

# 文書管理システム

文書管理システム（英: Document management system, DMS）は、電子文書や紙の文書をデジタイズしたものを格納・管理するコンピュータ上のシステムである。コンテンツマネージメントシステムと重複する部分があり、エンタープライズコンテンツ管理システムの一部とされることが多く、デジタル資産管理/文書の画像化/ワークフローシステム/記録管理システムとも関連する。

## 目次

- 1 概要
- 2 歴史
- 3 文書管理とコミュニケーション
- 4 構成要素
- 5 フリー/オープンソースの文書管理システム
- 6 外部リンク

## 概要

文書管理手法は以下のように分類される。

### 配置

文書をどこに格納するか。文書にアクセスするにはどこに行けばよいか。ファイリングキャビネットやファイル室への物理的移動と、文書管理システムを使う際に必要な画面上の手続きとは類似性がある。

### ファイリング

文書をどのようにファイルするか。後の検索の補助となる構成や索引付けをどうするか。文書管理システムはファイリング情報を格納するのにデータベースを使うことが多い。

### 検索

文書をどのように見つけ出すか。一般に検索は、文書群を順次見ていく方法と特定の情報を鍵として検索する方法がある。

### セキュリティ

文書の安全性をどのように保つか。無許可の人が文書を読んだり、改ざんしたり、破壊したりするのをどう防ぐか。

### 災害対策

火事、洪水、その他の自然災害から、どうやって文書を守るか。

### 維持

文書をどれだけの期間維持できるか。これは、その組織が情報や文書の維持（特にその期間）をどう

定義するかというポリシーの問題である。これは組織としての記録管理の慣習による規則である。

## 保存

将来の必要性を考慮して文書をどう保持するか。保存（アーカイビング）とは、文書と関連するメタデータを（その組織の定義に従って）ある期間を経た後に動的リポジトリから除去し、別の場所にまとめて保存しておくことである。電子文書でも紙の文書でも、アーカイブは文書の移動を伴う。紙の文書であれば文書用の倉庫、電子文書であればオフラインの記録媒体に移される。

## 配布

文書を必要とする人が利用できるようにする方法。

## ワークフロー

文書が人から人へと渡される場合、そのフローはどうあるべきか。

## 作成

文書作成方法。複数の人間が共同作業を行う場合、特に重要である。このため、バージョン管理システムなどが生まれた。

## 認証/認可

政府や企業が文書管理システムを使って電子化を行う場合、文書の原本性の保証、署名や押印に相当する機能をどう実現するかが問題となる。

# 歴史

文書管理は管理対象によって2種類に分類される。1980年代、複数のベンダーが紙の文書を管理するシステムを開発し始めた。それらのシステムは、印刷物だけでなく、写真その他も含めた紙の文書を管理する。コンピュータによって管理するため紙上の情報をデジタイズし、画像ファイルを管理するものであり、電子文書管理(EDM)システムなどと呼ぶ。

その後、別の種類のシステムとして電子文書に特化したシステムが生まれた。これは、コンピュータ上で作成された文書（ファイル）を管理するものである。初期のシステムは管理できるファイルフォーマットが独自のものだったり、限られていた。

これらの2つの系統が統合され、管理対象のファイルフォーマットも多様となり、ネットワーク上に文書を分散配置できるようになった。他のアプリケーションとの連携も強化され、セキュリティ機能も強化されていった。

# 文書管理とコミュニケーション

Carzaniga と Wolf (2001年) の論文“Content-based networking: a new communication infrastructure”において、電子文書管理について考察している。そこで、コミュニケーション基盤としてのコンテンツベースのネットワークが提案されている。ユーザーは関心のあることを表明し、それをネットワークにメッセージとして入力する。そして、ネットワークが適切な人々に適切な情報を配信する。Sprague (1995年) では、より詳細な文書管理方式が提案されており、それを電子文書管理（electric document management、EDM）と称した。彼の定義によれば、文書管理とは「組織の目的に適った文書の生成/格納/構成/転送/検索/操作/更新/廃棄」（pp.32）であり、EDM によって人々のコミュニケーションが改善されるとした（pp 42-43）。

EDMとコミュニケーションの関係については他にも例がある。Hansen と Haas (2001年) は、電子文書について情報提供者と情報利用者の役割の観点論じている。また、Thorpe と Mead (2000年) でも EDMとコミュニケーションの明確な関係を研究している。彼らは、EDMシステムがコミュニケーションのパターンを変革することを示した。彼らが研究した3つのプロジェクトでは、そのうちの2つがEDMを中心としていた(残る1つは3カ月で中止となった)。Howard と Pettersen (2001年) の研究対象となったあるプロジェクトでは、コミュニケーション手段としてEDMを三番目の手段としていた(上位手段は電話と会議、下位手段は電子メール、紙の文書、ファックスなど)。Rene Brohm (2005年) は学位論文でシアターモデルを提案した。シアターモデルとは、文書管理システムを劇場のステージに対比させて説明するものである。すなわち、ステージ上の演劇での相互作用と文書管理システムの機能の類似性を論じている。

Marin と Poulter (2004年) によれば、全てのデータや情報が集中型のデータベース(あるいはイントラネット)に置かれ、それを組織の全員が使う場合、ITと情報配布に明確な関係が生じる。彼らは、情報へのアクセスが容易であればあるほど、組織内に流通する情報量が増えるとした。

このコミュニケーション・ツールを改良する別の方法もある。Hansen と Haas (2001年) はEDMを競争のある市場に見立てた。それによると、情報提供者は情報共有の戦略を持つべきであり、ユーザー(従業員)にシステムを使わせる戦略も必要であるとした。

Yan と Garcia-Molina (1999年 pp.2) が提案した方法は、人々が必要な(興味のある)情報を表すクエリ群を含んだ長期的プロファイルを提示するというものである。これにより、得られる情報の質が向上するとした。つまり、事前に明らかになっているプロファイルに基づいて、情報が配布されるのである。それによって検索時間やコストが削減される。

## 構成要素

文書管理システムには、格納、バージョン、メタデータ、セキュリティ、索引、検索機能といった構成要素がある。以下では、これらについて解説する。

### メタデータ

メタデータは一般に各文書に付随して格納される。例えば、文書がシステムに格納された日付、格納を行ったユーザー名(ID)などがメタデータとして含まれる。文書内容からメタデータを自動的に抽出するシステムもあれば、ユーザーにメタデータを追加させるシステムもある。文書をスキャンした画像ファイルに対して光学文字認識を行うシステムもある。内容をコンピュータが認識できるテキストとすることで、そこから検索用のキーワードを抽出したり、全文検索を可能にするなどの可能性が生まれる。抽出したテキストもメタデータの一部として元の画像ファイルごとに格納されるか、検索用に集中管理される。

### 連携

多くの文書管理システムは、他のアプリケーションとの連携が可能であり、アプリケーションを通して文書を検索して取り出し、修正を加えた文書を文書管理システムに戻すといったことが可能である。連携が可能なアプリケーションとしては、オフィススイートや電子メールやグループウェアなどが一般的である。

### 画像入力

紙文書のデジタル画像化にはイメージスキャナや複合プリンタが使われる。光学文字認識ソフトもよく使われ、デジタル画像をコンピュータが内容を処理できるテキストに変換する。

### 索引

電子文書を特定するのに使われる。単なる識別子（番号）でもよいが、一般にもっと複雑な形式が使われ、メタデータによる分類や内容から抽出したキーワードなどによる索引付けが行われる。索引は検索のためにある。

## 格納

電子文書を格納する。格納には文書を管理するという意味も含まれ、格納位置の決定、格納期間の決定、記憶媒体から記憶媒体への移動、文書廃棄などを行う。

## 検索

格納された文書を検索する。ある文書を検索するというのは簡単なようだが、電子的な検索は非常に複雑で強力なものとなりうる。単なる個々の文書の検索は、一意な文書識別子を指定することでなされ、システムがその識別子に対応した文書を取り出す。より柔軟な検索としては、識別子やメタデータの一部と考えられるキーワードを指定することで行われる。この場合は一致した文書の一覧が表示されることが多い。複数のキーワードを使ったブール式を指定できるシステムや、内容の一部を指定できるシステムもある。検索は事前に作成された索引を使って行われる場合もあるし、より時間のかかる全文検索が行われることもある。

## セキュリティ

文書のセキュリティは文書管理システムにとっては不可欠である。文書の種類ごとにアクセス制限をどうするかが複雑に決められていることが多い。

## ワークフロー

ワークフローは複雑な問題であり、文書管理システムによってはワークフロー・モジュールが組み込まれている。ワークフローにはいくつかの種類があり、EDMが適用される環境によってワークフローが異なる。手動ワークフローは、ユーザーが文書を見てそれを誰に送るかを判断する。ルールベースのワークフローは、管理者がフローを規定するルールを作成する。例えば、請求書は決済プロセスを経て支払い部門に送られる。各部門がルールを作成して固有のワークフローを構築できるものもある。例えば、送られてきた請求書の金額によって最終的な送付先を変えるなどのルールが考えられる。

## 共同

EDM は複数の人間が共同で利用する。各文書は適切なユーザーが処理しなければならない。そのため、あるユーザーがある文書を処理中の場合、他のユーザーが同じ文書にアクセスするのをブロックしなければならない。

## バージョン

文書は変更が加えられるたびにバージョン（版）が付与される。これによって更新にミスがあったときに、古いバージョンから処理を再開できる。

# フリー/オープンソースの文書管理システム

- Alfresco
- KnowledgeTree (<http://www.knowledgetree.com/>)
- LogicalDOC (<http://www.logicaldoc.com/>)
- NemakiWare (<http://www.nemakiware.com/>)

- Nuxeo (<http://www.nuxeo.com/>)
- O3Spaces (<http://www.o3spaces.com/>)
- OpenKM (<http://www.openkm.com/>)
- WebFileSystem (<http://sourceforge.net/projects/webfilesystem/>)

## 外部リンク

- Miles L. Mathieu, Ernest A. Capozzoli (2002年) (PDF). *The Paperless Office: Accepting Digitized data* ([http://www.cabinetng.com/media/Paperless%20Office\\_Troy%20State.pdf](http://www.cabinetng.com/media/Paperless%20Office_Troy%20State.pdf)). Troy State University.
- Chris Ryan (2007年). *Getting Started on a Paperless Office* (<http://www.springcmmkt.net/GoPaperless>). SpringCM.

「<https://ja.wikipedia.org/w/index.php?title=文書管理システム&oldid=54449038>」 から取得

- 
- 最終更新 2015年2月16日 (月) 13:54 （日時は個人設定で未設定ならばUTC）。
  - テキストはクリエイティブ・コモンズ 表示-継承ライセンスの下で利用可能です。追加の条件が適用される場合があります。詳細は利用規約を参照してください。