

汎用アノテーション記述言語 MAML

伊藤 一成[†] 斎藤 博昭[†]

[†] 慶應義塾大学 大学院理工学研究科 〒223-8522 神奈川県横浜市港北区日吉 3-14-1
E-mail: {k_ito,hxs}@nak.ics.keio.ac.jp

あらまし 近年、メディアデータを効率よく検索したり要約する手段として、メタデータ技術が注目を集めている。我々は、メディアデータのための汎用アノテーション記述言語 MAML (Multimedia Annotation Markup Language) を策定した。MAML では、アノテーションの対象は実ファイルだけではなく、人物、オブジェクト、イベント等までも含めて、画一的な仕様で記述可能な点が大きな特徴である。いうまでもなく、ファイルの種類やフォーマットも限定しない。さらに、アノテーションデータの生成者及び利用者たる人間が、理解・記述しやすい文章中心の表現構造を採用している。MAML を適用することにより、メディアの種類別に依存しない様々な応用アプリケーションの実現が期待できる。

キーワード メタデータ, アノテーション, XML

MAML: A Generalized Annotation Description Language

Kazunari ITO[†] and Hiroaki SAITO[†]

[†] Department of Science and Technology, Keio University
Hiyoshi 3-14-1, Kouhoku-ku, Yokohama, Kanagawa, 223-8522 Japan
E-mail: {k_ito,hxs}@nak.ics.keio.ac.jp

Abstract A meta data technology calls our attention as means of retrieving or summarizing multimedia information efficiently. This paper proposes MAML (Multimedia Annotation Markup Language), a generalized annotation description format for media data. The annotated subjects in MAML include not only a real file but a person, an object, an event, etc., all of which can be annotated by a united specification which is independent of media type or format. MAML allows annotation in semi natural language style so that the user of the annotated data, as well as an annotator, can easily understand the data. A variety of applications will be realized by using MAML.

Key words meta data, annotation, XML

1. はじめに

情報化社会の今、膨大な量のコンテンツの中から目的とするデータを探しだし、能動的にアクセスできる仕組みが必要となってきた。近年、これらマルチメディア情報を効率よく検索したり要約する手段として、メタデータ技術が注目を集めている。これは、コンテンツの内容に関する特徴を予め記述しておき、その記述データを直接の処理対象とすることで処理を代替しようというもので

ある。ここで、記述データはメタデータと呼ばれる。特に注釈の意味でつけられるメタデータをアノテーションデータと呼ぶ。現在、テキスト、画像、音声、動画をはじめとした様々な形態のメディアデータが混在しており、またメディア毎に見ても、非常に多くのファイルフォーマットが存在する。これらを一括に利用・処理するには、メディアデータの種類に依存しない、統一的なアノテーション仕様が必須である。また多量のアノテーションデータが生成されていくためには、データ構造の理解が一般

の人々でも容易な記述仕様でなければならない。この理念に基づいたアノテーション記述仕様を策定することには大きな意義があると考えられる。

2. 背景, 目的

本章では、メディアデータのための汎用アノテーション記述言語 MAML (Multimedia Annotation Markup Language) の提案に至る背景と目的について概説する。

近年新しく規格化される仕様の中には、アノテーション(メタ)データを意識したものがいくつか見受けられる。動画などの非テキストデータのコンテンツ記述を目的とした MPEG-7 [1] はその代表例である。MPEG-7 は、非テキストファイルに対するメタデータ付与が前提のため、例えば映像ファイルとテキストファイルへのアノテーションを同じ仕様で且つ同一ファイル上に記述することはできない。

他には、Web 上に点在する様々なデータに統一の意味情報を付与するのは不可能であるという前提から、個々のユーザが独自に定義し記述したメタデータを付与し、人間と機械とのコミュニケーションの実現を目指す Semantic Web [2] が挙げられる。Semantic Web では、RDF [3] によってリソースのプロパティやリソース間の関係を記述するので、スキーマを別途定義しなければならない。さらに自由にスキーマを設定できるということは、処理段階においてオントロジーの整備も十分になされていないとすればならず、それは現状ではまだ難しい。

ところで近年、URI によって書籍の ISBN コードなど、インターネット上に存在しないリソースを指し示すことが可能になり、また RFID の登場により人やオブジェクトを一意に認識し、またさらにその行動や関係までコンピュータで容易に追跡、管理出来るようになってきた [4]。よって、メタデータを付与する対象は実ファイルに限らず、想定しなければならない領域は広がるばかりである。実世界の人物に関するメタデータを記述し、そのつながりを公開、共有するためのプロジェクトはその一例である [5]。このような情勢の中、様々な記述能力を有するメタデータ規格が次々と発表されているが、それらの中から目的や用途に応じて選択し、その仕様を理解して初めて生成のための作業が可能であり、その初期コストがアノテータにとって大変な負担となっている。また、アノテーションデータは、人間が直接活用することを主目的とするもので、機械による高度な処理を視野に入れたデータ構造にすることは絶対条件ではない。機械的にそれが生成されるケースもあるが、主として人間による付与作業が基本であるので、人間が記述及び理解しやすい表現

構造を念頭におくのが望ましい。一般的に、機械が解析しやすいメタデータは可読性に乏しく、概して記述も難しい。その逆も然りで、相反するものである。よって、タグなどによるデータの構造化や情報の意味記述は最小限にし、自然文章中心の構造に主眼を置くのが妥当であろう。この理念に基づいて設計された記述フォーマットに Dublin Core [6] が挙げられる。Dublin Core では任意のファイルフォーマットに付与可能な情報として、基本 15 項目を設定している。しかしながら、項目は“作成者”、“日付”、“資源タイプ”をはじめとしたファイルの属性に関する記述と内容記述に限定されている。

MAML は、上記の点を考慮した汎用アノテーション記述言語である。MAML は、“如何なるデータ、事象、物体、人物や行為に対して同一の仕様でアノテーションできる”ことを目標に掲げている。

3. MAML の概要

現在世の中に流通しているファイルは、通常データ形式によって分類される。例えば画像、映像、音声、映像と音声を合わせたメディアである動画、テキスト、HTML、XML などが挙げられる。MAML ではデータ形式やファイルフォーマットによるファイルの区分は行わず、すべてのファイルは音声情報 (audio)、映像情報 (visual)、内容情報 (contents) を有するメディアとみなす。ここでいう音声情報とは人の聴覚から得られる情報、映像情報とは人の視覚から得られる情報、内容情報とはコンテンツやテキストデータに対して人の知識、主観、推論によって導かれる情報をいう。MAML では、メディアの分類は、この 3 情報の中でどの情報をアノテーション可能であるかにより決定する。記述可能な情報と、定義されるメディアの種類を対照を表 1 に示す。例えば、音声 (audio) メディアは、音声情報と内容情報を、画像 (image) メディアであればメディア時間情報を有しない映像情報と内容情報をアノテーション可能なファイルと定義する。ここでメディア時間とは、メディアの再生経過時間のことをいう。この規則によりすべてのファイルは、表 1 に示した 6 種類のメディア種別のいずれかに一意に分類されるのは明らかである。また、人物、書籍や出来事などのオブジェクトやイベントも仮想ファイルと考え、アノテーションの対象とする。それぞれのメディアファイルが付与可能な情報も表 1 に併記してある。どの情報をアノテーション付与対象とみなすかで、同一のファイルでもメディアの種類が異なることもありうる。例えば書籍を考えた場合、書かれている文章だけに注目すればテキスト (text) メディアとなり、形状や外観にまで対象に含めれば画像

表 1 メディアが有する情報と名称の定義

メディア時間	包含する情報			メディアの種類	記述可能な情報			実ファイルの例	仮想ファイルの例
	音声	映像	テキスト		音声情報	映像情報	内容情報		
	x	x	/x	movie	x			動画ファイル	ビデオテープ中の動画
			/x	video				映像ファイル	ビデオテープ中の映像
			/x	audio		x		音声ファイル	カセットテープ中の音声、人物の声
x	x	x	/x	image	x			画像ファイル	写真、カセットテープの形状、人物の外見
x	x			text	x	x		テキストファイル	文書、書籍
x	x	x	x	null	x	x		圧縮ファイル	人物の性格、イベント

(image) メディアとなる。また、時間に関する情報はアノテーションデータを処理する上で重要な役割を果たす。MAML ではメディア時間に加え、コンテンツ時間、アノテーション時間の計 3 種類の時間情報を定義し、これを記述可能とする。コンテンツ時間情報とは、アノテーションの対象が実世界上の時刻で何時のことを表現しているかを記述するものである。例えば、複数ファイルから要約生成を行う場合、その時間的順序を加味する必要性があり、コンテンツ時間を基に整合性を保持する。一方、アノテーション時間情報とは、アノテーションの付与が実世界上の時刻で何時になされたかを記述したものである。あるファイルに対して、多人数が長期にわたって感想や意見などのアノテーションデータを追記していくようなケースも MAML は想定しており、個々の情報が生成された時間順序がデータを解析する上で非常に重要になるからである。メディアに付加されるアノテーションデータは基本単位（エレメントと呼ぶ）の集合体として構成する。以上の概念に基づいて、複数のメディア情報を一元的に記述・処理可能な環境の構築を目的とする。

4. MAML の記述仕様

MAML は XML [7] で表現する。XML は、人間が読むことができるシンプルなタグを使って、データをマッピングするために使われる汎用的なシンタックスを定義した規格である。最上位に maml タグ、その下層に media タグを、さらにその下層にアノテーションの基本単位となる element タグを列挙していく。図 1 に MAML の記述例を示す。

図 1 は動画ファイルと暑中見舞いはがきを対象メディアとしている。この例では、動画ファイルのメディア時刻 32.56 秒から 38.11 秒にかけて、赤い帽子をかぶった女性が映っており、この女性が“あけましておめでとうございます”と発話している。また、このシーンに対して k_ito@nak.ics.keio.ac.jp がデータを付与している。さらに、そのアノテーションデータに対して、kato@nak.ics.keio.ac.jp が書いた暑中見舞いはがきの文面を引用する形でアノテートされている。

4.1 メディア情報の記述

メディアに関する情報は、media タグの属性として記

表 2 media タグの属性

属性名	説明
type	メディアの種類
maml-location	MAML ファイルの存在場所を URI で記述
media-location	対象ファイルの存在場所を URI で記述
media-word	対象ファイルを特定する単語（文節）を記述
creator	対象ファイルの作成者を記述
annotator	アノテータを記述
contents-basetime	コンテンツ時間の基準値
annotation-basetime	アノテーション時間の基準値
duration	メディア時間長

表 3 element タグの属性

属性名	説明
id	エレメントを特定するための識別子
begin	メディア時間におけるエレメントが対象とする開始時刻
end	メディア時間におけるエレメントが対象とする終了時刻
contents	コンテンツが表現している実世界上の時刻
annotation	エレメントが生成された実世界上の時刻
target	アノテーション対象の指定

述する。表 2 に属性の種類を示す。type 属性値にはメディアの種類、maml-location 属性値には、MAML ファイルの URI を記述する。対象の特定方法としては、URI で指定できるものについては media-location 属性値でそれを記述する。対象を特定づける単語あるいは文章の形式で media-word 属性値として表わすことも可能である。コンテンツ時間の基準値を contents-basetime 属性値に、アノテーション時刻の基準値を annotation-basetime 属性値として記述できる。個別のエレメントがここで指定した時刻と異なる場合は、次節で解説するエレメント記述方式の方でさらに時間指定をする。また、対象メディアがメディア時間を有する場合は duration 属性にその全長も併記しておく。

また media タグを列挙することで、複数のメディアのアノテーション情報を同一 MAML ファイル内に記述できる。

4.2 エレメントの記述

エレメントは、最上層に element タグを記述する。表 3 に属性の種類を示す。各エレメントを識別するための id 属性、さらに対象メディアがメディア時間を有する場合、エレメントの対象開始時刻を begin 属性で、対象終了時刻を end 属性で記述する。また、target 属性値によって図 2 に示す 4 通りのアノテーションを定義できる。第一に、target 属性がない場合は、media-location 又は media-word で指定されるメディアに対する直接的なアノテーションとみなされる（図 2 の (1)）。第二に、target 属性値が id 番号のみの場合は、同一 MAML ファイル中

```

<?xml version="1.0" encoding="UTF-8"?>
<maml>
  <media type="movie"
    maml-location="http://www.xxx.jp/a.maml"
    media-location="http://www.xxx.jp/a.mpg"
    contents-basetime="2003-05-07"
    duration="01:42.56">
    <element id="1">
      :
    </element>
    :
    <element id="10" begin="32.56" end="38.11"
      contents="2003-01-01/2003-01-02">
      <audio>
        <utterance id="p1">
          あけましておめでとうございます。
        </utterance>
      </audio>
      <visual>
        <character id="p1">
          赤い帽子をかぶった女性。
        </character>
      </visual>
    </element>
    <element id="11" target="10"
      annotation="2003-05-06T05:30"
      annotator="k_ito@nak.ics.keio.ac.jp">
      <contents>
        <supplementation>
          このシーンに映っている女優と結婚しました。
        </supplementation>
      </contents>
    </element>
  </media>
  <media type="text"
    maml-location="http://www.xxx.jp/a.maml"
    media-word="加藤君から来た暑中見舞いはがき"
    creator="kato@nak.ics.keio.ac.jp"
    annotation-basetime="2003-07-18">
    <element id="12" target="11">
      <contents>
        <copy>
          先日、MAML Editor/Viewer で例の映画見ました。
          ご結婚されたんですね。おめでとうございます。
        </copy>
      </contents>
    </element>
  </media>
  :
</maml>

```

図 1 MAML の記述例
Fig. 1 An example of MAML.

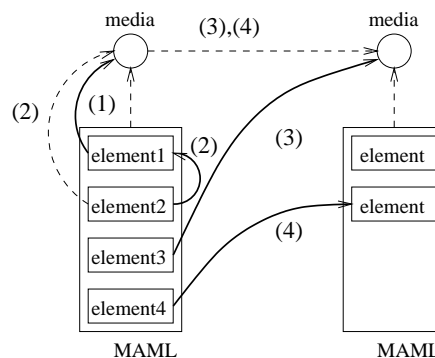


図 2 target 属性によるエレメントの意味付け

の特定のエレメントに対するアノテーションとみなされる。つまり、対象メディアへの間接的なアノテーションを定義していることになる(図2の(2))。第三に、target 属性値が他のメディアの URI の場合は、対象メディアに関する何らかの記述が、同時に target 属性値で指定されるメディアへのアノテーションになっていることを示す。これはつまり、メディアとメディア間の直接的な関係をアノテーションデータで定義することになる(図2の(3))。最後に、target 属性値が URI+"#" +id 番号の書式になっている場合は、アノテーションの対象が、他のメディアを対象とする MAML ファイルの特定のエレメントであることを意味する。言い換えれば、対象メディアに関する何らかの記述が、target 属性値で指定される MAML ファイルの対象メディアに関する記述へのアノテーションとして定義される。これはつまり、メディアとメディア間の間接的な関係をアノテーションによって定義することに相当する(図2の(4))。

4.3 エレメント内容の記述

エレメントタグの下層に音声、映像、内容情報を区別するための種別タグ、さらにその下層にそれぞれの種別毎にアノテーションの内容をさらに分類するためのクラスタグを記述する。クラスの種類を表4に示す。クラスに基づく具体的な内容は、クラスタグの下層に自然言語で記述する(以下アノテーションテキストと呼ぶ)。日常で使用している自然言語で主観に基づいて記述するので、当然専門的知識は全く必要としない。また表現の自由度が飛躍的に高まる。単に時間情報やクラス分類の情報だけを設定したいのであれば、アノテーションテキストは空白でもよい。また、ひとつのエレメント内に複数のクラスの情報をまとめて記述することも可能である。

4.3.1 音声情報の記述

音声情報は、人が発生した言葉をそのまま記述する“発話”クラス、その他の音情報で言葉に直接変換できる場

表 4 情報の種別

種別	クラス
音声情報 (audio)	発話 (utterance)
	擬音 (imitation sound)
	音楽 (music)
映像情報 (visual)	物体 (object)
	人物 (character)
	文字 (letter)
	背景 (background)
内容情報 (contents)	場所 (place)
	標題 (subject)
	説明 (explanation)
	属性 (attribute)
	補足 (supplementation)
	要約 (summary)
	複製 (copy)

合、それを記述する“擬音”クラス、言葉に直接変換出来ない音情報はすべて“音楽”クラスとしてその音情報に関する文章表現を記述する。

4.3.2 映像情報の記述

映像情報は、移動可能なオブジェクトに関しては人とそれ以外に分類し、人に関する情報を“人物”クラスで、それ以外を“物体”クラスで記述する。移動不可能なオブジェクトまたはその集合を“背景”クラスで定義する。また映像中の文字情報をテキストで抽出する“文字”クラス、背景の地理情報を記述するための“場所”クラスを設定している。

4.3.3 内容情報の記述

内容情報は Dublin Core の基本 15 項目にほぼ準拠する形となっている。コンテンツのタイトルを記述する“標題”クラス (Dublin Core の“Subject”に相当)、コンテンツの内容を自由に記述する“説明”クラス (Dublin Core の“Description”に相当)、対象ファイルの属性に関する情報を記述する“属性”クラスがある。テキストメディアに含まれるテキストの一部を MAML ファイルに転記する場合、それを“複製”クラスのエレメントで表現する。他のクラスのエレメント内で記述されたテキストの要約表現を“要約”クラスで表現する。例えば、クローズドキャプションは発話文を要約して表示しているが、そのような用途を想定している。対象メディアに関する直接の要約表現は、先に解説した“説明”クラスで表現する。また、アノテーションデータは、コンテンツの製作段階で付与される映像・音声転記や、流通の段階で必要な属性の情報に限らず、ユーザがコンテンツを活用する段階で付与されるデータも含まれる。この段階のデータというのは、個々のユーザがデータを管理する上で必要な補足情報や、一時的なコメントなどが挙げられる。それらを“補足”クラスで表現する。

ちなみに Dublin Core の一部は内容情報のクラス以外でも定義できる。“Date”は element タグの annotation 属

性で、“Coverage”は place クラスあるいは element タグの contents 属性で、“Identifier”は media タグの media-location か media-word 属性で、“Creator”は media タグの creator 属性で、記述可能である。このように、MAML では異なる情報の分類形態を用いているが、Dublin Core で定義される 15 項目すべてを記述可能である。

5. MAML の特徴

MAML は、一般の人々でも容易にアノテーションできることを主目的としているので、クラスの名称は、名前からその意味を直感的に把握でき、またクラスの分類数は最小限にとどめている。前章の説明や図 1 の例から、MAML のクラス体系がいかに単純かが理解可能であろう。また、各クラスタグが属性を持つことも極力避けた。現段階では、対象を特定したい場合の記述に id 属性を持つことだけが許容される。分類規則を多層化・複雑化すればするほど、その決定に専門的知識と時間を要し、さらに選定ゆれも発生するからである。タグセットを限定することで情報の詳細な区分が設定できない場合でも、自然文章中に何に関する情報であるかを含めることで、情報の種別を細分化できると考えるからである。ところで、アノテーションとはメディアの内容に関する情報だけではなく、それに含まれる映像や音声の転記情報についても記述対象とするものであり、MAML もそれを踏襲している。ここで注意しなければならないのは、自然文章中で種別情報まですべて記述しようとすると、逆に何についてのアノテーションが限定出来なくなる場合が生じてしまう。例えば、“おはようございますという会話がある。”とアノテートした場合、<utterance> おはようございます。</utterance>, <character> おはようございますという会話がある。</character>, <explanation> おはようございますという会話がある。</explanation> のいずれにも解釈できてしまう。これを解決するためにはアノテーションテキストの表現構造に制約を与えなければならなくなる。そのルールの定義は非常に困難であり、さらにアノテータへそのルールを強いることによる負担は非常に大きい。MAML で定義されているクラスはこれを区別するために、タグとして明記する最低限のセットである。

また、MAML では画像の一部領域やテキストの一部をタグや属性値によって指定することは想定していない。明示したければ、アノテーションテキスト中に自然言語で対象領域を表記するか、“複製”クラスを用いることとなる。MAML で時間領域の情報について定義し、空間領域の情報を定義しないのは、アノテーションデータと対

象メディアを同時にユーザに対して提示することを考えた場合、ユーザによってその情報の対象範囲を検出するための時間的コストに大きな差異があるからである。ある女性が映像に映っていたとして、メディア時間の指定がされていなければ、ユーザは全ビデオ区間を順再生して確認しなければならない。一方、女性が映っている画像が提示されたとして、その女性がどの画像領域のどの範囲に映っているかは、アノテーションテキストにその女性の特徴が書かれていれば、瞬時に把握できる訳である。文書に関しても、同様である。

実際のところ、どの書式を採用するのが望ましいかは、主として人間が読むのか、機械処理も念頭に置くのかによって異なり、MAML ではその両者の表現記述を可能とする。時間情報に関して例を挙げれば、前者だけを考えるのであればアノテーションテキスト中に“平成 15 年 3 月 26 日”などを書いておけばいい。機械処理を前提とした記述方式にするならば、begin 属性値、end 属性値、contents 属性値等で記述すればよい。具体的には、国際標準 ISO8601 に基づく方式で記述する。

MAML の最大の利点は、アノテータは対象とするメディアの種類や目的に応じて異なる仕様で記述しなければならないという煩わしさから解放されることにある。アノテーション作業において、専門的知識を全く必要としない記述体系も大きな特徴である。さらに、応用システムを構築する上で MAML を採用することにより、メディアに依存することなく画一的に処理を行うことが可能になる点も見逃せない。

MAML データから TEI (Text Encoding Initiative) や DocBook などの文書構造記述や XHTML, XSL-FO, RTF, TEX に代表される文書整形記述, MPEG-7 MDS, Dublin Core, RSS, FOAF, RDF など他 XML 文書への変換を要する場合、XSLT によるタグ変換をベースとするが、自然言語中心の構造ゆえに、アノテーションテキストに対する自然言語解析とオントロジーの知識を含めた変換処理までも想定しなければならず、今後の検討課題である。

6. アノテーションデータの構造化

アノテーションデータの処理プロセスを考える上で、どのようにデータが構造化されているかという点が重要になってくる。MAML ファイルに含まれる要素は、メディアに対する直接のアノテーションと、他の要素に対するアノテーションとに大別される。ある要素を子とし、そのアノテーションの対象を親とする関係を定義できる。あるいは、要素間に時間

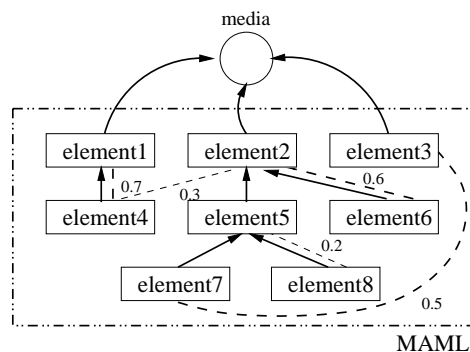


図 3 エレメント間の関連性の定義

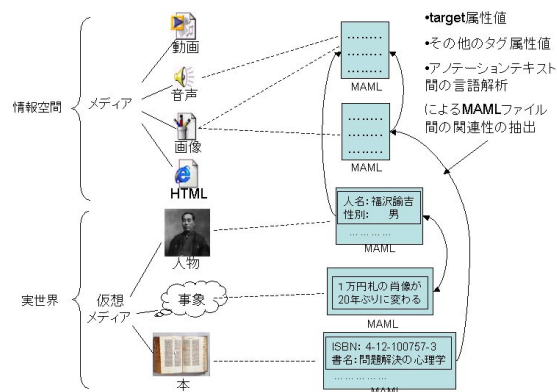


図 4 MAML によって構築されるメディアデータ間の関連性

的な包括関係が成り立つ場合が存在し、それも親子関係と見なす。すると、要素全体はメディアをルートとするツリー構造とみなすことができる(図3の矢印)。また、各要素のアノテーションテキスト間での単語の共起関係等によって関連性を定義することも出来る(図3の点線)。タグ構造に基づく記号処理的なアプローチだけではなく、アノテーションテキストを対象にした自然言語解析と併用することにより、異なった尺度で密結合された非常に有用なアノテーションデータ群が定義されることになる。何も関連付けは MAML ファイル内部だけに限られず、メディアファイル間の関連性も定義されていくことを意味する(図4参照)。実際に我々は、この概念を基に、MAML データを用いたメディア検索システムおよび要約システムを構築している[8][9]。システムでは、条件を基に図3のグラフ構造から、個々の要素のスコアを算出し、検索結果の提示や要約表現の生成を行っている。

7. まとめ

本論文では、メディアの種類やフォーマットに依存しない統一的なアノテーション記述仕様 MAML について

解説した。メタデータ仕様に関しては、以下に述べる3点が重要である。

(1) 一般の人々がデータを生成可能であること。つまり生成過程が容易であること。

(2) データが広く流通し、応用アプリケーションが数多く生成されること。

(3) その結果、ユーザがアノテーションしたことによる恩恵が多分に得られる環境が実現されること。

この3要素はスパイラルの関係にあり、すべてを満たさなければ、その仕様は受け入れられないであろう。

MAMLは様々な適用方式が想定される。今後はMAMLによる高度メディア利用環境の実現を目指して、仕様の再検討を行っていく予定である。

謝 辞

本成果は、平成14年度IPA未踏ソフトウェア創造事業「未踏ユース」及び平成15年度IPA未踏ソフトウェア創造事業の一部である。IPA（情報処理振興事業協会）、プロジェクトマネージャーの電気通信大学 竹内郁雄教授及び京都大学 田中克己教授に深く感謝いたします。

文 献

- [1] The MPEG Homepage,
<http://mpeg.telecomitalia.com/>
- [2] Berners-Lee, James Hendler and Ora Lassila: The Semantic Web, Scientific American May 2001, (2001).
- [3] W3C RDF Homepage,
<http://www.w3.org/RDF/>
- [4] 椎尾一朗, 早坂達: モノに情報を貼り付ける: RFID タグとその応用 情報処理学会誌, Vol. 40, No. 8, pp. 846-850, (1999)
- [5] The 'friend of a friend' project: FOAF Homepage,
<http://rdfweb.org/foaf/>
- [6] Dublin Core Metadata Initiative(DCMI) Homepage,
<http://dublincore.org/>
- [7] Extensible Markup Language(XML) Homepage,
<http://www.w3.org/XML/> <http://www.google.co.jp/imghp/>
<http://web.resource.org/rss/1.0/spec>
- [8] 伊藤一成, 斎藤博昭: メタデータ解析に基づくメディア検索システム, 情報処理学会研究報告, DBS131-69, pp. 515-520, (2003).
- [9] 伊藤一成, 酒井康旭, 斎藤博昭: メタデータ解析と自然言語処理を併用した要約動画の生成, 情報処理学会研究報告, DBS132-6 (2004).