

第 編

全体構成

M E M O

1. 物理フォーマット (公報情報)

物理フォーマットは、『JIS X 6241-1997 120mm D V D - 再生専用ディスク (ISO/IEC16448)』に準拠する。

2. 論理フォーマット

2. 1 公報情報

論理フォーマットは、『TR X 0006-1998 D V D - 再生専用ディスクのボリューム構造及びファイル構造』に準拠する。

情報交換による制約、さらに本仕様で追加した制約を以下に示す。

- (1) 論理ブロック長は、2 0 4 8 バイトとする。
- (2) ボリューム集合の大きさは、1 とする。
- (3) 各ファイルは、1 個のファイル分割からなる。
- (4) ファイル名は、ファイル拡張子と区切り文字も含めて、3 0 文字以内の英数字とする。(注1)
- (5) ファイル版数番号は、1 とする。
- (6) ディレクトリ名は、3 1 文字以内の英数字とする。(注2)
- (7)ブリッジフォーマット (JIS X 0609およびJIS X 0606(ISO 9660)) を前提とする。

ブリッジボリューム認識列のボリューム記述子は、基本ボリューム記述子とボリューム記述子集合終端子、主ボリューム記述子列におけるボリューム記述子は基本ボリューム記述子だけ用いる。基本ボリューム記述子の項目の内、本仕様で規定する内容を表2 - 1および表2 - 2に示す。

- (8) 拡張属性レコードは使用しない。

(注1) JIS X 0609-1998でのファイル識別子を、本仕様ではファイル名と称す。

英数字の文字コードは、JIS X 0201-1976とする。

(注2) JIS X 0609-1998でのディレクトリ識別子を、本仕様ではディレクトリ名と称す。

英数字の文字コードは、JIS X 0201-1976とする。

表2 - 1 DVD - ROMボリューム記述子集合基本ボリューム記述子 (JIS X 0606)

| No | 記述子内の バイト位置 | 名 称 | 内 容 |
|----|----------------|---------------|-------------------------|
| 1 | 9 - 40 | システム記述子 | スペースを記録する |
| 2 | 41 - 72 | ボリューム識別子 | (注1) |
| 3 | 121 - 124 | ボリューム集合の大きさ | 1 |
| 4 | 125 - 128 | ボリューム順序番号 | 1 |
| 5 | 129 - 132 | 論理ブロック長 | 2048 |
| 6 | 319 - 446 | 出版者識別子 | “ JAPAN PATENT OFFICE ” |
| 7 | 447 - 574 | データ編集者識別子 | “ JAPAN PATENT OFFICE ” |
| 8 | 575 - 702 | 応用システム識別子 | スペースを記録する |
| 9 | 703 - 739 | 著作権ファイル識別子 | “ COPYRIGHT ” |
| 10 | 740 - 776 | 抄録ファイル識別子 | “ ABSTRACT_CSV ” |
| 11 | 777 - 813 | 書誌ファイル識別子 | スペースを記録する |
| 12 | 814 - 830 | ボリューム作成日付及び時刻 | 作成日：DVD - ROMの発行日 時刻：0 |
| 13 | 831 - 847 | ボリューム更新日付及び時刻 | 更新日：DVD - ROMの発行日 時刻：0 |
| 14 | 848 - 864 | ボリューム失効日付及び時刻 | 指定しない(数値0) |
| 15 | 865 - 881 | ボリューム発効日付及び時刻 | 指定しない(数値0) |

備考1 上記の内容の記録形式は、JIS X 0606-1998に従う。

備考2 上記以外の項目は、JIS X 0606-1998に従う。

表2 - 2 基本ボリューム記述子 (JIS X 0609)

| No | 記述子内の バイト位置 | 名 称 | 内 容 |
|----|----------------|-------------|-----------------|
| 1 | 25 - 56 | ボリューム識別子 | (注1) |
| 2 | 57 - 58 | ボリューム順序番号 | 1 (DVD再生専用ディスク) |
| 3 | 59 - 60 | 最大ボリューム順序番号 | 1 (DVD再生専用ディスク) |
| 4 | 61 - 62 | 交換水準 | 2 (DVD再生専用ディスク) |
| 5 | 63 - 64 | 交換最大水準 | 2 (DVD再生専用ディスク) |
| 6 | 65 - 68 | 文字集合リスト | 1 |
| 7 | 69 - 72 | 文字最大集合リスト | 1 |

備考1 上記の内容の記録形式は、JIS X 0609-1998に従う。

備考2 上記以外の項目は、JIS X 0609-1998に従う。

(注1) ボリューム識別子は次のフォーマットで記録する。

| 項目番号 | | | | |
|-------|-----|-----|---------|-------|
| 長 さ | 2 | 2 | 4 | 3 |
| 内 容 例 | J P | H _ | 2 0 1 5 | 9 9 9 |

特許庁の識別：“ J P ”を記録する。

公報・公報情報種別の識別：

《公開公報・公表公報・再公表の場合》

“ G _ ”を記録する。

《特許公報の場合》

“ H _ ”を記録する。

《登録実用新案公報の場合》

“ U _ ”を記録する。

“ _ ”は(5 F)¹⁶。

西暦年： 西暦年を4桁で記録する。

年間の一連番号： 発行する媒体ごとに年間の一連番号を0 0 1から始まる3桁で記録する。(公報情報(公開公報、登録実用新案公報、公表公報、再公表、特許公報)は、インターネットを利用し発行する公開公報、登録実用新案公報、公表公報、再公表、特許公報の値を記録する。)

2.2 公報

- (1) インターネットからダウンロード可能な公報のファイルは2種類存在する。
- (2) 2種類のファイルは、Z i p形式(注1)及びT a r形式(注2)でアーカイブする(2種類のファイルはアーカイブ形式が異なるのみで内容については同一)。
- (3) ファイル名は、次のフォーマットで記録する。

| 項目番号 | | | | |
|------|-----|-----|---------|-------|
| 長さ | 2 | 2 | 4 | 3 |
| 内容例 | J P | U _ | 2 0 1 5 | 9 9 9 |

| 項目番号 | | |
|------|---|-------|
| 長さ | 1 | 3 |
| 内容例 | . | Z I P |

特許庁の識別：“ J P ”を記録する。

公報種別の識別：

公開公報・公表公報・再公表の場合 “ G _ ”を記録する。

特許公報の場合 “ H _ ”を記録する。

登録実用新案公報の場合 “ U _ ”を記録する。

西暦年：西暦年を4桁で記録する。

年間の一連番号：年間の一連番号を0 0 1から始まる3桁で記録する。

ピリオド“ . ”を記録する。

拡張子：

アーカイブ形式がZ i pの場合 “ Z I P ”

アーカイブ形式がT a rの場合 “ T A R ”

- (4) アーカイブされたファイルには、公報を構成するファイル及びディレクトリの情報が含まれている。
- (5) アーカイブされたファイルを展開すると『3. ファイル構成』で説明している構成になる。
- (6) アーカイブされたファイルに含まれるファイルのファイル名、ディレクトリ名は、「公報情報」のファイル名、ディレクトリ名の基準に準ずる。

- (注1) Z i p形式は、パーソナルコンピュータを含むコンピュータで広く普及しているアーカイブ形式の一つ。
- (注2) T a r形式 (Tape Archival and Retrieval format) は、U N I X系 O Sで広く普及しているアーカイブ形式の一つ。

例 1 公報種別の識別が登録実用新案公報の場合のファイル名

アーカイブ形式が Z i p の場合 JPU_2015999.ZIP

アーカイブ形式が T a r の場合 JPU_2015999.TAR

例 2 公報種別の識別が特許公報の場合のファイル名

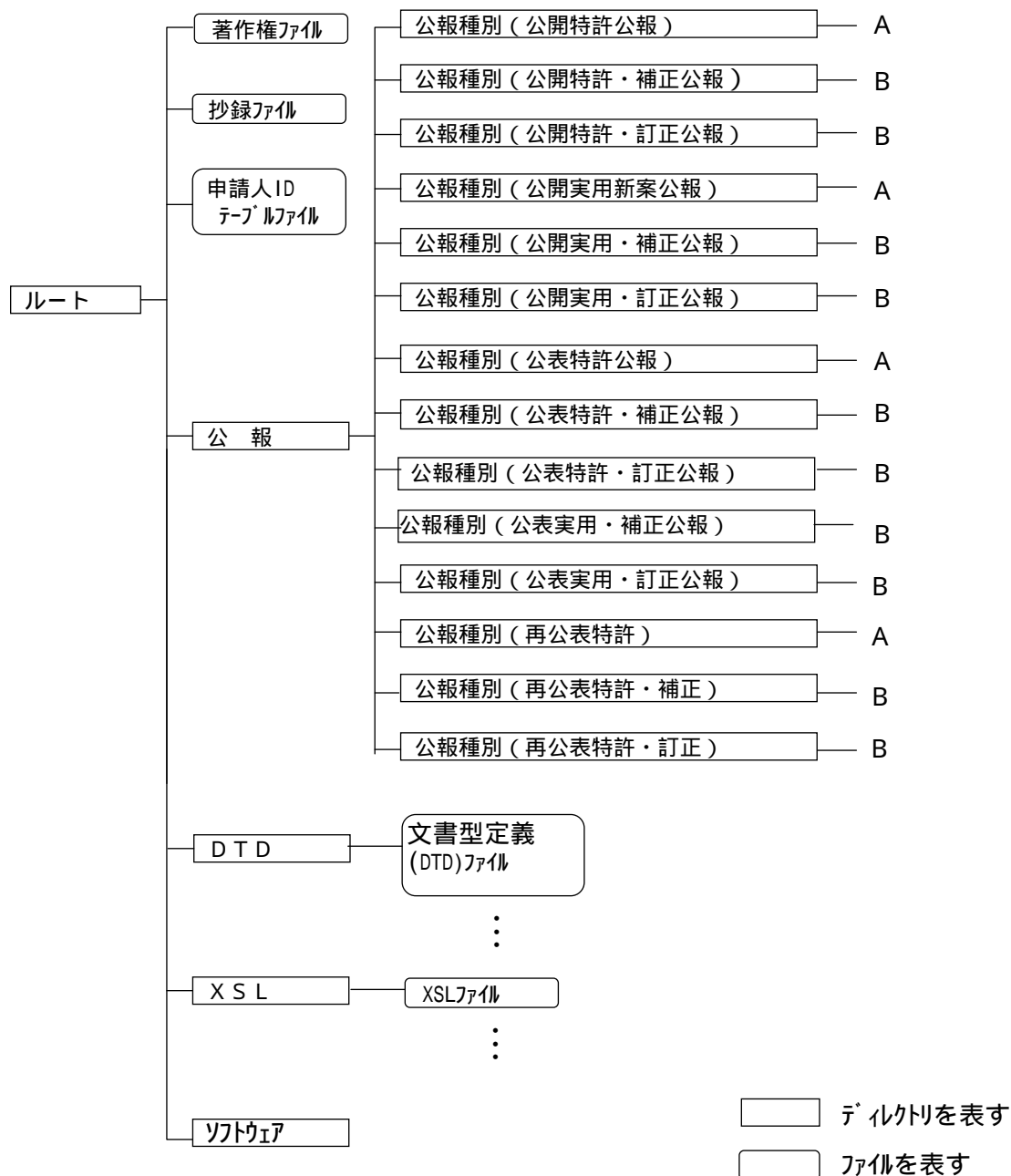
アーカイブ形式が Z i p の場合 JPH_2015999.ZIP

アーカイブ形式が T a r の場合 JPH_2015999.TAR

3. ファイル構成

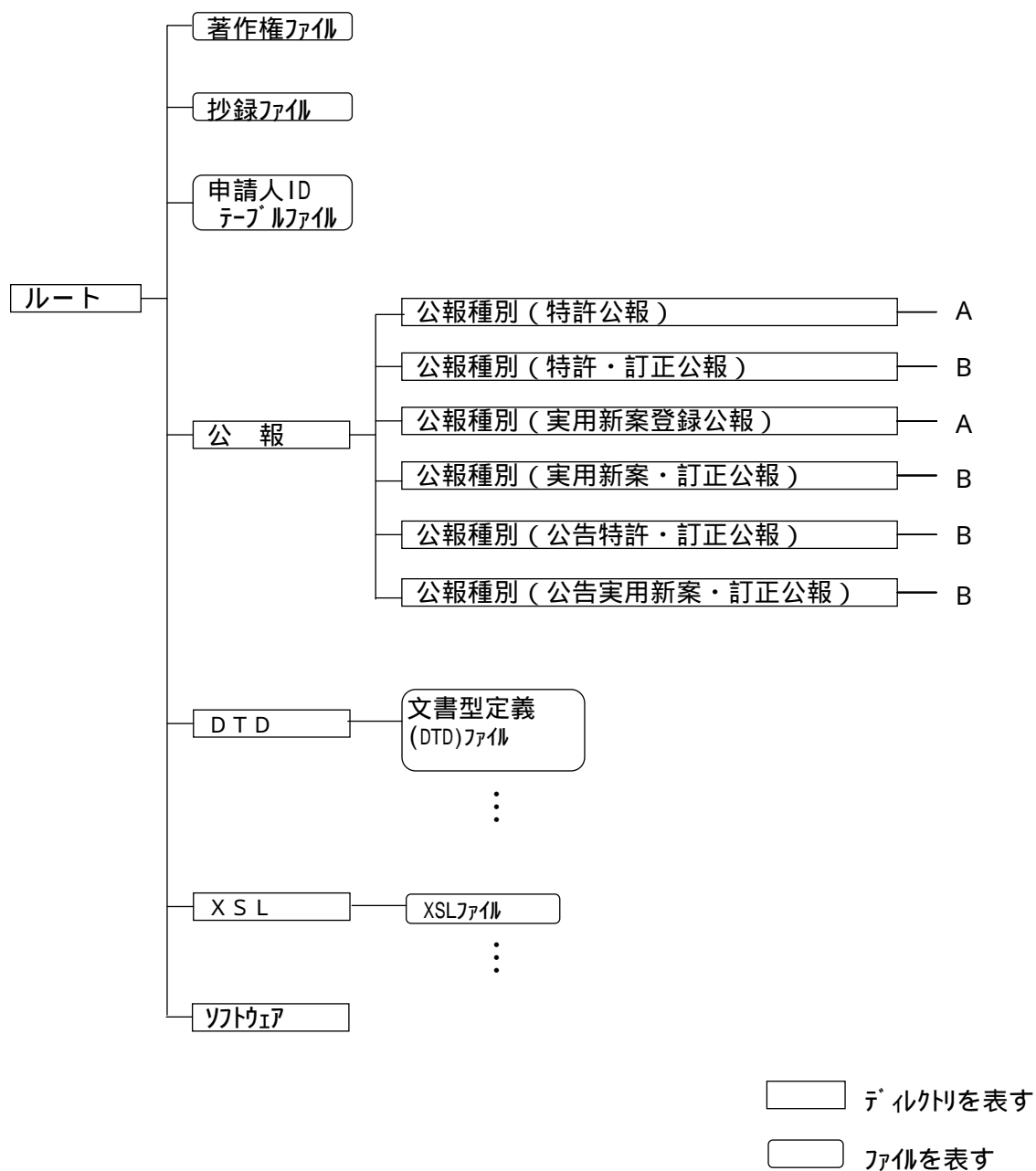
公開公報・公表公報・再公表のファイル構成を図3 - 1に、特許公報のファイル構成を図3 - 2に、登録実用新案公報の構成を図3 - 3に示す。各公報種別の下位ディレクトリのファイル構成を図3 - 4に示す。参照ファイルの構成を図3 - 5に示す。

図3 - 1 公開公報・公表公報・再公表のファイル構成



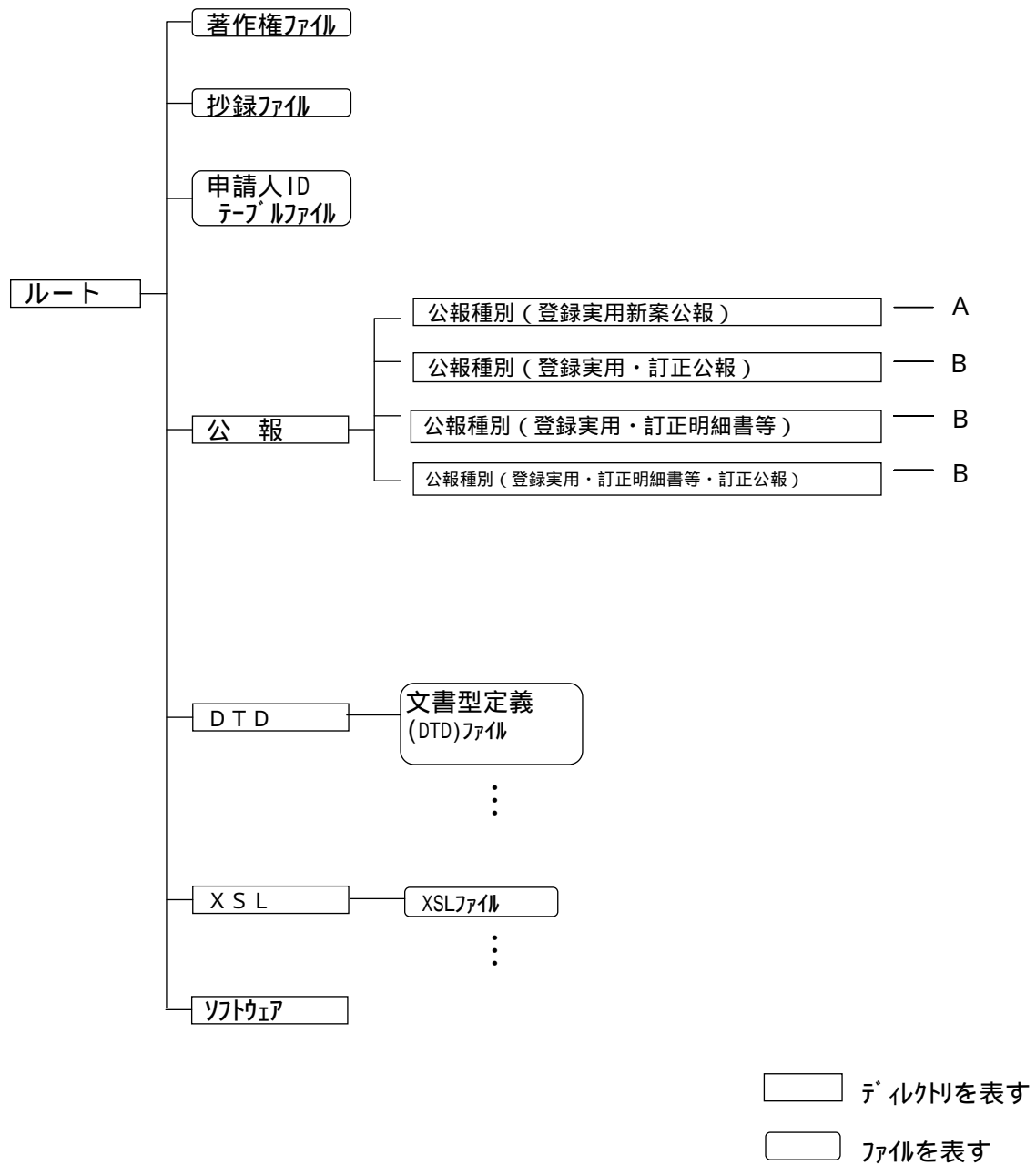
(注1) A Bの構成を図3 - 4に後述する

図3 - 2 特許公報のファイル構成



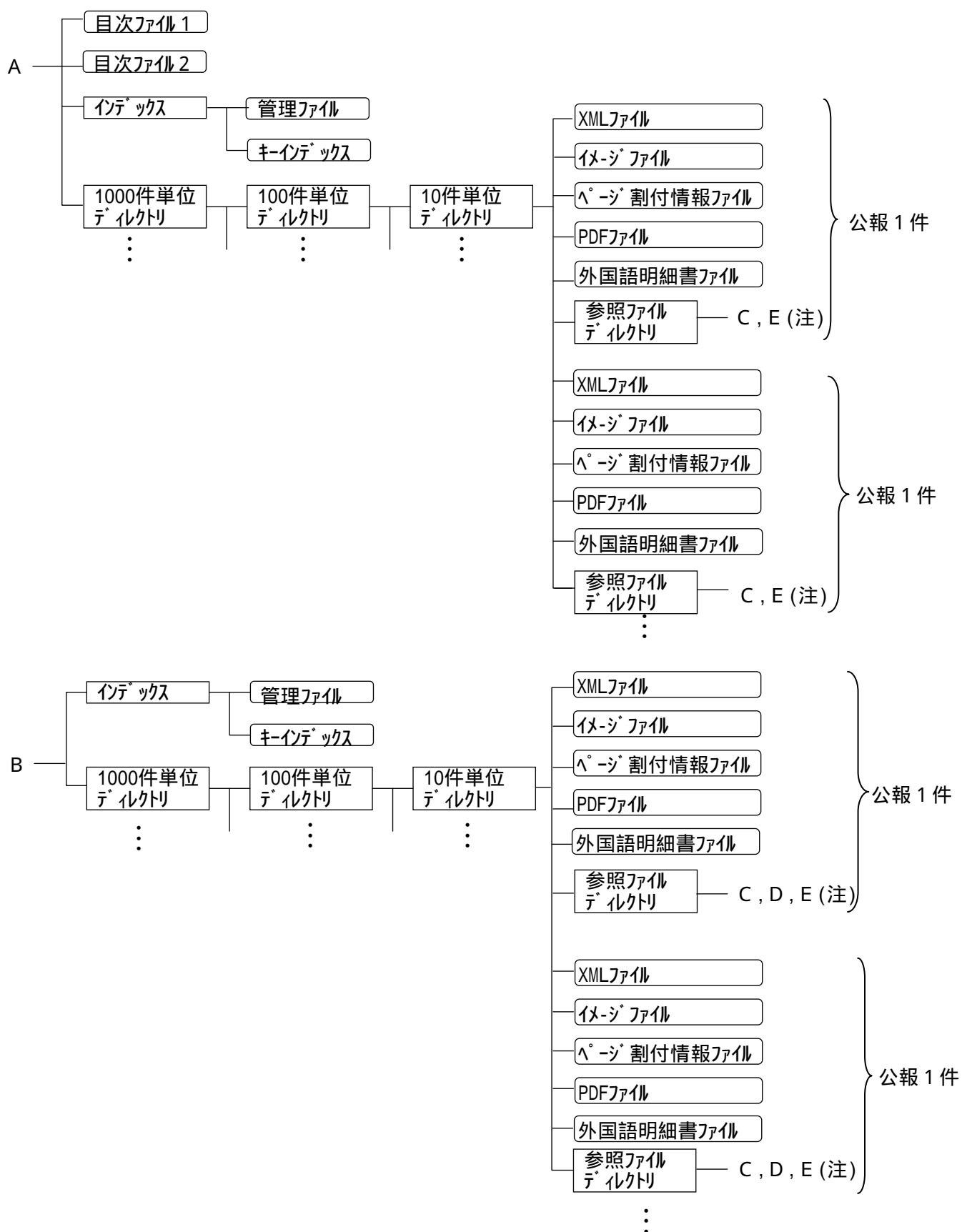
(注1) A B の構成を図3 - 4に後述する

図3 - 3 登録実用新案公報のファイル構成



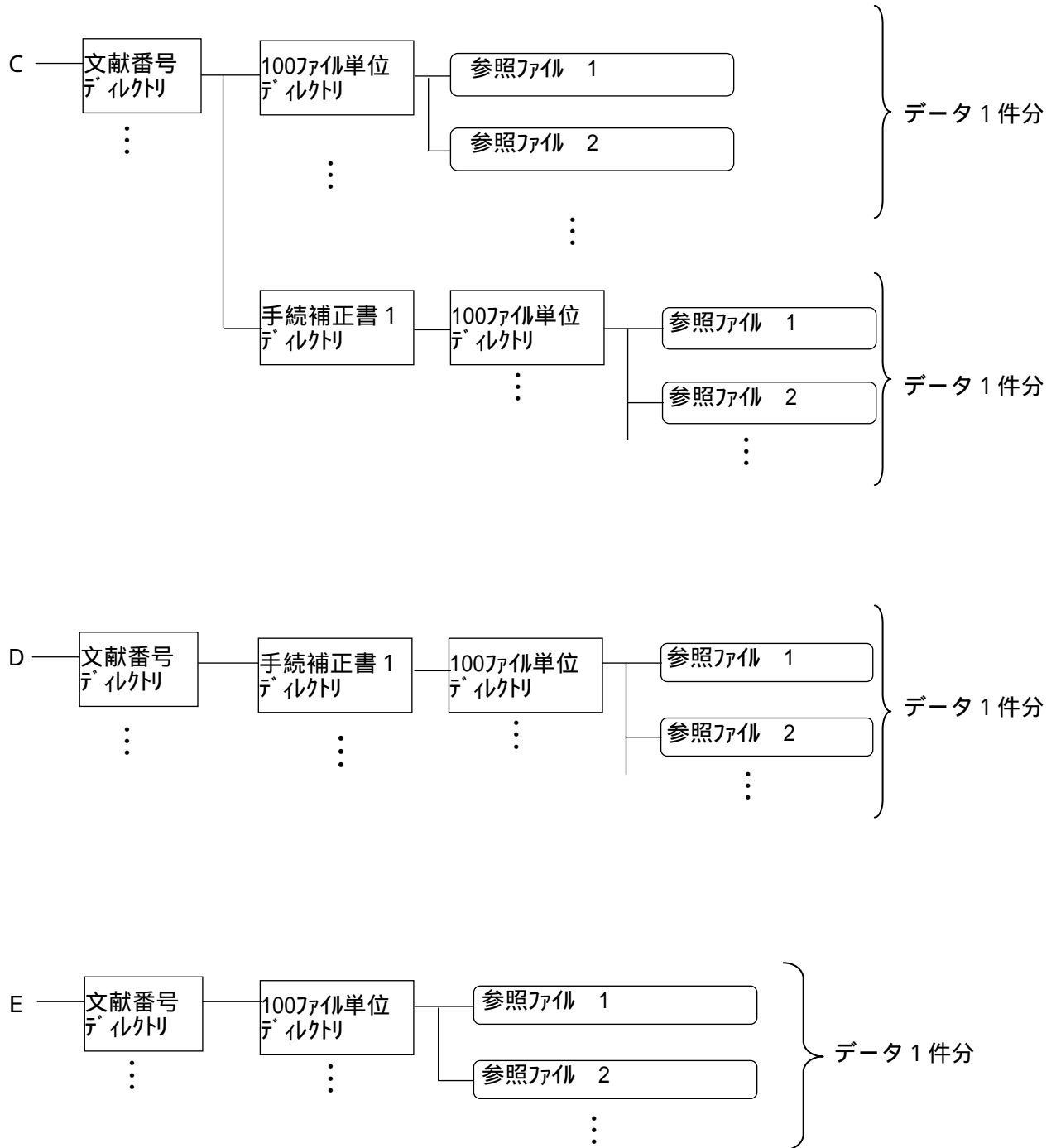
(注1) A B の構成を図3 - 4に後述する

図3 - 4 公報種別の下位ディレクトリのファイル構成



(注) 参照ファイルディレクトリの階層構造は図3 - 5を参照

図3 - 5 参照ファイルの構成



(注) 公開・登録実用新案・登録実用(訂正明細書等の掲載)・公表公報・再公表及びこれらの訂正公報においては“C”を参照。
 ただし、補正書に参照ファイルがない場合は“E”に同じ。また、本体に参照ファイルがなく、補正書で参照ファイルが追加された場合は“D”に同じ。
 上記の補正公報においては“D”を参照。
 特許公報及びその訂正公報、公告の訂正公報においては“E”を参照。

4. 各ディレクトリ及びファイルの名称と概要

『3. ファイル構成』で示した、各ディレクトリ及びファイルの概要を以下に示す。

ディレクトリ名とファイル名を〔 〕内に示す。ディレクトリ名とファイル名は、すべて1バイトコードの英数字を使用する。

(1) 著作権ファイル〔COPYRIGHT〕

著作権について記録するファイル。

(2) 抄録ファイル〔ABSTRACT.csv〕

公報仕様のバージョン、公報の発行日、公報を一意に識別するボリューム番号、公報発行当初からの一連番号、公報種別、文献番号の範囲、公報の件数等を記録するファイル。

(3) 申請人IDテーブルファイル〔IDTABLE.csv〕

公報に出現する申請人の名称と識別番号との対応を記録するファイル。

(4) 文書型定義(DTD)ディレクトリ〔DTD〕

文書型定義(DTD)ファイルを記録するディレクトリ。

(5) 文書型定義(DTD)ファイル〔xxxxxxxxxx.dtd〕

『TR X 0008-1999 XML 1.0』に準拠した、文書型定義(DTD)を記録するファイル。公報種別毎に必要となるDTDファイルについては表4-1を参照。

(6) XSLディレクトリ〔XSL〕

XSLファイルを記録するディレクトリ。

(7) XSLファイル〔nnnnnnnnnn.xml〕

XMLファイルを表示するためのスタイルシートファイル。公報種別毎に必要となるXSLファイルについては表4-1を参照。

(8) 公報ディレクトリ〔DOCUMENT〕

公報データを記録するディレクトリ。

(9) 公報種別ディレクトリ

公報の種別を示すディレクトリ。ディレクトリ名は公報の種別毎に異なる。公報種別とディレクトリ名の対応を表4-2に示す。記録する公報がある公報種別のみ存在する。

(10) 目次ファイル〔CONTENTS1.csv、CONTENTS2.csv〕

目次データを記録するファイル。公報種別が公開特許(実用新案)公報、登録実用新案公報、公表特許公報、再公表特許及び特許(実用新案登録)公報の場合に存在する。公報1件が1レコードに相当し、欠号公報のレコードも記録する。

CONTENTS1.csvは複数の国際特許分類及び出願人情報を記録するが、CONTENTS2.csvでは先頭

の情報のみを記録する。

(11) インデックスディレクトリ〔INDEX〕

検索用のインデックスを記録するディレクトリ。

(12) 管理ファイル〔CONTROL〕

キーインデックスファイルの情報を管理するファイル。インデックスディレクトリ下に存在する。

(13) キーインデックスファイル〔KEYINDEX〕

検索時に使用するインデックスファイル。インデックスディレクトリ下に存在する。

(14) 1 0 0 0 件単位ディレクトリ

公報 1 0 0 0 件毎に存在するディレクトリ。下位ディレクトリに存在し得る公報の文献識別番号の内、最も小さい番号をディレクトリ名とする。このため、ディレクトリ名の下 3 桁は、0 0 1 とする。

文献識別番号のフォーマットは以下のとおりとする。

《公開公報・公表公報・再公表、公告公報（訂正）の場合》

| 文 献 識 別 番 号 | |
|-------------|-----------------------|
| 西暦 4 桁 | 公開・公表・国際公開・公告番号の下 6 桁 |
| 4 桁 | 6 桁 |

《登録実用公報・特許公報の場合》

| 文 献 識 別 番 号 |
|-----------------|
| 特許番号または実用新案登録番号 |
| 1 0 桁 |

〔例〕 ディレクトリ名“0002500001”の最下層には、文献識別番号 0002500001 から 0002501000 の公報を記録する。

同様に、ディレクトリ名“0002501001”の最下層には、文献識別番号 0002501001 から 0002502000 の公報を記録する。

(15) 1 0 0 件単位ディレクトリ

公報 1 0 0 件毎に存在するディレクトリ。下位ディレクトリに存在し得る公報の文献識別番号（(14)参照）の内、最も小さい番号をディレクトリ名とする。このため、ディレクトリ名の下 2 桁は、0 1 とする。

〔例〕 ディレクトリ名“0002500001”の最下層には、文献識別番号 0002500001 から

0002500100の公報を記録する。

同様に、ディレクトリ名“0002500101”の最下層には、文献識別番号0002500101から0002500200の公報を記録する。

(16) 10件単位ディレクトリ

公報10件毎に存在するディレクトリ。本ディレクトリ下に存在し得る公報の文献識別番号((14)参照)の内、最も小さい番号をディレクトリ名とする。このため、ディレクトリ名の下1桁は、1とする。

[例] ディレクトリ名“0002500001”の下層には、文献識別番号0002500001から0002500010の公報を記録する。

同様に、ディレクトリ名“0002500011”の下層には、文献識別番号0002500011から0002500020の公報を記録する。

(17) XMLファイル {nnnnnnnnnn.xml}

公報1件分の文字データを記録するファイル。欠号公報のXMLファイルは存在しない。ファイル名の内“nnnnnnnnnn”は文献識別番号((14)参照)と一致する。

(18) イメージファイル {nnnnnnnnnn.tif}

公報1件分のイメージデータを記録するファイル。欠号公報のイメージファイルは存在しない。ファイル名の内“nnnnnnnnnn”は文献識別番号((14)参照)と一致する。

(19) ページ割付情報ファイル {nnnnnnnnnn.pos}

公報1件分の位置情報を記録するファイル。欠号公報のページ割付情報ファイルは存在しない。ファイル名の内“nnnnnnnnnn”は文献識別番号((14)参照)と一致する。

(20) PDFファイル {nnnnnnnnnn.pdf}

公報1件分のレイアウトを表示するファイル。欠号公報のPDFファイルは存在しない。ファイル名の内“nnnnnnnnnn”は文献識別番号((14)参照)と一致する。

(21) 外国語明細書ファイル {nnnnnnnnnnmmmmm.pdf}

1公報に関連する外国語明細書のファイル。公開特許公報の場合のみ存在する。外国語明細書のファイルの内、PDFファイルのみを対象とする。

ファイル名の内“nnnnnnnnnn”は文献識別番号((14)参照)と一致する。

“mmmmm”は000001からの通番である。

(22) 文献番号ディレクトリ(参照ファイルディレクトリ) {nnnnnnnnnn}

1公報に関連する参照ファイルが存在するディレクトリ。ディレクトリ名の“nnnnnnnnnn”は文献識別番号((14)参照)と一致する。

(23) 100件単位ディレクトリ {nnnnnnnnnnmmmmm} (参照ファイル用)

1公報に関連する参照ファイルが100件毎に存在するディレクトリ。

ディレクトリ名の内 “ nnnnnnnnnn ” は文献識別番号 ((14)参照) と一致する。

“ mmmmmm ” は000001からの通番である。

[例] ディレクトリ名 “ 0002500001000001 ” の下層には、ファイル名0002500001000001から0002500001000100の参照ファイルを記録する。

同様に、ディレクトリ名 “ 0002500001000101 ” の下層には、ファイル名0002500001000101から0002500001000200の参照ファイルを記録する。

(24) 手続補正書ディレクトリ [AMENnnnnn]

手続補正書を示すディレクトリ。ディレクトリ名の内 “ nnnnn ” は、 5 桁の数字とする。

(25) 参照ファイル [nnnnnnnnnnnmmmmm.app (又はpdf、xml、tif、jpg 等)] (注)

P D F ファイルからリンクが貼られている塩基配列又はアミノ酸配列表等のデータを記録するファイルの総称。公報 1 件に対して 0 ~ 複数存在する。欠号公報の参照ファイルは存在しない。ファイル名の内 “ nnnnnnnnnn ” は文献識別番号 ((14)参照) と一致する。

“ mmmmmm ” は000001からの通番である。

(注) A P P ファイル、P D F ファイル、X M L ファイル、T I F F ファイル又はJ P E G ファイル等が参照ファイルとなる。

(26) ソフトウェアディレクトリ [SOFTWARE]

特許庁提供の汎用ソフトウェアを収納するディレクトリ。

[例] 文献識別番号 0002500001 から 0002503000 までの 3,000 件の特許公報が記録される場合の、1,000 件単位のディレクトリ、100 件単位のディレクトリ、10 件単位のディレクトリ、X M L ファイル、イメージファイル、ページ割付情報ファイル、P D F ファイル、参照ファイルの関係を図4 - 1に示す。

表4 - 1 公報種別と D T D ファイル、X S L ファイルとの関係 (1 / 2)

| NO | 公報種別 | DTD | XSL |
|----|--------------------|---|----------------------------|
| 1 | 全公報種別共通 | wipo.ent mathml2.dtd soextblx.dtd | application-body.xsl |
| | | sequence-list.dtd | sequence-list.xsl |
| | | - | gat-JS-com.xsl |
| | | - | gat-com.xsl |
| 2 | 公開特許公報 公開実用新案公報 | gat-a.dtd gat-a-com.dtd | gat-a.xsl gat-a-com.xsl |

表4 - 1 公報種別とD T Dファイル、X S Lファイルとの関係 (2/2)

| NO | 公報種別 | DTD | XSL |
|----|----------------------------|------------------------------|------------------------------|
| 3 | 公開特許公報（補正） 公開実用新案公報（補正） | gat-a5.dtd gat-a-com.dtd | gat-a5.xsl gat-a-com.xsl |
| 4 | 公開特許公報（訂正） 公開実用新案公報（訂正） | gat-a6.dtd gat-a-com.dtd | gat-a6.xsl gat-a-com.xsl |
| 5 | 登録実用新案公報 | gat-u9.dtd gat-u-com.dtd | gat-u9.xsl gat-u-com.xsl |
| 6 | 登録実用新案公報（訂正） | gat-uc.dtd gat-u-com.dtd | gat-uc.xsl gat-u-com.xsl |
| 7 | 登録実用新案・訂正明細書等 | gat-u7.dtd gat-u7-com.dtd | gat-u7.xsl gat-u7-com.xsl |
| 8 | 登録実用新案・訂正明細書等 （訂正） | gat-ub.dtd gat-u7-com.dtd | gat-ub.xsl gat-u7-com.xsl |
| 9 | 公表特許公報 | gat-t.dtd gat-t-com.dtd | gat-t.xsl gat-t-com.xsl |
| 10 | 公表特許公報（補正） 公表実用新案公報（補正） | gat-t5.dtd gat-t-com.dtd | gat-t5.xsl gat-t-com.xsl |
| 11 | 公表特許公報（訂正） 公表実用新案公報（訂正） | gat-t6.dtd gat-t-com.dtd | gat-t6.xsl gat-t-com.xsl |
| 12 | 再公表特許 | gat-s.dtd gat-s-com.dtd | gat-s.xsl gat-s-com.xsl |
| 13 | 再公表特許（補正） | gat-s5.dtd gat-s-com.dtd | gat-s5.xsl gat-s-com.xsl |
| 14 | 再公表特許（訂正） | gat-s6.dtd gat-s-com.dtd | gat-s6.xsl gat-s-com.xsl |
| 15 | 特許公報 実用新案登録公報 | gat-b9.dtd gat-b-com.dtd | gat-b9.xsl gat-b-com.xsl |
| 16 | 特許公報（訂正） 実用新案登録公報（訂正） | gat-bc.dtd gat-b-com.dtd | gat-bc.xsl gat-b-com.xsl |
| 17 | 公告特許公報（訂正） 公告実用新案公報（訂正） | gat-b6.dtd | gat-b6.xsl |

表4 - 2 公報種別とディレクトリ名との対応

| 公 報 種 別 (日 本 語 名 称) | | 公報種別 | ディレクトリ名 |
|--------------------------|---------------------|--------|---------|
| 公 開 公 報 | 《公開特許公報関連》 | | |
| | 公開特許公報 | A | A |
| | 補正の掲載（公開特許公報）(注1) | A5 | A5 |
| | 訂正（公開特許公報） | A6 | A6 |
| | 《公開実用新案公報関連》 | | |
| | 公開実用新案公報 | U | U |
| | 補正の掲載（公開実用新案公報）(注2) | U5 | U5 |
| | 訂正（公開実用新案公報） | U6 | U6 |
| 新 登 案 録 実 用 公 報 | 《登録実用新案公報関連》 | | |
| | 登録実用新案公報 | U | U9 |
| | 訂正（登録実用新案公報） | U6 | UC |
| | 訂正明細書等の掲載（注4） | U7 | U7 |
| | 訂正（訂正明細書等の掲載） | U6 | UB |
| 公 表 公 報 | 《公表特許公報関連》 | | |
| | 公表特許公報 | A | T |
| | 補正の掲載（公表特許公報）(注1) | A5 | T5 |
| | 訂正（公表特許公報） | A6 | T6 |
| | 《公表実用新案公報関連》(注3) | | |
| | 補正の掲載（公表実用新案公報）(注1) | U5 | TA |
| | 訂正（公表実用新案公報） | U6 | TC |
| 再 公 表 | 《再公表特許関連》(注3) | | |
| | 再公表特許 | A1 | S |
| | 補正の掲載（再公表特許）(注1) | A5 | S5 |
| | 訂正（再公表特許） | A6 | S6 |
| 公 告 公 報 | 《公告特許公報関連》 | | |
| | 訂正（公告特許公報） | B6 | B6 |
| | 《公告実用新案公報関連》 | | |
| | 訂正（公告実用新案公報） | Y6 | Y6 |
| 特 許 公 報 | 《特許公報関連》 | | |
| | 特許公報 | B2, B1 | B9 |
| | 訂正（特許公報） | B6 | BC |
| | 《実用新案登録公報関連》 | | |
| | 実用新案登録公報 | Y2, Y1 | Y9 |
| | 訂正（実用新案登録公報） | Y6 | YC |

備考：“補正の掲載”は、本仕様中で“補正公報”と称することがある。

“訂正”は、本仕様中で“訂正公報”と称することがある。

(注1) 特許法第17条の2の規定による補正の掲載。

(注2) 旧実用新案法第55条第2項において準用する特許法第17条の2の規定による補正の掲載。

- (注3) 公表実用新案公報、再公表実用、補正の掲載（再公表実用関係）、訂正（再公表実用関係）は2000年以降発行されない。
- (注4) 実用新案法第14条の2の規定による訂正明細書等の掲載。

図4 - 1 文献識別番号 0002500001 から 0002503000 までの特許公報の例

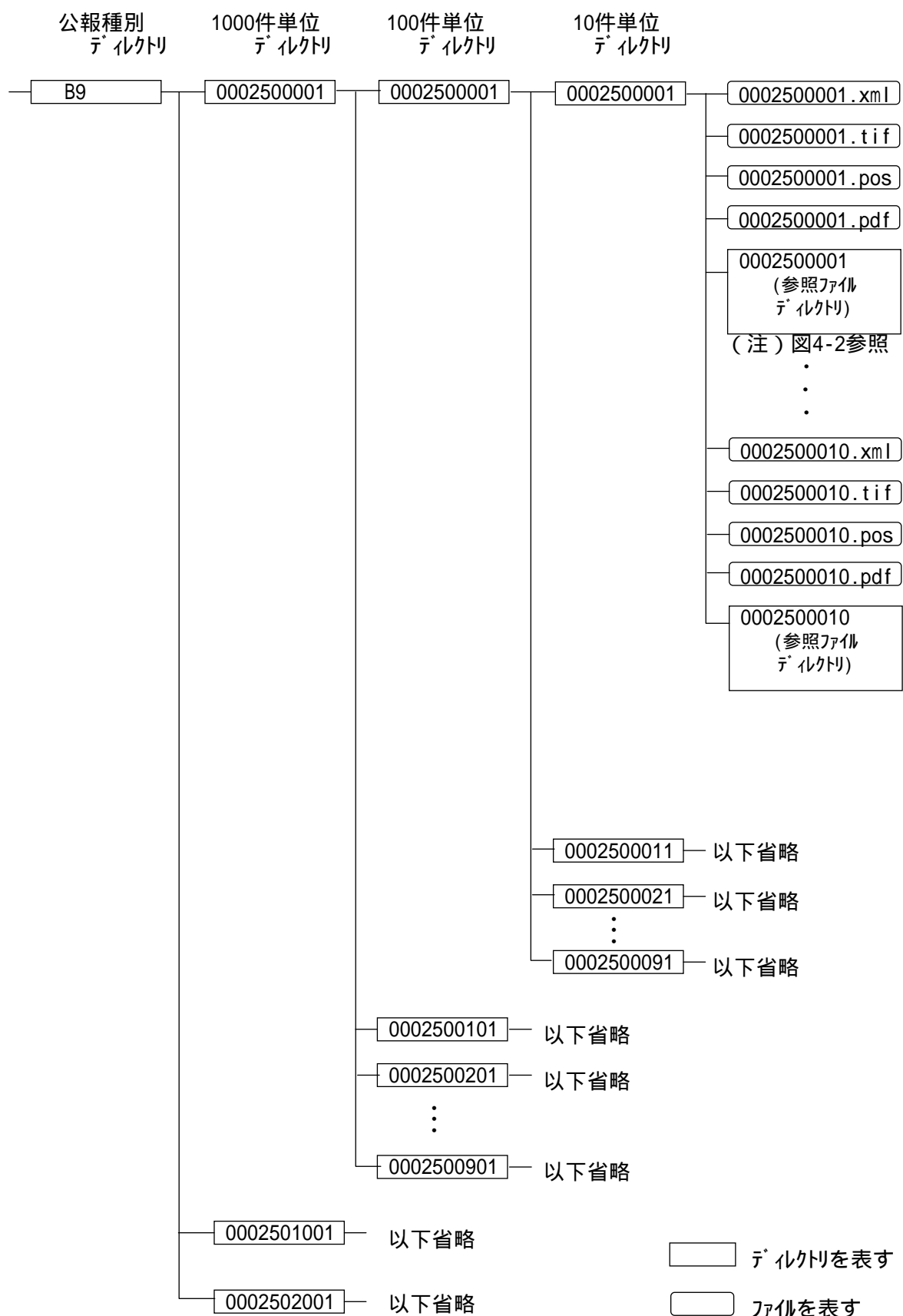
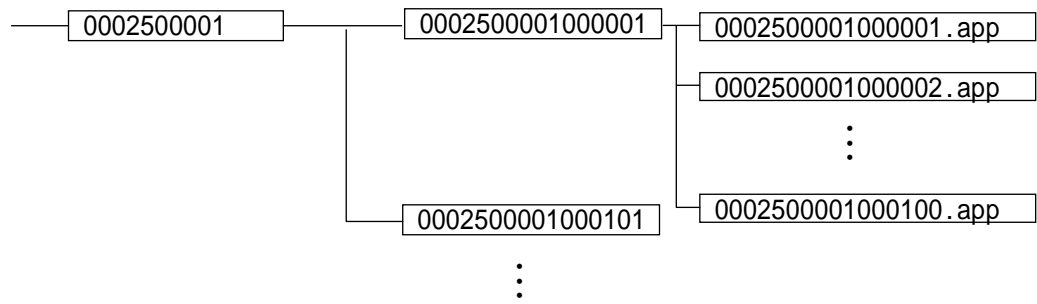


図4 - 2 参照ファイルディレクトリ構成の例



5. ファイル形式

公報内のファイルは、テキストファイル形式（CSVファイル形式含む）またはバイナリファイル形式のいずれかのファイル形式とする。

5.1 テキストファイル形式

ファイル内のデータを、『6. 文字コード』で定めるコードデータだけで構成するファイル形式をいう。

著作権ファイル、抄録ファイル、申請人IDテーブルファイル、目次ファイル、XMLファイル、XSLファイル、文書型定義（DTD）ファイル、参照ファイルは、テキストファイル形式である。

5.2 バイナリファイル形式

ファイル内のデータを、『6. 文字コード』で定めるコードデータと、バイナリデータで構成するファイル形式をいう。

バイナリデータの記録順序は、最下位バイト先頭形式とする。すなわち、16ビット数の数値 $(wxyz)_{16}$ は2バイト欄に $(yzwx)_{16}$ と記録し、32ビット数の数値 $(stuvwxyz)_{16}$ は4バイト欄に $(yzwxuvst)_{16}$ と記録する。

[例]

- ・10進数4660 $((1234)_{16})$ は、 $(3412)_{16}$ と記録する。
- ・10進数305419896 $((12345678)_{16})$ は、 $(78563412)_{16}$ と記録する。

管理ファイル、キーインデックスファイル、イメージファイル、ページ割付情報ファイル、PDFファイルは、バイナリファイル形式である。

（注）参照ファイルに関しては、テキストファイル形式、バイナリファイル形式の両方の形式が存在する。

6. 文字コード

6.1 文字コード

図形文字は、8単位コード系とする。

(1) 1バイトコード

『JIS X 0201-1976 7ビット及び8ビットの情報交換用符号化文字集合』のローマ文字用図形文字集合（JIS X 0201-1976 表2）をG0集合に指示し、列2～7へ呼び出したものを使用する。

(2) 2バイトコード

『JIS X 0208-1997 7ビット及び8ビットの2バイト情報交換用符号化漢字集合』をG1集合に指示し、列10～15へ呼び出したもの（JIS X 0208-1990 の図形文字符号の各バイトの8ビット目を“1”にしたもの）を使用する。

（注1）2バイトコードの間隔（スペース）は用いず、1バイトコードの間隔（スペース）を使用する。

（注2）外字は使用しない。外字が使用された部分に対しては、置き換え文字が使用される。置き換え文字以外は黒ベタ（ ）となる。置き換え文字は、“ ”と“ ”の間に置き換える文字を置くことにより表現する。

（注3）旧Xフォーマットによる手続分の文字変形制御コードを利用した半角文字表記の2バイト文字コードについては、一律に制御コードを除外して全角文字表記の2バイト文字コードで収録から、対応する1バイト文字コードに変換して収録へ変更する（片仮名及び仮名記号は、全角文字表記の2バイト文字コードのまま収録する。）。

表6.1 - 1 文字コード

| 名 称 | 記 号 | 符号化表現 |
|----------|-----|---------------------------|
| 黒ベタ | | (A 2 A 3) ₁₆ |
| 置き換え文字開始 | | (A 2 A 5) ₁₆ |
| 置き換え文字終了 | | (A 2 A 7) ₁₆ |

6.2 制御文字コード

復帰（CR）、改行（LF）、TAB（HT）を用いる。

表6.2 - 1 制御文字コード

| 名 称 | 記 号 | |
|-----|-----|-----------------------|
| 復帰 | CR | (0 D) ₁₆ |
| 改行 | LF | (0 A) ₁₆ |
| TAB | HT | (0 9) ₁₆ |

（注4）TABは参照ファイルにのみ使用する。

6.3 空白 (WhiteSpace) の扱いについて

X M L 規格では、半角スペース (20)₁₆、タブ (09)₁₆、L F (0A)₁₆、C R (0D)₁₆の4文字を総称して空白 (WhiteSpace) と呼ぶが、公報データ、公報情報データのX M Lでは、空白を半角スペース、L F、C Rだけに限定し、挿入可能位置も下記の場所のみとする。

X M L のセパレータとして認識される位置

(セパレータとはタグ名と属性名の区切りなどに使用されるものの総称)

開始タグの直前直後

終了タグの直前直後

空エレメントタグの直前直後