

UCCF 仕様案

1. 概要

電子出版の配信プラットフォームや電子書店は、多くの事業者により構築されている。それらの配信プラットフォームでは、種々の書式の電子出版コンテンツをそれぞれ独自の DRM により暗号化して配信している。そのため、これらの電子出版コンテンツは、対応する UI 環境でのみ解読、表示することができるという構造になっている。そのため、電子出版コンテンツはその対応 UI に紐づけられて管理されるので、UI と独立して管理、流通することができない。そのため、UI が対応しない端末環境では閲覧することができないなど、読者の利便性を損なう結果となっている。

そこで、電子出版コンテンツに、そのコンテンツの内容を示すメタデータを常に附属させ、そのメタデータは暗号化することなく、いつでも取得、表示できるようにすることで、共通化した管理を実現することのできる UCCF (Universal Content Container Format) を提案する。このフォーマットが普及すれば、DRM 方式を含めオープン型電子出版 UI の実現が容易になる。複数の端末で実行することのできるオープン型電子出版 UI が実現されれば、読者が所有する端末の電子出版コンテンツを一覧表示し、目的とするコンテンツにアクセスすることを容易とすることが可能になる。

2. UCCFの仕様案

UCCF は、概要に述べたように、どのような形式で格納されているコンテンツでも、その内容を示すメタデータと共に格納し、取り扱うことを可能とするフォーマットである。具体的には、図のように、電子出版コンテンツと、その内容をあらわすメタデータを共に ZIP ファイル形式で格納して取り扱うものである。

電子出版コンテンツは、そのフォーマットは問わないため、DRM などの技術により暗号化されていても良い。メタデータには、改ざんを検出するための電子署名が付加されているので、改ざんされていないことを保証することができる。また、メタデータと電子出版コンテンツの対応関係を保証するために、電子出版コンテンツのハッシュ値をメタデータに持ち、そのメタデータが、確かに対応する電子出版コンテンツのものであることを保証する。

メタデータは暗号化せず、平文で持ち、また ZIP ファイルシステムの最初のファイルとすることで、容易に取得することができるので、その UCCF が格納している電子出版コンテンツがどのようなものであるかを知って適切に取り扱うことができる。

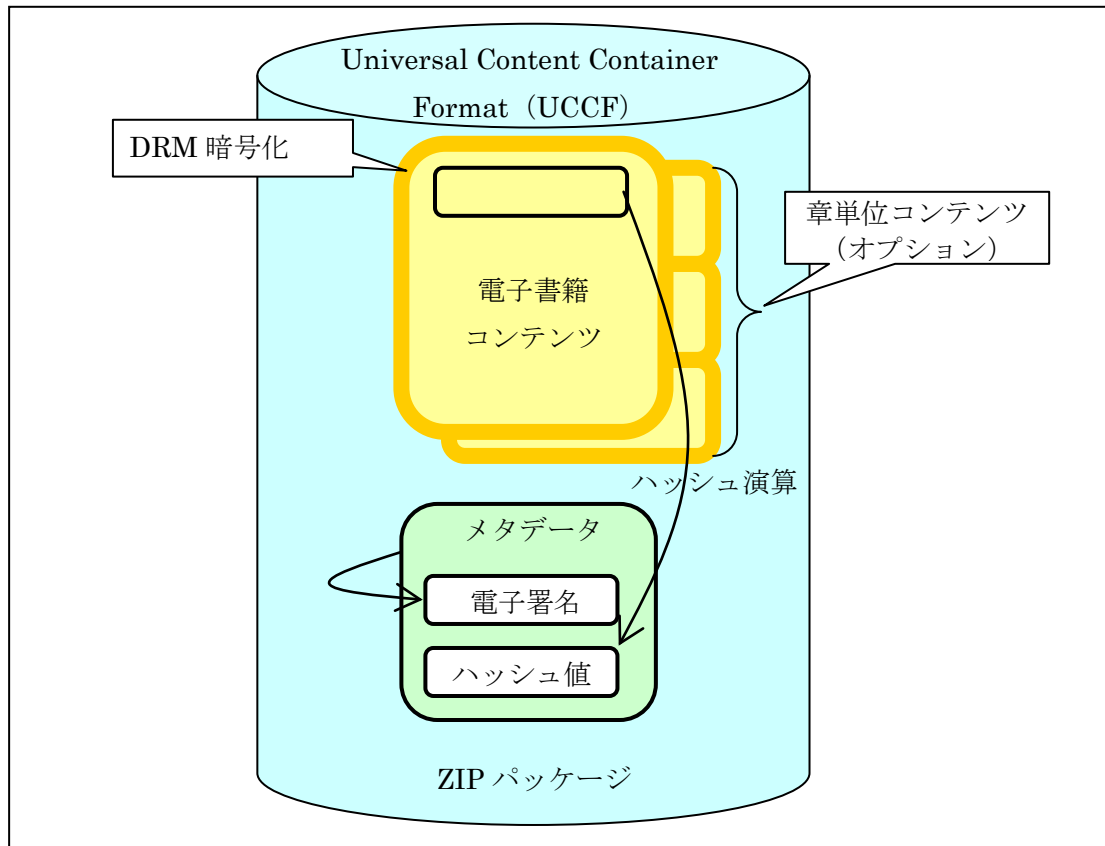


図 1 UCCF の概要

3. メタデータフォーマット仕様案

UCCF のメタデータは XML により記述され、以下の構造を持つ。

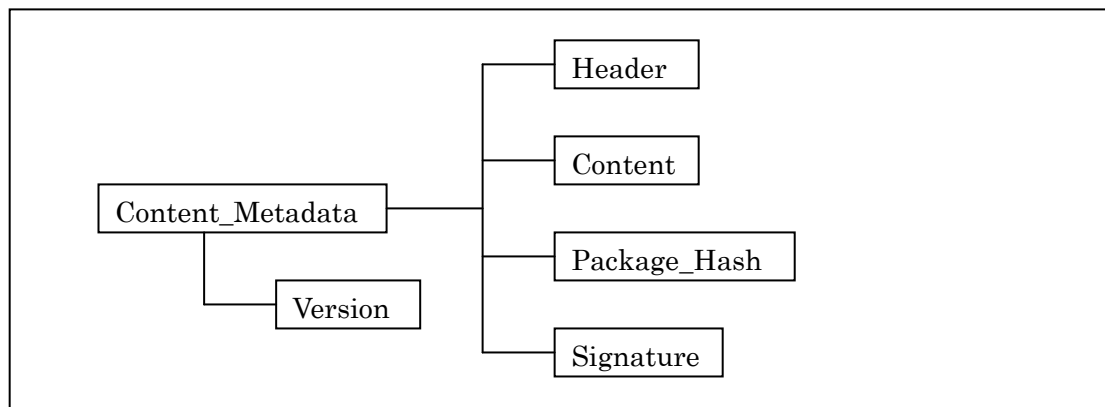


図 2

メタデータは、「content_metadata.xml」のファイル名で、UCCF の最初のファイルとして非圧縮で ZIP 化する。このことにより、メタデータの開始位置を特定できるので、UCCF 全体を UNZIP しなくても、メタデータを取り出すことができる。

ルート要素である「Content_Metadata」は以下の子要素と「Version」パラメータを持つ。

- ・ Header
- ・ Content
- ・ Package_Hash
- ・ Signature

表 1

名称	型	必須か	意味
Version	文字列	必須	本メタデータのバージョンを表す

(1) Header

Header は以下のパラメータを持つ。

表 2

名称	型	必須か	意味
NumberOfContents	正整数	必須	Content の個数を表す
platform	文字列	必須	配信プラットフォームの名前を表す

例：

<Header NumberOfContents="1" platform="オープン型 DRM / UI 実証実験"/>

(2) Content

Content は Dublin Core に準拠する書誌情報の子要素としてもち、以下のパラメータを持つ。

表 3

名称	型	必須か	意味
Chapter	正整数	必須	記述する章の番号を表す。0 の場合は、コンテンツ全体のメタデータであることを示す。
file_name	文字列	必須	対応するコンテンツのファイル名

複数の章のファイルによりコンテンツが構成されている場合、Chapter は章番号を表し、Content はその章のメタデータであることを示す。Chapter が 0 の場合は、Content はコンテンツ全体のメタデータであることを示す。

Content は Dublin Core に準拠する以下の子要素を持つ。

- ① identifier (ID)
- ② title (書名)
- ③ creator (著者)
- ④ contributor (貢献者)
- ⑤ publisher (出版者)
- ⑥ subject (主題)
- ⑦ description (内容の説明)
- ⑧ coverage (リソースの範囲もしくは対象)
- ⑨ format (リソースの形態)
- ⑩ type (ジャンル)
- ⑪ date (日時)
- ⑫ relation (関連するリソースへの参照)
- ⑬ source (リソースの派生元リソースへの参照)
- ⑭ rights (権利)
- ⑮ language (言語)

(3) Package_Hash

Package_Hash は、このメタデータが対応する電子出版コンテンツのものであることを証明するため、対応する電子出版コンテンツデータのハッシュ値を格納する。

Package_Hash は、次のパラメータを持つ。

表 4

名称	型	必須か	意味
start	16 進数	必須	ハッシュ値の計算領域の開始アドレス
length	16 進数	必須	ハッシュ値の計算領域の長さ
interval	16 進数	必須	次のハッシュ値計算領域までの距離
repeatCount	16 進数	必須	ハッシュ値計算領域の数。0 の場合はファイルの最後まで繰り返す
type	文字列	必須	ハッシュ値計算方式

電子出版コンテンツの容量が大きい場合に、計算負荷を下げるためにハッシュ値計算対象領域のサイズを小さくし、かつ、コンテンツ全域にわたってハッシュ値を計算できるようにするために、ハッシュ値を計算する対象領域を指定できる。

start、length、interval、repeatCount の値がすべて 0 の場合はコンテンツファイル全体を指定するものとする。

対象とするコンテンツファイルは、chapter が 0 のメタデータで指定されたものとする。

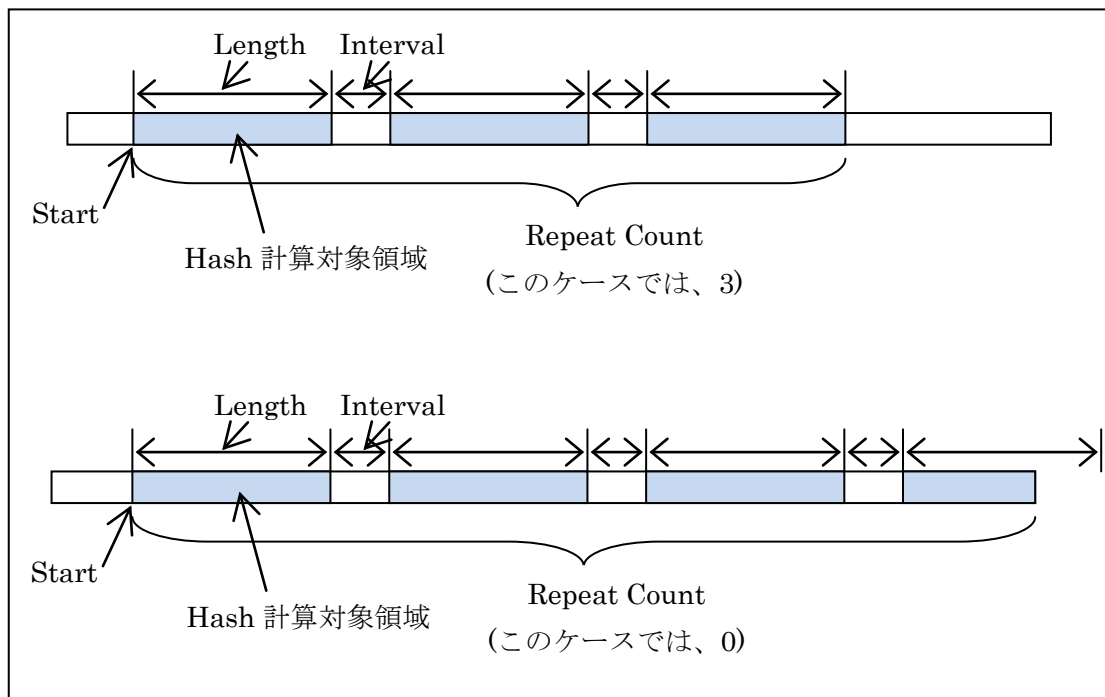


図 3

例：ハッシュ計算領域が、対象ファイルの先頭から1024バイトのみで、ハッシュ値が16進数で「AAAAAAAAAAAA」の場合。

```
<Package_Hash start="0" length="400" interval="0" repeatCount="1"
type="MD5">AAAAAAAAAAAA</Package_Hash>
```

(4) Signature

メタデータ自身の電子署名を Enveloped 署名形式で格納する。メタデータが改ざんされていないことを確認するために使う。

4. 仕様案 (例)

以上にあげた例をまとめたものを以下に示す。

7 行目の行の URL は仮アドレスである。

```
<?xml version="1.0" encoding="UTF-8"?>
<Content_Metadata xmlns="urn:ucf:ucf:schema:2011"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:ucf:ucf:schema:2011
http://www.impress.co.jp/2011/ucfContentMetadata.xsd" Version="1.0">
  <Header NumberOfContents="1" platform="オープン型 DRM / UI 実証実験"/>
  <Content Chapter="1" file_name="report.xhtml">
    <dc:identifier>4844321234</dc:identifier>
    <dc:title>サンプル書籍</dc:title>
    <dc:creator>インプレスR&D</dc:creator>
    <dc:contributor></dc:contributor>
    <dc:publisher>インプレスR&D</dc:publisher>
    <dc:subject>電子出版</dc:subject>
    <dc:description>電子出版に関するサンプル</dc:description>
    <dc:coverage></dc:coverage>
    <dc:format>EPUB</dc:format>
    <dc:type>Text</dc:type>
    <dc:date>2011-02-06</dc:date>
    <dc:relation></dc:relation>
    <dc:source></dc:source>
    <dc:rights>インプレスR&D</dc:rights>
    <dc:language>jp</dc:language>
  </Content>
  <Package_Hash start="0" length="400" interval="0" repeatCount="1"
type="MD5">AAAAAAAAAAAA</Package_Hash>
  <Signature>AAAAAAAAAAAA</Signature>
</Content_Metadata>
```