 DAX40-20 【課題テーマ】今後の図書館サービスシステムの構築を考える

# 改版履歴

## 2020年3月12日 DAX40-01の詳細を分冊化

# **はじめに**

## 今後10年を見据えると、 図書館やデジタルアーカイブを取り巻く環境が大きく変革する。知識情報資源を保有する機関の情報や、 業種業態を越えた各機関が提供するサービスのオープン化が進み、 AIが実サービスに組み込まれる段階に到達する。そのような社会の変革に図書館サービス、 デジタルアーカイブは、 何を目標に、 関係機関全体でどのようなサービスを実用レベルに持っていくか？ そのために各機関はどのような分担でサービスを構築していくか？ 具体的な構築に当たってはAIをどのようにして活用していくか等を 検討していく必要がある。

## どのような内容にすべきかは、 ここでのまとめで議論していきたい。

## 参考になる社会の変革、 国等の機関での方向性、 学会等での議論をのちに列挙する。

## ・・・・

# **【今後10年を見据えて】図書館を取り巻くサービスの変革への対応**

## 情報及びシステムのオープン化

* + 1. オープンデータ
       - オープンデータ（LOD）
    2. オープンシステム
       - オープンシステム
       - オープンソースソフトウェア（OSS）
       - オープンOS
    3. オープンAPI
       - メタデータ検索API
       - メタデータ交換API
       - 画像データ表示API
       - Webサービス連携API
    4. オープンAI
       - AIサービス連携API

## AIの実用化

# **図書館サービスシステムとして何を目指すか？**

## 理念・目標（？）

* + 1. 理念
       - 新たな知識の創造と還流
       - 社会・経済的な価値の創出
    2. 文献情報検索から事実検索へ
       - レファレンスサービスにおいて、 「どの本に書いてある」ではなく「信頼性の高い情報は、 どの本のどこにどのように書いてあるので、 質問の解答は何々である」

## 何をするサービスか（サービス要件）

* + 1. デジタルアーカイブジャパンの一員としてデジタルコンテンツを蓄積し、 デジタルアーカイブジャパンを活用して、 利用者からのレファレンスに対して事実に基づいた的確な情報を回答する
    2. レファレンスに必要な情報の蓄積
       - 図書館等において精緻に作成された過去のレファレンス事例、 辞書、 書誌データを教師データとしてAIに投入する
    3. 利用者からのレファレンスに対して回答
       - 利用者からの問合わせに対して、 AIシステムを活用して、 可能な限り「事実情報」を提供する
       - 利用者からの問合わせに対して、 参考になりそうな文献候補を提示する
       - その文献候補の所蔵先へ案内する
         * デジタル化された情報があれば、 閲覧サービスまでナビゲート
         * デジタル化された情報がなければ、 貸出、 購入サービスへナビゲート

## これにより何がどのように良くなるか（事前評価）

* + 1. 利用者が信頼性が高く、 より的確な情報を入手できるようにする？
    2. 利用者が入手した情報によって、 知識の再生産をするための工数、 経費を削減する？

## 業務とシステムはどのように分担してサービスを構築するか

* + 1. サービス全体概要（ユースケース図、 データフローダイアグラム（DFD）、 エンティティリレーション図（ER図））
    2. レファレンスに必要な情報の蓄積
       - 自館しか所蔵していない貴重な資料の可視化 （デジタル化、 電子書籍化、 本文テキスト化、 メタデータ付与）
       - 自館しか所蔵していない貴重な資料の知識化 （人もしくはAIシステムで組織化）
       - 他機関が提供する資料と併せて知識として統合利用化 （Japanサーチ等のサービスの利用）
    3. 利用者からのレファレンスに対して回答
       - 入力機能（GUI）
         * 利用者からの問い合わせを、 キーワードのみならず、 自然言語で受け付けられるようにする
       - 処理する機能
         * キーワード、 文脈から、 統合利用化された知識を検索し、 内容を取り出す
       - 出力機能（GUI）
         * 問い合わせに対して事実が明確な場合は、 その事実内容の紹介する
         * 問い合わせに関連した記述された参考文献があれば、 その文献の記述部分を紹介する
         * 問い合わせに関連した参考文献の候補と所蔵場所を紹介する

# **5～10年後に適用するシステムのイメージ**

## Subtopic

# **図書館サービスシステムでの人工知能の活用の可能性【作成中】**

## データ発生・データ収集

* + 1. 情報・データに関する事項
       - 永続的識別子
       - メタデータ
       - 目次・索引データ
       - 関連データ
       - 画像データ
       - 全文テキストデータ仕様
    2. 創造的活動の成果物のデジタルコンテンツ構築
       - 原資料からのデジタル化
         * デジタル化方法
         * スキャニング方法
         * 品質検査仕様
         * 画像データ仕様
         * コンテナ形式
         * メタデータ
         * テキストデータ仕様
       - 電子書籍化要件定義
         * 最終成果物の形式

構造化テキスト（マスター原稿）

リフロー型電子書籍（文字主体の本）

FIX型電子書籍（ビジュアル指向の本）

Webページ（HTML5+CSS3）

PODによるペーパーバック本

* + - * + 作成するメタデータ記述要素・記述規則

書誌的事項のメタデータ

本文埋め込みメタデータ

* + - * + 指定するビューア依存形式
        + 利用するコンテンツ制作環境サービス
        + 維持管理に利用するコンテンツ管理用・制作用ツール
        + 適用するコンテンツ制作用ガイドライン・テンプレート
    1. IoT等のシステム機器から生成されるデータ
    2. 情報を特定するための情報の作成（二次情報）
       - 情報に関する情報（メタデータ）
       - 調べ方案内
       - レファレンス情報（文献情報提供）
    3. 利用者の問いに関しての知識の収集（事実情報）
       - 予測調査資料
       - ・・・

## 学習フェーズ

* + 1. データ分析
       - 受付
       - 加工
       - 保管
       - 分析
         * 統計解析
         * 機械学習

## 識別フェーズ

* + 1. データ利用