

平成30年7月

資料デジタル化研修

資料デジタル化の基礎

国立国会図書館 関西館電子図書館課

構成

1. はじめに
 2. デジタル化に関する基礎知識
 3. デジタル化実務のポイント
- おわりに

1. はじめに

- (1) この講義の狙い【3】
- (2) 資料デジタル化の流れ【4】
- (3) 参考情報【5】

1. はじめに

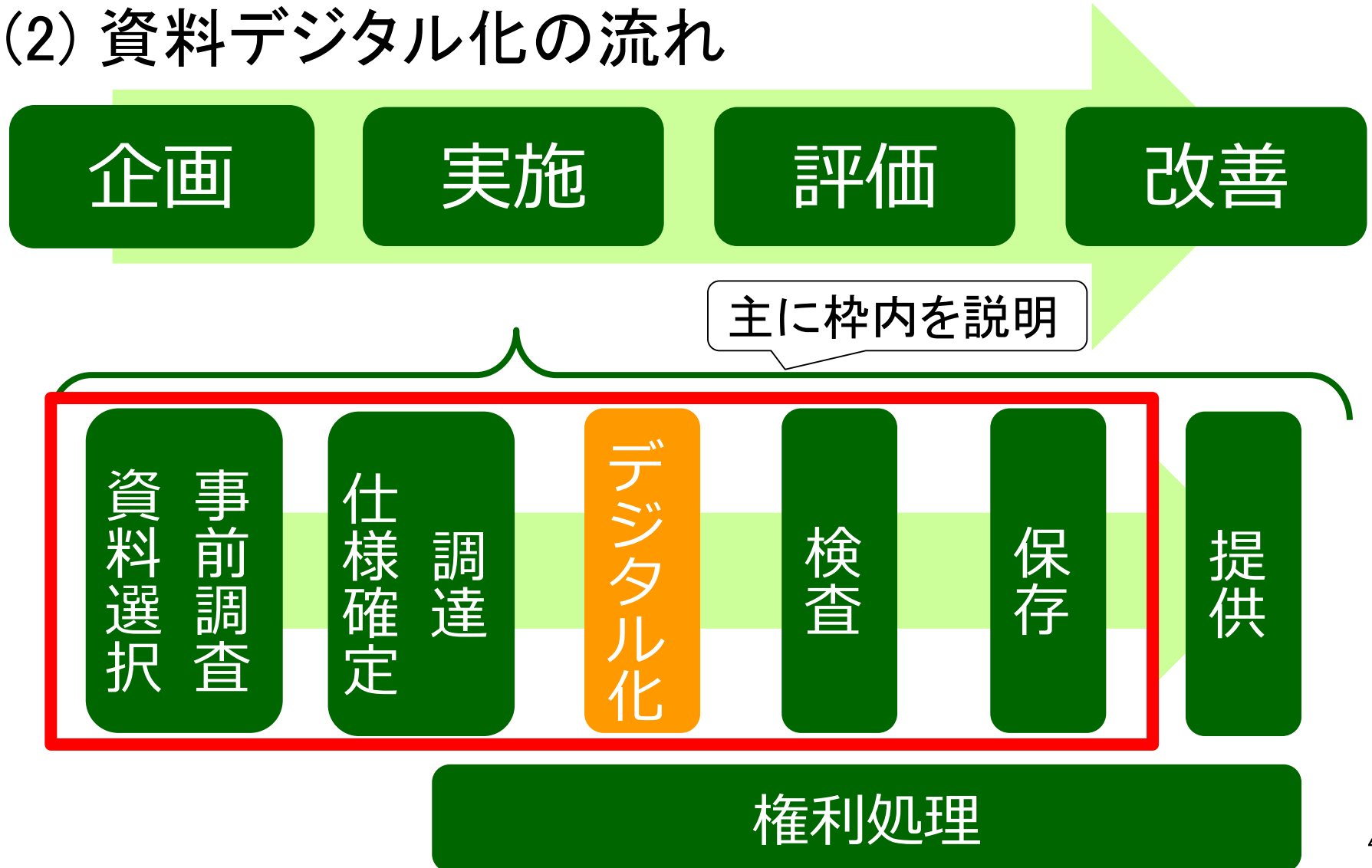
(1) この講義の狙い

「資料デジタル化に関する業務経験・知識が少ない新任担当者」が
「仕様書の作製や工程管理のために必要な基礎知識・実務上の注意点を理解すること」

- ・新たにデジタル化を実施することになったが組織内にノウハウがない
- ・既にデジタル化を進めているところ、新たに担当者になり、知識習得が必要
- ・自分の経験を新人に伝える必要があるが一から引き継ぎ資料を作るのが大変 etc...

1. はじめに

(2) 資料デジタル化の流れ



1. はじめに

(3) 参考情報

- ・「国立国会図書館資料デジタル化の手引」
(以下「手引」)

(<http://www.ndl.go.jp/jp/aboutus/digitization/guide.html>)

→国立国会図書館(以下「NDL」)が過去のデジタル化事業から得た知見等を整理(2017年改訂)

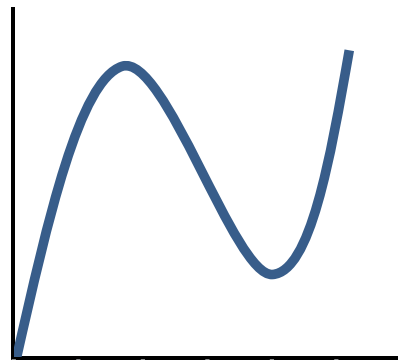
2. デジタル化に関する基礎知識

- 2. 1 「デジタル化」とは【7】
- 2. 2 画像データと解像度【10】
- 2. 3 階調・色深度【17】
- 2. 4 ファイルサイズ・フォーマット【22】
- 2. 5 色空間・カラーマネジメント【28】
- 2. 6 スキャナの種類【35】

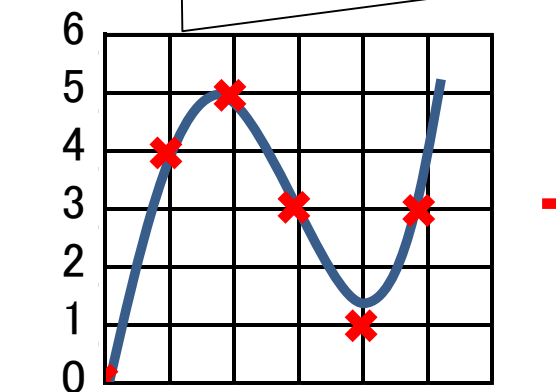
2. 1 「デジタル化」とは

- 「アナログ」(連続的な変化)の情報を「デジタル」(区切りをつけた数値表現)の情報に変換すること

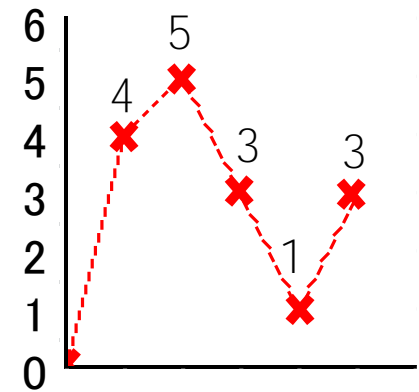
区切ることを「標本化」
数値に置き換えることを「量子化」と呼ぶ



アナログ情報



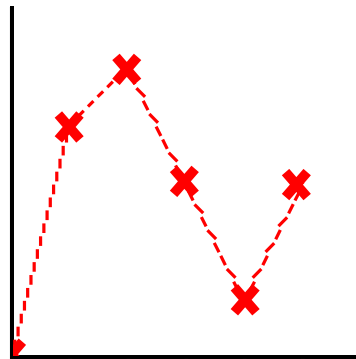
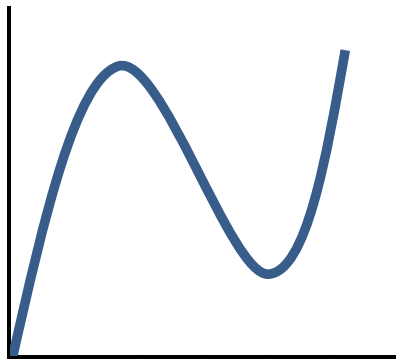
一定間隔に区切って
最寄りの交点の
数値に置き換え



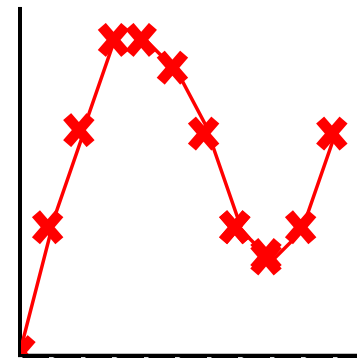
デジタル情報

2. 1 「デジタル化」とは

- 数値で表現する(デジタル化する)とコンピュータで扱えるようになる
- 細かく区切ると元の情報に近くなる



6分割



11分割

2. 1 「デジタル化」とは

・図書館資料のデジタル化の例

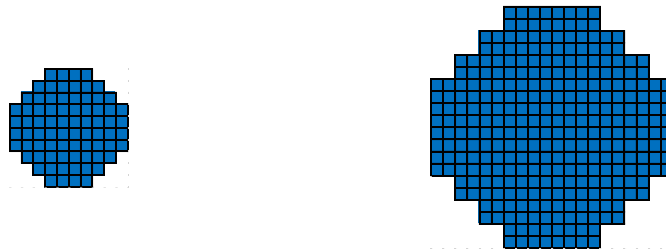
- ①紙やマイクロフィルム上の文字・図表・写真等
→画像データ化
- ②録音資料(カセットテープ、レコード等)
→音声データ化
- ③映像資料(ビデオテープ、レーザーディスク等)
→映像データ化
- ④目次や本文の文字情報
→テキストデータ化(例:青空文庫)

2.2 画像データと解像度

(1) 画像データの種類

① ペイント形式

- ・画像を「点の集まり」として記録
 - ・様々な色合い・濃さの点を組み合わせ
写真や複雑な図形を表現できる
- デジタル化画像、デジタル写真等で使用



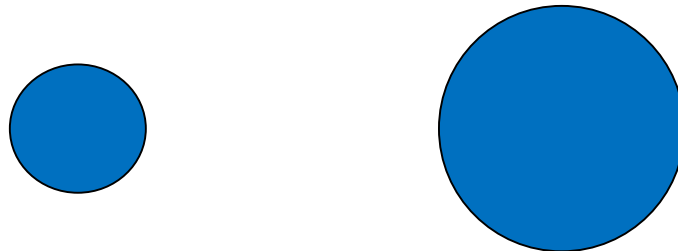
2.2 画像データと解像度

(1) 画像データの種類

②ベクタ形式

- ・画像を「図形」(点、線、面の情報)として記録
- ・数値計算して描画するので、拡大・縮小しても画像が劣化せず、ファイルサイズもあまり変わらない
- ・複雑な図形の表現には向かない

→文字フォント、MS Officeの図形等で使用



2.2 画像データと解像度

(2) 「解像度」とは
＝画像データを構成する「点」の密度
「点」の考え方は2種類ある

①ピクセル(pixel)、画素

・ペイントデータを構成する情報としての最小単位

②ドット(dot)

・プリンタ、ディスプレイ、スキャナ等の機器で
画像を表現・読み取る際の物理的な最小単位

2. 2 画像データと解像度

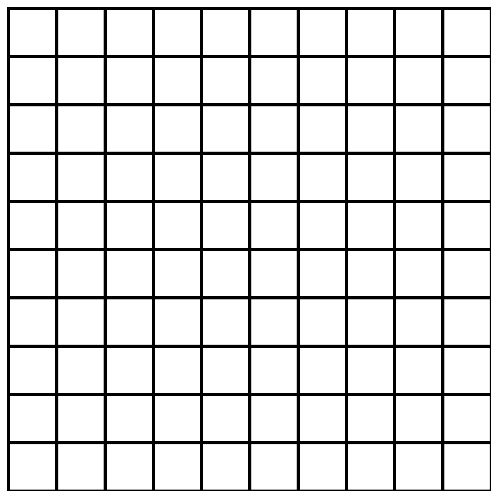
(3) 「絶対解像度」と「相対解像度」

①「絶対解像度」

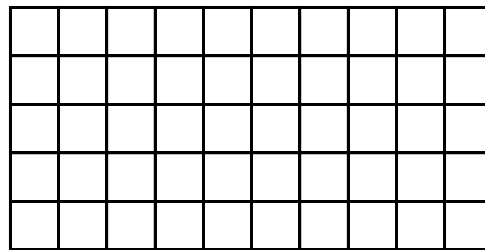
＝画像の大きさ、画像を構成する点（画素）の数

単位：ピクセル (pixel)

デジタルカメラ等の〇〇万画素



100ピクセル(10×10)



50ピクセル(5×10)

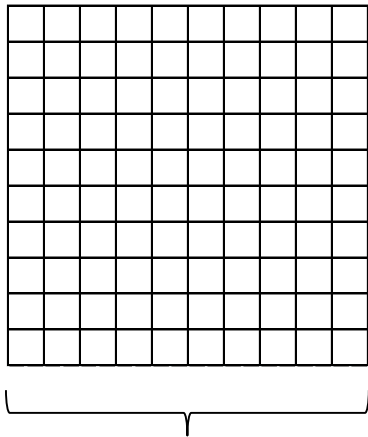
2.2 画像データと解像度

(3) 「絶対解像度」と「相対解像度」

②「相対解像度」=画像の密度

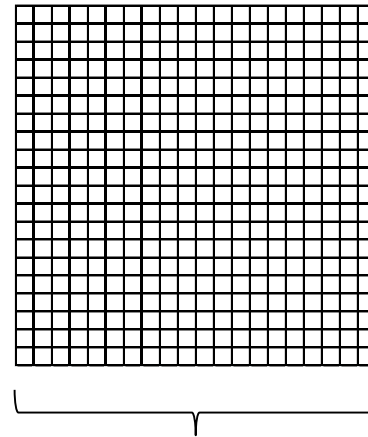
1インチ(25.4mm)当たりの点の数

単位:dpi (dots per inch), ppi (pixel per inch)



1インチ当たり10個
=10dpi

スキャナの
読み取り性能の
指標



1インチ当たり20個
=20dpi

2. 2 画像データと解像度

(4) 「光学解像度」と「補間解像度」

- 「光学解像度」= 実際にスキャナが読み込める解像度
 - 「補間解像度」= 画像補正によって疑似的に実際の光学解像度以上の画質を表現する解像度
- スキャナの性能は、光学解像度で評価するのが望ましい

2.2 画像データと解像度

(5) 解像度の設定

近年NDLでは400dpiが基本
(300-400dpi程度とする機関が多い)

- ・解像度とファイルサイズは比例する
→バランスを考慮して設定する

(参考) 白黒画像の解像度目安

フォントサイズ・線画の有無	判読できる解像度
10 ポイント以上	150dpi
8 ポイント以上	200dpi
6 ポイント以上、線画混じり	300dpi
4 ポイント以上、線画中心	400dpi

2.3 階調・色深度

(1) 「階調」・「色深度」とは

- ・表現できる濃淡の段階、色の数を表す
- ・図書館資料のデジタル化では
次の3種類のいずれかを用いることが多い

- ①白黒2値(1ビット)
- ②グレイスケール(8ビット等)
- ③カラー(24ビット等)

2.3 階調・色深度

(2) ビット(bit)とは

→コンピュータの情報処理の最小単位

「0」と「1」の2通りの状態を表現できる(2進数)

→2ビットならば4通り(00,01,10,11)

3ビットならば8通り

(000,001,010,011,100,101,110,111)

...nビットならば2のn乗通りの状態を表現

2.3 階調・色深度

(3) 白黒2値(1ビット)

- ・「白」と「黒」の2色だけで表現する
→ 中間の階調を表現できないため、
写真・図表や細かな文字を含む資料には不向き

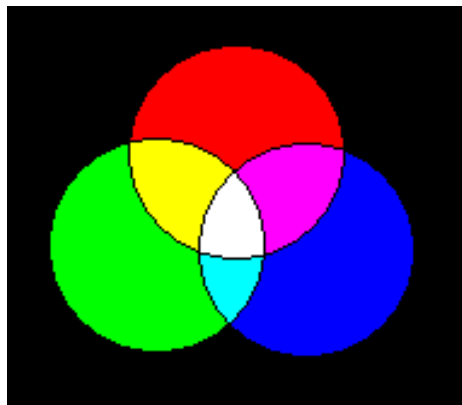
(4) グレイスケール(8ビット=256階調の場合)

- ・「真っ黒」を「0」、「真っ白」を「255」として、
その間を白黒の割合を変えた灰色で表現する

2.3 階調・色深度

(5) カラー(24ビット=16,777,216色の場合)

- ・赤色 (Red)、緑色 (Green)、青色 (Blue) にそれぞれ8ビット=256階調を割り当てその組み合わせで多様な色を表現できる



- ・光の三原色 Red, Green, Blue (RGB)
すべて混ぜる $(R, G, B) = (255, 255, 255)$
→ 白色
すべての要素が「0」 $(R, G, B) = (0, 0, 0)$
→ 黒色

2.3 階調・色深度



白黒2値
(1ビット)



グレイスケール
(8ビット)



カラー
(24ビット)

2.4 ファイルサイズ・フォーマット

(1) ファイルサイズ

- ・ファイルサイズは絶対解像度(画像の大きさ)と階調・色深度に比例する
 - ・ファイルサイズが大きいと表示速度が遅くなり、必要となるストレージ容量も増え、保存コストが高まる
- 用途に応じて解像度・階調を選択する
必要に応じて画像圧縮する

2.4 ファイルサイズ・フォーマット

(2) 用途別の画像種別

①保存用画像

- ・将来の利用や保存のために作製する画像

②提供用画像

- ・利用者への提供用に作成する画像

保存用画像を利用しやすいサイズに圧縮し作製

③サムネイル画像

- ・画像を一覧する際等に用いる小さな画像

2. 4 ファイルサイズ・フォーマット

(3) ファイルの圧縮

- ① 可逆圧縮 = 圧縮前のデータに戻せる
 - ・圧縮率は低めだが、画像が劣化しない
 - 「保存用画像」に向く

- ② 非可逆圧縮 = 圧縮前のデータに戻せない
 - ・画像が劣化しやすいが、圧縮率が高い
 - 「提供用画像」「サムネイル画像」に向く

2.4 ファイルサイズ・フォーマット

(4) 主要なファイルフォーマット

フォーマット (拡張子)	圧縮	用途	備考
TIFF (.tif(f))	基本的に 非圧縮	保存用	スキャニング時の保存用画像として利用
JPEG2000 (.jp2)	可逆	保存用	JPEGよりも高品質・高圧縮、高機能 ウェブブラウザではまだ標準対応していない
	非可逆	提供用	
JPEG (.jp(e)g)	非可逆	提供用	圧縮率に優れる ウェブサイトの画像等に標準的に利用
GIF (.gif)	可逆	提供用	ウェブサイトのアイコンとして利用(8bit)
PNG (.png)	可逆	提供用	GIFに替わる提供用画像として利用(~48bit)
PDF (.pdf)	可逆/ 非可逆	提供用	OS, フォントの違いに影響されず表示可能

☞「2.2.1 画像フォーマットの選択」『手引』pp.14-15

2. 4 ファイルサイズ・フォーマット

(5) ファイルフォーマット検討時のポイント

- ある程度普及したフォーマットを使うのが安全
- デジタルデータを公開するシステムで問題なく扱えるフォーマットを選ぶ

- 保存用画像でよく用いられるフォーマット
→ TIFF (非圧縮)、JPEG2000 (可逆圧縮)
- 提供用画像でよく用いられるフォーマット
→ JPEG2000 (非可逆圧縮)、JPEG、PDF

2.4 ファイルサイズ・フォーマット

(6) NDLで使用しているファイルフォーマット

①保存用画像:JPEG2000(可逆圧縮)

②提供用画像:JPEG2000(非可逆圧縮)

→近年は1/50程度に圧縮

(過去には1/100程度の時期もあり)

③サムネイル画像:JPEG(非可逆圧縮)

2.5 色空間・カラーマネジメント

(1) 「色」についての考え方

- ・原資料の色を忠実に再現するためには色の表現を調整する必要がある

「見栄えよく調整したい」
「読めればよい」等の考え方もありうるが
NDLは忠実再現を目指している

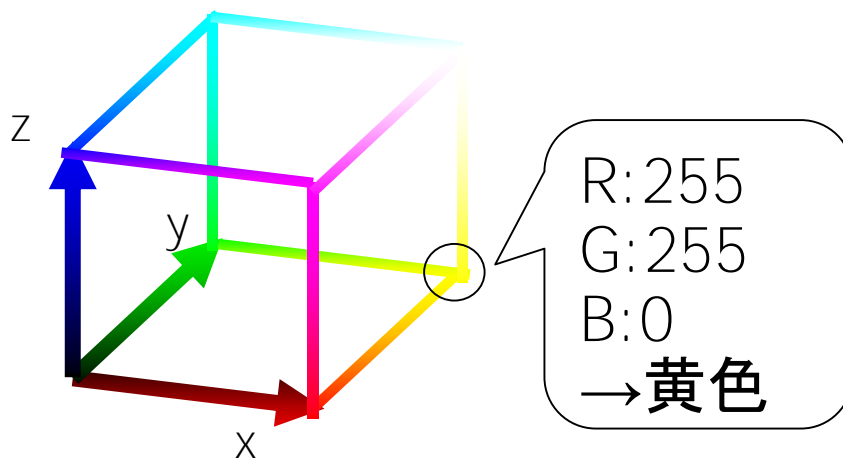
2.5 色空間・カラーマネジメント

(2) 色空間(カラースペース)

・「三次元で記述された色の座標空間」

例: RGB色空間

x軸を赤色、y軸を緑色、z軸を青色に割り当て
RGBの組み合わせで色を表現する

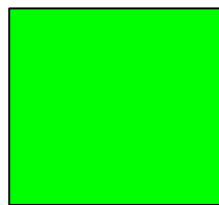


2.5 色空間・カラーマネジメント

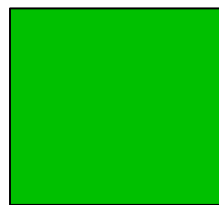
「色域」=色空間の中で表現できる色の範囲
→色域が広い方が資料の色合いを正確に再現できる

- ただし、素の状態では機器ごとに色域が異なるため、同じRGB値でも違う色に見える

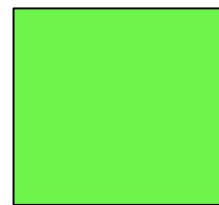
イメージ : $(R, G, B) = (0, 255, 0)$ (緑色)



機器A



機器B



機器C

2.5 色空間・カラーマネジメント

- ・色域を定義した色空間にあわせればどの機器でも同じように色が見える

①sRGB(standard RGB)

色域がAdobeRGBに比べて狭い

国際標準規格として多くの機器が対応

②AdobeRGB

色域がsRGBに比べて広い(特に緑、青)

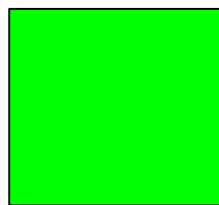
sRGBに比べると対応ディスプレイ・機器が少ない

- ・NDLでは①sRGBを採用

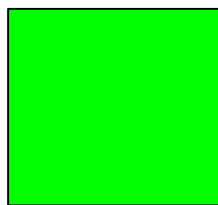
2.5 色空間・カラーマネジメント

- ・「カラープロファイル」
=どの色空間で画像を作製したか、という情報
スキャニング時に画像ファイルに付加できる
- ・sRGBのカラープロファイルを持つ画像を
sRGB用に設定したディスプレイで表示する
→どのディスプレイで見ても同じ色に見える

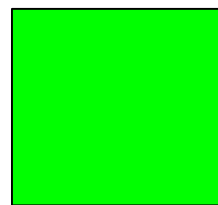
イメージ : $(R, G, B) = (0, 255, 0)$ (緑色)



機器A



機器B



機器C

2.5 色空間・カラーマネジメント

(3) カラーマネジメント

- ・スキャナの性能、外光条件等により色は変わる
→原本の色を忠実に再現するためには、機器の色表現を定期的に調整する必要がある

①定性的な調整

見た目の感覚で調整する

②定量的な調整

色に関する数値(色差)を測定して調整する

→NDLでは②定量的な調整(色差)を採用

2.5 色空間・カラーマネジメント

(4) 定量的な調整のイメージ

- ・複数の色を並べたカラーチャートをスキャンする
- スキャンした画像の各色の数値を測定する
- 公表されている各色の基準値と比較する

コダックカラーセパレーションガイド&グレースケール

X-rite カラーチェッカークラシック

例：X-rite Classicの基準値

http://xritephoto.com/documents/literature/en/ColorData-1p_EN.pdf

2.6 スキャナの種類

(1) スキャナの種類別の特徴

種類	特徴
①フラットベッド スキャナ	<ul style="list-style-type: none">・一般的なスキャナ(業務用～市販用)・下向きに資料を押しつけて撮影(劣化資料は不向き)・解像度が高い
②オーバーヘッド スキャナ	<ul style="list-style-type: none">・資料を上向きに置き、上から撮影・非接触型のため、資料を傷めない
③デジタルカメラ	<ul style="list-style-type: none">・資料形態による制約が少ない (立体物、厚みのある資料、大判の資料も撮影できる)・RAW現像処理が必要
④シートフィーダ型 スキャナ	<ul style="list-style-type: none">・解体した冊子等を連続スキャンできる・劣化資料、貴重資料には向かない
⑤フィルムスキャナ	<ul style="list-style-type: none">・マイクロフィルム等からのスキャンに用いる・フラットベッドスキャナよりも高解像度でスキャンできる

2.6 スキャナの種類

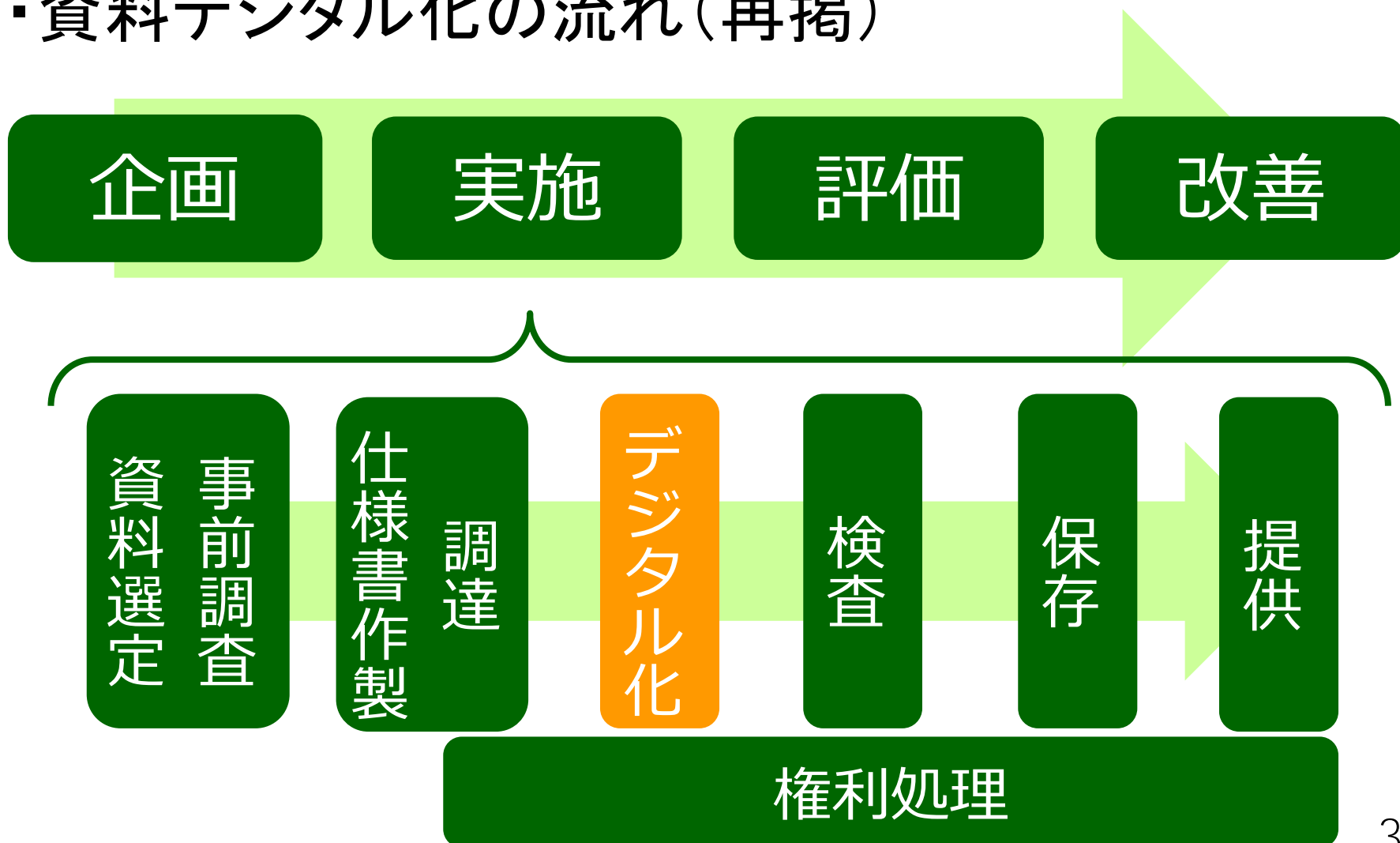
(2) スキャナ選定時のポイント

- 画像仕様や資料の特性に合致するものを選ぶ
→ 光学解像度、カラーマネジメント、
資料のサイズ・劣化状況 等
- 故障の際の修理体制・代替機の準備状況、
進捗遅延時の追加可否も確認しておくとい

3. デジタル化実務のポイント

- 3. 1 企画【39】
- 3. 2 資料選定・事前調査【40】
- 3. 3 仕様書作製・調達【42】
- 3. 4 デジタル化(1)：工程管理【47】
- 3. 5 デジタル化(2)：画像作製【49】
- 3. 6 デジタル化(3)：メタデータの作製【55】
- 3. 7 納品物検査【63】
- 3. 8 保存【70】
- 3. 9 (参考)提供に向けての検討事項【74】
- 3. 10 事業総括【76】

資料デジタル化の流れ(再掲)



3. 1 企画

- 事業目的(何のために)、対象資料(何を)、事業期間(いつまでに)、事業内容(どのように・どのくらい)等を明らかにする
 - 利用者ニーズ、資料価値・独自性、劣化状況、予算、事業化の難易度等を勘案して決定する
- NDL「資料デジタル化基本計画 2016-2020」
(http://www.ndl.go.jp/jp/aboutus/digitization/digitization_plan2016.pdf)
 - 多様な資料の大規模デジタル化を目指す

3. 2 資料選定・事前調査

- ・資料特性、分量を確認し、対象資料を選定する

(1) 外形特性

- ・判型、厚さ、折り込みページの量、製本状況、劣化状況、解体可否等
- ・デジタル化に向かない資料があれば除外する

劣化が激しい、形態が特殊 等

(2) 内容特性

- ・文字の大きさ・濃さ、図版・写真の量、目次ページの有無、分量、言語等

3. 2 資料選定・事前調査

(3) 作業量

- ・想定撮影コマ数、目次項目数・文字数等

(4) 作業環境(自館で作業する場合)

- ・書庫との間の動線、外光・温湿度コントロール、
機材設置用のスペース・電源等

→単価、作業工程の組み方等に影響するため
サンプル調査等を通じてできるだけ正確に
算定するのが望ましい

3.3 仕様書作製・調達

(1) 仕様書作製時の留意点

- ・作業内容を過不足なく、明瞭に記述する

→仕様書外の作業はできない

解釈にぶれが生じない書きぶりにする

- ・過度に複雑な内容にならないよう留意する

→性格の異なる資料群を一つの案件にまとめると

仕様が複雑になり、単価や工程管理コストが増加するおそれがある

3.3 仕様書作製・調達

(2) 仕様書作製にあたっての参考情報

・各種ガイドライン類

①「国立国会図書館資料デジタル化の手引」

仕様書サンプルを掲載

2017年版: 古典籍資料(冊子体)のデジタル化

2011年版: 戦前期刊行図書(冊子体)のデジタル化

②公益社団法人日本文書情報マネジメント協会「小規模団体での簡易版デジタルアーカイブガイドライン」

2018.1

(https://www.jiima.or.jp/pdf/digitalachive_guideline_shokibodantai.pdf)

3.3 仕様書作製・調達

- ③ IFLA「貴重書及び手稿コレクションのデジタル化計画のガイドライン」2014.9(NDL仮訳)

(http://www.ndl.go.jp/jp/aboutus/preservation/pdf/ifla_guideline_jp_2017.pdf)

・他館の先行事例・動向

- ①「公共図書館におけるデジタルアーカイブ事業の優良事例調査」2011.5

(<http://www.ndl.go.jp/jp/aboutus/kokyo.html#a01>)

- ②カレントアウェアネス・ポータル

(<http://current.ndl.go.jp/>)

→「デジタル化」「デジタルアーカイブ」等で検索

3.3 仕様書作製・調達

- ・できる範囲で試してみる

個人的に作ってみる、サービスを使ってみる、
小規模に試行してみる

- ・「チェックシート」を作製して考え方を整理する
→ 原本の貴重性・破損等のレベルはどうか？
対象資料の文字の大きさは？
資料等を運ぶ動線は確保されているか？ 等

3.3 仕様書作製・調達

(3) 調達時の留意点(外部委託)

- ・仕様書の内容について

応札者・受託者と認識を一致させておく

例:仕様書の内容に対する質疑応答

提案書の審査

受託後サンプルの確認

画質設定(明暗調整、色差等)、撮影方法、
撮影品質(傾き、トリミング、文字の鮮鋭度等)の
意識をあわせておく

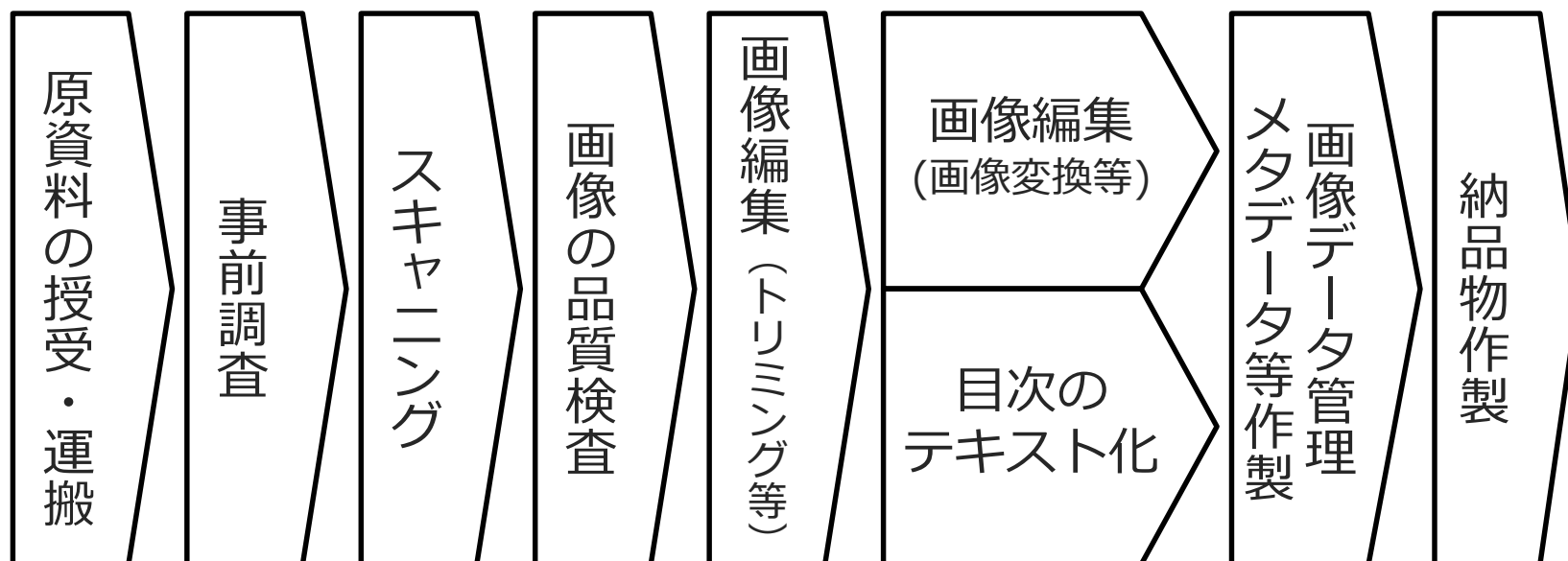
3.4 デジタル化(1) : 工程管理

(1) デジタル化の作業工程

- ・画像データ作製(スキャニング)以外にも
様々な工程がある

→ 工程管理(プロジェクトマネジメント)が重要

☞ 「5 デジタル化のプロジェクト管理」『手引』pp.55-67



3.4 デジタル化(1) : 工程管理

(2) NDLにおける工程管理の例

項目	内容
作業計画書確認・承認 (受託後10営業日以内)	作業計画書(スケジュール、実施手順等を記載) 品質検査計画書等を提出させ、内容を確認する
キックオフミーティング (作業開始前)	コミュニケーションルール、使用する書式、資料搬出の具体的な日時・手順等を確認・調整する
疑義照会対応 (随時)	スキャニング方法、目次データの採録方法、資料の解体可否等の疑義照会に回答する
週次報告 (週1回)	書面にて前週の作業実績、課題等を報告させる 進捗遅れ等があれば対案提示を求める
受託者作業現場視察 (1回)	資料の保管状況・取扱い、セキュリティ対策等 書面確認が難しい点も含めて現場を確認する
進捗状況報告会 (月1回)	対面で1月分の作業実績確認、リスク・課題管理を行う

3.5 デジタル化(2)：画像作製

(1) 原資料の授受・運搬(外部搬出する場合)

- デジタル化対象資料リストを作成する
- 資料に貼付したバーコード等の資料個体固有の情報とリストを突合し、必要な資料のみ抜き出す
- 運搬時や作業場所での保管時に
水濡れ・急激な温湿度変化がないよう注意する

保管庫、作業場所の環境は、書面では確認しづらいため、
作業現場視察が有効

3.5 デジタル化(2)：画像作製

(2) 事前調査

- ・資料の判型、劣化状態、製本状態
折り込み図や付属資料の分量等、
作業時に注意が必要な点を確認する
→疑問点がある場合は疑義照会へ

「製本が分厚く、資料の開きが悪いため、ノド元付近の文字が見切れてしまう。解体してよいか？」

「アンケート用紙が挟まっているがスキャン対象か？」等

3.5 デジタル化(2)：画像作製

(3) スキャニング

・資料サイズ、状態に応じて撮影方法を定めておく

①本文

見開き撮影か、片ページ撮影か

②折り込み図

折りたたんだ状態、裏面の撮影要否 等

③大判図

一回で撮影するか、分割撮影するか
分割する場合はどういう順番で撮影するか

3.5 デジタル化(2)：画像作製

(3) スキャニング

④ 挟み込み物

正誤表、別紙等が挟まっている場合の対応を決めておく(コマの順番、撮影方法等)

⑤ 写し込み物

カラーチャート、巻尺、名票等

3.5 デジタル化(2)：画像作製

(3) スキャニング：マイクロフィルムの場合

- ・マイクロフィルムがある場合は、原資料保存の観点から、マイクロフィルムからのデジタル化を検討する
- ・マスターフィルム保護のため、複製フィルム（DD(Direct Duplicating)ネガ）を作製し、それをデジタル化するのが一般的
- ・リールごとの濃度にばらつきがある場合は品質調整が難しい（白飛び、黒潰れ）

3.5 デジタル化(2)：画像作製

(4) 納品前の品質検査

- 納品前に、仕様書で要求する品質を満たしていることを確認させる
→ 納品後に発注者側でも検査する(3.7参照)
- NDLでは、品質検査項目、検査方法を規定した「品質検査計画書」の提出を求めている
→ 発注者、受託者双方の認識を共有しておく

3.6 デジタル化(3) : メタデータの作製

(1) 「メタデータ」の種類と整備方針

- ①書誌データ(タイトル、著者等)
 - ②原資料の個体データ(欠ページ、付属資料等)
 - ③提供のためのデータ(目次、本文テキスト等)
 - ④管理のためのデータ(資料ID、受託者名等)
 - ⑤権利に関するデータ(権利者情報等)
 - ⑥技術に関するデータ(撮影日時、機器等)
- 様々なメタデータのうち、何を整備・公開するのか？

3.6 デジタル化(3) : メタデータの作製

- ・データの見せ方に合わせて整備するメタデータ項目を検討する

例1) 自館ウェブサイトページを追加して見せる
書誌データ等、最低限の情報でもよい

例2) OPACの検索結果からリンクさせたい
原本の書誌データと、デジタルデータのメタデータを関連付ける

例3) 全文検索したい
本文のテキスト情報を整備する 等

3.6 デジタル化(3) : メタデータの作製

(2) メタデータ整備のポイント

①標準的なメタデータ項目を押さえる

例)「ダブリンコア」のメタデータ基本記述要素
(Simple DC)→15の基本語彙セット(次スライド)

②原本の書誌データ等、既存のデータを活用する

③最初はメタデータ項目を限定し、 公開後に情報を付加していくことも可能

ただし、デジタルデータの量が増えると
メタデータの修正・遡及整備が難しくなる可能性あり

3.6 デジタル化(3) : メタデータの作製

・ダブリンコアメタデータ基本記述要素 (Simple DC)

- Title (タイトル)
- Creator (作成者)
- Subject (キーワード)
- Description (内容記述)
- Publisher (公開者)
- Contributor (寄与者)
- Date (日付)
- Type (資源タイプ)
- Format (記録形式)
- Identifier (資源識別子)
- Source (出所)
- Language (言語)
- Relation (関係)
- Coverage (時空間範囲)
- Rights (権利管理)

3.6 デジタル化(3) : メタデータの作製

(3) NDL作製のメタデータ(4種類)

①(狭義の)メタデータ

- ・資料の提供、システムへの登録に必要な情報
→請求記号、書誌ID、タイトル、出版年等
- ・ダブリンコアをベースとしたDC-NDL(国立国会図書館ダブリンコアメタデータ記述)に準拠
(<http://www.ndl.go.jp/jp/aboutus/standards/meta.html>)
- ・「国立国会図書館サーチ」経由でNDL書誌から原本の書誌情報を補完

3. 6 デジタル化(3) : メタデータの作製

(3) NDL作製のメタデータ(4種類)

②目次データ

- 目次情報をテキスト化することで検索が可能に
→ 論題、著者名、掲載ページ、リンク先のコマの
画像ファイル名等
- 「雑誌記事索引」のデータを活用

3. 6 デジタル化(3) : メタデータの作製

(3) NDL作製のメタデータ(4種類)

③管理データ[1]

- ・個々の資料を管理するための情報

→書誌ID、資料貼付ID(バーコードラベル情報)

目次の有無、備考(欠ページ情報等)等

④管理データ[2]

- ・案件全体に関する情報

→受託者名、使用機器の製品名等

3.6 デジタル化(3) : メタデータの作製

(4) メタデータ作製時の注意点

①簡便な入力手順を目指す

外注業者は必ずしも
「図書館の専門家」ではない

- ・既存書誌データや「電子化対象リスト」の活用
- ・原資料を見て入力する項目については、情報源と判断基準を明確に定義しておく

②OCR(光学文字認識)によるテキスト化

- ・文字潰れ・薄れのある劣化資料等では読み取り精度が落ちる
- 人手による修正方針を定めておく

3.7 納品物検査

(1) 検査方法

- 案件の内容、各組織のリソース等に応じて目視検査とツール検査を組み合わせる

3.7 納品物検査

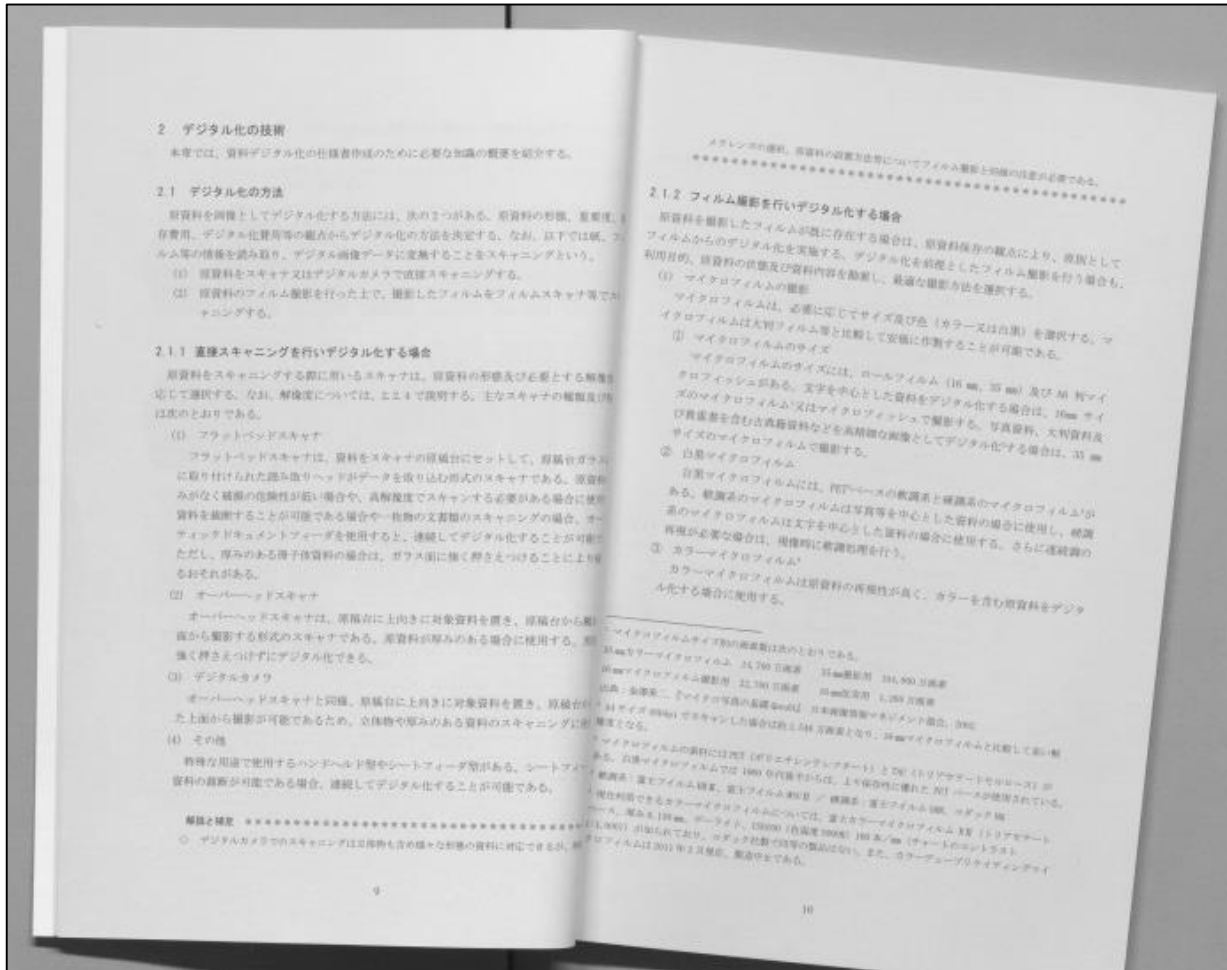
(2) 目視検査の項目例

- ・撮影漏れ・コマの重複
- ・ノド元の判読性
- ・傾き・トリミング
- ・カラーチャートの位置
- ・裏写り・間紙の使い方
- ・ピンボケ
- ・ゴミ・汚れ
- ・モアレ、ジャギー、偽輪郭(後ほど説明) etc...

☞「4.2.1.2 原資料の画像の目視検査」『手引』p.47

3.7 納品物検査

・問題画像の例①

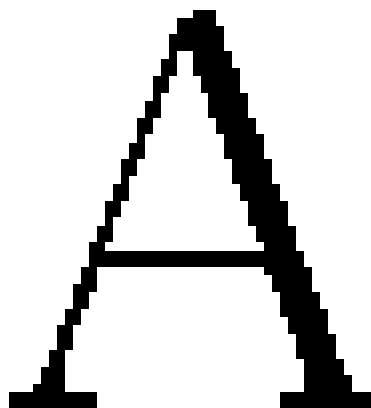


3.7 納品物検査

・問題画像の例②



モアレ



ジャギー



偽輪郭(空の部分)

→検査担当者の中で、合否の判断基準となる
サンプル(限界見本)を共有しておくとい

3.7 納品物検査

(3) ツール検査

- 大量に検査する場合、目視検査では限界がある
→ 検査ツール、プログラムの利用も検討する

例①: Jpylyzer (<http://jpylyzer.openpreservation.org/>)

- JPEG2000のメタデータチェックツール
→ JPEG2000の解像度、圧縮率等を検査できる

3.7 納品物検査

(3) ツール検査

例②: 独自に開発する

- NDLではPython, Perl等のプログラム言語を用いて検査ツールを作製している
- ファイルの格納構造、撮影漏れ、データ間の整合性、誤字、文字コード等、様々な検査に活用

→ ノウハウの引継ぎが重要

現在はOJTで維持・改良

組織としての人材育成も必要

3.7 納品物検査

(4) 瑕疵対応規定の準備

- 検査時にすべての問題を把握できるとは限らない
(履行期限後に問題が発覚する可能性あり)
 - 再度資料を搬出し、デジタル化するのは高コスト
- 仕様書に瑕疵対応規定を設けるとよい
(NDLの場合は3年間設定)

ただし、瑕疵対応のためには受託者側で作業環境の維持が必要になるため、調達コストに跳ね返る
(納品物のバックアップ保管、作業再開のための人員・環境準備等)
→ 納品物の分量が少なく、無理なく十分な検査ができる場合は、瑕疵対応期間を短縮(1年間等)するとよい

3.8 保存

(1) 保存方針の検討

・何を、いつまで、どう保存するのか検討しておく

①提供用画像

②保存用画像

③デジタル化した原本

→再度デジタル化することが難しい資料や、
高品質画像の使用想定（印刷物への使用等）が
ある場合は、②保存用画像の長期保存が必要

3.8 保存

(2) 長期保存上の課題

- ①障害・災害等による保存媒体・システムの損傷
→バックアップの作製、分散管理を検討
- ②媒体・ファイルフォーマット・再生機器の旧式化
→広く普及している技術を用いるのが安全
- ③保存媒体の劣化
→定期的に検査し、読み出し不能となる前に新しい媒体に変換(マイグレーション)する

3.8 保存

▪ JIS Z 6017「電子化文書の長期保存方法」

☞ 日本工業標準調査会ウェブサイトで検索・閲覧可能

(<http://www.jisc.go.jp/app/jis/general/GnrJISSearch.html>)

→ 記録媒体のハード・利用システム、
見読性の維持、媒体移行手順、廃棄等を規定

6.2 長期保存用途に使用する記録媒体及びドライブ装置
高品位な書き込みドライブ、媒体の使用などを規定

6.4 初期品質検査

作製されたBD-R等の品質評価の基準値を規定

→NDLでは仕様書に反映済み

3.8 保存

(3) NDLにおける保存

保存対象	保存方法
①提供用画像	「NDLデジタルデポジットシステム」のサーバ内に保存 近年は圧縮率1/50程度、JPEG2000非可逆圧縮
②保存用画像	BD-R(ブルーレイディスク)に記録し 温湿度管理された書庫内で保存(約17万枚) JPEG2000可逆圧縮
③デジタル化した原本	引き続き温湿度管理された書庫内で保存 (利用提供は停止する)

- ・現時点ではマイグレーションは未着手

3.9 (参考)提供に向けての検討事項

(1) 提供方法を考える

- ・見せ方や予算に応じて選択する

- a) 各館のウェブサイトにアップロードする
→経済的だが、量が多いと作業負荷が高まる
- b) 既存のウェブサービス・パッケージソフトの活用
→導入・運用コストが発生する
突然サービスが終了する可能性もあり
- c) 独自にシステムを開発
→自由度は高いが、導入・運用コストが膨大

3.9 (参考) 提供に向けての検討事項

(2) 問合せ対応の体制を整備しておく

- ・デジタル化によって資料にアクセスしやすくなる
→ 問合せの可能性も増える
- ・問合せの受理・回付・回答等の体制整備が必要

例「○ページの画像がない。撮影漏れか？ 欠ページか？」

- ① スキャニングはしたが登録に失敗していた
 - ② 原資料には存在していたがスキャニング漏れだった
 - ③ 元々欠ページだった
- 保存用画像、原資料等の確認が必要

3.10 事業総括

- 事業完了後に、事業実績や発生した問題点、今後の課題を総括する
- 必要に応じて、次回の仕様書等に反映する

おわりに

おわりに

- 「スキヤニングだけすれば終わり」ではない
企画から提供まで、体制づくりが重要
- 自組織の目的にあう形で仕様を検討する
各種ガイドラインは参考になるが万能ではない
- まずは無理のない範囲で試行する
実施した結果を次回の事業に活かす

参考文献(図書)

- 山田宏尚著『はじめてのデジタル画像処理：図解でわかる：画像処理技術を基礎から体系的に学べる』増補改訂版，技術評論社，2018.3.
- 日本文書情報マネジメント協会文書情報管理士検定試験委員会編『文書情報マネジメント概論：文書情報管理士指定参考書』日本文書情報マネジメント協会，2017.10.
- 水嶋英治，谷口知司，逸村裕編著『デジタルアーカイブの資料基盤と開発技法：記録遺産学への視点』晃洋書房，2016.4.

参考文献(ウェブ情報源)

- ・国立国会図書館関西館電子図書館課編「国立国会図書館資料デジタル化の手引」2017年版, 2017.4
(<http://www.ndl.go.jp/jp/aboutus/digitization/guide.html>)
- ・公益社団法人日本文書情報マネジメント協会「小規模団体での簡易版デジタルアーカイブガイドライン」
2018.1
(https://www.jiima.or.jp/pdf/digitalachive_guideline_shokibodantai.pdf)
- ・IFLA「貴重書及び手稿コレクションのデジタル化計画のガイドライン」2014.9(NDL仮訳)
(http://www.ndl.go.jp/jp/aboutus/preservation/pdf/ifla_guideline_jp_2017.pdf)

参考文献(ウェブ情報源)

- ・「公共図書館におけるデジタルアーカイブ事業の優良事例調査」2011.5
(<http://www.ndl.go.jp/jp/aboutus/kokyo.html#a01>)
- ・総務省「震災関連デジタルアーカイブ構築・運用のためのガイドライン 第4章 デジタルデータの長期保存・利用について」2013.3
(http://www.soumu.go.jp/main_content/000225130.pdf)
- ・総務省「デジタルアーカイブの構築・連携のためのガイドライン」2012.3.26
(http://www.soumu.go.jp/main_content/000153595.pdf)