
XML과 Voice XML의 효율적인 통합 방안 제시

장민석 · 홍용택

군산대학교

A Suggestion of Efficient Method for Integrating XML and Voice XML

Min-seok Jang · Yung-taek Hong

Kunsan National University

E-mail : msjang@cs.kunsan.ac.kr

요 약

본 논문에서는 현재 표준화가 진행중인 XML 관련 문서를 음성 통신서비스를 위한 Voice XML로 변환 및 통합할 수 있는 효율적인 방법을 제시하고자 한다. 가까운 미래의 웹 환경은 HTML 방식에서 XML 방식으로 변환 또는 통합될 것이 거의 확실시되어지고 있을 뿐만 아니라 이러한 환경에서 더욱 다양한 단말기로 자료를 접할 수 있는 방법이 요구되고 있다. 그 중에서 현재 최고의 대안으로 대두되고 있는 Voice XML을 이용함으로써 자료의 접근 방식을 저비용으로 웹에서 유·무선 단말기로 전환하고자 한다. 따라서 표준 XML을 이용하여 현재 구성되어 있는 가상의 시스템을 설정하고 이를 효과적으로 Voice XML로 변환 또는 통합하는 방안을 제시하고자 한다. XML은 특성상 확장성을 염두해 두고 만들어진 언어인 만큼 이를 이용해 효과적인 Voice XML 문서로의 변환 또는 통합하는 효율적인 방법이 제시된다면 많은 파장을 불러일으킬 것으로 확신한다.

ABSTRACT

In this paper we suggest a method for translating (or integrating) XML documents into Voice XML documents in order to provide voice communication service. In the forthcoming web environment, XML will certainly overwhelms HTML. At this situation, the method of accessing data through the more various types of terminal machines is required. The best way is to use Voice XML by which the data accessing method is able to change from Web to the wired or/and wireless terminal at low costs. Thus we suggest a method for integrating the XML-based system into the Voice XML-based one.

1. 서 론

현재 웹 환경은 HTML의 한계에 다달았다. 데이터와 스타일시트의 복합적인 구성으로 일관화 되어있는 현재의 웹 환경은 효율적인 정보처리를 힘들게 하고 있다. 이러한 단점을 보완하고자 문서의 구조적 정보 및 내용을 정의하는 마크업 언어인 SGML(Standard Generalized Markup Language)과 이의 장점을 그대로 수용하면서 보다 간소화한 XM

L(eXtensible Markup Language)이 등장하게 되었다. SGML 및 XML은 하드웨어 및 소프트웨어의 종류에 관계없이 이기종간에서도 장치에 독립적으로 문서를 교환할 수 있다. XML은 현재까지는 SGML의 대표적인 단점들을 대부분 보완하였고, 특별한 단점이 발견되지 않고 있어 전자문서를 사용하는 대부분의 응용분야에서 XML을 전자문서의 표준으로 적용하려는 움직임이 활발하다. 또

한 XML의 유연한 확장성으로 인해 채널정의포맷(CDF; Channel Definition Format), 화학마크업언어(CML; Chemical Markup Language), OFX(Open Financial eXchange) 등의 많은 파생언어가 생성되어지고 있다. Voice XML 또한 그러한 맥락에서 생성된 언어이다[1].

이러한 Voice XML은 음성 인식 기술이 실용화 수준으로 발전함에 따라 음성 포털을 비롯한 각종 음성인식 분야에 접목하고자 하는 방법이 연구되어지고 있는 실정이다[2]. 하지만 이러한 XML에서 파생된 언어들 은 각기 자신들의 영역에 특화되어 서로간의 호환성이 결여되어 있는 상황이다. 현재의 웹 환경은 더욱 다양한 단말기로의 접근을 시도하고 있다. 따라서 본 논문에서는 그 중에서 가장 보급률이 높은 유무선 단말기(전화기) 등을 통해 데이터를 접근 가능하게 할 수 있도록 표준 XML 문서를 Voice XML 문서로 변환 또는 통합하는 효율적인 방법을 제시하고자 한다.

II. XML, Voice XML 기술 및 표준화 현황

2.1 XML

XML은 1996년 W3C(WWW Consortium)에서 제안한 것으로서 기존의 HTML과 SGML이 갖는 단점을 보완하여 웹 상에서 구조화된 문서를 전송 가능토록 설계된 마크업 언어이다. XML은 구조화된 문서를 정의하여 태그를 자유롭게 정의 할 수 있는 SGML의 장점과 인터넷상에서 쉽게 하이퍼미디어 문서를 제공할 수 있는 HTML의 장점을 모두 가질 수 있도록 제안한 웹 표준문서 포맷이다. XML은 기술적으로 다음과 같은 특성을 가진다.

-인터넷 환경을 고려하여 만들어 졌기 때문에 XML에는 인터넷상의 정보 전달을 위한 매커니즘이 내재되어 있다.

-일반사용자들이 DTD(Document Type Definition)에 대한 지식이 없이도 유연하게 만들 수 있는 적절한(Well-formed) 문서를 지원한다.

-XML문서의 스타일을 표현하기 위한 스타일 시트(XSL; XML Stylesheet Language)를 사용한다.

-HTML에 비해 강력한 링크 및 포인터를 제공한다.

-메타 데이터 관리가 가능하다.

-문서와 DTD에 대한 표준 인터페이스를 지원 한다.

HTML은 웹 상에서 손쉽게 하이퍼미디어 문서를 만들 수 있고 이식성이 뛰어나지만 한정된 구조와 문서만을 만들 수 있으며 문서의 구조 정보 및 임의의 레이아웃을 가지기 어렵다. XML은 이러한 단점들을 보완하여 차세대 웹을 주도할 언어로 각광받고 있다.

현재는 XML/EDI를 주축으로 많은 응용분야가 개발되어지고 일부 사용되어지고 있다[3].

2.2 Voice XML

Voice XML은 AT&T, IBM, 루슨트 테크놀로지, 모토롤라 등 정보통신 분야의 4개 기업체에 의해 설립된 Voice XML 포럼에서 제안한 마크업 언어로서, 음성 애플리케이션 개발을 위해 고안된 XML 문서 형식의 일종이다. Voice XML 포럼은 1999년 8월 Voice XML 0.9버전을 발표 후 작년 3월 이를 크게 보완한 버전 1.0을 정식 제안했다. 세계 인터넷 환경을 주도하고 있는 W3C 컨소시엄에서는 Voice XML 포럼의 제안을 받아들여 작년 5월 22일 Voice XML을 웹의 대화형 마크업 언어 표준으로 공인했다[4].

표 1은 현재 전화기를 통해 서비스를 실시하고 있는 다른 언어들과 Voice XML의 특성을 비교한 것이다. 표에서 보는바와 같이 현재 타 언어에 비해서 많은 기술을 지원하고

있어 앞으로 표준만 확립된다면 폭발적인 수요가 이루어 질 것이라고 생각되어진다[5]. Voice XML은 XML에서 파생된 다른 언어들과 마찬가지로 음성을 위한 태그들이 존재한다. 이러한 태그들은 현재 W3C의 Voice XML 스펙에 자세히 기술되어 있다[1].

표 1 Voice XML의 특성비교

Technology	WTE	WAP	Voice XML
Voice-enable existing Web pages	Maybe		Maybe
Re-engineer entire sites		Partial	○
Primary aural interaction possible	○		○
Support simple telephony features	○	○	○
Support advanced telephony features		Partial	○
Support wireless devices		○	
Support plain old telephones	○		○
Hands-free operation			○

하지만 개발업체 마다 자신들 고유의 태그들이 난무하고 있는 상황이어서 아직까지는 공통된 방법으로 Voice XML을 개발하기 어려운 실정이다. 본 논문에서는 Motorola사의 VoxML을 이용하여 XML과 Voice XML의 통합방안을 모색해보았다.

표 2는 W3C에서 제시하는 Voice XML (ver1.0) 엘리먼트들과 Motorola사에서 제시하는 엘리먼트들의 차이점들을 나타내고 있다[6].

표 2 W3C와 Motorola사의 Voice XML의 차이

W3C Voice XML	Motorola의 VoxML
<assign>, <audio>, <block>, <break>, <catch>, <choice>, <clear>, <disconnect>, <div>, <dtmf>, <else>, <elsief>, <emp>, <enumerate>, <error>, <exit>, <field>, <filled>, <form>, <goto>, <grammar>, <help>, <if>, <initial>, <link>, <menu>, <meta>, <noinput>, <nomatch>, <object>, <option>, <param>, <prompt>, <property>, <pros>, <record>, <reprompt>, <return>, <sayas>, <script>, <subdialog>, <submit>, <throw>, <transfer>, <value>, <var>, <vxml>	<ACK>, <AUDIO>, <BREAK>, <CANCEL>, <CASE>, <CLASS>, <DIALOG>, <EMP>, <ERROR>, <EXIT>, <GOODBYE>, <HANGUP>, <HELP>, <INPUT>, <META>, <OPTION>, <OPTIONS>, <OR>, <PROMPT>, <PROS>, <RENAME>, <REPEAT>, <RESPONSE>, <STEP>, <SWITCH>, <TRANSFER>, <VALUE>

그림 1은 이와 같은 엘리먼트를 이용하여 VoxML 문서를 작성한 간단한 예이다.

VoxML 또한 XML에서 파생된 언어인 만큼 유니코드를 지원하여 원활한 어플리케이션 작성을 돕는다. 예와 표에서 보는바와 같이 Motorola사의 VoxML 1.2버전의 엘리먼트들

은 전화기에 최적화 되어있어 W3C에서 제시하는 엘리먼트들과 많은 차이점을 나타내고 있다.

```
<?xml version="1.0" encoding="euc-kr"?>
<DIALOG>
<STEP NAME="init">
<PROMPT>음료수를          콜라주세요
</PROMPT>
<HELP>콜라 사이다가 있습니다</HELP>
<INPUT          TYPE="OPTIONLIST"
NAME="drink" NEXT="#confirm">
<OPTION          VALUE="콜라">콜라
</OPTION>
<OPTION          VALUE="사이다">사이다
</OPTION>
</INPUT>
</STEP>
<STEP NAME="confirm">
<PROMPT>당신은<VALUE NAME
="drink"/>를주문하셨습니다.</PROMPT>
</STEP>
</DIALOG>
```

그림 1. VoXML 문서 예

그림 2는 일반적인 Voice XML을 이용한 서비스 방법을 나타내고 있다[1].

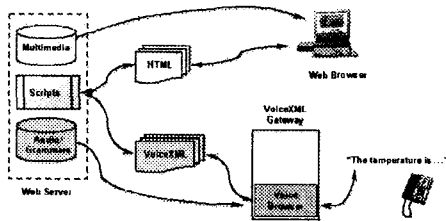


그림 2. Voice XML을 이용한 서비스 방법

이러한 기술들은 음성 인식 기술이 실용화 수준으로 발전함에 따라 음성 포털을 비롯한 각종 음성인식 분야가 새로운 이슈로 떠오르

고 있다. 차세대 꿈의 통신을 실현시킬 것이라고 하여 현재 업계에서 관심이 집중되고 있는 IMT-2000 사업의 준비 과정에서도 음성인식 관련 기술은 상당히 중요한 주제로 자리를 차지하고 있다. 이 같은 분위기 속에서 최근 Voice XML 문서 형식이 음성 합성과 음성 인식을 결합한 음성 서비스를 실현하기 위한 가장 유력한 방안으로서 기대를 모으고 있다.

III. XML과 Voice XML의 통합방안

현재 구축되어 있는 시스템들은 모바일 환경과 XML 기술의 표준화 및 실용화에 따라 현재의 시스템을 그대로 또는 적은 비용으로 접근 환경의 변환을 피하고자 노력하고 있다. 이러한 시점에서 본 논문이 제시하는 XML과 Voice XML의 통합방안에서는 현재 안고 있는 문제점을 파악하고 효과적인 통합 방안을 제시하고자 한다.

본 논문에서는 각각 XML 1.0과 Voice XML 1.0을 기준으로 가상의 시나리오를 바탕으로 통합방안을 제시한다. 그림 3은 현재 XML로 구축되어있는 웹사이트를 Voice XML로 통합하기 위한 개략적인 구성이다. 이와 같이 통합을 하기 위해서는 우선 XML 문서가 필요하다. 이문서는 효율적인 XML 문서관리를 위해서 데이터베이스에 저장되어 사용된다. 현재의 XML 문서 저장기술들은 각 벤더들에 의해 원하는 자료를 자동으로 XML 문서로 전환되어지고 또한 이를 효과적으로 저장하고 있다. 이러한 방법을 이용하면 기존의 데이터 또한 본 논문에서 제시하는 모듈을 이용한다면 표준만 마련이 된다면 손쉽게 Voice XML 문서를 제작할 수 있으리라 생각된다. 이 변환모듈에서는 저장된 XML 문서를 변환모듈을 통해서 Voice XML GateWay의 Voice XML Browser를 통해서 일반 전화기와의 접속을 수행한다. 이 변환 모듈의 기능을 요약하면 다음과 같다. 변환 모듈은 저장되어져 있는 표준

XML 문서를 Voice XML 문서로 필요시 자동 변환되어지게 하고 또한 이 자체를 파싱하여 사용 가능한 문서인지를 판단하며 필요시 사용자가 임의의 변경을 할 수 있도록 구성하였다. 또한 만들어진 Voice XML은 DB에 저장되어져 계속적인 문서 요구시 빠른 제공을 할 수 있다. 추가적으로 이 변환 모듈을 하나의 독립된 서버로 구축한다면, 유동적으로 변하는 인터넷 환경에서 Voice XML 및 XML 분산환경에 적합한 모델이 될 것으로 생각되어진다.

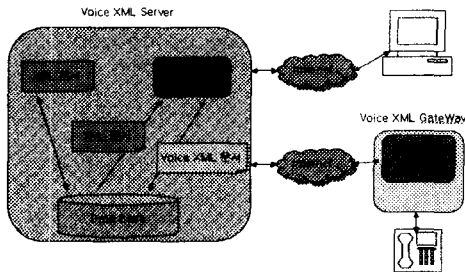


그림 3. Voice XML로의 통합 구성도

하지만 현재 Voice XML은 XML과 마찬가지로 아직까지 완전한 표준이 마련되어 있지 않았으며, 각 벤더들마다 자신들의 사업영역에 맞추어서 Voice XML을 바라보고 있기 때문에 원활하게 사용되기 위해서는 앞으로 많은 시간이 흘러야 할 것이다. 또한 현재의 추세가 단순 텍스트와 음성에서 멀티미디어로 전환되어지고 있어 Voice XML은 이러한 미디어를 수용하는 방법 또한 모색되어야 할 것이며 만일 일반유선 전화기 전용으로 발전한다면 기타 미디어와의 통합 방안 또한 연구되어야 할 것이며, 아직까지 음성인식율이 높지 않기 때문에 효과적인 Voice XML 문서 처리를 위해서는 음성인식분야의 발전이 선행되어야 할 것이다[2].

IV. 활용에 대한 건의

본 연구에서는 현재 구축되어 있는 시스템들

이 유무선 단말기 등으로 접근 가능케 하는데 많은 시간과 비용이 소요되는 문제를 Voice XML을 통해서 해결하고자 하였으며, XML 문서를 이용한 시스템에 한해서 Voice XML을 이용한 통합방안을 제시하였다. 따라서 앞으로 개발되어질 거의 모든 XML 시스템을 타 접근 장치를 이용하여 구축하려고 할 때 기초자료로 활용될 수 있으며, 본 연구에서 제시된 XML과 Voice XML의 통합방안을 토대로 하여 기업 또는 정부에서 향후 추진하려는 XML기반의 모든 사업에서 효율적인 정보제공을 위해 참조할 수 있을 것이다.

참고문헌

- [1] 월드와이드 웹 컨소시엄(World Wide Web Consortium, W3C), <http://www.w3c.org>
- [2] 마이크로 소프트웨어, 2000. 9
- [3] 한국전산원, "SGML, XML, EDI 통합 및 연계방안", 1999.6.
- [4] Voice XML 포럼, <http://www.voicexml.org>
- [5] http://webreview.com/2000/09_15/developers/09_15_00_1.shtml
- [6] 모토로라사 홈페이지, <http://www.motorol-a.com>