のような聞き取りを行った。以下では、その回答を示す。

事後アンケートで、No.5の「変数の意味を掘り下げる」について、B「起こることを知っている」と回答した理由について

回答 今回の可視化で取り組んだ内容に関しては、変数については解釈違いが起こらないシンプルなデータだった。そうでない場合では問題が起こることが想像できた。

事後アンケートで、No.12の「全体を眺める」について、B「起こることを知っている」と回答した理由について

回答 記述されている「事前に収集した情報だけに頼ると、そこで明らかになっていない注目すべきポイントを見落とす。」そういうことはあるんだろうな、でも今回はそういうことはあまりなかったなと思った。

以上の聞き取りから、Nは回答するにあたってこれまでの全ての経験を振り返ってはならず、直前に経験したプロジェクトだけを振り返って回答したと考えられる。また、パターンを運用した結果、事後にその問題への認識が変化した可能性も考えられる。そのため、単純な点数の前後比較からパターンの影響を推測することが難しいことがわかった。

一方、事前・事後の回答を比較せず個別に検討すれば、選択肢A（想像できない）の回答はごく僅かであり、各々のパターンに記述された状況が経験者が実務の上で想像しやすい内容であると考えられる。

8.4.3 自由記述での回答

事前アンケートの設問2での回答を示す。

「状況」「その状況において」にあるある！と思うものが多かったです。Assemblingのフェーズはまだ到達したことがないものが多かったですが、いずれはそういう壁にぶつかることがあるのか、と心の準備ができました。

事後アンケートの設問3での回答を示す。

可視化を進めている中で、次どうしたらいいんだろう？と手が止まったときに参考にさせていただき、ヒントを得て一歩先に進むきっかけを得ることができました。例えば、初期にスケッチしていたもの形になったとき、次に何に手を付けていいか迷いました。そんな時にパターン・ランゲージを拝読し、07聞き込み調査か

らヒントを得て、データ提供元の担当者にヒアリングし、データからは読み取れなかった外的要因や、担当者目線での可視化したいポイントを聞き取りました。そのおかげで、可視化の次のステップを思いつくことができ、先に進めることができました。また、可視化がある程度進んで再度担当者と打ち合わせる際、22 言葉によるガイドを参考に、予め自分で気になるポイントを文章にまとめて共有しました。前回の聞き込み調査では、担当者も何にコメントしたらいいかわからず、見た目的な視点 (特に UI に関わる部分) のコメントから始まってしまい、うまく聞きたいことを導き出せませんでした。しかし、2 回目は文章化したことでコメントがほしい部分を明確化でき、担当者もコメントしやすかったように思えます。

8.4.4 自由記述での回答の考察

事前アンケートから、N は各パターンに記述された内容を読み自身の記憶と対比しながら思い当たる経験を想起した様子が見える。また、自身が経験したことのない状況についてイメージを掴んだ様子が見える。

事後アンケートから、N がパターンを参照しながら立ち現れた問題の本質を掴み、具体的な方策を立て解決した様子が見える。また、初期段階において進め方に戸惑った様子を「次どうしたらいいんだろう? と手が止まったとき」と表し、次のステップを決めるために参考にした上でヒアリングを実施していることから、DEXAS モデルを参照して行動を選択したと考えられる。

8.5 作業の記録と考察

8.5.1 パターンを活用したエピソード

N の作業記録から、パターンを解釈し取り組んだエピソードを時系列に沿って抜き出した。

No.8 「予想のスケッチ」浮かんだアイデアを、スケッチで外部化する

考えていたものを絵に起こしてみる。地図平面と時間軸の 3 次元空間の散布図をスケッチすることで、言葉で考えている時よりも具体的に想像でき、必要な機能を整理することができた。

No.9 「視覚にさらす」自分の視覚にさらすことで、心に浮かぶものを捉える。

多くの変数を一度にひとつの空間に表示してみることを試した。複雑になりすぎたため、プロットにマウスで触れることで情報を表示できるようにした。また、時間軸に1ヶ月毎のグリッドを表示し目安をつきやすくした。その上で、特定のユーザの活動をプロットを線で繋ぐことで表現してみることや、曜日毎に集計してプロットし、休日と平日の変化を表現してみるといった多角的な見方を検証した。

No.7 「聞き込み調査」専門家とのパートナーシップは、プロジェクトの質を高める。

プロジェクトメンバーへ可視化を見せたうえでの聞き取りを実施した。運営ユーザのデータが混じることで印象が変わることがわかり、一般ユーザとデータを分けた表示へ更新した。

No.11 「全体と細部を行き来する」同じデータでも、マクロに見る時とミクロに見る時には別の物語を発見できることがある。

全体を眺めることはできたが、細部への意識が向いていないので、もっとクローズアップした状態でデータを眺めてみようという次のアクションにつながった。

No.14 「ディテールを活かす」細部をできるだけ潰さずに、全体像を眺める。

8K に映して眺めてみたら新しい発見がないかと着想し、検証した。

No.22 「言葉によるガイド」言語化が不十分だと、見た目に流されやすい。

プロジェクトメンバーへプロトタイプを見せながら聞き取りを実施した際、見た目やUIについての指摘に偏り、議論が深まらなかったと感じ、2回目の聞き取りでは見てほしいポイントを言語化したメモを用意した。

DEXAS について

現状では、Diving と Exploring に該当する部分は DEXAS に沿って進めることができたが、Assembling については実施できておらず、知見が足りていないと感じる。今後は社内外の展示や発表を設けるなどの機会が実現した際には参考にしたい

と感じた。

8.5.2 エピソードの考察

本プロジェクトは、Nがプロトタイプを繰り返し行い、多角的な表現から捉えた解釈を運営メンバーにレビューすることで確認しながら、その結果をプロトタイプに反映するプロセスを繰り返していることがうかがえる。No.8では初期アイデアをスケッチすることでイメージを確認し機能を整理している（図8.5はNが制作したスケッチ）。

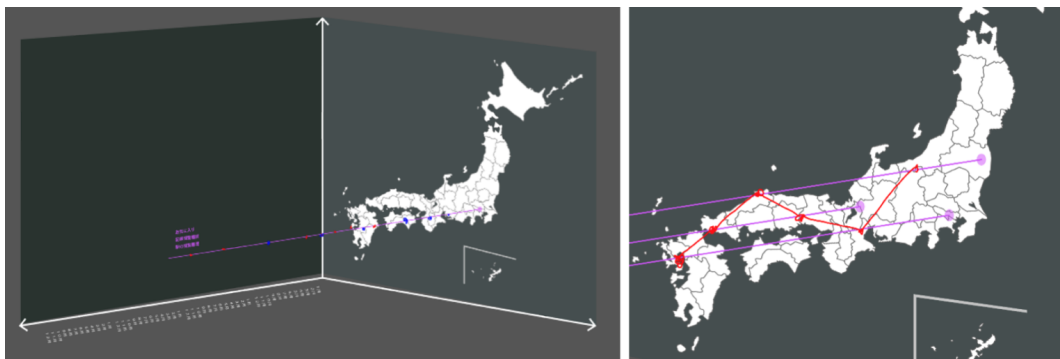


図 8.5 Nが制作した初期スケッチ

その後、No.9では多くの変数をひとつの空間で可視化することを試している（図8.6）。また、特定のユーザーの行動や曜日ごとの変化（図8.7）など視点をさまざまに変えながら多角的に試行した様子もうかがえる。その後を実施した運営メンバーへのレビューで得られた情報をもとに、データをフィルタリングし可視化に反映させている（No.7）。以上のステップは、DEXASモデルのうち、DivingとExploringの最初のレビューまでのプロセスに相当する。

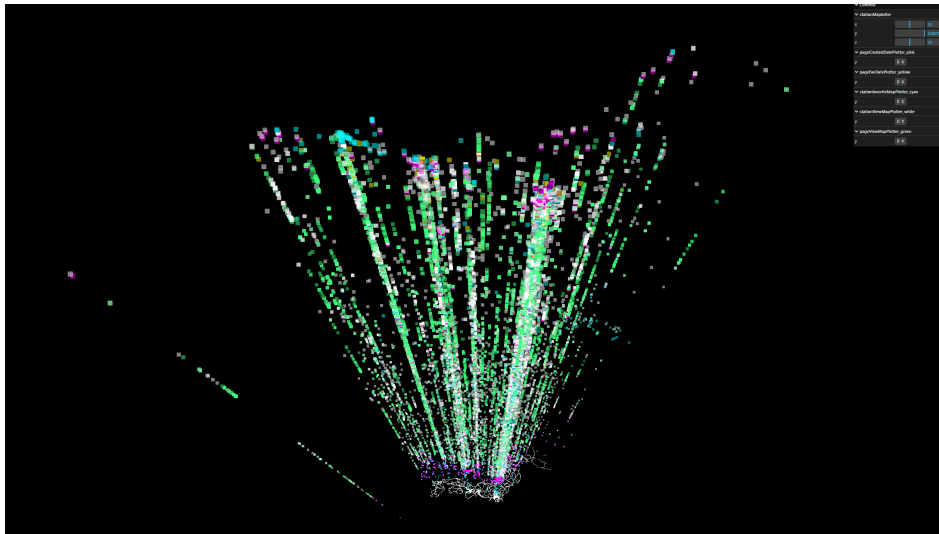


図 8.6 プロットの色に多様な変数を割り当てている（ピンク：記事作成日、シアン：施設お気に入り登録日、黄色：記事お気に入り登録日、白：閲覧日）プロットにマウスを当てると詳細情報が表示される。

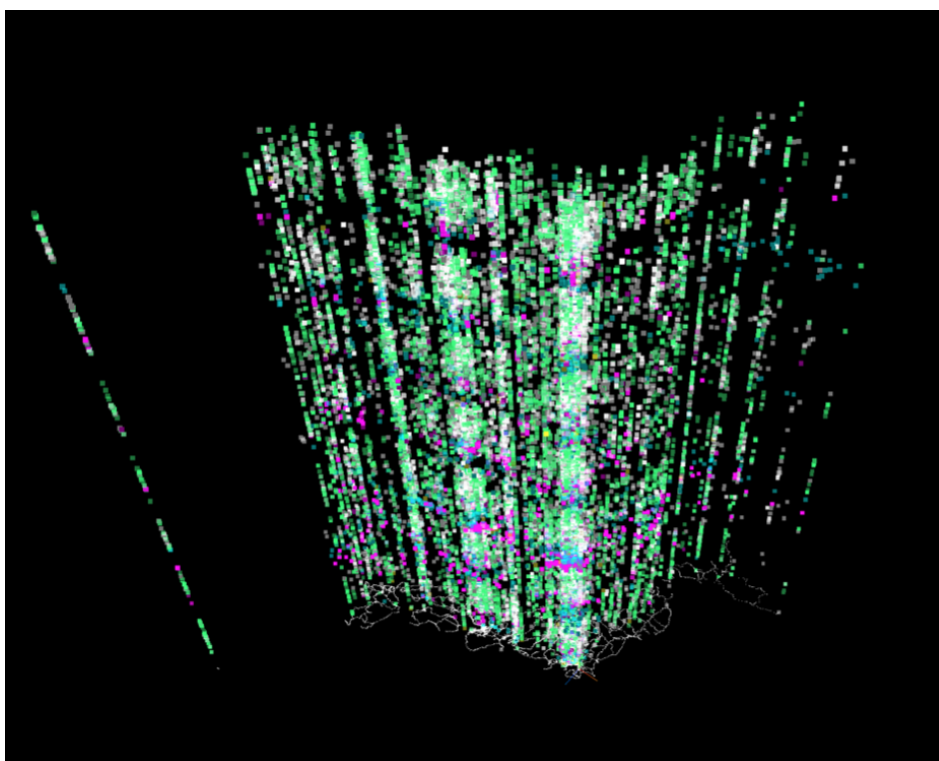


図 8.7 1週間ごとに重ねて表示し曜日ごとの変化を観察している

No.11、No.14 は更なる意味の発見を目指し、データの眺め方を工夫している。No.22 では、2 回目のレビューに向けて 1 回目のレビューの問題点を振り返り、パターンを参照して改善を目指している。

以上から、N がパターンを活用して複数のデータビジュアライゼーションを作成し、関係者との対話の結果を取り入れながら DEXAS のステップに沿ってプロトタイピングを進めたことが確認できた。ただし、本調査期間内に Assembling のステップは実施できなかった。

8.6 まとめ

本章では、Data-Viz デザインパターンが実務においてどのように機能するかを考察した。

調査開始時と終了時の選択肢式設問の結果から、各々のパターンの記述は、デザイナー自身が過去の経験に照らし合わせて想像しやすい内容であると考えられる。

事前の自由記述式設問の結果から、DEXAS モデルとパターンの内容を読むことで、デザイナー自身が未経験のプロセスについて、その全体像をイメージできることが示唆された。

事後の自由記述式設問の結果から、デザイナーがパターンを参照しながら立ち現れた問題の本質を掴み、具体的な方策を立て解決したことが確認できた。また、プロジェクトの初期段階で DEXAS モデルを参照して、次に行うべき対応を適切に選択したことが示唆された。

エピソードの記録から、デザイナーがパターンを活用して複数のデータビジュアライゼーションを試作し、関係者との対話の結果を取り入れながら DEXAS のステップに沿ってプロトタイピングを進められたことが確認できた。ただし、Assembling のステップには至っておらず、この段階での活用は確認できなかった。

以上の考察から、Data-Viz デザインパターンは実務における運用において、デザイナーのプロトタイピングを支援できることが確認できた。一方、調査したプロジェクトではプロトタイプを作品として完成させる段階には至っておらず、この段階での効果を考察するには更なる調査を要する。

第9章

結論

本章では、本研究の議論をまとめる。

9.1 本研究のまとめ

序論では本研究の背景として、今日の我々を取り巻くいわゆるビッグデータとそれを利用するデータ駆動型社会のあり方について公平性の問題を取り上げた。データから何らかの意味が取り出され社会に活かされる営みの中で、複雑で大規模なデータほどその意味が取り出されるプロセスが不透明になり、多様な意味解釈に一般の人々がアクセスすることはできない。そこで本研究では大規模で複雑なデータや専門的なデータを可視化し、一般の鑑賞者にとって現象を多様な視点から眺め、洞察へと導くことができるデータビジュアライゼーションに注目した。情報可視化やデータビジュアライゼーションは、コンピュータの登場以降、計算機科学の Human-Computer Interaction (HCI) 領域において幅広く研究され、現在では、複雑なデータを扱うタスクの効率化を必要とするさまざまな分野に普及した。一方、観測や計量によって得られる数の集合から複雑な自然や経済のメカニズムを理解しようとした科学革命や統計学の興隆以来、数の視覚的表現はグラフやチャートの発明のみならず、新たな世界像の獲得を目指して人間が模索してきた経緯があることを述べた。さらには情報爆発の時代を迎え、固定的なメディアから動的でパーソナルなメディアへ移行すると、一般の人々に対して文字や画像や動画だけでなく情報可視化技術を駆使してストーリーを伝える、インフォグラフィクス、データ・ジャーナリズム、データを使用した芸術など、コミュニケーションとしての情報可視化も活用されている。このように、現在我々が触れる情報可視化やデータビジュアライゼーションは、複数の文脈を持ちその用途も多様であるため多義的な状況となっていることから、序論ではその流れを整

理した。その上で、今後より一層複雑化・大規模化することが予見されるデータの多様な意味解釈の可能性を探るという観点から、本研究の目的を、デザイナー、アーティストなど視覚表現を行うクリエイターによる、大規模なデータを扱う情報可視化メディアのデザインを支援することを前提として、データビジュアライゼーションの制作プロセスを形式知化することに定めた。

2章では、情報可視化の制作プロセスに関連する既往研究を述べ、本研究の位置付けを明確化した。HCIの領域で議論されてきた問題解決型の情報可視化システム開発フレームワークと、データ・ジャーナリズムやインフォグラフィクスなど、一般に向けたコミュニケーションを目的とした可視化開発フレームワークについて検討し、本研究が目指すプロセスモデルとの差異を論じた。本研究が目指すグラフィックデザインに精通したエンジニアが開発に深く関与するプロセスとして Ben Fry が提唱した「Computational Information Design (CID)」に注目し、その成果物とプロセスを検討した上で問題点と未解明な点を議論した。

3章では、本研究の手法を述べた。手法は次のとおりである。グラフィックデザイナーでありエンジニアでもある筆者が実際のプロジェクトで CID を適用し、プロセスを記録したデータを分析することにより形式知化を行う。形式知化の手法は、プロセスのモデル化とパターン・ランゲージの制作を選択した。それらをデザイナーやクリエイターが読むことで援用可能な小冊子「データビジュアライゼーション・デザインパターン」にまとめ、ワークショップを実施して検証を行うものである。

4章では、筆者が観察したプロジェクトに CID を適用した結果を報告し、プロセスを分析しモデル化した。

5章では、筆者が観察したプロジェクトに CID を適用した結果から、パターン・ランゲージを制作した。

6章では、データビジュアライゼーション・デザインパターンについて制作目的の明確化、想定読者や構成を検討し成果物を報告した。

7章では、デザインやアートを学ぶ学生を対象としたワークショップを企画し、データビジュアライゼーション・デザインパターンを運用した結果を報告した。

8章では、データビジュアライゼーション・デザインパターンが経験のあるデザイナーの実務において機能することを報告した。

本研究の成果は次のとおりである。CID を適用してデータビジュアライゼーションの制作に取り組んだ筆者の参与的観察の分析結果から、データビジュアライゼーションのためのデザインプロセスモデル「DEXAS」を考案した。また、DEXAS を運用する上での経験則を共有し制作物の質を高めるためのパターン・ランゲージを制作した。この二つの

形式知をまとめたデータビジュアライゼーション・デザインパターンを執筆した。これはデータを多様な視覚表現に変換するスキルを持つクリエイターやデザイナーを想定読者として、彼らのデータビジュアライゼーション制作を支援し動機を高めるとともに学び方の指針になる図書である。制作の経験が少ない学生を対象としたワークショップでの運用結果からその有用性を示した。また、経験のあるデザイナーがデータビジュアライゼーションのプロトタイピング実務において同書を利用してプロジェクトを進めた結果から、その有用性を示した。

9.2 今後の課題と展望

9.2.1 実践の機会創出と認知の向上

本研究では、データビジュアライゼーション開発プロセスのモデルを提案し、成果の質を高めるための実践知をパターン・ランゲージとして記述する手法を提案した。それらを編纂した小冊子「データビジュアライゼーション・デザインパターン」（7章参照）では3つの目的を掲げた。すなわち、1. 制作経験がない場合の学びの支援、2. 制作過程における支援、3. 制作を経験した後の学びの支援である。1及び2については本研究のワークショップを通じて有用性を確認できたが、3についての検証は残されている。3の効果を検討するためには、データビジュアライゼーションが単発で一回性のプロジェクトではなく、継続的な取り組みとして中長期に学ばれ実践されるように社会の中で認知を獲得していく必要もあるだろう。筆者は本論文の執筆時点において、国内の大学及び大学院にてデータビジュアライゼーションの実践的な指導を目的とした講座を担当している。データサイエンス教育需要の高まりから毎年各大学から新規の講座開設の打診があるものの、とても全てをお引き受けすることはできず断腸の思いでお断りせざるを得ない状況である。また企業や行政関係からの講演依頼も確実に増えており、本研究の成果を基盤として認知を拡大し、実践者を増やし支援することでさまざまな実践知が集まり本研究をより発展させていくことにつながると考える。

9.2.2 創造的なプロセスモデルはどうあるべきか

本研究を論文としてまとめるにあたり、そのタイトルに「創造的」を加えた。ここで改めて、創造とは何か、創造的なプロセスモデルはどのようにあるべきかについて議論したい。

「創造的なデザイン」というフレーズは今や日常に溢れかえっている。その背景には、

昨今の厳しいビジネス競争環境において新たな付加価値や斬新なビジュアルコミュニケーション戦略、顧客体験の創出、市場の拡大といった生存戦略を勝ち抜くために、デザインに寄せる期待と願望が垣間見える。特に我が国においては、「創造的」とは何かこれまでとは違う洗練された商品やアイデアを指す意味で使われることが多く、カタカナの「クリエイティブ」に至っては、そのような物や物を生み出す職能自体をまるで個人の才能や熟練の技のような意味に用いられることもある。

本研究で述べたデータビジュアライゼーションやデータアートの優れた作品の造形や様態を表層的に捉えれば、それらは確かに前述の意味での「創造的」な産物と言えよう。しかし本研究の創造的とはそのような意味とは一線を画す。

デザインにおける創造の本質について久保田らは、日本創造学会における創造の定義「人が問題を異質な情報群を組み合わせて統合して社会あるいは個人レベルで新しい価値を生むこと」[78]を引き、デザイナーが異質な情報を深く探索する分析と、その組み合わせを模索する統合を畳み込み、明らかになっていない新しい価値を仮置きしつつ漸次的に進む実践的な行為であると指摘する[79]。本研究における創造の解釈もこれに沿うものである。

データビジュアライゼーションにおいても、創造的に物事を生み出す本質は、個人の技能や直感、あるいは従前に考えた緻密な計画ではなく藁をも掴むような試行錯誤の連続である。このことは第4章に示した通りである。本研究で参照した卓越した作品の多くは、そのようなプロセスの末に行き着いた結果が形となったものである。

このような「創造的」の意味をめぐる社会での齟齬はいかにして起こるのか。おそらくその理由の一つとして、そのプロセスが理解されにくいことが挙げられるだろう。通常そのような泥臭いプロセスは詳らかにされることはなく、ごく一部の成功事例のうちのさらに一部の事例において、そのプロセスがあたかも“計画的に”実行されたかのように語られるのみである。これは嘘や捏造を行っているのではなく、プロジェクトが辿った膨大な試行錯誤やおびただしい数の議論をありのまま記述して取り出し、ひらめきや発想に至る機微を他者が理解できる形で提示することがほとんど不可能であるためである。

創造的なプロセスの本質を知る最善の方法は、自らそれに取り組むことである。まるで先行きが見えない暗い森に分け入るように少しずつ足下を確かめながら進み体得するほかない。本研究のプロセスモデルは、そのような手探りな状況そのものを避けることはできないが、その状況の中でわずかながらでも確かな前進感を持ち続けるために重要なのである。

創造的プロセスの研究は社会からの誤解を受けやすい。プロセスモデルが提案されるとその利用者はモデルをマニュアルのように形式的に踏襲してしまい、目の前で起きている

この本質を観取しないままプロセスが進行することが往々にして起こる。結果として予定調和な作品やプロジェクトが破綻する事態に見舞われる。昨今の「デザイン思考」への批判や、パターン・ランゲージを利用したオレゴン大学キャンパス計画の失敗などにも類似した構造が見て取れる [80]。

本研究における今後の発展を考えるにあたり、このような先行例を踏まえた上で、モデルだけが一人歩きするのではなく実践を通じて多くのデザイナーが学びを得られる環境を準備することが普及の要になるだろう。一例として、大学等でのデザイン教育のみならず現場で活動するデザイナー向けのワークショップなどを企画し、実践者や教育者の育成にも力を注ぎたい。

9.2.3 データビジュアライゼーションの良い質とは何か

本研究はデータビジュアライゼーションのデザイン手法について作り手の視点で論じたが、完成した作品に対する一般鑑賞者（視聴者）の反応、あるいは作者と鑑賞者以外の関係者や社会への影響、すなわち、データビジュアライゼーションの良い質がもたらす効果について事例を収集することも残された課題である。本論文では触れていないこれらの点について、本研究の過程で得られた示唆的な情報を示すことで、今後の展望の足がかりとしたい。

鑑賞者の反応

4章で分析したプロジェクトである「新型コロナウイルスゲノム系統樹の3次元可視化」を一般観覧者に向けて提示し、解説員による案内をつけたセミナー形式のイベントで行ったアンケート結果について示す。セミナーの内容は4.4.3項「フェーズ2: 第2回公開までの記録」を参照されたい。

アンケートは質問紙形式で行い来場者約50名のうち10代未満から70代までの男女16名から回答を得た。映像表現についての質問では各質問項目について、そう思う、そうは思わない、どちらとも言えないの選択肢からひとつ回答を求めた。視覚要素についての質問では、効果的だった、理解を妨げた、どちらとも言えないの選択肢からひとつ回答を求めた。印象に残った内容についての質問では、1. コロナウイルスの名前の由来、2. コロナウイルスの構造、3. 増殖や遺伝子変異（コピーミス）の仕組み、4. 感染がどのようなルートで世界各地に広がった、5. 性質が変化したウイルスが世界に広がったこと、から印象に残った内容を複数選択を許容して回答を求めた。なお、選択肢のうち1は写真を使って説明した内容、2と3は模型を使って説明した内容、4と5はデータビジュアライズで

説明した内容となっている。以下、結果をそれぞれ表 9.1、表 9.2、表 9.3 に示す。

映像表現についての結果（表 9.1）は、質問に対して「そう思う」と回答した人数は多い順に、知らないことがわかった 16 名（100%）、わかりやすい 15 名（93.8%）、興味が湧いた 15 名（93.8%）、知っていたことがさらによくわかった 14 名（87.5%）、美しい 13 名（81.3%）であった。情報の理解の点で十分な満足感を得たことに加えて、美しさという感情的な側面での満足度も高い結果といえる。

視覚要素についての結果（表 9.2）では、要素に対して理解を妨げたと回答した人は、「線や物体の色」が 2 名（12.5%）と最も多かった。変異株に色を割り当てたビジュアライゼーションと変異量に色を割り当てたビジュアライゼーションの 2 種類を提示したことから混乱が生じた可能性がある。効果的だったと回答した人は「線や物体の形」が 16 名（100.0%）と最も多かった。生命進化の変遷を表す系統樹がウイルス変異のストーリーを理解するための形状として相応していたと考えられる。また、一般的に 3 次元の情報の表現は物体同士の重なりや奥行きが生じて理解を困難にすると考えられるが、結果からは理解を妨げる要因にならずむしろ効果的であった。これは視点移動やアニメーション表現によって 3 次元表現の欠点を補完できたためと考えられる。

印象に残った内容についての結果（表 9.3）では、写真で説明した 1 を選択した人は 5 名、模型で説明した 2 または 3 を選択した人は 14 名（平均 7 名）、ビジュアライゼーションで説明した 4 または 5 を選択した人は 18 名（平均 9 名）となった。他の手法による説明に比べてビジュアライゼーションで説明した内容が高い結果となったが、写真や模型での基礎的な説明がビジュアライゼーションで説明した内容の理解を促進し印象を高めた可能性もある。鑑賞者の印象や記憶への定着を図る上では、ビジュアライゼーション以外の方法との組み合わせも効果が高い可能性がある。

複雑なデータビジュアライゼーションや、作品性の高いデータビジュアライゼーションは鑑賞者に多様な洞察を示唆する一方で何が伝わったのかの確認や検証が乏しいことがたびたび指摘される。本研究で提示したデータビジュアライゼーションの質を高める実践知をさらに集積するためには鑑賞者の反応やフィードバックを反映させる必要がある。ここで例示した評価方法は一例に過ぎないがこのような事後検証も取り入れた質の検討は欠かせないだろう。

表 9.1 映像表現について回答人数とパーセンテージ

質問項目	そう思う	そうは思わない	どちらとも言えない
わかりやすい	15(93.8%)	1(6.3%)	0(0.0%)
美しい	13(81.3%)	1(6.3%)	2(12.5%)
興味が湧いた	15(93.8%)	0(0.0%)	1(6.3%)
知らないことがわかった	16(100%)	0(0.0%)	0(0.0%)
知っていたことがさらによくわかった	14(87.5%)	2(12.5%)	0(0.0%)

表 9.2 視覚要素についての回答人数とパーセンテージ

質問項目	効果的だった	理解を妨げた	どちらとも言えない
線や物体の色	14(87.5%)	2(12.5%)	0(0.0%)
線や物体の形	16(100.0%)	0(0.0%)	0(0.0%)
線や物体の大きさ・太さ	13(81.3%)	0(0.0%)	3(18.8%)
線や物体の位置や疎密	13(81.3%)	0(0.0%)	3(18.8%)
線や物体の関係（重なりや接続）	13(81.3%)	0(0.0%)	3(18.8%)
3次元的な表現	14(87.5%)	0(0.0%)	2(12.5%)
視点移動やアニメーション	13(81.3%)	0(0.0%)	3(18.8%)
アニメーションの速さ	12(75.0%)	0(0.0%)	4(25.0%)
地図	13(81.3%)	0(0.0%)	3(18.8%)
画質	13(81.3%)	0(0.0%)	3(18.8%)

表 9.3 印象に残った内容について回答人数とパーセンテージ

選択肢	人数	%
1. コロナウイルスの名前の由来	5	31.3
2. コロナウイルスの構造	4	25.0
3. 増殖や遺伝子変異（コピーミス）の仕組み	10	62.5
4. 感染がどのようなルートで世界各地に広がったか	10	62.5
5. 性質が変化したウイルスが世界に広がったこと	8	50.0

共同製作者の反応

本研究はデータビジュアライゼーションのデザインを対象としているが、質とはプロセスから生まれるアウトプットの質だけでなくプロセスを進める上で深い関係を築き議論を重ねるパートナーやプロジェクトメンバーとの関係と捉えることも重要である。「新型コロナウイルスゲノム系統樹の3次元可視化」プロジェクト全期間において制作に関わった

プロデューサーへのインタビューをここに紹介する。

当初、系統樹の地図上での3次元化にチャレンジしようというアイデアが出たときには、その本当の価値を私たちは理解していませんでした。数週間デザイナーや専門家、ディレクターとの検討を経て、初めて可視化されたときにそれが見え、遺伝情報の変異から国境や大陸をまたいだヒトの動き（感染伝播）がわかることに驚きました。このデータから新しい価値が生み出され、後に4K8Kの番組や報道番組にまで利用されていくことが明確にイメージできました。

ここにはアイデアの一つであった系統樹の3次元可視化がプロジェクトメンバー内の議論と試作を通じてデータの意味を表現する手法としての確信へと変わり、利用価値が生まれていく様子が振り返られている。しかしプロジェクト開始当初においては筆者を含め全員がそこから何が見えどのような価値が生まれるか全く想像できていなかった。大規模なデータをビジュアライズするプロジェクトは、データにどのような意味があるのか何が見えるのかについて暗中模索の中で進むことを余儀なくされ、時には数ヶ月にわたってそのような状況が続く場合もある。これはゴールが見えず非常にストレスフルではあるがDivingやExploreingのプロセスでは通常起こり得る状態である。DEXASはチームが置かれている状況を俯瞰的に捉えながら方向を見失わないように進む指針を提供し、パターンは陥っている状況を共通の認識にするための言葉として機能する。その意味でDEXASはデザイナーだけが知っておくべき情報ではなく、プロジェクトに関わるメンバー全員が知ることでプロジェクト自体の質を高めることに貢献すると考えられる。

9.2.4 データビジュアライゼーションを社会に位置づけるために

本研究では、データビジュアライゼーションという具象物のデザインプロセスについて着目した。しかし、本研究の背景（1.2節）で述べたとおり、AIをはじめ、大規模なデータを社会基盤に活用するにあたって、人々の関心を惹起し透明性を高めるためには、データビジュアライゼーション単体のデザインだけでなく変化する社会状況の中でその位置づけを考慮する必要がある。

そこで、ここではデータビジュアライゼーションを取り巻く社会のいくつかの相に注目し、データビジュアライゼーションとそれらが結ぶ関係にどのような状態が求められるかを検討したい。

まずひとつ目は、社会の変化である。これは、データビジュアライゼーションの動機となる社会的ニーズの観点である。ふたつ目は、市民の変化である。これは、データビジュア

ライゼーションの潜在的な視聴者や鑑賞者である市民が、メディアからの情報にどのように触れているかという観点である。最後に、学術研究の変化である。これは、データビジュライゼーションの対象となるような大規模なデータとその科学研究についての観点である。

社会の変化 – 大規模データに翻弄される社会

序章でも述べたとおり、およそ 2000 年以降、情報通信処理速度の目覚ましい向上やスマートフォンの普及に後押しされる形で、社会に流通するデータ量は急速に増え続け、あらゆる社会的活動がデータとして記録できるようになった。あまりにも膨大なデータは、それを収集・保持・分析・活用する能力を有するものが一部の特定企業などに偏り、情報の独占や個人の不利益につながる不適切な使用に対する懸念の声が高まっている。また情報流通やメディアにも大きな変化があった。ブログやソーシャルメディアによって個人が情報を発信・流通させることが可能になり、自分の趣味趣向にあったニュースを取捨選択して接触する傾向が強まった。相対的にマスメディアの地位が下がり、多様な意見に接する機会が減少した結果、フィルターバブルやフェイクニュースなど、意見の対立が先鋭化しやすくなる構造が顕在化した。それらが招く社会の分断やメディア不信は、今や大国の世論を二分するほどに問題化している [81]。

2022 年末に登場した「ChatGPT」*¹はその対話能力、汎用性の高さから耳目を集める AI である。人間との自然な受け答えや問題解決の能力はこれまでに人々が AI に抱いていた限定的な能力しか持たない AI のイメージを一変させた。その応用範囲は広く、さまざまな分野の IT 企業がこれを取り入れたサービスの開発競争に乗り出しており、まもなくこれまでの仕事、教育、趣味など日常風景は大きく変わっていくだろう。一方、これらのブレイクスルーをもたらした人工知能技術である大規模言語モデルの汎化性獲得原理についてはまだ完全に解明されておらず、事実と異なる事柄を事実のように語る「幻覚（ハルシネーション）」や、誤った情報を鵜呑みにする利用者による誤情報の拡散など、さまざまな社会的リスクが指摘されている [82]。

このように、増え続ける大規模なデータの情報量とは反比例するように、我々の社会は見通しにくく不確実なものになっている。この傾向は今後も続くと考えられる。

*¹ 米国 OpenAI 社が提供する汎用的な対話型人工知能によるチャットサービス <https://chat.openai.com/>

市民の変化 – 情報過多の中で何をみるべきか

情報を受け取る側の市民の意識はどうだろうか。2018年にNHK放送文化研究所が実施した世論調査報告「情報過多時代の人々のメディア選択」[83]^{*2}からその様相を検討する。

その結果「今の社会は情報が多すぎると思う」と考える人の割合が全ての世代において8割を超えている。年齢層にかかわらず社会の情報過多を感じており、特に若年層では「自分が知りたいことだけ知っておけばいい」という意識をもつ人が多く、またそのような意識を持っている場合、情報源は少数でよいと考え、情報収集にも時間をかけず複数の情報源を見比べない人が多いことが報告されている。

また、それら若年層の情報収集はマスメディアとインターネット系メディアの利用率に大きな差がなく、受け取る情報をなるべくコントロールしさまざまなメディアから「自分が知りたいものだけを知る」状況が垣間見えると報告している。

学術研究の変化 – 大規模データを利用する研究

学術界ではデータやシミュレーションを駆使した研究は自然科学や工学分野だけでなく、社会科学や人文科学にも波及してきている。こうした研究成果の多くは、論文として日の目を見ることになる。しかし、論文という限られたスペースの中で、大規模なデータを可視化した図を掲載することは難しく、また、読者に簡潔に伝えるためには、論旨に合うよう簡潔に表現された図に利がある。個別の専門領域では専門家同士で齟齬なく理解できる可視化を使用すべきであるが、一般社会との科学コミュニケーションやアウトリーチを考える場合においては、研究対象としている大規模なデータの一部ではなく全体について、視覚的な工夫を凝らして可視化することによって、一般市民の興味を惹きやすくするメリットがある。図9.1は、筆者らが制作したデータビジュアライゼーションと同種のデータを専門家が論文掲載用に作図した図を比較したものである。しかし、研究者はこのようなニーズに特化した知識やツール持たず、また、そうした十分な時間もないため、研究利用される大規模データの多くはその魅力を十分に引き出されることなく一般社会から隔絶された状態と言えるだろう。

^{*2} 調査は住民基本台帳から無作為に抽出した全国の16～69歳の男女3600人に対し実施し、2369人から有効な回答を得たもの

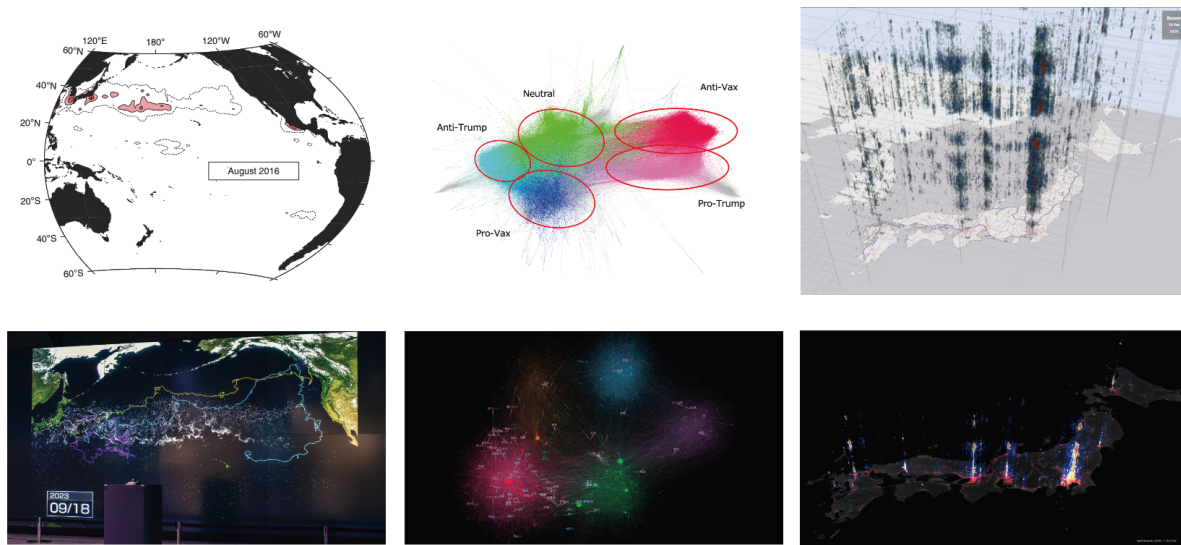


図 9.1 論文用の作図とデータビジュアライゼーションの比較。上段が専門家による作図。画像出典：左：Abundance of non-conservative microplastics in the upper ocean from 1957 to 2066[84] / 中：Characterizing the Anti-Vaxxers' Reply Behavior on Social Media[85] / 右：COVID-19 時空間発生マップ [86]

データビジュアライゼーションの公共的価値

以上、大規模なデータを背景とした3つの相（社会の相、市民の相、学術研究の相）を取り上げた。データビジュアライゼーションがこの3つの相の中心にあってどの相とも連携している時、それは社会に強く根差し、高い公共性とインパクトを持つものになると言えるだろう（図 9.2）。すなわち、社会の相ではテクノロジーの進化を諸手をあげて礼賛したり、不確実な情報に翻弄されることのないよう、そのリスクや課題を見つめ直す「眼」としての役割を果たすこと、市民の相では「情報過多」の意識が定着し「知りたいものだけ」を選び好む接触スタイルが一般化しつつある中、興味を惹き出し、社会全体で議論すべき重要な関心を提示できること、学術研究の相では、研究者や科学者だけでは答えを得られない地球規模の課題や倫理的な課題について、最新研究の動向を市民社会へ開く架け橋となることなどが挙げられる。全てのデータビジュアライゼーションがそうした目的で作られる必要はないが、本研究の背景に示したデータ社会の課題に切り込んでいくには、重要な観点であると考えられる。

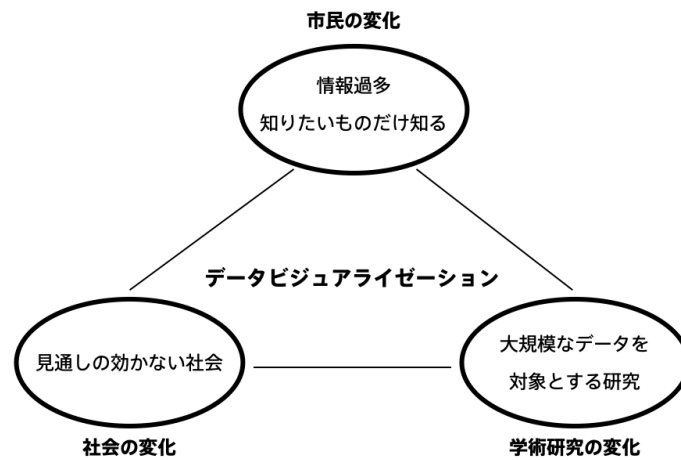


図 9.2 データビジュアライゼーションの社会での位置づけ

改めて本研究を振り返ると、筆者らが制作した多くのデータビジュアライゼーションは、この3つの観点で社会とのつながりを持つことに制作チームの意識が向けられていたように思う。テーマを選ぶ段階では長期的な社会の関心が高くシンプルな答えがない不確実な事象として、COVID-19が候補となった。また、視覚造形の際、非専門家でも読み取れるだけでなく、圧倒的で斬新な視覚体験で鑑賞者の心を掴むことに力を注ぎ、インタラクティブ展示や報道など、使用されるメディアやシチュエーションに応じて、どんなストーリーをどんな形式で提示すれば鑑賞者（視聴者）に届きやすいかを吟味した。

第4章で分析した「新型コロナウイルス3次元系統樹」が幸いにも多くの受賞の機会を得ることができたことは、データビジュアライゼーション単体としてのデザイン性や機能性における評価だけではなく、その周囲にある社会との関係が高く評価されたことも一因となっている。このように、データビジュアライゼーション単体のデザインの外にある、プロジェクトのデザインや関わりのデザインについても、広義のデータビジュアライゼーションデザインと捉えて今後の研究の射程に加えたい。

9.3 おわりに

不確実性が高まる現代の社会において、データの究極の役割とは人間社会と地球環境にとってより豊かな未来を導くために活用されることにある。昨今、データサイエンスやAI技術の重要性はますます高まり、データを扱う人材の教育や便利で多機能な分析ツールの環境は整いつつあり、オープンデータも積極的な公開が活発化している。大量のデータの普及と分析の高度化はまさに破竹の勢いである。分析の正確さや予測精度は今後も飛

躍的に進化し続けるだろう。一方で、人々の行動変容を促すためには技術的精度の向上だけでなく、取り組みに対する幅広い共感と深い理解の機会を生み出せるかが鍵となる。

ビジュアライゼーションによって大規模なデータを見る体験は人々にそのような思いを起こさせるものである。

筆者らが制作したビジュアライゼーション作品が、鑑賞者にインスピレーションを与えたひとつの事例をここに紹介したい。

現在、自然エネルギーを活用した持続可能な社会を実現するために帆船型ドローンの研究開発を進めるスタートアップであるエバーブルーテクノロジーズ株式会社^{*3}の野間恒毅代表は、慶應義塾大学 SFC 協田玲研究室と東京大学大学院経済学研究科の研究 [87] から派生した、18 世紀の商船航海記録データのビジュアライゼーション作品「NEW SYNERGETICS」^{*4}を展示会場で目にし、無数の帆船が自然エネルギーだけで外洋を行き交う様子を見たことが事業立ち上げのきっかけだったと語った。一般の目には触れにくい研究用途の巨大なデータがデータビジュアライゼーションによって思いもよらない形でインスピレーションとなったことを物語っている。

最後に、本論文の執筆を終えようとしていた矢先の 2022 年 11 月、OpenAI 社の「ChatGPT」が公開され、社会に大きな衝撃と動揺が走った。対話という人にとって最も原始的なインタラクションによって、まるで人と変わらぬ知性的で自然な文章や映像を生成する。ビジネス界は即座に色めき立ったが、冷静に考えれば、人類は史上初めて自律的に学習し人と見紛うような知性を感じさせる人ではないものと出会い、今後は長期にわたってそれと向き合うという、とてつもなく大きな課題を背負い込んだといえる。本研究の舞台となった COVID-19 は、未知のウイルスと人類が長期にわたって対峙した現象という観点では、今後、大規模言語モデルが社会にもたらすであろう現象と相似形を成すと言える。本研究が一区切りしたいま、筆者は幸いなことにこの大きな課題について、専門家との協力関係の下、デザイナーとしてデータビジュアライゼーションを通じて社会に問いを立てる研究を行う機会を得た。今後は本研究を糧として、この取り組みを進めていきたい。

社会にはまだ人の目に触れずに眠っている巨大なデータが数多くある。優れた分析に役立てることはもちろんのこと、データビジュアライゼーションによって多くの人々がデータを前に対話する光景がより多くの場所で生まれることを切に願う。

^{*3} <https://www.everblue.tech/>

^{*4} <http://akirawakita.com/#works/newsyn>

謝辞

本研究を進めるにあたっては実に多くの方のご協力とご指導を賜りましたことに感謝を申し上げます。お一人ずつお名前を挙げて感謝の意を表すところではありますが、紙幅に限りがあるために全ての方のお名前を記すことができないことをどうかご容赦ください。

慶應義塾大学大学院政策・メディア研究科の脇田玲教授には主査としてご指導いただいただけでなく、研究の道とは無縁だった筆者を社会人博士課程として受け入れてくださり、本研究につながるデータビジュアライゼーションについて多くの企業や研究者の方々と実践的な共同プロジェクトを行う機会を提供してくださいました。

同大学院の田中浩也教授、石川初教授、鳴川肇准教授には直前のお願いにも関わらず副査として本研究に対する貴重なご助言をいただきました。2022年度まで多摩美術大学教授を務められた永原康史先生には副査としてグラフィックデザイン史やメディアデザインの広範な視点から貴重なご助言を賜りました。加えて、多摩美術大学 TUB 第 12 回企画展「Visible x Invisible ——ビッグデータと次世代の情報表現」にて制作したデータビジュアライゼーションを展示する機会を提供してくださいました。

東海大学医学部中川草准教授は「新型コロナウイルスゲノム系統樹の3次元可視化」の共同開発者として、生命科学の専門的見地から幅広くご助言いただきました。

武蔵野美術大学造形構想学部の長谷川敦士教授には、本研究の評価方法を検討するにあたり貴重なご助言をいただきました。

明星大学デザイン学部の武藤努教授、千葉工業大学創造工学部の稲坂晃義准教授、東京都立大学システムデザイン学部の杉本達應准教授には本研究のリリースにご協力いただきました。

名古屋工業大学コミュニティ創成教育研究センターの五十嵐康伸客員准教授には、情報可視化の応用に関する見識に基づいた貴重なご助言をいただきました。

慶應義塾大学 SFC 脇田玲研究会の学部生、大学院生の皆様には本研究についての議論を通じて、さまざまな示唆に富むご意見をいただきました。同研究会の修了生である中野

亜樹人氏、長島禎氏、計良風太氏はデータビジュアライゼーション開発の実務や技術面で多大なるご協力をいただきました。

2017年、多摩美術大学での演習講座を皮切りに、情報科学芸術大学院大学、武蔵野美術大学にて筆者が担当するデータビジュアライゼーション演習を履修した学生の多くは熱意を持って取り組み、筆者が想定した以上の成果作品を発表しました。このことは筆者が本研究で掲げた創造的なデータビジュアライゼーション手法の体系化を目指すにあたって、貴重なインスピレーションとモチベーションになりました。

筆者の力量不足から、博士研究と仕事と家庭の両立がうまくいかない時間も多く、周囲には体力的にも精神的に負担を強いていたにも関わらず、常に筆者を励まし博士研究を最後までやり抜く機会を与えてくれた家族、親族、教員、同僚の皆様の寛容さには感謝の言葉が見つかりません。

今後も本研究の成果を基盤として社会における創造的なデータビジュアライゼーションの担い手を増やすべく、研究教育活動を昇華させていきたいと考えます。そして、機会が許すなら本研究でお世話になった皆様お一人お一人に直接感謝の言葉を贈りたいと思います。

参考文献

- [1] 唐沢かおり. データ駆動型社会における「人間中心」に向けた課題. 横幹, Vol. 14, No. 1, pp. 24–32, 2020.
- [2] Nathan Shedroff. 11 information interaction design: A unified field theory of design. *Information design*, p. 267, 2000.
- [3] Yun Wang, Adrien Segal, Roberta Klatzky, Daniel F. Keefe, Petra Isenberg, Jörn Hurtienne, Eva Hornecker, Tim Dwyer, and Stephen Barrass. An emotional response to the value of visualization. *IEEE Computer Graphics and Applications*, Vol. 39, No. 5, pp. 8–17, 2019.
- [4] Nick Cawthon and Andrew Vande Moere. The effect of aesthetic on the usability of data visualization. In *2007 11th International Conference Information Visualization (IV '07)*, pp. 637–648, 2007.
- [5] IEEE. IEEE Vis Arts Program. <https://visap.net/>, 2022.10.30 アクセス.
- [6] Data Visualization Society. IIB Awards. <https://www.datavisualizationsociety.org/about-iib-awards>, 2022.11.30 アクセス.
- [7] 朝日新聞社. 見えない交差点. <https://www.asahi.com/special/jiko-kosaten/>, 2022.12.11 アクセス.
- [8] National Geographic. These are the best and worst states for women. <https://www.nationalgeographic.com/magazine/article/these-are-the-best-and-worst-states-for-women>, 2022.12.11 アクセス.
- [9] 内閣府. V-RESAS. <https://v-resas.go.jp/>, 2022.12.11 アクセス.
- [10] Andres Ramirez Gaviria. When is information visualization art? determining the critical criteria. *Leonardo*, Vol. 41, No. 5, pp. 479–482, oct 2008.
- [11] Andrea Lau and Andrew Vande Moere. Towards a model of information aesthetics in information visualization. In *2007 11th International Conference Informa-*

- tion Visualization (IV '07), pp. 87–92, 2007.
- [12] 増井俊之. 情報視覚化の最近の研究動向 (〈特集〉情報の視覚化). 情報の科学と技術, Vol. 54, No. 11, pp. 558–567, 2004.
- [13] G. Judelman. Aesthetics and inspiration for visualization design: bridging the gap between art and science. In *Proceedings. Eighth International Conference on Information Visualisation, 2004. IV 2004.*, pp. 245–250, 2004.
- [14] Robert Plot. Observations of the wind, weather, and height of the mercury in the barometer, through out the year 1684; taken in the musaeum ashmoleanum at oxford, by robert plot, lld. to which is prefixt a letter from him, to dr. martin lister, f. of the r. s. concerning the use of this and such like historys of the weather. *Philosophical Transactions (1683-1775)*, Vol. 15, pp. 930–943, 1685.
- [15] Michael Friendly and Howard Wainer. *A History of Data Visualization and Graphic Communication*, p. 21. Harvard University Press, 2021.
- [16] 大田暁雄. 視覚言語による社会革命 – オットー・ノイラートの「社会と経済」アトラス. オーム社, 2021.
- [17] 桑田学. 経済的思考の転回——世紀転換期の統治と科学をめぐる知の系譜, 第 2 章. 以文社, 2014.
- [18] 伊原久裕. アイソタイプからピクトグラムへ (1925-1976) : オットー・ノイラートのアイソタイプとルドルフ・モドレイによる図記号標準化への影響に関する研究. PhD thesis, 九州大学, 2014.
- [19] 伊原久裕. 思想家によるデザイン – オットー・ノイラートの経済思想とデザインをめぐって. デザインに哲学は必要か, pp. 75–101, 2019.
- [20] 永原康史. インフォグラフィックスの潮流: 情報と図解の近代史. 誠文堂新光社, 2016.
- [21] 鳴川肇. AuthaGraph. <http://www.authagraph.com/>, 2023.1.10 アクセス.
- [22] Joachim Krausse, Claude Lichtenstein, 神奈川県立近代美術館, 愛知県立美術館, ワタリウム美術館 (編). *Your Private Sky R. バックミンスター・フラワー アート・デザイン・サイエンス*, pp. 269–175. Lars Müller, 2001.
- [23] Richard Saul Wurman. *Understanding USA*. Rswinc, 1999.
- [24] Brian Johnson and Ben Shneiderman. Tree-maps: A space filling approach to the visualization of hierarchical information structures. In *Proceeding Visualization '91*, pp. 284–291. IEEE, 1991.
- [25] Christopher Ahlberg and Ben Shneiderman. Visual information seeking: Tight

- coupling of dynamic query filters with starfield displays. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'94)*, pp. 313–317. Addison-Wesley, 1994.
- [26] Tamara Munzner. H3: Laying out large directed graphs in 3d hyperbolic space. In *Proceedings of VIZ'97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium*, pp. 2–10. IEEE, 1997.
- [27] Stephen G Eick, Joseph L Steffen, Eric E Sumner, et al. Seesoft-a tool for visualizing line oriented software statistics. *IEEE Transactions on Software Engineering*, Vol. 18, No. 11, pp. 957–968, 1992.
- [28] SENSORIUM. <https://archive.aec.at/prix/showmode/31699/>, 2022.11.30 アクセス.
- [29] John Maeda. SHISEIDO 30-YEAR ANNIVERSARY POSTER. <http://maedastudio.com/1995/shisposter/>, 2022.11.30 アクセス.
- [30] Fernanda B Viégas, Martin Wattenberg, Jesse Kriss, and Frank Van Ham. Talk before you type: Coordination in wikipedia. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*, pp. 78–78. IEEE, 2007.
- [31] Aaron Koblin. Flight Patterns. In *ACM SIGGRAPH 2006 Computer animation festival*, p. 203. 2006.
- [32] Sensorium. Web Hopper. <http://eto.com/1996/WebHopper/>, 2022.11.30 アクセス.
- [33] Edward R Tufte. *The visual display of quantitative information*. Graphics Press, 1983.
- [34] Edward R Tufte. *Envisioning information*, Vol. 126. Graphics Press, 1990.
- [35] Edward R Tufte. *Visual Explanations*. Graphics Press, 1997.
- [36] Scott Bateman, Regan L Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. Useful junk? the effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 2573–2582, 2010.
- [37] Colin Ware. *Information visualization: perception for design*. Morgan Kaufmann, 4th edition edition, 2019.
- [38] William S. Cleveland and Robert McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, Vol. 79, No. 387, pp. 531–554, 1984.

- [39] Stephen Few. *Show me the numbers*. 2004.
- [40] Thomas MJ Fruchterman and Edward M Reingold. Graph drawing by force-directed placement. *Software: Practice and experience*, Vol. 21, No. 11, pp. 1129–1164, 1991.
- [41] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*, pp. 364–371. Elsevier, 2003.
- [42] Stuart K Card, Jock D Mackinlay, and Ben Shneiderman. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.
- [43] Riccardo Mazza. *Introduction to information visualization*. Springer Science & Business Media, 2009.
- [44] 伊藤貴之. 意思決定を助ける 情報可視化技術, p. 6. コロナ社, 2018.
- [45] Michael Sedlmair, Miriah Meyer, and Tamara Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics*, Vol. 18, No. 12, pp. 2431–2440, 2012.
- [46] David Lloyd and Jason Dykes. Human-centered approaches in geovisualization design: Investigating multiple methods through a long-term case study. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 17, No. 12, pp. 2498–2507, 2011.
- [47] Tamara Munzner. *Visualization analysis and design*. CRC press, 2014.
- [48] Tamara Munzner. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 15, No. 6, pp. 921–928, 2009.
- [49] Andy Kirk. *Data visualisation: A handbook for data driven design*. Sage, 2016.
- [50] Donald Norman. A. 2004, emotional design: Why we love or hate everyday things, 2004.
- [51] Visible language workshop special collection, the art, culture, and technology program at mit. <https://act.mit.edu/special-collections/vlw-archive/>, 2023.1.13. アクセス.
- [52] Ishantha Lokuge and Suguru Ishizaki. Geospace: An interactive visualization system for exploring complex information spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '95*, pp. 409–414, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

- [53] Gerhard Bachfischer and Toni Robertson. From movable type to moving type-evolution in technological mediated typography. 12 2022.
- [54] David Reinfurt and Robert Wiesenberger. *Muriel Cooper*, pp. 183–185. MIT Press, 2017.
- [55] Lisa Strausfeld. Financial viewpoints: using point-of-view to enable understanding of information. In *Conference companion on Human factors in computing systems*, pp. 208–209, 1995.
- [56] Benjamin Jotham Fry. *Computational information design*. PhD thesis, Massachusetts Institute of Technology, 2004.
- [57] Processing.org. <https://processing.org/>, 2022.11.30 アクセス.
- [58] Aaron Koblin and Valdean Klump. Flight Patterns: A Deep Dive. *Beautiful visualization: Looking at data through the eyes of experts*, pp. 91–102, 2010.
- [59] Michael Young and Nick Bilton. A Day in the Life of the New York Times. *Beautiful visualization: Looking at data through the eyes of experts*, pp. 271–290, 2010.
- [60] Hadley Wickham et al. ggplot2. <https://ggplot2.tidyverse.org/>, 2023.5.12 アクセス.
- [61] The Matplotlib development team. matplotlib. <https://matplotlib.org/>, 2023.5.11 アクセス.
- [62] Michael Bostock et al. D3.js Data-Driven Documents. <https://d3js.org/>, 2023.5.12 アクセス.
- [63] three.js authors. three.js. <https://threejs.org/>, 2023.5.12 アクセス.
- [64] Urban Computing Foundation. deck.gl. <https://deck.gl/>, 2023.5.12 アクセス.
- [65] Masaki Yamabe, So Nakagawa, and Akira Wakita. Time-space-based visualization of sars-cov-2 phylogeny. *VIZBI*, 2021.
- [66] Bette Korber, Will M Fischer, Sandrasegaram Gnanakaran, Hyejin Yoon, James Theiler, Werner Abfalterer, Nick Hengartner, Elena E Giorgi, Tanmoy Bhattacharya, Brian Foley, et al. Tracking changes in sars-cov-2 spike: evidence that d614g increases infectivity of the covid-19 virus. *Cell*, Vol. 182, No. 4, pp. 812–827, 2020.
- [67] 国立感染症研究所. 新型コロナウイルス感染症の直近の感染状況等 (2021年5月26日現在). <https://www.niid.go.jp/niid/ja/2019-ncov/10397-covid19-ab36th.html>, 2022.10.31 アクセス.

- [68] 感染・伝播性の増加や抗原性の変化が懸念される 新型コロナウイルス (sars-cov-2) の新規変異株について (第 9 報) . <https://www.niid.go.jp/niid/ja/2019-ncov/2484-idsc/10434-covid19-43.html>, 2022.12.15 アクセス.
- [69] Danny Holten. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Transactions on visualization and computer graphics*, Vol. 12, No. 5, pp. 741–748, 2006.
- [70] 井庭崇, 中埜博, 竹中平蔵, 江渡浩一郎, 中西泰人, 羽生田栄一. パターン・ランゲージ: 創造的な未来をつくるための言語. 慶應義塾大学出版会, 2013.
- [71] 井庭崇, 古川園智樹. 創造社会を支えるメディアとしてのパターン・ランゲージ. 情報管理, Vol. 55, No. 12, pp. 865–873, 2013.
- [72] 鈴木寛, 岩瀬直樹, 今井むつみ, 市川力, 井庭崇. クリエイティブ・ラーニング: 創造社会の学びと教育. 慶應義塾大学出版会, 2019.
- [73] Takashi Iba and Taichi Isaku. A pattern language for creating pattern languages: 364 patterns for pattern mining, writing, and symbolizing. In *Proceedings of the 23rd Conference on Pattern Languages of Programs, PLoP '16, USA, 2016*. The Hillside Group.
- [74] 井庭崇. 創造的な対話のメディアとしてのパターン・ランゲージ: ラーニング・パターンを事例として. *Keio SFC journal*, Vol. 14, No. 1, pp. 82–106, 2014.
- [75] 井庭崇. ソフトウェアパターン-時を超えるソフトウェアの道-:5. パターンランゲージ 3.0 -新しい対象 × 新しい使い方 × 新しい作り方-. 情報処理, Vol. 52, No. 9, pp. 1151–1156, aug 2011.
- [76] 井庭崇, 梶原文生. プロジェクト・デザイン・パターン 企画・プロデュース・新規事業に携わる人のための企画のコツ 32, p. 25. 翔泳社, 2016.
- [77] 長田尚子. 試行錯誤を享受する実践コミュニティのエスノグラフィ: PBL の経験を記述するパターン・ランゲージの開発を通じて. ヒューマンインタフェース学会論文誌, Vol. 23, No. 3, pp. 287–302, 2021.
- [78] 日本創造学会. 創造の定義. <http://www.japancreativity.jp/definition.html>, 2023.6.20 アクセス.
- [79] 久保田拓朗, 小田裕和, 串田隼人, 長尾徹, 田隈広紀, 八馬智. デザイン教育における創造的学習プラットフォームの提案. 国際 P2M 学会誌, Vol. 10, No. 1, pp. 35–52, 2015.
- [80] 長坂一郎. クリストファー・アレグザンダーの思考の軌跡. 彰国社, 2015.
- [81] 林 香里. メディア不信: 何が問われているのか. 岩波新書. 岩波書店, 2017.

- [82] 岡野原大輔. 大規模言語モデルは新たな知能か：ChatGPT が変えた世界. 岩波書店, 2023.
- [83] 保高 隆之. 情報過多時代の人々のメディア選択. 放送研究と調査, Vol. 68, No. 12, pp. 20–45, 2018.
- [84] Atsuhiko Isobe, Shinsuke Iwasaki, Keiichi Uchida, and Tadashi Tokai. Abundance of non-conservative microplastics in the upper ocean from 1957 to 2066. *Nature communications*, Vol. 10, No. 1, p. 417, 2019.
- [85] Kunihiro Miyazaki, Takayuki Uchiba, Kenji Tanaka, and Kazutoshi Sasahara. Characterizing the anti-vaxxers’ reply behavior on social media. WI-IAT ’21, pp. 83–89, New York, NY, USA, 2022. Association for Computing Machinery.
- [86] 東北大学環境科学研究科, 株式会社 JX 通信社. Covid-19 時空間発生マップ. <https://nakaya-geolab.com/covid19-stkd/japan/>, 2023.6.19 アクセス.
- [87] VEHA: Visualizing Economic History of Asia. <http://www.veha.e.u-tokyo.ac.jp/>, 2023.1.12 アクセス.

付録 A

付録

本研究で作成した「データビジュアライゼーション・デザインパターン」を付録として収録する。

データビジュアライゼーション
デザインパターン

「わかる」より「感じる」可視化のデザイン

はじめに

【大規模なデータ利用の拡大に潜む影】

情報技術の進展によって、現代ではかつてない大規模なデータの蓄積と分析が可能になりました。機械学習やデータサイエンスなどを通じてデータの利用が促進され、私たちの日常はデータによって便利さを増しています。

一方で、複雑で膨大なデータを分析して利用できる形(モデル)を導くことができるのは、高度に専門的な知識を持つ人々に限られます。そもそも大規模なデータの全容は人間の知覚を通して理解することが難しく、緻密な分析作業や深層学習などコンピュータの計算力を通して理解できるものです。そのため、モデルやモデルを使用したサービスを扱う一般の人々が、そのプロセスを理解することは困難です。

日常生活に関係する物事の予測や分類にさまざまなモデルが普及しはじめています。それらのモデルの精度が高く感じられたとしても、信頼に足るのか、正当な評価をしているのかという不安を私拭する確かな方法はありません。私たちは、いくらモデルを説明されたとしても、元のデータの複雑な有り様や多様な解釈の幅を俯瞰的にみて、その妥当性を判断することはできないからです。もしかしたらそこには隠れたバイアスが存在しているかもしれないかもしれません。しかし今のところ、一般の利用者の立場としてはモデル(とその分析者)を信用するしか方法がありません。

このような不安は「データ独占」への批判にもつながっています。データを集め分析する高い能力を持った企業と、個人の属性や行動パターンを提供する消費者という、情報の非対称が拡大していることが背景にあります。その問題を緩和しようと、法律の整備が国際的に進んでいます。企業が個人の同意を得ずに情報収集することに歯止めをかけるこうした動きは、企業間競争の公正を保ち、個人が意図しないデータ提供を防ぐという意味で人々の不信感を解消する一定の効果は望めます。しかし、企業や

組織のデータ収集と蓄積が何を目的にしているのか、それによって個人あるいは社会にどんな利益をもたらそうとしているのかを理解可能にする効果はあまり期待できません。つまるところ、データは常に隠されています。複雑なデータは利用者にとっては見ることも触れることもできず、データを持つものを持たざる者の間にある深刻なコミュニケーション不足が、不安や不信感の根本にあります。こうした問題が広がれば、やがて利用者はただモデルに従うだけで、真実を見通せない不可視な社会を生きていることを余儀なくされてしまいます。

【データビジュアライゼーションの役割と可能性】

複雑で大規模なデータを一般の人も見ても何かを感じられるようにする方法はないのでしょうか。本書ではこうした大規模なデータをめぐる問題に目を向けて、複雑なデータから、多様なあり方が感じられるような表現、すなわち「データビジュアライゼーション」について可能性を論じ、その詳細と作り方を紹介するものです。

データビジュアライゼーションに近い言葉として、「情報可視化(Information Visualization)」があります。情報可視化は数値をグラフやチャートなどの視覚的な表現に置き換えたり、地図にマップピングすることで傾向の把握や、比較を容易にします。データビジュアライゼーションも数値を視覚的に表現するという意味では同じですが、その狙いは異なります。情報可視化は大規模で複雑なデータを理解しやすいシンプルな状態に整理し、誰もが一目瞭然に理解できることを主眼に置きます。これを認知負荷の軽減と言います。

一方、データビジュアライゼーションは、大規模なデータに特有な複雑さや多様な情報の切り口を、できるだけありのままに表現しようとしています。前段の話になぞらえると、情報可視化は、複雑なデータからのはっきりと見出せるシンプルな傾向を切り出した

り、モデルを説明することに向いています。一方、データビジュアライゼーションは、大規模なデータを前に「はっきりとはわからなければデータ全体の全体像はこんな形をしているのか」、「こうした見え方の特徴は、現実のどの問題に関連しているのだろうか?」という具合に、興味や関心を生み出す、いわば何かを予見したり感じさせることに主眼を置きます。

このような、データビジュアライゼーションと情報可視化の役割による定義は学術的なコンセンサスが確立してのものではありません。文献で使われている場合にはこのような意味が適合しないことが多いです。しかし、「データ」の意味が解釈され「情報」になる前の生に近い状態で見ることと考えると、データビジュアライゼーションという呼び方で区分することは理に適っていると言えます。

料理に例えるならば、情報可視化は、データという素材を調理済みの美味しい食事としていただくことに似ています。一方、データビジュアライゼーションは、まだ調理の途中で、素材がぐつぐつと煮えている鍋をみんなで覗き込んであれこれ言うような状況に似ていると言えます。前者のように、美味しくいただくための情報可視化に関しては、すでに多くの手法の提案や書籍が出版されています。

一方、「データの顔」を覗き込むことで、専門家と非専門家が対話できる環境を作り出すようなデータビジュアライゼーションの作成方法について解説するものは、残念ながらほとんど見つけられません。本冊子は、そのような問題意識から筆者自身がデザイナーとして従事したデータビジュアライゼーションプロジェクトを振り返って分析し、そのプロセスをまとめたものです。本冊子を通じて、多くのデザイナーやクリエイターが専門家とともに開かれたデータビジュアライゼーションのデザインを実践することに役立っていただければ幸いです。

目次

はじめに	2
[I] データビジュアライゼーションとは何か	7
データビジュアライゼーションの目的	8
作品紹介1	10
作品紹介2	12
作品紹介3	14
[II] データビジュアライゼーションをつくる	17
データビジュアライゼーションの難しさ	18
DEXASでデータビジュアライズを共創する	19
DEXASは誰の視点で作られている？	20
DEXASの詳細	21
DEXASのチーム構成	30
[III] DEXAS バターン・ランゲージ	31
DEXAS バターン・ランゲージとは	32
バターン・ランゲージの使い方	32
バターン・ランゲージの全体像	33
バターンの読み方	35
バタートーン一覧	37
カテゴリ0：コアバタートン(No.1)	38
カテゴリ1：データにまつわる情報収集(No.2～No.8)	39
カテゴリ2：データを可視化する(No.9～No.18)	46
カテゴリ3：可視化をレビューする(No.19～No.26)	56
カテゴリ4：作品を組み立てる(No.27～No.33)	64

I データビジュアライゼーションとは何か

データビジュアライゼーションの目的

本冊子ではデータビジュアライゼーションの目的を次のように整理します。

- 複雑なデータを多様なあり方が感じられるような表現で可視化する
- 多面的な視点でデータを眺めることができる
- 興味を喚起したり、問題について幅広く議論する場を生み出す

このように整理すると、データビジュアライゼーションはアートの文脈に近いと感じるかもしれません。現に、「データ・アート」と呼ばれる芸術ジャンルの中にもこうした作品を見ることが出来ます。データ・アートとデータビジュアライゼーションは明確に区別できるものではありません。どちらも美術館で展示されるものもあれば、インターネット上で公開され、ブラウザで閲覧できるタイプのものもあります。鑑賞者が操作できるものもあれば、ただ眺めながら鑑賞するものもあり、その様態は多岐にわたります。

データ・アートはデータの可視化を介して表現すると言う点ではデータビジュアライゼーションと共通していますが、本来、目には見えないマクロやミクロの自然現象の様相や、人間社会の有り様を描き出そうとするものが多く見受けられます。その多くは生身の体では知覚できない複雑な構造や、変化のダイナミズムを表象し、鑑賞者の感性に訴えかけるように設計されており、具体的に数値を読み解いたり比較することは難しいものもあります。

またデータビジュアライゼーションは「データ・ジャーナリズム」とも近い関係にあります。データ・ジャーナリズムもデータビジュアライゼーションやデータ・アートと同様に、一般の人々が見ることを想定し、データの可視化を伴って問題の核心に迫り、時には幅広い話題を提供します。しかし、データ・ジャーナリズムは作者が主張を語る上の裏付けや補強のために、実証的にデータの可視化を利用するのに対して、データビジュアライゼーションは、作者の主張が言葉によって前面に出るのではなく、鑑賞者

自身がより能動的に、データの背後にある物語を読み取れるように設計されています。

このように、データビジュアライゼーションはデータ・アートやデータ・ジャーナリズムと共通点を持ちます。しかし、制作者が特定の問題についてどれだけ深く具体的に主張するのか、あるいは、解釈の幅を鑑賞者にどれだけ持ち持たせるのかと言う点では差があり、データビジュアライゼーションは両者の中間に位置するという見方も成り立ちます。

本冊子では、データビジュアライゼーションの作り方の枠組みを紹介するのですが、データ・アートやデータ・ジャーナリズムに関する文献の中にも多くのヒントを発見することが出来ます。以下では、データビジュアライゼーションの事例を紹介します。

作品介绍1

Flight patterns (2005)

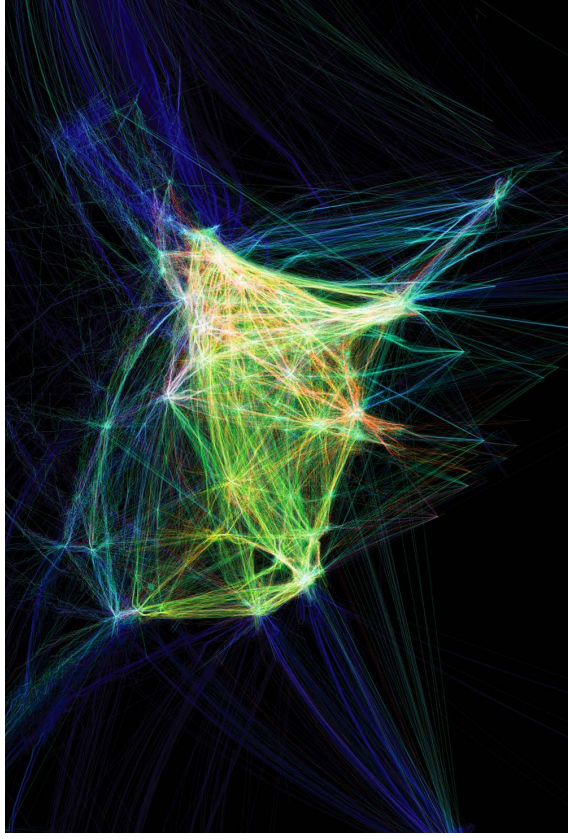
作者：Aaron Koblin, Valdean Klump

米国連邦航空局(FAA)が収集した2005年3月19日17:00～翌20日20:00までの約14万作のフライトデータを可視化したものです。データには、航空機の約3分間隔の飛行位置(緯度、経度、高度)、時刻、機種、機体メーカー、フライト番号など、多数の変数を含んでいます。これらの変数を組み合わせ、さまざまな切り口で描くことで、現代の航空交通システムの緻密さや、大規模な人の移動のリリテイを感じさせることに成功しています。時間ごとに変化するアニメーションでは、夕方から夜にかけて次第に飛行の軌跡が減り、また朝には時差的に早いヨーロッパから大西洋を越えて大量の航空機が飛来してくる様子がわかります。機種によって色分けされた軌跡からは、大型機が就航している路線と中型機が就航している路線の違いによって、北米の大都市の分布と中規模都市の関係を強く想起させます。ズームしてみると、混雑している大規模空港周辺でも離陸機と着陸機の軌跡が整然と並び、管制システムの緻密さをうかがわせます。

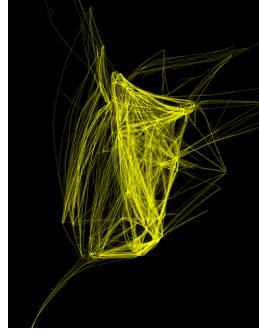


下層中の航空機(オレンジ色)、上昇中の航空機(青)

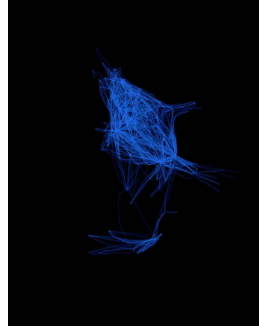
10 | データビジュアライゼーションとは何か



"Flight Patterns" Aaron Koblin, Valdean Klump (2005)



ボーイング737(大型機)の軌跡



エンブラエルERJ 145(中型機)の軌跡

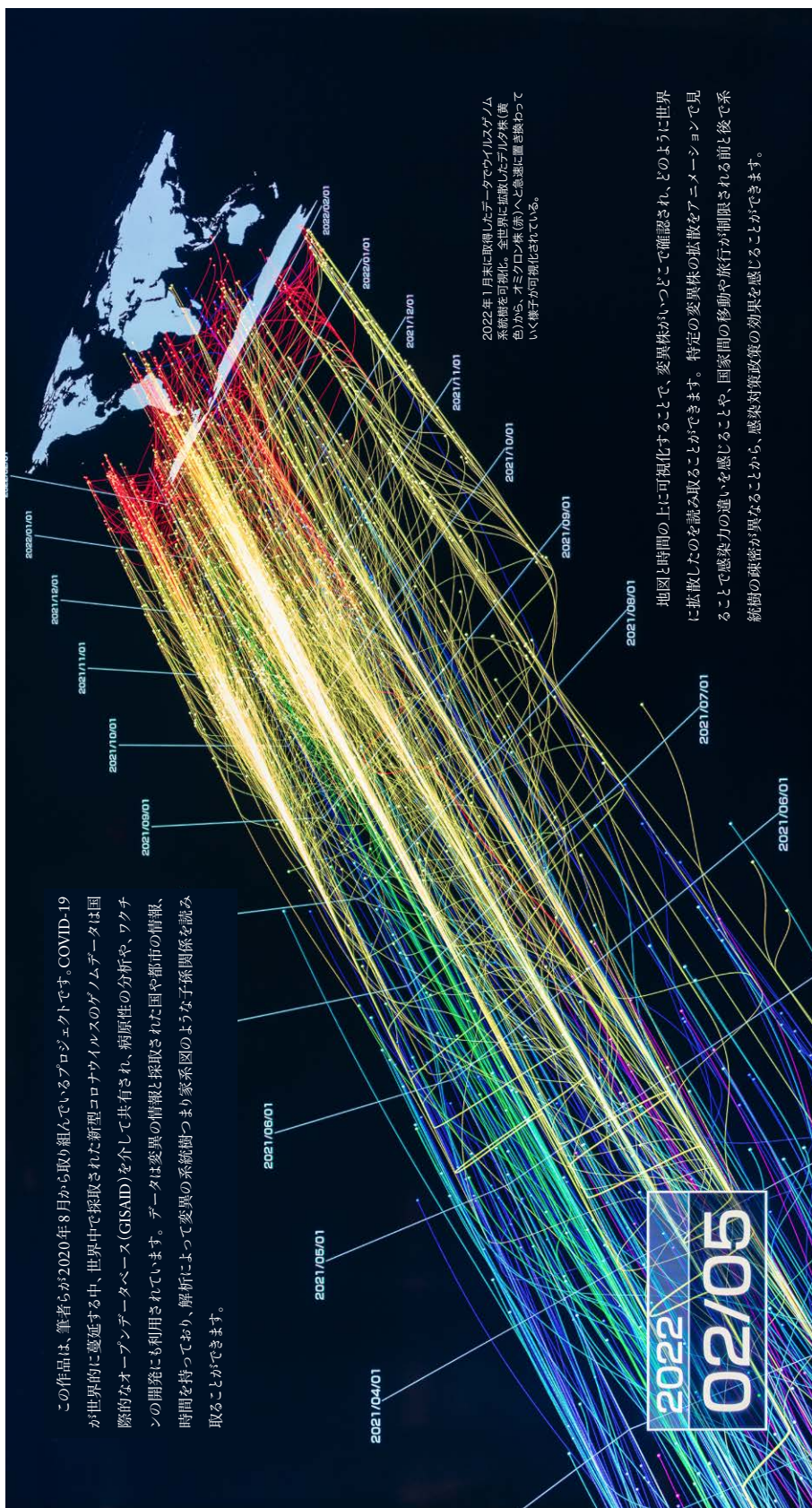
同作品はウェブサイトで見ることができます。また作者のインタビューも公開されています。

<http://www.aaronkoblin.com/work/flightpatterns/>

https://www.ted.com/talks/aaron_koblin_visualizing_ourselves_with_crowd_sourced_data

(画像出典：Julie Steele, Noah Illinsky, データビジュアライゼーション、オライリー・ジャパン、2011、p86、90-91)

1 | データビジュアライゼーションとは何か | 11



©NHK

作品介绍2

新型コロナウイルスゲノム系統樹 (2020)

作者：山辺真幸、中川草 映像制作：NHK エドキュケーション

可視化の動画は公開されています
<https://fimespacephylogeny.xyz/>

作品介绍3

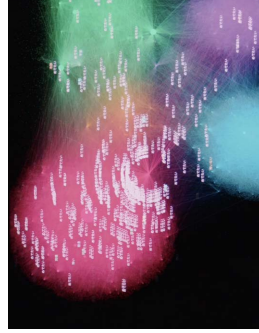
コロナ禍のツイート—感情はどう広がるのか？(2022)

作者：NHK ビジュアライゼーション | 山辺真幸 データ提供・監修：笹原俊和

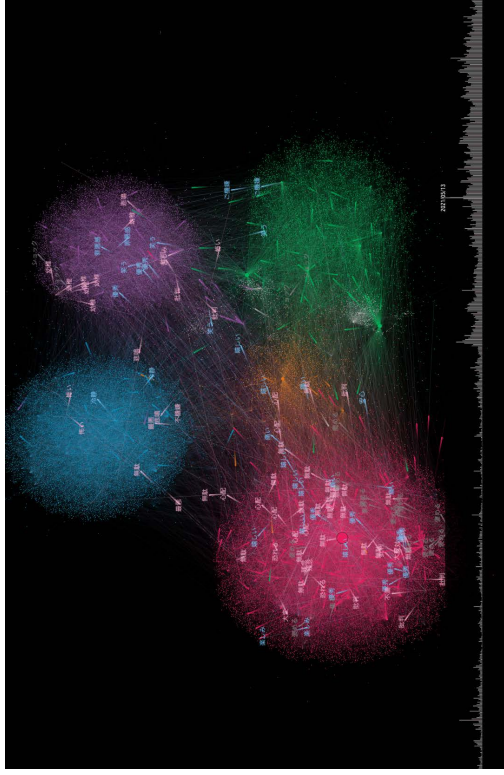
コロナ禍の約2年間に「Twitter」に投稿された「新型コロナ」や「ワクチン」に関連するコメントやツイート約550万件を分析した結果のうち、約28万件を8Kで可視化した作品です。

ひとつひとつの点はアカウントを表し、あるアカウントのツイートを別のアカウントがリツイートするとアカウント同士が線で結び合います。ツイートの内容を解析した結果から、似通った意見を持つアカウントは同じ色で近くに配置されます。この図では6色にグループ分けされています。赤はワクチンに積極的な反応を示すアカウント、青は否定的なアカウントを示します。ツイート内のテキストから分析された感情は、アカウント間を飛び交う文字と、粒子の色で表現されます(肯定的な場合は寒色系、否定的な場合は暖色系)。下部には横軸に時間を、縦軸にリツイート数をとった棒グラフが表示されており、鑑賞者は時間を操作することで、ワクチン関連のニュースが時系列に変わっていく中で、人々がどのように反応したのかを読み取ることができます。

可視化の動画は公開されています
<https://vimeo.com/688760254>



ツイートが大量にリツイートされ拡散していく様子



II データビジュアライゼーションをつくる

データビジュアライゼーションの難しさ

〔感性に響く性質、理解を促す性質〕

優れたデータビジュアライゼーションはどのように作ることができるのでしょうか。これまでの話を踏まえると、データビジュアライゼーションはデータ・アートのように理屈を超えて感性に響く性質と、データ・ジャーナリズムのように実証的な洞察(客観的に見て成り立つ理解)を伝える性質が、ひとつの作品の中に備わっているとと言えます。言葉にするとはシンプルですが、この二つの性質を同時に成立させようとすると巨匠な問題が浮かび上がります。例えば、一般的に実証的な洞察を得るためには条件を設定して、少ない変数に絞って可視化したり論を組み立てるとよい結果につながります。しかし、それを追求しすぎると、現実の複雑さや生々しい質感といった「リアルさ＝感性に響く性質」は失うこととなります。逆に、一目で引き込まれるようなビジュアルを持つた作品は興味を惹きつけます。しかし、過剰な演出は論外としても、多くの変数を可視化してより多元的に表現しようとするれば、少ない変数で説明できるはずの個別の現象を把握しづらくなります。このように、この二つの性質はデータビジュアライゼーションを作るさまざまな局面でトレードオフに陥りやすいのです。

〔異なる専門性—視覚デザインとドメイン知識〕

相反する二つの性質をひとつの作品にまともなレベルを難しくする、別の問題があります。それは異なる専門性が要求されることです。例えば、あるデータを可視化して実証的な洞察を提示しようとするれば、ただ可視化するだけでは不十分で、可視化からその現象を説明するに足る専門的な知識(ドメイン知識)が必要になります。一方、感性に訴えかける能力的な表現を実現するには視覚デザインや、創造的なプログラミングの知識や技術力が要求されます。

優れたデータビジュアライゼーション制作で高い評価を受けている Ben Fry は

「Computational Information Design」と題した2004年の論文の中で、巨大なビッグデータデータのビジュアライゼーション制作した経験から、データビジュアライゼーション制作にとって、この異なる専門性を統合することが重要であることを示しました。コンピュータサイエンティストであり情報デザインの専門家でもある Fry は、まさにこの統合を体現しています。

しかし、それから20年近く経った現在、データをめぐる社会状況は大きく移り変わり、今や、巨大で複雑なデータはあらゆる分野で蓄積されるようになりました。データサイエンスが注目を集め、科学界でも産業界でも多くの分野で高度なデータ分析によって新しい知見が生まれています。このような状況で、個人の努力で統合を目指すこと、すなわち、それぞれの分野の専門家が高度な視覚デザインの技能を磨いたり、デザイナーが専門家と同じように特定分野の知識を深めることは、より一層困難になってきたと言えます。

DEXASでデータビジュアライゼーションを共創する

こうした困難さを乗り越える良い方法は、デザイナー(あるいは視覚表現を得意とするデジタルクリエイター)と専門家が協働してデータビジュアライゼーションに取り組むことです。実際に、世の中で注目を集める優れたデータビジュアライゼーションは、専門性の異なるメンバーのチームワークで制作されている場合が多いです。しかしその具体的な方法論は形式化されていません。

そこで、筆者(デザイナー)は、過去にさまざまな分野の専門家と協働でデータビジュアライゼーション制作した実例から、両者の間やりとり(あるいは関係性)を分析し、前述の問題を解決する制作の進め方を形式化しました。それを「DEXAS」と名づけています。DEXASは「データにまつわる情報収集」を意味する「Diving」、「データを可視化して、意味を読み解く」過程を繰り返す、さまざまな物語を抽出することを意味する「Exploring」、そして「物語を束ねひとつの作品としてまともなレベルを上げる」ことを意味する

「Assembling」の頭文字をとって名付けられています。この3つのステップを意識することで、デザイナーと専門家が互いの強みを生かしながら、前述した二つの性質をよいバランスで取り入れることにつながり、その結果、質の高いデータビジュアライゼーションが生まれやすくなります。

DEXASは誰の視点で作られている？

DEXASの詳細について説明する前に、DEXASを利用する主体について説明します。前述の通り、データビジュアライゼーションはデザイナーと専門家の協働で進めますが、DEXASはデザイナーの視点で整理しています。これはプロジェクトそのものを終始デザイナーがリードして進めていくことを意味します。一見、データ分析や可視化に詳しい専門家がリードして、デザイナーに指示した方が都合がよさそうに思えるかもしれませんが、なぜデザイナーがリードするのか、その理由は2つあります。

ひとつは、大規模なデータといえども、専門家が見るポイントはかなり絞られているためです。なぜなら、専門家は自らの問題意識から生まれた研究の「問い」すなわち関心事に対して、その答えや洞察を得べくデータを分析します。逆に言えば、自らの関心から外れる部分や関連が薄い部分についてはあまり見ないのです。関心のある部分であっても必要以上に複雑な可視化は行いません。データがあるからといって闇雲にあらち眺き込んで調べたりしないのは、研究者としてごく自然な振る舞いであるからです。

一方、データビジュアライゼーションでは一つの視点からではなく、データのより多面的な見方を探る必要があります。これはひとつの問いを深く研究する専門家の使命とすれど、プロジェクトのゴールとうまく結びつきません。ですから、デザイナーが専門家の視点を借りるようにして情報を聞き出し、リサーチしながらその外側へと探索していく必要があるのです。

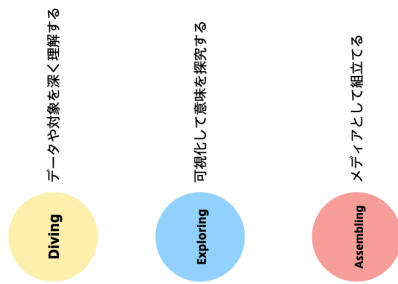
もうひとつの理由は、専門家とデザイナーが常に密接にコラボレーションできるのは限らないためです。すでに書いた通り、さまざまな視点でデータを可視化するのはデザイナーの仕事であり、専門家の仕事ではありません。専門家の本来の仕事は別にあります。ですから、専門家はデザイナーからリサーチされる対象として、またデザイナーが探索した可視化から読み取れる現象の科学的な解釈や、リサーチを深めるための助言をする立場として、要所で力を発揮する存在として位置付ける方法でなければ、プロジェクトの円滑な前進は望めません。とはいえ、デザイナー側にとって都合の良いところだけ専門家の力を借りるようではうまくいきませんから、専門家とのよい関係や環境づくりも大切な仕事として心がける必要があります。

筆者の経験では、よい関係性が築けたプロジェクトの最終段階では、データビジュアライズを見た専門家から「こんなふうに見えるとは意外でした」というコメントをもらうことがあります。デザイナーの感性と専門家の知見がきっちり噛み合った時には、専門家にとっても新しい視点を見出す機会として捉えられるかもしれません。

DEXASの詳細

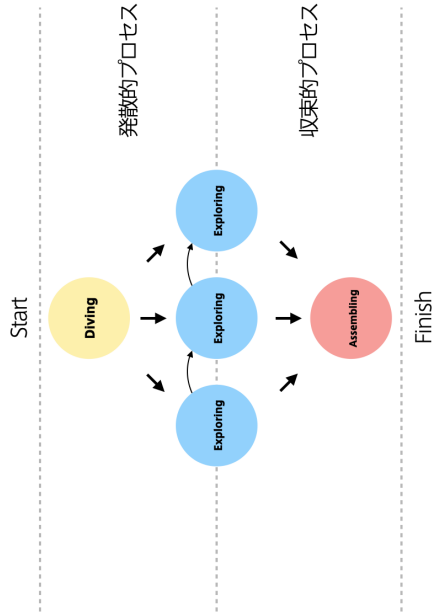
[3つのパートで進める—Diving, Exploring, Assembling]

ここからは、いよいよDEXASの中身について紹介します。DEXASでは、初めに何らかの複雑で大規模なデータが存在している状態でスタートすることを想定しています。図1を見てください。前述の通り、DEXASはDiving, Exploring, Assemblingの3つのパートで構成されています。Divingでは、データにまつわる情報収集をおこなって十分な前提知識を蓄えます。Exploringではデータを可視化してみてもさまざまな意味を読み解きます。AssemblingではExploringで得られた結果を束ね、ひとつの作品としてまとめ上げます。



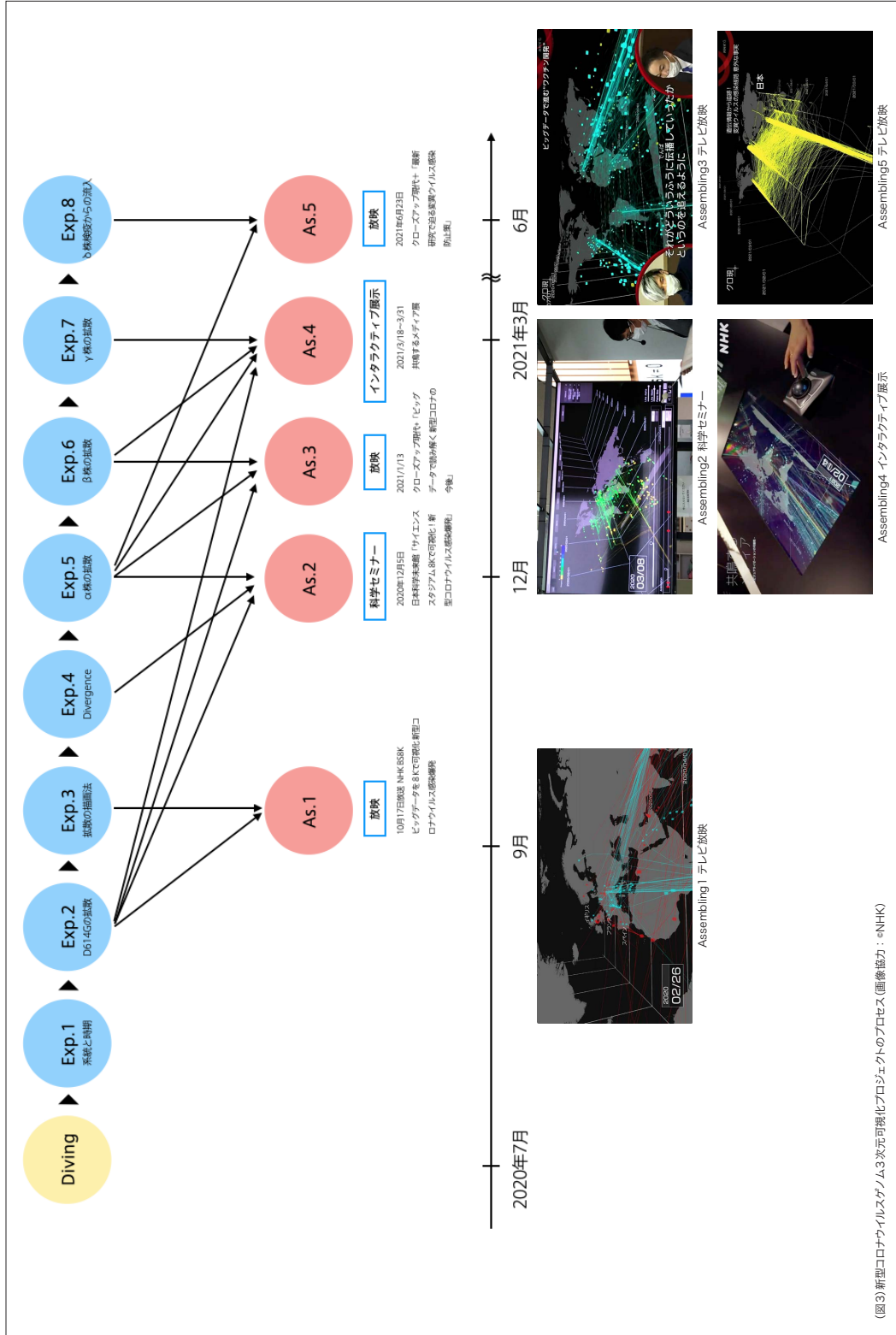
(図1) DEXASを構成する3つのパート

Diving, Exploring, Assemblingの3ステップは、おおむね前から順番に行いますが、必ずしも一直線に進んで完了するものではありません。図2を見て下さい。Divingで獲得した前提知識を手がかりとして、Exploringで多くの可視化の試作をくりかえし実際に目で見てその意味を確認していきます。この試作を重ねてデータが持つ多面性を十分に採り取ってきたら、作品として語るべき内容を検討し、Assemblingでひとつの作品にまとめ、最後にアウトプットするメディアの特性に合わせてデザインしていきます。DivingからExploringに向かうときはできるだけ多くの可能性を探る発散的なプロセスである一方、ExploringからAssemblingに向かうプロセスは内容を厳選する収束的なプロセスであると言えます。



(図2) 発散的プロセスと収束的プロセス

DEXASを具体的なプロジェクトの実例で見てください。図3は前節で紹介した、新型コロナウイルスゲノム3次元可視化プロジェクトの2020年7月から2021年6月までの活動を、DEXASの各パートごとに整理して時系列で表したものです。図から、時問とともにExploringを重ねて多面的な読み取り方が蓄積されていくのがわかります。そして、その蓄積の中から選択されたものがテレビ番組用のCG映像やインタラクティブ展示など、機会に応じて組み立てられているのがわかります。また、この図では省略していますが、Exploringでわかったことをさらに深掘りするためにもう一度Divingに戻ることもあります。COVID-19のように状況が刻々と変化し、それにつれてデータが更新されていくような現象を対象とする場合は、このようにプロセスを反復的に繰り返すことで、持続的で発展的なデータビジュアライゼーションを作ることができます。



(図3) 新型コロナウイルスゲノム3次元可視化プロジェクトのプロセス (画像協力: eNHK)

[3つのパートの進め方]

Diving, Exploring, Assemblingの各プロセスの中身を見ていきましょう。プロセスの中身を示した図4を見て下さい。太枠で囲まれた部分に注目すると、Divingは情報収集パート、Assemblingは組立パートに対応しますが、Exploringは二つのパートに別れています。前半部分である可視化と後半部分であるレビュー、可視化した結果から解釈できる意味をレビューによって精査します。レビューの結果を踏まえて、さらに可視化に磨きをかける必要があるため、この二つのパートは反復的となります。この繰り返しを何度か行い、可視化の質が高まったところで1回のExploringは終了となります。1回のExploringであまり多くの視点や意味を求めすぎはいけません。ここでは限られた意味をうまく表現する可視化をすることに注力し、他の視点に移る時には、Exploringのまともめを行ってから2回目のExploringに進みます。まともめでは、必ず可視化と意味のセットを記録するようにしましょう。情報収集とレビュー（*印）についてはデザイナーと専門家のコミュニケーションが特に重要になるパートです。デザイナーはうまく専門家の力を引き出す必要があります。

[詳細タスクの定義]

次に、図4の細枠で囲まれた部分を説明しましょう。ここは太枠のパートと違い必ず実行する必要があるものとは限りません。太枠のパートでの目的が達せられていけば行わない場合もあります。これらの各タスクの定義は、前述したBen Fryが「Computational Information Design」の中で示した当時のものをベースとして、現在の状況にあうように筆者が修正したものです。

[DIVINGのタスク]

取得

データは解析しやすい形式で提供することが定着しつつありますが、全てがそうであるとは限りません。専門的な用途で収集されたデータは特殊な形式である場合も多く取得後に調査が必要です。

解析

取得したデータから読み取れることは何かを検討し予想を立てるプロセスです。記録されている変数がどのように収集されたもので、どのような意味を持つのかについて理解を深めるとともに、対象にしたい現象の文献等をリサーチすることでドメイン知識を深めることが大切です。また、統計的な手法でデータの全体像を掴んだり、専門家の意見を聞いて可視化の方針を立てることも重要です。

[Exploringのタスク]

フィルタリング

可視化の方針を定めた後、その方針に沿って必要な変数群をデータから抽出します。比較的シンプルな構造のデータであれば表計算ソフトで行える場合がありますが大規模なデータではプログラミングで行う場合があります。これらはデータの前処理に

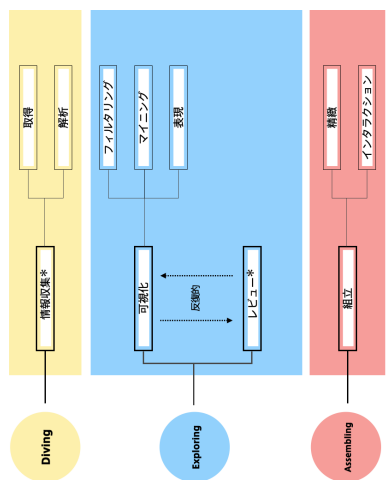


図4 各パート内のタスク

「Assemblingのタスク」

精緻

精緻とは大きく2つのポイントを意識し可視化を微調整することです。ひとつは、鑑賞者が可視化から得る洞察を妨げないように調整することです。鑑賞者が洞察を得るのに全体の構造が大事なのか、特定の部分の比較しやすさが大事なのか目的に応じで取りうる方策は変わります。この点については情報可視化における視覚属性の割り当て方について学ぶと技術の向上が望めます。もうひとつは、驚きや情動を感じられるかという感性的な調整です。大規模な可視化の場合、ひとつの線の微細な太さや色の透明度など微妙な調整で大きく結果が変わります。まさに創造的な視覚表現を行うデザイナーやクリエーターにとっては腕の見せ所です。この点については色彩や構図に関するグラフィックデザインの知識や、CG表現、ビジュアルなプログラミングの知識を学ぶことから技術の向上が望めます。

インタラクション

インタラクションは、2つのポイントを意識して計画します。ひとつは鑑賞者が興味を持った部分についてより深く知るための手段を提供することです。よく知られた方法としてはズームインや、ツールチップなどがあります。もうひとつはデータを別の視点で見るときの手段を提供することです。色や大きさで表現される変数を変更するためのボタンやプルダウン、空間的な表現では視点を変更するためのマウスドラッグなどが一般的ですがこの限りではありません。鑑賞者はその場で多くの操作を学習することができないため、多すぎるインタクティブな逆効果になる場合もあります。

ついて学ぶことで必要な技術を得ることができます。

マイニング

対象となる変数群の状態や概要を把握し、可視化の方針に従って可視化しやすい形式に変換するプロセスです。状態を把握するとは、各変数の記述統計量（最大値、最小値、平均値、中央値、分散、標準偏差など）を調べたり、異常値や欠損値への対応、ヒストグラムや変数同士の相関を見るなどが挙げられます。また、可視化のための変換とは、例えば名義尺度（男性か女性か）などの変数をプログラムで処理しやすいように数値に置き換えたり、グループごとにクラスタリングするなどといった、元データにはない情報へと変換することを言います。これらの知識はデータ分析の基礎を学習することで身に付きます。

表現

前の4つのパートでのリサーチや、得られた結果を念頭に方針を定め、実際にデータを可視化します。ここで選択できる可視化手法は様々ですが、変数の種類や可視化から理解したい内容に応じて適切に選択する必要があります。いきなり独自性の高い可視化手法を模索するよりは、よく知られた手法で手がかりを掴みます。

レビュー

レビューの目的は客観性と感性の両方の質を高めることです。まずは、可視化から読み取れる情報を科学的に解釈することを試みます。視覚的な特徴が現実のどんな現象と対応しているのか、合理的に説明できる状態にすることが大切です。ここでは専門家の知見が大きな役割を果たします。場合によってはDivingに戻ってさらにリサーチしてみるのもよいでしょう。現象の理解が進んだら、可視化の視認しやすさ、美しさ、驚きなど感性を軸とした評価も行います。レビューはデザイナーのひとりで行うものではなく、チームメンバー全員で意見を出し合って行います。

DEXASのチーム構成

DEXASのチーム構成の基本は「専門家とデザイナー」です。ただし、ここでいうデザイナーとは、「情報可視化と視覚表現の知識を持ち、プログラミンクを通してそれらをコントロールできる技術があること」が要件であり、特にデザイナーという肩書きで普段から活動していることは必須要件ではありません(もちろんグラフィックデザイナーとしての実績があればよりよい作品が作れるでしょう)。例えば、クリエイティブ・コーダーと呼ばれる人々の中にも近いスキルを持つ人が見受けられます。チームの基本構成は専門家とデザイナーですが、プロジェクトの規模や、扱うデータの規模、必要な専門性などに応じて構成は変化します。いくつかのパターンを紹介します。

まず最も小さなプロジェクトは、デザイナーのみの場合です。これはデザイナー自身が専門家の視点や知識を持つことが可能な場合に限られます。例えば、デザイナー自身が収集したデータを扱い、客観的な分析ができる場合、あるいは他人のデータを利するならばその分析と情報収集が十分できる場合です。その際、デザイナーはデザイナーの立場と専門家の立場の両方を意識してプロジェクトを進めなければなりません。データを可視化して不特定多数の人々に提示する行為には、それなりの社会的な責任が伴います。ですから「データがあるから可視化して作品にしてみたい」という安易な気持ちで取り組むのは、データビジュアライズとしては不適切です。(あくまで観賞用として、意味が読み取れなくてよい作品を作るのであれば別です)

基本構成にメンバーが加わる場合としては、Divingで情報収集をサポートするリサーチャー、Assemblingで作品に組み込む内容を選択しテーマを明確にするディレクター、Assemblingでのインタラクションの実装や、ハイスベックなデジタル表現を行う場合に、実装を担当するエンジニアやインタフェースデザイナーなどが挙げられます。これらのメンバーは個別の人物である必要はなく、デザイナー自身や、複数のパートを担える人物が担当する場合もあり得ます。大切なことはメンバーが増えてもチームをリードする役目を担うのはデザイナーであり全てのパートに関わることです。

III DEXASパターン・ランゲージ

DEXASパターン・ランゲージとは

DEXASについて、その全体像やメンバーの関わり方について説明しました。ここからは読者の皆さん自身がDEXASのDiving, Exploring, Assemblingの3つのパートを実践しながら学んでいくために、さまざまなデータビジュアライズプロジェクトの現場で得られた経験をもとに、それらに共通する経験則を「パターン・ランゲージ」にまとめました。パターン・ランゲージは、プロジェクトを進行する上で生じやすい典型的な問題について、その状況、解決方法、結果をセットにした「パターン」としてまとめました。個々のパターンには覚えやすい名前がつけられており、プロジェクトを重ねるたびに、そのパターンについて事前に対策を立てたり、メンバー同士で議論することで経験を共有しやすくなります。

パターン・ランゲージの使い方

パターン・ランゲージはマニュアルではなく「コツ」の集合のようなものです。そのため、プロジェクトごとに異なるような固有の性質によらず、似た状況が悪い浮かぶように、内容はあえて具体的すぎない抽象度で書かれています。それぞれのパターンを自分の状況に当てはめてみて、コツにもとづいた解決方法を自ら考え実践に移すことで、自ずと課題が解消され、プロジェクトの質が高まります。そのような「言葉になりにくい経験則」を共有することを目的としているのがパターン・ランゲージです。

パターンは全部で33個あり、プロジェクトの進行とおおむね一致するように順に書かれています。パターンを全て実施する必要はありません。まずは一度、読み物のように前から順番に読むことで、プロジェクトの全体像やそれぞれのパートの状況を生き生きと感じられるはずです。全体を読んでおくと、いざ、その問題が顕在化した時に、どこにその解決の糸口があるのかすぐにイメージできるはずです。また、プロジェクト開始前にメンバーで読み合わせしておくことで、具体的な進め方を考える手がかりにもなります。

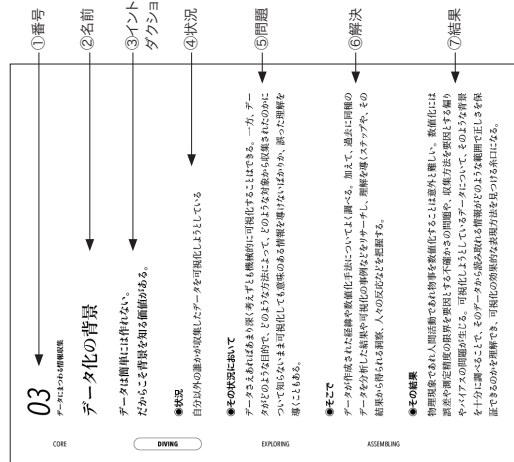
プロジェクトの途中で問題が生じてきたら、似た状況のパターンを再び読み返してみてください。プロジェクトが終了したら、まだ記述されていないパターンがあることに気づくかもしれません。そのような場合は、新たなオリジナルパターンを記述して、加えることでパターン・ランゲージを拡張できます。

パターン・ランゲージの全体像

パターンは5つのカテゴリに分類されつつ、階層的な構造を持っています。図5を見てください。中心にある「カテゴリ0:コア」は全てのパターンが目指すべき方針について書かれています。「カテゴリ1:データにまつわる情報収集」はDivingのため、パターンのため、パターンをまとめられています。「カテゴリ2:データを可視化する」はExploring前半の可視化のためのパターンをまとめられています。「カテゴリ3:可視化をレビューする」はExploring後半のレビューのためのパターンをまとめられています。「カテゴリ4:作品を組み立てる」はAssemblingのためのパターンをまとめられています。1~4の各カテゴリにはそのカテゴリで目指すべきことを代表するパターンが1つあり、カテゴリ内の他のパターンと接続しています。全体として、中心から外側に向かって従って、パターンは具体的な問題について描く構造になっています。

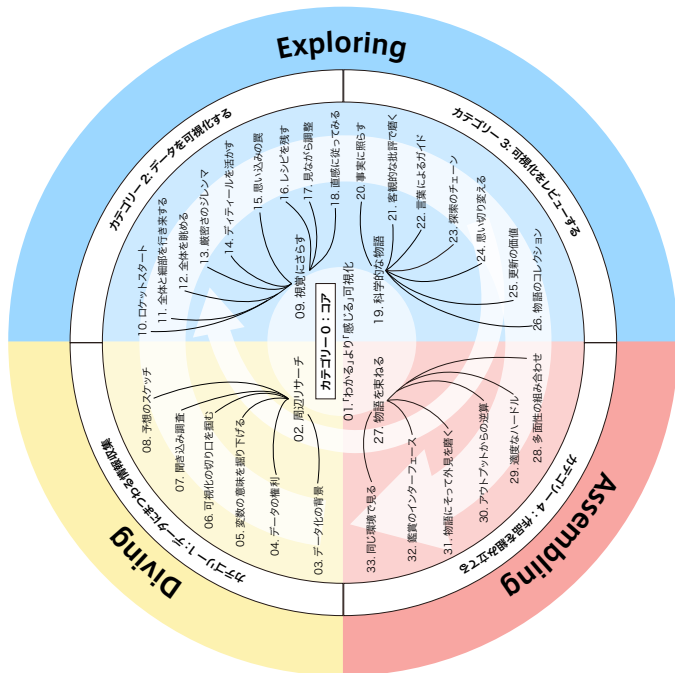
パターンの読み方

パターンは図6に示す形式で書かれています。パターンには①番号、②名前、③イントロダクション、④状況、⑤問題、⑥解決、⑦結果の各要素があり、ここでそれぞれの役割を説明します。



(図6)パターンの各要素と役割

①番号は各パターンの番号です。②名前はチーム内でパターンを共通の理解とする場合のことを考えて、そのパターンを象徴し記憶しやすい短い言葉でつけられています。③イントロダクションはそのパターンで取り上げる事柄の概要をイメージしやすくする短い文章です。次に④状況はこのパターンが示す問題がどのような状況で起こり得るのかを示しています。⑤⑥⑦を讀むと、パターンが取り上げようとしているコトが



(図5)DEXASパターン・ランゲージの全体像

活きる状況が把握できます。これを手がかりに、読者自身の状況に照らし合わせます。
 ⑤問題はその状況で起りやすい問題を説明し、⑥解決はその問題を解決するための考え方が書かれています。コツはあくまでもコツであり、具体的な対処や手段に落とし込むためには、読者の皆さん自身の工夫や現場に合わせた方法が必要です。⑦結果はコツを具体的な行動に移して問題が解消された場合にどのような結果が期待されるかが書かれています。

パターン一覧

CORE

カテゴリー-0:コア

01.「わかる」より「感じる」可視化

DIVING

カテゴリー-1:データに基づいた情報収集

02.周辺リサーチ

03.データ化の背景

04.データの権利

05.変数の意味を掘り下げる

06.可視化の切り口を掴む

07.聞き込み調査

08.予想のスケッチ

EXPLORING 1

カテゴリー-2:データを可視化する

09.視覚にさらす

10.ロケットスタート

11.全体と細部を行き来する

12.全体を眺める

13.厳密さのジレンマ

14.ダイアールを活かす

15.思い込みの罠

16.レンビを残す

17.見ながら調整

18.直感に従ってみる

EXPLORING 2

カテゴリー-3:可視化をレビューする

19.科学的な物語

20.事実にはらす

21.客観的な批評で磨く

22.言葉によるガイド

23.探索のチェーン

24.思い切り変える

25.更新の価値

26.物語のコレクション

ASSEMBLING

カテゴリー-4:作品を組み立てる

27.物語を束ねる

28.多面性の組み合わせ

29.適度なハードル

30.アウトプットからの逆算

31.物語にそって外見を磨く

32.鑑賞のインターフェース

33.同じ環境で見る

(下線は各カテゴリーの軸になるパターン)

「わかる」より「感じる」可視化

誰かが「わかった」ことを本当に自分もわかるためには、誰かの「わかった」を理解するだけでなく、なぜわかったのかを「感じる」ことが大切。

●状況

多様なデータの大規模な記録が可能となり、その分析結果が社会のさまざまな場面で意思決定に重要な役割を持つ時代を迎えている。

●その状況において

データを分析したり機械に学習させた結果として得られるモデルの性能が重視される一方、元データの素性や、モデルが導かれたプロセス、データの多様な解釈の仕方について知る機会は少ない。そのため情報の非対称性は拡大し、多くの人が無意識のうちにモデルに左右される社会になりかねない。

●そこで

データビジュアライズによって、大規模で複雑なデータから魅力的な鑑賞体験を伴った可視化作品をつくる。その複雑性や多様性を顕在化させ、鑑賞者をいくつもの洞察（物語）へと導く。

●その結果

興味と関心を高め、惹き込まれるデータビジュアライズの鑑賞体験を通じて、データの素性や、データに蓄積された対象に対する多様な解釈の可能性について、他者と共有したり議論することができる。それにより、データに内在する問題や解釈の妥当性について、専門家から一方的に押し付けられるのではなく、専門家と一般市民あるいは市民同士が対話する場が形成され、人々がモデルに対して従属的でなく主体的な立場で関わるようになる。

周辺リサーチ

データがなぜ存在するのか。理由や背景を探ることが作品の種になる。

●状況

対象としたデータがある

●その状況において

データを分析したり可視化しても、その意味を読み解いたり、確かさに自信を持つことができないことが多い。

●そこで

データに向き合う前に、そのデータがどんな目的でどのように集められたか、どのように利用されているか、どのようなストーリーを見出せそうかなど、データに関する情報を収集する。

●その結果

人は、単に多くのデータがあれれば信頼に足る事実と近づくことができると思い込んでしまいがちである。しかし、世の中には簡単に数値化できない情報は多く、また、データの対象には偏りや誤差もある。そのため、大量のデータをただ分析したり可視化しても現実を映し出しているとは限らない。データにまつわる情報を丹念に調べることでデータから導かれる結果と現実の差を埋める手助けになる。

03

データにまつわる情報収集

データ化の背景

データは簡単には作れない。
だからこそ背景を知る価値がある。

●状況

自分以外の誰かが収集したデータを可視化しようとしている。

●その状況において

データさえあればあまり深く考えずとも機械的に可視化することはできる。一方、データがどのような目的で、どのような方法によって、どのような対象から収集されたのかについて知らないまま可視化しても意味のある情報を導けないばかりか、誤った理解を導くこともある。

●そこで

データが作成された経緯や数値化手法についてよく調べる。加えて、過去に同種のデータを分析した結果や可視化の事例などをリサーチし、理解を導くステップや、その結果から得られる洞察、人々の反応などを把握する。

●その結果

物理現象であれ人間活動であれ物事を数値化することは意外と難しい。数値化には誤差や測定精度の限界を要因とする不確かさの問題や、収集方法を要因とする偏りやバイアスの問題が生じる。可視化しようとしているデータについて、そのような背景を十分に調べることで、そのデータから読み取れる情報がどのような範囲で正しさを保証できるのかを理解でき、可視化の効果的な表現方法を見つげる糸口になる。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

04

データにまつわる情報収集

データの権利

公開されているデータでも、
著作者や権利の確認を怠らない。

●状況

公開されているデータを利用したい

●その状況において

インターネットで公開されているデータであっても利用方法に制限があったり、データ作成者の利用許諾が必要な場合がある。

●そこで

データの作成者や公開されているサイトを確認し、利用許諾についての条文をよく確認する。

●その結果

近年、データは扱いやすい形式で公開される場合が多く、データがあげばすぐに分析や可視化を開始できる。とはいえ、作成者の利用許諾に反したり、由来が不明確なデータを扱うことは、その分析や可視化結果に責任を負えず大きな問題になる。事前に確認を怠らなければこうした問題を避けられる。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

05

データにまつわる情報収集

変数の意味を掘り下げる

わかったつもりでは、道を誤ることも。

●状況

変数が多いデータを可視化しようとしている。

●その状況において

ログデータのように変数が多いデータは、データから何が読み取れるのか想像がつきにくい。しかし、数値が並んでいるとすぐに分析や可視化に手をつけがちである。例えば気温であっても、その瞬間の気温なのか、1時間の平均気温なのかで解釈の仕方は変わる。

●そこで

データに付随するドキュメントを読み込んだり、データ作成者への聞き取りを行い、変数について理解を深める。

●その結果

その変数がどのような尺度で、どのような方法で数値化されたか知ることで、分析の手法の選択や異常値や欠損値に適切に対処することができる。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

06

データにまつわる情報収集

可視化の切り口を掴む

基礎的な解析から、アプローチ方法が見えてくる。

●状況

データに記録されている個々の数値の意味を理解できた。

●その状況において

データにどんな傾向があるのかわからず、どの変数を可視化したらよいか手がかりが掴みづらい。

●そこで

統計学の知識を応用してデータの傾向を掴む。例えば、代表値(平均値、中央値、最頻値)を調べる。必要であればさらに標準偏差などの散布度や変数ごとの相関を調べてみる。

●その結果

変数が多いところから可視化すべきかの選択が難しく、全てを可視化するのは現実的でない。そのため代表値や散布度などデータの概要、変数同士の関係を探ること、そのデータが持っている意味を想像しやすくなり、探索の手がかりを得やすくなる。また、極端な偏りがある場合など、可視化に適しているかも吟味できる。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

聞き込み調査

専門家とのパートナーシップは、プロジェクトの質を高める。

●状況

専門家から提供されたデータを可視化する。

●その状況において

一般の人にはわからない専門的なデータの場合、どのように可視化したり分析したらよいかわからない。

●そこで

専門的なデータは専門外の人間が調べて理解するには限界がある。その分野の専門家やデータの作成者から直接ヒアリングすることがベスト。データ形式、分析手法や可視化手法は分野ごとに標準化されている場合が多く、その分野に特有のツールで可視化される場合も多い。そうした手法を真似ることも切り口の発見につながる。

●その結果

専門家は、自分が関わった研究が多くの人に知られたり役に立つことは好意的に捉えているため、デザイナーがそうしたことに貢献するつもりで質問をすれば手がかりを掴みやすい。また、専門家は自分の関心事に沿ってデータを分析したり可視化するため、専門家が気づいていないデータの切り口が残されている可能性も少なくない。そのような発見はデザイナーと専門家が共同作業する大きなメリットの一つ。

予想のスケッチ

浮かんだアイデアを、スケッチで外部化する。

●状況

データから読み取りたいことや、可視化したい変数がおおむね絞り込めた。

●その状況において

注目する変数が決まっても、どのように可視化すべきか方針が立てられない。

●そこで

可視化した状態を手描きのスケッチに描き出してみる。描き出すことで、可視化するために必要なデータの処理や加工手段が把握しやすくなる。チームで作業を進める場合であれば、スケッチを作ることで事前の議論もスムーズに行える。

●その結果

データの状態や概要が把握できていれば、可視化後のイメージをスケッチすることで、そのデータから読み取れる物語を予想することにつながる。予想を立てておくことで、可視化の際にプログラミングの間違いに気づきやすくなる効果もある。

視覚にさらす

自分の視覚にさらすことで、心に浮かぶものを捉える。

●状況

世の中には大量のデータがある。データに眠っている意味に人々の意識を向け発展的な議論に結びつけたい。

●その状況において

技術革新によって、膨大なデータが蓄積が可能となり、その分析結果を活用することに注目が集まる一方、分析過程や元データの性質について人々の関心を向ける取り組みは少なく、そのような担い手も少ない。そのため、問題のある分析結果やモデルが普及してしまうこともあり得る。

●そこで

データをできるだけ多くの切り口で可視化して、興味を喚起する形式で提示する。

●その結果

データを多くの切り口で可視化することにより、大規模なデータが持つ多様な解釈についてさまざまな立場の人が認識したり議論できるようになる。また、関心の低かった人々も問題に関心を寄せるきっかけとなり得る。

ロケットスタート

序盤では荷をてらわず、手早く確かな方法で可視化する。

●状況

データにまつわる情報収集を終え、可視化に着手する。

●その状況において

情報収集の段階で可視化すべきさまざまな視点到に気づいた。しかし、どのような可視化からスタートしてよいか戸惑ってしまう。

●そこで

最初は複雑な手法や取扱いの難しい変数を避け、わかりやすい変数を軸に、素早く手堅くできる可視化から始める。専門家の手法や過去の事例に倣ったり、自動の可視化ツールを使うのもよい。

●その結果

いきなり複雑な可視化や先行事例のない可視化を行おうとすると、プログラムのバグを生みやすく、検証も難しくなる。まず、可視化の出力結果を予測しやすい変数に限定して可視化し、検証することでしっかりとした土台を築く。それにより、より高度な可視化への展開が望める。また、シンプルなアウトプットでも可視化されたものがあると、チーム内の雰囲気がよくなり、議論も進めやすい。

11

データを可視化する

全体と細部を行き来する

同じデータでも、マクロに見る時とミクロに見る時には別の物語を発見できることがある。

●状況

これからデータの多面的な切り口を探索しようとしている。

●その状況において

データの特徴的な部分や、はっきりとした傾向だけに関心が向きすぎて多面的な発見ができない。

●そこで

データから掴みやすい傾向や部分的な特徴だけを注視するのではなく、全体を見渡し、その規模感や大きな構造を感じ取ってみる。

●その結果

規模が大きなデータでも、はっきりとした特徴や変数同士の明らかな関係を追求すると、それほど多くの発見は得られないことも少なくない。見方を変えて、全体像から得られる特徴と、細部からわかる特徴を洗い出すと、異なる物語の発見につながりやすい。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

12

データを可視化する

全体を眺める

何かを語るにも、全体の景色を見なければ説得力を欠く。

●状況

初期段階で可視化した結果を見ながら興味深いポイントを探している。

●その状況において

事前に収集した情報だけに頼ると、そこで明らかになっていない注目すべきポイントを見落とす。

●そこで

可能な限り全体的な変数の全量を可視化し、全体像の中で視覚的に特徴的な部分に注目しながらその意味を探求する。全体的な変数を可視化する事が難しければインタラクティブに切り替えられるようにすることも効果的。

●その結果

初期段階の全量可視化には大きく3つの効果がある。①データの規模感と分布を知覚できる。これによりデータの全体像とデータに映し出された現象の多様さを実感できる。②事前に収集した情報や予想した特徴が表れているか確認できる。独自の可視化を行っている場合は、予想と結果を照らし合わせることでプログラムにミスがないか確認できる。③予想外の特徴を把握できる。これらはより深く探索する価値のある「種」に当たる部分であり、可視化を進展させ多くの気づきを得られる可能性がある。

13

データを可視化する

厳密さのジレンマ

厳密さを求めすぎると、見えなくなるものがある。

●状況

大規模なデータの全体像を可視化しようとしている。

●その状況において

全体像を可視化してみたが、データが多いためひとつひとつの数値を読み取りたり比較することが困難になる。厳密に表現できない可視化をしていいのか不安になる。

●そこで

ひとつひとつは厳密に表現できなくても、全体像としてみた場合に実態をうまく映しているかどうかを意識して見極める。

●その結果

特にビッグデータをを用いてデータの全体を可視化するとき、細部の厳密な正確性にこだわらざる方がよい場合がある。そもそもデータは誤差や離散的な要因で現実とは異なる。そのためひとつひとつは厳密に表現できなくても全体像としてみた場合には現実をうまく映し出すことがある。一方、ひとつひとつの誤差が累積的に全体に影響するような場合は、やはり全体像も現実とは異なるため、注意が必要。

CORE

14

データを可視化する

ディテールを活かす

細部をできるだけ潰さずに、全体像を眺める。

●状況

データの全体像を可視化しようとしている。

●その状況において

データ量が多いとメモリ容量や画面の領域が足りなくなり、全体を可視化できなくなってしまう。

●そこで

データを分割して部分的に可視化してから合成し、高解像度のディスプレイで全体を表示する。

●その結果

No12のように、全体を眺めることは多くの洞察を得るきっかけになるが、計算の負荷が高く大きな表示領域も必要となる。一度に表示できない場合は、データを時間や空間を基準に複数のセクションに分割して出力し(画像にしておくのもよい)、結合した状態を高解像度のディスプレイに映し出す。データを間引いて可視化する(複数の値を集計したり、代表する数値に置き換える)と微細な特徴が損なわれやすい。また、画面に一部分の領域だけを表示しスクロールで可視化の全体を見る方法ではなく、全体を一覧できる方法で表示すると全体を把握することにつながりやすい。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

15

データを可視化する

思い込みの畏

大規模なデータの可視化ほど、間違いに気づきにくい。思い込みを捨て慎重に点検する。

●状況

データを可視化するプログラムを書く。

●その状況において

データを色や形に変換するアルゴリズムを独自にプログラミングする時、はじめから大量のデータに対して適用すると結果が正しいか確認することが困難になる。

●そこで

少量のテスト用データを用意して、入力データと出力された可視化を比べて、アルゴリズムが意図した通りに実装できているかを慎重に確認する。

●その結果

特に実装の初期の段階では、変数を少なく絞り、正しく表現されているかひとつひとつを確認するよう気もちで、慎重に確認する。そうすることで、ミス把握しないまま開発が進むことを防ぐ。変数やデータが多くなるほどミスに気づきにくくなるため、こまめな確認を怠らないことで信頼できる可視化のプログラムが実装できる。

CORE

16

データを可視化する

レジビを残す

再現性がなくては、科学的に語れない。

●状況

可視化を探索する途中で興味深い見え方に出会った。

●その状況において

可視化を探索する途中で興味深いイメージを発見したのだが、表示した変数の処理プロセスや可視化のパラメータがわからず再現できない。

●そこで

興味深いイメージに出会ったらスクリーンショットを撮り、画像として記録するだけでなく、その時に使用したデータの処理手順や、色や大きさに変換するパラメータなど、再現に必要な情報も漏れなく記録する。

●その結果

興味深い結果に出会っても、それがどのような意味を含んでいるかについて検討するためには、可視化を再現するための情報が必要となる。それらの情報はさらに探求を深めるためにも必要であるだけでなく、最終的な作品として公開した後でも、追検証のために必要な情報であり、可視化を通じて更なる有益な議論を導く上でも重要な情報となる。

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

17

データを可視化する

見ながら調整

パラメータを連続的に変えながら、スライドスポットを探す。

●状況

感性に響く魅力的な見え方を探索する。

●その状況において

視覚的に魅力的な状態を探る時に、線の微妙な太さや色や形の細かな調整など、なかなかベストな設定(スライドスポット)が見つからない。

●そこで

スライドスポットを探すには繊細な数値の調整を要するため、視覚的な要素に関わる多くのパラメータをインタラクティブに操作するためのスライダー等の GUI を実装する。

●その結果

データから意味や理解を引き出すために必要な探索的なインタラクション(ズームやフィルタ)と、可視化をより魅力的に見せるための視覚調整を行うインタラクション(調整のスライダーなど)は目的が異なる。関心の低い鑑賞者も引き込まれるような感性に訴える見た目は後者の用途のインタフェースを実装することでスライドスポットを探し出しやすくなる。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

18

データを可視化する

直感に従ってみる

理屈でなく「見たい」と思う先には何かがある。

●状況

いろいろな可視化を作る過程で、事前リサーチャーや専門家の意見では触れられていない思いがけない見せ方のアイデアが浮かんだ。

●その状況において

思い浮かんだ可視化から何が読み取れるか確信が持てず、手を動かしてやってみる価値があるのか躊躇してしまう。

●そこで

「どうなるかわからないけど見たい」という好奇心や衝動に素直に従ってみる。

●その結果

労を惜しまず積極的に可視化することのメリットは2つある。ひとつはプロジェクトが進むにつれ、チームメンバーの中でデザイナー自身がデータの特性や可視化表現の可能性に一番敏感な存在になっているため、専門家やチームメンバーが可能性を感じていなくてもデザイナーが直感した可視化には価値がある可能性がある。もうひとつは、多様な視点を持った可視化作品を作るにはできるだけ多くの表現への挑戦が必要であり、選択肢を広げる意味でも、多くの試作が価値を持つ。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

科学的な物語

可視化から読み取れる情報を検証し、科学的な推論に基づいて物語にする。

●状況

データの特徴や傾向を読み取れる可視化ができた。

●その状況において

可視化から読み取れることが、現実とどう対応しているのかうまく説明できない。

●そこで

可視化から読み取れる意味を科学的な視点から検証し、視覚と意味の両面の価値を掘り下げる。

●その結果

驚きや美しさという主観的な価値を持つだけでなく、視覚的な特徴が現実と対応して理解や想像を促すという客観的な価値も持つ可視化を作ることができる。この両面性を備えた可視化をできるだけ多く試作することで、それらを束ね、深みのあるデータビジュアライズ作品を作ることができる。

事実に照らす

可視化から読み取れる情報が、そのまま現実を映し出しているとは限らない。

●状況

可視化の特徴的な部分についてその意味するところを理解しようとしている。

●その状況において

可視化の中で他とは違う状態(疎密や急な変化)が生じる理由をうまく説明できない。

●そこで

想像や思い込みを捨て、事実に照らして合理的に解釈できるまで徹底的に向き合う。

●その結果

可視化の中の特徴的な部分について、「なぜそう見えるのか」という疑問に客観的で合理的な理由を説明できれば、作品全体の説得力を高めることができる。こうした部分を多く含んだ作品ほど、物事の多面性やリアリティを感じさせる質の高いものになる。

客観的な批評で磨く

作り手の思い、苦勞、感情はいつたん脇に置くことも大事。

●状況

試作した可視化とそこから感じ取れる意味をチームでレビューする。

●その状況において

自分では価値があると思つた可視化でも、他のメンバーからの疑問や意見に晒されると、萎縮したりギスギスしてしまう。

●そこで

レビューは作者を批判するものではなく、可視化を客観的な目線で検証して質を高める協働作業だと考える。

●その結果

客観性を大切にすることで、批評する側もポジティブな姿勢で臨むことができる。提起された疑問には感情を脇に置いて科学的な態度で対応する。その場で疑問に答えられなくても、負い目を感じる必要はなく、仮説を立て検証し再度メンバー内でレビューすることが大切。このような繰り返しは回りくどいように感じるかもしれないが、放置すると最終的な作品の鑑賞者も同じような疑問を抱くため、レビューを疎かにせず徹底的に向き合うことで説得力が増し、決して無駄にはならない。

言葉によるガイド

言語化が不十分だと、見た目に流されやすい。

●状況

制作した可視化をチームメンバーにレビューしようとしている。

●その状況において

可視化のレビューの時、視覚的な要素(見た目)にばかり話題が集中し議論が散漫になり、議論の質が高まらない。

●そこで

可視化から読み取れる重要なポイントについて、言葉でも説明できるように準備する。

●その結果

単にグラフィックを見せるだけでなく、表現から読み取れる現象について言葉や数値で説明してグラフィックと共に共有するのがよい。言葉や数値による説明は、可視化のどこに注目すべきか、議論を推進するガイドの役目を果たす。

23

可視化をレビューする

探索のチェーン

レビューで得られた課題や意見は、次の可視化の糧になる。

●状況

レビューを終え、さまざまな意見を受け取った。

●その状況において

レビューの結果を踏まえて、どのように次のステップに進んでよいかかわからない。

●そこで

まずはレビューから得られた問題や指摘を解決するように表現する。その上で、視覚的なパラメータを調整して、その周辺の魅力的な可視化を探索する。

●その結果

問題や指摘を解決するだけでなく、それらを講じた状態ですらに魅力的な表現を探索する「探索のチェーン」によって、グラフィックの魅力と物語の魅力の双方を高めることにつながる。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

24

可視化をレビューする

思い切り変える

伝えたい物語が浮き彫りになったら、固執せず、思い切り表現を変えてみるのもよい。

●状況

可視化をやレビューを終えて、伝えるべき事柄が浮き彫りになってきた。

●その状況において

これまでに作成した可視化では、伝えたい情報が十分に伝わらない可能性に気づいてしまった。

●そこで

作った可視化に固執せず、伝えたい意味に即した表現方法に思い切り変えてみる。

●その結果

探究を繰り返したことで主題が明確になると、それまでの可視化手法に内在する問題が発見されるのは自然なことである。可視化の表現方法次第で視認できる精度に差があるため、狙った主題を描き出すには表現手法そのものを変えることで、飛躍的に理解が促進される場合がある。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

25

可視化をレビューする

更新の価値

データを差し替えることで、さらに探求が深まる。

●状況

内容が更新される可能性があるデータを可視化している。

●その状況において

プロジェクトの途中で新しいデータが利用可能になり、新しいデータに切り替えるかどうか迷ってしまう。

●そこで

データの前処理の手順を自動化しておくなど、データ更新の負担が少なくなるようになるように工夫しておく。

●その結果

近年では、リアルタイム性が高く頻繁に更新されるデータが少なくなない。更新が容易であれば、新旧のデータを比較してその変化から現象の意味を読み解くことができるため、より深い洞察を得ることができる。

CORE

26

可視化をレビューする

物語のコレクション

可視化とそこから見出した物語を、セットにして収集する。

●状況

ひとつの探索を終えて、次の探索に移ろうとしている。

●その状況において

ひとつのデータでも、変数が多いといくつもの可視化ができ、探索を進めるうちに可視化とそこから得られた意味が曖昧になってしまう。

●そこで

ひとつの探索を終えたら、可視化のスクリーンショット、色や形などのパラメータ、そこから読み取れる意味をひとまとめにしてストックしておく。

●その結果

探索の結果を可視化と意味の両方を記録しストックしておくことで、後で見返すことが容易になる。そのようにして得られたストックは、Assemblingで物語を束ねる際に重要な材料になる。

CORE

DIVING

EXPLORING 1

EXPLORING 2

ASSEMBLING

物語を束ねる

得られた可視化を、多面的な示唆を持つデータビジュアライズ作品にまとめ上げる。

●状況

探索(Exploring)結果をコレクションし、多様な見方が得られた。

●その状況において

探索結果をひとつずつバラバラの作品にすると、それぞれの意味同士が関連し合うことが見落とされてしまう。

●そこで

コレクションの意味を吟味して、魅力的な外觀とインターフェースを持ったひとつの作品にまとめ上げる。

●その結果

複数の視点から得られる物語をひとつの作品として束ねることで、単純にわかりやすさを追求する可視化とは異なり、物事の多面性をより深く考える糸口を提供する作品となる。

多面性の組み合わせ

さまざまな面があるからこそ、リアリティを感じる。

●状況

探索の結果で得られた物語を作品にまとめ上げる。

●その状況において

深い専門知識を持たない鑑賞者に、短い時間の中で多くのことを感じ取ってもらうことは難しく、コレクションの中からどのように取舍選択して作品にまとめめるか迷う。

●そこで

コレクションのリストを作り、専門家も交えたチーム全体で議論してみる。似通ったものだけでなく鑑賞者にとって対象の見方が変わるような組み合わせを選ぶ。

●その結果

作品の中に組み込める物語の数には限界がある。その組み合わせ方によっても印象は大きく変わり、意図しないメッセージをミスリードする場合もあることに注意し、さまざまなパターンを吟味する中で決定する。

適度なハードル

鑑賞者が想像する余地を作る。

●状況

作品から読み取れる情報を、鑑賞者の目線で検討している。

●その状況において

鑑賞者にとって、わかりやすい内容になっているか心配になる。

●そこで

わかりやすさだけを重視するのではなく、鑑賞者にとって、もっと知りたくなるようなきっかけとなる程度の「適度なハードル」をイメージしてみる。

●その結果

鑑賞者の経験や知識と照らしてかけ離れすぎるとよくないが、わかりやすすぎるのもよくない。外見のインパクトが先行して印象だけしか残らないのもよくない。適度なハードルをイメージすることで、意外性があり、より深く知りたくなる興味を引き出す、議論したくなるような作品を目指すことができる。鑑賞者が知識として知っている情報であっても、データの可視化を通して、現象の規模感や複雑性をももたせてそれらが意識に登ると、単純に理解していたことと現実のギャップに気づくことが多い。

アウトプットからの逆算

メディアの特性から可視化の組み立て方を考える。

●状況

テーマを決定し複数の視点を持った作品にまとめようとしている。

●その状況において

多面的な視点を取り入れようとしてもどれも捨て難く、適切な数の見当がつかないため絞ることが難しい。

●そこで

最終的な展示メディア(形式)から考え、そのメディアの特性や鑑賞者のインタラクティブ方法などを基準に検討する。

●その結果

展示方法や発表するメディアの形式が決まっている場合は、そのメディアの特性に合わせて作品を構成する必要がある。そのため、メディアの特性をよく理解していることが大切。映像、インタラクティブメディア、グラフィックデザインなど、そのメディアの専門家の見取り込むのもよい。

31

作品を組み立てる

物語にそって外見を磨く

読み取ってほしいポイントを明確にして、ルック&フィールを磨く。

●状況

作品が完成間際まで近づいた。

●その状況において

鑑賞者の目を引くために、視覚表現を必要以上に華美にしてしまう。

●そこで

視覚的に磨き上げる場合、「可視化から読み取ってほしい意味」がより豊かに感じられるかどうかに注意して慎重に行う。

●その結果

豊かに感じられるとは、単純にわかりやすいかどうかというだけでなく、複雑な現象はその複雑さが可視化から滲み出るようなことを指す。指針を持って磨き上げること、「精緻」という言葉のニュアンスに引きずられて過剰に整理したり、必要以上に強調することを防ぐ。ここまで制作に携わったメンバー以外の人に鑑賞してもらい、新鮮な目で見た判断を取り入れることも有効。

CORE

DIVING

EXPLORING

EXPLORING 2

ASSEMBLING

32

作品を組み立てる

鑑賞のインターフェース

制作者が思うほど、鑑賞者は操作してくれない。

●状況

鑑賞者が操作できる展示を計画している。

●その状況において

探索の過程で実装したインターフェイス(ズーム、フィルタ等)が、鑑賞者にとっては操作が難しく、展示してもあまり利用してもらえない。

●そこで

鑑賞用に別のインターフェースデザインを実装し、体験の満足度を高める。

●その結果

鑑賞とデータ探索は全く異なるユーザー体験を伴うため、同じUIを流用できるとは限らない。例えば、鑑賞者にとっては作品に触ることだけでも敷居が高く、さらに、その場で操作を学習するという心理的なコストもあるため、PCで探索する時のような複雑な操作はできない。ユーザー体験のデザイン知識があれば、作品をどのような環境に置くか、鑑賞者はどんな人か、どのように触れられるかについて吟味し、鑑賞体験に即したUIをデザインすることができる。

CORE

DIVING

EXPLORING

EXPLORING 2

ASSEMBLING

33

作品を組み立てる

同じ環境で見る

新鮮な目で、現場で鑑賞者と同じ視線に立って見る。

●状況

作品が完成し公開を控えている。

●その状況において

展示してみると、デバイスの微妙なスペックの違いや、設置される場所の照明環境により、可視化の印象が想定と異なる。

●そこで

鑑賞者と同じ状況で最終確認を行う。現場で可視化の色味や線の太さなどを微調整するため、鑑賞者には見せない調整用のUIを準備しておく。

●その結果

可視化を展示する機材の違いによって印象が変わることは少なくない。特に大量のデータを使って繊細なグラフィックを表示する場合はそうした問題が起こりやすい。最終的に鑑賞者が見る可視化を確認して調整することは、意図しないメッセージを伝えることを防ぐために可視化を作るものとしての責任でもある。

CORE

DIVING

EXPLORING

EXPLORING 2

ASSEMBLING

データビジュアライゼーション デザインパターン
Ver. 1.02
2023/7/1