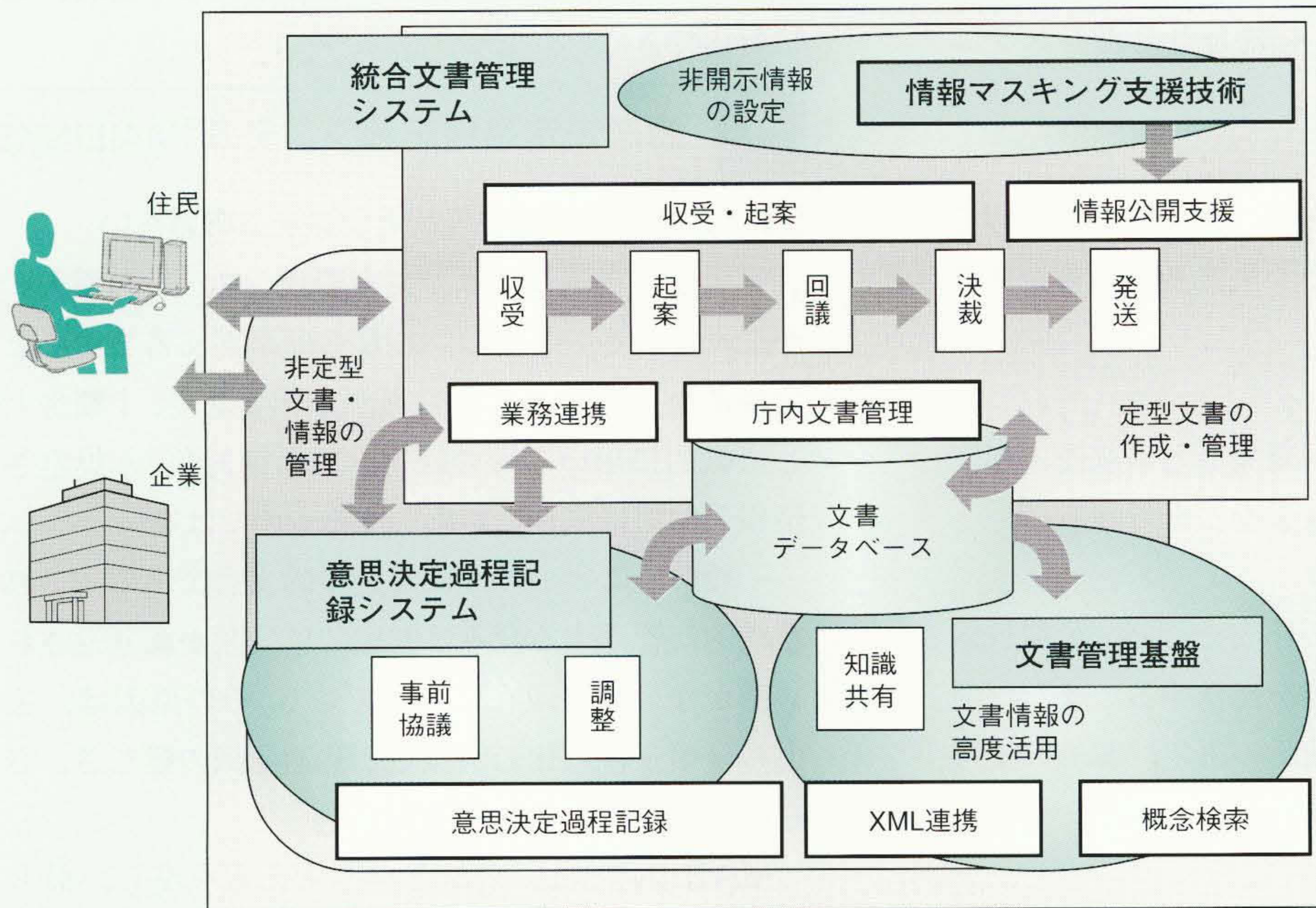


情報公開を見据えた文書情報マネジメントシステム

Document Management System for Information Disclosure

眞利裕之 Hiroyuki Mari 松田純一 Jun'ichi Matsuda
伊藤泰樹 Yasuki Itô 柴原節男 Setsuo Shibahara



注：略語説明
XML (Extensible Markup Language)

文書情報マネジメントの概要

文書情報マネジメントシステムは、情報公開への対応、事務処理の効率化、行政サービスの高度化を推進するための基盤である。

日立製作所の文書情報マネジメントシステムは、行政での文書情報の発生から廃棄までのライフサイクル管理や情報公開などを支援する基盤機能であり、(1) 收受・起案から決裁・発送までの定型的な管理、(2) 調整・事前協議なども含む意思決定過程の記録、(3) 概念検索、(4) XML (Extensible Markup Language) 連携、(5) 非開示情報のマスキングなどの要素機能で構成している。

これらの機能は、統合文書管理システム、意思決定過程記録システム“MAILINKS”，文書管理基盤“DocumentBroker”，および情報マスキング支援技術によって提供される。このシステムにより、文書情報管理の電子化だけでなく、情報公開への対応が可能となる。

1 はじめに

2001年4月から施行される情報公開法は、行政機関の職員が組織的に用い、保管している「行政文書」を国民の請求に従って開示することを定めたものである。情報公開法を適正かつ円滑に運用するためには、行政機関が保持する膨大な文書を適正に管理することが必要となる。総務庁は「行政文書の管理方策に関するガイドラインについて」の中で、各行政機関で可能なかぎり統一的な仕様で文書を管理する方策を示した。

ここでは、情報公開に対応するために必要な文書情報マネジメントシステムに関する製品、技術の開発状況、

および今後の展開について述べる。

2 統合文書管理システム

2.1 システムのねらい

日立製作所の統合文書管理システムは、官公庁や自治体の行政文書の起案から廃棄までのライフサイクルを一元的に管理し、行政文書の有効活用と、業務効率の向上を目的に開発したパッケージソフトウェアである。

保管した行政文書と管理情報(目録)を公開情報として利用したり、紙文書をスキャナで読み込むイメージデータへの対応や、他業務システム(例えば、財務会計システム)との連携など、拡張性のあるシステムとしている。

2.2 統合文書管理システムの構成と機能

統合文書管理システムは、文書管理基盤“Document-Broker”の適用により、概念検索やXML(Extensible Markup Language)に対応している。また、意思決定過程記録システム“MAILINKS”とも連携させ、文書管理に必要な以下のサブシステムを統合させたシステムである。

(1) 収受・起案管理システム

文書の収受(受付)から起案、保存、廃棄までの管理を行う。決裁ルートに登録でき、かつ決裁者と順番を任意に設定できる。

(2) 庁内文書管理システム

行政文書としての会議録や、例規集などの収受・起案文書以外の文書の作成、保管、検索などを行う。

(3) 情報公開支援システム

行政文書から管理情報(目録)の作成を行い、公開目録のインターネット上での検索機能を提供する。また、開示請求の受け付けから開示・非開示の決定、審査の支援と進捗(ちよく)状況の管理を行う。

(4) 業務連携システム

他業務システムと連携できるインタフェースを提供する。

2.3 システムの特徴と導入効果

(1) ウェブアプリケーションを採用することにより、TCO(Total Cost of Ownership)の削減を図ることができる。

(2) 機能・目的別にサブシステムを選択することによって段階的導入が可能となり、初期導入費用を抑制できる。

(3) 文書操作にきめ細かく権限を設定できることから、

高度なセキュリティ(安全性)対策ができる。

(4) システム導入効果

文書作成から保管まで一貫した管理ができるため、文書共有化による知識情報の高度化が可能となる。また、きたるべき情報公開法施行による煩雑な事務に、効率的に対応できる。

3 意思決定過程記録システム“MAILINKS”

3.1 文書情報記録の現状とシステムのねらい

官公庁・自治体では、文書管理システムの導入・普及が進んでいる。しかし、これらは起案文書を中心とした定型的な文書ファイル管理が主体であり、「職務上作成し組織的に用いるものはすべて公開対象」と規定された情報公開法に完全に対応しうるものではない。

一般に、文書の多くは非定型的文書である。これらを組織的に管理する試みは歴史的に失敗が繰り返され、決め手がないまま現在に至っている。その原因は、主に文書の分類・体系化の困難さと共通認識の難しさ、さらに登録の手間によるものと考えられる。

MAILINKSは、登録という考え方を廃し、最も身近で組織に浸透した情報交換ツールである電子メールを活用することにより、非定型な文書とその作成過程、すなわち意思決定過程を同時に記録し、組織的に活用することを目的として開発したものである。

3.2 MAILINKSの構成と機能

MAILINKSはメールソフトウェアに外付けするシステムであり、特定のメール(組織共有メール：本文、添付文書)を抽出して記録する。送信するメールをMAILINKSで記録する場合は、送信者がタイトルにタグ

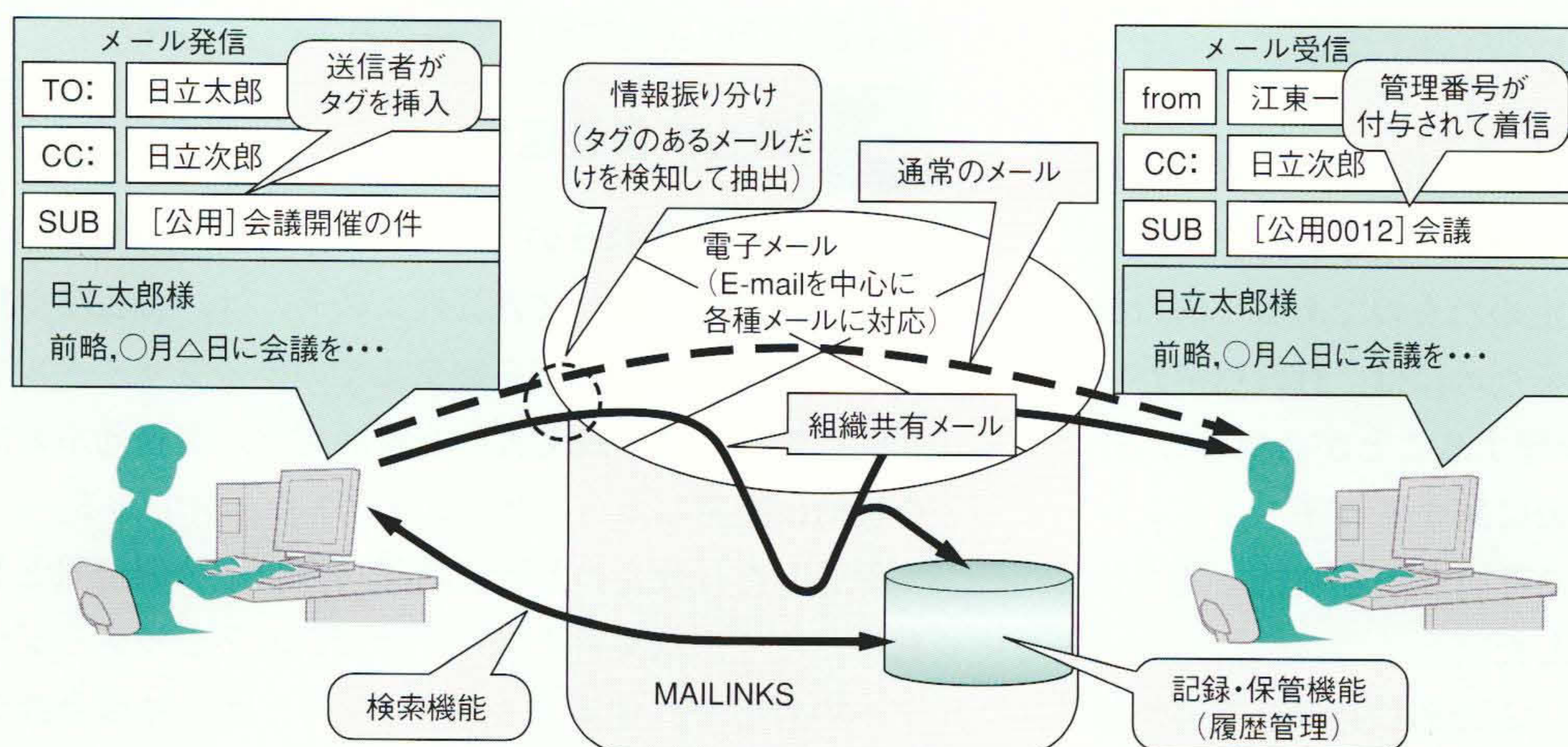


図1 MAILINKSの機能構成

MAILINKSは、情報振り分け、記録・保管(文書版履歴管理)、および検索の機能で構成する。

(特定の文字列)を挿入すると、MAILINKSがこれを検知して抽出、格納する(図1参照)。

抽出したメールをMAILINKS内に格納する際、転送、返信されたメールについては元のメールとの関連づけを行う。また、返信、転送された際、添付文書の書き換えを検知すると、版を更新して保管する。

これらのメール情報は、格納時に取得したメール属性と添付文書属性、さらにそれらの組合せ情報を用いて検索することができる。

3.3 実務上のメリット

3.3.1 意思決定過程の記録

このシステムにより、個々のメール情報だけでなく、関連メールをたどることによって連続的な意思決定過程が把握でき、情報を組織的に共有、活用することができる。

3.3.2 非定型文書の管理

非定型文書の登録を組織に強いるのは好ましいことではない。しかし、メール操作の範囲内での添付文書格納が実現すれば、登録作業に対する抵抗がなくなるうえ、メール送信によって付加される情報(送受信者、タイトル、送受信日時など)がそのまま文書属性情報として活用でき、所望の文書を探す確率は格段に向上する。さらに、属性情報の組合せにより、メールのやり取りの「さま」をキーに用いた検索(特定人物とのメールの往復回数を検索条件として指定するなど)が可能となる。これを「ふるまい検索」と呼ぶことにした。

4 文書管理基盤“DocumentBroker”

統合文書管理システムの基盤となるミドルウェア“DocumentBroker”では、実業務の中で、文書をいっそうアクティブに活用するための基本機能を部品としてそろえている。それらの部品を業務要件に応じて組み合わせることにより、ウェブ画面をクライアントとする文書管理システムを構築することができる。

4.1 統合文書管理システムでの適用要件

総合文書管理システムの基盤として、(1)利用中の文書をそのまま扱えること、(2)多様な条件で検索でき、(3)関係する文書を収集し、(4)別文書を生成できること、さらに(5)文書のセキュリティが確保できることを満たす必要がある。

4.2 開発のコンセプト

次の五つを開発の柱とした。

(1) コンテンツフリー

日常利用しているワープロや表計算ソフトウェアで作成したデータをそのまま格納、利用できる。

(2) ドキュメントコーディネーション

文書情報を業務の流れに応じて管理する。

(3) サーチ

業務遂行に必要な文書だけを確実に検索、収集する。

(4) オープン

業界標準に対応したシステムを構築する。

(5) セーフティ

データの保全性を確保しつつ、不正アクセスや情報の漏えいを防止する。

4.3 特徴的な機能

DocumentBrokerの特徴的な機能は以下のとおりである。

(1) XML対応

WWW(World Wide Web)の共通記述言語として注目されているXML形式のデータに対し、タグの特徴を生かして文書を利用できる。

(2) ライフサイクル支援

文書運用と業務プロセスを有機的にリンクできる。

(3) アクセス制御

文書やフォルダなどの単位で利用権限を定義できる。

(4) 概念検索

全文検索や属性検索に加え、任意に指定した文字や文章を基に、その条件と類似した概念(内容)を含む文書を検索できる。

(5) マルチレンディション

同じ内容を持つ複数形式の文書[Microsoft Word^{※)}、PDF(Portable Document Format)、SGML(Standard Generalized Markup Language)など]を同一名称で扱える。アクセス権限によってアクセスできるデータ形式を変えるなどといった運用が可能である。

(6) バージョン管理機能

文書の履歴を保存できる。フォルダにもバージョンを持たせて管理できる。

5 情報マスキング支援技術

5.1 情報マスキング支援技術開発のねらい

情報公開法では、個人情報や国家機密など幾つかの情

※) Microsoft Wordは、米国Microsoft Corp.の商品名称である。

報を公開しないことができる。文書全体が非公開な場合もあるが、人名や口座番号、住所などのプライバシーにかかわる部分だけを黒く塗りつぶして公開する、いわゆる「マスキング」が行われることがある。この場合、公開事務を行う職員が文書の内容をチェックしてマスキング処理を行うことになる。しかしこれには、(1) チェックに時間と手間がかかり行政サービスの低下を招く、(2) マスキング漏れによるプライバシー侵害の可能性があるという二つの問題がある。そのため、マスキング部分を特定するための支援技術を開発した。

5.2 機能概要

この技術による操作の流れは以下のとおりである。(1) ワード編集画面上で対象文書を開く、(2) マスキングボタンを押すことにより、マスキング候補部分が自動的に反転表示される、(3) マスキング候補部分をチェックし、マスキングすると判断した部分で「マスキングする」ボタンを押すことによって文字を記号に置き換え、黒く塗りつぶす。これにより、マスキング済みの文書が作成される(図2参照)。

マスキング候補部分を特定する処理は以下のとおりである。対象文書から個人情報などを抽出するため、まず、各文を単語に分割し、品詞などの属性情報を設定する形態素解析処理を行う。このシステムでは、3万語余りの固有名詞辞書を持ち、辞書に登録されている人名や法人名などを、この時点で検出することができる。一方、希少な名前や新規法人名などは辞書に登録していないこと

がある。また、「林」のように、名前なのか一般名詞なのかあいまいな場合もある。この場合は、構文意味解析を行い、マスキングパターンを参照してマスキング候補部分を特定する。

マスキング情報は、(1) 人名や法人名の後に付く接尾語、(2) 品詞の並び、(3) 主語が人間になりやすい動詞などから推定できる。

マスキングパターンの例を以下にあげる。

- (1) 「X課長」→X(人名接尾語の直前)をマスキング
- (2) 銀行名+数字→数字をマスキング(口座番号をマスキング)
- (3) 「Yが出席」→Y(「出席」の主語)をマスキング

マスキングパターンでは、任意の品詞、単語の並びや構文構造を記述できるので、マスキング基準が変わっても迅速に対応することが可能である。

6 おわりに

ここでは、情報公開への対応、事務処理の効率化、および行政サービスの高度化を推進するための基盤となる「文書情報マネジメントシステム」について述べた。

今後は、電子認証基盤や、セキュリティ基盤との連携強化を図っていく考えである。

執筆者紹介



眞利裕之

1991年日立製作所入社、金融・通流システムグループ
ビジネスソリューション開発本部 所属
現在、官公庁・自治体分野での文書管理システム導入計画策定コンサルティングに従事
E-mail: mari @ bisd. hitachi. co. jp



伊藤泰樹

1985年日立製作所入社、金融・通流システムグループ
ソフトウェア事業部 所属
現在、文書管理・知識管理ソリューション開発に従事
E-mail: itoyasuk @ gm. soft. hitachi. co. jp



松田純一

1984年日立製作所入社、中央研究所 所属
現在、自然言語処理の研究開発に従事
情報処理学会会員、人工知能学会会員
E-mail: j-matusd @ crl. hitachi. co. jp



柴原節男

1982年日立製作所入社、公共システムグループ 公共システム事業部 所属
現在、官公庁・自治体分野での文書管理システム開発に従事
E-mail: shibahara @ jkk. hitachi. co. jp

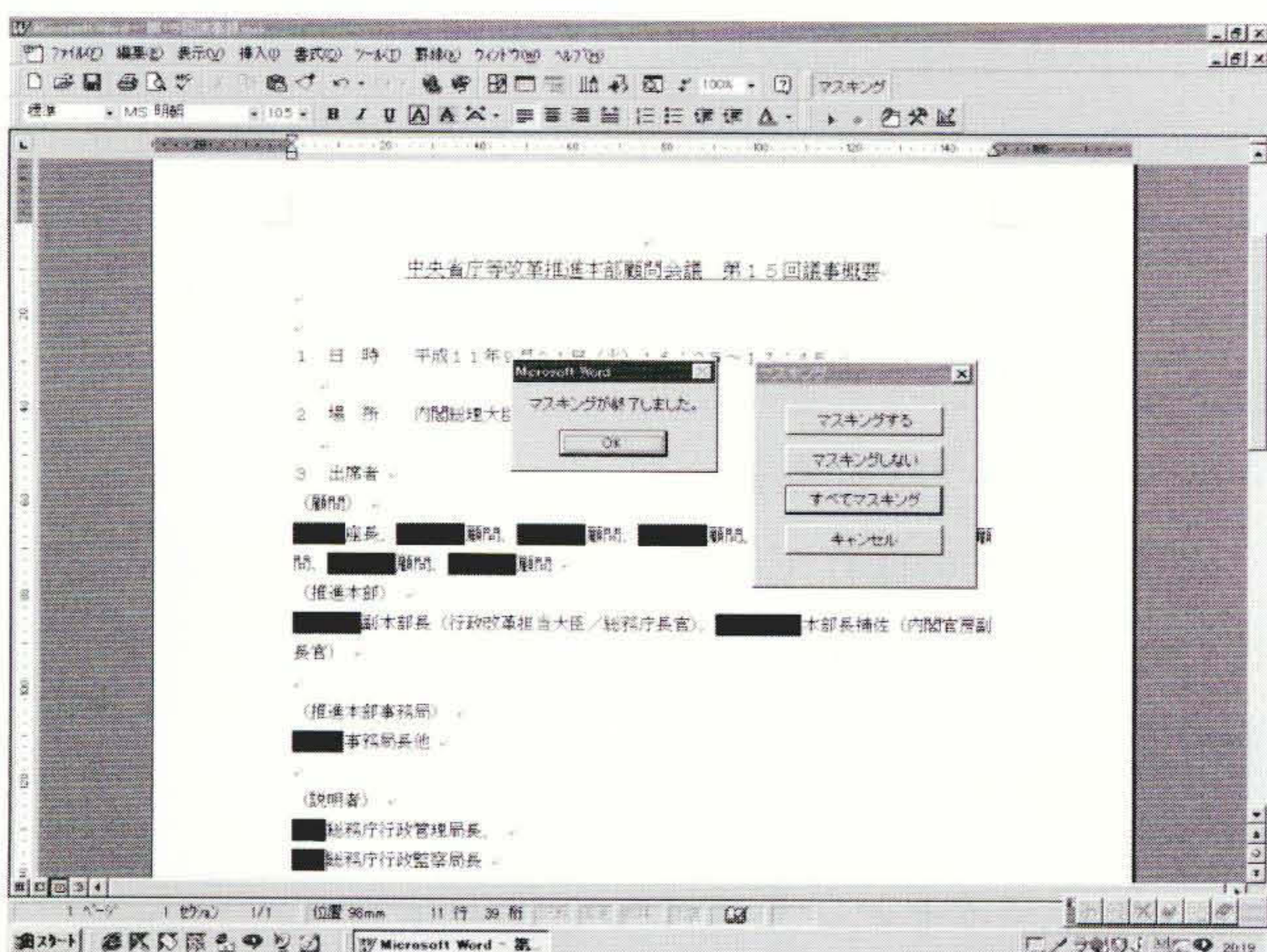


図2 マスキング処理の画面例

マスキング候補部分を指定し、マスキングすると判断した部分で「マスキングする」ボタンを押すことによって文字を記号に置き換え、黒く塗りつぶす。これにより、マスキング済み文書が作成される。