

文書格納・検索(DFR)応用の標準化動向

久保田 浩司 春田 勝彦

NTT 通信網総合研究所

ISO/IEC JTC 1/SC 18/WG 4 で 分散オフィス環境における文書の格納及び検索の機能を提供する Document Filing and Retrieval (DFR) の標準化が行われている。本稿では、DFR の標準化動向、概要及び課題について紹介する。

The Development of DFR Standardization

Koji Kubota Katsuhiko Haruta

NTT Telecommunication Networks Laboratories

1-2356 Take Yokosuka-shi Kanagawa 238-03, Japan

The Document Filing and Retrieval (DFR) application is standardized by ISO/IEC JTC 1/SC 18/WG 4, in order to provide the functionality of document filing and retrieval in a distributed office system. This paper describes the development of DFR standardization, its abstract and its required future studies.

1. はじめに

近年、オフィスにおける文書の量が増大し、文書のオフィスに占めるスペースが大きくなってきた。また、オフィスにおける紙の使用量も急速に増え、森林資源の破壊につながっているとの批判も大きくなっている。このような問題を解決するために、オフィスにおけるペーパーレス化が叫ばれ、大容量の電子/光ファイリングシステム又はファイルサーバが開発され、普及しつつある。

電子/光ファイリングシステムは、検索機能及び表示機能などが充実しているが、その多くは、入力が紙ベースである、格納形式がイメージだけである、各ベンダシステム間の互換性がない、ネットワークを介して専用端末からしかアクセスできないなどの問題点がある。

ファイルサーバは、ネットワークを介して既存の端末からのファイルアクセスを可能としているが、ファイルを格納するだけであって検索などの機能に乏しい。

このような背景から、オフィス文書を対象とした格納及び検索機能を提供する通信アプリケーションである Document Filing and Retrieval(DFR)の標準化が、ISO/IEC JTC 1/SC 18/WG 4で行われている。

本稿では、DFRの標準化の経緯、DFRの概要及び実現に対する技術的課題について紹介する。

2. 標準化の経緯

DFRは、分散オフィス環境において、複数の利用者に大規模な不揮発性記憶装置を用いた文書の格納及び検索の機能を提供することを目的に、1987年5月に SC18 NWI (New Work Item) として検討が開始された。以降、

1987年 9月 WD(Working Draft)

1988年10月 DP(Draft Proposal)

1989年12月 DIS(Draft International Standard)

と標準化作業が進められ、1991年6月に IS (International Standard) 化作業が終了し、現在では IS 出版待ちとなっている。

標準化の検討の途中、DFR と FTAM (File

Transfer , Access and Management)[3]との類似性が疑問視され、FTAM の担当 WG である ISO/IEC JTC 1/SC 21/WG 5 と DFRの担当 WG との間で1989年3月(ニューヨーク)及び5月(ストックホルム)に合同会議を開催し、国際レベルの調整を繰り返した。結果、それぞれの抽象化レベル並びに目的及び要求条件が異なるため、共通する一般的な格納モデルの開発の必要はないということで合意し、別々に規格化作業を進めることとなった。

3. DFRの概要

DFRは、分散オフィスアプリケーションのモデルを規定している DOAM (ISO/IEC 10031-1 ~ 2 ; Distributed Office Application Model) [4] の設計ガイドラインに基づいた規格の一つであり、OSI 基本参照モデルの最上位層である応用層の応用サービス要素 (ASE : Application Service Element) の一つとして位置付けられる。DFRは、分散オフィス環境において複数の利用者に大規模な不揮発性記憶装置を用いて文書の格納及び検索の機能を提供する目的で、そのモデル、文書の格納モデル、文書を格納しているサーバへのアクセス方法等を規定している。

3.1 機能概要

DFR は、文書の格納及び検索の機能に関連した様々な機能を提供する。主な機能を次に示す。

- (1)分散オフィスシステムにおいて複数の利用者が利用できる。
- (2)文書をグループ化して格納し、階層的に管理する。
- (3)様々な型(フォーマット)の文書を格納する。
- (4)文書及びグループへの参照(リンク情報)を管理する。
- (5)文書などを順序を付けて格納する。
- (6)文書などの内容(実体)と独立に文書などに関連する属性を付加する。
- (7)属性を用いた探索、順序付け一覧機能を行う。
- (8)文書などの変更、削除、移動、読取りを行う。
- (9)属性を利用して文書の版管理を行う。
- (10)サーバへのアクセス認証を行う。

- (11)属性を利用して文書などへのアクセス権を管理する。
- (12)属性を利用して文書などへの複数利用者の同時アクセスを制御する(予約)。

3.2 抽象モデル

DFRのモデルを図1に示す。DFRは、DOAMで規定されたクライアント・サーバモデルに従い、DFR サービスを要求する側(DFR-User)及びDFR抽象サービスを提供する側(DFR-Server)からなる。DFR-UserとDFR-Serverとの間で通信するために使用するプロトコルをDFRプロトコルという。

DFRポートは、DFR-User及びDFR-Serverに対してそれぞれに定義される。このDFRポートを用いて、DFR-Userは、DFR-Serverと通信することができる。このDFRポートによって提供される機能が、DFRサービスである。DFRサービスは、非対称であり、DFR-Userによって起動され、DFR-Serverによって実行される。

3.3 情報モデル

DFRでは、実在する多種多様な文書管理を統一的に扱えるように、キャビネット、フォルダなどの概念を盛り込んだ論理的な階層構造の文書の格納

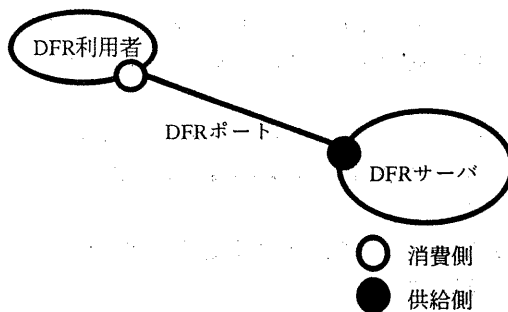


図1 DFRモデル

モデルを規定している。

3.3.1 DFR-Document-Store(DS)

DFRでは、管理する情報のかたまり(例えば文書)をDFR-Objectと呼び、DFR-Document-Store(DS)に論理的な階層構造の形で格納される。DFR-Objectには、DFR-Document、DFR-Group、DFR-Reference及びDFR-Search-Result-Listがあり、さらにDFR-Groupには、DFR-Root-GroupとDFR-Proper-Groupがある。

図2にDSの構造例を示す。DFRでは、DSの構造として、次の三つの形態の指定を許している。

- (1) Flat Store Set : DFR-Root-Group の下に

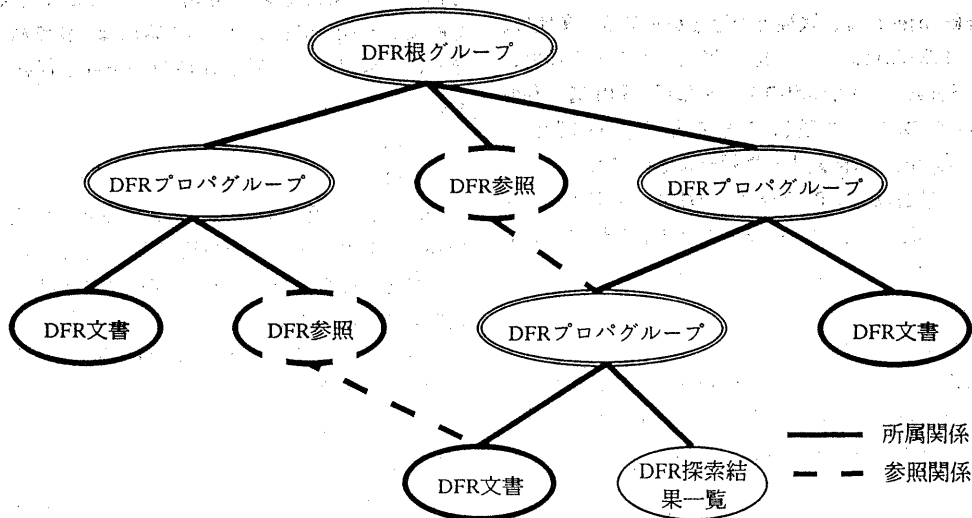


図2 DFR-Document-Storeの構造例

DFR-Document、DFR-Reference 又は DFR-Search-Result-List を格納する形態で、二階層の形態である。DFR-Proper-Group は存在しない。

- (2) Pre-defined Store Structure Set : DFR-Group の構造が予め決められている形態である。
- (3) Full Set : 構造を自由に設定及び変更ができる形態である。

DS は、一つの DFR-Server に一つである。DFR-Server は、DS 中の DFR-Object への操作手順を DFR-User に提供する。

DFR-Object は、DS 中では階層的な位置付けを表す属性 (DFR-Pathname 及び DFR-Parent-Identification) を付加した DFR-Entry の形で格納される。

DFR-User は、DS の中で一意に識別される UPI (DfrUniquePermanentIdentifier) 又は DFR-Root-Group からの経路を示す DFR-Pathname を介して DFR-Object をアクセスできる。UPI は、DFR-Server によって割り当てられ、DFR-Object が存在する間は変化しない。

3.3.2 DFR-Object

DFR-Object は、属性と内容とからなる。属性は、各 DFR-Object の内容に関連したデータ項目、DFR-Server の管理用のデータ項目、利用者の利便の目的のデータ項目などである。内容は、DFR-Object に格納する情報そのものであり、この情報の性質は、DFR-Object の種別に依存する。次に各 DFR-Object の概要を示す。属性については、3.3.3 に示す。

(1) DFR-Document

DFR-Document は、DS に格納される文書であり、その内容は文書そのものである。DFR-Server は、DFR-Document の内容を解釈しない。

(2) DFR-Group

DFR-Group は、メンバと呼ぶ DFR-Object の集合

である。DFR-Group の内容は、その DFR-Group を構成するすべての DFR-Object の UPI の列からなる。

DFR-Group は、DFR-Root-Group 又は DFR-Proper-Group のいずれかである。DFR-Root-Group は、最上位の DFR-Group であり、一つの DS に一つだけ存在する。DFR-Proper-Group は、DFR-Root-Group 以外の DFR-Group であり、常にその他の DFR-Group に属している。DFR-User は、DFR-Root-Group を作成及び削除できない。

(3) DFR-Reference

DFR は、他の DFR-Object への参照(リンク)を作成し、DS 内に格納する機能を提供する。この参照情報を管理する DFR-Object が DFR-Reference である。DFR-Reference の内容は、参照している DFR-Object へのポインタである。

参照先として指定できる DFR-Object は、DFR-Document、DFR-Proper-Group 又は DFR-Search-Result-List のいずれかだけである。

(4) DFR-Search-Result-List

DFR-Search-Result-List は、探索結果の保存及び再利用を目的とする DFR-Object である。DFR-User は、DS 内に DFR-Search-Result-List を作成し、探索結果を格納することができる。DFR-Search-Result-List の内容には、探索結果のほかに、探索した日時、領域及び条件も保存される。

3.3.3 属性

DFR では、必須である DFR 基本属性集合と ODA で規定する ODA 文書概要 [5] から取り出すことができる任意選択の DFR 拡張属性集合の二つの属性集合を規定する。

DFR で規定する属性と DFR-Object の対応を表 1、表 2 に示す。

3.4 抽象操作

3.1 で述べた DFR が提供するサービス/機能は、すべて抽象操作を用いて実行される。DFR で規定する抽象操作を次に示す。また、各抽象操作と

表1 DFR基本属性集合

Attribute -Type	DFR-Object				Attribute -Type	DFR-Object			
	DO	GR	RE	SR		DO	GR	RE	SR
DFR-UPI	M	M	M	M	DFR-Version-Root	C		O	
DFR-Object-Class	M	M	M	M	DFR-External-Location	O	O	O	
DFR-Document-Type	M		O		User-Reference	O	O	O	O
DFR-Title	M	M	M	M	User-References-To -Other-Objects	O	O	O	O
DFR-Pathname	C	C	C	C	DFR-Attributes-Create -Date-And-Time	M	M	M	M
DFR-Parent -Identification	M	M	M	M	DFR-Content-Create -Date-And-Time	C	C	C	C
DFR-Referent-Deleted			C		DFR-Created-By	M	M	M	M
DFR-Membership -Criteria		O			DFR-AttributesModify -Date-And-Time	C	C	C	C
DFR-Ordering		O			DFR-Content-Modify -Date-And-Time	C	C	C	C
DFR-Resource-Limit	O	O	O	O	DFR-Attributes -Modified-By	C	C	C	C
DFR-Resource-Used	M	M	M	M	DFR-Content-Modified- By	C	C	C	C
DFR-Number-Of-Group -Members		M			Document-Date-And -Time	O		O	
Version-Name	C		O		DFR-Reservation	C	C	C	C
DFR-Previous-Version	C				DFR-Reserved-by	C	C	C	C
DFR-Next-Version	C				DFR-Access-List	C	C	C	C
備考 DO=DFR-Document, GR=DFR-Group, RE=DFR-Reference, SR=DFR-Search-Result-List: C=条件付き必須, M=必須, O=任意選択									

表2 DFR拡張属性集合

Attribute-Type	DFR-Object				Attribute-Type	DFR-Object			
	DO	GR	RE	SR		DO	GR	RE	SR
Other-Titles	0	0	0	0	Preparers	0		0	
Subject	0	0	0	0	Owners	0		0	
Document-Type	0				Authors	0		0	
Document-Architecture-Class	0		0		Status	0		0	
Keywords	0	0	0	0	User-Specific-Codes	0	0	0	0
Creation-Date-And-Time	0	0	0	0	Superseded-Documents	0		0	
Purge-Date-And-Time	0	0	0	0	Number-Of-Pages	0		0	
Revision-Date-And-Time	0	0	0	0	Languages	0		0	
Organizations	0		0						

備考 DO=DFR-Document, GR=DFR-Group, RE=DFR-Reference, SR=DFR-Search-Result-List:
C=条件付き必須, M=必須, 0=任意選択

表3 DFR抽象操作とDFR-Objectとの対応

DFR-Operation	DFR-Document		DFR-Group		DFR-Reference		DFR-Search-Result-List	
	Attributes	Content	Attributes	Content	Attributes	Content	Attributes	Content
DfrBind								
DfrUnbind								
Create	○	○	○		○	○	○	
Delete	○	○	○	○	○	○	○	○
Copy	○	○	○	○	○	○	○	○
Move	○	○	○	○	○	○	○	○
Read	○	○	○		○	○	○	
Modify	○	○	○		○	○	○	
List				○				○
Search	○		○		○		○	○
Reserve	○	○	○	○	○	○	○	○
Abandon								

備考 ○:適用可能

DFR-Object の対応を表 3 に示す。

(1) DfrBind

DFR-User と DFR-Server との間のアソシエーションを確立する。このとき、DFR-Server は利用者の認証を行う。また、DFR-Server がサポートしている拡張属性、名前などの制約、DS の構造などを折衝することもできる。

(2) DfrUnbind

DFR-User と DFR-Server との間のアソシエーションを解放する。

(3) Create

DS 内の指定した DFR-Group の配下に新たに DFR-Entry を格納する。このとき、格納先の DFR-Group の予約(同時アクセス制御機能)状態、DFR-Group への利用者のアクセス権、DFR-Entry の DFR-Group の所属基準に対するチェックなどが行われる。格納時には、DFR-Server によって UPI が割り当てられる。

(4) Delete

DS 内の指定した DFR-Entry を削除する。このとき、DFR-Entry 及び所属先の DFR-Group の予約状態、DFR-Entry 及び DFR-Group への利用者のアクセス権のチェックが行われる。

(5) Copy

指定した DFR-Entry を指定した DFR-Group へ複写する。このとき、DFR-Entry 及び複写先の DFR-Group の予約状態、DFR-Entry 及び DFR-Group への利用者のアクセス権のチェック、並びに複写結果となる DFR-Entry の DFR-Group の所属基準に対するチェックが行われる。複写時には、DFR-Server によって複写結果である DFR-Entry に新規に UPI が割り当てられる。

(6) Move

指定した DFR-Entry を指定した DFR-Group へ移動する。このとき、DFR-Entry 所属する DFR-Group 及び移動先の DFR-Group の予約状態、それ

ら DFR-Object への利用者のアクセス権のチェック、並びに DFR-Entry の移動先の DFR-Group の所属基準に対するチェックが行われる。移動時には、DFR-Entry の UPI は変わらない。

(7) Read

指定した DFR-Entry の属性又は内容を読み取る。このとき、DFR-Entry の予約状態及びアクセス権のチェックが行われる。

(8) Modify

指定した DFR-Entry の属性又は内容を変更する。このとき、DFR-Entry の予約状態及びアクセス権のチェックが行われる。また、変更に伴い、所属している DFR-Group の所属基準に矛盾しないかのチェックも行われる。

(9) List

指定した DFR-Group に属する又は DFR-Search-Result-List 中に記録されている DFR-Entry の属性を一覧表示する。このとき、DFR-Group 又は DFR-Search-Result-List 、及び各 DFR-Entry の予約状態、アクセス権のチェックが行われる。予約及びアクセス権で読み取れない DFR-Entry は、一覧の中に含まれない。この操作では、一覧に含まれる DFR-Entry の数が増えることを想定し、その数(容量)に制約を設定することができる。DFR-Server は、このとき、List 操作を中断する。再度継続要求があれば、継続結果を応答する。

(10) Search

指定した条件に適合する属性値をもつ DFR-Entry を一覧表示する。この結果を DFR-Search-Result-List の中に格納することができる。また、既に DS 内に格納されている DFR-Search-Result-List にある探索条件を利用することもできる。List 時同様に、予約及びアクセス権で読み取れない DFR-Entry は、一覧の中に含まれない。この操作では、一覧に含まれる DFR-Entry の数が増えること及び探索時間が長くなることを想定し、その数(容量)及び探索時間に制約を設定することができる。DFR-Server は、このとき、Search 操作を中断す

る。再度継続要求があれば、継続結果を応答する。

(11) Reserve

指定した DFR-Entry への同時アクセス制御レベルを設定する。このとき、DFR-Entry の予約状態及びアクセス権のチェックが行われる。

(12) Abandon

DFR-Server へ要求している抽象操作並びに中断されている List 及び Search 抽象操作を中止する。中止の扱いについては、DFR-Server の実装に依存する。

3.5 DFR プロトコル

DFR-User と DFR-Server との間のプロトコルのモデルを図 3 に示す。DFR サービスは、DFR サービス要素 (DFRSE) によってサポートされる。DFRSE は、非対称の応用サービス要素 (ASE) であり、DFR サービスの消費側である DFRSEc と DFR サービスの供給側である DFRSEs から構成される。DFRSE は、その他の ASE によってサポートさ

れる。

遠隔操作サービス要素 (ROSE)[6] は、DFR のモデルにおける DFR ポートで発生する抽象操作の要求 / 応答の機能をサポートする。高信頼転送サービス要素 (RTSE)[7] は、DFR-User と DFR-Server との間の操作のパラメータを含む応用プロトコルデータ単位 (APDU) を高信頼で転送する際に任意選択で使用される。アソシエーション制御サービス要素 (ACSE)[8] は、DFR-User と DFR-Server との間の応用アソシエーションの確立及び解放をサポートする。DFR-User と DFR-Server との間のアソシエーションは、DFR-User だけが確立可能であり、アソシエーションを確立した起動者だけがそれを解放できる。

DFRSE は、それをサポートする各 ASE へ写像する機能を提供する。RTSE への写像も任意選択で許しているが、一般的ではないので、ここでは ACSE 及び ROSE への写像について述べる。DFRSE 間でアソシエーションを確立/解放するために ACSE を利用し、DFR の他の抽象操作を実現するために ROSE を利用する。表 4 に各抽象操作の各 ASE サービスへ

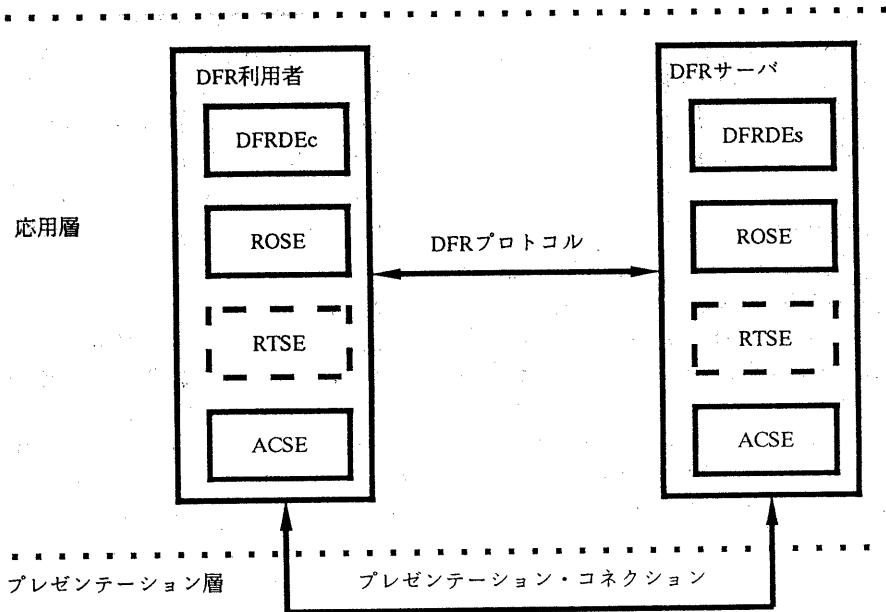


図3 DFRプロトコルモデル

表4 DFR-Operationの低位への写像例

DFR-Operation	ASE	ASEサービス
DfrBind	ACSE	A-ASSOCIATE
DfrUnbind	ACSE	A-RELEASE
Create	ROSE	RO-INVOKE
Delete	ROSE	RO-INVOKE
Copy	ROSE	RO-INVOKE
Move	ROSE	RO-INVOKE
Read	ROSE	RO-INVOKE
Modify	ROSE	RO-INVOKE
List	ROSE	RO-INVOKE
Search	ROSE	RO-INVOKE
Reserve	ROSE	RO-INVOKE
Abandon	ROSE	RO-INVOKE

の写像の例を示す。

4. 課題

DFRの規格は、汎用的な文書の格納形態及び格納・検索手順を規定している。従って、実際にDFRを用いたドキュメントファイリングシステムを実現するには、既存のドキュメントファイリングシステムとの関わりなどから、様々な課題があると考えられる。ここでは、その中の主なものを次に紹介する。

4.1 分散化

DFRは、個々のDFR-Serverに対するプロトコルを規定している。すなわち、DFR-Userと単一のDFR-Serverとの間の相互作用を管理するものである。今後、各オフィス及びオフィス間のネットワーク環境が整備されると、ドキュメントファイリングシステムの分散化及び分散したドキュメントファイリングシステムへのアクセス要求が高まってくる。従って、分散したDFR-Serverシステムの機能、DFR-Server間のシステムプロトコルの検討が必要である。

4.2 文書内容操作

DFRでは、DFR-ServerはDFR-Documentの内容

を解釈しないと規定している。そのため、DFR-Documentへのアクセスは、文書単位となっている。すなわち、文書をまるごと読み取る、変更する、複写する、移動することになる。しかし、既存のドキュメントファイリングシステムの多くは、文書をページ単位で管理しているため、ページ単位の読み取り、変更、版管理、検索などができる。今後は、ODAのように、ページだけではなく、さらに構造化された文書が増える。それに伴い文書の内容 / 構造に関わる操作が必要となる。DFRでは、これらの文書内容操作との関係を明確化し、それらとの相互動作などについて検討する必要がある。

4.3 属性の柔軟性

3.3.2に示した様に、DFRでは、属性としてDFR基本属性集合とDFR拡張属性集合とを規定している。そのうち、拡張属性集合は、文書型(ODAやSGMLなど)に応じてシステムに設定できるよう自由度をもたせている。しかし、多くのドキュメントファイリングシステムでは、利用者がグループ別又は文書種類別に自由にキーワード名を設定できる様になっている。このことに対応するため、DFRにおいても利用者が名前を設定でき、それを用いて検索できるような属性を追加する必要がある。

4.4 管理操作

現在のDFRの規格では、利用者への文書の格納及び検索のサービス提供に関する操作だけを規定している。しかし、実際にシステムを運用するためには、DFR-Root-Groupの設定、DS構造の設定又は結合可能な利用者の追加・削除及びそれに伴う処理などの管理操作が必要となる。これらの管理操作は、DFR-Serverにローカルな機能であるとも考えられるが、将来的には遠隔より操作出来ることが望ましいと考える。

5. おわりに

本稿では、分散オフィス環境において文書の格納及び検索の機能を提供するために ISO/IEC JTC

1/SC 18/WG 4 で標準化が行われているDFRについて、標準化の経緯、概要及び課題について紹介した。DFR を用いたドキュメントファイリングシステムの実現には、本稿で示したほかにも課題は多く存在すると思われるが、ドキュメントファイリングシステム間での相互接続 / 文書交換の観点より、今後普及していくと考える。今後は、標準化機関における残された課題の検討 / 解決、実装規約の制定などが必要と考えられる。

[謝辞]

最後に、本稿を執筆するに当たりご指導・ご鞭撻を賜りました NTT 通信網総合研究所ネットワークインテグレーション研究部 山縣部長、木下グループリーダー並びに関係各位に深く感謝致します。

[参考文献]

- [1] ISO/IEC JTC 1/SC 18/WG 4 N1683;"Revised Text of ISO/IEC DIS 10166-1 , Information Technology , Text and Office Systems - Document Filing and Retrieval (DFR) - Part 1: Abstract Service Definition and Procedures" ; 1991.7
- [2] ISO/IEC JTC 1/SC 18/WG 4 N1684;"Revised Text of ISO/IEC DIS 10166-1 , Information Technology , Text and Office Systems - Document Filing and Retrieval (DFR) - Part 2: Protocol Specification";1991.7
- [3] ISO 8571-1~5:1988;Information Processing Systems - Open Systems interconnection - File transfer, Access and Management
- [4] ISO/IEC 10031-1~2;Information technology - Text and office systems - Distributed-office-application model
- [5] ISO 8613-4:1989;Information processing - Text and office systems - Office Document Architecture (ODA) and interchange format - Part 4:Document profile
- [6] ISO/IEC 9072-1-2:1989; Information processing systems - Text communication - Remote Operations1
- [7] ISO/IEC 9066-1-2; Information processing

systems - Text communication - Reliability Transfer

[8] ISO 8649:1988;Information processing systems - Open Systems Interconnection - Service definition for the Association Control Service Element

[9] ISO 8650:1988;Information processing systems - Open Systems Interconnection - Protocol specification for the Association Control Service Element