

公文書館等における
デジタルアーカイブ・システムの標準仕様書

独立行政法人国立公文書館

<改訂履歴>

平成21年3月	作成
平成24年10月	一部改訂
平成30年3月	改訂

目次

I 概要

1 デジタルアーカイブ・システム標準仕様一覧	1
2 システム概要図	5
3 システム構成例	6
3-1 基本的な構成	6
3-2 標準的な構成	8
3-3 発展的な構成	10

II デジタルアーカイブ・システム標準仕様

1 はじめに	12
1-1 デジタルアーカイブ・システムの導入の背景と目的	12
1-2 デジタルアーカイブ・システムの標準仕様の考え方	13
1-3 システム概要図	15
2 デジタルアーカイブ・システム標準仕様	16
2-1 目録データ・デジタルコンテンツの準備	16
2-1-1 目録データ	16
(1) 記述方式、電子符号化方式	16
(2) 横断検索のために必要な最小限の目録データの項目	17
(3) 階層構造の表現	20
(4) 文字コード	20
(5) 外字の扱い	21
2-1-2 デジタルコンテンツ	22
(1) 画像データ	22
(2) 動画・音声データ	23
2-2 目録データ・デジタルコンテンツの登録と管理	25
2-2-1 目録データ	25
(1) 目録データの登録・管理機能	25
2-2-2 デジタルコンテンツ	27
(1) デジタルコンテンツの登録・管理機能	27
2-3 利用者向け機能	29

2-3-1	検索機能	29
(1)	検索機能	29
(2)	検索支援機能	30
2-3-2	デジタルコンテンツ閲覧機能	32
(1)	デジタルコンテンツ閲覧機能	32
(2)	画像配信機能	34
(3)	動画・音声配信機能	34
2-3-3	ユーザ・インターフェース	35
(1)	アクセシビリティ	35
(2)	スマートデバイスのためのユーザ・インターフェース	35
2-4	利活用の促進	36
2-4-1	横断検索機能	36
(1)	横断検索プロトコル	36
2-4-2	一括提供機能	36
2-4-3	外部提供インターフェース	36
2-4-4	永続的アクセス	37
2-4-5	ライセンスの表示	38
2-5	利用者支援	39
2-5-1	システムの利用案内	39
2-5-2	レファレンス機能	39
2-6	業務向け機能	40
2-6-1	利用状況の把握機能	40
2-6-2	閲覧業務関連機能	40
2-7	システム環境	41
2-7-1	システム環境検討のポイント	41
(1)	運用形態の検討	41
(2)	構成の検討	41
(3)	システム規模の検討	44
(4)	インフラ基盤の検討	44
2-8	セキュリティ	46
2-8-1	ウイルス対策	46
2-8-2	外部攻撃への対策	46
2-8-3	不正アクセス対策	47
2-9	システム運用	48
2-9-1	運用・保守	48
2-9-2	バックアップ機能	48
2-9-3	システム監視機能	48
2-9-4	データ移行	49
2-10	性能要件	50

III システム構成例

1 基本的な構成	51
1-1 対象データ	51
1-2 機能構成	52
1-3 仕様書記載事項の例	53
2 標準的な構成	59
2-1 対象データ	59
2-2 機能構成	60
2-3 仕様書記載事項の例	61
3 発展的な構成	68
3-1 対象データ	68
3-2 機能構成	69
3-3 仕様書記載事項の例	70

I 概要

1 デジタルアーカイブ・システム標準仕様一覧

表 I-1は、デジタルアーカイブ・システムの標準仕様を一覧として整理したものである。これら標準仕様の考え方については、「Ⅱ デジタルアーカイブ・システム標準仕様書」に取りまとめている。

表 I-1 デジタルアーカイブ・システム標準仕様一覧

【凡例】 ◎：標準項目 ○：推奨項目

大項目	中項目	小項目	標準・推奨	仕様
目録データ・デジタルコンテンツの準備	目録データ	記述方式、電子符号化方式	◎	目録データ項目は ISAD(G) に基づいて設定する。
			◎	電子符号化の方式は、再利用、データ変換、保守の容易さを考慮した方式とする。
		横断検索のために必要な最小限の目録データの項目	◎	Dublin Core の要素を参考に、横断検索のために必要な最小限の目録データの項目として、「ID」「年代（作成年度）」「資料名」「作成者（部署名）」「場所」「備考等」の 6 項目を設定する。
			◎	公文書館等の目録データ項目と横断検索のために必要な最小限の目録データ項目（6 項目）との対応表を定義する。
		階層構造の表現	◎	公文書等に関する階層構造は、公文書等の作成／移管元の組織や業務を反映したものとする。
		文字コード	◎	UCS-4 を表現できる文字コードとする。
			◎	符号化形式は UTF-8 または UTF-16 とする。
		外字の扱い	◎	外字は指定せず、代替可能な UCS 規格内の文字（代替文字）を用いる。
○	外字を用いる場合は、「文字図形番号」（情報処理学会試行標準 IPSJ—TS0002：2014）を用い、EGIX で記述する。			

大項目	中項目	小項目	標準・推奨	仕様
	デジタルコンテンツ	画像データ	◎	保存用の画像形式として TIFF（非圧縮）、公開用の画像形式として JPEG または PDF をサポートする。
			○	保存用と公開用を兼ねた画像形式として JPEG 2000 をサポートする。
			○	公開用画像は、画像サイズを抑えつつも劣化が視認できないレベルの画質に保つ。
		動画・音声データ	○	公開用の動画データの形式として、MP4 をサポートする。
			○	公開用の音声データの形式として、MP3 をサポートする。
目録データ・デジタルコンテンツの登録と管理	目録データ	目録データの登録・管理機能	◎	目録データ等の電子符号化方式に合わせて、適切なデータベースや検索エンジンを採用する。
			◎	目録データを個別及び一括で登録、編集、削除、出力する機能を設ける。
			◎	非公開の目録データと一般公開する目録データを分離し、アクセス制御する。
	デジタルコンテンツ	デジタルコンテンツの登録・管理機能	◎	デジタルコンテンツとその管理データを個別及び一括で登録、編集、削除する機能を設ける。
◎			非公開のデジタルコンテンツと一般公開するデジタルコンテンツを分離し、アクセス制御する。	
利用者向け機能	検索機能	検索機能	◎	任意の項目を対象とした、キーワード及び条件指定による検索を可能とする。
			◎	項目を限定せずに、キーワードによる全文検索を可能とする。
			◎	複数の条件及びキーワードの組み合わせによる検索を可能とする。
			◎	目録データの階層構造表示から、指定した階層に属する目録データの一覧表示を可能とする。
			◎	目録データの階層構造表示から、指定した階層以下に属する目録データに限定した検索を可能とする。
		検索支援機能	○	辞書検索機能、ファセット検索機能、異体字検索機能等の検索支援機能を設ける。

大項目	中項目	小項目	標準・推奨	仕様
	デジタルコンテンツ 閲覧機能	デジタルコンテンツ閲覧機能	◎	目録データに対応するデジタルコンテンツが存在する場合には、目録データの検索結果一覧及び詳細表示にデジタルコンテンツへのリンクを表示し、当該リンクからデジタルコンテンツを閲覧可能とする。
			◎	画像データの閲覧方式は、利用者が特別なソフトウェアを必要としない方式とする。
			◎	画像ビューアは、直観的な操作で拡大、縮小、移動、ページ送り等が行えるものとする。
			○	デジタルコンテンツのサムネイル画像を表示する。
			○	デジタルコンテンツを個別及び一括でダウンロードする機能を設ける。
		画像配信機能	○	画像の相互運用性に考慮した配信方式（例：IIIF）を採用する。
		動画・音声配信機能	○	動画・音声の公開にはストリーミング配信方式を採用する。
	ユーザ・インターフェース	アクセシビリティ	○	アクセシビリティを考慮したユーザ・インターフェースを提供する。
スマートデバイスのためのユーザ・インターフェース		○	スマートデバイス用のユーザ・インターフェースを提供する。	
利活用の促進	横断検索機能	◎	SRU/SRW または OpenSearch による検索に対応する。	
	一括提供機能	○	OAI-PMH による目録データの一括提供機能（リポジトリ機能）を実装する。	
	外部提供インターフェース	○	情報を外部提供するインターフェースとして Web API（SPARQL 等）を実装する。	
	永続的アクセス	○	目録データやデジタルコンテンツの URL のリンク切れを防止する。	
	ライセンスの表示	○	利用条件を明示するための枠組みを用いて、デジタルコンテンツ等の利用条件を明示する。	

大項目	中項目	小項目	標準・推奨	仕様
利用者支援	システムの利用案内		◎	デジタルアーカイブ・システムの利用案内（利用方法、動作環境等）を明示する。
			○	資料（原本）の閲覧方法を利用者が把握しやすいように案内を設ける。
	レファレンス機能		○	利用者から目録データやデジタルコンテンツに関するレファレンスを受け付ける仕組みを設ける。
業務向け機能	利用状況の把握機能		◎	Web アクセス状況等の統計情報を作成し、利用状況を把握可能とする。
	閲覧業務関連機能		◎	閲覧請求書等の印刷機能を設ける。
システム環境			◎	システム導入は、業務の在り方やニーズ等に応じて、運用形態、機能と構成、データ量、インフラ基盤等を検討したうえで実施する。
セキュリティ	ウイルス対策		◎	ウイルス対策ソフトウェアを導入する。
	外部攻撃への対策		◎	ファイアウォールを導入する。
	不正アクセス対策		○	暗号化通信に対応する。
システム運用	運用・保守		◎	運用・保守等を適切に実施する。
			◎	定期的にセキュリティ更新プログラムを適用する。
	バックアップ機能		◎	バックアップ取得機能を導入する。
			○	遠隔地にバックアップデータを保管する。
	システム監視機能		○	利用者向けサービスの稼働状況を監視する。
データ移行		◎	既存システムからのデータ移行、次期システムへのデータ移行に留意する。	
性能要件			◎	目録データの検索、デジタルコンテンツの配信・閲覧に係る速度は、利用者が快適に利用できる水準とする。
			◎	目録データの登録、デジタルコンテンツの登録等に係る所要時間は、業務に支障をきたさない水準とする。

2 システム概要図

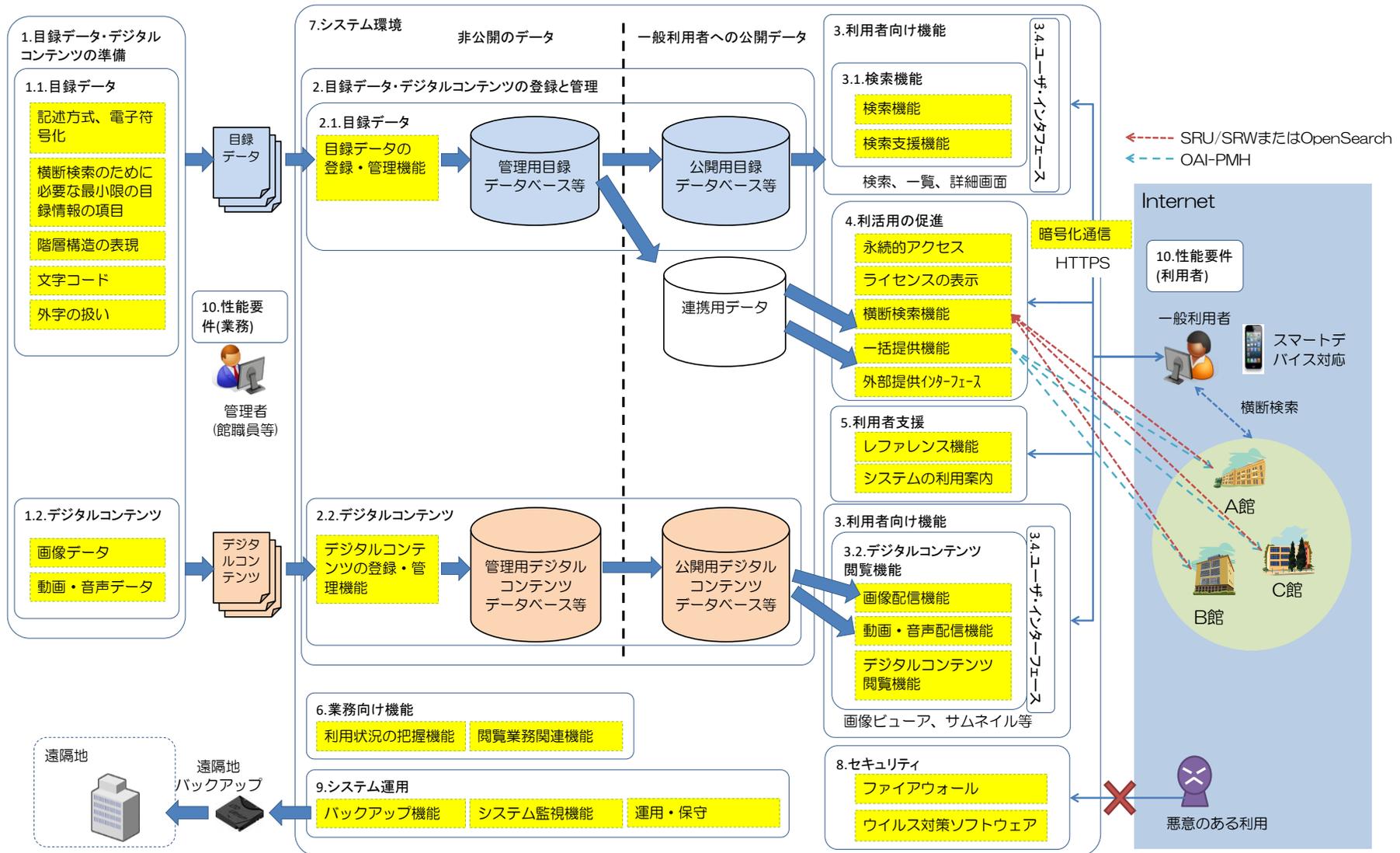
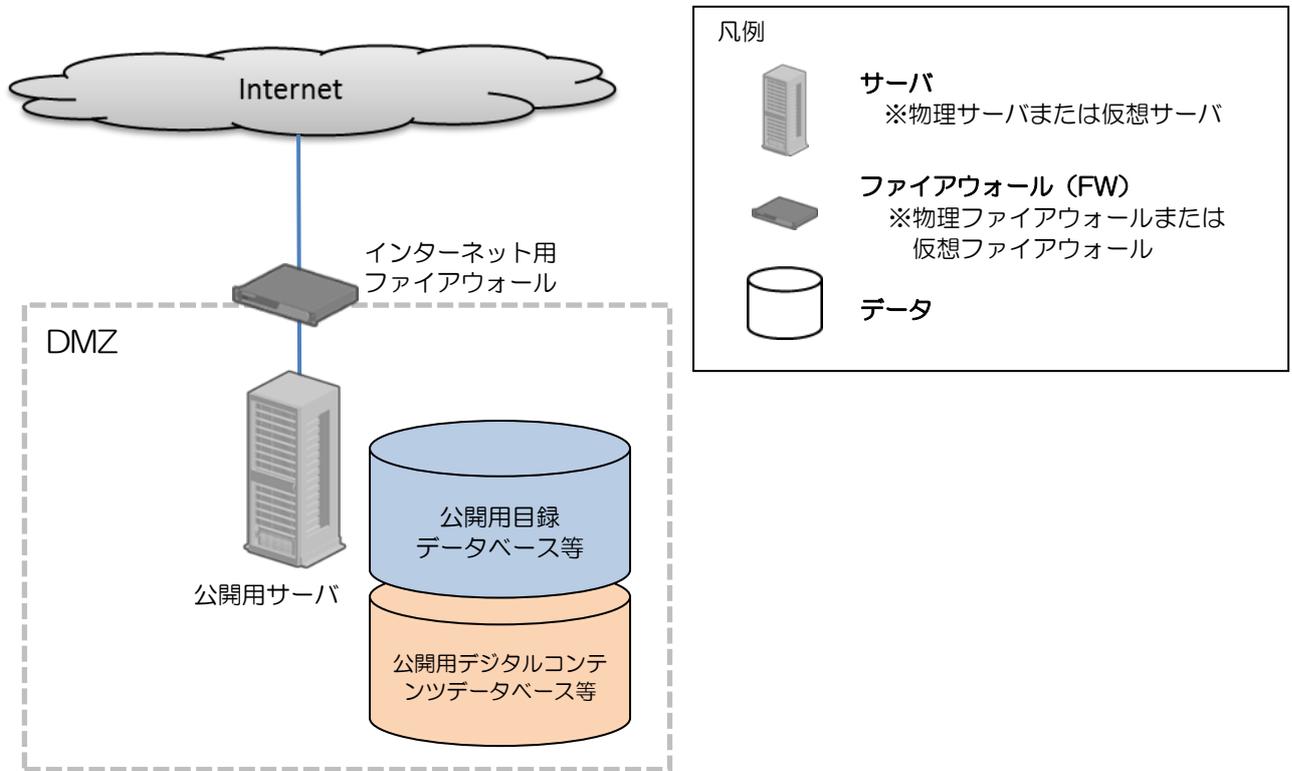


図 I-1 デジタルアーカイブ・システム概要図

3 システム構成例

3-1 基本的な構成

公開データを扱う上で最小限必要と考えられる機能からなる構成を「基本的な構成」とする。



DMZ:セキュリティ強化のため、インターネットと内部ネットワーク（館内ネットワーク）の間に設けられた、ファイアウォール等により隔離される領域。

図 I-2 基本的な構成におけるシステム構成図

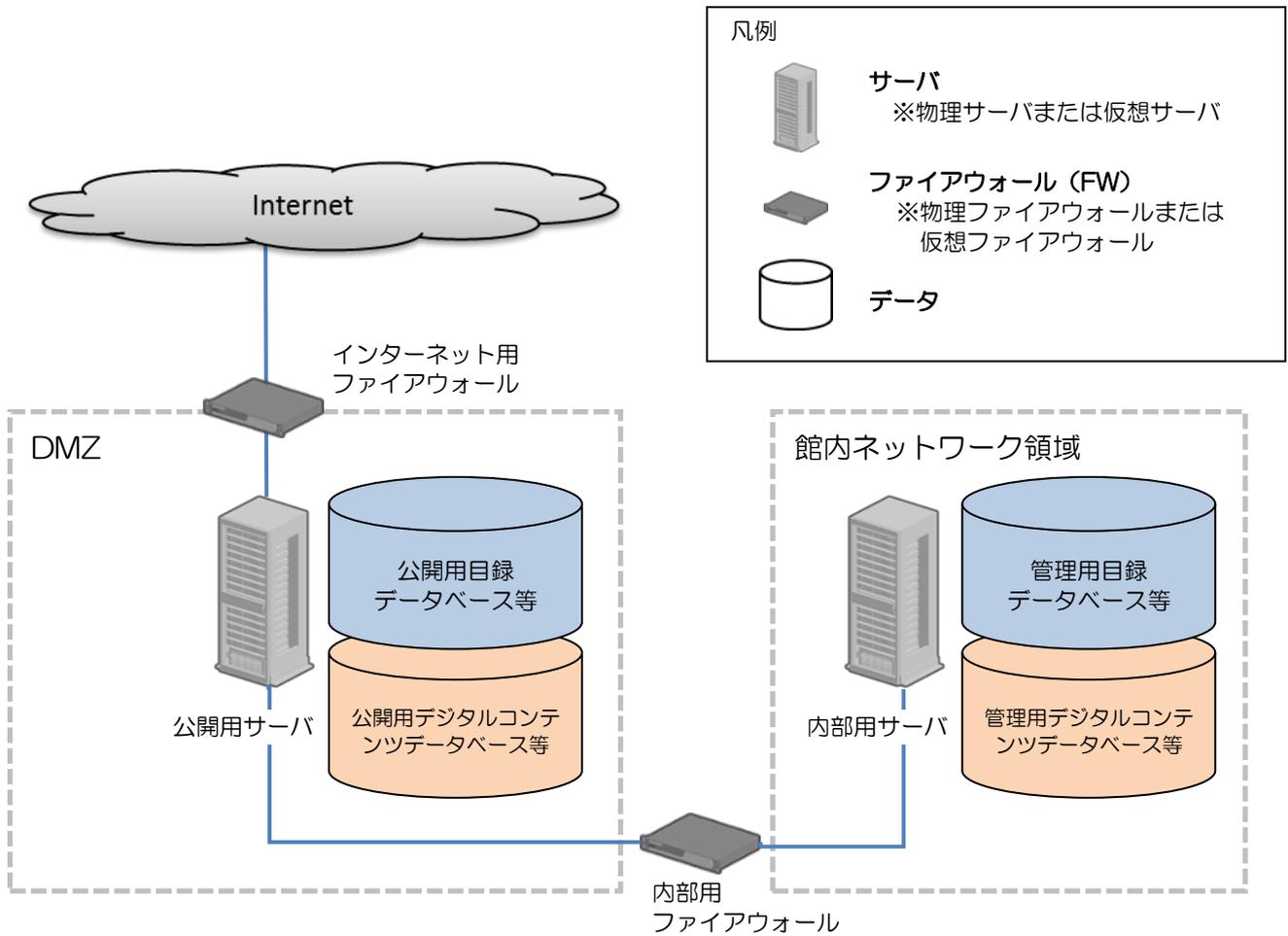
基本的な構成においては以下の機能を想定する。



図 I-3 基本的な構成における機能構成図

3-2 標準的な構成

公開データ及び非公開データの両方を扱う上で最小限必要と考えられる機能からなる構成を「標準的な構成」とする。



DMZ: セキュリティ強化のため、インターネットと内部ネットワーク（館内ネットワーク）の間に設けられた、ファイアウォール等により隔離される領域。

図 I-4 標準的な構成におけるシステム構成図

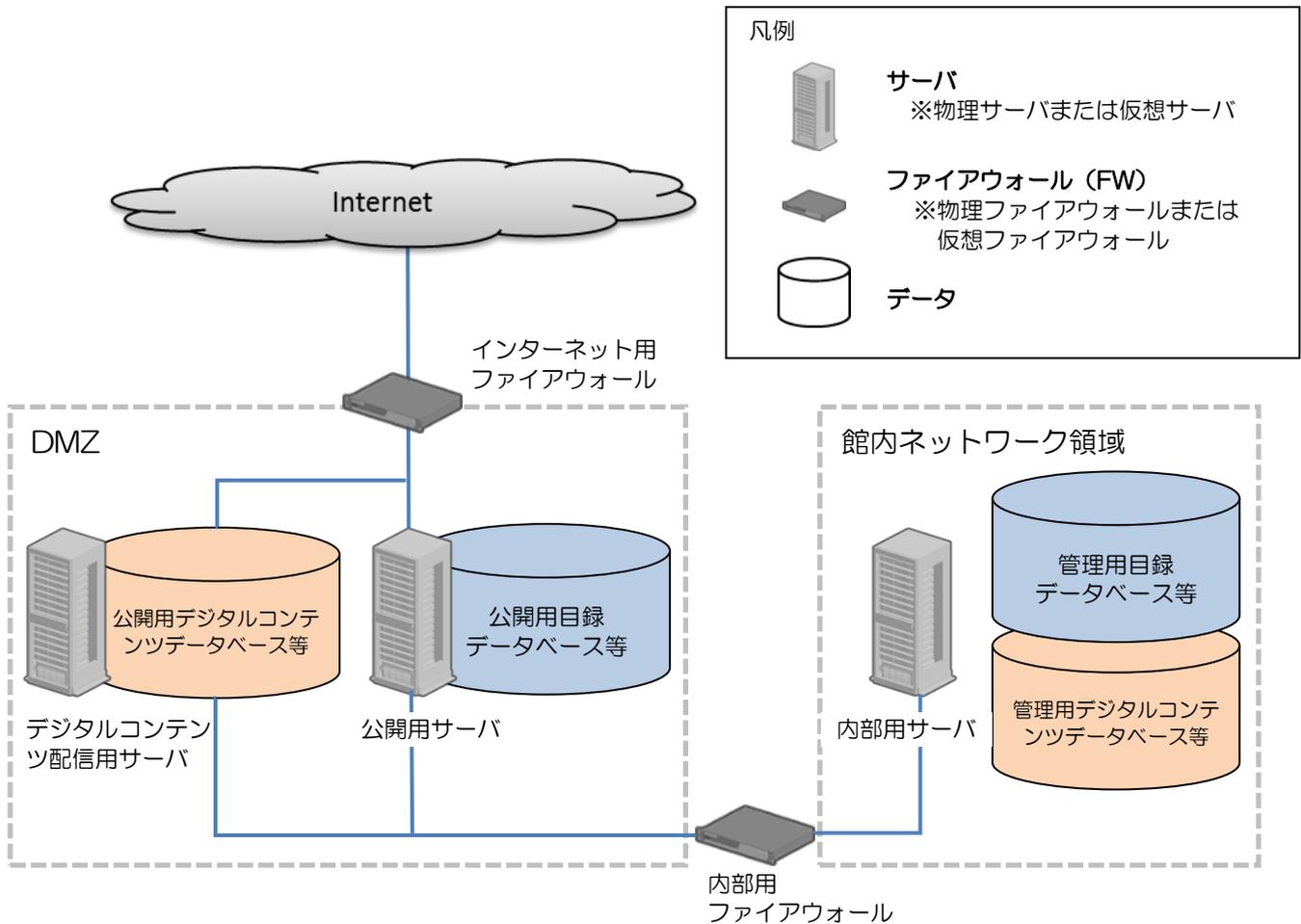
標準的な構成においては以下の機能を想定する。



図 I-5 標準的な構成における機能構成図

3-3 発展的な構成

「標準的な構成」に加えて、推奨機能及び付加的価値のある機能を含む構成を「発展的な構成」とする。



DMZ: セキュリティ強化のため、インターネットと内部ネットワーク（館内ネットワーク）の間に設けられた、ファイアウォール等により隔離される領域。

図 I-6 発展的な構成におけるシステム構成図

発展的な構成においては以下の機能を想定する。

凡例：下線のうち、実線は推奨機能、点線は推奨機能以外の付加的価値のある機能を示す。

<p>目録データの登録・管理機能</p>	<p>利用者向け機能</p>	<p>利活用の促進</p>
<ul style="list-style-type: none"> ● 目録データの登録・管理 ● 公開／非公開設定 ● <u>外字の扱い</u> 	<ul style="list-style-type: none"> ● 検索機能 ● <u>検索支援機能（辞書、ファセット、異体字）</u> ● デジタルコンテンツ閲覧 ● <u>サムネイル表示</u> ● <u>ダウンロード機能</u> ● <u>高機能画像配信機能（IIIF等）</u> ● <u>動画・音声配信機能</u> 	<ul style="list-style-type: none"> ● 横断検索（SRU/SRW、Open Search） ● <u>一括提供（OAI-PMH）</u> ● <u>外部提供インターフェース（Web API）</u> ● <u>永続的アクセス</u> ● <u>ライセンス表示</u> ● <u>SNS連携</u>
<p>デジタルコンテンツの登録・管理機能</p>		
<ul style="list-style-type: none"> ● デジタルコンテンツの登録・管理 ● 公開／非公開設定 		
<p>業務向け機能</p>	<p>利用者支援</p>	<p>ユーザ・インターフェース</p>
<ul style="list-style-type: none"> ● 利用状況の把握 ● 閲覧請求書等印刷 	<ul style="list-style-type: none"> ● システムの利用案内 ● <u>レファレンス機能</u> 	<ul style="list-style-type: none"> ● <u>アクセシビリティ</u> ● <u>スマートデバイス対応</u>
<p>セキュリティ</p>		
<ul style="list-style-type: none"> ● ウイルス対策 ● <u>不正アクセス対策（暗号化通信）</u> ● 外部攻撃への対策（ファイアウォール） ● <u>外部攻撃への対策（WAF、侵入検知）</u> 		
<p>システム運用</p>		
<ul style="list-style-type: none"> ● バックアップ機能 ● <u>遠隔地バックアップ</u> ● <u>システム監視機能</u> 		

図 I-7 発展的な構成における機能構成図

Ⅱ デジタルアーカイブ・システム標準仕様

1 はじめに

1-1 デジタルアーカイブ・システムの導入の背景と目的

全国の公文書館等（以下、「公文書館等」という。）は、国及び地域の歴史を伝える重要な公文書等を保存し、利用提供する責務を負う機関である。そのような公文書館等においては、知的資源としての公文書等が社会に有効に活用されるよう、情報通信技術の活用を通じて、所蔵する公文書等のデジタルアーカイブ化を推進することが求められる。

公文書館等にデジタルアーカイブ・システムが導入されることで、公文書等の目録データやデジタルコンテンツは、「いつでも」「どこでも」「誰でも」「自由に」利用することが可能となる。また公文書館等においては、デジタルアーカイブ・システムを導入し、横断検索連携等を実現させることで、情報発信力の強化や所蔵資料の利活用促進が可能となる。

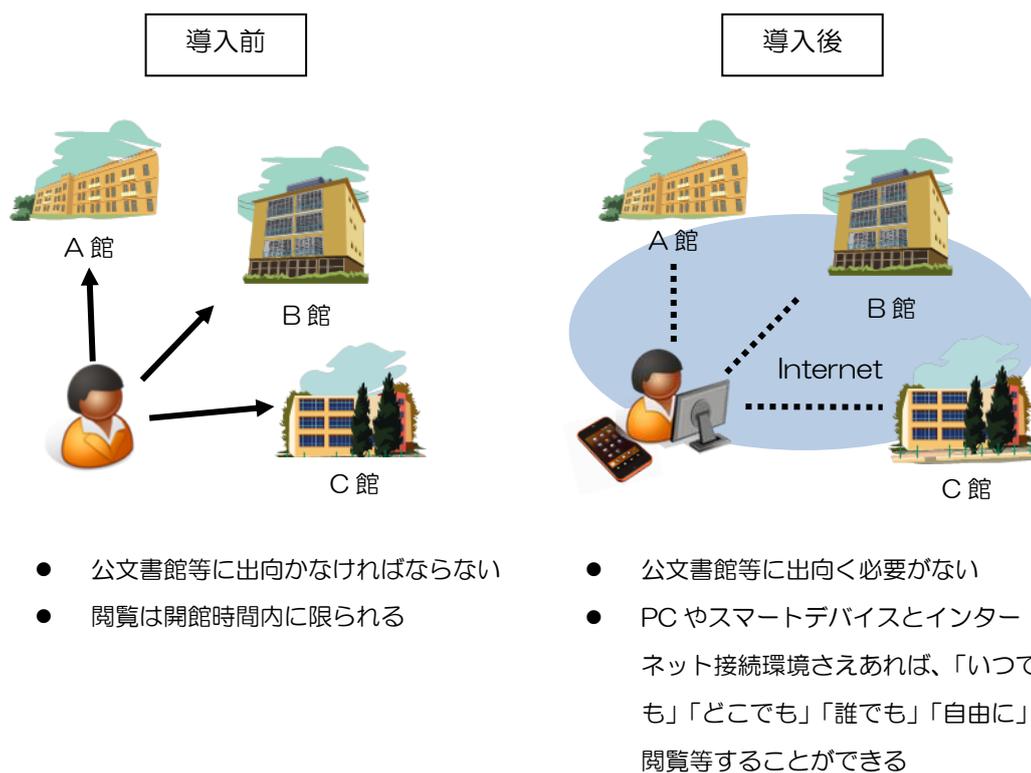


図 II-1 デジタルアーカイブ・システムの導入イメージ

1-2 デジタルアーカイブ・システムの標準仕様の考え方

本書は、公文書館等がデジタルアーカイブ・システムを構築するうえで必要となる技術的情報等を、標準仕様として整理し紹介することを目的としている。公文書館等では、デジタルアーカイブ・システムの構築にあたり本書を参照することで、システム導入に係る事前の調査・検討に要するコストの削減を図ることが可能となる。

本書では、以下の基本的な考え方に基づきながら、次の表に示す項目ごとに具体的な標準仕様等についての検討及び取りまとめを行っている。

<基本的な考え方>

① デジタルアーカイブ・システムの在り方

- (ア) 公文書館等において現実的に導入可能であり、また保守・運用が可能なものであること。
- (イ) 公文書館等の個別の状況に応じて、柔軟に機能や構成が選択可能であること。
- (ウ) 公文書館等において導入されるシステムとして適切な水準のセキュリティが確保されていること。

② 国際標準等の採用

デジタルアーカイブ・システムの標準仕様には、国際標準または事実上の標準の採用を基本とすること。

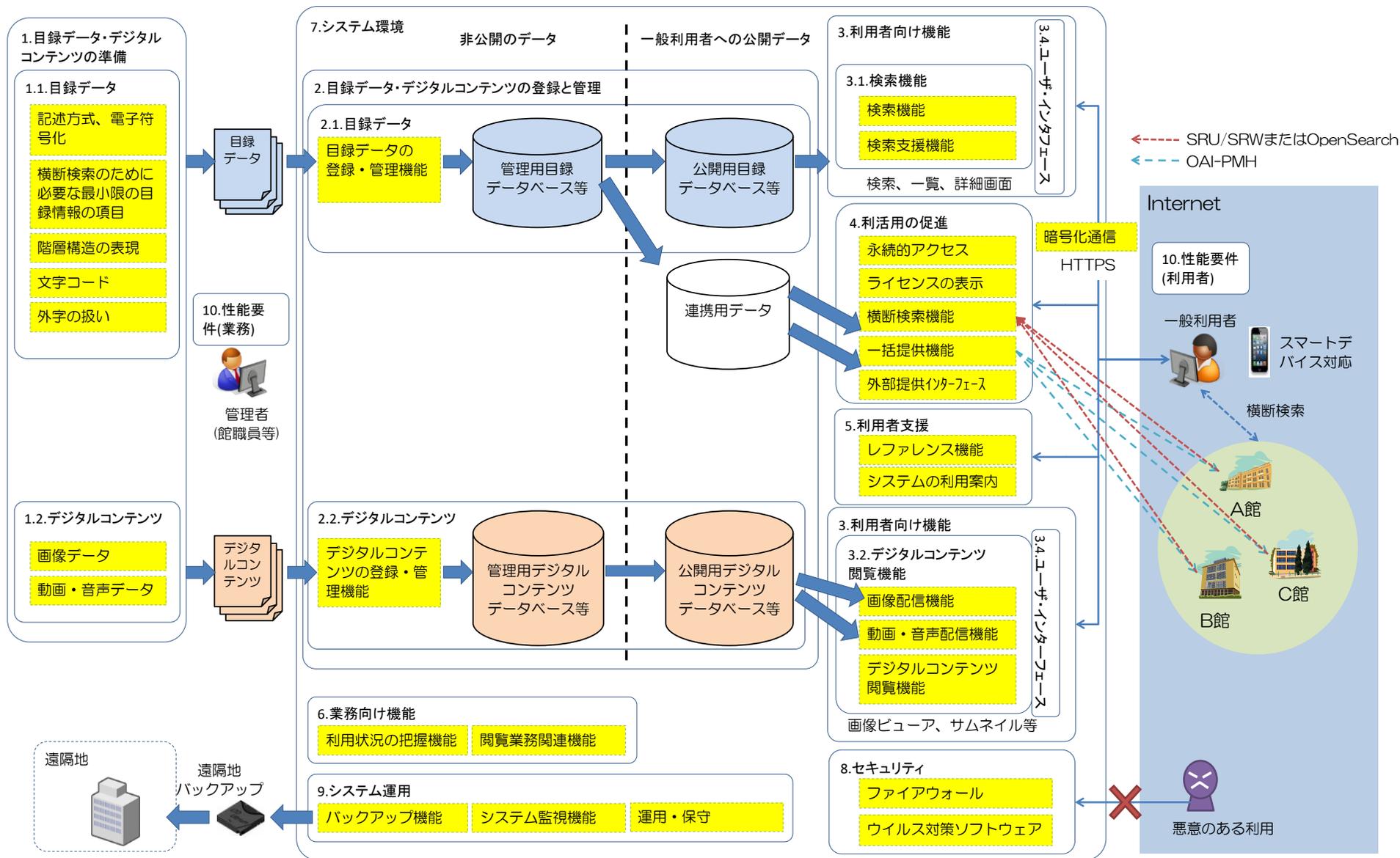
③ システム構成例

- (ア) 公開データを扱う上で最小限必要と考えられる機能からなるデジタルアーカイブ・システムの構成を「基本的な構成」とする。
- (イ) 公開データ及び非公開データの両方を扱う上で最小限必要と考えられる機能からなるデジタルアーカイブ・システムの構成を「標準的な構成」とする。
- (ウ) 「標準的な構成」に加えて、推奨機能及び付加的価値のある機能を含むデジタルアーカイブ・システムの構成を「発展的な構成」とする。

表 II-1 デジタルアーカイブ・システムの標準仕様項目

目録データ・デジタルコンテンツの準備	目録データ	記述方式、電子符号化方式
		横断検索のために必要な最小限の目録データの項目
		階層構造の表現
		文字コード
		外字の扱い
	デジタルコンテンツ	画像データ
		動画・音声データ
目録データ・デジタルコンテンツの登録と管理	目録データ	目録データの登録・管理機能
	デジタルコンテンツ	デジタルコンテンツの登録・管理機能
利用者向け機能	検索機能	検索機能
		検索支援機能
	デジタルコンテンツ閲覧機能	デジタルコンテンツ閲覧機能
		画像配信機能
		動画・音声配信機能
	ユーザ・インターフェース	アクセシビリティ
スマートデバイスのためのユーザ・インターフェース		
利活用の促進	横断検索機能	
	一括提供機能	
	外部提供インターフェース	
	永続的アクセス	
	ライセンスの表示	
利用者支援	システムの利用案内	
	レファレンス機能	
業務向け機能	利用状況の把握機能	
	閲覧業務関連機能	
システム環境		
セキュリティ	ウイルス対策	
	外部攻撃への対策	
	不正アクセス対策	
システム運用	運用・保守	
	バックアップ機能	
	システム監視機能	
	データ移行	
性能要件		

1-3 システム概要図



2 デジタルアーカイブ・システム標準仕様

2-1 目録データ・デジタルコンテンツの準備

2-1-1 目録データ

(1)記述方式、電子符号化方式

デジタルアーカイブ・システムでは、公文書や古文書等の目録情報を登録、管理、検索等できることが求められる。公文書館等が作成する目録情報の最小単位としては、簿冊単位または件名単位を考えることができ、デジタルアーカイブ・システムではこれら最小単位の目録情報を扱える必要がある。また、目録情報の記述方式及び電子符号化方式は、標準化されることで相互運用性を高めることができる。

公文書館等における目録情報の記述のための国際標準としては、国際公文書館会議（ICA：International Council on Archives）が制定したISAD(G)（General International Standard Archival Description）がある¹。ISAD(G)は、公文書館等の所蔵資料のための記述項目や記述に係る原則をまとめた標準であり、1994年の初版公開の後、2000年に第2版が公開されている。

また、目録情報の電子符号化方式としては、XML（Extensible Markup Language）形式によるEAD（Encoded Archival Description）があり²、ISAD(G)とのクロスワーク（対応表）も整備されている。EADは、公文書館等の目録情報の電子符号化方式における事実上の国際標準であり、目録情報を階層構造で表現することが可能となっている。EADの1.0版は1998年に公開され、2002年に1.0版からEAD2002へ改版された後、2015年8月にはEAD3(1.0)が正式に公開された。

デジタルアーカイブ・システムにおける目録データの項目は、公文書館等の所蔵資料の特性等を踏まえつつ、国際標準であるISAD(G)に基づいて設定することが適切と考えられる。また、目録情報の電子符号化方式については、再利用、データ変換、保守の容易さを考慮した方式とする必要がある。デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会による「デジタルアーカイブの構築・共有・活用ガイドライン」（平成29年4月）では、表形式のデータ作成時に機械可読性を担保するための留意事項がまとめられており、公文書館等における目録データ作成の参考となる³。なお、電子符号化方式としてXMLを採用する場合は、EADに準拠することが望ましい。

【標準仕様の考え方】

標準	目録データ項目はISAD(G)に基づいて設定する。
標準	電子符号化の方式は、再利用、データ変換、保守の容易さを考慮した方式とする。

¹ International Council on Archives. “ISAD(G): General International Standard Archival Description - Second edition”. 2011-09-01.

<https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>, (参照 2017-12-25).

² Library of Congress. “Encoded Archival Description Official Site”. 2017-12-13. <http://www.loc.gov/ead/>, (参照 2017-12-25).

³ デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会. “デジタルアーカイブの構築・共有・活用ガイドライン”. 内閣府知的財産戦略推進事務局. 2017, p.39.

https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/guideline.pdf, (参照 2017-12-25).

表 II-2 ISAD(G) 項目一覧

ISAD(G)エリア	エレメント	
1 識別要素	1.1	レファレンス・コード
	1.2	タイトル
	1.3	資料作成年月日
	1.4	記述レベル
	1.5	数量
2 コンテキスト	2.1	作成者名称
	2.2	組織歴または個人履歴
	2.3	伝来
	2.4	入手先
3 内容・構造	3.1	資料内容
	3.2	評価・廃棄・保存年限
	3.3	追加受入
	3.4	編成
4 公開・利用条件	4.1	法的位置づけ
	4.2	公開条件
	4.3	使用言語
	4.4	物的特徴
	4.5	検索手段
5 関連事項	5.1	オリジナル資料
	5.2	複製
	5.3	関連資料
	5.4	出版書誌情報
6 注記	6.1	注記
7 記述作成情報	7.1	記述注記
	7.2	記述規則
	7.3	記述作成日

(2) 横断検索のために必要な最小限の目録データの項目

現在、各公文書館等の目録データの横断的な検索は、一部の公文書館等と国立公文書館との間の実現に留まっている。しかしながら、公文書館等の所蔵資料の利活用が促進されるためには、横断検索連携の普及によるデジタルアーカイブ・システムのネットワーク化の推進が欠かせない。

目録データの詳細は各公文書館等により異なることが想定され、そのままでは横断検索において有効な結果を得ることが難しい。効果的な横断検索を実現するためには、各公文書館等の目録データのうち検索対象とする最小限の項目を共通化する必要がある。

検索対象項目の共通化にあたっては、国際標準のメタデータ記述フレームワークであるDublin Core (ISO15836) を参考にすることができる。Dublin Coreとは、Web上の情報資源のメタデータ記述のた

めに策定された、15の基本要素を核とする記述定義である。複数のメタデータ体系間の相互互換性を高めるためにも広く活用されている。

検索行動に関するある調査では、Dublin Coreの基本要素のうち、検索項目として主に用いられる項目は「Title」「Creator」「Subject」「Description」「Type」等であることが報告されている⁴。また、国立公文書館が行った調査では、公文書館等の横断検索のための項目として、「ID」「年代（作成年度）」「資料名」「作成者（部署名）」「備考等」を挙げる公文書館等が多くあった。この他、前述の「デジタルアーカイブの構築・共有・活用ガイドライン」においては、「タイトル（ラベル）」「作者（人物）」「日付（時代）」「場所」「管理番号（表内で重複しない恒久的な識別）」の5つが必須の項目として求められている⁵。さらに、同ガイドラインでは、デジタルコンテンツのサムネイル/プレビューのURLについても、メタデータ項目とすることが期待されている。

以上を踏まえると、横断検索に必要な最小限の目録データ項目としては、「ID」「年代（作成年度）」「資料名」「作成者（部署名）」「場所」「備考等」の6項目を考えることができる⁶。また、利用者によるデジタルコンテンツへのアクセスを容易化し、利活用を促進するうえでは、「デジタルコンテンツ（サムネイル/プレビューを含む）のURL」をメタデータ項目に含めることも有益と考えられる。

【標準仕様の考え方】

標準	Dublin Core の要素を参考に、横断検索のために必要な最小限の目録データの項目として、「ID」「年代（作成年度）」「資料名」「作成者（部署名）」「場所」「備考等」の6項目を設定する。
----	---

表 II-3 Dublin Core 要素一覧⁷

要素名	ラベル	説明
(1)タイトル	Title	オブジェクトの名前
(2)著者あるいは作者	Creator	情報資源の内容に関して責任を持つ人または組織
(3)主題およびキーワード	Subject	情報資源に述べられたトピック
(4)内容記述	Description	アブストラクトやイメージデータの説明など内容に関する記述
(5)公開者（出版者）	Publisher	情報資源を現在の形態にしたもの（出版社、大学など）
(6)寄与者	Contributor	著者ではないが文書の内容の作成に関わった人または組織（編集者や翻訳者等）

⁴ 松村敦, 古川沙希子, 宇陀則彦. 検索行動調査に基づく検索エレメント設計に関する一考察. 情報知識学会誌. 2007, vol. 17, no. 1, p. 15-31. <http://jglobal.jst.go.jp/public/20090422/200902241846828027>, (参照 2017-12-25).

⁵ デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会. “デジタルアーカイブの構築・共有・活用ガイドライン(平成 29 年 4 月)”. https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/guideline.pdf, (参照 2017-12-25).

⁶ 最小限の6項目のうち「ID」については、各公文書館等において資料特定のために用いられる番号（所蔵資料番号等）であって、検索項目としては適さない可能性もあることから、表示項目としてのみ使用することも考えられる。

⁷ 杉本重雄. “Dublin Core Metadata Element Set について — 現在の状況と利用例” 表1「Dublin Core の15エレメント」より. http://www.dl.slis.tsukuba.ac.jp/DLjournal/No_14/1-sugimoto/1-sugimoto.html, (参照 2017-12-25).

要素名	ラベル	説明
(7)日付	Date	現在の形で利用できるようになった日付
(8)資源タイプ	Type	ホームページ, 小説, 詩, 辞書といった情報資源の型
(9)形式	Format	Postscript ファイルやWindows 実行形式といった, 情報資源のデータ形式
(10)資源識別子	Identifier	情報資源を一意に識別するための番号あるいは名前
(11)情報源(出処)	Source	情報資源の出所となった情報資源を一意に示す番号もしくは文字列
(12)言語	Language	情報資源の内容を記述している言語
(13)関係	Relation	他の情報資源との関連づけ
(14)対象範囲(空間的・時間的)	Coverage	地理的場所や時間的な内容に関する情報資源の特性
(15)権利管理	Rights	著作権記述などの権利に関する記述や利用条件に関する記述へのリンク(URL もしくは何らかのURI)

実際に横断検索連携を行うにあたっては、各公文書館等において自館の目録データ項目と横断検索に必要な最小限の目録データ項目の対応表を定義することが必要となる。参考として、横断検索に必要な最小限の目録データ項目とISAD(G)との対応表を表 II-4に示す。

【標準仕様の考え方】

標準	公文書館等の目録データ項目と横断検索のために必要な最小限の目録データ項目(6項目)との対応表を定義する。
-----------	--

表 II-4 横断検索のために必要な最小限の目録データ項目(6項目)とISAD(G)との対応表⁸

横断検索のために必要な最小限の目録データ項目	ISAD(G)項目
ID	1.1 レファレンス・コード
年代(作成年度)	1.3 資料作成年月日
資料名	1.2 タイトル
作成者(部署名)	2.1 作成者名称
場所	—
備考等	—

複数の公文書館等の所蔵資料があわせて検索される横断検索においても、自館の所蔵資料のIDを一意に保つためには、他の公文書館等と重複しないIDを使用することが考えられる。そのような場合に採用できる識別子(ID)体系の標準としては、ISO 20247:2018として定められた「国際図書館資料識別子(International Library Item Identifier: ILII)がある。ILIIでは、公文書館等で既に使用しているID(ローカル資料識別子)を用いながら、以下の形式で識別子を記述することとされている。

⁸ Crosswalks: Dublin Core - ICA-AtoM, [https://wiki.ica-atom.org/Crosswalks:_ISAD\(G\)](https://wiki.ica-atom.org/Crosswalks:_ISAD(G)), (参照2018-02-01). を参考に作成。

(3) 階層構造の表現

効率的な資料検索のためには、所蔵資料の全体的な階層構造を見ながら、目当ての資料が含まれる可能性の高い階層に絞って検索を行うことが可能な階層検索が有効である。階層検索を可能とするためには、デジタルアーカイブ・システムの画面上で、目録データを階層的に表現するための構造情報を整備する必要がある。

資料の出所等に関する情報からは、資料の種類、時代背景、集合としての目的や意味を明らかにできることが多い。そのため公文書等の階層構造については、作成／移管元の組織構成（例えば、局→部→課→係（担当））や業務に基づき表現することが適切である。

なお、利用者に対して、最新の組織構成に基づく階層構造を提供するとした場合には、階層構造の適切な維持・管理を行う必要が生じる。また、階層構造の変更履歴をデータ化して管理することにより、組織構成の変遷を辿ることが可能な仕組みを用意することも考えられる。

【標準仕様の考え方】

標準	公文書等に関する階層構造は、公文書等の作成／移管元の組織や業務を反映したものとす。
-----------	---

(4) 文字コード

日本語を含む世界の主要な言語で使われる文字については、そのほとんどが、ISO/IEC10646として定められた文字集合であるUCS（Universal multi-octet coded Character Set）によって表現可能となっている。UCSには31ビット（4バイト）のUCS-4と、その部分集合である16ビット（2バイト）のUCS-2があるが、UCS-2の2バイト表記においては最大65,536文字までしか扱うことができない。そのため、中国語・日本語・韓国語において同等の意味やルーツを持つ漢字については、全て同じ文字とみなされ、同一のコードが割り当てられることから、適正な表現ができないことがある。このことから、デジタルアーカイブ・システムでは、UCS-4を表現できる符号化方式を採用する必要がある。

UCSの符号化方式としては、一般的にはUTF（Unicode Transformation Format）が使用されている。代表的なUTFとしては、UTF-8とUTF-16があり、それぞれ文字を表現する際の単位が異なっている。UTF-16では、2バイトまたは4バイトで文字を表現することとなっており、英数字及び日本語文字については全て2バイトで表現される。一方、UTF-8においては、英数字は1バイトで表現され、日本語は3バイトで表現されるようになっている。

【標準仕様の考え方】

標準	UCS-4を表現できる文字コードとする。
標準	符号化形式はUTF-8またはUTF-16 とす。

(5) 外字の扱い

デジタルアーカイブ・システムにおいて、古文書等でみられる特殊な文字を扱いたい場合、外字として登録する方法がある。ただし利用者にとっては、外字が使用されることで、キーワード検索が難しくなる（意図したデータがヒットしない）ことも考えられる。利用者の利便性を考慮すれば、可能な限り外字は用いず、UCS規格に含まれる代替可能な文字（代替文字）を用いることが求められる。

またUCS規格では、漢字における異体字を表現するための仕組みとして、漢字異体字シーケンス（Ideographic Variation Sequences）が規定されている。IVSでは、字体/字形を限定する字形選択子を付加することで、字体/字形を区別して扱うことが可能となっている。IVSは、これをサポートするOSやアプリケーションの普及状況にもよるものの、異体字を取り扱う方式の一つとして有用と考えられる。

外字を使用する場合には、情報処理学会試行標準IPSJ—TS0002：2014として整理されている「文字図形番号」⁹を用い、EGIX（Embedding Glyph Identifiers In XML Documents）により記述することが望ましい。EGIXは、W3Cが公開しているXMLによる外字記述方式であり、国内ではJIS X 4166（XML文書へのグリフ識別子の埋込み）としてJIS規格化されている。

【標準仕様の考え方】

標準	外字は指定せず、代替可能なUCS規格内の文字（代替文字）を用いる。
推奨	外字を用いる場合は、「文字図形番号」（情報処理学会試行標準IPSJ—TS0002：2014）を用い、EGIXで記述する。

```
<html xmlns="http://www.w3.org/1999/xhtml">
<body xmlns:glyph="http://www.xml.gr.jp/PRE/Reference">
<p><span glyph:name="ISO/IEC 10036/RA//Glyphs:10003290"
>吉</span>田茂</p>
</body>
</html>
```

図 II-2 EGIX による外字の記述例¹⁰

⁹ 情報処理学会. “文字図形識別情報 Character Shapes Identification (情報処理学会 試行標準 IPSJ-TS 0002:2014)”. <https://www.itscj.ipsj.or.jp/ipsj-ts/ts02-2014/charsp-id.htm>, (参照 2018-3-15).

¹⁰ W3C. “Embedding Glyph Identifiers in XML Documents”. 2002-12-20. <http://www.w3.org/TR/2002/NOTE-EGIX-20021220/>, (参照 2017-12-25).

2-1-2 デジタルコンテンツ

(1)画像データ

(ア)画像形式

デジタルアーカイブ・システム上で公開する画像データについては、利用者が特別なソフトウェア等を用いずとも閲覧可能な形式である必要がある。他方で、公開用画像のマスタデータとして公文書館等で長期的に保存する画像については、保存に適した形式とする必要がある。

事実上の標準または国際標準である画像等の形式としては、TIFF (Tagged Image File Format)、BMP (Bitmap)、PNG (Portable Network Graphics)、JPEG (Joint Photographic Experts Group)、JPEG 2000、PDF (Portable Document Format) があり、それぞれの特徴は下表のとおりとなっている。このうちPNG、JPEG、JPEG 2000、PDFはISOによって規格化されている。

表 II-5 画像等の形式

形式	特徴
TIFF	汎用性のある画像形式であり、様々な圧縮方式に対応。サポート環境が多く、画像データの保存形式としても普及している。
BMP	非圧縮の画像形式で、Windowsにおける標準的な画像形式。
PNG	可逆圧縮の画像形式。圧縮率は低い (JPEGの約5~10倍)。
JPEG	非可逆圧縮の画像形式。圧縮率を上げるとノイズが生じやすくなるが、インターネット上では主流の静止画規格である。
JPEG 2000	非可逆圧縮と可逆圧縮の両方に対応。圧縮効率が良いことから、圧縮をしてもブロックノイズ等の画像劣化が少ない。
PDF	閲覧環境によらず元のレイアウトを再現することができる。画像だけでなくテキストデータや図形を含めることも可能。

表 II-5記載の特徴を踏まえると、デジタルアーカイブ・システムのために準備する画像等の形式には、保存用としてTIFF (非圧縮)、公開用としてJPEGまたはPDFを採用することが適当と考えられる。また、保存・公開の両用としてはJPEG 2000を採用することが考えられる。

なお、JPEG 2000については、これを標準サポートしているOS等が少なく、閲覧にあたってはJPEG 2000対応のビューワの用意やブラウザへのプラグイン等が必要となる点に留意が必要である。

【標準仕様の考え方】

標準	保存用の画像形式としてTIFF (非圧縮)、公開用の画像形式としてJPEGまたはPDFをサポートする。
推奨	保存用と公開用を兼ねた画像形式としてJPEG 2000をサポートする。

(イ)解像度、色数等

各公文書館等においてデジタルアーカイブ・システムのために用意する画像については、用途に応じて適切な解像度を選択することが望ましい。例えば保存用の画像には、小さな文字や線でも視認可能なように

400dpi程度の解像度とする一方で、公開用の画像については、画像サイズを抑えながらもディスプレイ上では劣化が視認できない程度の解像度にするといったことが考えられる¹¹。

国立国会図書館が公開している「国立国会図書館 資料デジタル化の手引 2017年版」では、保存用画像、提供用画像、サムネイル画像といった用途別に画像の解像度の基準が示されており、デジタルアーカイブ・システムのための画像作成の参考とすることができる¹²。

また、画像色数の設定については、白黒の文書のデジタル画像化（マイクロフィルムからデジタル画像を作成する場合など）であれば8ビットグレースケール、カラー画像作成の場合は24ビットカラーとすることが適当と考えられる。

【標準仕様の考え方】

推奨	公開用画像は、画像サイズを抑えつつも劣化が視認できないレベルの画質に保つ。
----	---------------------------------------

(2)動画・音声データ

(ア)動画データ

デジタルアーカイブ・システムで動画を配信する場合、動画データは利用者が特別なソフトウェア等を用いずとも閲覧可能な形式であることが望ましい。

動画データは、動画及び音声データの符号化方式（コーデック）とそれらのデータが格納されるコンテナ（形式）という2つの要素から構成される。表 II-6は、代表的な動画データの形式（コンテナ）と格納されるコーデックの組み合わせである。

表 II-6 代表的な動画データ形式

形式 (コンテナ)	格納される代表的なコーデック		説明
	映像	音声	
AVI	WMV7/WMV8/WMV9 Motion JPEG (MJPEG) H.264/MPEG4-AVC H.265	LPCM WMA MP3 AAC	AVIは、Microsoftが開発した動画フォーマットであり、様々なコーデックをサポートしている。
MP4	MPEG-2 Video H.264/MPEG4-AVC H.265	MP3 AAC Apple Lossless	MP4はMPEG-4 Part14の略称。ISO/IECの動画コーデックの国際標準規格であるMPEG-4の第14部で策定された動画フォーマットである。

¹¹ スマートフォンやタブレット等のスマートデバイスが急速に普及したことにより、利用者がデジタルアーカイブ・システム上の画像をスマートデバイスから閲覧するケースが増えることが予想される。スマートフォンの画面解像度(デバイスピクセル)は320~375×568~667が6割近く、タブレットでは768×1024が6割近くを占めているが、高解像度ディスプレイが採用されている場合はこの2倍程度の画像解像度でなければ画像がぼやけることがある。そのため、スマートデバイスでの閲覧を想定する場合は、画像の撮影解像度に注意が必要である。

¹² “国立国会図書館 資料デジタル化の手引 2017年版”, p.17. 「2.2.3 画像別の解像度」. 国立国会図書館.
http://dl.ndl.go.jp/view/download/digidepo_10341525_po_digitalguide170428.pdf?contentNo=1&alternativeNo= (参照 2017-12-25).

インターネットにおける一般的な動画データ形式としては、現在はMP4が普及している。また、MP4に格納される動画コーデックとしては、圧縮効率の良いH.264/MPEG4-AVCが主に使用されている。MP4は、主要なWEBブラウザやスマートフォンにおいて特別なソフトウェアを用いずとも再生可能であることから、デジタルアーカイブ・システムで配信する動画データ形式としてはMP4が望ましいと考えられる。

(イ) 音声データ

動画データの場合と同様、音声データについても利用者の環境を選ばない形式であることが望まれる。

音声データも、符号化方式(コーデック)とそれをまとめるコンテナという2つの要素から構成されるが、音声データの場合は1つのコーデックが1つの形式に対応していることが多い。表 II-7は、現在の代表的な音声データ形式である。

表 II-7 代表的な音声形式

形式	説明
WAV	MicrosoftとIBMにより開発されたコンテナ。一般的には非圧縮の音声データ(リニアPCM)が格納される。
AAC	国際標準の非可逆圧縮フォーマット。地上デジタル放送やBSデジタル放送の音声に使用されている。携帯音楽プレーヤーやスマートフォンにおいてもよく使用されているフォーマットのの一つである。
MP3	国際標準の非可逆圧縮フォーマット。主要なOSやWEBブラウザがサポートしていることから、非可逆圧縮の音声データにおける事実上の標準フォーマットとなっている。

普及度や再生環境を選ばない点などを踏まえると、デジタルアーカイブ・システムで配信する音声データ形式としてはMP3が望ましいと考えられる。

【標準仕様の考え方】

推奨	公開用の動画データの形式として、MP4をサポートする。
推奨	公開用の音声データの形式として、MP3をサポートする。

2-2 目録データ・デジタルコンテンツの登録と管理

2-2-1 目録データ

(1)目録データの登録・管理機能

(ア)データベースと検索エンジンの考え方

デジタルアーカイブ・システムでは、扱う目録データの符号化方式や各公文書館等の業務に合わせて、適切なデータベースや検索エンジンを採用する必要がある。

また、目録データの登録や管理にあたっては、後述の一括登録、一括更新、一括出力などの方法により、データの編集や抽出を効率的に行うことが求められる。この点においては、リレーショナル型のデータベースの導入も有効と考えられる。

【標準仕様の考え方】

標準

目録データの電子符号化方式等に合わせて、適切なデータベースや検索エンジンを採用する。

(イ)目録データ登録・管理機能

デジタルアーカイブ・システムへ登録される目録データの件数や種類は、各公文書館等により異なることが考えられる。目録データの登録・管理に係る作業の内容や規模により、目録データを1件ごとに登録、更新、削除する機能から、大量の目録データを一括して登録する機能、既に登録されている目録データを出力し表計算ソフト等で修正した上で一括して反映(更新)する機能まで、様々な方法により目録データを登録・管理する機能が必要とされる。

【標準仕様の考え方】

標準

目録データを個別及び一括で登録、編集、削除、出力する機能を設ける。

デジタルアーカイブ・システム上で取り扱う情報については、その特性に応じて、一般利用者へ公開するか、職員のみ利用に限定するか等、公開範囲を整理する必要がある。また、非公開の情報を含む目録データの管理には、一般公開するデータの管理と比べて強固なセキュリティが求められる。そのため、職員の業務においてのみ利用される非公開の目録データについては、一般公開する目録データとは物理的な格納先を分離する、非公開の目録データを扱う領域へのアクセスを内部ネットワークからのみに制限するといった仕組みが求められる(デジタルアーカイブ・システムとは別のものとして、非公開の目録データやデジタルコンテンツを扱うシステムを構築することも考えられる)。

【標準仕様の考え方】

標準

非公開の目録データと一般公開する目録データを分離し、アクセス制御する。

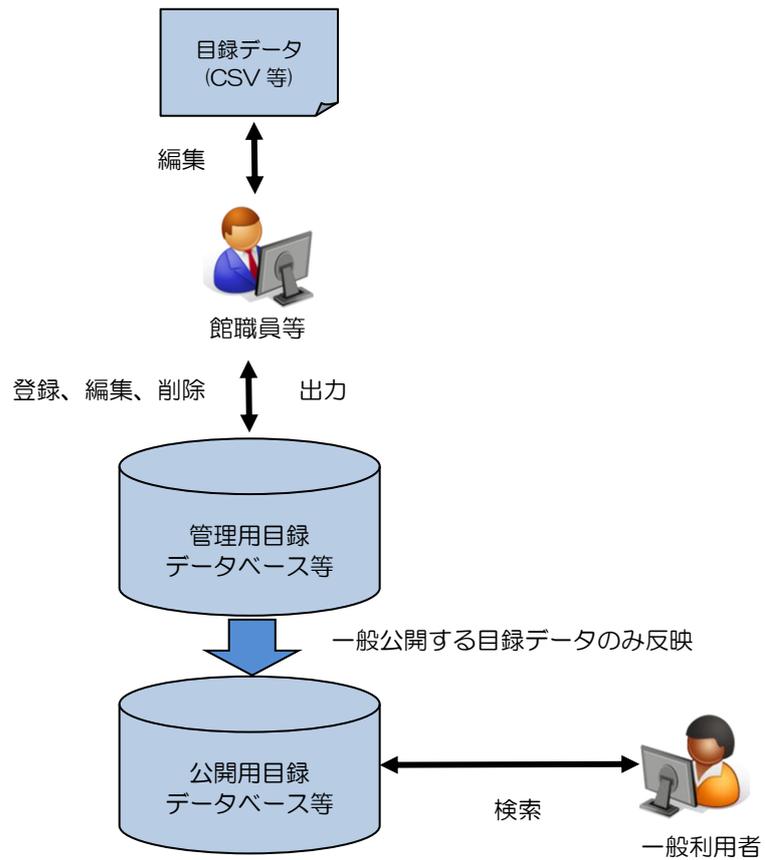


図 II-3 目録データの登録・管理と一般利用のイメージ

2-2-2 デジタルコンテンツ

(1) デジタルコンテンツの登録・管理機能

デジタルアーカイブ・システムではデジタルコンテンツの効率的な登録・管理を行うため、デジタルコンテンツとその管理データを個別に登録、編集、削除する機能、また一括して登録、編集、削除する機能が必要となる。なお、ここでいう編集機能とは、デジタルコンテンツファイルの差し替えや、目録データとデジタルコンテンツの関連付けの変更等を行うための機能を指す。

デジタルコンテンツの管理データの項目としては、デジタルコンテンツを管理するためのファイル名、目録データとの関連付け情報、ファイル形式、ファイルサイズ等が想定される。

【標準仕様の考え方】

標準

デジタルコンテンツとその管理データを個別及び一括で登録、編集、削除する機能を設ける。

非公開のデジタルコンテンツの管理については、非公開の目録データと同様のセキュリティレベルが求められるため、非公開のデジタルコンテンツと一般公開するデジタルコンテンツの分離格納、非公開のデジタルコンテンツを扱う領域へのアクセス制御といった対応が必要となる。

【標準仕様の考え方】

標準

非公開のデジタルコンテンツと一般公開するデジタルコンテンツを分離し、アクセス制御する。

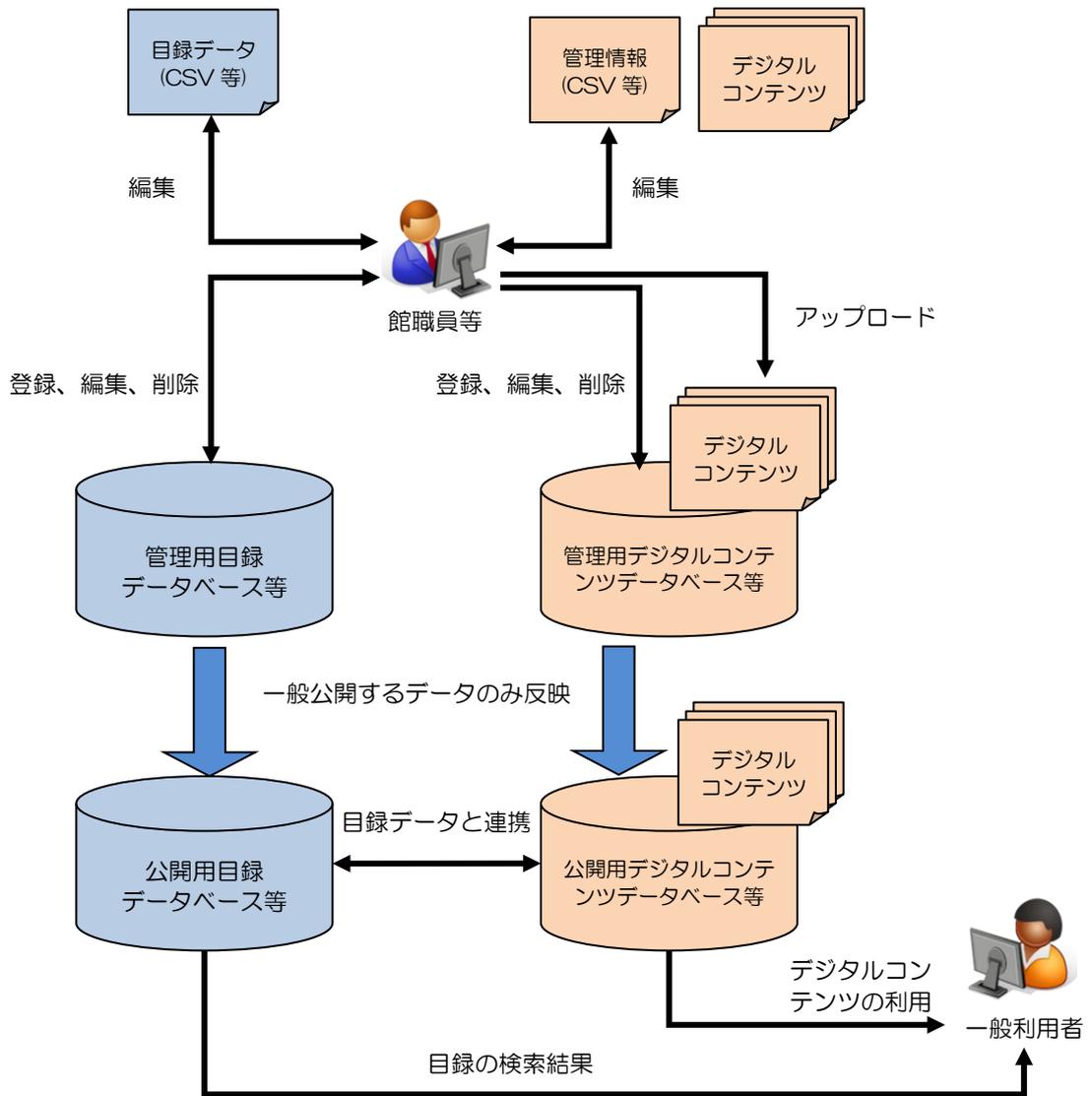


図 II-4 デジタルコンテンツの登録・管理と一般利用のイメージ

2-3 利用者向け機能

2-3-1 検索機能

(1)検索機能

利用者が資料へ辿り着くための手段であるデジタルアーカイブ・システムでは、整備された目録データを効率的に検索できることが求められる。基本的な検索方式として、任意の項目を対象としたキーワード及び条件指定（画像の有/無等の条件、年代の範囲指定等）による検索、キーワードによる項目を限定しない全文検索、また、複数の条件及びキーワードの組み合わせによる検索を行える必要がある。

また、資料の階層構造を参照可能とすることや、特定の階層以下に限定した検索（階層検索）を可能とすることで、利用者の利便性を高めることができる。特に検索結果に関連する資料を探すには、階層を辿る方法が容易であることから、検索結果に示される資料が属する階層を把握できるようにすることも重要である。

【標準仕様の考え方】

標準	任意の項目を対象とした、キーワード及び条件指定による検索を可能とする。
標準	項目を限定せずに、キーワードによる全文検索を可能とする。
標準	複数の条件及びキーワードの組み合わせによる検索を可能とする。
標準	目録データの階層構造表示から、指定した階層に属する目録データの一覧表示を可能とする。
標準	目録データの階層構造表示から、指定した階層以下に属する目録データに限定した検索を可能とする。



図 II-5 階層構造をツリー型で表現している事例（国立公文書館デジタルアーカイブ）

(2) 検索支援機能

デジタルアーカイブ・システムでは、キーワードや階層構造による検索以外の検索支援機能を設けることで、公文書等に馴染みのない一般的な利用者から専門的知識を持つ利用者まで、幅広い利用者にとっての利便性を向上させることが可能となる。

主な検索支援機能の例としては、辞書検索機能、ファセット検索機能、異体字検索機能があり、それぞれの概要は以下のとおりである。

表 II-8 検索支援機能の例

辞書検索機能	利用者が検索を行う際、公文書等に用いられる専門用語を検索ワードとしての確に入力できるとは限らない。システム上に同意語彙の辞書を登載し、検索時に利用できるようにすることで、よりの確な検索を支援することができる。また、辞書に関連する機能としては、利用者が辞書の使用を選択できる機能、利用者が検索に利用した語彙を把握できる機能なども考えられる。
ファセット検索機能	検索結果を特定項目ごとに分類し、それぞれの項目ごとにヒット件数を表示することで、利用者が容易に絞り込みを行える機能。例えば、デジタルコンテンツの種類（画像、文書、動画、音声等）、年代、作成/移管元機関等の項目により結果を分類することが考えられる。
異体字検索機能	漢字における異体字表記（「高」と「𨳊」等）、ひらがな・カタカナ、英数字の全角・半角や表記のゆれなど、コンピュータ上では区別される文字や表現をあわせて検索する機能。

【標準仕様の考え方】

推奨 辞書検索機能、ファセット検索機能、異体字検索機能等の検索支援機能を設ける。



図 II-6 ファセット検索の例(秋田県公文書館デジタルアーカイブ)

2-3-2 デジタルコンテンツ閲覧機能

(1) デジタルコンテンツ閲覧機能

デジタルアーカイブ・システム上で、目録の検索結果からデジタルコンテンツを閲覧可能とするためには、デジタルコンテンツと目録データを連携させる機能を実装する必要がある。

【標準仕様の考え方】

標準

目録データに対応するデジタルコンテンツが存在する場合には、目録データの検索結果一覧及び詳細表示にデジタルコンテンツへのリンクを表示し、当該リンクからデジタルコンテンツを閲覧可能とする。

また、デジタルアーカイブ・システムでは、誰もがデジタルコンテンツにアクセスできるよう、特別なソフトウェアを必要とせずにデジタルコンテンツを閲覧できることが求められる。現在の主要なWEBブラウザでは、HTML5やCSS3/4（Cascading Style Sheets, level 3/4）への対応が進んでいることから、画像データのための高機能なユーザ・インターフェースをWEBブラウザのみで実現することができるようになっている。

なお画像ビューアには、画像の拡大、縮小、表示位置の移動等の操作を利用者が直観的に行えること、また、複数ページにより構成される画像の場合は、前後のページや先頭・最後のページへの移動が容易に行えることが求められる。



図 II-7 画像ビューアの例（国立公文書館デジタルアーカイブ）

【標準仕様の考え方】

標準	画像データの閲覧方式は、利用者が特別なソフトウェアを必要としない方式とする。
標準	画像ビューアは、直観的な操作で拡大、縮小、移動、ページ送り等が行えるものとする。

デジタルアーカイブ・システム上で、デジタルコンテンツの縮小画像（サムネイル）や部分抽出（プレビュー）を表示することは、利用者が目的の資料を見つけやすくするための方法として有効である。

また、デジタルコンテンツを利用者が個別及び一括でダウンロードできるようにすることも、利用者の利便性向上及び利活用促進の観点から望ましいと考えられる。このダウンロード機能の実装とあわせて、二次利用に関するライセンス情報の明示を行うことで（後述）、利用者によるデジタルコンテンツの利活用をサポートすることが可能となる。

【標準仕様の考え方】

推奨	デジタルコンテンツのサムネイル画像を表示する。
推奨	デジタルコンテンツを個別及び一括でダウンロードする機能を設ける。



図 II-8 サムネイル表示の例（アジア歴史資料センター）

(2)画像配信機能

公文書館等のデジタルアーカイブ・システム以外でも、インターネット上では図書館や博物館等の様々な機関によって画像データが公開されている。これらの画像について活発な相互利用が行われるためには、画像データの配信方式等が標準化されていることが望ましい。

近年では、画像の相互運用性を高めることを目的とする国際的な画像配信等のためのフレームワークとして、IIIF (International Image Interoperability Framework) の普及が進んでいる¹³。IIIFのAPIを実装するサーバ上の画像を閲覧する場合、利用者は画像閲覧に用いるIIIF対応ビューアを自由に選択することができる。具体的には利用者側では、公文書館等のデジタルアーカイブ・システムで提供されるビューア以外のIIIF対応ビューアを用いて、IIIFに基づく配信を行っている公文書館等や博物館等の画像群を比較しながら閲覧するといった利用が可能となる。

デジタルアーカイブ・システムにおける画像配信機能については、利用者の利便性の向上等の観点から、IIIFのような画像の相互運用性に考慮した方式を採用することが望ましい。

【標準仕様の考え方】

推奨

画像の相互運用性に考慮した配信方式（例：IIIF）を採用する。

また、高精細画像やフルカラーの大判画像を公開する場合には、大容量の画像を効率的に配信する仕組みが求められる。こうした大容量画像の配信には、元の画像を細分化したタイル画像を配信するためのサーバ機能を実装することも考えられる。

(3)動画・音声配信機能

現時点では、動画・音声の配信を行っている公文書館等は少数にとどまるものの、動画・音声の資料を公開したいというニーズは少なからずある。

動画・音声は、画像と比べるとデータサイズが大きくなる傾向があるため、一般利用者がストレス無く視聴できるためには、ストリーミング配信方式を採用することが望ましい。ストリーミング配信方式においては、専用の配信サーバソフトウェアを用いて必要なデータを都度配信することで、ネットワークを効率的に利用することができる。

また近年では、YouTubeに代表される動画配信サービスが広く普及しており、動画配信を容易に行える環境が整っている。公文書館等において動画配信サービスの利用が認められる場合には、デジタルアーカイブ・システムにおける動画・音声の配信の手段として検討することも考えられる。

【標準仕様の考え方】

推奨

動画・音声の公開にはストリーミング配信方式を採用する。

¹³ IIIF における画像配信に関わる API としては IIIF Image API があり、配信画像フォーマットは JPEG 形式とされている。公文書館等で、IIIF Image API 実装サーバに JPEG 以外の画像形式を格納する場合でも、当該サーバに画像を JPEG 変換する機能があれば、IIIF による JPEG 配信を行うことができる。

2-3-3 ユーザ・インターフェース

(1) アクセシビリティ

「誰でも」利用できるデジタルアーカイブ・システムを実現するためには、アクセシビリティを考慮したユーザ・インターフェースを提供することが望ましい¹⁴。

【標準仕様の考え方】

推奨

アクセシビリティを考慮したユーザ・インターフェースを提供する。

(2) スマートデバイスのためのユーザ・インターフェース

平成29年3月末現在の政府統計によると、タブレットの普及率は28.7%、スマートフォンの普及率は60.3%となっており、59歳以下の年齢ではスマートフォンの普及率がPCの普及率を上回っている¹⁵。デジタルアーカイブ・システムでも、PCによる閲覧に加えて、今後はタブレット等のスマートデバイスによる閲覧が増加することも予想される。

スマートデバイスは、従来のPCとは異なるディスプレイサイズや解像度を持ち、またタップやスワイプといったPCには存在しない操作があることから、デジタルアーカイブ・システムにスマートデバイスのためのユーザ・インターフェースを用意することが望ましい。スマートデバイスにも適したユーザ・インターフェースを提供することで、利用者層の裾野の拡大を期待することができる。

【標準仕様の考え方】

推奨

スマートデバイス用のユーザ・インターフェースを提供する。

¹⁴ 総務省では公的機関がウェブアクセシビリティの確保・維持・向上に係る取組の支援を目的として「みんなの公共サイト運用ガイドライン」を公開しており、2016年のJIS X 8341-3の改正に合わせて改訂を行っている。自治体等では同ガイドラインを元に独自のウェブアクセシビリティガイドラインを設けているケースもある。デジタルアーカイブ・システムの構築にあたっては、このようなガイドラインを参考とすることが考えられる。

¹⁵ 内閣府景気統計部、「消費動向調査 平成29年3月調査」、政府統計の総合窓口(e-Stat)。

<https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00100405&tstat=000001014549&cycle=0&year=20170&tclass1=000001095275&tclass2=000001099216&second2=1>, (参照 2018-01-05)

2-4 利活用の促進

2-4-1 横断検索機能

(1)横断検索プロトコル

各公文書館等の目録データ等の横断検索を行うためのプロトコルとしては、Z39.50の後継でHTTPをベースとするSRU/SRW（SRU：Search/Retrieve via URL、SRW：Search/Retrieve Web Service）の普及が進んでいる。また、SRU/SRWと同様のHTTPベースの検索プロトコルとして、OpenSearchの採用も見られるようになってきている¹⁶。デジタルアーカイブ・システムで他の公文書館等をはじめとする外部機関との横断検索連携を実現するためには、SRU/SRWまたはOpenSearchによる検索に対応する必要がある。

【標準仕様の考え方】

標準

SRU/SRWまたはOpenSearchによる検索に対応する。

2-4-2 一括提供機能

目録データやデジタルコンテンツの利活用促進には、利用者の要求に応じて目録データ等のメタデータを一括して提供等する仕組みも有効と考えられる。このようなメタデータ交換のための国際標準プロトコルとしては、OAI-PMH（Open Archives Initiative Protocol for Metadata Harvesting）が存在し、国内外でも学術機関を中心に、OAI-PMHによるメタデータ提供機能の実装が増えている。OAI-PMHでは、メタデータを収集するクライアントにあたる「ハーベスタ」と、メタデータを提供するサーバにあたる「リポジトリ」の機能が定義されている。

公文書館等においても他機関との横断検索連携の取組を視野に入れ、OAI-PMHリポジトリ機能を実装することが望ましい。

【標準仕様の考え方】

推奨

OAI-PMHによる目録データの一括提供機能（リポジトリ機能）を実装する。

2-4-3 外部提供インターフェース

近年では、公的機関が作成するデータの利活用拡大のため、国規模のオープンデータ化が推進されており、デジタルアーカイブ・システムについても、機械可読形式で目録データ等の提供を行うことが期待されている。

¹⁶ SRU/SRW は、OASIS(Organization for the Advancement of Structured Information Standards)において searchRetrieve 1.0 として標準化されており、現在は米国議会図書館が標準の管理を行っている。OpenSearch は Amazon.com の子会社 A9 によって開発され、クリエイティブ・コモンズ・ライセンス下で A9 により提供されている。SRU/SRW は図書館や学術機関での採用が多く、OpenSearch は民間・商用での採用が多い。国立国会図書館の国立国会図書館サーチ(NDL サーチ)ではSRU/SRW、OpenSearch の両方に対応している。

現在、Web上で普及している機械可読のデータ形式としては、RDF（Resource Description Framework）を代表的なものとしてあげることができる。RDFは、Web上の情報資源を記述するための標準的フレームワークであり、情報資源はRDFで記述されることで多様なアプリケーションによる取り扱いが可能となる。また、RDFで記述された情報資源を検索するためのAPIとしてはSPARQL（SPARQL Protocol and RDF Query Language）があり、W3Cによって標準化されている。

デジタルアーカイブ・システム上の目録データ等の情報資源を他の異なるアプリケーションからも利用可能とするためには、目録データ等のRDF化を行いつつ、SPARQLのようなWeb APIを実装することが望まれる。

【標準仕様の考え方】

推奨	情報を外部提供するインターフェースとしてWeb API（SPARQL等）を実装する。
----	--

現在では、インターネット上で情報発信するための手段として、SNS（ソーシャル・ネットワーキング・サービス）が広く普及している。デジタルアーカイブ・システム上の目録データやデジタルコンテンツの情報を、利用者がSNSを用いて容易に発信できる仕組みを実装することで、利活用の促進と利用者の拡大を期待することができる。

2-4-4 永続的アクセス

目録データやデジタルコンテンツへのアクセスURLは、利用者によりブックマーク登録されることや、Webサイト、論文等の参考文献として記載されることなどが想定される。また、他機関との連携においても、URLによって連携先システムへの遷移が行われることが想定される。このことから、目録データやデジタルコンテンツのURLについては、リンク切れが起こらないような仕組みとすることが望ましい。

リンク切れは、システム内で固定のURLを設定することにより防止することができる。また、システム移行時においては、可能な限り固定のURLを維持すること、またURLが変更となる場合でも、そのルールを明示することにより変更後のURLを容易に追跡可能とすることが望ましい。

【標準仕様の考え方】

推奨	目録データやデジタルコンテンツのURLのリンク切れを防止する。
----	---------------------------------

Web上での永続的アクセスを保証する仕組みとしては、パーマリンク（Permanent Linkの略）がある。パーマリンクとは、Webサイト内のページやコンテンツに対して永続的に与えられるURI（URLとする場合も考えられる）を指す。デジタルアーカイブ・システム上の目録データ等について、長期的に安定したアクセスを提供するためには、パーマリンクを導入することが理想的である。ただし、パーマリンクの導入にあたっては、将来における目録データ等の増加やURLの維持にかかる費用を想定し、システム構築時にURLを慎重に決定する必要がある。

2-4-5 ライセンスの表示

デジタルアーカイブ・システムで公開しているデジタルコンテンツ等については、その利用条件を明示することで、二次利用など利活用の促進を期待することができる。

インターネット上のコンテンツ等の利用条件を明示するための枠組みとしては、例えばパブリック・ドメイン・マークやクリエイティブ・コモンズ・ライセンスがある。クリエイティブ・コモンズ・ライセンスの枠組みでは、著作権者によりコンテンツの再利用に係る条件が明示されることによって、コンテンツの適正な再利用を促進することがねらいとされている。デジタルアーカイブ・システムで提供するデジタルコンテンツ等についても、このような枠組みを用いてライセンス情報を分かりやすく提示することが望ましい。

なお、デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会の「デジタルアーカイブの構築・共有・活用ガイドライン」の補足資料では、利用条件表示一覧としてパブリック・ドメインやクリエイティブ・コモンズ・ライセンスについての整理があり、具体的なライセンス表示を行う際の参考とすることができる¹⁷。

【標準仕様の考え方】

推奨	利用条件を明示するための枠組みを用いて、デジタルコンテンツ等の利用条件を明示する。
----	---

¹⁷ デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会、「デジタルアーカイブの構築・共有・活用ガイドライン(平成 29 年 4 月)」, p. 32.

https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/guideline.pdf, (参照 2017-12-25).

2-5 利用者支援

2-5-1 システムの利用案内

各公文書館等で構築されるデジタルアーカイブ・システムについては、操作方法や動作環境がそれぞれ異なることが想定される。また、目録データやデジタルコンテンツの提供の在り方・方針も、各公文書館等で異なることが考えられる。様々な利用者が円滑にデジタルアーカイブ・システムを利用できるためには、デジタルアーカイブ・システム上に利用案内のページ（ヘルプページ等）を設けて、それらのページへのリンクを分かりやすい場所に設置することが求められる。

【標準仕様の考え方】

標準

デジタルアーカイブ・システムの利用案内（利用方法、動作環境等）を明示する。

また、デジタルアーカイブ・システムの検索結果から、原本の閲覧を希望する利用者に対しては、原本の利用手続きが分かりやすい形で示されていることが望ましい。

公文書館等では資料の閲覧方法をホームページ等で案内していることも多いことから、それらのページへのリンクを適切な場所に設置するといった方法が考えられる。

【標準仕様の考え方】

推奨

資料（原本）の閲覧方法を利用者が把握しやすいように案内を設ける。

2-5-2 レファレンス機能

デジタルアーカイブ・システムで提供する目録データやデジタルコンテンツに関しては、利用者へのレファレンスサービスを行うことも想定される。想定されるレファレンス事項や、頻繁に繰り返されるレファレンス事項については、FAQ（Frequently Asked Questions）としてあらかじめデジタルアーカイブ・システムに掲載することで、利用者の資料検索等を支援することが可能となる。また、FAQでは解決しない内容のレファレンスを受け付けるためには、専用の入力フォーム等の仕組みを設けることが望ましい。

レファレンス専用の入力フォーム等については、対象となる資料の識別情報（ID等）や問い合わせの種類（質問、誤りの指摘等）といった定型的な入力項目を設けることで、レファレンスサービス実施に必要な情報を得ることが容易になる。

【標準仕様の考え方】

推奨

利用者から目録データやデジタルコンテンツに関するレファレンスを受け付ける仕組みを設ける。

2-6 業務向け機能

2-6-1 利用状況の把握機能

利用者ニーズを適切に踏まえながら、目録データやデジタルコンテンツの充実化を図るにあたっては、デジタルアーカイブ・システムの利用状況を把握するための機能が必要となる。

また、デジタルアーカイブ・システムへのアクセス状況等の情報は、各公文書館等における年報や利用実績報告のなかで活用することが可能である。

【標準仕様の考え方】

標準

Webアクセス状況等の統計情報を作成し、利用状況を把握可能とする。

2-6-2 閲覧業務関連機能

公文書館等においては一般的に、特定の様式による閲覧請求書等が用意されている。原本の来館利用に際しての利便性向上のためには、デジタルアーカイブ・システムの検索結果から、閲覧請求書等の書式を印刷できる必要がある。

また、デジタルアーカイブ・システムにおいて個人情報を安全に取り扱うための情報セキュリティ対策を十分に実施できる場合には、閲覧利用者管理機能やオンラインによる利用請求機能を設けることも考えられる。

【標準仕様の考え方】

標準

閲覧請求書等の印刷機能を設ける。

2-7 システム環境

2-7-1 システム環境検討のポイント

各公文書館等ではデジタルアーカイブ・システムを導入するにあたり、その個別の状況やニーズ等に応じて、システム環境を適切に選択する必要がある。各公文書館等でデジタルアーカイブ・システムのシステム環境について検討を行う際のポイントとしては、以下のような点があげることができる。

(1) 運用形態の検討

従来は、それぞれの公文書館等が単独でデジタルアーカイブ・システムを構築し、運用する形態のみが想定されていた。しかしながら近年では、クラウドサービスを利用することで、複数の公文書館等や類縁機関等が共同でデジタルアーカイブ・システムを運用することも可能となっている。

デジタルアーカイブ・システムの構築にあたっては、単館運用及び共同運用のそれぞれの特徴を踏まえながら、採用する運用形態を検討する必要がある。

表 II-9 デジタルアーカイブ・システムにおける運用形態の特徴

運用形態	メリット	デメリット
単館運用	自館の業務や利用者ニーズにあわせたアプリケーションを構築できるため、サービス面において自館の独自性を出すことが比較的容易である。	構築・運用コストを単独で確保する必要がある。
共同運用	共同運用者の間で構築・運用コストを分担することにより、各館あたりの構築・運用コストの抑制を期待できる。	共同運用する機関との間で、実装機能、構築・運用コスト、運用・保守等に係る調整が必要となる。

(2) 構成の検討

各公文書館等ではデジタルアーカイブ・システムの構築にあたり、システムで扱うデータの種類（公開/非公開）、及びシステムの機能構成について検討を行う必要がある。表 II-10では、システムで扱うデータの種類及び本書で示した推奨機能の有無という点から、デジタルアーカイブ・システムの構成例を「基本的な構成」、「標準的な構成」、「発展的な構成」の3種類に整理した。各公文書館等ではこの3つの構成例を参考に、自館に適したデジタルアーカイブ・システムの構成についての検討を行うことが可能である。

表 II-10 デジタルアーカイブ・システムにおける構成の例

構成	扱うデータの種類		推奨機能等
	公開データ	非公開データ	
基本的な構成	○	-	-
標準的な構成	○	○	-
発展的な構成	○	○	○

○：あり -：なし

(※) 推奨機能等とは、本標準仕様書で「推奨」とする機能及び本標準仕様書で紹介した付加的価値のある機能を指している。

上記の3種類の構成（基本的な構成、標準的な構成、発展的な構成）に対応する具体的なシステム構成例は、図 II-9～図 II-11のとおりとなる。

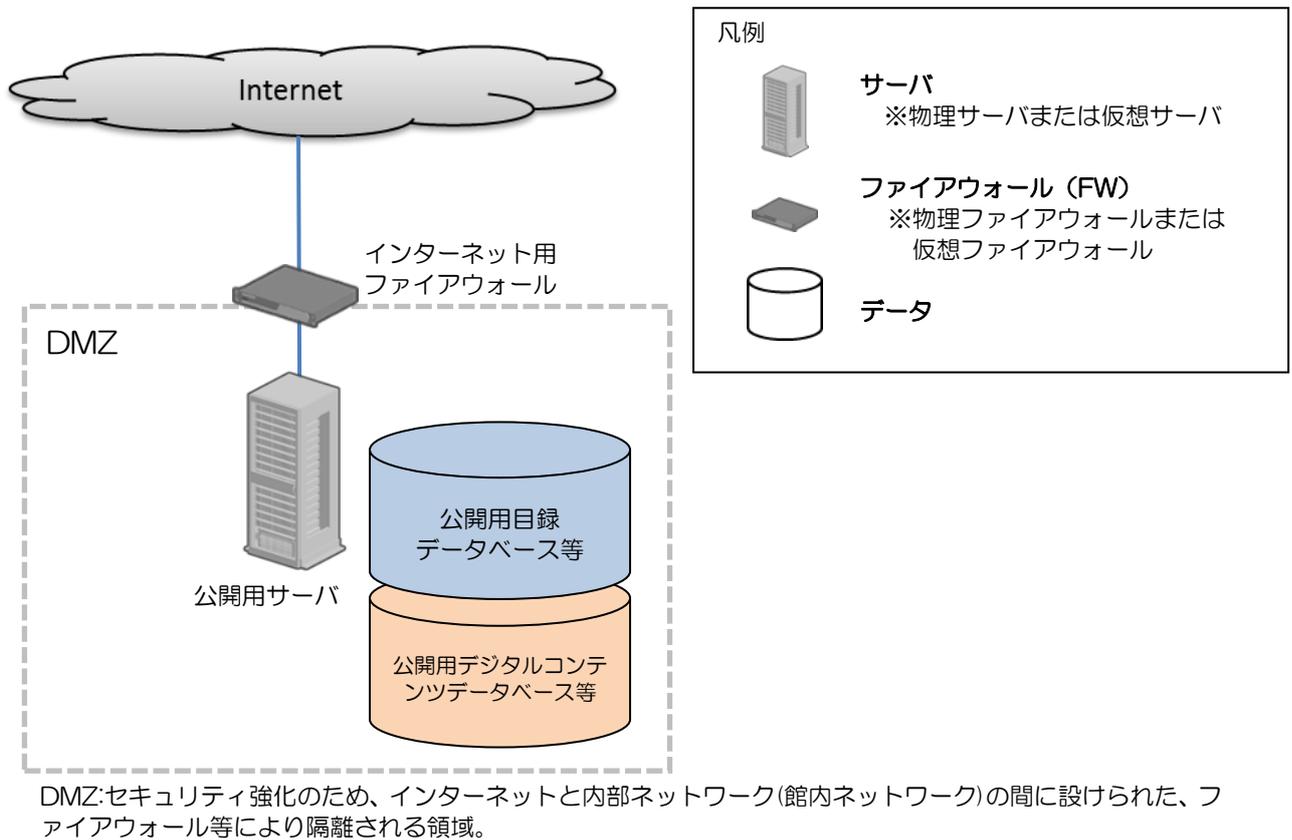


図 II-9 基本的な構成でのシステム構成図

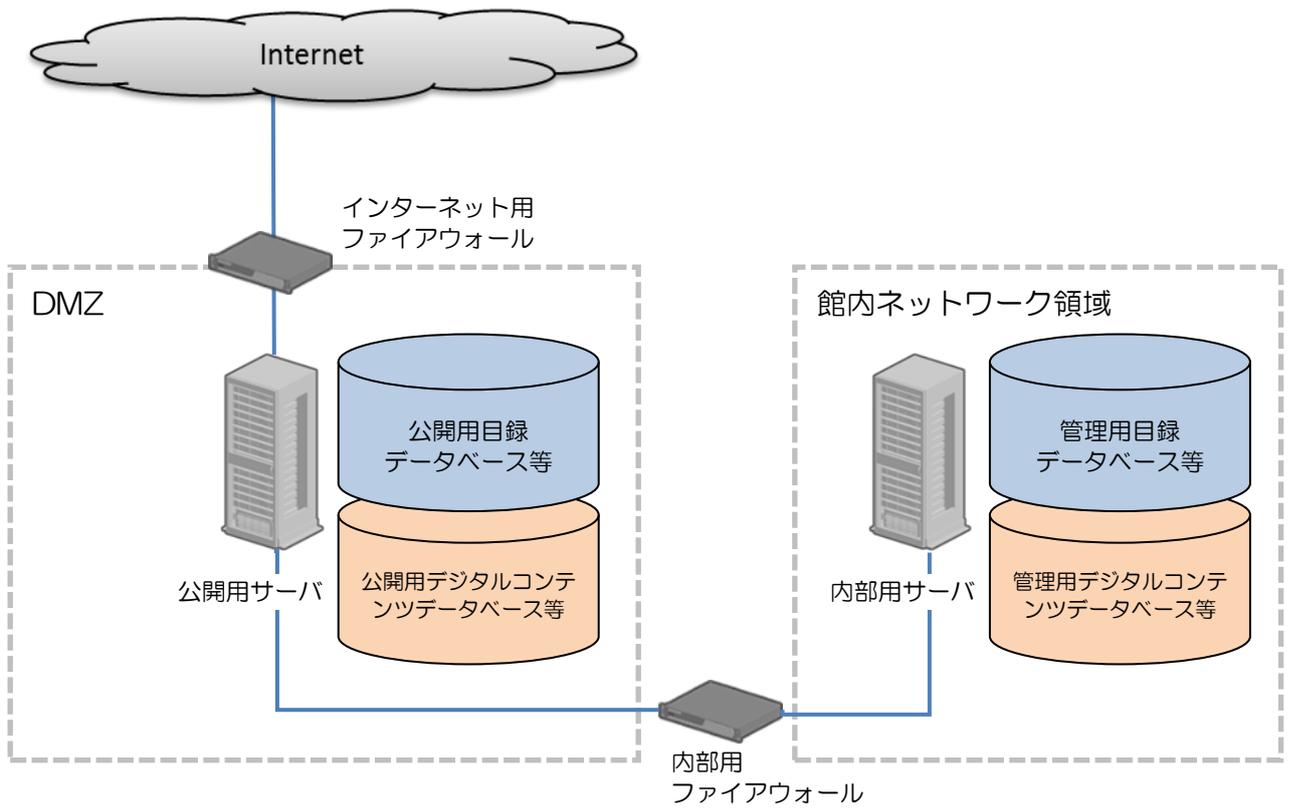


図 II-10 標準的な構成でのシステム構成図

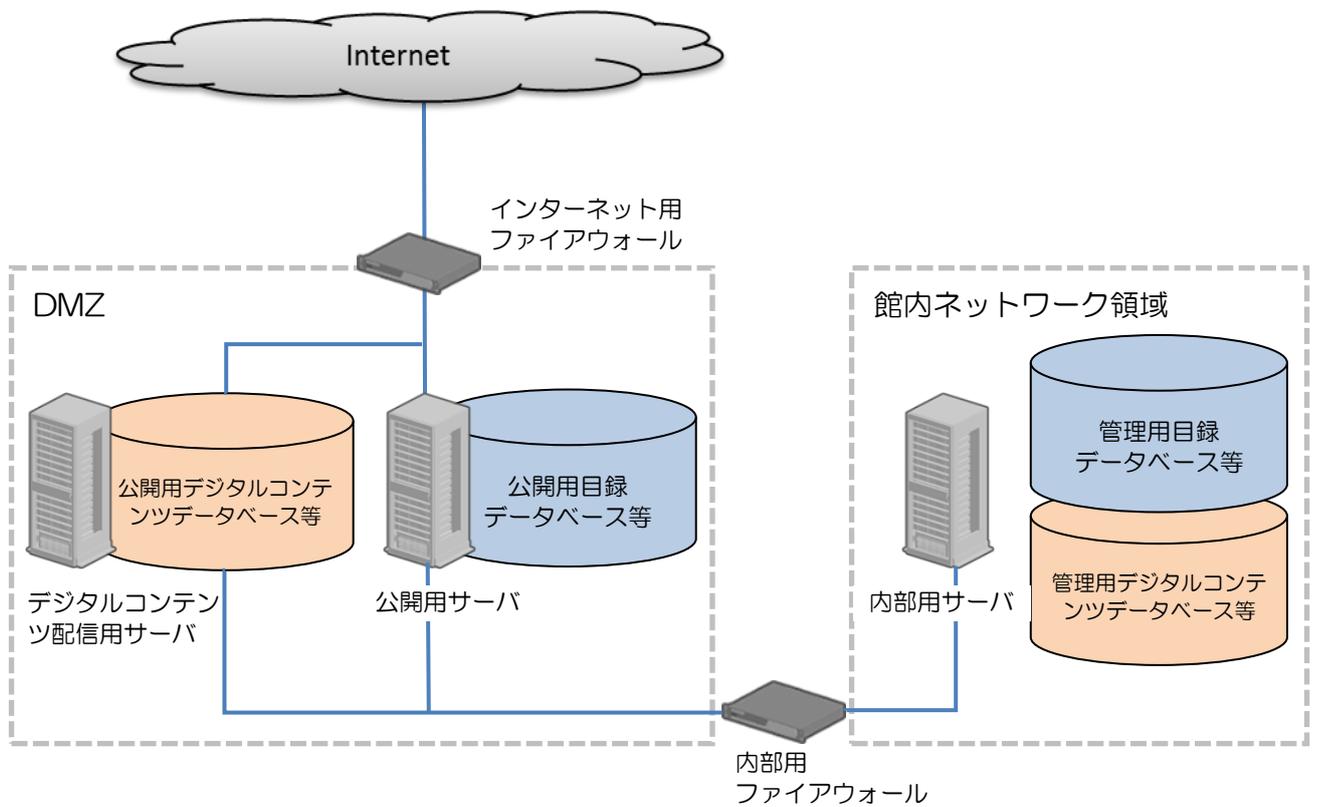


図 II-11 発展的な構成でのシステム構成図

(3) システム規模の検討

デジタルアーカイブ・システムを構成するサーバ機器の台数やディスク容量等、システムの規模については、各公文書館等で扱う目録データ及びデジタルコンテンツの件数・容量を基に検討を行う必要がある。

(4) インフラ基盤の検討

公文書館等における従来型のデジタルアーカイブ・システム構築では、サーバ機器やネットワーク機器等のハードウェアを自前で調達（購入、リース契約等）し、それらを公文書館等が占有等する施設や区画内に設置する形がとられていた。この従来型のインフラ基盤形態は、一般的にオンプレミスと呼ばれる。一方で近年では、クラウド技術の進展により、クラウドサービスを利用してデジタルアーカイブ・システムを構築できるようになっている。クラウドサービスを利用したシステム構築・運用では、初期構築コスト及びインフラ運用コストの軽減を期待することができるため、予算と人員に制約のある公文書館等においては選択肢の一つとなりうる。

なお、クラウドサービスでは、システムを構成する機器等が複数のサービス契約者によって共有される形がとられる。そのため、一般公開しない非公開データをクラウド上で管理する場合には、クラウドサービスのセキュリティ面のサービス品質について十分に確認を行う必要がある。

デジタルアーカイブ・システムにおけるオンプレミス形態及びクラウド形態のそれぞれの概要は、表 II-11のとおりである。

表 II-11 デジタルアーカイブ・システムにおけるインフラ基盤の例

インフラ基盤	概要
オンプレミス	公文書館等は、サーバ機器、ストレージ機器、ネットワーク機器等のハードウェアを自前で調達する。設置場所は公文書館等の施設内である場合と外部のデータセンター等の施設内である場合とがある。 なお、外部のデータセンターを利用したインフラ設置形態は、コロケーションとも呼ばれる。
クラウド	公文書館等は、クラウドサービスの提供者が有するサーバ機器、ストレージ機器、ネットワーク機器等のハードウェア資源の一部を、利用契約を結ぶことにより利用する。サービス形態によっては、ハードウェア資源だけでなく、OS やミドルウェアも利用することが可能である。 なお、一般的なクラウドサービスの形態としては、IaaS (Infrastructure as a Service)、PaaS (Platform as a Service)、SaaS (Software as a Service) の3種類があるが、本標準仕様書においてデジタルアーカイブ・システムへの利用を想定しているのは、IaaS と PaaS の2種類である。SaaS は事業者により提供される機能をサービスとして利用する形態であって、サービス利用者側で機能仕様等の追加や変更を行うことが基本的には出来ないものであることから、本標準仕様書では SaaS の利用を想定していない。

クラウドサービスを利用する場合は、自前でサーバ機器やネットワーク機器を設置する必要がなくなるため、設置に係る作業や工事が発生せず、その分のコスト削減が可能となる。また、インターネット回線につ

いてもサービス内で提供されることが一般的であることから、回線の用意に係る作業・工事が不要となり、その点のコストも抑制することができる。

クラウドサービスの選定にあたって確認・検討すべき点としては、課金方式（定額制/従量制）、バックアップサービスの有無及びバックアップ範囲、サービスレベルとして保証される稼働率（年間のサービス停止時間）、メンテナンスの事前通知タイミング等をあげることができる。

各公文書館等では上記（１）から（４）に示されるような検討を経たうえで、自館の業務の在り方やニーズ等に応じたシステムを導入する必要がある。

【標準仕様の考え方】

標準

システム導入は、業務の在り方やニーズ等に応じて、運用形態、機能と構成、データ量、インフラ基盤等を検討したうえで実施する。

2-8 セキュリティ

デジタルアーカイブ・システムに求められるセキュリティ対策の水準は、取り扱うデータの内容や実装機能により異なってくる。具体的には、一般公開可能な目録データやデジタルコンテンツのみを登録・管理する場合と、これに加えて業務利用限定の非公開データも管理する場合、さらには利用者情報などの個人情報等も管理する場合とでは、必要とされるセキュリティ対策の水準が大きく異なる。

以下に示すセキュリティ対策は、デジタルアーカイブ・システム上で一般公開データのみを管理する場合であっても対策を講じることが望ましい（あるいは必要とされる）ものである。

セキュリティ対策は何重にも講じることが理想的であるものの、システム導入・運用コストが高くつく可能性もあるため、各公文書館等では予算やシステム構成に応じた対策を選択的に導入する必要がある。

2-8-1 ウイルス対策

システムセキュリティ上の大きな脅威のひとつに、コンピュータウイルス（以下「ウイルス」とする）がある。デジタルアーカイブ・システムがウイルスに感染すると、データの消失や改ざん等の被害が生じ、利用者への安定したサービスを提供することができなくなる。また、感染したデジタルコンテンツをインターネットで公開することで、利用者の端末にもウイルスを拡散させる恐れがある。

このようなリスクを回避するためには、デジタルアーカイブ・システムにウイルス対策ソフトウェアを導入することが必要となる。

【標準仕様の考え方】

標準

ウイルス対策ソフトウェアを導入する。

2-8-2 外部攻撃への対策

デジタルアーカイブ・システムの運用にはインターネットへの接続が必須となる。インターネットと各公文書館等の内部ネットワーク（LAN）との間には、インターネットを経由した外部からの不正な通信を遮断するためのファイアウォールを導入する必要がある。

【標準仕様の考え方】

標準

ファイアウォールを導入する。

一方でファイアウォールにおいては、通信内容の検査は行われなため、Webアプリケーションの脆弱性を標的とした攻撃までは検出することができない。そこで近年では、WAF（Web Application Firewall）と呼ばれる仕組みの導入が増えてきている。WAFは、ハードウェアまたはソフトウェアとして導入する形と、クラウドサービスとして提供される形とがある。WAFについては、独立行政法人情報処理推進機構（IPA）がWAFの導入・運用事例を紹介した「Web Application Firewall（WAF）読本」を公開しており、システ

ムのセキュリティレベル向上に有益な情報が提供されている¹⁸。

また万が一、外部からの攻撃や不正侵入を受けた場合であっても、早期に発見し対処することで、被害を最小限にとどめることが可能である。外部攻撃の早期発見に有効な仕組みとしては、侵入検知システム（Intrusion Detection System: IDS）や改ざん検知ソフトウェア等がある。侵入検知システムとは、システムに対する侵入／侵害を検出・通知するシステムである。また、改ざん検知ソフトウェアについては、Webコンテンツのファイル更新を監視するもの、主要な攻撃パターンとの一致を検出するもの、WEBブラウザ上での動作から改ざんを検知するもの等、多くの種類がある。

2-8-3 不正アクセス対策

公的機関においては近年、Webサイトに対する信頼性を高めるため、公開情報の配信に際しても常に暗号化通信が用いられる傾向（常時SSL/TLS化）がある。このような近年の傾向も踏まえれば、公開データのみの取り扱いに限定されるデジタルアーカイブ・システムであっても、暗号化通信に対応することが望ましい。また、オンラインによる利用請求機能など利用者に個人情報等を入力させる機能を実装する場合には、通信の暗号化は必須となる。

現在、インターネットにおける通信上のなりすまし、改ざん、盗聴への対策として用いられる主な暗号化通信方式としては、HTTPSが普及している。

【標準仕様の考え方】

推奨	暗号化通信に対応する。
----	-------------

¹⁸ 独立行政法人 情報処理推進機構 セキュリティセンター. Web Application Firewall (WAF) 読本. 改訂第2版, 2011, <https://www.ipa.go.jp/files/000017312.pdf>, (参照 2017-12-25)

2-9 システム運用

2-9-1 運用・保守

デジタルアーカイブ・システムの運用・保守の具体的な内容としては、主に機器やソフトウェアの定期点検、バックアップ、障害時における対処、職員に対する操作研修や問い合わせへの対応等が考えられる。特にデジタルアーカイブ・システムのセキュリティレベルを維持するためには、OSやソフトウェア等の脆弱性についての情報収集を行い、システムに対する影響を確認のうえ、定期的にセキュリティ更新プログラムを適用することが必要となる。

デジタルアーカイブ・システムの規模が大きくなる場合には、運用・保守における負荷も増大することとなる。このような場合に円滑な運用を確保するためには、デジタルアーカイブ・システムの運用・保守に係る仕様を定め、外部委託を行うことも考えられる。

【標準仕様の考え方】

標準	運用・保守を適切に実施する。
標準	定期的にセキュリティ更新プログラムを適用する。

2-9-2 バックアップ機能

デジタルアーカイブ・システムにおいては定期的なバックアップを実施することで、システム障害が発生した場合でも、その後のリカバリを速やかに行うことが可能となる。採用するバックアップ方式やメディアについては、データ種別ごとの重要性、更新頻度、データ量、データ量増加予測を踏まえながら、検討を行う必要がある。

また広域災害に備えるに当たっては、データの重要性やコスト等を勘案のうえ、バックアップデータを遠隔地に保管することが望ましい。

デジタルアーカイブ・システムの構築・運用にクラウドサービスを利用する場合には、サービスメニューの一部（またはオプション）としてのデータバックアップサービスの有無を確認しておく必要がある。クラウドサービスの中には、遠隔地のデータセンター等へのバックアップサービスが提供されているケースもある。データバックアップサービスが提供されないクラウドサービスの場合や、データバックアップサービスを利用しない場合には、公文書館等で独自にバックアップデータの取得・保管等を行うことが必要となる。

【標準仕様の考え方】

標準	バックアップ取得機能を導入する。
推奨	遠隔地にバックアップデータを保管する。

2-9-3 システム監視機能

デジタルアーカイブ・システムが「いつでも」利用可能なものであるためには、利用者向けサービスの稼働

が安定している必要がある。システム異常の早期発見や迅速な復旧等を可能とするためには、システムの自動的な死活監視の仕組み、また、問題がある場合に公文書館等職員やシステム保守担当者へメール等により通知が行われる仕組みを用意しておくことが望ましい。

また異常の発生を未然に防ぐためには、サービス稼働に必要となるプロセスの状態や、サーバ機器のリソースの状況（CPU使用率、メモリ使用率、ディスク使用率等）を監視する仕組みを導入し、サービス異常の兆候、アクセス数やデータ量の増加によるシステムリソースの逼迫等を早期に発見可能とすることも有効と考えられる。

【標準仕様の考え方】

推奨

利用者向けサービスの稼働状況を監視する。

2-9-4 データ移行

公文書館等に目録データやデジタルコンテンツを扱う既存のシステムが存在する場合は、新たに構築するデジタルアーカイブ・システムへのデータ移行に際し様々な留意事項が存在する。具体的には、既存システム上の目録データ項目の移行先、データ変換の必要性の有無、既存システムから出力されるデータの形式、マスタデータの移行方法等について確認・検討を行い、データ移行時にデータの欠落や不整合が生じないようにする必要がある。特に既存システムで業務利用限定の非公開データが扱われている場合は、非公開データを含めた移行データ件数の確認、非公開項目の移行結果の検証、移行データの受け渡し方法に係るセキュリティ確保等も留意すべき点となる。

デジタルコンテンツは一般的に、目録データに比べてデータ容量が大きいことから、既存システムからのデータ抽出や新システムへの格納には時間を要することが考えられる。また、移行に際して目録データやデジタルコンテンツのデータ変換を行う場合は、変換処理にも時間を要する点に留意が必要である。

公文書館等によっては、目録データやデジタルコンテンツ以外のデータ（ホームページ掲載の情報や閲覧業務関連データ、アクセスログ等）の移行が行われることも想定されるが、その場合は、データの移行範囲や移行方法をあらかじめ明確にしておくことが必要となる。

デジタルアーカイブ・システムを一度構築した後、数年を経て次期システムへの更新を行う際にも、上記と同様の点に留意する必要がある。デジタルアーカイブ・システムの構築に際して、目録データやデジタルコンテンツを汎用的な形式で出力できるようにしておくこと、次期システムへのデータ移行を円滑に進めることが可能となる。

【標準仕様の考え方】

標準

既存システムからのデータ移行、次期システムへのデータ移行に留意する。

2-10 性能要件

デジタルアーカイブ・システムの構築に際しては、どの程度の速度で、どの程度のデータ量进行处理できるかといった性能目標の設定が見落とされがちである。

デジタルアーカイブ・システムの性能要件については、一般利用者向けのサービスと公文書館等職員が利用する業務機能の2つの観点から考えることができる。一般利用者向けのサービスについては、目録検索やデジタルコンテンツの閲覧がストレスなく利用できることが目安となると考えられる。

また、公文書館等職員による目録データやデジタルコンテンツの登録等の所要時間については、同時に利用する職員数等を考慮したうえでなお、業務に支障をきたさない範囲とすることが望ましい。

これら性能要件は、各公文書館等がデジタルアーカイブ・システムで公開する目録データ件数やデジタルコンテンツの容量、想定するアクセス数、業務内容により異なってくるものと考えられる。そのため、各公文書館等では、自館のデジタルアーカイブ・システムの運用及び利用状況（想定を含む）を適切に見積もったうえで性能要件を定めることが望ましい。

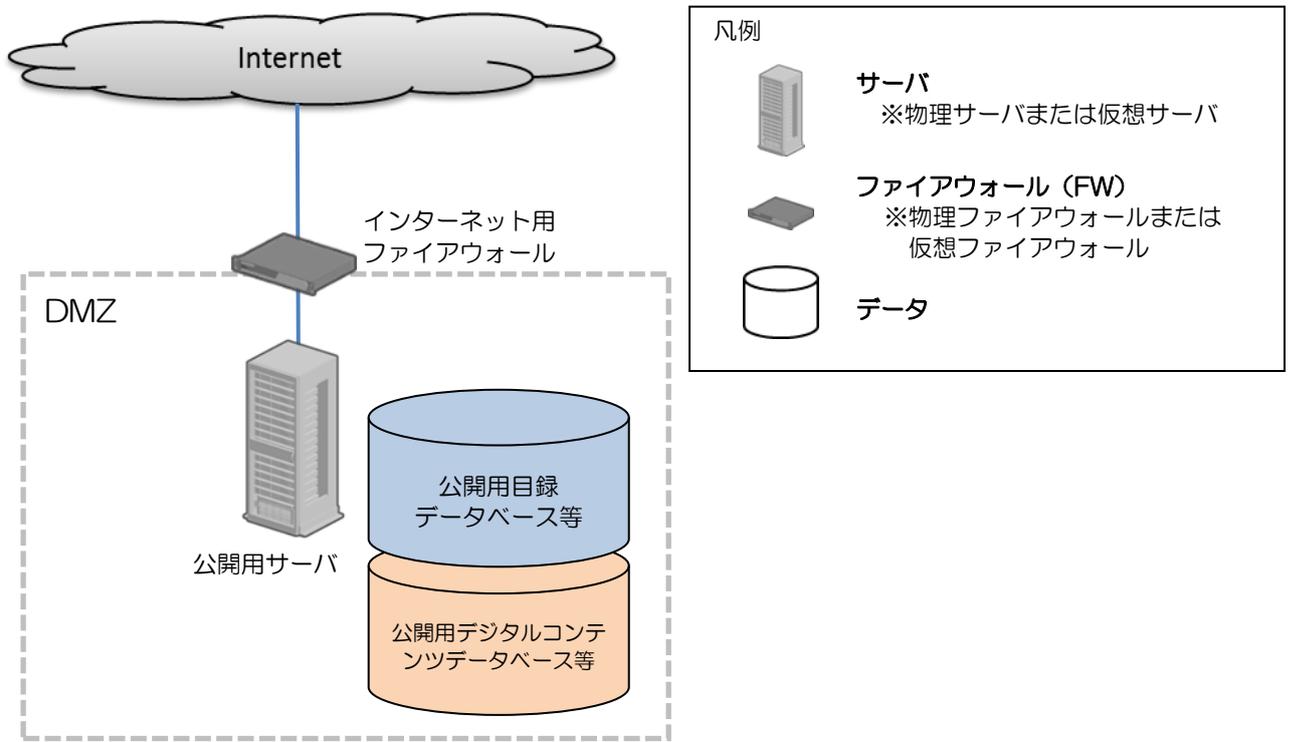
【標準仕様の考え方】

標準	目録データの検索、デジタルコンテンツの配信・閲覧に係る速度は、利用者が快適に利用できる水準とする。
標準	目録データ及びデジタルコンテンツの登録等に係る所要時間は、業務に支障をきたさない水準とする。

Ⅲ システム構成例

1 基本的な構成

デジタルアーカイブ・システムにおいて、公開データを扱う上で最小限必要と考えられる構成を基本的な構成として例示する。



DMZ:セキュリティ強化のため、インターネットと内部ネットワーク（館内ネットワーク）の間に設けられた、ファイアウォール等により隔離される領域。

図 Ⅲ-1 基本的な構成のシステム構成

1-1 対象データ

基本的な構成で取り扱うデータは以下の通りとする。

目録データ	一般公開	
デジタルコンテンツ	一般公開	閲覧用画像
		印刷用画像

1-2 機能構成

基本的な構成においては、以下の機能を想定する。

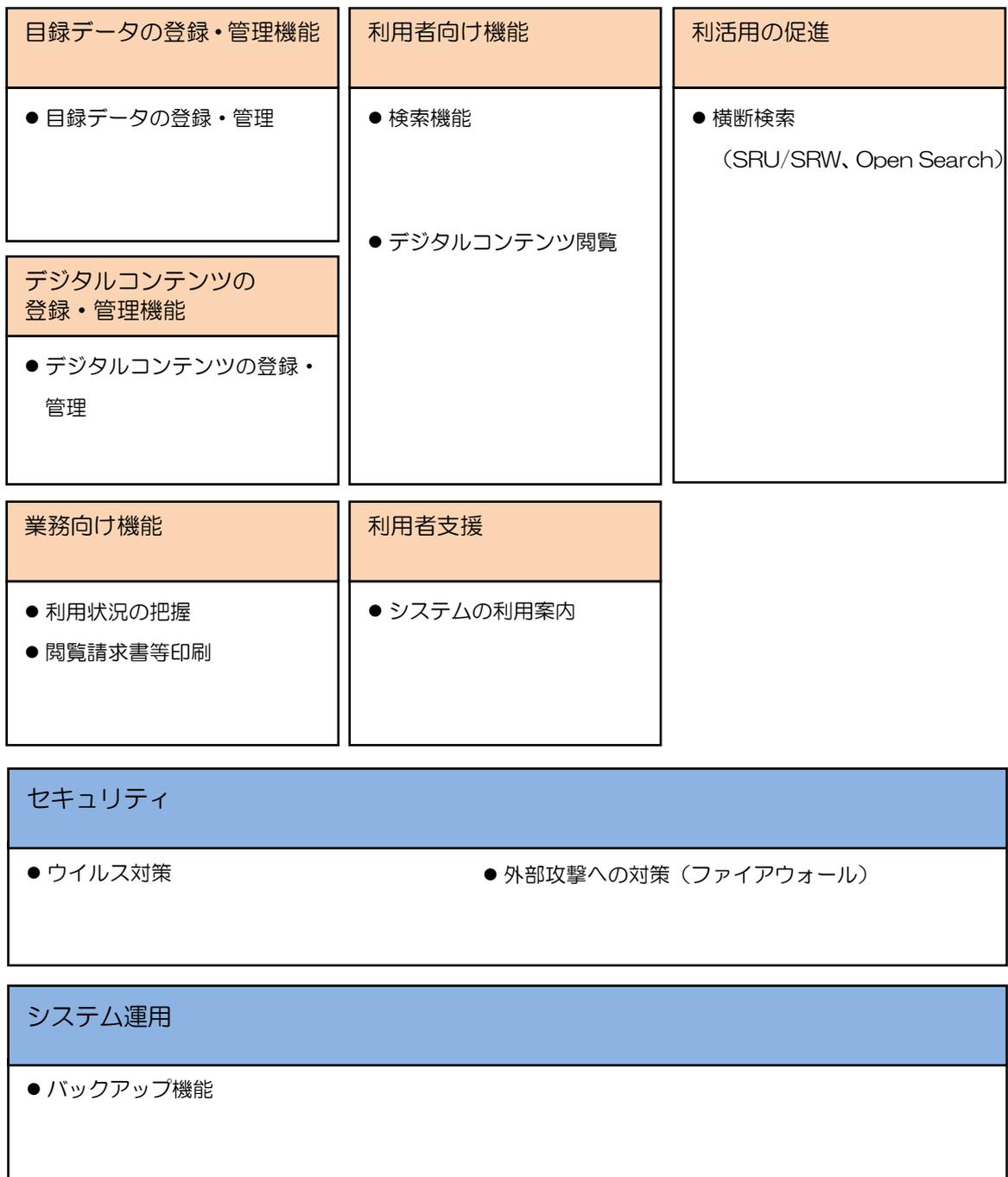


図 Ⅲ-2 基本的な構成における機能構成図

1-3 仕様書記載事項の例

本構成において想定される仕様書記載事項の一例を以下に示す。

項番	項目	仕様記載内容(案)	記載	備考		
1		目録データ・デジタルコンテンツの準備				
	1	目録データ				
	1	記述方式、電子符号化方式	本システムへ登録する目録データはCSV形式とする。	いずれか選択	汎用的な形式としては、CSV、TSVが考えられる。	
			本システムへ登録する目録データはTSV形式とする。			
			本システムへ登録する目録データはEAD形式とする。			XML形式の場合はEAD(最新版はEAD3)に準拠することが望ましい。
			本システムへ登録する目録データは表計算ソフトXXXバージョンXXの形式とする。			表計算ソフトの形式を指定する場合はバージョンも記載する。
	2	横断検索のために必要な最小限の目録データの項目	横断検索に使用する目録データの項目は別紙XXのとおりとする。	必須	別表においてDublin Coreとのマッピング案を示すことが望ましい。	
	3	階層構造の表現	目録データは階層構造を用いた管理ができること。	必須		
			階層構造に係る情報の管理ができること。	推奨	各階層に対してはタイトルや解説等の情報を付与することが望ましい。	
	4	文字コード	目録データの文字コードについては、文字集合はUCS-4とし、符号化形式はUTF-8またはUTF-16とする。	必須		
	2	デジタルコンテンツ				
	1	画像データ	公開用画像として、印刷用画像及び閲覧用画像を用意する。	任意	保存用形式はデジタルアーカイブ・システム上では扱わない。	
			公開用画像の画像形式はJPEG形式を取り扱えること。	いずれかまたは複数選択		
			公開用画像の画像形式はPDF形式を取り扱えること。	択		
			公開用画像の解像度はXXdpiを閲覧できること。	いずれかまたは複数選択		印刷用、閲覧用を分ける場合はそれぞれについて記載する。
			公開用画像(カラー)の色数は24bitフルカラ	択		

項番			項目	仕様記載内容（案）	記載	備考
				連付け情報、ファイル形式、ファイルサイズ等）が含まれること。		
				1.2 に示す形式のデジタルコンテンツと管理データにより、本システムのデジタルコンテンツを一括更新できること。		
				本システムに登録済みのデジタルコンテンツを一括削除できること。		
				本システムに登録済みのデジタルコンテンツの管理データを一括出力できること。		
				本システムに登録済みの管理データに対して条件を指定して検索を行い、結果を一覧表示できること。	必須	
				デジタルコンテンツ及び管理データを個別に登録できること。	必須	
				本システムに登録済みの管理データを個別に編集できること。また、編集した管理データを基にデジタルコンテンツのファイルを更新（差し替え）できること。		
				本システムに登録済みのデジタルコンテンツ及び管理データを個別に削除できること。		ユーザ権限設定において削除の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる。
3			利用者向け機能			
	1		検索機能			
		1	検索機能	任意の項目を対象とした、キーワード及び条件指定による検索ができること。	必須	
				項目を限定せずに、キーワードによる全文検索ができること。		
				複数の条件及びキーワードの組み合わせによる検索ができること。		
				目録データの階層構造表示から、指定した階層に属する目録データの一覧表示ができること。		
				目録データの階層表示から、指定した階層以下に属する目録データに限定した検索ができること。		
				検索結果を一覧形式で表示できること。	必須	
				検索結果一覧から、特定の目録データ項目による並び替え、条件を追加した絞り込み検索が行えること。		一般的な検索システムに含まれると考えられるため、必須とする。

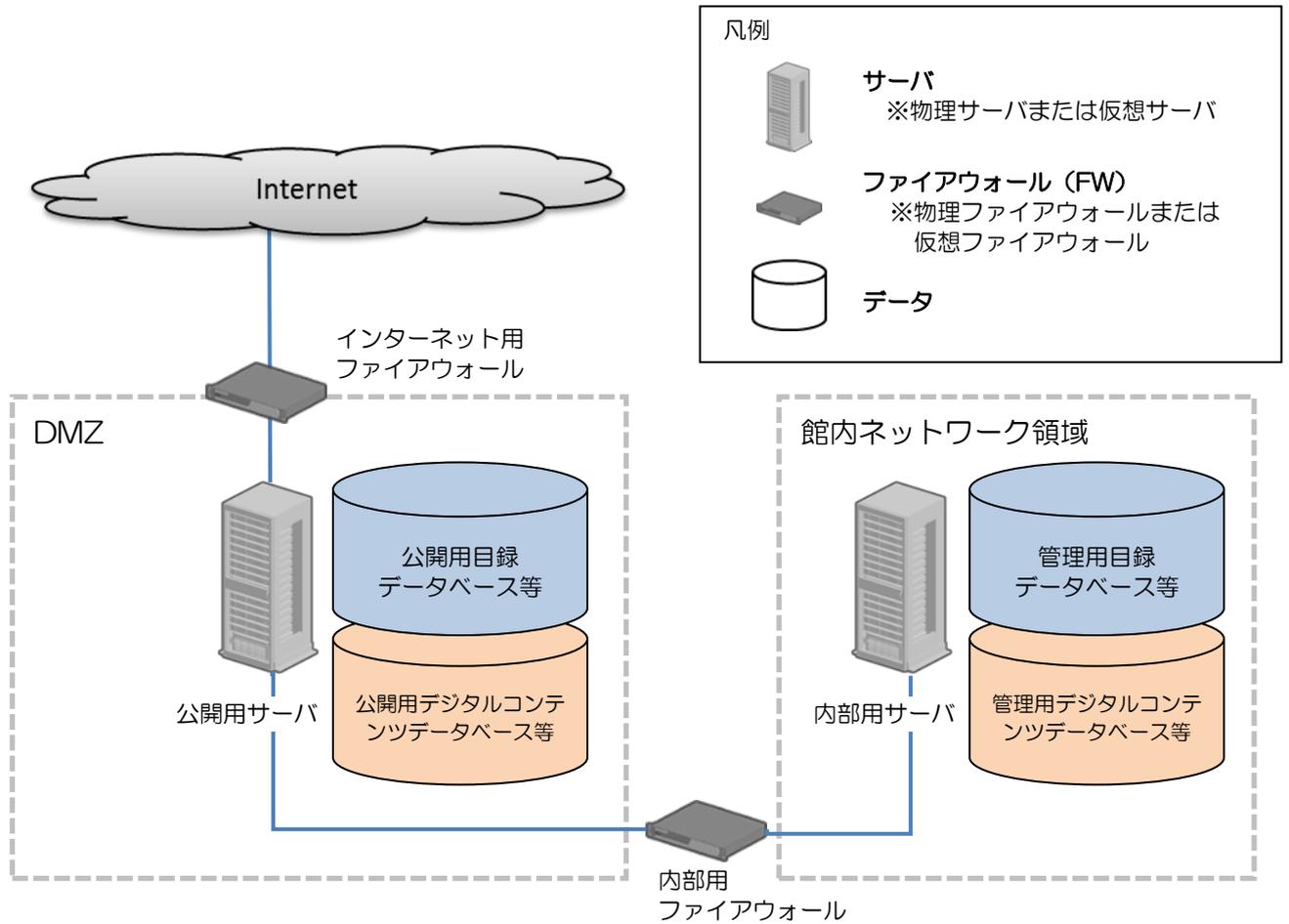
項番		項目	仕様記載内容（案）	記載	備考
			検索結果一覧から、指定された目録データの詳細を表示できること。		
	2	デジタルコンテンツ閲覧機能			
		1	デジタルコンテンツ閲覧機能	必須	
			検索結果の目録データに対応するデジタルコンテンツが存在する場合には、検索結果一覧上にデジタルコンテンツへのリンクを表示し、閲覧画面へ遷移可能とすること。		
			3.2.1に加え、目録データの詳細表示からもデジタルコンテンツへのリンクを表示し、閲覧画面へ遷移可能とすること。		
			画像データの閲覧方式は、利用者が特別なソフトウェアを必要としない方式とすること。	必須	「画像データの閲覧は、特別なソフトウェアを必要とせず、Webブラウザのみで可能なこと」等、具体的な表現も可能。
			画像データを閲覧するためのビューア機能を備えること。ビューアでは、直観的な操作により拡大、縮小、移動、ページ送りを行えること。	必須	
4		利活用の促進機能			
	1	横断検索機能	SRU（Search/Retrieve via URL）/SRW（Search/Retrieve Web Service）プロトコルに対応したサーバ機能を備えること。	いずれか選択	
			OpenSearchに対応したサーバ機能を備えること。		
5		利用者支援			
	1	システムの利用案内	本システムの操作方法、動作環境等に係る説明を含むヘルプページを用意し、利用者が分かりやすい場所に当該ページへのリンクを表示すること。	必須	
6		業務向け機能			
	1	利用状況の把握機能	本システムへのアクセス状況等をアクセスログファイルに記録すること。	必須	
			アクセスログファイルから、目録データの検索回数やデジタルコンテンツへのアクセス回数等の統計情報を取得できること。		
	2	閲覧業務関連機能	目録データの検索結果一覧から、選択された資料の閲覧請求書等を出力し、印刷できること。	必須	「閲覧請求書」は各公文書館等で使用する様式の名称に合わせる。 様式については仕様書別紙
			閲覧請求書等には選択された目録データの内容が自動的に出力されること。		

項番			項目	仕様記載内容（案）	記載	備考
						として提示することが望ましい。
7			システム環境			システム構成例を参考に、インフラ基盤を選定し別紙に記載する。
				本システムが稼働するインフラ基盤は別紙 XX のとおり。	いずれか選択	既存のインフラ基盤を利用する場合。
				本システムは別紙 XX に示す要件を満たすクラウドサービス上で稼働すること。		クラウドサービスを利用する場合。
				本システムは別紙 XX に示す要件を満たす機器により構築されたインフラ基盤上で稼働すること。		オンプレミスで機器を導入する場合。
8			セキュリティ			
	1		ウイルス対策	ウイルス対策ソフトを導入し、本システムに格納されるすべてのファイルの検疫を行うこと。	必須	
				ウイルス対策ソフトのパターン定義ファイルは、本システムの運用期間中、定期的に自動更新されること。		
	2		外部攻撃への対策	当館ネットワーク上に設置されたファイアウォールに、本システムへのアクセスを設定するために必要な情報を提供すること。	いずれか選択	既存ネットワークを利用する場合。
				ファイアウォールサービスを導入し、本システムへの不正・不要なアクセスを遮断すること。		クラウドサービスを利用する場合。
				ファイアウォールを導入し、本システムへの不正・不要なアクセスを遮断すること。		オンプレミスで機器を導入する場合。
9			システム運用			
	1		運用・保守	本システムは原則 24 時間 365 日（メンテナンス等による計画的停止を除く）稼働させること。	必須	
				本システムの OS、ミドルウェア、アプリケーションに対して年 X 回、セキュリティ更新プログラムを適用すること。ただし、緊急性の高い脆弱性が発見された場合は、適時セキュリティ更新プログラムを適用すること。		
	2		バックアップ機能	システムのバックアップを任意の時点で取得できること。	必須	セキュリティ更新プログラムの適用時等にシステムのバックアップを取得することで、問題発生時の切り戻しが可能となる。
				目録データ、デジタルコンテンツデータのバック		

項番	項目	仕様記載内容（案）	記載	備考
		アップを定期的に取り得ること。		
		目録データ、デジタルコンテンツデータのバックアップをX世代分保持できること。		保持する世代数に応じてバックアップメディアを用意する。
3	データ移行	既存システム上の目録データのすべてを本システムへ移行すること。 既存システム上の目録データ項目、目録データ件数、出力データ形式は別紙XXのとおり。また、別紙YYに従ってデータ変換を行うこと。	いずれかまたは複数選択	別紙YYにはデータ変換の仕様を記述する。
		既存システム上のデジタルコンテンツのすべてを本システムへ移行すること。 既存システム上のデジタルコンテンツ数、データ形式は別紙XXのとおり。また、別紙YYに従ってデータ変換を行うこと。		別紙YYにはデータ変換の仕様を記述する。
		別紙XXに示すデータを本システムへすべて移行すること。移行にあたっては、別紙YYに従ってデータ変換を行うこと。		目録データ、デジタルコンテンツ以外のデータ移行を含める場合。 別紙YYにはデータ変換の仕様を記述する
		本システムの運用終了時に、次期システムへのデータ移行を円滑に行うため、目録データ及びデジタルコンテンツデータは汎用的な形式により全件出力可能であること。	必須	
10	性能要件	本システム（登載される目録データ件数の目安：X件）のキーワード検索は、ネットワーク等の外部影響を受けない状態で、Y秒以内に完了することを目標とする。	必須	データ規模やアクセス数の想定を踏まえてレスポンス時間を決定する。
		本システム上のデジタルコンテンツ（1ファイルあたりの容量の目安：Xメガバイト）は、ネットワーク等の外部影響を受けない状態で、Y秒以内に表示されることを目標とする。		
		本システムの目録データ登録は1件あたりX秒以内、更新は1件あたりY秒以内に完了することを目標とする。最大同時アクセス数はZユーザとする。	必須	1件あたりの処理時間の目安を設定する。 デジタルコンテンツについてサムネイル生成やフォーマット変換を行う場合は、その処理時間も考慮する。
		本システムのデジタルコンテンツ登録は1件あたりX秒以内、更新は1件あたりY秒以内に完了することを目標とする。最大同時アクセス数はZユーザとする。		

2 標準的な構成

デジタルアーカイブ・システムにおいて、公開データ及び非公開データの両方を扱う上で最小限必要と考えられる機能からなる構成を標準的な構成として例示する。



DMZ:セキュリティ強化のため、インターネットと内部ネットワーク（館内ネットワーク）の間に設けられた、ファイアウォール等により隔離される領域。

図 Ⅲ-3 標準的な構成のシステム構成

2-1 対象データ

標準的な構成で取り扱うデータは以下の通りとする。

目録データ	一般公開	
	非公開	
デジタルコンテンツ	一般公開	閲覧用画像
		印刷用画像
	非公開	画像

2-2 機能構成

標準的な構成においては、以下の機能を想定する。

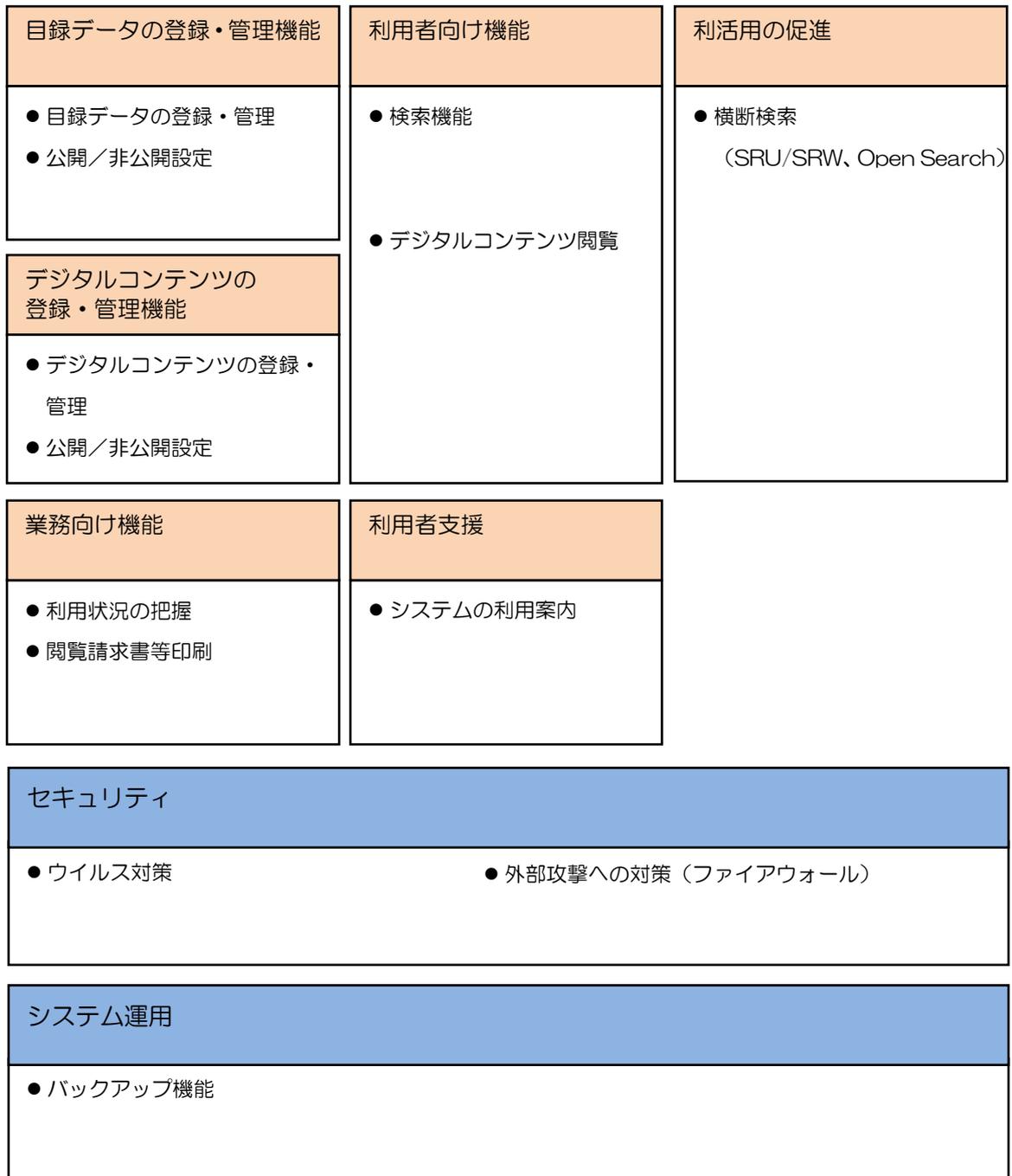


図 Ⅲ-4 標準的な構成における機能構成図

2-3 仕様書記載事項の例

本構成において想定される仕様書記載事項の一例を以下に示す。

項番	項目	仕様記載内容（案）	記載	備考	
1	目録データ・デジタルコンテンツの準備				
	1	目録データ			
	1	記述方式、電子符号化	本システムへ登録する目録データは CSV 形式とする。	いずれか選択	汎用的な形式としては、CSV、TSV が考えられる。
		本システムへ登録する目録データは TSV 形式とする。			
		本システムへ登録する目録データは EAD 形式とする。	XML 形式の場合は EAD（最新版は EAD 3）に準拠することが望ましい。		
		本システムへ登録する目録データは表計算ソフト XXX バージョン XX の形式とする。			
	2	横断検索のために必要な最小限の目録データの項目	横断検索に使用する目録データの項目は別紙 XX のとおりとする。	必須	別表において Dublin Core とのマッピング案を示すことが望ましい
	3	階層構造の表現	目録データは階層構造を用いた管理ができること。	必須	
			階層構造に係る情報の管理ができること。	推奨	各階層に対してはタイトルや解説等の情報を付与できることが望ましい。
	4	文字コード	目録データの文字コードについては、文字集合は UCS-4 とし、符号化形式は UTF-8 または UTF-16 とする。	必須	
	2	デジタルコンテンツ			
	1	画像データ	公開用画像として印刷用画像及び閲覧用画像を用意する。	任意	保存用形式はデジタルアーカイブ・システム上では扱わない。
		公開用画像の画像形式は JPEG 形式を取り扱えること。	いずれかまたは複数選択		
		公開用画像の画像形式は PDF 形式を取り扱えること。			
		公開用画像の解像度は XXdpi を閲覧できること。	いずれかまたは複数選択	印刷用、閲覧用を分ける場合はそれぞれについて記載する。	
		公開用画像（カラー）の色数は 24bit フルカラ			

項番		項目	仕様記載内容(案)	記載	備考
			一を閲覧できること。		スマートデバイスに対応する場合は、画面解像度に留意する。
			公開用画像(白黒)の色数は8bitグレースケールを閲覧できること。		
2		目録データ・デジタルコンテンツの登録と管理			
	1	目録データ			
		目録データの登録・管理機能	各種データを管理するためのRDBMSを備えること。	いずれかまたは複数選択	各館のニーズに応じた記載を行う。
			各種データを管理するためのデータベースシステムを備えること。		
			目録データを効率的に検索するための全文検索エンジンまたは同等の機能を備えること。		
			1.1に示す形式の目録データを本システムへ一括登録できること。	必須	ユーザ権限設定において一括登録等の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる
			1.1に示す形式の目録データにより、本システムに登録済みの目録データを一括更新できること。		
			本システムに登録済みの目録データを一括削除できること。		
			本システムに登録済みの目録データを、1.1に示す形式の目録データとして一括出力できること。		
			本システムに登録済みの目録データに対して条件を指定して検索を行い、結果を一覧表示できること。	必須	
			目録データを個別に登録できること。	必須	ユーザ権限設定において削除の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる。
			本システムに登録済みの目録データを個別に編集・更新できること。		
		本システムに登録済みの目録データを個別に削除できること。			
		目録データに対して公開/非公開の設定ができること。		ユーザ権限設定において公開/非公開設定の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することがで	

項番		項目	仕様記載内容（案）	記載	備考
					きる。
			一般公開する目録データと非公開の目録データは、格納領域を物理的に分離すること。また、それぞれの格納領域へのアクセスは適切に制御できること。		
	2	デジタルコンテンツ			
		1	デジタルコンテンツの登録・管理機能	必須	ユーザ権限設定において一括登録等の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる。
			1.2 に示す形式のデジタルコンテンツを、デジタルコンテンツの管理データに基づき本システムへ一括登録できること。管理データには本システム上でデジタルコンテンツを適切に管理するうえで必要な情報（ファイル名、目録データとの関連付け情報、ファイル形式、ファイルサイズ等）が含まれること。		
			1.2 に示す形式のデジタルコンテンツと管理データにより、本システムに登録済みの目録データを一括更新できること。		
			本システムに登録済みのデジタルコンテンツを一括削除できること。		
			本システムに登録済みのデジタルコンテンツの管理データを一括出力できること。		
			本システムに登録済みの管理データに対して条件を指定して検索を行い、結果を一覧として表示できること。	必須	
			デジタルコンテンツ及び管理データを個別に登録できること。	必須	
			本システムに登録済みの管理データを個別に編集できること。また、編集した管理データを基にデジタルコンテンツのファイルを更新（差し替え）できること。		
			本システムに登録済みのデジタルコンテンツ及び管理データを個別に削除できること。		ユーザ権限設定において削除の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる。
			デジタルコンテンツに対して公開/非公開の設定ができること。	必須	ユーザ権限設定において公開/非公開設定の実行可否についての設定を適切

項番		項目	仕様記載内容（案）	記載	備考	
					に行うことで、誤操作のリスクを軽減することができる。	
			一般公開するデジタルコンテンツと非公開のデジタルコンテンツは、格納領域を物理的に分離すること。また、それぞれの格納領域へのアクセスは適切に制御できること。			
			関連付けられた目録データの公開/非公開設定に連動して、デジタルコンテンツの公開/非公開設定が自動的に設定・更新されること。	推奨	目録データとデジタルコンテンツの公開/非公開設定に不整合が生じることを回避するため。	
3		利用者向け機能				
	1	検索機能				
		1	検索機能	任意の項目を対象とした、キーワード及び条件指定による検索ができること。	必須	
			項目を限定せずに、キーワードによる全文検索ができること。			
			複数の条件及びキーワードの組み合わせによる検索ができること。			
			目録データの階層構造表示から、指定した階層に属する目録データの一覧表示ができること。			
			目録データの階層構造表示から、指定した階層以下に属する目録データに限定した検索ができること。			
			検索結果を一覧形式で表示できること。	必須	一般的な検索システムに含まれると考えられるため、必須とする。	
			検索結果一覧から、特定の目録データ項目による並び替え、条件を追加した絞り込み検索が行えること。			
			検索結果一覧から、指定された目録データの詳細を表示できること。			
	2	デジタルコンテンツ閲覧機能				
		1	デジタルコンテンツ閲覧機能	検索結果の目録データに対応するデジタルコンテンツが存在する場合には、検索結果一覧上にデジタルコンテンツへのリンクを表示し、閲覧画面へ遷移可能とすること。	必須	
			3.2.1 に加え、目録データの詳細表示からもデジタルコンテンツへのリンクを表示し、閲覧画面へ遷移できること。			

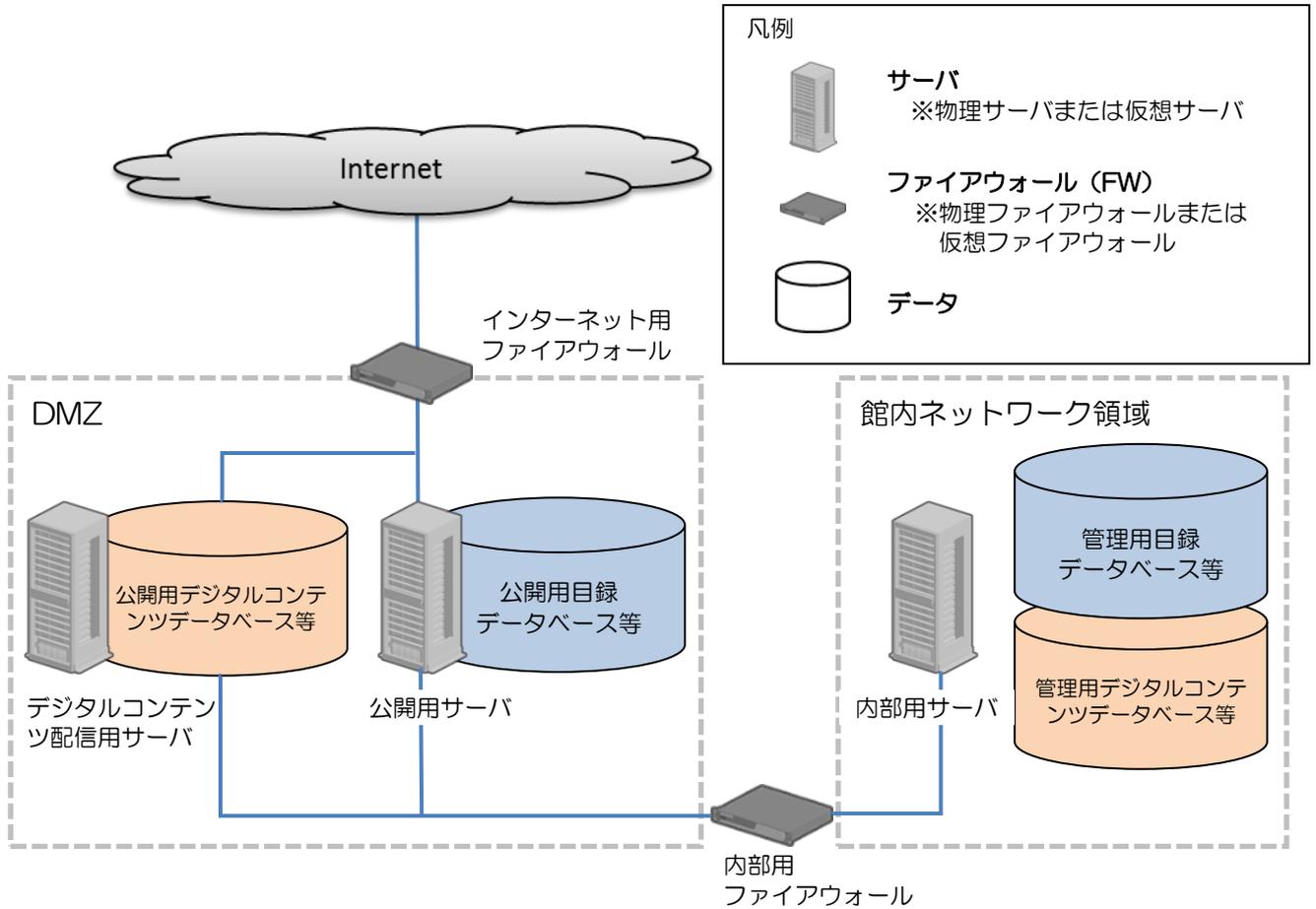
項番		項目	仕様記載内容（案）	記載	備考
			画像データの閲覧方式は、利用者が特別なソフトウェアを必要としない方式とすること。	必須	「画像データの閲覧は、特別なソフトウェアを必要とせず、Webブラウザのみで可能なこと」等、具体的な表現も可能。
			画像データを閲覧するためのビューア機能を備えること。ビューアでは、直観的な操作により拡大、縮小、移動、ページ送りが行えること。	必須	
4		利活用の促進			
	1	横断検索機能	SRU (Search/Retrieve via URL) /SRW (Search/Retrieve Web Service) プロトコルに対応したサーバ機能を備えること。	いずれか 選択	
			OpenSearch に対応したサーバ機能を備えること。		
5		利用者支援			
	1	システムの利用案内	本システムの操作方法、動作環境等に係る説明を含むヘルプページを用意し、利用者が分かりやすい場所に当該ページへのリンクを表示すること。	必須	
6		業務向け機能			
	1	利用状況の把握機能	本システムへのアクセス状況等をアクセスログファイルに記録すること。	必須	
			アクセスログファイルから、目録データの検索回数やデジタルコンテンツへのアクセス回数等の統計情報を取得できること。		
	2	閲覧業務関連機能	目録データの検索結果一覧から、選択された資料の閲覧請求書を出力し、印刷できること。	必須	「閲覧請求書」は各公文書館等で使用する様式の名称にあわせる。 様式については仕様書別紙として提示されることが望ましい。
			閲覧請求書には選択された目録データの内容が自動的に出力されること。		
7		システム環境			システム構成例を参考に、インフラ基盤を選定し別紙に記載する。
			本システムが稼働するインフラ基盤は別紙 XX のとおり。	いずれか 選択	既存のインフラ基盤を利用する場合。
			本システムは別紙 XX に示す要件を満たすクラウドサービス上で稼働すること。		クラウドサービスを利用する場合。
			本システムは別紙 XX に示す要件を満たす機器		オンプレミスで機器を導

項番		項目	仕様記載内容（案）	記載	備考
			により構築されたインフラ基盤上で稼働すること。		入する場合。
8		セキュリティ			
	1	ウイルス対策	ウイルス対策ソフトを導入し、本システムに格納されるすべてのファイルの検疫を行うこと。	必須	
			ウイルス対策ソフトのパターン定義ファイルは、本システムの運用期間中、定期的に自動更新されること。		
	2	外部攻撃への対策	当館ネットワーク上に設置されたファイアウォールに、本システムへのアクセスを設定するために必要な情報を提供すること。	いずれか 選択	既存ネットワークを利用する場合。
			ファイアウォールサービスを導入し、本システムへの不正・不要なアクセスを遮断すること。		クラウドサービスを利用する場合。
			ファイアウォールを導入し、本システムへの不正・不要なアクセスを遮断すること。		オンプレミスで機器を導入する場合。
9		システム運用			
	1	運用・保守	本システムは原則 24 時間 365 日（メンテナンス等による計画的停止を除く）稼働させること。	必須	
			本システムの OS、ミドルウェア、アプリケーションに対して年×回、セキュリティ更新プログラムを適用すること。ただし、緊急性の高い脆弱性が発見された場合は、適時セキュリティ更新プログラムを適用すること。		
	2	バックアップ機能	システムのバックアップを任意の時点で取得できること。	必須	セキュリティ更新プログラムの適用時等にシステムのバックアップを取得することで、問題発生時の切り戻しが可能となる。
			目録データ、デジタルコンテンツデータのバックアップを定期的に取得できること。		
			目録データ、デジタルコンテンツデータのバックアップを X 世代分保持できること。		保持する世代数に応じてバックアップメディアを用意する。
	3	データ移行	既存システム上の目録データのすべてを本システムへ移行すること。既存システムの目録データ項目、目録データ件数、出力データ形式は別紙 XX のとおり。また、別紙 YY に従ってデータ変換を行うこと。	いずれか または複数 選択	別紙 YY にはデータ変換の仕様を記述する。

項番		項目	仕様記載内容（案）	記載	備考
			既存システム上のデジタルコンテンツのすべてを本システムへ移行すること。既存システムのデジタルコンテンツ数、データ形式は別紙 XX のとおり。また、別紙 YY に従ってデータ変換を行うこと。		別紙 YY にはデータ変換の仕様を記述する。
			別紙 XX に示すデータを本システムへすべて移行すること。移行にあたっては、別紙 YY に従ってデータ変換を行うこと。		目録データ、デジタルコンテンツ以外のデータ移行を含める場合。 別紙 YY にはデータ変換の仕様を記述する。
			本システムの運用終了時に、次期システムへのデータ移行を円滑に行うため、目録データ及びデジタルコンテンツデータは汎用的な形式により全件出力が可能であること。	必須	
10		性能要件	本システム（登録される目録データ件数の目安：X 件）のキーワード検索は、ネットワーク等の外部影響を受けない状態で、Y 秒以内に完了することを目標とする。	必須	データ規模やアクセス数の想定を踏まえてレスポンス時間を決定する。
			本システム上のデジタルコンテンツ（1 ファイルあたりの容量の目安：X メガバイト）は、ネットワーク等の外部影響を受けない状態で、Y 秒以内に表示されることを目標とする。		
			本システムの目録データ登録は 1 件あたり X 秒以内、更新は 1 件あたり Y 秒以内に完了することを目標とする。最大同時アクセス数は Z ユーザとする。	必須	1 件あたりの処理時間の目安を設定する。 デジタルコンテンツについてサムネイル生成やフォーマット変換を行う場合は、その処理時間も考慮する。
			本システムのデジタルコンテンツ登録は 1 件あたり X 秒以内、更新は 1 件あたり Y 秒以内に完了することを目標とする。最大同時アクセス数は Z ユーザとする。		

3 発展的な構成

デジタルアーカイブ・システムにおいて、公開データ及び非公開データの両方を扱い、推奨機能を含む構成を発展的な構成として例示する。



DMZ:セキュリティ強化のため、インターネットと内部ネットワーク（館内ネットワーク）の間に設けられた、ファイアウォール等により隔離される領域。

図 Ⅲ-5 発展的な構成のシステム構成

3-1 対象データ

発展的な構成で取り扱うデータは以下の通りとする。

目録データ	一般公開	
	非公開	
デジタルコンテンツ	一般公開	閲覧用画像（JPEG 2000含む）
		印刷用画像
		動画
		音声
	非公開	画像（JPEG 2000含む）
		動画
音声		

3-2 機能構成

発展的な構成においては、以下の機能を想定する。

凡例：下線のうち、実線は推奨機能、点線は推奨機能以外の付加的価値のある機能を示す。

<p>目録データの登録・管理機能</p>	<p>利用者向け機能</p>	<p>利活用の促進</p>
<ul style="list-style-type: none"> ● 目録データの登録・管理 ● 公開／非公開設定 ● <u>外字の扱い</u> 	<ul style="list-style-type: none"> ● 検索機能 ● <u>検索支援機能（辞書、ファセット、異体字）</u> ● デジタルコンテンツ閲覧 ● <u>サムネイル表示</u> ● <u>ダウンロード機能</u> ● <u>高機能画像配信機能（IIF等）</u> ● <u>動画・音声配信機能</u> 	<ul style="list-style-type: none"> ● 横断検索 (SRU/SRW、Open Search) ● <u>一括提供（OAI-PMH）</u> ● <u>外部提供インターフェース（Web API）</u> ● <u>永続的アクセス</u> ● <u>ライセンス表示</u> ● <u>SNS連携</u>
<p>デジタルコンテンツの登録・管理機能</p>		
<ul style="list-style-type: none"> ● デジタルコンテンツの登録・管理 ● 公開／非公開設定 		
<p>業務向け機能</p>	<p>利用者支援</p>	<p>ユーザ・インターフェース</p>
<ul style="list-style-type: none"> ● 利用状況の把握 ● 閲覧請求書等印刷 	<ul style="list-style-type: none"> ● システムの利用案内 ● <u>レファレンス機能</u> 	<ul style="list-style-type: none"> ● <u>アクセシビリティ</u> ● <u>スマートデバイス対応</u>
<p>セキュリティ</p>		
<ul style="list-style-type: none"> ● ウイルス対策 ● <u>不正アクセス対策（暗号化通信）</u> ● 外部攻撃への対策（ファイアウォール） ● <u>外部攻撃への対策（WAF、侵入検知）</u> 		
<p>システム運用</p>		
<ul style="list-style-type: none"> ● バックアップ機能 ● <u>遠隔地バックアップ</u> ● <u>システム監視機能</u> 		

図 Ⅲ-6 発展的な構成における機能構成図

3-3 仕様書記載事項の例

本構成において想定される仕様書記載事項の一例を以下に示す。

項番		項目	仕様記載内容（案）	記載	備考	
1		目録データ・デジタルコンテンツの準備				
	1	目録データ				
	1	記述方式、電子符号化	本システムへ登録する目録データは CSV 形式とする。	いずれか 選択	汎用的な形式としては CSV、TSV が考えられる。	
			本システムへ登録する目録データは TSV 形式とする。			
			本システムへ登録する目録データは EAD 形式とする。			XML 形式の場合は EAD（最新版は EAD 3）に準拠することが望ましい。
			本システムへ登録する目録データは表計算ソフト XXX バージョン XX の形式とする。			表計算ソフトの形式を指定する場合はバージョンも記載する。
	2	横断検索のために必要な最小限の目録データの項目	横断検索に使用する目録データの項目は別紙 XX のとおりとする。	必須	別表において Dublin Core とのマッピング案を示すことが望ましい。	
	3	階層構造の表現	目録データは階層構造を用いた管理ができること。	必須		
			階層構造に係る情報の管理ができること。	推奨	各階層に対してはタイトルや解説等の情報を付与できることが望ましい。	
	4	文字コード	目録データの文字コードについては、文字集合は UCS-4 とし、符号化形式は UTF-8 または UTF-16 とする。	必須		
	5	外字の扱い	外字は「文字図形番号」（情報処理学会試行標準 IPSJ—TS0002：2014）を用い、EGIX で記述する。	推奨	既存の目録データ内の外字について、事前に文字図形番号による表現に変更するか、移行時に変換するかを検討する必要がある。	
	2	デジタルコンテンツ				
	1	画像データ	公開用画像として、印刷用画像及び閲覧用画像を用意する。	任意	保存用形式はデジタルアーカイブ・システム上では扱わない。	
			公開用画像の画像形式は JPEG 形式を取り扱えること。	いずれか または複		

項番		項目	仕様記載内容（案）	記載	備考	
			公開用画像の画像形式は PDF 形式を取り扱えること。	数選択		
			公開用画像の画像形式は JPEG 2000 形式を取り扱えること。			
			公開用画像の解像度は XXdpi を閲覧できること。	いずれか または複	印刷用、閲覧用を分ける場合はそれぞれについて記載する。 スマートデバイスに対応する場合は、画面解像度に留意する。	
			公開用画像（カラー）の色数は 24bit フルカラーを閲覧できること。	数選択		
			公開用画像（白黒）の色数は 8bit グレースケールを閲覧できること。			
	2	動画・音声データ	公開用の動画データの形式は MP4 形式を取り扱えること。	いずれか または複		動画・音声を扱う場合、必要に応じて記載する。
			公開用の音声データの形式は MP3 形式を取り扱えること。	数選択		
2		目録データ・デジタルコンテンツの登録と管理				
	1	目録データ				
	1	目録データの登録・管理機能	各種データを管理するための RDBMS を備えること。	いずれか または複	各館のニーズに応じた記載を行う。	
			各種データを管理するためのデータベースシステムを備えること。	数選択		
			目録データを効率的に検索するための全文検索エンジンまたは同等の機能を備えること。			
			1.1 に示す形式の目録データを本システムへ一括登録できること。	必須	ユーザ権限設定において一括登録等の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる。	
			1.1 に示す形式の目録データにより、本システムに登録済みの目録データを一括更新できること。			
			本システムに登録済みの目録データを一括削除できること。			
			本システムに登録済みの目録データを、1.1 に示す形式の目録データとして一括出力できること。			
			本システムに登録済みの目録データに対して条件を指定して検索を行い、結果を一覧として表示できること。	必須		
			目録データを個別に登録できること。	必須		
			本システムに登録済みの目録データを個別に編			

項番		項目	仕様記載内容（案）	記載	備考
			集・更新できること。		
			本システムに登録済みの目録データを個別に削除できること。		ユーザ権限設定において削除の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる。
			目録データに対して公開/非公開の設定ができること。	必須	ユーザ権限設定において公開/非公開設定の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる。
			一般公開する目録データと非公開の目録データは、格納領域を物理的に分離すること。また、それぞれの格納領域へのアクセスは適切に制御できること。		
	2	デジタルコンテンツ			
	1	デジタルコンテンツの登録・管理機能	1.2 に示す形式のデジタルコンテンツを、デジタルコンテンツの管理データに基づき本システムへ一括登録できること。管理データには本システム上でデジタルコンテンツを適切に管理するうえで必要な情報（ファイル名、目録データとの関連付け情報、ファイル形式、ファイルサイズ等）が含まれること。	必須	
			1.2 に示す形式のデジタルコンテンツと管理データにより、本システムに登録済みの目録データを一括更新できること。		
			本システムに登録済みのデジタルコンテンツの管理データを一括出力できること。		
			本システムに登録済みの管理データに対して条件を指定して検索を行い、結果を一覧として表示できること。	必須	
			デジタルコンテンツ及び管理データを個別に登録できること。	必須	
			本システムに登録済みの管理データを個別に編集できること。また、編集した管理データを基にデジタルコンテンツのファイルを更新（差し替え）できること。		

項番		項目	仕様記載内容（案）	記載	備考
			本システムに登録済みのデジタルコンテンツ及び管理データを個別に削除できること。		ユーザ権限設定において削除の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる。
			画像または動画のサムネイルを登録できること。	いずれか または複数選択	サムネイルを使う場合に記載する。
			画像または動画の登録に際して、サムネイルを自動的に生成・登録できること。		
			デジタルコンテンツに対して公開/非公開の設定ができること。	必須	ユーザ権限設定において公開/非公開設定の実行可否についての設定を適切に行うことで、誤操作のリスクを軽減することができる。
			一般公開するデジタルコンテンツと非公開のデジタルコンテンツは、格納領域を物理的に分離すること。また、それぞれの格納領域へのアクセスは適切に制御できること。		
			関連付けられた目録データの公開/非公開設定に連動して、デジタルコンテンツの公開/非公開設定が自動的に設定・更新されること。		
3		利用者向け機能			
	1	検索機能			
		1 検索機能	任意の項目を対象とした、キーワード及び条件指定による検索ができること。	必須	
			項目を限定せずに、キーワードによる全文検索ができること。		
			複数の条件及びキーワードの組み合わせによる検索ができること。		
			目録データの階層構造表示から、指定した階層に属する目録データの一覧表示ができること。		
			目録データの階層構造表示から、指定した階層以下に属する目録データに限定した検索ができること。		
			検索結果を一覧形式で表示できること。		
			検索結果一覧から、特定の目録データ項目による並び替え、条件を追加した絞込み検索が行え		

項番			項目	仕様記載内容（案）	記載	備考
				ること。		
				検索結果一覧から、指定された目録データの詳細を表示できること。		
		2	検索支援機能	辞書による同義語・関連語等の検索が行えること。辞書の登録・更新・削除は権限を付与されたユーザのみが実行できること。	いずれか または複数選択	
				指定した項目によるファセット検索が行えること。		
				異体字検索が行えること。		ひらがな/カタカナ、英数字の全角/半角、表記のゆれ等、具体的に記載する必要があるれば記載する。
	2		デジタルコンテンツ閲覧機能			
		1	デジタルコンテンツ閲覧機能	検索結果の目録データに対応するデジタルコンテンツが存在する場合には、検索結果一覧上にデジタルコンテンツへのリンクを表示し、閲覧画面へ遷移可能とすること。	必須	
				3.2.1 に加えて、目録データの詳細表示からもデジタルコンテンツへのリンクを表示し、閲覧画面へ遷移可能とすること。		
				画像データの閲覧は、利用者が特別なソフトウェアを必要としない方式とすること。	必須	「画像データの閲覧は、特別なソフトウェアを必要とせず、Web ブラウザのみで可能なこと」等、具体的な表現も可能
				画像データを閲覧するためのビューア機能を備えること。ビューアでは、直観的な操作により拡大、縮小、移動、ページ送りが行えること。	必須	
				目録データの検索結果一覧及び詳細表示では、目録データに関連付けられたデジタルコンテンツのサムネイル画像を表示できること。	必須	サムネイルを使う場合は必須とする。
				デジタルコンテンツの閲覧画面から、当該デジタルコンテンツを個別及び一括でダウンロード可能とすること。	推奨	
		2	画像配信機能	III F に基づく画像配信機能を備えること。	推奨	
		3	動画・音声配信機能	動画・音声データをストリーミング方式により配信できること。	推奨	

項番		項目	仕様記載内容（案）	記載	備考
	3	ユーザ・インターフェース			
	1	アクセシビリティ	一般利用者向けのユーザ・インターフェースは、総務省の「みんなの公共サイト運用ガイドライン」を参考とすること。	いずれか 選択	自治体等のアクセシビリティガイドラインが存在しない場合。
			一般利用者向けのユーザ・インターフェースは、「〇〇アクセシビリティガイドライン」を参考とすること。		自治体等のアクセシビリティガイドラインが存在する場合。
	2	スマートデバイスのためのユーザ・インターフェース	スマートデバイス（スマートフォン、タブレット）用に最適化されたユーザ・インターフェースを用意すること。利用者のデバイスを判別し、ユーザ・インターフェースの切り替えを自動的に行うこと。	推奨	
4		利活用の促進			
	1	横断検索機能	SRU（Search/Retrieve via URL）/SRW（Search/Retrieve Web Service）プロトコルに対応したサーバ機能を備えること。	いずれか 選択	
			OpenSearchに対応したサーバ機能を備えること。		
	2	一括提供機能	OAI-PMH（Open Archives Initiative Protocol for Metadata Harvesting）に対応したサーバ（リポジトリ）機能を備えること。	推奨	
	3	外部提供インターフェース	SPARQL プロトコルに対応したサーバ機能を備え、利用者が Web API として利用できること。	いずれか 選択	
			目録データ、デジタルコンテンツへアクセスするための Web API を備えること。		
			目録データの詳細表示やデジタルコンテンツの閲覧画面上に、SNS へのリンクを表示すること。	任意	連携する SNS についてはサービスの利用者数等を勘案して決定する。
	4	永続的アクセス	個別の目録データ及びデジタルコンテンツへアクセスするための固定の URL を設定できること。	いずれか または複数 選択	
			既存システムの目録データ及びデジタルコンテンツへアクセスするための URL 体系を維持すること。URL を変更する場合は一定のルールにより新しい URL を生成できること。		既存システムからの移行後も URL を維持する場合。
	5	ライセンス表示	デジタルコンテンツの利用条件について記述したページを用意し、当該ページへのリンクをデ	いずれか 選択	デジタルコンテンツ全体に対して同一の利用条件

項番		項目	仕様記載内容（案）	記載	備考
			デジタルコンテンツ閲覧画面に表示すること。		を一律に設定する場合。
			デジタルコンテンツの管理データとして設定された利用条件に従い、デジタルコンテンツ閲覧画面にパブリック・ドメイン・マークまたはクリエイティブ・コモンズ・ライセンスの表示を行うこと。		個々のデジタルコンテンツに対して異なる利用条件を設定する場合。
5		利用者支援			
	1	システムの利用案内	本システムの操作方法、動作環境等に係る説明を含むヘルプページを用意し、利用者が分かりやすい場所に当該ページへのリンクを表示すること。	必須	
			資料（原本）の閲覧手続き等を案内するページを用意し、利用者が分かりやすい場所に当該ページへのリンクを表示すること。	いずれか 選択	
			利用者が分かりやすい場所に、〇〇ホームページにおいて別途提供される利用案内ページへのリンクを表示すること。		既存のページを使用する場合の例。
	2	レファレンス機能	利用者からの目録データやデジタルコンテンツに関するレファレンスを受け付けるための入力フォームを用意し、利用者が分かりやすい場所に当該フォームへのリンクを表示すること。	推奨	
6		業務向け機能			
	1	利用状況の把握機能	本システムへのアクセス状況等をアクセスログファイルに記録すること。	必須	
			アクセスログファイルから、目録データの検索回数やデジタルコンテンツへのアクセス回数等の統計情報を取得できること。		
	2	閲覧業務関連機能	目録データの検索結果一覧から、選択された資料の閲覧請求書を出力し、印刷できること。	必須	「閲覧請求書」は各公文書館等で使用する様式の名称に合わせる。 様式については仕様書別紙として提示することが望ましい。
			閲覧請求書には選択された目録データの内容が自動的に出力されること。		
7		システム環境			システム構成例を参考に、インフラ基盤を選定し別紙に記載する。
			本システムが稼働するインフラ基盤は別紙 XX のとおり。	いずれか 選択	既存のインフラ基盤を利用する場合。

項番		項目	仕様記載内容（案）	記載	備考
			本システムは別紙 XX に示す要件を満たすクラウドサービス上で稼働すること。		クラウドサービスを利用する場合。
			本システムは別紙 XX に示す要件を満たす機器により構築されたインフラ基盤上で稼働すること。		オンプレミスで機器を導入する場合。
8		セキュリティ			
	1	ウイルス対策	ウイルス対策ソフトを導入し、本システムに格納されるすべてのファイルの検疫を行うこと。	必須	
			ウイルス対策ソフトのパターン定義ファイルは、本システムの運用期間中、定期的に自動更新されること。		
	2	外部攻撃への対策	当館ネットワーク上に設置されたファイアウォールに、本システムへのアクセスを設定するために必要な情報を提供すること。	いずれか 選択	既存ネットワークを利用する場合。
			ファイアウォールサービスを導入し、本システムへの不正・不要なアクセスを遮断すること。		クラウドサービスを利用する場合。
			ファイアウォールを導入し、本システムへの不正・不要なアクセスを遮断すること。		オンプレミスで機器を導入する場合。
			Web Application Firewall を導入し、本システムへの攻撃を意図したアクセスを遮断すること。	任意	クラウドでサービスとして提供されている場合もある。
			改ざん検知装置を導入し、本システムの Web ページ、目録データ、デジタルコンテンツへの改ざんが発生した場合にシステム管理者へ通知すること。	任意	
	3	不正アクセス対策	TLS 等の暗号化通信プロトコルにより情報発信を行えること。	推奨	
9		システム運用			
	1	運用・保守	本システムは原則 24 時間 365 日（メンテナンス等による計画的停止を除く）稼働させること。	必須	
			本システムの OS、ミドルウェア、アプリケーションに対して年 X 回、セキュリティ更新プログラムを適用すること。ただし、緊急性の高い脆弱性が発見された場合は、適時セキュリティ更新プログラムを適用すること。		
	2	バックアップ機能	システムのバックアップを任意の時点で取得できること。	必須	セキュリティ更新プログラム適用前等にシステムのバックアップを取得することで、問題発生時の切

項番		項目	仕様記載内容（案）	記載	備考
					り戻しが可能となる。
			目録データ、デジタルコンテンツデータのバックアップを定期的に取得できること。		
			目録データ、デジタルコンテンツデータのバックアップをX世代分保持できること。		保持する世代数に応じてバックアップメディアを用意する。
			システムバックアップ、データバックアップを格納したメディアを遠隔地に保管すること。	推奨	
	3	システム監視機能	本システムの一般利用者向けページ（URL）に対して、一定間隔で死活監視を行うこと。	いずれか または複数選択	対象は一般利用者向けサービスに限定する。
			本システムを構成するサーバ機器のリソース使用状況（CPU使用率、メモリ使用率、ディスク使用率等）を監視すること。監視結果に異常が見つかった場合、システム管理者へメールで通知すること。		
	4	データ移行	既存システムの目録データのすべてを本システムに移行すること。既存システムの目録データ項目、目録データ件数、出力データ形式は別紙XXのとおり。また、別紙YYに従ってデータ変換を行うこと。	いずれか または複数選択	別紙YYにはデータ変換の仕様を記述する。
			既存システムのデジタルコンテンツのすべてを本システムに移行すること。既存システムのデジタルコンテンツ数、データ形式は別紙XXのとおり。また、別紙YYに従ってデータ変換を行うこと		別紙YYにはデータ変換の仕様を記述する。
			別紙XXに示すデータを本システムへすべて移行すること。移行にあたっては、別紙YYに従ってデータ変換を行うこと。		目録データ、デジタルコンテンツ以外のデータ移行を含める場合。 別紙YYにはデータ変換の仕様を記述する。
			本システムの運用終了時に、次期システムへのデータ移行を円滑に行うため、目録データ及びデジタルコンテンツデータは汎用的な形式により全件出力が可能であること。	必須	次期システムへの移行を踏まえて記載する。
10		性能要件	本システム（搭載される目録データ件数の目安：X件）のキーワード検索は、ネットワーク等の外部影響を受けない状態で、Y秒以内に完了することを目指す。	必須	データ規模やアクセス数の想定を踏まえてレスポンス時間を決定する。

項番			項目	仕様記載内容（案）	記載	備考
				本システム上のデジタルコンテンツ（1 ファイルあたりの容量の目安：X メガバイト）閲覧は、ネットワーク等の外部影響を受けない状態で、Y 秒以内に表示されることを目標とする。		
				本システムの目録データ登録は 1 件あたり X 秒以内、更新は 1 件あたり Y 秒以内に完了することを目標とする。最大同時アクセス数は Z ユーザとする。	必須	1 件あたりの処理時間の目安を設定する。 デジタルコンテンツについてサムネイル生成やフォーマット変換を行う場合は、その処理時間も考慮する。
				本システムのデジタルコンテンツ登録は 1 件あたり X 秒以内、更新は 1 件あたり Y 秒以内に完了することを目標とする。最大同時アクセス数は Z ユーザとする。		