

1. 以下の問いに答えよ。

(ア) 現在の WWW では HTML がページ記述言語として用いられている。このことの利点を述べよ。

「ブラウザ・オペレーティングシステム・情報発信者・情報獲得者」

HTML は人々が情報を発信・検索する上での環境に依存しない、つまりシステムやデバイスなどに依存しない情報の表現方法を定義している。よってどんな環境でも閲覧できるウェブサイトを作ることができる。現在では様々な OS が存在するが、OS にかかわらず HTML ファイルを開くことができる。また、ブラウザは HTML 文書を解釈して実装するが、ブラウザの種類によっては多少異なって実装されてしまう。しかし、HTML 文書は文字のサイズや位置などを細かく定義できるので、特定のブラウザに依存することなく、ほとんどのブラウザでほぼ同じ出力結果を得ることができる。(サイズやフォントを指定していない文字などについては、ブラウザによって表示の仕方が異なってくる) このことで情報発信者が意図した実装と情報獲得者が得るページの実装とで、大きな差異はなく誤解が生じることなどほとんどない。よって、国際的に効率的なコミュニケーションを可能にし、技術的な進歩とビジネス上の協業を支えるなど大きな利点がある。また、HTML 言語は情報を表示することを目的にしているだけなので、タグの種類がはっきりと決まっている。

2. HTML を用いている限りセマンティックウェブを実現することが困難である。その理由を述べよ。

まず、HTML は文章の構造を定義するものであった。つまり、HTML 文書の主な使用法は情報を表現することである。これに対し、セマンティックウェブを実現するということは、文書の含む意味を形式化し、コンピュータがその文書の内容を理解する。それにより、コンピュータによる自動的な情報の収集や分析へのアプローチが可能になることを期待される。つまり、情報の意味を記述したタグが存在して、そのタグによって文章を構造化したり文章に意味を付加することで、コンピュータに文章を理解させる。しかし、HTML 文書を使用するということは、文章に対して内容だけでなくセマンティックウェブを実現するのに不必要な書式まで設定しなければならず、セマンティックウェブを実現するのが困難になってしまう。HTML は文書の各部分とそれらの相互関係についての情報を含んでいない。

また、HTML は要素をただの文字列としてしか理解していない。そのために、例えば同じ意味を持つ異なる表記の単語があっても、これらの単語に対して全く関係性を見出すことができない。逆に同じ表記の単語でも異なる意味をもつものが存在するが、HTML 文書で

はこの単語を同じものとしてしか理解することができない。

3. 問2の理由をふまえ、セマンティックウェブを実現するうえでXMLが必要とされる理由を述べよ。

セマンティックウェブを実現するためには、問2で述べたように文章の内容をコンピュータに理解させなければいけない。そして、コンピュータに文章の意味を理解させるためのタグが必要であった。それがXMLである。XML文書には意味タグと構造がある。

まず、XMLはHTMLとは異なり、内容と書式とを分離して考えることができる。XMLのタグは情報の意味を記述していて、文章の各部分の情報が説明している。このことで例えば、HTMLと違って、異なる文字列で表現されている同じ情報も、はっきりと同じものだと識別が出来る。このようにXMLで情報を定義することで、機械にも利用できる情報の表現を可能にしている。

XMLは固定した種類のタグをもたないが、ユーザーが自分のためのタグを定義することができる。

4. セマンティックウェブを実現するうえでDTDやXMLスキーマが必要となる理由を述べよ。

DTDやXMLスキーマを用いることで、各部分の情報の関係などが入れ子構造によって定義される。つまり、物事のルールが定義される。よってXML文書によって定義された情報がルール通りの正しいものか・意図どおりに記述されているかなどこれらをもって判断することができる。また、これらを用いることでコンピュータに内容を理解させるのが容易になる。

異なる組織間での共通のデータ交換に大変役に立つ。つまり、情報のやりとりを円滑に行うことができる。ただし、他者とデータを交換する際に、異なるタグで定義されたXML言語の情報を交換しても理解はできない。よって、DTDやXMLスキーマを存在することでお互いの情報交換が可能になる。