

# 琉球政府文書デジタル・アーカイブズ推進事業における資料画像 —利用目的から考えるファイルフォーマット—

大田 文子<sup>†</sup>

はじめに

- 1 原資料の補撮影
- 2 ウェブ公開用画像の作成
  - 2-1 作業用 PDF の作成
  - 2-2 PDF から TIFF へ
  - 2-3 タイル分割 Jpeg
- 3 閲覧室閲覧用画像 PDF/A の作成

おわりに

附録：撮影及び画像データ作成仕様

はじめに

沖縄県が 2013 年度（平成 25）より 9 年計画で進めている「琉球政府文書デジタル・アーカイブズ推進事業」（以下「本事業」とする。）においてデジタル化した資料画像は、2020 年（令和 2）10 月末現在、101,327 点となった。筆者は、「琉球政府文書デジタル・アーカイブズ推進事業における資料画像の長期保存及び公開システムに関するレポート」で、本事業における、2013 年度（平成 25）から 2015 年度（平成 27）までの事業初期（以下「初期」とする。）と、2016 年度（平成 28）から 2018 年度（平成 30）までの事業中期（以下「中期」とする。）のデジタル化の撮影仕様と画像データ作成仕様を紹介し、初期の撮影仕様には、A3 サイズ対応のスキヤナで 2 分割を超える大判サイズの文書や、沖縄県公文書館管理規則第 4 条により利用が制限される資料の一部分の撮影が省略されていることを記した<sup>1</sup>。また、中期のはじめ数か月間に撮影及び画像データの作成が行われた一部の資料には、大判サイズの文書が省略されていた資料画像がある。事業年度ごとの詳しい撮影仕様及び画像データ作成仕様は附録を参照されたい。

資料の一部分の撮影が省略されて作成された資料画像は、記録としては不完全なものである。本事業の資料画像データの長期保存用媒体の作成およびウェブ公開に係る業務を受託している公益財団法人沖縄県文化振興会は、そうした不完全な資料を補って長期保存、ウェブ公開している（以下、撮影を省略した部分を補うことを「補撮影」とする。）。本稿は、補撮影の撮影仕様と画像データの作成仕様を紹介し、これらデジタル化した資料画像データを使った活用例とその選択過程を記録することを目的とする。

## 1 原資料の補撮影

表 1 は、補撮影の仕様及びデジタルデータ作成仕様をまとめたものである。原資料の本撮影<sup>2</sup>はオー

---

<sup>†</sup> おおた あやこ 公益財団法人沖縄県文化振興会公文書管理課 公文書主任専門員

<sup>1</sup> 大田 文子「琉球政府文書デジタル・アーカイブズ推進事業における資料画像の長期保存及び公開システムに関するレポート」『沖縄県公文書館研究紀要第 21 号』（沖縄県文化振興会 2019）p.24

<sup>2</sup> 本稿では、琉球政府文書デジタル・アーカイブズ推進事業の資料のデジタル化に係る業務で原資料の撮影をすることを「本撮影」という。

バーヘッド式ブックスキャナであるが、補撮影は、大判文書はA0サイズに対応したシートスルースキャナ、A3サイズまでの文書の場合はフラットベッドスキャナを用いてスキャニングしている。他に、本撮影は資料を綴じたまま撮影しているが、補撮影時は綴じを外して一枚ずつスキャニングしていることが特筆すべき違いといえる。

表1 補撮影時の撮影仕様及び画像データ作成仕様

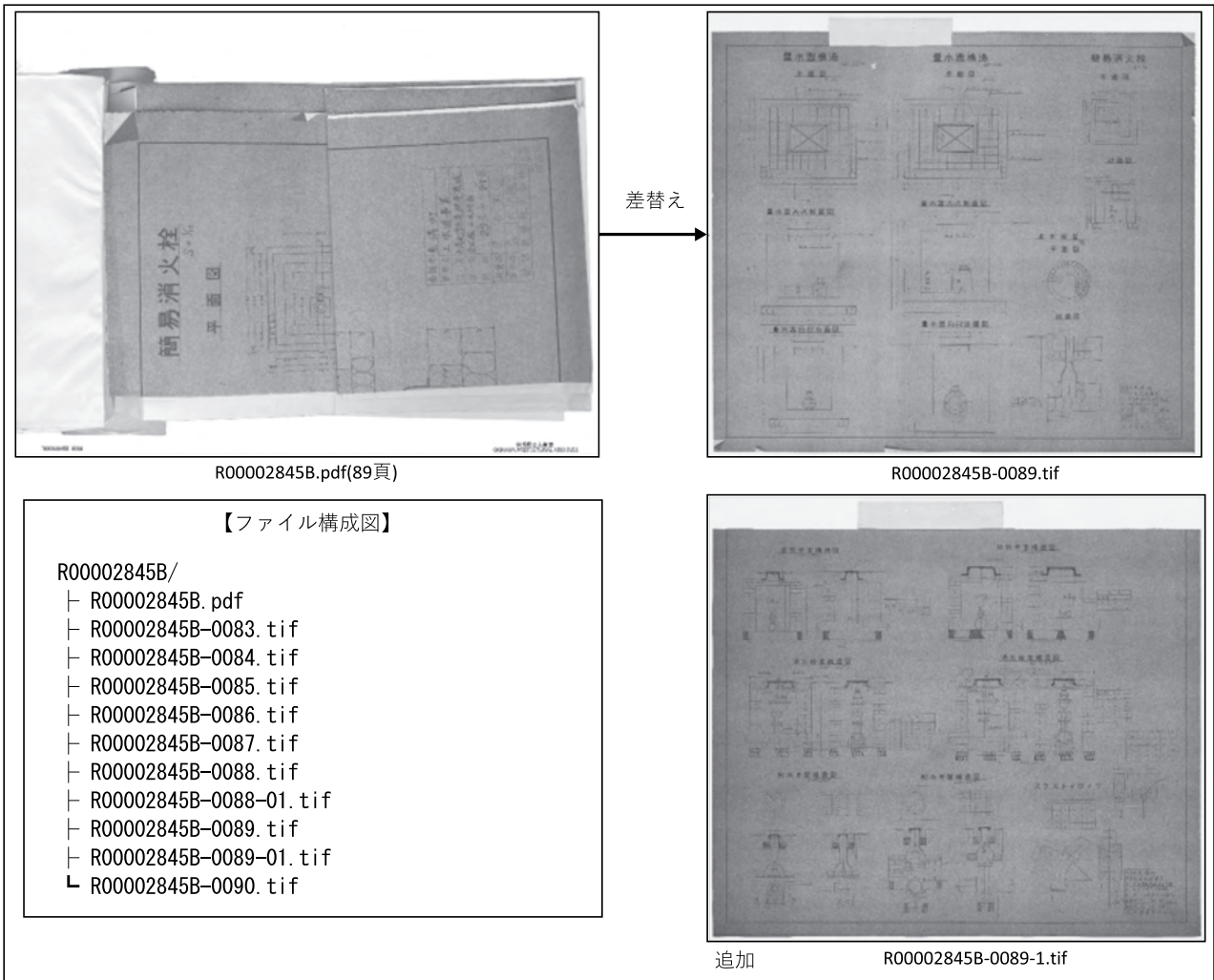
		大判 (A3 サイズを超える) 省略資料	A3 サイズ以下の省略資料
補撮影仕様	撮影機材	A0サイズ対応 シートスルースキャナ	A3 フラットベッドスキャナ
	撮影解像度	300dpi	300dpi
	撮影対象文書	デジタル化事業において撮影を省略または後回しされた文書 (部分) すべて撮影を省略された文書を含む資料に2分割で撮影された大判サイズの文書 綴じを外し、利用制限情報である袋掛け及び付箋を取り除きすべて撮影する。 白紙のページ、同一資料他、すべてスキャニングする。 A0サイズを超える場合、分割してスキャニングする	
画像データ作成仕様	データフォーマット	TIFF (非圧縮) / 拡張子「.tif」	
	ファイル名	原資料の資料コード・(挿入する) ページ番号-00 (01 から始まる 2桁の整数 / ただし、100以上の挿入がある場合は3桁)	
	トリミング	文書ではない背景はトリミングする。	
	背景色	スキャン時の背景のまま (白色)	
	資料画像下部	なし	
	メタタグ	作成者：沖縄県公文書館	

画像データ作成仕様については、ファイルフォーマットの違いが挙げられる。補撮影したデータの作成は本事業中期の本撮影デジタルデータと同様、非圧縮のTIFF (Tagged Image File Format) であるのに対し、初期の本撮影デジタルデータはPDF (Portable Document Format) である。長期保存、ウェブ公開用の資料画像は、撮影および作成方法の違うデータを一括りにする必要がある。一括りにするときは、本撮影時に作成された資料画像データに変更や削除などの改編は加えない。

そこで、補撮影画像データを追加する際、補撮影画像データのファイル名を、初期の資料データの場合はPDFの余白部分に付されている「原資料の資料コード+ページ番号」に、中期の資料データの場合はファイル名である「原資料の資料コード+ページ番号」に枝番号を付け、資料画像を挿入する位置を特定できるようにした。補撮影画像データには、補撮影対象の資料のうち本撮影時に作成された資料画像に「個人情報等があるため閲覧できません」と記したターゲットが含まれるものや、大判サイズの文書を広げずに折りたたんだまま撮影している資料、大判サイズの資料の一部が本撮影時に2分割で撮影されている資料画像を分割せずに補撮影したものなど、本撮影時の画像データの一部を差替えるものがある。差替えの場合の補撮影画像データのファイル名は、原資料の資料コードに

差し替えて文書のページ番号のみをファイル名とした。図1は、本撮影時、大判サイズの文書を2枚重ねて折りたたんだまま撮影した資料画像データと補撮影した資料画像データを一括りにした資料画像のサンプルとファイル構成図である。

図1 本撮影画像と補撮影画像及びそのファイル構成

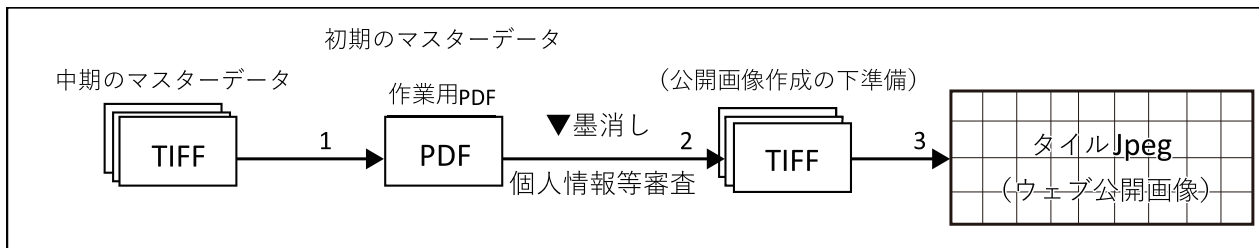


## 2 ウェブ公開用画像の作成

本章では、本事業においてデジタル化した画像データ（以下「マスターデータ」とする。）を用いたウェブ公開用画像の作成工程と各工程で使用する画像データの変換仕様や構造について順を追ってみていく。

まず、マスターデータから、ウェブ公開用画像データを作成するまでの工程を図2に示す。資料画像は、ウェブ公開するために、個人情報等の利用制限情報の有無を審査し、その結果、利用が制限される情報がある場合にはその部分をウェブ非公開情報としてマスキングする必要がある。TIFFはファイルサイズが大きいため、PCのディスプレイに表示するだけでも時間がかかり、審査担当者の負担が大きいため作業用の画像として最適でない。また、ファイルサイズが大きいと配信速度に影響することから、TIFFは公開用の画像としても適さない。そのため、マスターデータTIFFは、まず、個人情報等の審査作業用のファイルフォーマット(PDF)に変換し、その後、さらに公開用画像に変換する。

図2 ウェブ公開用画像データ作成工程



### 2-1 作業用 PDF の作成

最初に、マスターデータ TIFF から作業用の PDF を作成する。個人情報等、利用が制限される情報のマスキングは、画像の上から黒いスタンプなどを貼り付けるだけではなく、はがれないように画像そのものを削除する必要がある。そこで、Adobe Acrobat Pro (以下「Acrobat」とする。)を使い、マスターデータを PDF にフォーマット変換して作業用資料画像データファイルを作成し「墨消し」ツールで資料画像から非公開の情報を部分的にマスキングする方法を採った。Adobe Acrobat マニュアルによると、「墨消し」ツールは「文書から表示されているテキストやグラフィックを完全に削除する処理」で、「表示されているテキストやグラフィックを完全に削除する処理」であるため、マスキングには最適な機能である<sup>3</sup>。また、PDF にフォーマット変換することにより、PC の表示速度を上げることができるため作業効率も上がる。

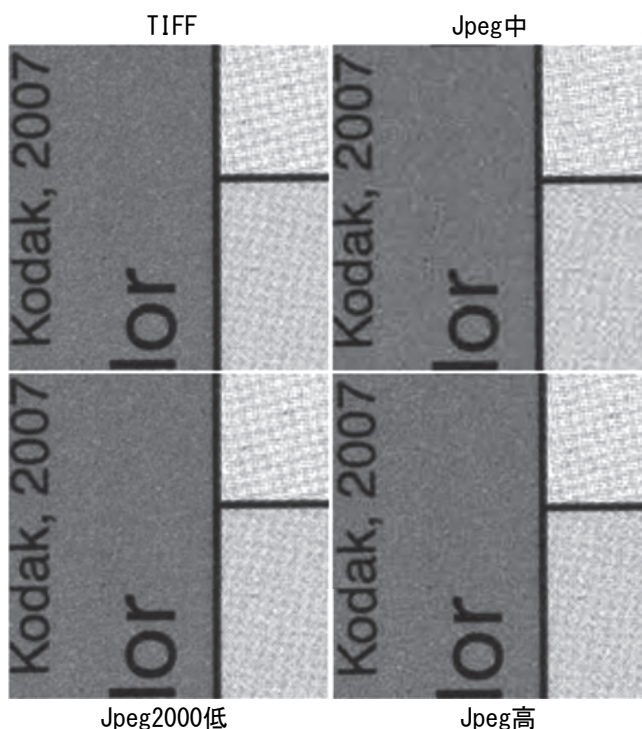
しかし、一言に「PDF にフォーマット変換」といっても画像データのフォーマット変換時には、どのような圧縮をするか選択をしなければならない。PDF は「PDF」というコンテナの中に Jpeg (Joint Photographic Experts Group) などのビットマップ画像を格納するイメージで、TIFF をどのファイルフォーマットに圧縮(変換)して格納するかを決めなければならない。Acrobat で TIFF から PDF へファイル変換する際、圧縮に係る選択肢は、「ZIP」「Jpeg」「Jpeg2000」の3つのフォーマットがあり、そのうち「Jpeg」「Jpeg2000」には細かく「最高」「高」「中」「低」「最低」、さらに「Jpeg2000」は「画質劣化なし」の12パターンの圧縮レベルがある。そこで、資料画像データ数が259枚の標準的な中期の資料画像 TIFF をそれぞれの圧縮方法でファイルサイズと変換時間を計った(表2)。その結果、ZIP 圧縮、Jpeg2000 圧縮はかなりの時間がかかり、高画質になるとファイルサイズもかなり大きいことがわかった。

<sup>3</sup> Adobe Acrobat マニュアル「PDF 内の機密情報の削除」(<https://helpx.adobe.com/jp/acrobat/using/removing-sensitive-content-pdfs.html> 2020.8.11)

表 2 ファイルサイズと変換時間

圧縮	ファイルサイズ (単位 /MB)	変換時間 (H:M)
TIFF	15,080.71	-
ZIP	9,395.29	0:17
JPEG	最高	723.85
	高	293.23
	中	148.21
	低	96.58
	最低	66.69
Jpeg2000	劣化なし	4,840.87
	最高	1,014.75
	高	685.68
	中	520.49
	低	355.22
	最低	188.04

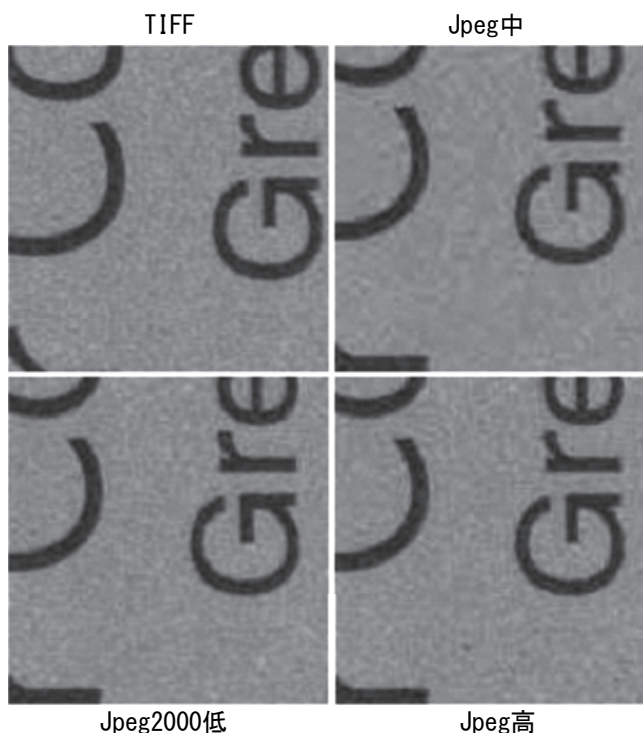
図 3 PDF 圧縮フォーマット別 PDF 画像



次に、作成した PDF のファイルサイズの中央値である 438MB 付近の Jpeg 「高」「中」と Jpeg2000 「低」の画像を比べた (図 3)。その結果、Jpeg2000 「低」の方が TIFF に近く高画質で、Jpeg 画像には少しノイズが発生しているが、どの画像も特別な差は感じられない。しかし、Jpeg2000 を格納した PDF に変換するには、1 資料の変換に 24 分から 30 分かかる。本事業でのデジタル化資料数は、130,000 点を超える計画であり、マスターデータが PDF である 2015 年度 (平成 27) までのデジタル化資料数 51,155 点を除くと約 78,800 点の資料画像データの変換をすることになる。単純に計算しても、この一工程だけで 39,400 時間かかることになる。不必要な時間をかけることは避けたい。一方で、作業用の画像は、後にウェブ公開用の画像に変換する。つまり、ここでの選択が公開用画像の画質に影響する。ウェブ公開用の画像データのファイルサイズは小さいほどいいが、閲覧に耐えうる画質でなければならない。

そこで、最終的なウェブ公開画像の画質を比較、検討することにした (図 4)。Jpeg 「中」は特にノイズが発生しているものの、文字情報を閲覧する画像としては可読でき、閲覧に耐えうる画質であると言える。ファイルサイ

図 4 PDF 圧縮フォーマット別公開用画像



ズ、画質、変換時間などを総合的に勘案して検討した結果、Jpeg「中」画質圧縮を採用することとした。なお、初期のマスターデータはPDFファイルであるため、作業用資料画像データファイルの作成は省略することができるが、閲覧に最適なJpeg圧縮画像データを格納したPDFであるため、図4で見たJpeg「中」画質圧縮画像データを格納したPDFよりノイズが目立つ資料画像もある。また、1章で述べた、補撮影資料画像データがある場合には、補撮影資料画像データをTIFFからJpeg「中」画質圧縮画像データを格納したPDFに変換し、該当箇所に挿入する作業が必要となる。

## 2-2 PDF から TIFF へ

次に、公開用画像データを作成するための下準備を行う。作業用のPDFは、複数の画像ファイルを1つに結合しているため、個人情報等の利用制限情報の審査を終えたPDF資料画像から1ページずつ分離した画像ファイルフォーマットに書き出す必要がある。書き出すファイルフォーマットとして、Jpegの方がTIFFよりファイルサイズが軽いことから、フォーマット変換にかかる時間も短縮できるが、Jpegは保存する度に劣化するファイルフォーマットである。審査し、マスキングするときに1回、PDFからJpegへ書き出すときに1回、後述する公開用画像データを作成するときにさらにもう1回と3度画質を劣化させることになる。一方、TIFFへ書き出す場合は、PDFから書き出す際の劣化を防ぐことができる。また、先に述べたが、初期のマスターデータは既に閲覧に最適なファイルサイズになるように圧縮したJpegを格納しているPDFである。これ以上の劣化は避けたい。これらの理由により、PDFからの書き出しはTIFFとすることとした。

## 2-3 タイル分割 Jpeg

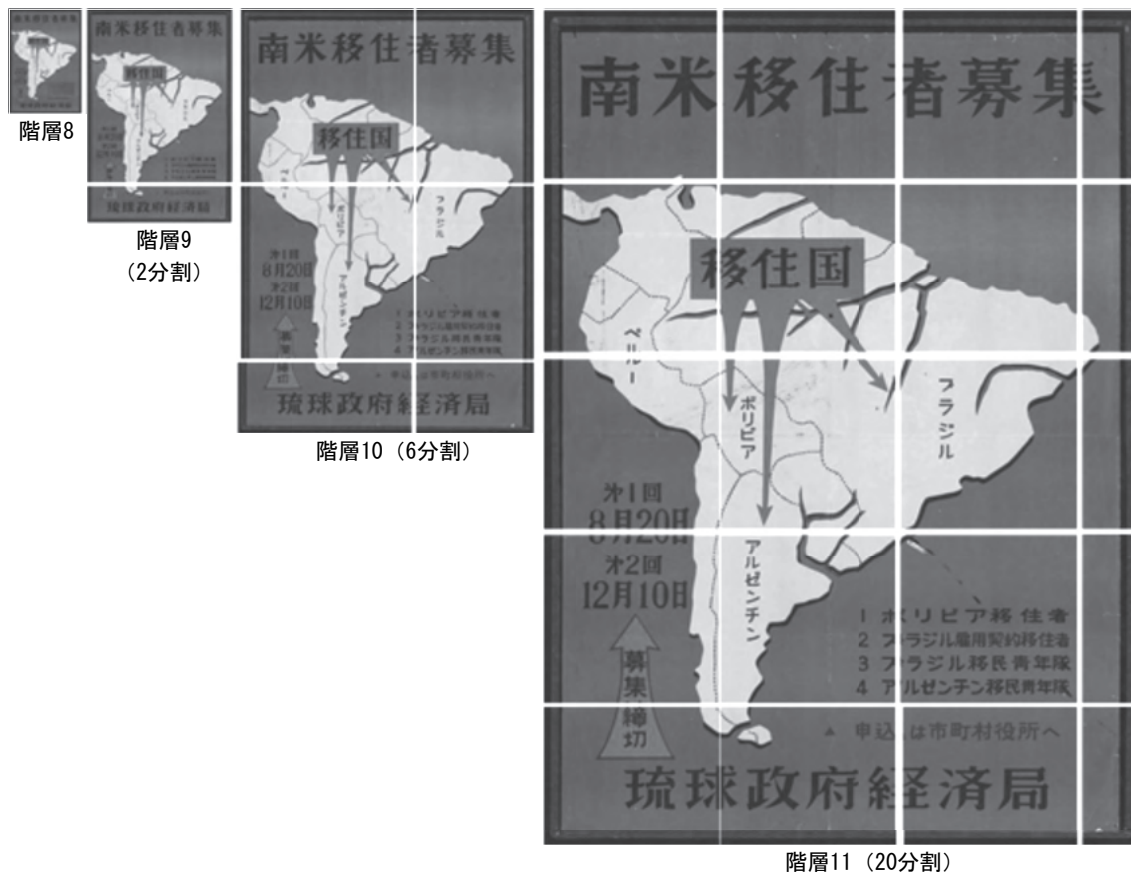
ここでようやく、公開用画像データの作成に取りかかることができる。本事業のウェブ公開資料画像は、2016年度（平成28）に改修した画像ビューア「Contents View FLEX」を介してウェブ公開しており、資料画像は、専用のコンバータを使って作成する<sup>4</sup>。「Contents View FLEX」は、画像の拡大縮小レベルに応じて必要な部分のみを読み込み、表示させる仕組みである。この仕組みに応じた画像を生成する専用コンバータは、1ファイルの画像データを拡大縮小レベルごとに階層化したJpeg群とその階層構造を管理し、Jpeg群を画像ビューアに表示させるために必要なjsファイルやxmlファイル、画像ビューアをブラウザに表示するためのhtmlを作り出す。Jpeg群は、高解像度の資料画像データ1に対し、2分の1、4分の1とだんだんに1ピクセル四方になるまで縦横解像度を低減させたいくつもの画像から成るフォルダで構成されている。そのため、1ピクセル四方の画像データが1ファイルのみを格納したフォルダから256pixel四方の画像データが2,684ファイル以上もあるフォルダもあり、変換前の資料画像の解像度が高いほど階層が増える。『南米移住国移住地地図 南米移住者募集ポスター』（R00054241B 沖縄県公文書館所蔵）3頁のウェブ公開資料画像（図5）を例にみると、変換前は横6,791pixel×縦9,803pixelの資料画像データで14階層、1,053分割である。ウェブ公開画像を閲覧する機器やその拡大縮小率によって読み込まれる階層は異なるが、図5の8階層目はスマートフォンなどで資料画像ビューアを閲覧したときのサムネイル部分に表示されることが多く、14階層目の画像は拡大率が最大のときに活躍する。

以上がマスターデータを用いたウェブ公開用画像の作成工程と、大量の資料画像を配信するための

<sup>4</sup> 大田 文子「琉球政府文書デジタル・アーカイブズ推進事業における資料画像の長期保存及び公開システムに関するレポート」、前掲,p.28

変換仕様及びその検討過程である。

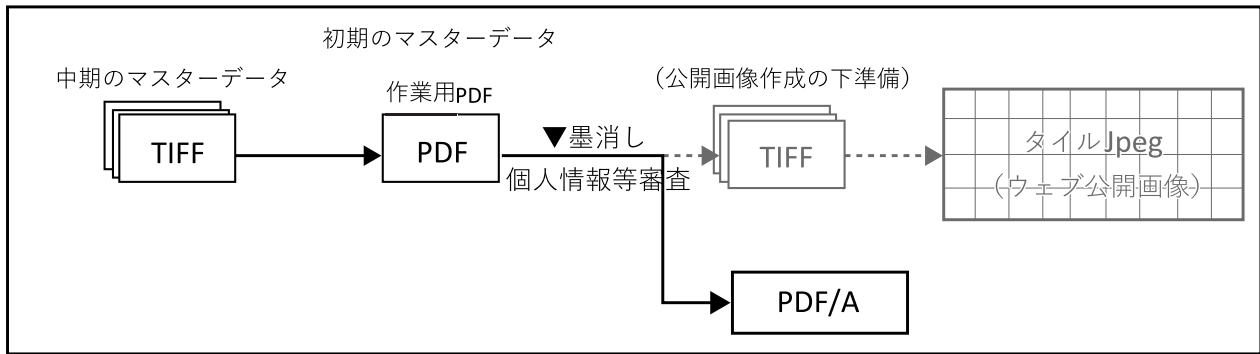
図5 ウェブ公開資料画像 (R00054241B 沖縄県公文書館所蔵) 3頁



### 3 閲覧室閲覧用画像 PDF/A の作成

現在、沖縄県公文書館閲覧室での資料の閲覧は、マイクロフィルムや光ディスクなどの複製物がない場合、原資料を利用に供している。原資料は紙媒体であることが多く、閲覧に供するほど原資料は劣化する。琉球政府文書は新しいものでも作成から約 50 年は経過しており、その紙質は脆く朽ちてしまうものや感熱紙の資料で光にあたるほど情報を失ってしまうものもある。そのため、原資料の保存の観点から、本事業では、ウェブ公開用のデータだけでなく、閲覧室でも提供できるよう、閲覧室閲覧用（以下「閲覧用」とする。）の資料画像も作成している。ウェブ公開用の資料画像データは、前述したように大量の画像データで構成されており、ウェブ公開には適している。しかし、ローカルドライブで閲覧する、必要な頁を印刷する場合などは OS に依存しないファイルフォーマットである PDF が最適であると考えた。PDF は、ドキュメントや図表などを印刷に最適な形に保つことができ、また、無償のアプリケーションソフトが配布されていることから、閲覧室で利用するファイルフォーマットとしては扱いやすい。そこで、閲覧用 PDF データは、ウェブ公開用画像を作成する前の個人情報等の審査を終えた PDF を利用することとした。変換工程は図 6 の通りである。

図6 閲覧用PDF変換工程



PDFは、Adobeによって開発され、現在では国際標準(ISO32000-1:2008)化されているファイルフォーマットである。PDFにはいくつかの規格があり、それぞれ国際標準化機構(ISO)の認定を受けている。まず、閲覧用として最適なPDF規格を選定し、選定した規格の準拠レベルを決める必要がある。PDFの規格のうち、文書に関連する規格とその内容を表3に示す<sup>5</sup>。PDF/Aは、視覚的な外観が維持されることを保証している。また、安定した長期アーカイブのための国際標準規格である。閲覧用という利用目的に鑑みるとPDF/Aに準拠することが適していると考え、この規格を採用することとした。

PDF/AにはPDF/A-1(ISO19005-1:2005)、PDF/A-2(ISO19005-2:2011)、PDF/A-3(ISO19005-3:2012)といった、3つのパートで構成されている。ISO19005-1によると、PDF/A-1の目的は、ツールやシステムに関係なく、長期間視覚的な外観を維持できるファイル形式を定義すること、このファイル形式に準拠し作成した

PDFファイルにはメタデータとして電子文書のコンテキストや履歴を記録する構造を提供すること、そして、論理構造やセマンティック<sup>6</sup>情報を表示させる枠組みを提供することにある<sup>7</sup>。PDF/A-2(ISO19005-2<sup>8</sup>)の目的は、PDF/A-1と同じである。両者の違いは、基準とするPDFの仕様の違

表3 PDFの規格

PDF	Adobeにより開発された規格。Portable Document formatの略。電子文書や画像をコンパクトに格納し、交換することができるファイル形式。
PDF/A (ISO19005)	電子文書を長期保管用に作成、表示および印刷するための仕様を規定。最終記録文書を自己完結型のファイルとして簡単に維持及び保護し、将来に渡る情報へのアクセスを確保できる。
PDF/X (ISO15930)	印刷・出力用頁の作成、表示および印刷するための仕様を規定。カラースペースやトラッピングといった印刷の重要な要素を左右するPDFを定義。
PDF/E (ISO24517-1)	エンジニアワークフローに用いられる文書を作成、表示及び印刷するための仕様を規定。技術図面・CADファイル等を変換、使用できるようにする。

<sup>5</sup> 小林 徳滋『PDFインフラストラクチャ解説 第1.1版』(アンテナハウス株式会社CAS電子出版 2017年)

<sup>6</sup> セマンティック(semantic)とは、一般に「意味」と訳される。IT用語として、コンピュータにデータの持つ意味を理解させ、処理をさせる技術のことをいう。

<sup>7</sup> ISO, Online Browsing Plat Form Foreword「Document management –Electronic document file format for long-term preservation –Part1: Use of PDF1.4(PDF/A-1)」(<https://www.iso.org/obp/ui/#iso:std:iso:19005:-1:ed-1:v2:en,2005,2020.11.24>)

<sup>8</sup> ISO, Online Browsing Plat Form Foreword「Document management –Electronic document file format for long-term preservation –Part2: Use of ISO 32000-1 (PDF/A-2)」(<https://www.iso.org/obp/ui/#iso:std:50655:en,2011,2020.11.24>)



いであり、PDF/A-1はPDF1.4の仕様を、PDF/A-2はISO32000-1 (PDF1.7)の仕様を基準としている。そこで、PDFの仕様の違いを確認した。PDF1.4<sup>9</sup>とISO32000-1の主な違いは、PDF1.4で禁止されていたJpeg2000圧縮による画像ファイルが使用可能となったことなど、一部の項目が緩和されたことが挙げられる。他に、ファイルサイズの限界値の違いがある。PDF 1.4にはファイルサイズ制限に関する記述がない。しかし、2GBを超える資料画像PDFをAcrobat「プリフライト」ツールを用いてPDF/A-1に適合しているかを解析すると、「2GBを超えている」というエラーメッセージが出現した。エラーの詳細には「Acrobat6まではPDFのファイルサイズが2GBに制限されていました。Acrobat7から2GBを超えるファイルサイズがサポートされています。」と表示される。Acrobat6はPDF1.5の仕様を、PDF/A-1はPDF1.4の仕様を基準としている。このことから2GBを超えるPDFファイルはPDF/A-1に適合しないことがわかる。一方、ISO32000-1 (PDF1.7)の仕様を基準としているPDF/A-2のファイルサイズについては、PDF Reference, Sixth Edition, 1.7に以下のように記されている。

PDF itself has one architectural limit: Because ten digits are allocated to byte offsets, the size of a file is limited to 1010bytes (approximately 10 gigabytes).<sup>10</sup>

つまり、PDF1.7は、ファイルサイズの制限は約10GBまでということがわかる。2章で述べたように、PDFファイルの作成の際にJpeg圧縮をしており、標準的な資料画像PDFファイルは2GBを超えない。しかし、450頁を超える資料画像や大判サイズの文書が含まれる場合、ファイルサイズが2GBを超えることがままあり、すべてをPDF/A-1に準拠したファイルを作成することは困難である。PDF/A-3 (ISO19005-3<sup>11</sup>)は、PDF/A-2と同様、ISO32000-1 (PDF1.7)の仕様を基準としているが、PDF/A-1またはPDF/A-2以外の任意ファイルの埋め込みが可能となった点に違いがある。閲覧用PDFの規格は、資料画像を格納したファイルであるため一番基本的なPDF/A-1を選択したいが、前述したように2GB制限に対応するには困難であることから、PDF/A-2を選択することとした。

さて、次は準拠レベルの選択である。PDF/A-2には「PDF/A-2a (accessibility)」「PDF/A-2b (basic)」「PDF/A-2u (Unicode)」の3種類ある。PDF/A-2aは、ISO19005-2に完全に準拠したファイルである。比べてPDF/A-2bは、PDF/Aの目的であるツールやシステムに関係なく長期間視覚的な外観を維持できるファイル形式を定義することに限定した一部準拠レベル、PDF/A-2uはPDF/A-2bに加えてPDF内テキストのUnicodeの取得を保証している。格納した資料画像すべてにタグを付けて構造化することは不可能に近く、また、閲覧用PDFを作成するためにPDF/Aに完全準拠することに重要な意味を持たないと考え、カラーの再現性が保証され、長期間視覚的な外観を維持できるPDF/A-2bに準拠したPDFを閲覧用資料画像として作成することとした。なお、作成したPDF/A-2bは、作成、改編した履歴や画像情報などがコンテキストやメタデータに記録される。

<sup>9</sup> Adobe Systems Incorporated 「PDF Reference third edition, version 1.4」 ([https://www.adobe.com/content/dam/acom/en/devnet/pdf/pdf\\_reference\\_archive/PDFReference.pdf](https://www.adobe.com/content/dam/acom/en/devnet/pdf/pdf_reference_archive/PDFReference.pdf), 2001, 2020.11.24)

<sup>10</sup> Adobe Systems Incorporated 「PDF Reference Sixth edition, version 1.7」 ([https://www.adobe.com/content/dam/acom/en/devnet/pdf/pdf\\_reference\\_archive/pdf\\_reference\\_1-7.pdf](https://www.adobe.com/content/dam/acom/en/devnet/pdf/pdf_reference_archive/pdf_reference_1-7.pdf), 2006, 2020.11.24)

<sup>11</sup> ISO, Online Browsing Platform Foreword 「Document management – Electronic document file format for long-term preservation – Part 3: Use of ISO 32000-1 with support for embedded files (PDF/A-3)」 (<https://www.iso.org/obp/ui/#iso:std:iso:19005:-3:ed-1:v1:en>, 2020.11.24)

## おわりに

本稿では、本事業でデジタル化した資料画像の活用例を紹介し、画像フォーマットを選択した背景を示してきた。今後残される課題として、1章で述べた、補撮影が必要な文書がまだ存在することが挙げられる。事業完了までに補えなかった場合、公開できる資料画像は一部を省略している旨を利用者に示してウェブ公開することを検討するが、ウェブ公開のためだけでなく不完全な資料は引き続き補う必要がある。また、初期に撮影した資料画像は利用目的の幅を狭めて作成しているため、研究者やマスコミが求める画質に対応しきれないことが想定される。それゆえ、本事業でデジタル化されているが、再撮影が必要となることがある。他に、本事業における個人情報等のマスキングする情報と閲覧室での利用を制限する情報が異なるため、3章で述べた、閲覧用PDFの再作成が必要となる場合がある。たとえば、琉球政府文書に「個人の財産」に関する情報が記録されていた場合、閲覧室では沖縄県公文書館管理規則第4条別表に則し「個人の秘密であって、当該情報を公にすることにより、当該個人の権利利益を不当に害するおそれのあるもの」として、50年の経過を待ってその部分が公開される。しかし、ウェブ公開画像はたとえ50年を経過していても、個人の財産に関する情報はウェブで公にすることが適当でない個人の秘密に該当するものとし、マスキングしている。閲覧用PDFを利用する閲覧室の利用者からマスキング部分の開示を求められたときは、マスターデータからPDFを再作成し、再審査の後、新たに閲覧用のデータを作り直す必要がある。

本事業は2021年度（令和3）に事業を完了する予定であり、完了時にはデジタル化資料は130,000点を超える。これらの膨大な資料画像は「複製物の作成」や「ウェブ公開」だけでなく、利用の目的に応じてフォーマット変換しながら活用の幅を広げて発信していく必要がある。本事業では、とにかく大量の資料画像を配信することに重きを置いてきた。たとえば重要文化財の巻物などをデジタル化する場合のように一つの資料を丁寧に公開する場合とは選択肢が違ってくることもあるだろう。事業完了後に残される課題もある。本事業が完了したら終了ではなく、本事業の成果は引き継がれ、改良されて活用され続ける。本稿が今後の選択の参考になればと思う。

附録：撮影及び画像データ作成仕様

	初期			中期		後期
	2013	2014	2015	2016	2017	2018 ~ 2021 (* 2021 は見込み)
<b>撮影仕様</b>						
撮影機材	オーバーヘッド式ブックスキャナ ・A3 サイズ対応			オーバーヘッド式ブックスキャナ ・A3 サイズ対応 ・大判対応 (A1 サイズ)		
撮影解像度	350dpi 以上			300dpi	300dpi 以上	
表表紙	資料コードが見える面を表表紙とする。			資料コードが見える面を表表紙とする。 表紙にカラーチャート及びスケールを入れて撮影。		
撮影対象文書	見開き状態で左右両ページを撮影できる場合、両ページを 1 コマとして撮影。					
	白紙のページは撮影しない。			白紙のページも撮影する。		
	大判文書 (A3 サイズ対応スキャナで文書の四隅が撮影ができない、かつ、2 分割を超える分割撮影が必要な文書) が少ない資料は、資料は撮影対象とするが、大判文書のみ撮影しない (大判対応スキャナ導入年度後に一部分撮影する)。			大判文書 (A3 サイズ対応スキャナで文書の四隅が撮影ができない、かつ、2 分割を超える分割撮影が必要な文書) は大判対応スキャナで撮影する。		
	A3 サイズ対応スキャナで文書の四隅が収まらないときで、撮影できない範囲に文字情報がない場合は、見切れていても分割しない。					
	綴じられていて開けない資料は綴じたまま撮影する。			綴じられていて開けない資料がある場合は、綴じを外す、または緩めて撮影する。		
	利用制限情報である袋掛け及び付箋は外さず、該当文書の撮影はしない。			利用制限情報である袋掛け及び付箋を取り除きすべて撮影する。		
ノド部の撮影	可読できればよし。			綴じを外す、または緩めて撮影する。		
同一資料	様式のみで同一資料が続く場合は同一資料の一部目を撮影し、その後は撮影しない。		すべて撮影する。			
<b>画像データの作成仕様</b>						
画像データフォーマット	マルチ PDF (Jpeg 圧縮)			TIFF (非圧縮)		
	圧縮率の定めなし					
ファイル名	原資料の資料コード			原資料の資料コード・ページ番号 (0000 から始まる 4 桁の整数)		
トリミング	文書ではない背景はトリミングする。					
背景色	白色に色調補正			撮影時の背景のまま		
資料画像下部	画像の下部に余白を作り、原資料の資料コード・ページ番号 (0000 から始まる 4 桁の整数) と、作成者名を付加。			なし		
メタタグ	-			作成者：沖縄県公文書館		
格納媒体	可搬ハードディスク					