

令和5年度 メディア芸術連携基盤等整備推進事業  
分野別強化事業

「クレジット情報のメタデータ処理向上とアナログ映像素材へのOCR適用と連絡調整会の組成」  
実施報告書

一般社団法人日本アニメーター・演出協会

令和6年2月

## 目次

第1章 事業概要 .....	3
1.1 目的（目的・背景） .....	3
1.2 実施内容 .....	3
1.3 実施体制 .....	5
1.4 実施スケジュール .....	5
第2章 成果・課題・評価 .....	7
2.1 実施結果 .....	7
2.1.1 自動化／高精度化／メタデータ処理 .....	7
2.1.2 アナログ素材のデジタル処理 .....	9
2.1.3 周知活動 .....	11
2.1.4 アニメ作品情報連絡調整会 .....	12
2.2 結果からの推測・課題 .....	14
2.2.1 利用システムのユニット化 .....	14
2.2.2 AI利用の可能性 .....	15
2.2.3 アナログ素材のデジタルに関する課題 .....	17
2.2.4 配信など多種映像メディアへの対応 .....	18
2.3 その他の特記事項 .....	19
2.3.1 アニメーションスタッフデータベースの意義再確認 .....	19
第3章 実施内容 .....	22
3.1 会議開催 .....	22
3.1.1 定例会議 .....	22
3.1.2 周知活動 .....	22
3.1.3 アニメ作品情報連絡調整会 .....	22
3.1.4 その他会議 .....	23

## 目次

3.2	高精度化／メタデータ処理 .....	23
3.2.1	機能要件・動作概要 .....	23
3.2.2	成果・効果・制約事項 .....	33
3.3	再配布可能なシステム .....	34
3.3.1	機能要件・動作概要 .....	34
3.3.2	成果・効果・制約事項 .....	35
3.4	アナログ素材のデジタル化 .....	36
3.4.1	機能要件・動作概要 .....	36
3.4.2	成果・効果・制約事項 .....	39
3.5	周知活動 .....	42
3.5.1	メカデミア国際学術会議（令和5年（2023年）5月27、28日、29日） .....	42
3.5.2	「マチ★アソビ」Vol.27（令和5年（2023年）10月28日） .....	44
3.5.3	「つづきみ」（令和5年（2023年）9月度より随時） .....	45
3.5.4	コミケットフリンジ研究会（令和5年（2023年）12月30日 .....	47
3.6	採録状況 .....	48
3.6.1	全録サーバ（杉並）実績 .....	48
3.6.2	全録サーバ（千代田）実績 .....	49
3.7	メディア芸術データベースへのデータ提供 .....	50
3.8	学術研究機関との連携検討 .....	51

## 第1章 事業概要

### 1.1 目的（目的・背景）

本事業は、アニメーション（以下、アニメ）のスタッフ情報を収集してデータベース（DB）に記録し、どのような人物や組織がアニメの制作・製作に携わっていたのかを、将来において判別・確認できるようにすることを目的としている。日本が世界に誇る文化であり、同時に重要な産業でもあるアニメが、どのような人材によりどのような環境で作られてきたかを知る上で、こうしたスタッフ名は最重要の情報となるからである。

このため、平成29年度では「リスト制作委員会における採録手法の研究」、平成31年／令和元年度では「アニメ番組の全録画」、令和2年度では「クレジット表示された画面を含む全画面のキャプチャ画像生成」「クレジット表示されたキャプチャ画像を使用してOCR（光学文字認識）を利用したテキスト変換」、令和3年度では「全録サーバから生成したクレジット情報テキスト識別の高精度化とメタデータ処理の自動化の試行」を行った。

令和4年度については、「クレジット情報テキスト識別の更なる高精度化（高識別精度化、高速化、システム化など）」と「メタデータ処理の汎用化（クレジット識別のためのシステム化）」を実施。これら一連の事業を積み重ねた結果、録画したアニメ番組から文字情報が出ている場面を選び出し、キャプチャを行った上でOCRにかけ、テキスト情報を抽出する作業を、高精度でかつ現実的な作業負担で行うシステムの構築・運用について、一定の目処〔めど〕が付いた。

抽出した個々の文字列の属性を判断し、メタデータとして記録する作業についても、辞書となる人名データを用意すれば実施可能との推定を得られた。ただ、近年著しく発達してきたAI（人工知能）を活用すれば、より高精度に文字列の属性を判断する手法の探索が可能との推測も生じていた。

こうした画像データからのテキスト抽出とは別に、元となる映像を全録サーバによる録画以外の、個人が収集・保管しているビデオテープなどのアナログ素材についても、テープや再生機材の劣化によって再生不可能となり、放送時のクレジットデータを抽出・遡及入力できなくなる前に、アナログ素材をデジタルデータ化しておく必要性も高まっていた。

さらに、元となるクレジットデータの入手経路を拡大するために、データを持つアニメ制作会社やアニメ作品宣伝会社、パッケージを販売するビデオソフト会社などとの関係を深め、情報を得やすくする必要もあった。

そこで、令和5年度は「クレジット情報のメタデータ処理向上とアナログ映像素材へのOCR適用と連絡調整会の組成」を試行した。

### 1.2 実施内容

令和5年度の本事業では、「クレジット情報のメタデータ処理向上」「アナログ映像素材へのOCR適用」「アニメ作品情報連絡調整会の組成」に取り組んだ。

放送されるアニメ作品に掲出されるスタッフやキャスト、企業などの名称をキャプチャによって

画面ごと切り出し、OCR によってテキストへと変換して記録しやすくする手法に取り組んでいる本事業においては、OCR の読み取り性能が採録した情報の精度に大きく関わってくる。大量の情報を処理するため、これに携わるコストも問題となってくる。

本事業では、OCR の性能が日進月歩で高性能化する中、クラウド上で提供されているサービスから、より高精度で安価な OCR を選定して組み込み、高品質のテキストデータを得られるよう取り組んできた。加えて、全録サーバに録画されたアニメ映像からキャプチャ画面を抽出し、OCR へ投入し取り出せたテキストを記録する流れを、それぞれのソフトウェアやサービスの結合によってユニット化し、作業手順の省略にも取り組んできた。

今年度もこうした作業を進め、録画データからテキストデータを抽出するまでの流れを、自動的に行えるようにして省力化の実現に前進した。テキスト化に当たっては、画面に段組されたり、複数のブロックで表示されたりする場合もある文字列を、レイアウトを崩さず採録できるようにした。これにより、役職名や役名が人名と関連付けられた形で表示されるようになり、テキストデータをコピーなどによって記録していく作業で、目視によって関連付けを行う負担が軽減された。

メタデータ処理は、抽出されたテキストの属性を辞書データから判別し、データベースへの格納の際に属性に適した場所に入れて検索・利用しやすいようにする上で、欠かせないプロセスとなっている。このとき、AI がキャプチャ画面上に表示されるアニメスタッフに関連する文字列を判別し、レイアウトなども考慮して役職名や人名を拾い出してテキスト化する処理を行う。

ここで、全録サーバの稼働によって採録されたデータや、アニメ研究家の原口正宏氏が主宰するリスト制作委員会によりテレビ放送されたアニメから目視によってスタッフ名を採録する作業によって作成された「リスト DB」を活用し、蓄積された膨大なデータを教師データとして、役職名や人名といった属性も含めて記録可能なメタデータとして処理を可能とする道筋は見えた。

一方で、急速に普及が進み、世界中からのデータ収集により性能も向上している汎用 AI を活用すれば、処理を迅速に行える可能性が浮上してきた。ただし、データが膨大になればその中に未確定の情報も多く含まれるようになり、結果に影響の出る可能性を排除できない。このため、汎用 AI の処理機能を活 [ い ] かしつつ、「リスト DB」など正確性の高いデータを教師データと使用する必要性が見えた。教師データの一般でも使用可能なものとしての配布には、依然として著作権的、コスト的な課題もあるため、解消に向けて取り組む必要性も改めて確認された。

こうした状況から、非営利で公共性のあるデータベースの構築という目的のため、アニメ関連で共通に利用できる情報ソースの確保が必要であるとも判明した。とりわけ遡及入力ソースの一つとして、アニメ番組を録画したビデオテープを始めとしたアナログ素材が、本事業が構築した OCR からのテキストデータ抽出のフローに投入可能となれば、有力と判断した。具体的には、リスト制作委員会の活動として原口氏が録画し保管してきたビデオテープをデジタル化して素材とし、順次テキストデータを抽出していくものとなる。これにより劣化が懸念されるアナログ素材からデータをレスキューできる。ただし、原口氏宅に保管されているテープ類は膨大な本数にのぼり、劣化の度合いや種類も異なっている。このため、今年度は、一部のテープを対象として、劣化が進んでいるのか、記録に問題はないか、デジタル化したデータからキャプチャした画面を OCR で処理できる

かを調査した。

一部のテープに、録画時のビデオデッキと条件が異なるデッキでは再生できないものが見られたが、デッキを調整すれば再生できると分かり、処理にかかる時間なども把握でき、本格的なデジタル作業への道が確認された。

事業ではまた、アニメ関連データベースのより一層の充実につながり、利活用の促進にも寄与する施策として、アニメ業界側からのビジネス情報（宣材やクレジットデータなど）を集約する組織を組成する必要があると判断し、こうした情報を持つアニメ制作会社や製作委員会会社、アニメ関連情報を採録している個人・団体、アニメ情報を発信している雑誌やウェブといった関係者に参加を呼びかけ、「アニメ作品情報連絡調整会」の創設に取り組んだ。アニメ関連イベントや講演や学会への出席を通じて関係者からの理解を得て、第1回の会合開催へと至った。

今年度はほかに、アニメの新番組を四半期ごとに取りまとめて紹介するイベント「僕たちは新作アニメのプロモーション映像を3時間かけて一気に観たらどのくらいつづきをみたくなるのだろうか？」（以下、「つづきみ」）と連係して、情報の交換を始めた。大学など学術研究機関や美術館・博物館などの施設との交流も深め、データ採録への理解推進に努めた。

さらに、アニメ関連の情報が網羅されたDBの存在意義を知ってもらうために、京都市で行われた国際学術会議「メカデミア」や、徳島県徳島市で定期的で開催されているアニメ関連イベント「マチ★アソビ」で講演などを実施。シーズンごとに数十本単位で増えていく新作アニメから、スタッフ・キャストなどの情報を採録するために必要な労力や、こうして集めた情報にも誤記が含まれる場合があるため、より広範にからデータを集める必要性などを話した。

### 1.3 実施体制

コーディネーター：高橋 望（ATAC（特定非営利活動法人アニメ特撮アーカイブ機構））

オブザーバー：大坪 英之（一般社団法人日本アニメーター・演出協会）

関係者：三好 寛（ATAC（特定非営利活動法人アニメ特撮アーカイブ機構））

須山 大介（合同会社 Prod.）

原口 正宏（リスト制作委員会）

磯部 正義（リスト制作委員会）

斎藤 之夫（北海道アーカイブセンター）

谷口 隆一（ライター）

### 1.4 実施スケジュール

本事業の実施スケジュールは以下のとおりである。

令和5年4月21日 記録作成のための方針会議

令和5年5月12日 第7回AI・人工知能EXPO【春】視察

令和5年5月25日 内部会議（第1回）

令和5年5月27日 京都「メカデミア国際学術会議」での聴講  
令和5年5月28日 京都「メカデミア国際学術会議」での講演登壇  
令和5年5月29日 京都「メカデミア国際学術会議」での聴講  
令和5年6月21日 デジタイズのための方針会議  
令和5年6月23日 内部会議（第2回）  
令和5年7月14日 横手市増田まんが美術館訪問  
令和5年7月28日 内部会議（第3回）  
令和5年8月25日 内部会議（第4回）  
令和5年9月より随時 「つづきみ」との情報連係  
令和5年9月29日 内部会議（第5回）  
令和5年10月16日 中間報告会  
令和5年10月25日 内部会議（第6回）  
令和5年10月28日 徳島「マチ★アソビ」でのアニメ DB 関連講演登壇  
令和5年11月15日 Inter BEE 2023 視察  
令和5年11月30日 内部会議（第7回）  
令和5年12月26日 内部会議（第8回）、デジタイズに関する技術会議  
令和5年12月30日 コミケットフリンジ研究会登壇  
令和6年1月12日 第1回アニメ関連情報連絡調整会開催  
令和6年1月17日 OCR 処理からのテキスト化とメタデータ化に関する技術会議  
令和6年1月25日 内部会議（第9回）  
令和6年2月1日 JAniCAにて宣材の受け入れ窓口を開設  
令和6年2月16日 内部会議（第10回）  
令和6年2月21日 最終報告会

## 第2章 成果・課題・評価

### 2.1 実施結果

テレビ放送などのアニメ作品を全録サーバに録画し、スタッフやキャストが表示されている画面をキャプチャしてOCRに投入し、テキストデータとして抽出する作業を続けている中、令和5年度の事業では、「クレジット情報のメタデータ処理向上とアナログ映像素材へのOCR適用と連絡調整会の組成」に取り組んだ。

過去の事業を積み重ねた結果、Excelのシートに記載された録画対象の作品タイトルやエピソードから、個別の録画データにアクセスしてクレジットが表示されている画面をOCRに投入し、テキスト化の処理を行った上で、画像やテキストデータをひとまとめにして保管。テキストデータを見れば記録する必要があるクレジットデータを確認できるよう、一連の作業フローを整備した。

OCRの精度については、1枚の画面の複数箇所ブロックとして表示されることがあるクレジットを、位置関係から個別のブロックごとにひとかたまりのものと判断し、解析を行ってテキスト化する技術をほぼ確立。これによって「役職名ー人名」といった関連性が分かる状態でテキストデータを抽出・閲覧できるようになった。

こうして抽出されたテキストデータに、役職名か人名かを判断した上で属性を付与し、データベース投入など利活用を容易にするために欠かせないメタデータ処理についても取り組んだ。AI（人工知能）の活用により、準備された辞書データを元に一定の属性付与は可能であると判明した。しかし元となる辞書データの構築や、利用の可否に関する判断が不十分である点、汎用AIの進化と普及がメタデータ処理の性能向上に役立つ可能性がある点などから、今後のAIの進展や利用可能な辞書データの構築を視野に入れながら、引き続き開発に注力していく必要が確認された。

今年度はまた、採録データの拡充及び劣化が懸念されるデータ救出を可能とするべく、ビデオテープに代表されるアナログ素材のデジタイズ試験に取り組んだ。アニメーションスタッフデータベースに関する周知活動としては、アニメ情報の採録活動とデータベースの構築に対する理解促進を目的に、アニメ関係のイベントや国際学術会議での講演を行った。さらに、より正確な情報を迅速に入手できる窓口を作り、全録サーバからの収集にとどまらない情報の採録により多面的な判断を下せる環境を整える目的で、アニメ業界側からのビジネス情報（宣伝材料（以下、宣材）やクレジットデータなど）を集約し、同様に採録を行っている団体・個人との連携を図れるようにする組織となる「アニメ作品情報連絡調整会」の組成に取り組み、第1回の会合開催へと至った。

#### 2.1.1 自動化／高精度化／メタデータ処理

全録サーバによるアニメ番組の録画、映像データをキャプチャしてクレジット情報が表示された画像だけを抜き出し、データの容量や作業量を軽減する技術、OCRによるテキスト化に当たってクレジットに関わる文字列だけを読み取った上で、画面上の表示位置を再現するようにテキスト化する技術などを構築してきた。今年度は、こうした一連の作業フローにおけるそれぞれの工程について、簡略化や自動化、高度化に取り組んだ。

簡略化については、全録サーバで録画されたアニメ作品のリストを Excel に列記していく形で把握できるようにした上で、このリストを基に、作業対象となる作品名なりエピソード数を指定すれば、キャプチャ作業へと移行できるよう作業フローをシンプルにした。プログラムで利用しやすい形式の XML (Extensible Markup Language) や JSON (JavaScript Object Notation) といったデータ形式でも処理の実行は可能だが、記載方法について Excel をベースにすれば、どのようなファイルが存在し、どのような処理が行われているかが、プログラミング知識を有しない者でも作業できる可用性の高い仕組みにできると判断した。

キャプチャされた画像については、アニメ番組の全場面の画像ファイルを保管しては、データ容量が膨大となって保存用の HDD を圧迫する。ここから文字列が表示されていない画像を除外し、クレジット情報が表示されている画面でも重複分は除いて OCR に投入し、処理量を減らす仕組みも搭載。こうしてテキスト化された文字情報が、処理済みの画像ファイルと同じフォルダに入って、開けばクレジット情報を閲覧できる流れを作り、録画からテキストの閲覧までの流れを把握しやすくした。

高精度化については、OCR の機能開発に引き続き取り組んだ。本事業では、汎用の OCR として昨年同様に Google Colaboratory を使用している。全世界から常に新しい情報を登録しており、単体でも高い精度でテキスト化を行えるが、対象がアニメクレジットの場合、役職名などで業界独特の語句が使われているため、より高い精度でテキスト化を行うためには、アニメ独自の仕組みを理解したデータセットを用意する必要があった。

そこで、OCR によってテキスト化されたクレジットデータをデータセットに投入できるようなツールを用意した。役職名や人名、団体名といったものを追加して記憶させていった結果、標準モデルでは誤読が生じる語句でもしっかりと把握し、テキスト化できるようになった。大量のデータを用意しなくても、放送された数話分だけである程度まで精度を出せることも判明した。

ただ、本命と目して取り組んできた、テキストデータとして記載された文字列から属性を把握し、人間が判断して振り分けなくても自動的に属性に対応したリストなり DB の必要箇所に格納できるようにする技術については、アニメに特化したより詳細な辞書データが必要となるため大きくは進展しなかった。ChatGPT を始めとした生成系 AI の登場が、属性までも含めたメタデータ解析に有効と見られる一方、進化が激しくどの段階で活用すべきかを見定める難しさ、AI を活用してもやはり正確性が担保された辞書データが必要となる課題などがあって、明確な道筋を策定できなかった。

このほか、画像認識の機能を利用して、キャプチャした画面に表示されるクレジットが幾つかのブロックに別れていても、ひとかたまりの意味ブロック (領域) として把握した上で、表示されているレイアウトを再現する形でテキスト化できるようにした。これにより、役職と名前が関連付けられた形で記載され、これを読み込みながら項目ごとに記載していく作業が効率化された。

### 2.1.2 アナログ素材のデジタル処理

令和 5 年度の本事業では、新たにアナログ映像素材のデジタルに取り組んだ。ここで言うアナログ素材とは、主に家庭用のビデオデッキでテレビ放送から録画されたアニメ作品を指す。

本事業で行っている、テレビ放送されたアニメ作品のクレジットデータを OCR によってテキスト化する作業で、採録されたクレジットデータの正確性などを検証してもらっているリスト制作委員会の原口正宏氏は、家庭用ビデオデッキの普及期から、アニメ作品を録画し、クレジット情報を目視によって記録していく作業を行ってきた。原口氏の作業場には、それらの録画済みビデオテープが現在も大量に保管されている。



図 2-1 リスト制作委員会主宰の原口正宏氏が所蔵するビデオテープ類

このテープの中には、後に再放送されたり、ビデオやレーザーディスク、DVD、ブルーレイといったパッケージとして商品化されたりした際に、クレジットが差し替えられたものもある。パッケージ化自体が行われていない作品もあるため、当時の録画映像からクレジットを採録し、携わった人名や企業名を採録する作業には、歴史をより正確に記録するアーカイブ推進の点から意義がある

と考えられる。

原口氏が所有するテープ類は、家庭用ビデオデッキが普及し始めた 1970 年代から 1980 年代全般、そして 1990 年代といった長期間にわたって録画されたもので、量的に膨大となる。これらを改めて再生しながらの目視でのクレジット確認は、作業量の面から難しい。ここで、本事業が構築に取り組んできた OCR 処理によるテキスト化の技術を使えば、比較的容易にクレジットの採録が可能になる。

そのためには、アナログ素材となっている映像をデジタル素材に変換する必要がある。ただし、原口氏が所有するテープ類を全てデジタル化する作業に当たって、どれくらいの時間がかかるのか、何か問題は発生しないかを検討する必要があった。そこで本事業では、従前よりアナログ素材のデジタル化に取り組んできた北海道アーカイブセンターの参加を仰ぎ、原口氏の作業場より選び出した 30 本ほどのテープ類を実際にデジタル化して、作業にかかる時間や問題点などを検証し、後の本格的なアーカイブ化に資するデータを収集した。



図 2-2 原口正宏宅でのデジタル化用ビデオテープ類の選別作業

ビデオテープに録画された映像素材の中には、録画時期から 40 年近く経 [た] ってビデオテープ

そのものに劣化や破損が生じているもの、録画されている情報に毀損が見られるものがある場合も想定される。アナログ素材のデジタル化作業では、原口氏から預かった素材に関してこうした劣化や毀損も検証。破損などには修繕を施し、録画データそのものの異常については、再生を行うビデオデッキの調整などを通して映像を見られる状態にまで回復させ、その上で専用機器を通してデジタル素材化して OCR が可能かを検証するステップへと回した。結果、デジタル化された映像素材を全録サーバで録画された映像と同様に、OCR からのテキスト化という作業フローに投入して、テキストデータを採録可能と判明した。

ここで、録画時にビデオデッキの録画ヘッドが微小ながら傾いていたため、当該のデッキ以外で再生しようとした場合、トラッキングが合わず再生できないケースがあると判明した。再生にはヘッドの角度を物理的に調整する必要があり、対応に時間を必要とした。

### 2.1.3 周知活動

2023年5月27日から29日まで開催された国際学術会議「メカデミア (MECHADEMIA)」で講演と質疑応答を行った。「メカデミア」は漫画・アニメ・ゲームなどアジアのポピュラーカルチャー全般を扱う国際学術ジャーナルであり、日本や韓国、アメリカなどで会議を開催してきた。今年度は京都精華大学及び同学が運営する京都国際マンガミュージアムと共同で開催。ここに、本事業でオブザーバーを務める JAniCA (一般社団法人日本アニメーター・演出協会) 事務局長の大坪英之が参加し、5月28日に登壇して講演を行い参加者との間で質疑応答を行った。

講演では、メディア芸術データベース (以下、MADB) の歩みや、日本でアニメ作品関連情報ほどの程度まで収集されているかについて内外から集まった学術関係者に説明した。日本で制作されるアニメ作品の数の多さにも言及し、これらからスタッフに関する情報採録の困難さについても紹介した。質疑応答では、30人ほど入る部屋が満席になるほど参加者が集まり、日本におけるアニメ作品に関して、海外の学術関係者が多大な関心を持っている状況が伺えた。

10月28日には、「マチ★アソビ」Vol.27内の企画として、「ufotable CINEMA」を会場に開催されたトークイベント「あなたの知らないアニメデータベースの世界 アニメーターから見る『クレジット』とは」に、大坪英之が JAniCA 理事でアニメーターの徳野悠我とともに登壇した。

「マチ★アソビ」とは、アニメやゲーム、漫画などエンターテインメントに関連する事業者が徳島県徳島市に集まり、イベントや上映会、講演などを開催してアニメファンに作品への関心を持ってもらうためのイベントであり、2009年から主に年に2回のペースで開催されてきた。本事業では、2022年10月15日に開催された「マチ★アソビ」Vol.25でも、トークイベント「あなたの知らないアニメデータベースの世界」を実施。今回の登壇はこれに続くものとなり、アニメーターがクレジットデータをどれくらい重要視しているかを説明して、参加したアニメファンやアニメ業界関係者のアニメスタッフデータベースに対する関心を誘った。

イベントでは、クレジットデータが記録されていると、クリエイターが自身の業務歴を客観的に確認できて、新しい仕事先にリストとして提出してどのような作品に関わってきたかを説明できる

メリットを紹介し、そのためにスタッフリストがデータベースされる意義が話された。一方で、リスト化されたデータも全体の一部に過ぎず、完全に近い履歴を整理するためには幅広く収集する必要があることも伝えた。

12月30日には、世界最大級の同人誌即売会「コミックマーケット103」の開催に合わせて、ポップカルチャーに関する研究発表会「コミックフリッジ研究会」がオンラインで開かれた。大坪英之がこれに参加し、「アニメ分野のデータセットと研究」のタイトルで講演を行った。こちらでも、アニメ分野に関する研究を行う上で必要となる作品情報の収集や記録、そして提供といったプロセスを通じてデータセットが日々蓄積されている現状を説明し、参加者にMADBの研究への活用を働きかけた。より深い研究のための基盤整備の必要性も訴え、データセットの効率的かつ網羅的な収集に役立つ技術を開発している本事業への理解を求めた。

本事業ではこのほか、クレジットデータの網羅的な収集を目的に、「僕たちは新作アニメのプロモーション映像を3時間かけて一気に観たらどのくらいつづきをみたくなるのだろうか?」（以下、「つづきみ」というイベントとの連携も開始した。「つづきみ」とは、四半期ごとに次のクールに放送されるアニメのプロモーション映像を集めて紹介し、新作への興味を誘うイベントとして開催されている。

「つづきみ」では開催に当たって、アニメ作品関連情報やプロモーション映像の提供を行っている企業に依頼して情報や素材を提供してもらっている。全録サーバや「リストDB」からの情報採録に加え、ビデオテープなどのアナログ素材からの採録にも取り組んでいる本事業が、更に広範囲から情報を収集できるようにする上で、作品提供者とのつながりを持つ「つづきみ」との関係が有効と判断した。

「つづきみ」の側でも、増加傾向が収まらないアニメ作品全ての網羅的な紹介に困難さを覚えていた。全録サーバへの録画を通して可能な限り全てのアニメ作品を録画しようと努めている本事業が持つ作品情報を得られれば、イベントの充実につながると判断し、関係が成立した。

#### 2.1.4 アニメ作品情報連絡調整会

令和4年度の事業で、アニメのスタッフクレジット収集に当たり、製作委員会の主幹事会社、制作会社からリスト収集を行っている公的・私的団体までが参加した緩やかな組織を作り、情報を補完し合うような関係を築くための「アニメ作品情報連絡調整会」を組成する準備を進めた。令和5年度は関係各所への声かけを経て、2024年1月12日にオンラインで第1回の会合を開き、正式に「アニメ作品情報連絡調整会」として発足させた。

本事業では、アニメに携わった人名や企業名などを記録し、日本が誇る文化であり産業としても重要なアニメがどのような人材なり組織によって制作されたのかが分かるようにして、アニメ産業の継続的な振興なりアニメスタッフの文化的な価値の永続的な賞賛が果たされることを目指している。

そのために、全録サーバによるアニメ作品の録画からスタッフクレジットの抽出を行ってきた。また、リスト制作委員会の活動として原口正宏氏が中心となって記録してきた「リスト DB」の活用にも務めてきた。

今年度は更に、ビデオテープのようなアナログ素材として残っているアニメ作品をデジタル化して、OCR を使ったスタッフクレジットのテキスト化にも取り組んだ。ただ、こうした活動とは別にアニメ制作会社や製作委員会主幹事会社から情報を入手するルートもあると考え、アニメ制作会社、パッケージメーカー、テレビ局等のプロデューサーらへのヒアリングを通して、クレジットデータを直接提供してもらうための受皿作りを検討の俎上 [そじょう] に載せた。

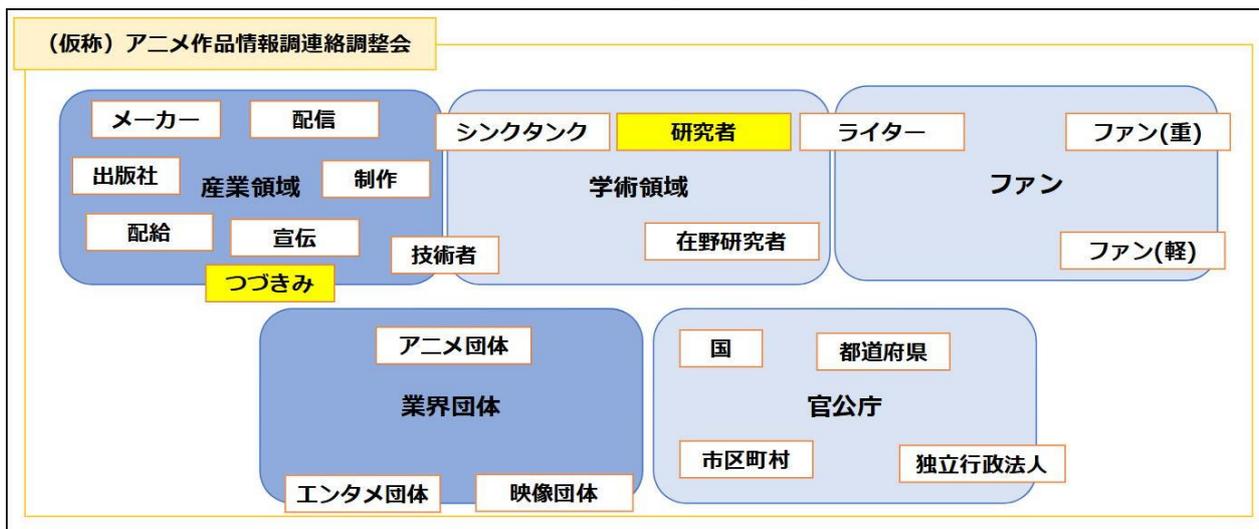


図 2-3 アニメ作品情報連絡調整会の想定参加メンバー

ここで、受皿として浮かんだ組織が「アニメ作品情報連絡調整会」となる。組織の目的は、オフィシャルのキービジュアルや各話スチールなどを私的なリスト採録・発信組織が使う際、公認とまでは行かずとも黙認してもらえる関係を作り、連絡が取り合えるようにすること。情報が滞るなどの緊急時に相互扶助を可能にすることも目的に含まれる。

結果、①メーカーや産業側 以前にヒアリングをお願いした法人、②ウェブや紙の媒体、③アニメ業界団体 JAniCA、ATAC、一般社団法人日本動画協会 (AJA)、④採録している個人・団体等に参加を仰ぐ形となり、JAniCA が発信元となって依頼状を発送。2024 年 1 月 12 日に開催した「アニメ作品情報連絡調整会」で、アニメ作品情報を持つ企業や個人、団体等に協力を呼びかけた。

会合ではまず、JAniCA の入江泰浩代表理事が挨拶し、「本事業がこれまで手がけてきたアニメスタッフデータベース作りという活動アニメ業界の方々に知っていただくことで、それぞれが自分ならこういう取組ができると考え行う行動が、伝播 [でんぱ] していくことが大切だ」と発言し、アニメスタッフの記録に向けた行動が広まることを期待した。その上で、「そうした活動のための場を作ることで、実現への手助けになる」と話し、連絡調整会発足の意義を語った。

アニメ作品情報連絡調整会は、緩やかな組織体として参加者や企業に義務として情報の提供を背負わせるのではなく、それぞれが日々の業務を通して生成した情報や知見を持ち寄り、紳士協定の下で相互に利用できる仕組みとし、それぞれの範囲にとどまっていた情報の流通を促す。JAniCA はアニメ関連の団体としてクリエイターに関する正確なポートフォリオを持っている。また、アニメ産業に関する正確な調査統計情報も収集・発信している。こうした事例を参考として示し、ほかの参加者にもそれぞれの業務範囲で手間のかからない範囲での情報提供が可能だと知ってもらった。

会合では、アニメ作品情報連絡調整会が目指すこととして、JAniCA をアニメの宣材を提供する先の一つとして追加してもらい、情報を流してほしいと依頼した。その後は、それぞれが提供してくれるデータについての解析も黙認してもらえそうな空気感も醸成して、貴重な情報が勝手に商業利用されるような事態への不安を払拭していきたいと説明をした。

その後、参加者からの質疑や要望を受けた。会合には「つづきみ」の代表が参加しており、作品の紹介に当たって、19 項目について回答を入れてもらっているが、宣伝担当者も多忙なためこれらの全てをチェックできない状況と紹介。アニメスタッフデータベースが充実して、19 項目のうちの10 項目について、連絡調整会からデータを得られるようになれば、宣伝担当者の労力も減ってより強い協力が仰げるとの見通しを語った。

今回の会合は、平日の開催でリアルタイムでの参加者が本事業関係者や「つづきみ」関係者に限られたが、会合の様相については今後、オンラインで映像として配信予定。蓄積されたアニメスタッフに関するデータの利活用により、新しい産業なり文化が生まれる可能性があるを知ってもらい、広範囲からの参加を募る。

この連絡調整会開催を経て、JAniCA では2月1日に宣材の受け入れ窓口を開設した。収集対象を「過去（既に公開済み）、未来（これから公開する）の日本で公開されるアニメ作品」とし、作品制作の初出情報（制作開始）、放送開始情報（放送局、時刻、特番等）、放送情報（話の公開直前/直後、放送スケジュール変更）、パッケージ販売情報（発売日、特典映像等）、配信情報（配信サービス、異編集版等）、メディア掲載情報（雑誌、出版等）、関連情報（2.5次元舞台、商品販売）、ニュース（受賞など）などを求めていくと告知した。キービジュアル、作品あらすじ（初出時、公開後、パッケージ時など）、各話あらすじ（公開前、公開後など）についても提供を呼びかけていく。

## 2.2 結果からの推測・課題

### 2.2.1 利用システムのユニット化

Google Colaboratory をプラットフォームに、全録サーバで録画した映像ファイルのキャプチャから不要なファイルの削除、OCR による処理後のテキストデータの抽出に至る一連の流れを構築し、

不要なファイルの削除や OCR 時の学習データを基にした役職名、人名等の把握、画像認識を利用した位置データの記録などを行い、一つのユニットとして全体のシステムを稼働させられるようになった。

ただし、完全に自動で稼働するものではなく、キャプチャ作業を行うための映像ファイルの指定や OCR への投入のための「定義情報の記述」や、処理後のテキストデータからのスタッフデータとしての「分類や同定作業を含めた採録」は人の手で行う必要がある。多くの人がクレジットデータの採録に関心を持ち、より広範囲からのデータ採取を現実的な作業量で実現可能にするためには、録画対象番組のリストアップから最終的なテキストデータのメタデータ化、及びデータベースなどへの搭載をある程度の省力化された状況で行えるようにする必要がある。

ボトルネックとなりそうなのがメタデータ化で、全録サーバで録画した番組を処理する過程で蓄積した学習データ、原口正宏氏が主宰するリスト制作委員会が構築してきた「リスト DB」を活用した学習データの利用によって前進できる目処は立っているが、こうした教師データは収集時に膨大な手間やコストがかかっていることや、公開に当たっての確認事項が数多くあることから一般への配布までの目途が立たず、ユニットそのものは構築できてもソフトウェアの真髓的な部分で課題が残る。

汎用 AI にアニメスタッフに関連する辞書データが搭載される状況になれば、これを組み込んで誰でもスタッフクレジットを取り出せるようにはなりそうだが、広範囲から学習用のデータを集めてくる AI の場合、精度の低い情報に悪影響を受けるため、機械的なメタデータ処理を行っても、信頼性の高い結果を得られないといった課題が残る。公的な事業として行われ得るメディア芸術データベース（以下、MADB）への投入は難しいが、個人が採録を楽しんでアニメのスタッフに関心を抱き、作品への興味を高めて次の作品制作につながるような動きが生まれれば、アニメにとってはポジティブな結果を生む。

こうした情報の信頼性と採録の機動性の双方を利用目的に応じて使い分けられるようになる上で、基本となるシステムのユニット化は大いに意味を持つと言えるだろう。

### 2.2.2 AI 利用の可能性

OCR を始め IT は技術の陳腐化が激しく、本事業でも作業を進めている途中でどんどんと革命的な技術が生まれて、手がけている手法が古くなっていく。これは AI にも当てはまる。現在、ChatGPT のような AI 技術が大きく伸びていて、性能面の向上も著しく進んでいる。言葉を翻訳する技術については、これまでの仕組みが ChatGPT によって駆逐されつつある。汎用性も広く拡張性も高いため、本事業でもこれをどのように使っていくかを検討する必要があるだろう。

ChatGPT は文章の作成や翻訳に限らず、OCR の機能も持っている。性能が向上すれば、これま

でのサービスと比較した上で、画像解析や画像認識に使わざるを得なくなると考えられる。ただし、これまでの事業で積み上げてきた、アニメに関連したデータセットをある程度活用した方が、より有効な結果を得られるものと見ている。

キャプチャ画面に登場していた役職名や人名などのクレジットを、テレビ画面に登場していた形そのままにテキスト化を行いたい場合がある。例えば「キャラクターデザイン/総作画監督 長田好弘」「サブキャラクターデザイン/総作画監督 宗圓祐輔」「総作画監督 坂本俊太」といった具合に、役職名と名前の間を半角ではなく全角で空けたいとき、役職と名前を分離する方法としては、役職と名前を区別できる辞書データの使用方法がある。

役職名を判断した上で、それ以外を名前として判断するなり、逆に名前と判明したもの以外を役職として判断するような処理を行うには、どうしても辞書が必要となる。ここで、辞書を用意しないで、テキストを役職なり名前として分離し抽出する方法について検討し、ChatGPT を試験的に使用した。

OCR から整形したテキストに対して、役職と名前を分離するよう ChatGPT に「Prompt : 役職と名前を抽出してテキストでください。」といった指示を行い、出力させたところ、「役職: キャラクターデザイン/総作画監督, 名前: 長田好弘」「役職: サブキャラクターデザイン/総作画監督, 名前: 宗圓祐輔」といった出力を得られた。

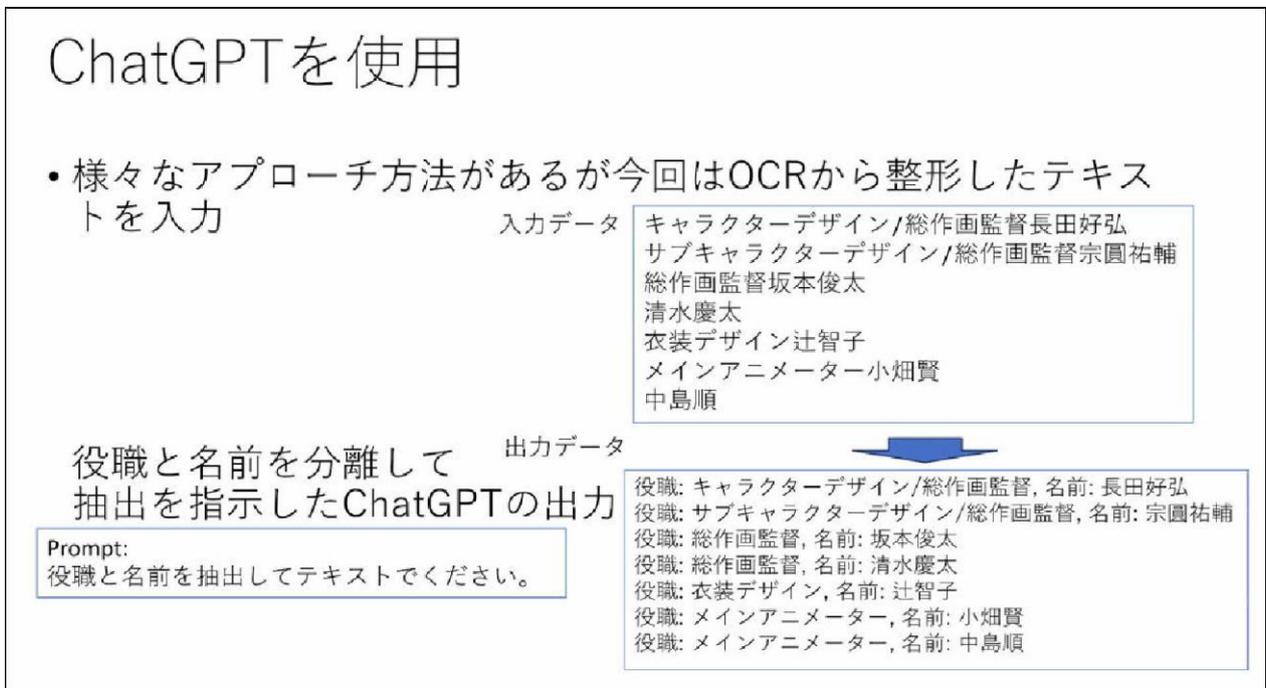


図 2-4 ChatGPT を利用した役職名と名前の分離例

これに加えて、「Prompt：名前を姓名に分けた物」として出力するよう ChatGPT に指示したところ、「役職：キャラクターデザイン／総作画監督，姓：長田，名：好弘」「役職：サブキャラクターデザイン／総作画監督，姓：宗圓，名：祐輔」といった形で出力された。さらに、チャット機能を使って「キャラクターデザインは誰ですか？」と ChatGPT に聞くと、「提供されたスタッフリストにおいて、キャラクターで事案として『長田好弘』が担当しています。」といった答えが返ってきた。姓と名を分けさせて、対話を積み重ねれば、正解に近い回答が得られると想定できた。

課題については、AI は単語の意味や文脈など相関を踏まえての処理は得意だが、レイアウトされた文字組みされた物理的な距離や大きさを考慮した処理は未解決となっている点がある。レイアウトされた文字なりクレジットの場合、その表示位置などからの関係性と意味、属性などの判断が、現状の AI では余り得意ではない。

ただ、画像から理解する能力について AI は大きく進んでいる。Google の AI ではハンバーガーの画像を入力しただけで、ハンバーガーの名前が出力されるようになっていた。こうしたシステムを利用し、ファイルをアップして「スタッフリストを出して」と命令すれば、より正解に近いものを出してくれるようになるまでに、それほど時間はかからないかもしれない。本事業の高度化や簡略化を企図する上で、ChatGPT も含めた AI の進化に期待がかかるところである。

### 2.2.3 アナログ素材のデジタル化に関する課題

原口氏宅から預かったアナログ素材のデジタル化において、その成果として録画状況が良いものは問題なくデジタル素材へと変換でき、そのまま OCR によるテキスト化へと移行可能と分かった。テープ類の破損についても、カセットシェルが劣化によって破損していたり、テープが切れていたりするような状況は発生しておらず、今後発生しても修繕によって対応が可能と推察できた。

ただ、本格的にアーカイブ事業として原口氏に保管されているテープ類をデジタル化していくと、分母が膨大となるため毀損されたものも一定量出てくると思われ、それらの対応に時間をかける必要が出てくる可能性も改めて確認された。

原口氏から預かったアナログ素材には、録画ヘッドが傾いた状態で録画されたため、通常のデッキでは再生できず、別に調整を施したデッキを投入する必要があった。こうしたデッキとの“相性”がごく一部のアナログ素材に限られるのか、当該のビデオデッキを使用していた期間全体にわたって録画された素材全般に言えるのかなども明らかにした上で、対応できる体制を構築する必要があると判明した。

再生は可能でも、クレジットが重なって見えるようなゴーストや、全体に明るく光って文字が読み取りづらくなるハレーションなどが発生する可能性も浮上した。こうした現象は、デジタル化の過程での緩和が可能だが、発生頻度は低くても全体量との兼ね合いで発生件数が多くなれば、処理などにかかる機材や人的な対処も必要となってくる。

全体として、アナログ素材のデジタル化から OCR 処理を経てのクレジットデータのテキスト化

という作業フローの構築が可能であると確認できた。今後は更に取り扱うアナログ素材の数を増やし、これまでに確認されていないような支障が発生するかを確認していく必要がある。それらの対策を講じた上で、本格的なデジタル化作業へと入っていけば、クレジットデータに限らず番組そのものがほかに現存を確認できず、希少性の高いものもあるアナログ素材のデジタル化も可能となり、アーカイブの推進と拡充という施策に資する。

#### 2.2.4 配信など多種映像メディアへの対応

本事業では、年間に制作・公開される作品のうち、本数で約 8 割を占めると推測されるテレビアニメ作品について全録サーバにて録画を行い、蓄積された映像からクレジットデータを採録する作業を行っている。採録に際しては、映像の中でクレジットデータが表示されている場面をキャプチャして静止画として記録し、OCR で処理してテキスト化している。令和 2 年度から取組を始め、テキスト化の精度向上と、工程を可能な限りの自動化に努めた。令和 5 年度までの取組で、テレビアニメのクレジットデータについては、自動化による省力化の下、高精度で採録できるようになった。

しかし、近年はテレビ放送以外のプラットフォームを使ってアニメ作品が提供されるケースが増えている。一つが劇場で、映画として公開されるアニメ作品が年間で 100 作品近くに及ぶ状況となっている。また、Netflix や AmazonPrime ビデオ、Disney+といった配信サービスで独占的に提供される作品もある。このような、テレビ以外のメディアプラットフォームから提供されている作品についても、法的・技術的・道義的な観点に配慮した上で、全録サーバに映像データを格納すれば、テレビアニメ作品と同様にクレジット情報の抽出は可能だと言える。

パッケージ化された劇場作品や、オリジナルビデオアニメ（OVA）などについては、これまでも原口正宏氏とリスト制作委員会がパッケージを入手した上で、目視と手作業でスタッフ・キャスト名を抽出してスタッフデータベースを作成してきた。本事業では、こうした個人のスキルに依存した体制の自動化と効率化により、広範かつ迅速にスタッフデータベースを構築する意味から、作業の一部を OCR によるクレジット情報の抽出に置き換えていく取組を始めている。

これにより、リスト化されるデータの範囲が広がり、データベースへの投入も迅速化されて、MADB を利用した調査・研究の促進や新規事業の展開に資すると期待できる。同時に、リスト制作委員会が負担してきた作業の軽減にもつながり、原口氏の持つアニメスタッフやアニメスタジオに関する知見の、学術的な分野での活用にもつながる。

配信作品については、本事業でもリスト制作委員会側でも人的リソース配分の問題や、法的・技術的・道義的な面への配慮から、全録サーバにデータを収録した上でのテキストデータの抽出といった作業は行っていない。しかしパッケージ化されない作品や、新規サービスの登場もあって配信ビジネスの拡大でアニメ作品の提供形態が大きく変化している状況で、網羅的なクレジットデータの採録のため、欠落部分を埋めるという課題解決に向けて、配信作品の採録についても本格的な検討が必要となっている。

本事業では、こうした配信や劇場での流通作品とは別に、個人宅に保管されている家庭用ビデオデッキを利用した録画作品についても、デジタル化を経て OCR 処理からのテキストデータ採録を検

討した。今年度から原口氏宅に保管されているビデオテープ類の一部をデジタル化し、ここからテキストデータ採録が可能かを検証するテストを行った。結果は良好で、今後は膨大な量となる録画テープを効率的に処理していく作業フローの確立や、録画状況が悪く所有の機材では再生が難しい映像の修復について、検討を進めていく。

本事業が開発・整備に取り組んでいる、「映像から文字情報を効率的に取り出す」技術はアニメ作品に限らず応用範囲が広く、ドラマやバラエティといった番組のデータベース構築などにも活用できる。こうした分野でもテレビ放送以外のプラットフォームでの提供が増えており、本事業を通してプラットフォームの拡大に向けた取組を進めることは、あらゆる映像から様々な文字情報をテキスト化して保存し、利活用できるようにする上で多大な意義があるものと考えられる。

## 2.3 その他の特記事項

### 2.3.1 アニメーションスタッフデータベースの意義再確認

全録サーバで録画した番組からのクレジットデータ採録、原口正宏氏が構築してきた「リスト DB」の変換によるクレジットデータの確保、及びアナログ素材のデジタル化によるクレジットデータ採録といった作業に道筋が付き、アニメ制作会社やパッケージ会社、アニメ作品宣伝会社から作品に関する情報を定期的に入手する「アニメ作品情報連絡調整会」の組成にも目処が立った。より正確で内容的にも広範なアニメーションスタッフデータベースを構築する上で、必要とされる作業体制はおおむね整ったと言える。

アニメーションスタッフデータベースが持つ意義は、作品に携わった人名なり企業名の記録により、将来において関係者の経歴を参照する際に詳細なリスト化を可能にする点にある。また、個々のスタッフが新しく仕事を得る上で、過去に携わった作品について公的な機関によって整えられたリストを示せば、より深い信頼を得られ、仕事の獲得につながるといった利点もある。



りに表示されるクレジットデータを参照して構築された場合、後々まで堀越氏の名前は番組スタッフとして登場しない事態となる。DBの正確性においても、個人の経歴を再確認する場合においてもこうした欠落はない方がよい。したがって、欠落を埋めるための作業が必要となるが、ここで本放送時に録画された映像を改めて再生して確認したくても、現存していなければ確認できない。

原口氏が保管している膨大な録画テープの中に含まれている可能性もあるが、そこに現存していなかった場合でも、広く呼びかければテープの発見につながる。その際に、アニメーションスタッフデータベースの存在や、それが実際に利活用されて個人のキャリアの再確認なり、次のキャリアの構築に役立っている点をアピールすれば、より前向きな協力が得られるだろう。

ここで、アニメのスタッフを漫然とリスト化するだけでなく、どうしてリスト化すべきなのかを改めて問い直す必要性を付記しておく。スタッフとしてクレジットに名前が掲載された人たちは、それぞれにアニメ作品の中で欠かすことのできない働きをしている。そうした働きぶりを理解した上で、個人としての名前を記録することで、スタッフと作品に対する敬意が生まれる。DBを作成するとは、そうした敬意の積み重ねであることを理解しておかなければ、信念を持ち長く作業の継続は難しくなる。

アニメーションスタッフデータベースの構築にはもちろん意義があるが、構築する行為そのものに意義があるのだと再確認した上で、より精緻で広範なアニメーションスタッフデータベースを作り残していくことが、アニメという日本にとって有数の文化であり産業が、これからも長く続く上で必要だと言える。

## 第3章 実施内容

### 3.1 会議開催

#### 3.1.1 定例会議

今年度は、遠距離からの参加者もあることから、ネット会議ソフトのZoomを利用して参加メンバーによる月次の定例報告会議を開催した。

令和5年5月25日 内部会議（第1回）  
令和5年6月23日 内部会議（第2回）  
令和5年7月28日 内部会議（第3回）  
令和5年8月25日 内部会議（第4回）  
令和5年9月29日 内部会議（第5回）  
令和5年10月25日 内部会議（第6回）  
令和5年11月30日 内部会議（第7回）  
令和5年12月26日 内部会議（第8回）  
令和6年1月25日 内部会議（第9回）  
令和6年2月16日 内部会議（第10回）

#### 3.1.2 周知活動

アニメのデータ収集とDB構築が持つ意義を知ってもらい、活動に理解を得るために、アニメファンが集まるイベント会場や、アニメ研究者が参加する会議に登壇した。また、四半期ごとに新作アニメのプロモーション映像を集めて紹介し、アニメファンの視聴を誘うイベント「つづきみ」に情報提供で協力した。

令和5年5月28日 京都「メカデミア国際学術会議」への登壇及び質疑応答  
令和5年10月28日 徳島「マチ★アソビ」でのアニメDB関連講演登壇  
令和5年12月30日 コミケットフリンジ登壇  
令和5年9月より随時 「つづきみ」との情報連係

#### 3.1.3 アニメ作品情報連絡調整会

全録サーバを活用しての放送作品からスタッフ・キャスト情報の採録とは別に、こうした情報を持つアニメ制作会社や製作委員会会社、宣伝会社から情報を入手できるように関係を構築するとともに、同様にアニメ関連情報を発信している雑誌やウェブ媒体、独自にアニメ関連情報を採録している団体・個人ともつながりを持ち、情報交換や緊急時の相互扶助を行えるようにして、それぞれに充実した活動を進めていける仕組みとなる「アニメ作品情報連絡調整会」の創設を企図。年度を通じて声かけや活動の周知を行い、参加者を得て2024年1月12日に第1回の会合を開催した。

令和 6 年 1 月 12 日 第 1 回アニメ作品情報連絡調整会開催  
令和 6 年 2 月 1 日 JAniCA にて宣材の受け入れ窓口を開設

### 3.1.4 その他会議

上記以外の会議については適宜行った。

令和 5 年 4 月 21 日 記録作成のための方針会議  
令和 5 年 6 月 21 日 デジタイズのための方針会議  
令和 5 年 10 月 16 日 中間報告会  
令和 5 年 12 月 26 日 デジタイズに関する技術会議  
令和 6 年 1 月 17 日 OCR 処理からのテキスト化とメタデータ化に関する技術会議  
令和 6 年 2 月 21 日 最終報告会

個々の作業においては、ネットワークを接続して、千代田サーバ、杉並サーバで録画した番組を基に、OCR 化の効率化や生成されるデータの精度向上に必要なプログラムの作成、それぞれに採集できたデータの検証、リスト DB 収録データのメディア芸術データベース向けデータへの変換プログラム作成、リスト制作委員側の要請による不足データの提供、原口氏宅に保管されているアニメ番組を録画したビデオテープの選別及び北海道アーカイブセンターへの送付などを行った。付随する相談、検討についても随時行った。

## 3.2 高精度化／メタデータ処理

### 3.2.1 機能要件・動作概要

全録サーバで録画されたアニメ作品のリストを Excel に列記していく形で把握できるようにした。作業時にはこのリストから、作業対象となる作品名なりエピソード数を選べば、キャプチャ作業へと移るようになった。

	A	B	C	D	E	F	G	H
1	title	broadcast	epno	edst	eden			
2	16bitセンセーション ANOTHER LAYER	AT-X	1					
3	ブルバスター	AT-X	1					
4	ブルバスター	AT-X	2					
5	ブルバスター	AT-X	3					
6	ブルバスター	AT-X	4					
7	ブルバスター	AT-X	5					
8	ブルバスター	AT-X	6					
9	ブルバスター	AT-X	7					
10	ブルバスター	AT-X	8					
11	ブルバスター	AT-X	9					
12	ブルバスター	AT-X	10					
13	ブルバスター	AT-X	11					
14	ブルバスター	AT-X	12					
15								
16								
17								

図 3-1 録画されたアニメ作品のリスト

キャプチャされた画像については、クレジット情報が出ている画面だけが残るようになっており、これを OCR に投入すれば、画像ファイル及びクレジット情報を抽出したテキストファイルが生成されるようにした。

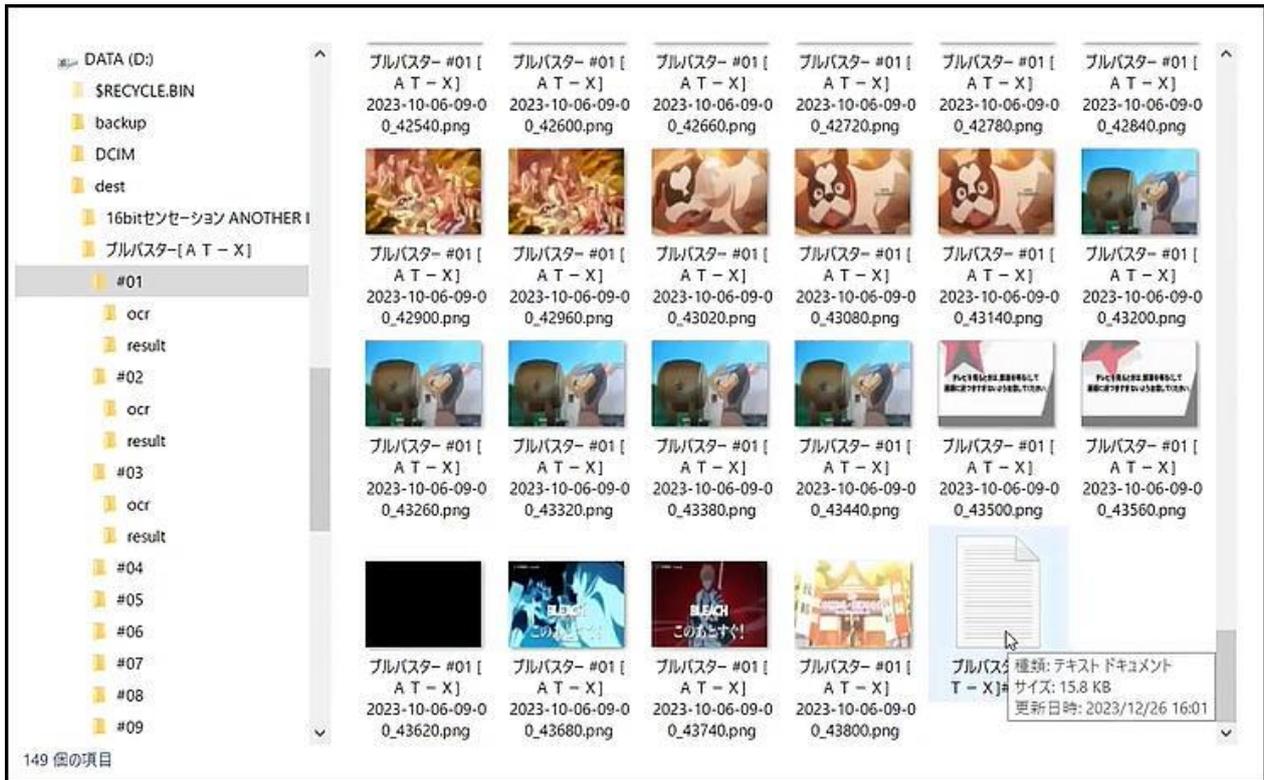


図 3-2 キャプチャされ OCR 処理された画像ファイルと抽出されたテキストファイル

こうした処理の途中で、同じような画面でも表示されているクレジットの内容が違うものが連続して処理される状況も起こるが、差分の見極めにより、同じように見える画面でも、違う文字列が認識できた場合はその旨を指摘するような仕組みを取り入れて、同じ画面だからといって見落としがないようにした。

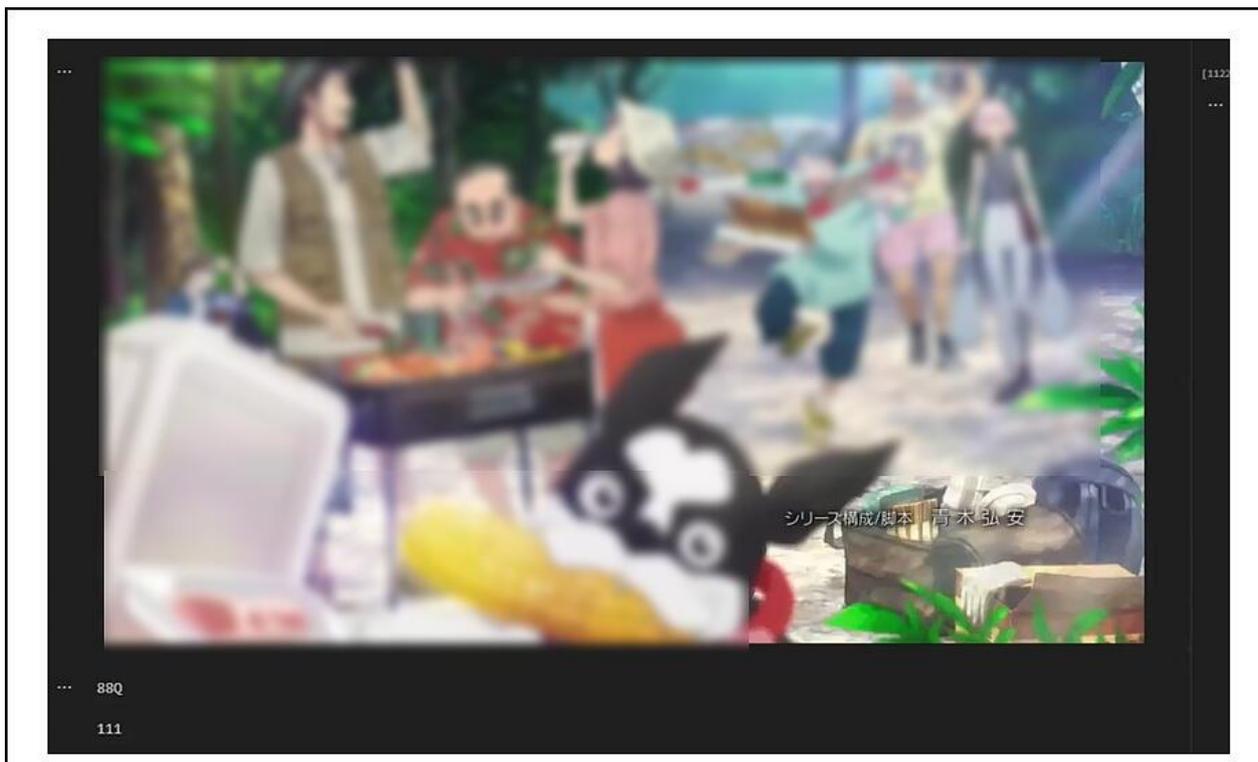


図 3-3 クレジットが表示された画面

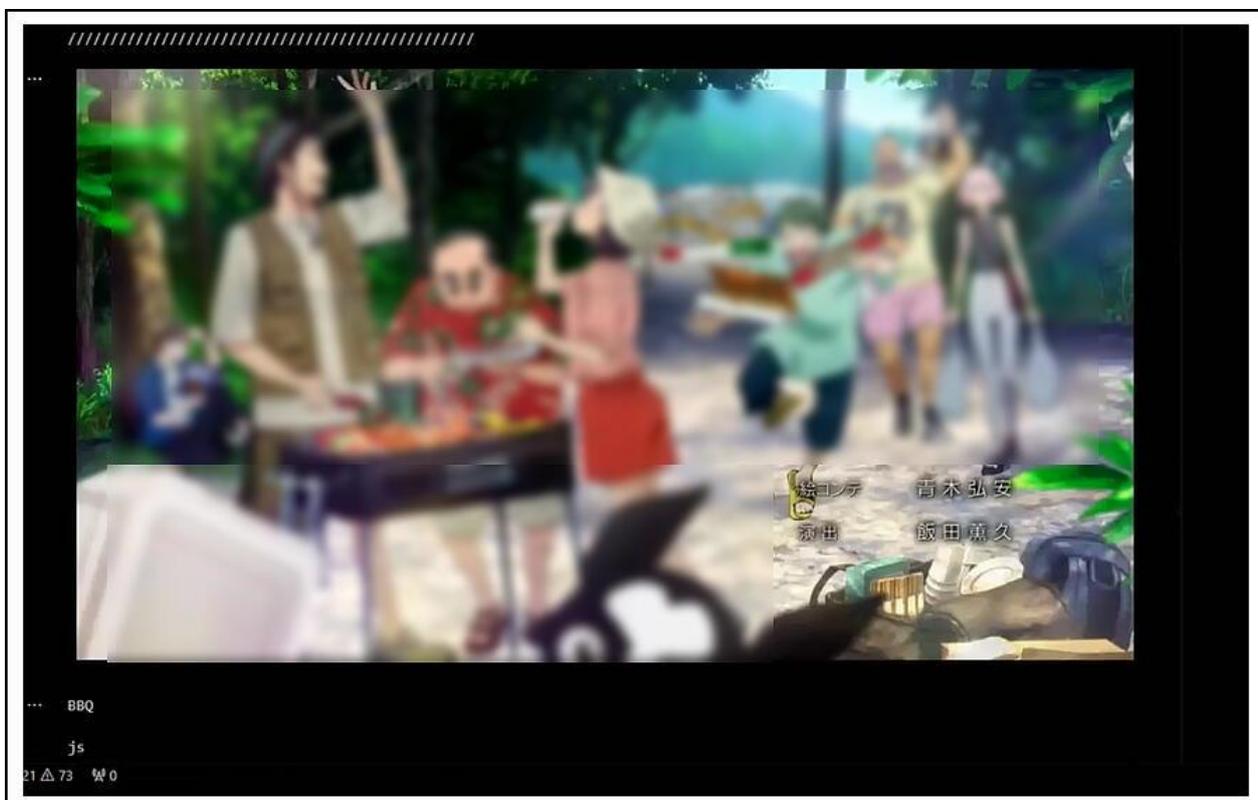


図 3-4 前掲のものと似ているがクレジット名が異なる画面



図 3-5 差分として抽出された文字列

こうしたある種のユニット化とインターフェースの簡素化によって、手順さえ理解していれば誰でも録画データからのテキスト抽出までの作業を行えるようになった。

一連の工程で重要なポイントとなる OCR の機能開発にも取り組んだ。本事業では汎用の OCR サービスを活用している。これ自体でかなりの精度は出せるものの、アニメのクレジットでより高い精度を出すには、アニメ独自の仕組みを理解したデータセットを用いる必要がある。そこで、OCR によってテキスト化されたクレジットデータをデータセットに投入できるようなツールを用意した。

## ツールに入力出来るように変換

- 誤字修正
- 属性を付加



図 3-6 ツールに入力できるように変換する

役職名や人名、団体名といったものを付加しておくことで、画面上に表示されたクレジットを属性によって切り分け、抽出できるようにした。下記の黒い画面の左側は学習を行わなかったもの、右が学習を行ったものとなる。添付の標準のモデルで読ませているが、未学習の左では、キャストなどはうまく読み取れていない。学習させれば右のように対応できるようになる。



図 3-7 未学習と学習済みツールを使った比較サンプル

[[[1129, 317], [1219, 317], [1219, 353], 'キャスト', 0.6313340663909912), ([[1048, 366], [1149, 366], [1149, 410], '音響馬', 0.1988976615169995), ([[1200, 362], [1344, 362], [1344, 410], '絵子', 0.20821562890757422), ([[1048, 413], [1148, 413], [1148, 453], '玉城杏夏歩', 0.369080599044812), ([[1202, 409], [1342, 409], [1342, 453], '坂舞弘', 0.12735896910946135), ([[1049, 457], [1149, 457], [1149, 497], '聖舞理', 0.9871746897697449), ([[1202, 452], [1342, 452], [1342, 500], '青天吉ゆこ士', 0.17946672939310587), ([[1024, 503], [1148, 503], [1148, 543], '宣編', 0.14536030744136095)	[[[1129, 317], [1219, 317], [1219, 353], 'キャスト', 0.9999804903111128), ([[1048, 366], [1149, 366], [1149, 410], '青天国春', 0.9978960751268886), ([[1200, 362], [1344, 362], [1344, 410], '鈴代紗弓', 0.9979516367745398), ([[1048, 413], [1148, 413], [1148, 453], '玉城杏夏', 0.9999151239622234), ([[1202, 409], [1342, 409], [1342, 453], '蟹沢萌子', 0.9983290154828316), ([[1049, 457], [1149, 457], [1149, 497], '聖舞理王', 0.9999932521723386), ([[1202, 452], [1342, 452], [1342, 500], '夏吉うこ', 0.9544081278826443), ([[1024, 503], [1148, 503], [1148, 543], '祇園寺雪音', 0.9999999999999999)
--	--

図 3-8 変換比較の拡大 (左が未学習で右が学習済み)

今回、データセットの投入はアニメ番組の 1 話から 2 話分くらいにとどめているが、それでもある程度まで読み取れた。アニメ作品の場合、話数が変わってもクレジットの段組・表示タイミングなどには余り変化はない。背景のタイミングが変わるなり、特殊回で変わる程度で、そういったものに追従できれば十分だろう。

キャプチャした画面に表示されている文字を平板なテキストとして出力する上で、役職と名前の並びなどを意識して、整形された形で出力できるようにした。作画監督なら作画監督といった単語単位でデータを持っていたり、1 文字ごとといったばらばらな単位での把握から、分解能を一旦ローレベルまで下げてデータを取得した上で組み立て直したりするようにして、そこからレイアウトに従ってテキストが採録できる段階へと引き上げた。結果として、レイアウトを踏まえたテキストを出力するようになったが、行間などについては省略されるため、元の並びのように出力できるように整えた。

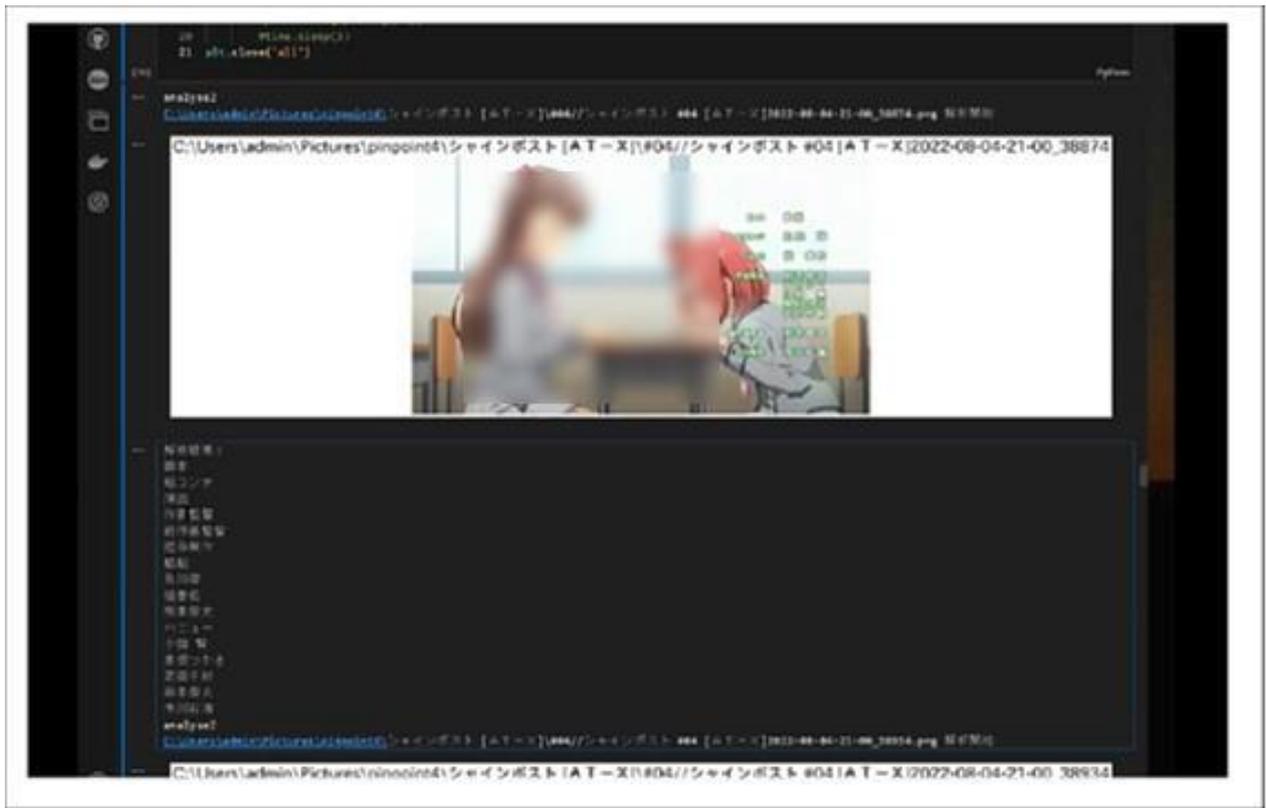


図 3-9 キャプチャ画面の文字情報がベタな段組で表示される例

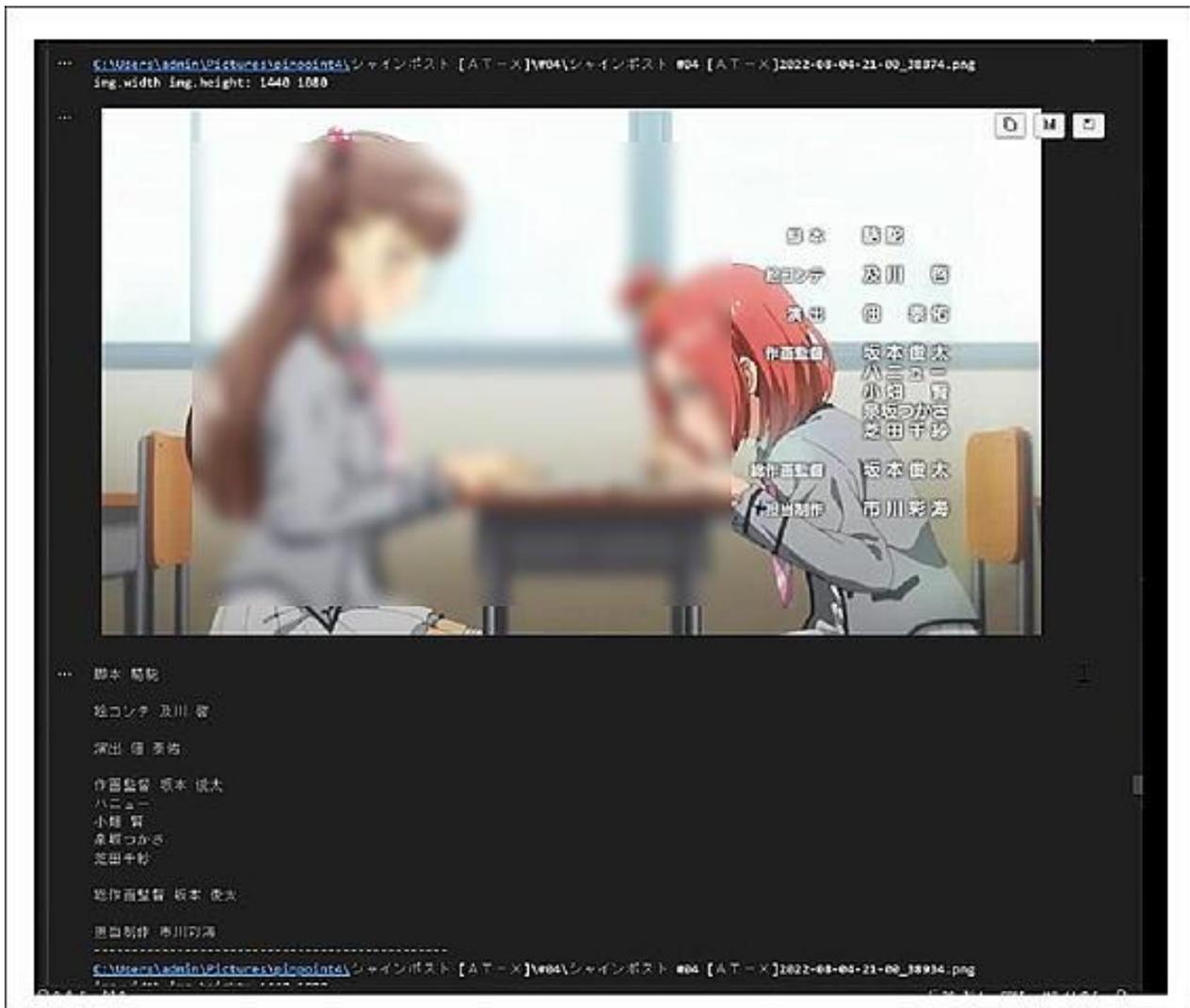


図 3-10 キャプチャ画面の文字情報と同じレイアウトでテキスト化される例

ただ、左右にクレジットのグループができる場合、ブロックごとにテキストをかたまりとして出力するところには至っていない。アニメの場合は二つのブロックや三つのグループでクレジット表示される場合がある。こうした作品ごとの癖を理解し、データとして用意しておけば、グループごとに読めるようになると考えている。

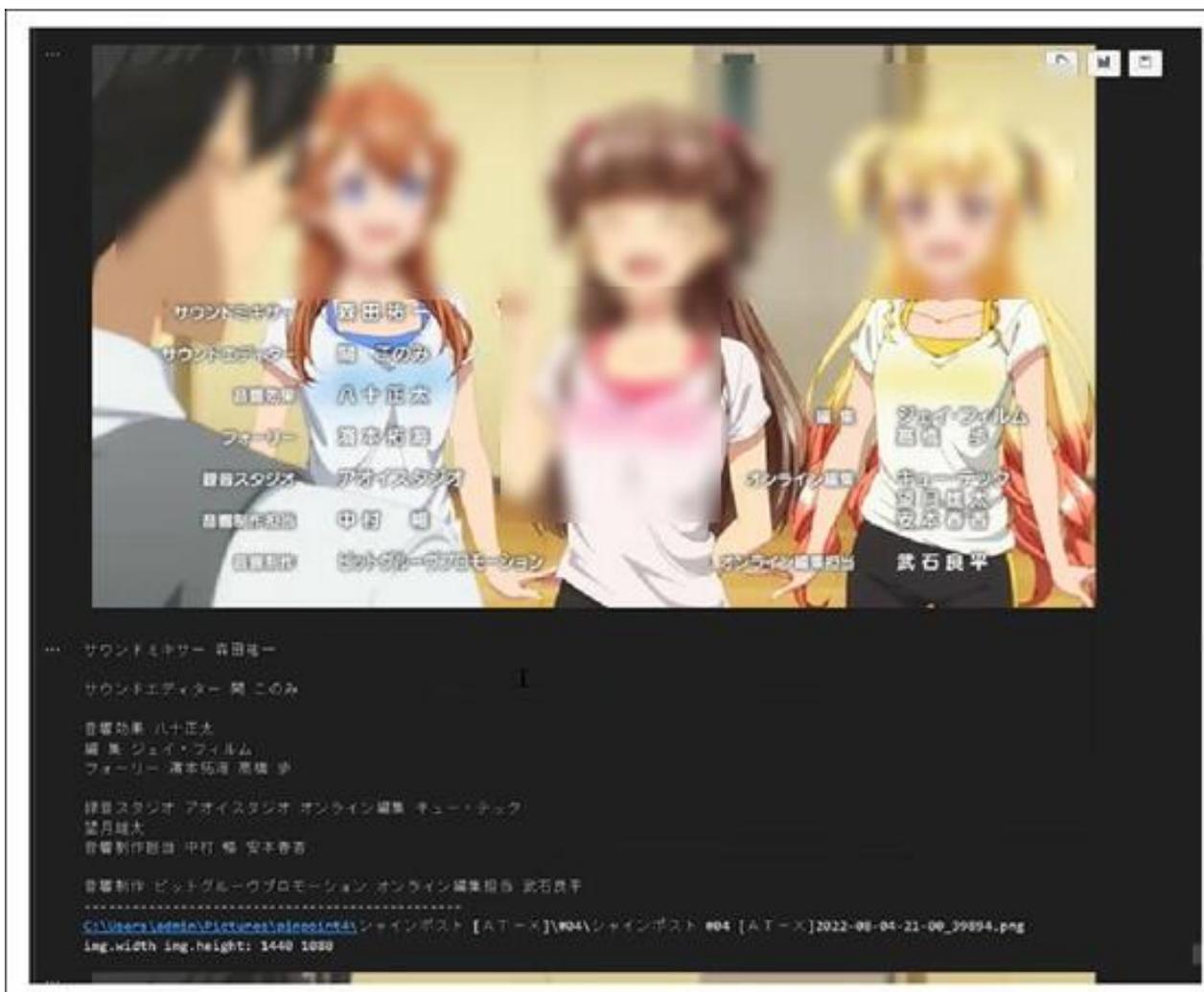


図 3-11 キャプチャ画面内にクレジットのグループが二つある場合のテキスト化例

### 3.2.2 成果・効果・制約事項

全録サーバによって録画された作品を Excel のシート上に列記し、処理が必要なタイトルを選び出して指示を与えキャプチャから OCR への投入、テキスト化などの処理を行い、キャプチャ済みの画像ファイルとともにテキストファイルを作成する一連の作業フローを構築した結果、映像からテキストを抽出するまでの一連の流れを理解しやすくなった。

映像ファイルからキャプチャまでの作業と、キャプチャ語の画像ファイルの OCR 投入、テキストファイルとして抽出された人名や役職名といったクレジット情報の採取などがそれぞれに個別の作業となるため、これらを接続した録画さえ行われれば、後はテキスト抽出までを一つの流れの中で行えるようになる。より使いやすいクレジットデータの採録システム構築まで、大きく前進したと言える。

録画する対象をリスト化して記載するだけで、一連の作業が完了まで向かうような仕組みが構築できれば、一般の人でも自宅等で全録サーバの稼働、キャプチャの実行、OCR によるテキスト化と

いった作業を手軽に行えるようになり、現状では首都圏しかカバーできていないクレジットデータの抽出を、地方でも推進してより詳細なクレジットデータの採録につなげられる。ある種のパッケージ化が可能なら、今後取り組むべき課題と言えるだろう。

ただし、テキストファイルに記載された文字列から、これは役職名である、これは人名であるといった属性を判断してメタデータとして抽出するには、アニメ関連にチューニングされた辞書データがやはり欠かせない。現状ではテキストデータを目視で判断し、作業者の知見で項目ごとに振り分ける必要がある。こうした作業の簡素化を目指してメタデータ処理の確立に取り組んだが、「リスト DB」等を活用して構築される辞書データを一般にも供する状況には至っておらず、著作権等を明らかにした上で利用可能な範囲を決めるステップが依然として必要であると改めて確認された。

ここで、進化が著しい汎用 AI の活用により、作成されたテキストデータを読み込ませるだけで自動的にメタデータ化されるような方法も検討の俎上に上がってきた。ただし、汎用の AI が教師データとして吸収してきたものが、リスト制作委員会による目視確認のように正確性を担保される情報かという点で疑問符が付く。汎用 AI の処理機能を活用しつつ、独自に正確性の高い辞書データを用意して組み込み、高精度で高品質なメタデータの生成が行える手法を探るべきだろう。

### 3.3 再配布可能なシステム

#### 3.3.1 機能要件・動作概要

全録サーバに記録された映像ファイルをキャプチャして画像ファイル化し、OCR への投入を経てクレジットデータを抽出し、テキスト化する作業を行う上で必要なソフトウェア群をモジュール化するプラットフォームとしては、引き続き「Google Colaboratory」を使用している。OCR 部分については、Google の Google Cloud Vision API を活用している。

録画した作品のファイル名を Excel 上に列記し、これを選択して当該のファイルに入っている映像データをキャプチャする作業、取り出された画像ファイルを OCR に投入してテキストデータを抽出する作業などをある程度ユニット化しており、慣れれば希望する作品のエピソード単体でも、1クール分 13 本まとめてでも適宜処理に移れる。シングルスレッドで処理しても、1クール分で 1 時間かからない程度であり、マルチスレッド化すれば更に処理時間は短縮できる。

ただし、Excel にファイル名をリストアップする作業や、キャプチャ作業を行うための映像ファイルの指定、OCR への投入のための「定義情報の記述」や、処理後のテキストデータからのスタッフデータとしての「分類や同定作業を含めた採録」は個別に行う必要がある。ここを自動的に行う仕組みが構築できれば、トータルの作業は更に省力化されるだろう。一方で、採録されたテキストデータからのメタデータ処理は、完全に自動で属性を付与して仕分を行う手法は確立されていないため、目視によってテキストを転記しながら属性に応じて振り分けていく必要がある。

将来的にアニメ関係に強い辞書データが構築されて利用できるようになるか、汎用 AI が発達して

アニメ関係に関する辞書を搭載し、クレジットデータのメタデータ処理にも利用できるようなれば、こうした人力による処理時間を軽減できる。正確性を期すために最終的な確認は人が行う必要があるものの、それ以前の段階で高精度なクレジット採録が行われていれば、検証にも多くの時間はかからない。

一連の作業について、どのタイミングで行うかも検討を要する。エピソードが放送されて映像ファイルができた段階で、逐次 OCR 化からテキストデータ抽出まで実行してしまう方法、1クールなり2クールで終了した作品を、全話まとめて OCR 処理する方法などが考えられる。メディア芸術データベースへの即応を考慮するなら、毎日の放送直後にテキスト化すべきなのかもしれないが、数話分であっても有効な学習データを得られるようになっている状況で、1クール分あればより高性能の学習データを構築できるのなら、シーズンが終了してから1作品についてまとめて処理する方が、効率的との考えもある。

### 3.3.2 成果・効果・制約事項

より簡便に OCR を利用する方法として、Google のアプリケーションとなる Google Drive の機能に付いている OCR の使用を試行した。映像データからスライスしてキャプチャした画像ファイルをこの OCR に投入すれば、単純なテキスト抽出は行える。ただし、段組で表示されるクレジットをそのままのレイアウトで把握し、関係性を維持したままテキスト化するにはできておらず、一つのファイルを選択して作業を指示する必要がある。したがって作業量の軽減につながらないため、この手法は本事業での使用対象から除外した。

ただし、シンプルでかつ容易にテキスト化を行える仕組みとしては活用可能。複数の画像ファイルを PDF にまとめてしまうと、作業手順を大きく簡略化できるのも魅力的だ。Google Drive の制約で 2MB 以上のファイルは投入できないが、個人が簡便にクレジットデータの OCR によるテキスト化を試すだけなら問題なく利用できそうだ。

本事業でも利用している Google Colaboratory を使用したシステムについては、バージョンアップを行い、画面に表示されているものと同じような形に、文字を整形して出力できるようにした。多少ノイズが混じる場合もあるが、ほぼ表示のままに出力できる状態が可能である。スペースについても、文字列が並んでいる場合では、列間の間の幅もそのまま再現して、レイアウト的に人が読める形に出力されるようになった。

こうしたシステムでも、汎用のものである以上、外部で独自にクレジットデータの採録を行っている団体や個人でも、アニメのクレジットに関する知見を持ち、コンピュータに関する知識を有していれば、構築し利用できる。利用には Google アカウントを取得が必要となり、作業に当たっても OCR の費用がかかるが、1回 API キーを取得すれば使用自体は難しくない。このシステムが普及して誰でもクレジットデータ採録して登録できるようになれば、本事業の目指すクレジットデータの

網羅にも大きな進展がある。

OCR の精度についても、現在放送されている番組なら 8 割くらいの番組はテキスト化できる。やはり課題として Google のアカウントでキーを取得する際に個人では手間がかかる点が挙げられるが、ここもいずれ取得しやすくなると見られる。現時点でシステムの要件を整理しておく、新しい採録者の参入を期待できる。

### 3.4 アナログ素材のデジタル化

#### 3.4.1 機能要件・動作概要

リスト制作委員会を主宰する原口正宏氏宅に保管されているアナログ映像素材のデジタル化に当たって、作業を行う北海道アーカイブセンターでは、再生用のビデオデッキ、アナログデータとして出力される映像データをデジタルデータに変換するハードウェアエンコーダ、デジタル化されたデータを記録する HDD（ハードディスクドライブ）など機材を準備した。北海道アーカイブセンターでは従前より同種の作業を行っており、ビデオデッキやエンコーダなどの機材は既存のものを使用した。

北海道アーカイブセンターは、ビデオデッキを従前より 20 台ほど設置しており、アナログ素材のデジタル化では複数台のビデオデッキでビデオテープをまとめて再生しながら、モニター上で再生される映像をチェックし、不具合があれば補正等の処理を行っている。本事業においては、原口氏宅から預かった 30 本ほどのテープ類の確認にとどまったため、複数台のデッキを長時間、同時に使用して再生するような作業は行っていないが、今後本格的なデジタル化作業を想定した場合には、ビデオデッキを 10 台ほど使用し、同時再生しながらデジタル化を行っていく運びとなるだろう。

通常モードで 2 時間分の録画が可能なビデオテープに、3 倍の分量を録画できるモードが使用されていた場合は、再生だけでテープ 1 本につき最低でも 6 時間が必要となる。前後してどのような番組が録画されているか、トラッキングのずれやヘッドの目詰まりといった異常が発生していないかを監視する必要があるため、1 日の作業時間は 6 時間から 8 時間でテープの本数で 10 本ほどとなることが推計される。



図 3-12 複数のビデオテープを同時再生しながらモニターで監視する

さらに、実際に発生したケースで録画に使用したビデオデッキに固有の問題からトラッキングにずれが生じていて、北海道アーカイブセンターが所有するデッキではすぐには再生できなかった。このため、ヘッドの角度を物理的に調整するなり新たにデッキを入手して録画時の条件に近づける作業を行って、映像だけでも再生可能な状態へと改善した。こうしたアクシデントの解消に作業にかかる時間も、全体の流れの中で考慮する必要があると判明した。

通常のビデオテープについては、ビデオデッキで再生しつつハードウェアエンコーダを介してデジタル化を行った。試行では補正等は行わず、出力された状態で HDD に記録していったが、ゴーストと呼ばれる文字が何重にもなって見えるような現象が生じていれば、クレジット部分をより見えやすくする補正を行った上で記録できる。

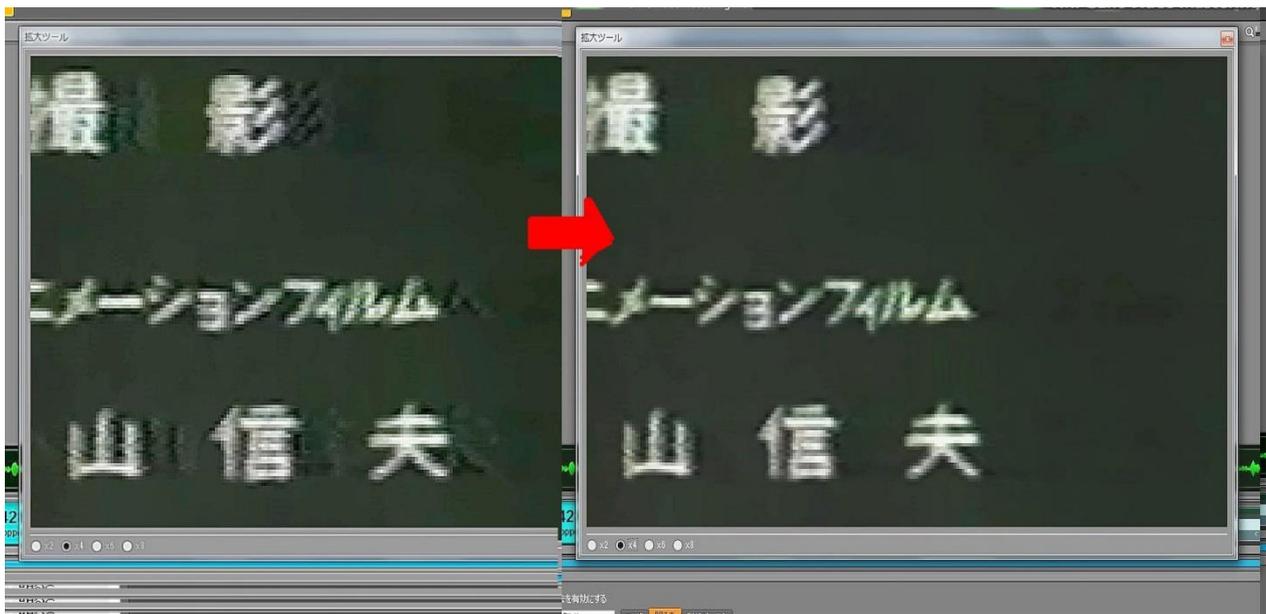


図 3-13 ゴーストが発生している文字（左）を補正して読み取りやすくする

こうして出力されたデジタル映像を OCR によるテキスト化の工程に回して、全録サーバに録画されたデジタルデータからの OCR 処理を経てのクレジットデータ採録と同様、テキストデータによるクレジットデータの採録を行えた。



図 3-14 デジタイズされた画像を OCR 処理してテキストを抽出

デジタイズに当たっては、動画ファイルに一般的に使われている MPEG (Moving Picture Experts Group) のフォーマットではなく DV (ビデオ規格) を使用した。理由は、一旦低画質となったものを改めて高品質化することは不可能であり、データの圧縮率が高い MPEG ではクレジットデータを OCR で読み取りづらい場合があるためである。

### 3.4.2 成果・効果・制約事項

原口氏宅に保管されているビデオテープ等のアナログ素材については、おおむねデジタル化から OCR 処理を経てクレジットデータを採録できると判明した。複数台のビデオデッキで同時にビデオテープを再生しながら、ハードウェアエンコーダを介してデジタル化を行い記録していくフローを確立しつつ、トラッキング異常やテープ及びシールの破損といったケースに個別に対応するラインを用意し、必要ならそちらを使うようにしていけば、アナログ素材のデジタル化を進めていけるだろう。

ただし、本事業で手がけたデジタル化は飽くまでも膨大な量保管されたビデオテープの一部であって、今後どれだけの確率で個別対応が必要な素材は出てくるかは判然としていない。データの記録に使われている磁性体の劣化なりテープのベースとなっているポリエステル素材の劣化が進んで、低温乾燥などの処理をして再生可能な状況とするケースも想定される。こうした想定すら超えたトラブルが起こる場合も考えられ、その時々で対応する必要があると言える。

原口氏が所有するアナログ素材には、原口氏が自ら録画して保管したもの以外に、原口氏以外の人物が同じように自宅で録画したビデオテープ類も含まれており、録画に使用されたビデオデッキなどの機材や保管環境にバラつきがある。比較的良好な状態で保管されていた原口氏による録画テープと違い、劣化や破損が進んでいるものがある事態も考えられるため、これについても改めてサンプルとして作業を行い、状態を確認する必要があるだろう。

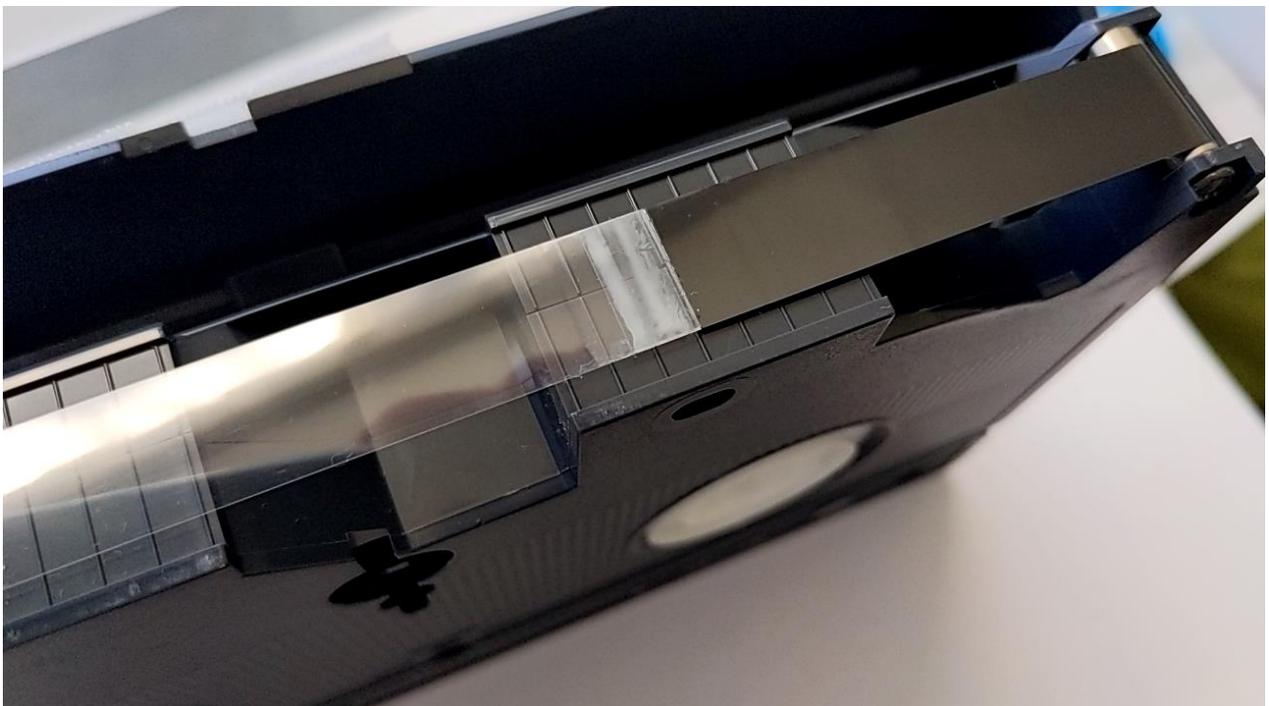


図 3-15 リーダーテープと磁気テープ部分が切れそうになったら補修してつなげる

再生に使用するビデオデッキについても、HDD や DVD、ブルーレイへの録画が進んだ 2000 年台移

行には新しい製品が登場しなくなっている。現行の作業で使用しているビデオデッキが劣化又は破損などの状況に陥り、部品の欠如で修理も難しくなっていく状況も将来的には考えられる。アナログ素材のデジタイズがアーカイブにおいて不可欠の作業と言えるものである以上、そうした作業を支える機材の調達・確保についても考慮する必要があると確認された。



図 3-16 北海道アーカイブセンターでストックしているビデオデッキ

これとは別に、ヘッドのずれなどがもたらす再生面での不具合を、ヘッド角度の調整などによって解消できないケースも考えられる。絶対に不可能なのか、データを映像として読み取る技術が別にあるのかを探りつつ、まずは可能な範囲からデジタイズを進め、クレジットデータ採録の元となるデジタル素材の蓄積に努めるべきだろう。

また、録画時に作品名をインデックス等に記録したのものもあるが、それらは飽くまで便宜的に付与されたものであり、作業用に付与された統一表記ではなく正式のタイトルを調べる必要がある。加えて、録画時期によっては番組全体を録画したもの、一部だけを録画したものなど状態がばらばらで、アニメ作品以外のドラマやバラエティ、歌番組なども録画されたビデオテープもあるため、常に監視してアニメ作品が何かを確認する必要性が生じた。

デジタイズ時における記録にも膨大な労力が必要と判明した。アナログ素材の種類を確認し、ビデオテープなら VHS であるかベータであるか、どのメーカーのものかを確認。テープに貼られているラベルに記載された録画内容とともに記録した上で、ビデオデッキで再生して実際に録画されている番組名を列記しつつ、いつ頃放送されたものかを確認するほか、本放送時のものか再放送時のものかを確認などしていく。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	列1	テープラベル表題	メディア 種	録画モ ード	記録メ ディア製 造年代	記録メ ディア メーカー	録画年代	実際の録画内容	実際の録 画映像本 数	問題点	物理損傷 対応	ノイズ程 度	必要な対 策	デジタイ ズ可否	再生機種	OCR可否	OCRの為 必要な後 処理
1	1	ドラえもん、ケン ちゃんの冒険、のび 太の宇宙開拓史	家庭用 VHS	3倍	1979	ビクター	1981	【映画ドラえもん、のび太の宇宙開拓史 (EX1981)テレビ放映】【ドラえもん、ケ ンちゃんの冒険(EX1981-07-01)国際障害 者年特番】(題は生きています(一部 分)EX1981)単発】【日曜洋画劇場、2001年宇 宙の旅(EX1981テレビ放映)】【母をたず ねて三千里#52(CX1981再)】【アルプスの 少女ハイジ#1(CX1981再)】【アクション大 魔王#1(TX1981再)】	6		有→テー プ冒頭の 接着劣化 により、 テープが カセット から外れ る	やや多め	【トラッキ ング不良= 経度=3流 モード再生 専用ヘッド 切替機構が デッキで対 応】テープ 冒頭外れ= ビデオテー プ修理キッ トでスライ ス補修】	済	松下NW- SB707	可	
2	2	ルパン三世 カリオ ストロの城	家庭用 VHS	3倍	1980	マクセル	1982	【白雪姫～トロンまで ディズニー映画 名場面集(詳細不明18分)】【特撮メカザブ ンガ#32(NB1982/09/18)】【氷壁ロード ショー (NTV1982/09/22)ルパン三世 カリ オストロの城-宮崎駿自身が放送枠に合 せ約7分をカット編集した放映初期 Ver】【ふしぎなメルモOPのみ(TX1982 再)】【特撮メカザブンガ #33(NB1982/09/25)】【ロボット8ちゃん OPEDのみ(CX1982/09/28)】【松田聖子の野 菊の墓(EX1982)消し残りの一部】	7		無	やや有		可	松下NV- SB707	可	
3	3	【超人ロック ロード ・レオン(1)】【風魔 の小次郎(5)8】【機動 戦士SDガンダムの逆 襲】【新・キャプテン 翼(4)】	家庭用 VHS	標準ダビ ング	1988	富士フイ ルム	1989	【超人ロック ロード・レオン (1)OVA1889/10】【風魔の小次郎 (5)OVA1888】【機動戦士SDガンダムの逆 襲】(OVA1889/07)】【新・キャプテン翼 (4)OVA1889/10】	4	制作会社等 から借りた 家庭用テー プをさらに 家庭用に分 ピンクした テープ。家 庭用同士の ダビングの ため劣化が 気になる。	無	やや有		可	松下NV- SB707	可	
4																	
5																	
6																	

図 3-17 アナログ素材デジタイズ時の記録項目と記録例

原口氏以外が録画して寄贈したテープ類は、更に個々人のセオリーによって録画や番組名の記載が行われているものもあると推定される。それらを一つ一つ確認して、正式な作品名共々記録していく作業には、アニメ作品に関するある程度の知見が必要となる。膨大かつ正確なアニメアーカイブの構築に当たって、こうしたアニメ作品に関する知見を持った人材の育成が求められていると言える。



図 3-18 原口正宏氏宅のビデオテープに記載された番組名

### 3.5 周知活動

#### 3.5.1 メカデミア国際学術会議（令和 5 年（2023 年）5 月 27～29 日）

京都精華大学及び同学が運営する京都国際マンガミュージアムで令和 4 年 5 月 27 日から 29 日まで、国際学術会議「メカデミア（MECHADEMIA）」が開催された。「メカデミア」とは漫画・アニメ・ゲームなどアジアのポピュラーカルチャー全般を扱う国際学術ジャーナルで、日本や韓国、アメリカなどで会議を開催してきた。2023 年はメカデミア・京都国際マンガミュージアム・京都精華

大学国際マンガ研究センターの共催で開催。「アフターマス (Aftermath)」をテーマに、それぞれの専門家による研究発表と基調講演が行われた。

本事業ではこの「メカデミア」に、オブザーバーを務める JAniCA 事務局長の大坪英之が参加。5月28日に講演を行い、テレビアニメや劇場アニメ、OVA といった媒体で発表される日本のアニメの多彩さを海外の研究者に知ってもらった。また、アニメ作品に関連した情報が集約されているメディア芸術データベース (以下、MADB) が 2001 年の文化芸術振興基本法の公布を経て事業として構築されてきた経緯を紹介し、日本にこうしたアニメスタッフに関する DB が存在していることを伝えた。

The image shows a screenshot of the Media Arts Database (MADB) website. The page title is "Accumulation and Development of Information on Animation Works in Japan" and "History of the Media Arts Database". The website interface includes a search bar and navigation links. To the right of the screenshot is a timeline of key events:

- December 7, 2001. "Basic Act for the Promotion of Culture and the Arts" was promulgated.
- From 2010 to 2014 "Project to Improve Networking and Archival Infrastructure of Media Arts" <Mori Building>
- From 2015 to 2016 "Project for the Promotion of Cooperation of Media Arts" <Kyoto Seika>
- From 2017 to the present "Project for the Promotion of Cooperation of Media Arts" <DNP>

The URL <https://mediaarts-db.bunka.go.jp/> is displayed at the bottom left of the screenshot.

図 3-19 MADB と連携事業の歴史

この MADB にアニメ作品のクレジット・キャスト情報を採録する上で、どのような取組が行われているかも説明した。アニメ作品に付随するデータには、役職名やスタッフ名、役名、キャスト名など様々な種類のものであって、採録する際のデータも多岐に及ぶ実態を話して、詳細な DB 構築にかかる作業量の多さを知ってもらった。参加者との間で質疑応答を開催。海外からの参加者が多く訪れ 30 名ほどの部屋が一杯になり、個々の作品への関心も含め、文化・産業として重要な日本のアニメに関する情報へのニーズが、海外でも高いことが伺えた。



図 3-20 メカデミア国際学術会議での質疑応答

### 3.5.2 「マチ★アソビ」 Vol.27 (令和 5 年 (2023 年) 10 月 28 日)

「マチ★アソビ」は、アニメやゲーム、漫画などエンターテインメントに関連する事業者が、徳島県徳島市に集まって上映会や講演などを開催し、アニメファンに作品への関心を持ってもらうためのイベントで、2009 年から年に 2 回のペースで開催されている。受け手となるアニメファンだけでなく、送り手となるアニメ制作会社や製作委員会会社、アニメ作品関連情報を伝える出版社やニュースサイト関係者も多く来場しており、ここでイベントを行うことで情報が広範囲に伝わるとともに、関係者とのつながりを構築できる場として、アニメ業界関係者も認識している。

本事業では、2023 年 10 月 7 日から 29 日にかけて、順次イベントが開催された「マチ★アソビ」Vol.27 に大坪英之が参加。クライマックスランとして大半のイベントが開催となる 28 日と 29 日の 2 日間のうち、10 月 28 日に徳島市内の映画館「ufotable CINEMA」を会場として開催されたトークイベント「あなたの知らないアニメデータベースの世界 アニメーターから見る『クレジット』とは」に、JAniCA 理事でアニメーターの徳野悠我とともに登壇した。

イベントでは、クレジットデータが記録されていることでクリエイターが自身の業務歴を客観的に確認できて、新しい仕事先にリストとして提出してどのような作品に関わってきたかを説明できるメリットを紹介し、そのためにスタッフリストがデータベースされる意義が話された。一方で、JAniCA のようなアニメ関連団体でも、会員の業務に関する全ての履歴を把握して発信している訳ではなく、バイオグラフィーの作成において十分な資料とはなり得ないこと、一方で「作画@wiki」のように有志がテレビ放送された番組などからスタッフリストを小まめに採録し、発信しているサイトには多くのデータが掲載されているが、こちらも把握可能な作品数に限りがある状況などが示さ

れ、完全に近い履歴を整理するためには、より幅広い収集の必要性が浮かび上がった。

リスト化された名前についても、「斉藤」と「齋藤」「齋藤」のように異体字が混在している場合は、検索しても全てがヒットする訳ではなく、完全な履歴を把握しづらい点や、テレビ放送時、あるいは劇場での上映時に表記されるクレジットに誤記や不備があって、パッケージ化の際に修正される事態が起こるケースも説明。どの段階のクレジット表記を正しいものと判断するかを検討する必要性も話された。

アニメーターがどこまでクレジットデータに関心を持ち、自分たちの業歴が記録として残ることに意味を感じているかについて、アニメーター本人から聞く機会が余りなく、アニメファンがアニメスタッフデータベースの存在や意義について知る機会も少ないため、作り手側と受け手側が集まり交流しながら話を聞く「マチ★アソビ」でのトークイベントは、スタッフデータ採録の必要性や課題などを知ってもらう上で意義があるものとなった。

### 3.5.3 「つづきみ」(令和5年(2023年)9月度より随時)

本事業では、クレジットデータの網羅的な収集を目的に、「僕たちは新作アニメのプロモーション映像を3時間かけて一気観したらどのくらいつづきをみたくなるのだろうか？」(以下、「つづきみ」というイベントとの連携を開始した。



図 3-21 「つづきみ」のサイトにあるイベント紹介画像

「つづきみ」は、1月、4月、7月、10月といったアニメ作品の新番組が始まる月の直前に、次のクール(3か月間)に放送されるアニメのプロモーション映像をまとめて紹介し、アニメファンに新作への興味を持ってもらうイベントとして開催されている。

「つづきみ」は、新作のアニメが1話も見られずに終了してしまうような事態を避けるため、ど

のような作品なのかをまず知ってもらう目的で、アニメ好きの有志が集まった「ゼロ話切り撲滅委員会」が中心となって2016年にスタートした。アニメに詳しいことで知られるニッポン放送の吉田尚記アナウンサーが第1回から司会を務め、新作アニメの放送時期に関する情報やプロモーション映像の提供を行っている企業に依頼して素材を提供してもらい、配信及び会場を借りてのイベント開催によって情報を発信してきた。

「つづきみ」は、2023年12月27日の開催が第30回となるまで続いてきた。こうした活動の積み重ねの中で、新作アニメの情報を持つ企業との連携も深まってきたが、作品数が増加している状況で、全てのアニメ作品を把握することが難しくなっていた。JAniCAではこの「つづきみ」に応援団として参画。全録サーバを稼働させている関係で、新作アニメの詳細な情報を把握している状況を「つづきみ」側に伝えて情報の網羅に協力することにした。

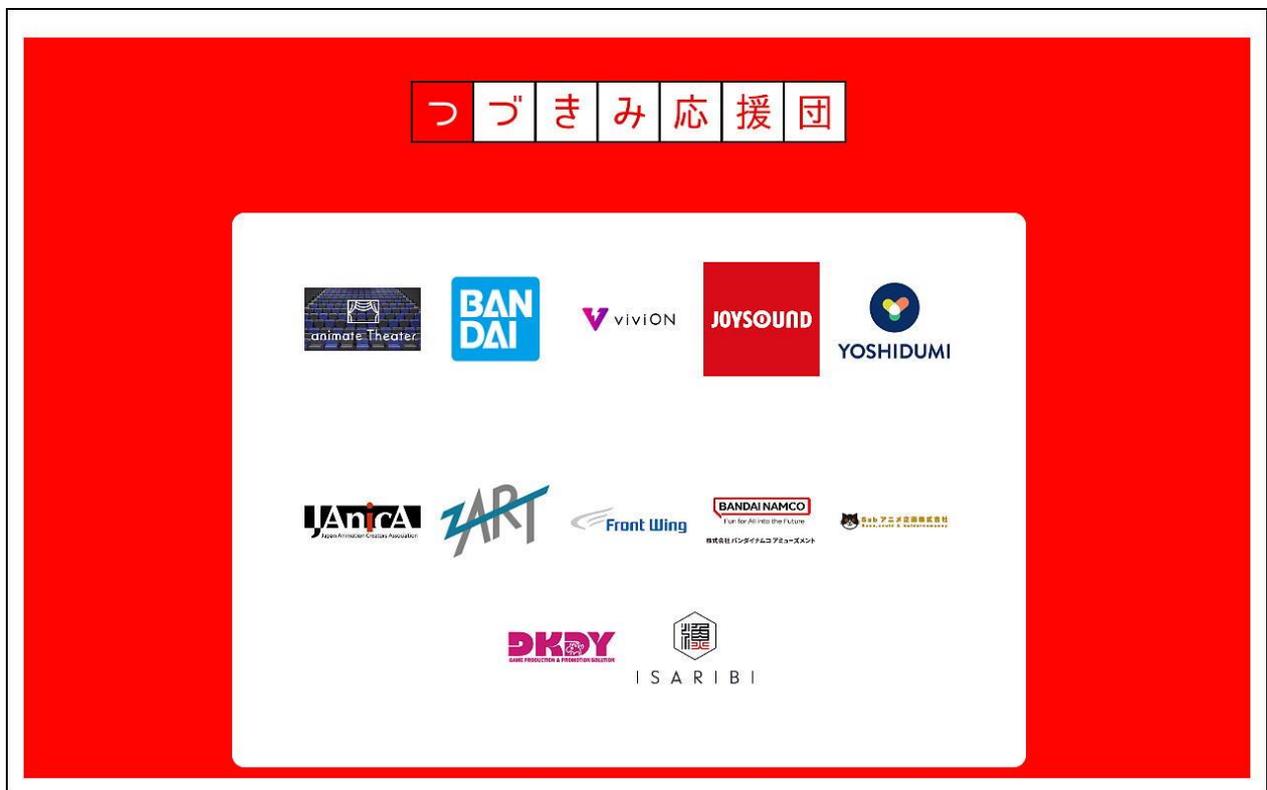


図 3-22 JAniCA を含む「つづきみ」応援団一覧

JAniCA側では「つづきみ」から、イベントで紹介するためにつながりを持った、新作アニメの宣材や情報の窓口について紹介を受け、組成に向けて動いている「アニメ作品情報連絡調整会」への参加を呼びかけて、スタッフ情報のより充実した収集に役立てることを目指した。協力関係の構築は、それぞれが持つリソースやノウハウを持ち寄れば、共に目的としているアニメ産業の振興でありアニメ文化のより一層の浸透に資するものとなる。

### 3.5.4 コミケットフリンジ研究会（令和5年（2023年）12月30日）

世界最大級の同人誌即売会「コミックマーケット」の開催に合わせ、漫画やアニメ、ゲームといったポップカルチャーの研究をしている人が集まり、意見交換を行う場として「コミケットフリンジ研究会」が、東京工業大学の出口弘名誉教授や中村仁准教授によって運用されてきた。今年度も「コミックマーケット 103」開催期間中の12月30日に実施され、大坪英之が参加して「アニメ分野のデータセットの研究」という演題で講演を行った。

前提として、アニメ分野の量的研究を行う場合に、余りにも多くのアニメ作品が放送されている、あるいは上映、発売されている状況では、時々刻々と生成されていく作品情報に把握が追いつかず、いつまでも全量が分からない問題があった。こうした状況を解決する手段として、既に、多くのアニメ作品情報が収録されたMADBを紹介。そこに現状、どれだけのデータがどのような形で存在しているかを話し、研究に役立つと指摘した。

MADBのデータには、テレビ作品については原口正宏氏らリスト制作委員会が採録して作成した「リストDB」や「年間パーフェクトデータ」、本事業が全録サーバの利用によって採録したアニメスタッフリストなどがある現状を説明。一方で、劇場作品については2017年以降のデータが未採録になっている点、近年増加が著しい配信作品については調査すら行われていない実態も紹介し、全量の把握に向けてこうした分野の採録方法を検討する必要性を訴えた。

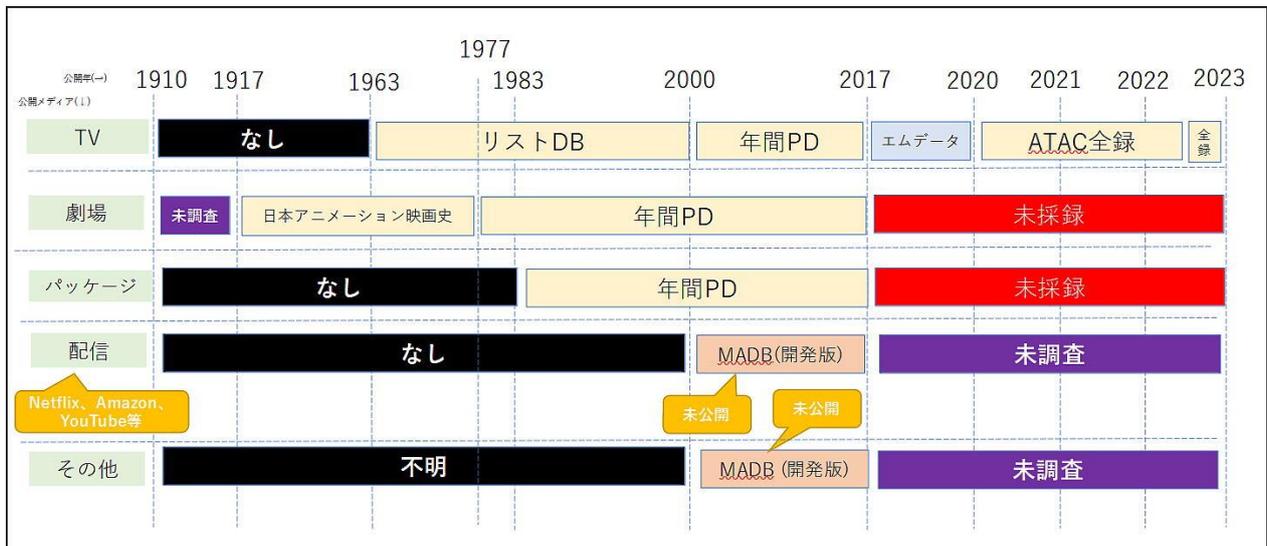


図 3-23 MADB へのアニメ作品の採録状況

アニメ分野における調査研究をこれからも続けるに当たっては、対象をより広範なものにするなり、より深く調査する場合、基盤となるMADBを始めとしたデータセットの充実が求められる。本事業におけるクレジットデータ採録の効率化はこれに資するものとなるが、現状では網羅できていない分野についても採録してデータ化し、MADBに投入できるような体制の整備にも大変な意義があると改めて示して、活動への理解を促した。

講演後には、参加者との質疑応答を行った。1989年から開催されてきた「CG アニメコンテスト」からは、新海誠監督や吉浦康裕監督ら、現在第一線で活動しているアニメ監督が大勢排出されており、アニメの歴史を語る上で重要なコンテストと位置づけられるが、一方で初期の記録が残っていない点が指摘された。これに関して大坪が、後に活躍するようになったクリエイターでも、その頃の活動については経歴的な空欄となってしまう可能性を話して、可能な限りデータを残してリスト化し、DB化していく必要を示した。

また、学術研究者からMADBの活用について、所属する学生は意欲的だが、ほかには余りMADBの活用例を見ず、どのように使えば良いか分からないため、何か事例はないかといった質問が寄せられた。これに大坪が答えて、実際にMADBが広く知られてない状況を挙げつつ、過去に大坪がアニメスタッフデータを元に、アニメ作品内に登場するモニター画面をCGなどで制作する仕事がどのように発生してきたかを調べ、アニメ研究者に提供した事例を紹介。明確な目的さえ与えられれば活用できるとして利用を働きかけた。

### 3.6 採録状況

本事業では、全録サーバを設置し、アニメに関連した番組を全て録画してクレジットデータを抽出して、スタッフリストを採録・研究している原口正宏氏が主宰するリスト制作委員会と情報共有したりする業務を行っている。全録サーバは杉並と千代田に設置して、それぞれに番組の録画を行っている。

アニメ全録 受信状況								
No	放送局	区分	原口(1階)	原口(2階)	杉並	千代田	大坪	須山
1	地上波局(関東広域圏)	無料	○	○	○	○	○	○
2	BSデジタル局	無料	○	○	○	○	○	○
3	TOKYO MX	無料	○	○	○	○	○	○
4	tvk	無料	○	○	×	×	○	×
5	チバテレ	無料	×	×	×	○	○	×
6	テレビ埼玉	無料	×	×	×	×	×	×
7	AT-X	有料	×	○	×	×	×	○
8	アニマックス	有料	?	?	×	×	○	?
9	WOWOW(3局)	有料	?	?	×	×	○	?

図 3-24 録画体制

#### 3.6.1 全録サーバ(杉並)実績

杉並サーバについては、地上波、BS、CS をカバーして録画している。アニメ専門チャンネルもAT-X のように最新の作品を放送しているチャンネルを録画している。単発で放送される番組もあるため、個別にチェックしてアニメ作品が見つければそれも録画している。令和5年1月1日から12月31日までの削除後の杉並サーバの録画ファイル数は12,700件余りとなった。

録画は 30 分の番組で 1 番組当たり 2GB から 3GB といったファイルサイズとなる。新番組でも再放送でも、基本的に放送されている番組は全て録画した上で必要なデータを残し、不必要なデータを削除する方針を採っている。1 週間に同じ作品を何度も放送するチャンネルの場合は、一度だけ録画する。再放送については録画していない。こうした録画状況では、HDD がすぐに一杯となってしまう。

この HDD の容量の問題は、地上波の番組だけを録画していたものを、BS にも範囲を広げた関係で、記録容量が一気に増えてしまったために発生した。随時、重複分を確認して不要なデータを削除しているが、人力では追いつかない部分もある。重複を見つけて削る作業だけで 1 日 1 時間から 2 時間が必要となる。

少し溜 [た] めてしまうと削除の作業に半日かかることもあるため、HDD の容量重複分を検出して削るプログラムを作成し、適用を始めた。これによって、1 日に 1 時間かかっていた作業時間を半分くらいに効率化できた。ただ、こうした削除を適宜行っても、録画対象となる番組本数が多いため、HDD の容量が一杯になるまでの時間は短い。容量がぎりぎりになってからでは、録画の連続性を維持するために HDD を追加するタイミングが失われてしまいかねない。先を見越して追加していく必要があることが分かった。

杉並サーバではまた、機材に浸水被害が発生した。録画用サーバが設置してある場所のそばにある通気口から水が入り、チューナーカードに水がかかって錆 [さび] が発生した。窓以外の場所からでも風雨による震災被害が発生するリスクを理解した上で、機材配置を考える必要があると分かった。

杉並サーバでは、使用電力の容量オーバーに伴うブレーカーのシャットダウンによって停電が発生する事態も発生した。こうしたヒューマンエラーや事故、故障などは、注意していても防げない場合があるため、多拠点で異なる方式で録画しておく必要性が、改めて確認された。

### 3.6.2 全録サーバ（千代田）実績

千代田サーバも、1 年で 13,900 件以上を録画した。録画対象は杉並サーバと同じだが、新作としてのアニメが中心で、再放送の録画は少ない。代わりに放送前や放送後、シーズン終了後、シーズン途中で放送される特番を追って録画していった。

千代田サーバでは今年度、受信に必要な B-CAS カードを読み取るリーダーの故障が発生して代替機器への入替えを行った。こうした不調の結果、事業年度では昨年度に当たる 2023 年 1 月中旬から、今年度の 5 月上旬までの録画に損失が発生している状況が判明した。不足分については杉並サーバの録画分から補完したが、本事業ではリスト制作委員会側が録画を逃した番組のクレジットデータを補完する作業も行っており、不足が重なる事態も考え複数の系統で録画しておく必要性が改めて確認された。

千代田サーバでは、データを記録する HDD にも不具合が生じた。A 系統、B 系統の二重体制で録画を行っているが、このうち B 系統の録画機で不具合が見つかったため、一旦録画系統を A 系に統合。その上で B 系統の録画機の復旧を試みて復旧させた。

録画対象となる番組の不足も発生した。一例では、テレビ東京の「ポケモン」関連番組に録画できていなかったものがあった。アニメ作品ではなくトーク番組で、途中にアニメが放送されるため事前に番組の存在を確認できなかった。こうした番組内番組とも言えるアニメ作品も含め、クレジットデータの採録に必要となる作品が、気付かないところで発生している可能性は常にある。全てを事前に把握するのは困難だが、注意深く情報を収集したり、アニメ関係者とのネットワークを広げて事前に情報を教えてもらったりするなどして、録画対象となる番組の把握に可能な限り努める必要があることが分かった。

### 3.7 メディア芸術データベースへのデータ提供

全録サーバによるクレジットデータのテキスト化、及びリスト制作委員会の活動として原口正宏氏が蓄積してきた「リスト DB」のコンバートなどによって、国立アトリエサーチが運用するメディア芸術データベース（以下、MADB）へのデータ登録が進んだ。全録サーバ由来のものが、作品数で 572 件、話数は 11,939 件となった。共通スタッフ数は延べ 0 件、各話スタッフ数は 0 件、各話声優数は 0 件となっている。

「リスト DB」から、MADB へのデータ投入については、作品数は 3,667 件、話数は 23,170 件に及んだ。共通スタッフ数（延べ）は 302,383 件、各話スタッフ数（延べ）は 71,690 件、各話声優数（延べ）は 17,900 件に達した。

今年度は主に、①1 回の放送分で 2 話や 3 話といった複数のエピソードが放送されるテレビアニメ作品、②劇場作品、③ビデオパッケージ作品について、MADB 側へとデータを引き渡した。

①の視聴機会（1 採録粒度）において複数エピソードで構成された作品については、MADB（開発版）では個々の複数エピソード単位での登録が行われており、MADB（開発版）から MADB（β 版）への移行時には 1 採録粒度として複数エピソードの情報を統合して登録が行われた。昨年実施したリスト DB から MADB（β 版）へのデータ反映時には該当作品数は少ないものの該当視聴機会（＝番組）数と該当エピソード数が多いためにデータ移行対象からは除外していた。本年では、MADB（β 版）の改修方針が定まったことと連動してデータを移行した。

②の劇場作品については、移行対象として特段の問題なく移行作業を行えた。

③のビデオパッケージ作品については、(a) 作品とパッケージ商品が 1 対 1 で対応する狭義のパッケージ（「OVA」と呼ばれる形式）と、(b) パッケージ商品の中に既に公開済みの作品と併せて新規収録される映像作品（いわゆる「特典映像」「新規映像」と呼ばれる形式）があり、前者は MADB（β 版）の採録粒度と同一であるが、後者は採録粒度よりも 1 段階細かい粒度の情報のため両者を

区別しながらデータを移行した。

### 3.8 学術研究機関との連携検討

本事業では、JAniCA、ATAC といったアニメや特撮に関連した民間の団体が中心となって、テレビ放送されたアニメ番組のスタッフクレジット採録や、個人によって録画され保管されていたビデオテープからのスタッフクレジット採録などに取り組んできた。こうして採録されたデータは、信頼性において十分に高いものであると確認されており、メディア芸術データベースへの投入も行われ始めている。

ただ、こうしたデータの所在が広く知られているとは言えない。このため昨年度の事業では、ドイツのシュトゥットガルトで開催された国際会議「**Japanese Visual Media Graph**」に出席し、日本におけるアニメーションスタッフデータベースの存在を訴えた。国内向けにもイベントなどでの講演によって、クレジット情報の採録と運用に関する取組の認知度向上に努めてきた。

今年度は更に国内の学術研究機関とも連携を図り、アニメの歴史や産業構造を研究している研究者にアニメスタッフデータベースの活用を働きかけ、利活用の幅を広げようと取り組んだ。具体的には、学術研究機関でアニメ関連の研究を行っている研究者のリストアップ、8月19日に日本アニメーション学会への参加（聴講）し、実際にアニメ関係の研究を行っている新潟大学との連携の模索などを行った。

ここで挙げた新潟大学は、経済科学部の石田美紀教授が主導し、キム・ジュニアン准教授とともにアニメ制作の過程で発生する「アニメ中間素材」の収集と利活用に取り組んできた実績がある。利活用では、2022年秋に企画展「原画から見る1980年代TVアニメ」を開催して、原画がどのような過程を経て映像になるのかを、展示やインスタレーションによって見せ、2023年3月に開催された第1回新潟国際アニメーション映画祭でも、同様の展示を関連イベントとして実施された。



図 3-25 新潟大学によるアニメアーカイブの展示イベント（2023年3月）

本事業では9月20日に石田教授、キム准教授らとウェブで会談を実施。新潟大学が中間素材の利活用を進める中で、中間素材に含まれる原画や修正原画、動画、レイアウト、絵コンテといった個々の素材に関しては、アニメ制作関係者の知見を得て、より精緻に分析して分類・保存する必要があると感じているとの所感を得た。

これに対して、中間素材のアーカイブを手がけている ATAC や、アニメ制作者の集まりである JAniCA が協力すれば、双方が研究や利活用の深度を深められると考えた。一連の状況から、双方が持つ知見を活かす協力関係を築けそうとの感触を得た。今後は中間素材に関する知見の提供で、新潟大学側の研究の一層の促進を目指す。一方で、中間素材を保管・研究する組織として学術研究機関も候補となり得る可能性を探っていく。

現在のアニメ中間素材に関するアーカイブは、アニメ制作会社が独自に行っているケースが多いが、企業活動の中では費用負担が問題となり、残す素材と廃棄する素材の取捨選択が行われる場合がある。学術研究機関なら、商業性とは違う視点から中間素材を収集・保管する場となり、将来に価値をつなげられる。新潟大学を一つの候補として、アニメに関する研究を進めたい学術研究機関との連携を探れば、少しでも多くのアニメ中間素材を残せるようになり、日本が世界に誇るアニメ

の学術的な研究も進捗すると考えられる。

本報告書は、文化庁の委託業務として、大日本印刷株式会社が実施した令和 5 年度「メディア芸術連携基盤等整備推進事業 分野別強化事業」の成果をとりまとめたものであり、第三者による著作物が含まれています。  
転載複製等に関する問い合わせは、文化庁にご連絡ください。