

Heroku PostgresとTableau Desktopを用いた ニコニコ動画データセットの可視化

輪島 幸治^{1,a),b),c)}

概要：ユーザが情報発信するCGM(Consumer Generated Media)が一般化してきており、ソーシャルメディアなどオンライン上でのコンテンツ共有が増加してきている。CGMは、近年、オンラインコミュニティにおける情報発信は、受信者の趣味やライフスタイルが多様化していることが指摘されている。したがって、受信者に対して、適切なパーソナライゼーションやレコメンドを行うため、網羅的な分析が不可欠とされている。本論文では、オンラインコンテンツ共有サイトの一つであるニコニコ動画における投稿動画のメタデータから、動画投稿者の活動時間に着目してオンラインコンテンツの分析を行う。提案手法では、投稿動画のメタデータを時系列分析タスク、可視化タスクで評価した。結果、時系列分析タスクにおいて、コミュニティの変化が明らかとなった。また、可視化タスクで、動画投稿者の活動特性が明らかとなった。結果を報告する。

1. はじめに

情報通信技術の進歩により、ユーザが情報発信するCGM(Consumer Generated Media)が台頭してきた。CGMの一つにインターネットにおけるソーシャルメディアの一つである動画共有サービスがある。個人を主体に作られ、ユーザが共通性、共同性、連帯性を持ち、形成されるCGMでは、近年、盛んな取り組みが行われている。

一方で、動画共有サービスを始めたとしたソーシャルメディアは、個人が主体で作られるが、そのサービス基盤は、有料会員や広告を始めたとしたビジネスで成り立っている。したがって、ソーシャルメディアを維持し続けるためには、ソーシャルメディアを形成する利用者側の傾向を把握することが必須である。本研究では、情報学研究所が提供しているデータセットである動画共有サービスのニコニコ動画に着目した。

本研究では、投稿動画のメタデータをクラウドデータベースHeroku Postgresに格納して、Tableau Desktopを用いて可視化を行う。分析データには、ニコニコ動画サービスのIR情報、投稿動画のメタデータを用いた。また、Tableau Desktopにおける分析軸には、カテゴリ名、FileType、年月日時間を用いた。集計値は、閲覧数、コメント数、マイ

リスト登録数、動画投稿数、動画の長さ、ファイルサイズである。

本研究における特徴は、メタデータを集計して分析することで、コミュニティにおける全体像を明らかにできること。また、投稿動画の投稿日時に着目することで、カテゴリ別のトレンドや投稿者の活動時間を明らかにできることである。全体像、トレンド、活動時間を明らかにすることで、コミュニティの特性を明らかにすることが期待できる。

評価の結果、ニコニコ動画における各カテゴリの投稿動画数、投稿された動画のファイルタイプの変化、各年ごとの投稿月、投稿日、投稿曜日、投稿時間、各カテゴリの時系列推移が明らかとなった。結果を報告する。

本論文では2章で、本研究の評価対象である、ニコニコ動画について述べる。3章で関連研究に関して述べる。4章では、提案手法の実装と評価対象について述べ、5章で実験結果を示す。最後に6章でまとめと今後の課題を示す。

2. 研究における対象

2.1 ニコニコ動画

ニコニコ動画は^{*1}、2007年に登場した動画共有サイトの一つである。ニコニコ動画は、ユーザが動画に対してコメントを投稿できる独自のコメント機能を持ち登場した。このため、多くのユーザのコメントがリアルタイムに動画に流れているように感じるサービスとして人気である。

¹ ロジカル・アーツ株式会社
Osaka, Osaka 541-0054, Japan

a) wajima@logical.co.jp

b) wajimak@nict.go.jp

c) kwajima@ce.slis.tsukuba.ac.jp

^{*1} niconico(ニコニコ動画) : <https://www.nicovideo.jp/>

ニコニコ動画は、2013 年から日本テレビ放送網株式会社と、日本電信電話株式会社の資本参画があり、2014 年から株式会社 KADOKAWA に経営統合された。近年におけるニコニコ動画サービスの登録会員数、プレミアム会員数を表 1 に示す。また、親会社である株式会社 KADOKAWA の IR 情報を表 2 および表 3 に示す。

表 1: 通期決算説明資料 (単位: 百万円)

年	登録 会員数	有料 会員数
2007	403 万人	14 万 5 千人
2008	982 万人	20 万 8 千人
2009	1,425 万人	51 万人
2010	1,895 万人	101 万人
2011	2,369 万人	139 万人
2012	2,946 万人	175 万人
2013	3,626 万人	211 万人
2014	4,320 万人	236 万人
2015	4,706 万人	244 万人
2016	5,541 万人	256 万人
2017	6,430 万人	243 万人
2018	7,222 万人	207 万人
2019	7,709 万人	180 万人
2020	7,867 万人	163 万人
2021	8,549 万人	153 万人

表 2: 決算情報 - 直近 4 年セグメント別売上高

	出版	映像	ゲーム	Web サービス	その他
2018	112,691	32,554	15,026	29,023	20,821
2019	115,958	30,816	17,534	25,848	22,143
2020	117,303	34,116	14,237	24,739	19,497
2021	129,576	31,314	16,636	22,008	17,463

表 3: 決算情報 - 売上および利益

	売上高	売上原価	売上総利益	売上総利益率
2018	206,785	152,795	53,990	0.261
2019	208,605	151,590	57,015	0.273
2020	204,653	139,793	64,860	0.317
2021	209,947	136,256	73,690	0.351

ニコニコ動画だが、表 1 から明らかだが、登場から現在まで登録会員数は増加しており、15 年で約 20 倍となっている。しかし、ニコニコ動画には快適に動画を閲覧できる有料会員サービスについては、2017 年以降は有料会員数が減少傾向にある。加えて、IR 情報によれば、表 2 のセグメント別売上高からも、ニコニコ動画が含まれる Web サービスの売上高は減少傾向にある。

さて、テレビ番組をインターネットで同時配信する際の権利処理を簡素化した改正著作権法が成立して、2022 年 1 月 1 日に施行された。近年においては、TVer^{*2}のように民放テレビ局の番組をインターネットで閲覧できるサービスも登場しており、従来のテレビだけでなく、アプリや PC など Web ブラウザを利用した動画サービスは広がりを見せている。表 3 から、株式会社 KADOKAWA は、直近 4 年間における売上高は維持できており、売上総利益や売上総利益は上がっている。したがって、登録会員数を多いことからプラットフォーム保有者側は改正著作権法が追い風となり、有料会員サービスが減少した場合でも、有料生放送やニコニコチャンネルなど異なる配信場面における課金機会が期待できる。ここで、一般的な広告は、クライアント、広告会社、メディア（メディア購入会社や紙媒体）と、広告の目標を決めておいてから予算が決まるものとされている [1]。広告における目標では短期的な商品の販売高とは限らず、多くの要素の蓄積を目標として販売目標を到達させる場合や、知名度を目標とする場合がある。ゆえに、プラットフォーム上の投稿動画の傾向分析は、長期的なメディアにおける利用者や要望の把握、プラットフォームの収益化の面で必要とされている。

2.2 Heroku および Tableau Desktop

本研究では、データ格納のためのプラットフォームおよび可視化における分析プロダクトに米国企業のセールスフォース・ドットコム (salesforce.com, Inc.) 社が提供する既存の Heroku および Tableau Desktop 商用プロダクトを用いる。Heroku は、PaaS(Platform as a service) 型のサービスである。Heroku を図 1 および図 2 に示す。

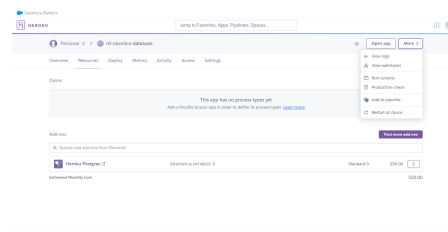


図 1: Heroku Resources

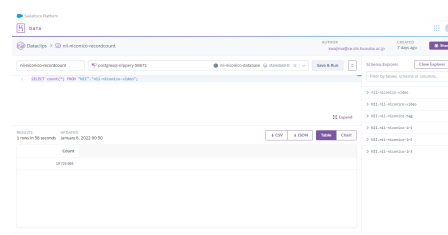


図 2: Heroku Postgres Dataclips

^{*2} 民放公式テレビポータル「TVer (ティーバー)」: <https://tver.jp/>

PaaS 上に各種アプリケーションを Add-ons として追加することで、OS から環境構築を行うことなく、Heroku 上にデータベースを始めとしたアプリ開発、運用するためのツールやサービスを追加でき、迅速に研究環境を構築できる。Heroku 社は 2007 年に創業された企業だが、2011 年 1 月にセールスフォース・ドットコム社に買収され、セールスフォース・ドットコム社のプロダクトとなった。本研究では、Heroku Postgres をニコニコ動画のデータを格納するデータベースとして用いる。

Tableau は、データを視覚化する Business intelligence (BI) アプリケーションである。本研究で用いる Tableau Desktop を図 3 および図 4 に示す。

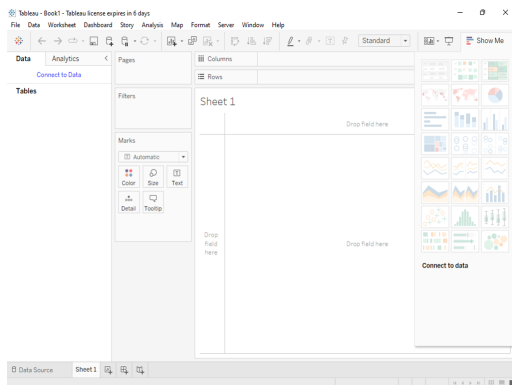


図 3: Tableau Desktop - Sheet1

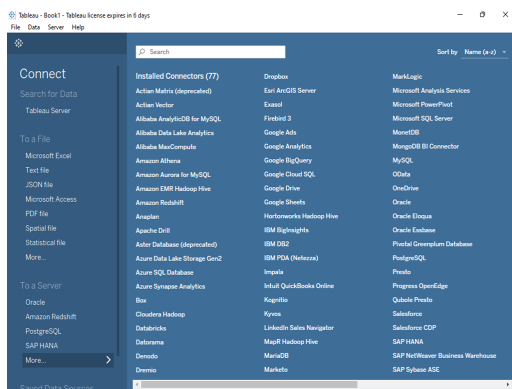


図 4: Tableau Desktop - Connect

Tableau はクライアントアプリケーションの Tableau Desktop だけでなく、Tableau Desktop の可視化結果を一元管理するサーバーアプリケーションである Tableau Server や、クラウド上で使用する Tableau Online なども提供されている。Tableau を用いることで、全 84 種類のデータソースからプログラミング不要かつ一意的なレコードデータの可視化が行える。Tableau Software 社は 2003 年に創業された企業だが、2019 年 8 月にセールスフォース・ドットコム社に買収され、セールスフォース・ドットコム社のプロダクトとなった。本研究では、Heroku に格納されたレコードを可視化するために用いる。

3. 関連研究

3.1 オンラインコミュニティに関する研究

オンラインコミュニティにおける既存研究では、インターネットにおける Web テクノロジーを用いた試みは、開始コストなど参入障壁が低いことから、スタートアップベンチャーなどによって、日々新たな Web サービスやアプリケーションが登場している。インターネットにおけるオンラインコミュニティの研究は、cQA サイト (Community based question-answering service) の研究など、従来より研究が行われている [2]。印象語などのテキスト情報を用いた評価の研究 [3] や検索やランキング [4] などの研究などがある。

3.2 コミュニティにおけるコンテンツに関する研究

コンテンツの共有サイトは、ゼネラル・メディアにおけるラジオ・テレビなど、トップダウン方式の発信ではなく、エンドユーザーが発信することができるボトムアップ方式の発信である。テキストデータから、自動で動画を作成するアプリケーションなども存在している [5][6][7]。3DCG やバーチャル YouTuber (VTuber) モデルで動画配信を行うクリエイター向けのモデル制作サービスも登場してきている*3。したがって、CGM において共有されるコンテンツは、が明確に定められたテレビ番組の制作などと同じ次元で議論できない。ゆえに、管理維持を目的とした収益構造の最適化のために、コンテンツにおけるメタデータの網羅的分析が必要不可欠であると言える。

3.3 ニコニコ動画に関する研究

情報学研究データリポジトリ (IDR) における、ニコニコデータセットの研究は、2020 年 5 月時点で 32 件ある。したがって、ニコニコデータセットにおける投稿動画のメタデータなどに基づいた分析は、盛んに分析されている。ニコニコ動画に投稿された動画におけるクリエイターを検索するシステムの提案 [8] や、相関行動を評価する関数 [9]、動画における特徴シーンを推定して、特徴コメントを推定する研究 [10] などがある。

また、投稿日や投稿されたシーンにおけるコメント数の時系列分析を用いて、不当に高く評価された不公正なビデオを、検出する方法の研究 [11] もある。加えて、データベースにおける全文検索エンジンの検証データセットとしても用いられている [12]。ニコニコ動画は、メタデータに基づいた、オンラインコミュニティとしての分析は、機械学習を含めた網羅的な分析が不足していることから、機械学習タスクや時系列分析タスクが、必要とされている。

*3 デジタル職人株式会社:<https://digishoku.co.jp/>

4. 提案手法

4.1 分析データセットの概要

本研究で用いる評価対象のデータセットは、株式会社ドワンゴがデータセット共同利用研究開発センター（DSC）から提供する 2007 年 3 月 6 日から、2021 年 9 月 30 日時点で削除/非公開されていない投稿動画約 2,000 万件のデータが対象である [13]。対象となっている投稿された動画のメタ情報の件数は、19,712,836 件である。対象となっている動画には、有料のチャンネル動画も含まれている。各取得項目を表 4 から表 5 に示す。

表 4: 投稿動画のメタ情報

No.	属性名	値
1.	video_id	sm9
2.	watch_num	20502606
3.	comment_num	4690391
4.	mylist_num	180576
5.	title	新・豪血寺一族 - 煩悩解放 - レッツゴー！陰陽師
6.	description	(表 5 参照)
7.	category	“null”
8.	tags	(表?? 参照)
9.	upload_time	2007-03-06T00:33:00+09:00
10.	file_type	”flv”
11.	length	320
12.	size_high	21138631
13.	size_low	17436492

表 5: 投稿動画のメタ情報 - Description

レッツゴー！陰陽師（フルコーラスバージョン）

表 6: 投稿動画のメタ情報 - Tag

No.	Tag 名
1.	“”3 月 6 日投稿動画”
2.	“ゲーム”
3.	“プロ野球禁止令”
⋮	⋮
11.	“音楽”

本研究では、提供データセットにおける各項目値をプログラム言語 Python の `json.loads`^{*4}を用いて、JSONL ファイルを分析データセットに変換処理して分析データセットを作成する。

^{*4} Python.org - json — JSON エンコーダおよびデコーダ:
<https://docs.python.org/ja/3/library/json.html>

4.2 提案手法

本研究における提案手法を示す。本研究では、投稿動画から CGM における投稿者の活動時間や年度ごとの活動時間の違いを分析を行うことを目的とする。本研究では、HighlightTable を用いた分析を行う。

このため、本研究では、`upload_time` に着目した。本研究では、`upload_time` から複数の分析軸を作成して、Tableau Desktop による分析に用いる。`upload_time` から作成した分析軸を表 7 に示す。

表 7: `upload_time` から作成した Tableau Desktop の分析軸

No.	Tag 名
1.	year
2.	month
3.	day
4.	hour
5.	weekday_name

提供データセットから取得できる各項目値を表 7 の分析軸で分析することで、各項目値の分析に加えて、投稿時刻に着目した網羅的な分析を行うことができる。具体的には、各カテゴリの動画が投稿される時刻が日中が多いのか夜間が多いのか。また、平日に投稿される動画が多いのか休日に投稿される動画が多いのか。加えて、夏休みの時期や冬休みの時期や月初月末の時期で投稿動画が多くなる時期があるのかなど CGM コミュニティにおける投稿者の活動時刻が明らかとなる。上記の分析軸を各年度で分析することでスマートフォン・タブレット端末の普及によるインターネット人口の増加や Flash 廃止などの要素技術の変化が投稿者の活動に変化があるかを明らかにすることができる。網羅的な投稿動画の分析軸では、コミュニティにおける活動の全体像を明らかにして投稿傾向を可視化する。

5. 実装

5.1 実験環境

本研究では、プログラム言語に Python,

データ集計に、`pandas`^{*5} 集計データを格納するデータベースに Heroku, 分析 BI ツールに Tableau Desktop を用いている。Tableau Desktop のインストール環境には、Windows11 を用いており、データを格納する Heroku Postgres は、レコード件数が約 2,000 万件あることから、“Standard 0”プラン (\$50/月) を使用している。^{*6}

^{*5} pandas - Python Data Analysis Library
<https://pandas.pydata.org/>

^{*6} Heroku 価格:<https://jp.heroku.com/pricing>

6. 評価

本研究では、ニコニコ動画に投稿される動画のメタデータから、オンラインコミュニティ上の投稿者の活動を可視化する。本研究で可視化する投稿者の活動は、動画の投稿カテゴリ、投稿年度ごとの動画の投稿数、投稿年度ごとの各動画のメタデータを集計した値である。

動画の投稿カテゴリからニコニコ動画を使用するコミュニティユーザの好む話題を明らかにする。また、各年度における動画の投稿数から、動画投稿が盛んな年度を明らかにする。本研究では、投稿時点からどの程度、閲覧数・コメント数・マイリスト登録数がされたかを明らかにすることで、以前に投稿された動画の影響度を明らかにする。

加えて、本研究では、動画投稿ユーザの各メタデータと活動年度、月初・月末などの活動日、時間帯、曜日からといった投稿時点の日時に着目して、可視化を行い、コミュニティの活動ユーザの特性を明らかにした。

6.1 ニコニコ動画に投稿される各動画のカテゴリ

まず、ニコニコ動画に投稿済みの動画の各カテゴリの集計値を表8に、ニコニコ動画に投稿される投稿年度ごとの動画のカテゴリを図5に示す。

表 8: カテゴリ別 投稿動画件数

CategoryName	Count	CategoryName	Count
game	9,538,943	dance	200,321
music_sound	2,893,025	r18	193,259
	2,678,377	animal	189,170
entertainment	1,141,324	commentary_lecture	175,142
other	806,155	vehicle	119,586
anime	671,363	cooking	93,740
radio	277,026	technology_craft	91,143
society_politics_news	263,818	traveling_outdoor	85,515
sports	244,312	nature	62,786

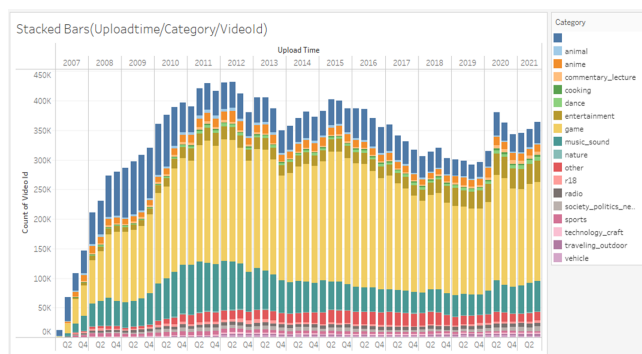


図 5: Stacked Bars Category Count VideoId

表8から、投稿上位のカテゴリは、ゲーム、音楽、エンターテイメントと娯楽に関する動画が多い。

図5の投稿年度ごとのカテゴリ投稿数を可視化しても、ゲームや音楽の投稿動画は毎年多い。このことから、ニコニコ動画に投稿される動画の多くは娯楽向けの動画である。

このため、動画投稿ユーザは、自分自身が好きな動画を投稿しているか、テレビ番組を提供している制作会社などがビジネスとして娯楽向けのニコニコ動画にコンテンツを供給していることが推定される。したがって、ニコニコ動画にコンテンツを供給するユーザの多くは娯楽向けのコンテンツの製作者であると言える。

6.2 投稿動画属性値の各年比較

次に、投稿年度ごとに動画の投稿数、投稿年度ごとのメタデータを集計した値を表9および表10に示す。また、アップロードされるファイルを図6に示す。表9の Watch Numは閲覧数, Comment Numはコメント数, Mylist Numはマイリスト登録数である。また、表10の Countは動画投稿数, Lengthは投稿動画の長さで単位は秒, Size Highはファイルサイズで単位はバイトある。

表 9: 投稿年度ごとの閲覧数, コメント数, マイリスト登録数

	Watch Num	Comment Num	Mylist Num
2007	6,047,262,140	632,077,005	59,408,424
2008	10,188,351,217	611,979,572	106,224,551
2009	9,366,433,767	445,624,931	105,845,937
2010	9,418,943,839	394,938,853	107,185,613
2011	9,152,389,805	392,113,488	107,564,925
2012	10,134,901,646	315,168,663	115,562,188
2013	10,157,609,723	240,106,471	92,703,197
2014	10,085,643,000	217,930,766	80,813,197
2015	11,032,523,515	213,509,729	81,410,423
2016	9,593,410,236	165,862,696	63,452,974
2017	7,741,991,776	133,319,315	49,854,841
2018	6,132,004,607	114,411,845	39,295,434
2019	4,624,193,973	91,813,336	30,563,520
2020	3,786,851,385	102,054,655	23,706,567
2021	1,926,841,880	54,300,490	8,648,065

表9から、閲覧数の総数については、動画投稿から5年以上経過で、閲覧数総数100億前後で頭打ちとなっている。同様にマイリスト登録数の総数については、8年以上経過で、登録数総数1億前後が頭打ちとなっている。コメントの総数については、頭打ちとなっておらず、投稿された年度が古ければ古いほどコメントの総数が多い。

頭打ちとなっている要因だが、閲覧数については投稿された動画を既に見ている。あるいはオンラインコミュニティのユーザが繰り返し見ることに飽きた。もしくは、学習コンテンツのように動画を見る必要がなくなったといった要因が考えられる。また、マイリスト登録数が頭打ちとなる要因は新規で参入してきたユーザが、既に投稿済みの古い動画に興味を持たないといった要因が考えられる。

表 10: 投稿年度ごとの投稿数, 長さ, サイズ

	Count	Length	Size High
2007	335,506	168,401,641	6,183,238,659,776
2008	995,470	542,610,196	20,213,361,419,647
2009	1,212,803	829,302,380	38,929,411,644,760
2010	1,523,573	1,048,816,731	58,477,882,323,908
2011	1,656,878	1,152,508,487	75,237,006,964,247
2012	1,655,115	1,185,827,903	86,070,015,365,246
2013	1,547,372	1,132,791,697	88,935,778,588,706
2014	1,483,333	1,096,871,236	92,644,072,195,620
2015	1,572,217	1,174,791,432	104,913,260,856,437
2016	1,498,823	1,145,609,777	104,156,141,944,062
2017	1,348,519	1,116,746,329	94,920,908,865,775
2018	1,243,173	1,084,577,857	43,163,778,168,642
2019	1,187,530	1,041,489,239	74,057,921,038
2020	1,402,231	1,283,116,333	185,560,304,729
2021	1,062,460	939,804,800	64,304,468,998

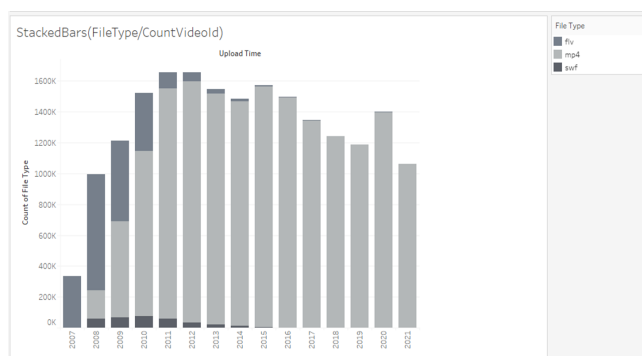


図 6: StackedBars - FileType/CountVideoId

表 9 の動画投稿数は, 毎年 120 万件から 160 万件ほどである. 最も動画投稿数が多いのは 2011 年である. 動画の長さについては, 2010 年以降は毎年約 11 億秒分, 2020 年の場合は約 14851 日, 約 40.7 年分の動画がアップロードされている.

アップロードされるファイルサイズについては, 2007 年から 2010 年までの期間と 2011 年以降を比較するとファイルサイズが大きくなっている. 一方で, 図 6 の結果から, 2010 年以降は拡張子が “flv” の Flash Video のコンテンツがほとんどアップロードされなくなり, Flash Video については, 2020 年 12 月 31 日をもって, アドビによるサポートが終了しており, 現在は Flash Player における Flash コンテンツの実行をブロックしていることから, プラットフォーム上にて技術的な要因による変更があったものと推定される. ファイルサイズについては, 2018 年から 2019 年以降に 43.2 テラバイトから 74.1 ギガバイトに急激に減少した. 一方で, 投稿される動画数に大きな変化はないことから, アップロードされるコンテンツのサイズが小さくなる変更が行われたと推定される.

6.3 HighlightTable 分析 - CountVideoId

投稿動画数を集計値として, 可視化した結果を図 7 および図 8 に示す. 図 7 は縦軸が年であり, 横軸は時間, 日, 月, 曜日の単位で集計している. 図 8 は横軸が曜日であり, 縦軸は時間, 日, 月, カテゴリで集計している.

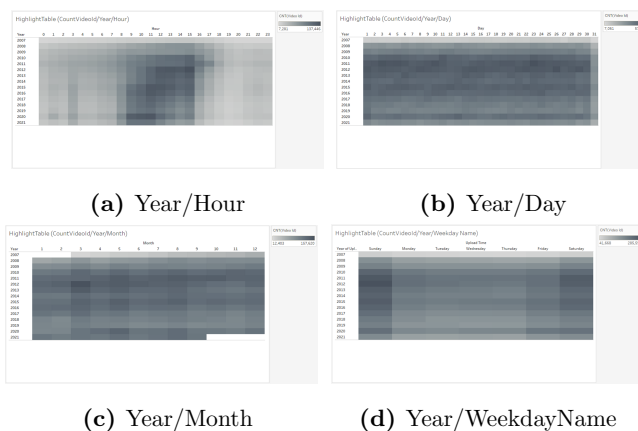


図 7: CountVideoId(1)



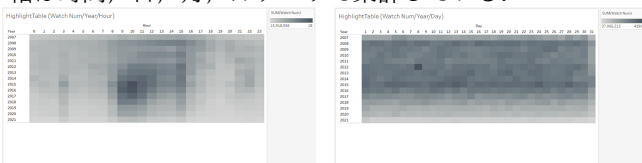
図 8: CountVideoId(2)

図 8a から図 7d の結果について考察する. 年と時間の集計した図 8a では, 朝の 8 時から夕方 17 時までの間の投稿が多く, 投稿年が進むにつれて投稿時間は早まる傾向にあった. 年と日で集計した図 8b では, 各年単位で集計では, 月末や月初など時期による違いは見受けられない. 年と月で集計した図 8c では, 3 月や 8 月については, 他の月と比較して投稿動画が多い. 年と曜日で集計した図 7d では, 土曜日と日曜日の投稿動画数が多い. このことから, 動画投稿の傾向や集中した時期を可視化することができると言える.

また, 図 8a から図 8d の結果について考察する. 時間と曜日で集計した図 8a では, 9 時から 15 時までの投稿が多く, 金曜日の 9 時から 12 時の間が最も投稿が盛んである. 日と曜日で集計した図 8b では, 土曜日の朝 8 時に動画投稿が集中していた. 月と曜日で集計した図 8c では, 各月土曜日と日曜日に動画投稿が集中しており, 3 月や 5 月が顕著だった. カテゴリと曜日で集計した図 8d では, ゲームの投稿動画が多いことが明らかである. このことから, 日や月では, 曜日特有の投稿状況を可視化することができた.

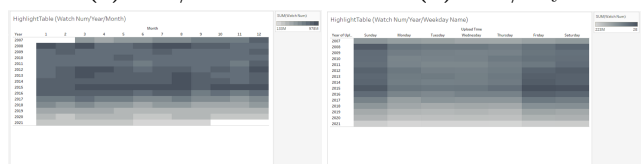
6.4 HighlightTable 分析 - WatchNum

閲覧数を集計値として、可視化した結果を図9および図10に示す。図9は縦軸が年であり、横軸は時間、日、月、曜日の単位で集計している。図10は横軸が曜日であり、縦軸は時間、日、月、カテゴリで集計している。



(a) Year/Hour

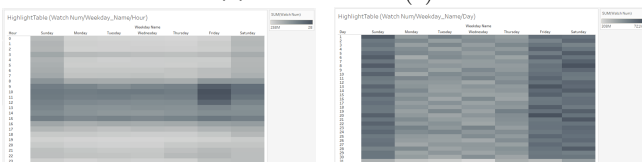
(b) Year/Day



(c) Year/Month

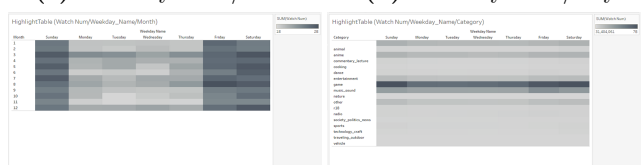
(d) Year/WeekdayName

図 9: WatchNum(1)



(a) WeekdayName/Hour

(b) WeekdayName/Day



(c) WeekdayName/Month

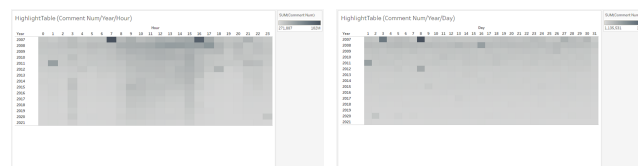
(d) WeekdayName/Category

図 10: WatchNum(2)

図9aから図9dの結果について考察する。年と時間の集計した図9aでは、2014年から2018年までの期間の9時から11時に投稿された動画の閲覧数が多くなる傾向にあった。年と日で集計した図10bでは、各年単位で集計では、月末や月初など時期による違いは見受けられない。年と月で集計した図10cでは、3月や8月に投稿された動画の総数が、閲覧数が多くなる傾向にあった。年と曜日で集計した図??では、金曜日から日曜日の3日間の投稿動画数が多い。閲覧数は、投稿時の話題となり、ピックアップされることによって、閲覧数が増える場合もある。このため、投稿日時で、閲覧数が増える時期などが、可視化できた。

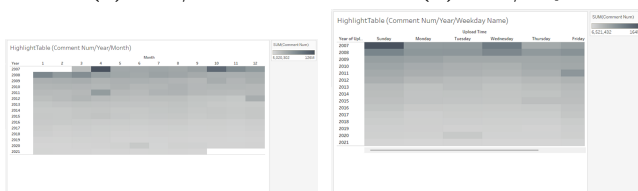
また、図10aから図10dの結果について考察する。時間と曜日で集計した図10aでは、9時から15時までの閲覧数が多く、金曜日の9時から12時の間が最も投稿が盛んである。日と曜日で集計した図10bでは、土曜日の朝8時に動画投稿が集中していた。月と曜日で集計した図10cでは、12月の土曜日に動画投稿が集中していた。カテゴリと曜日で集計した図10dでは、金曜日と日曜日に音楽の投稿動画が多い。このことから、曜日特有の投稿状況を可視化することができたとと言える。

6.5 HighlightTable 分析 - CommentNum



(a) Year/Hour

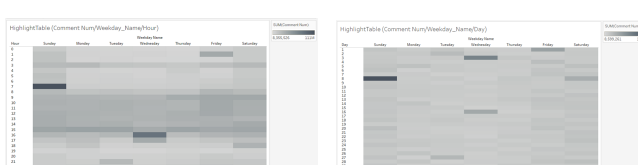
(b) Year/Day



(c) Year/Month

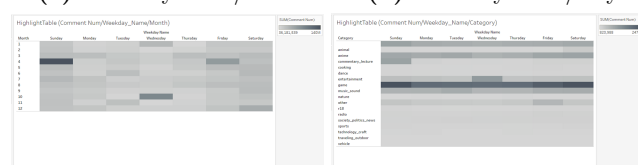
(d) Year/WeekdayName

図 11: CommentNum(1)



(a) WeekdayName/Hour

(b) WeekdayName/Day



(c) WeekdayName/Month

(d) WeekdayName/Category

図 12: CommentNum(2)

6.6 HighlightTable 分析 - MylistNum



図 13: MylistNum(1)

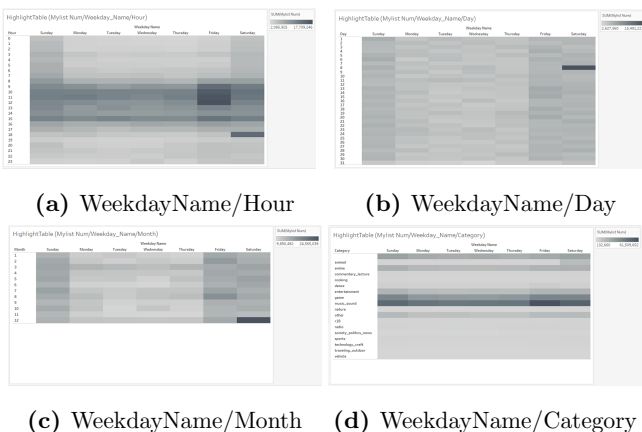


図 14: MylistNum(2)

7. まとめ

謝辞

本研究では、国立情報学研究所の IDR データセット提供サービスにより株式会社ドワンゴから提供を受けた「ニコニコ動画コメント等データ」を利用した。ここに記してデータ提供頂いた株式会社ドワンゴに感謝申し上げます。

参考文献

- [1] 信田和宏. 「電通」成長神話の秘密. 一般社団法人 経済産業調査会.
- [2] 奥村学. ソーシャルメディアを対象としたテキストマイニング. 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, Vol. 6, No. 4, pp. 285–293, 2013.
- [3] 横山友也, 宝珍輝尚, 野宮浩揮, 佐藤哲司. 文章の特徴量を用いた質問回答文の印象の因子得点の推定. 日本感性工学会論文誌, Vol. 12, No. 1, pp. 15–24, 2013.
- [4] 呉鍾勲, 鳥澤健太郎, 橋本力, 川田拓也, デサーガステイン, 風間淳一, 王軼謳. 意味的極性と単語クラスを用いた why 型質問応答の改善. 情報処理学会論文誌, Vol. 54, No. 7, pp. 1951–1966, jul 2013.
- [5] 林正樹. 人々が感動するものは何か. 映像情報メディア学会誌, Vol. 62, No. 7, pp. 1017–1021, 2008.
- [6] 林正樹. 放送 - 通信分野におけるインタラクティブコンテンツデザイン. 映像情報メディア学会誌, Vol. 68, No. 2, pp. 116–120, 2014.
- [7] Masaki Hayashi, Seiki Inoue, Mamoru Douke, Narichika Hamaguchi, Hiroyuki Kaneko, Steven Bachelder, and Masayuki Nakajima. [paper] t2v: New technology of converting text to cg animation. *ITE Transactions on Media Technology and Applications*, Vol. 2, No. 1, pp. 74–81, 2014.
- [8] 優末吉, 洋平関. 音楽のジャンルと印象を用いた vocaloid クリエータの検索. 人工知能学会誌, Vol. 32, No. 1, 2017.
- [9] 知伸尾崎. 分散表現を利用した特徴的相関ルールの抽出. 人工知能学会 知識ベースシステム研究会 (SIG-KBS), 第 B5 巻, pp. 53–58, 2018.
- [10] 智彦末長, 智一早川, 輝雄疋田. 視聴者の時刻同期コメントを用いた動画の特徴シーンの推定. データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2018) 論文集, 2018.
- [11] Time-series analysis of video comments on social media, 2017.
- [12] WG#1PostgreSQL エンタープライズコンソーシアム技術部会. 2017 年度 wg1 活動報告. PostgreSQL エンタープライズコンソーシアム技術部会, 2018.
- [13] 株式会社ドワンゴ. ニコニコ動画コメント等データ. 国立情報学研究所情報学研究データリポジトリ (データセット), 2020.