

The W3C Speech Interface Framework Media Types: application/voicexml+xml, application/ssml+xml, application/srgs, application/srgs+xml, application/ccxml+xml and application/pls+xml
[draft-froumentin-voice-mediatypes-02](#)

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of [Section 3 of RFC 3667](#). By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she become aware will be disclosed, in accordance with [RFC 3668](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on November 2, 2005.

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

Copyright Notice

Copyright (C) The Internet Society (2005).

Internet-Draft

W3C Speech Interface Media Types

May 2005

Abstract

This document defines the media types for the languages of the W3C Speech Interface Framework, as designed by the Voice Browser Working Group in the following specifications: the Voice Extensible Markup Language XML, the Speech Synthesis Markup Language (SSML), The Speech Recognition Grammar Specification (SRGS), Call Control XML (CCXML) and the Pronunciation Lexicon Specification (PLS).

Table of Contents

1.	Introduction	3
2.	Registration of application/voicexml+xml, application/ssml+xml, application/srgs+xml, application/ccxml+xml and application/pls+xml	4
2.1	Encoding considerations	4
2.2	Interoperability considerations	4
2.3	Published specifications	4
2.4	Applications which use these media types	4
2.5	Security Considerations	5
2.6	Additional Information	5
2.6.1	Magic numbers	5
2.6.2	File extensions	5
2.6.3	Fragment identifiers	5
2.6.4	Macintosh File Type Code	5
2.6.5	Person & email address to contact for further information	6
2.6.6	Intended usage	6
2.6.7	Change Controller	6
3.	Registration of application/srgs	7
3.1	Encoding considerations	7
3.2	Interoperability considerations	7
3.3	Published specifications	7
3.4	Applications which use this media types	7
3.5	Security Considerations	7
3.6	Additional Information	8
3.6.1	Magic numbers	8
3.6.2	File extensions	8
3.6.3	Macintosh File Type Code	8
3.6.4	Person & email address to contact for further information	8
3.6.5	Intended usage	8
3.6.6	Change Controller	8

4.	IANA Considerations	9
5.	IPR Disclosure Acknowledgement	10
6.	Normative References	10
	Author's Address	11
	Intellectual Property and Copyright Statements	12

[1.](#) Introduction

This specification defines the media types of the VoiceXML, SSML, SRGS, CCXML and PLS, the specifications of the W3C Speech Interface Platform.

VoiceXML is an XML language designed for creating audio dialogs that feature synthesized speech, digitized audio, recognition of spoken and DTMF key input, recording of spoken input, telephony, and mixed initiative conversations. The associated media type defined in this document is "application/voicexml+xml".

The Speech Synthesis Markup Language Specification (SSML) defines an XML-based markup language for assisting the generation of synthetic speech in Web and other applications. The essential role of SSML is to provide authors of synthesizable content a standard way to control aspects of speech such as pronunciation, volume, pitch, rate, etc. across different synthesis-capable platforms. The associated media type defined in this document is "application/ssml+xml".

The Speech Recognition Grammar Specification (SRGS) defines syntax for representing grammars for use in speech recognition so that developers can specify the words and patterns of words to be listened for by a speech recognizer. The syntax of the grammar format exists in two forms, an Augmented BNF Form and an XML Form. The respective media types defined in this document are "application/srgs" and "application/srgs+xml".

The Call Control eXtensible Markup Language (CCXML) is an XML language designed to provide telephony call control support for dialog systems, such as VoiceXML. The associated media type defined in this document is "application/ccxml+xml".

The Pronunciation Lexicon Specification (PLS) defines an XML syntax for specifying pronunciation lexicons to be used by speech recognition and speech synthesis engines in voice browser

applications. The associated media type defined in this document is "application/pls+xml".

- [2.](#) Registration of application/voicexml+xml, application/ssml+xml, application/srgs+xml, application/ccxml+xml and application/pls+xml

MIME media type name: application

MIME subtype names: voicexml+xml, ssml+xml, srgs+xml, ccxml+xml, pls+xml

Required parameters: none

Optional parameters:

"charset": This parameter has identical semantics to the charset parameter of the "application/xml" media type as specified in [RFC 3023](#) [[RFC3023](#)].

[2.1](#) Encoding considerations

Identical to those of "application/xml" as described in [RFC 3023](#) [[RFC3023](#)], [section 3.2](#).

[2.2](#) Interoperability considerations

There are no known interoperability issues.

[2.3](#) Published specifications

Voice Extensible Markup Language 2.0 [[VoiceXML2.0](#)]

Voice Extensible Markup Language 2.1 [[VoiceXML2.1](#)]

Speech Synthesis Markup Language (SSML) Version 1.0 [[SSML](#)]

Speech Recognition Grammar Specification Version 1.0 [[SRGS](#)]

Voice Browser Call Control: CCXML Version 1.0 [[CCXML](#)]

Pronunciation Lexicon Specification (PLS) Version 1.0 [[PLS](#)]

[2.4](#) Applications which use these media types

Various W3C Speech Interface Platform implementations use these media types.

Froumentin

Expires November 2, 2005

[Page 4]

Internet-Draft

W3C Speech Interface Media Types

May 2005

[2.5](#) Security Considerations

Several instructions in the cited specifications may cause arbitrary URIs to be dereferenced. In this case, the security issues of [\[RFC2396\]](#), [section 7](#), should be considered.

In addition, because of the extensibility features of those specifications, it is possible that the registered media types may describe content that has security implications beyond those described here. However, if the processor follows only the normative semantics of the specifications, this content will be ignored. Only in the case where the processor recognizes and processes the additional content, or where further processing of that content is dispatched to other processors, would security issues potentially arise. And in that case, they would fall outside the domain of this registration document.

[2.6](#) Additional Information

[2.6.1](#) Magic numbers

Although no byte sequences can be counted on to always be present,

XML MIME entities in ASCII-compatible charsets (including UTF-8) often begin with hexadecimal 3C 3F 78 6D 6C ("<?xml"), and those in UTF-16 often begin with hexadecimal FE FF 00 3C 00 3F 00 78 00 6D 00 6C or FF FE 3C 00 3F 00 78 00 6D 00 6C 00 (the Byte Order Mark (BOM) followed by "<?xml"). For more information, see [Appendix F](#) of [XML].

[2.6.2](#) File extensions

VoiceXML files: .vxml

SSML files: .ssml

SRGS files (XML syntax): .grxml

CCXML files: .ccxml

PLS files: .pls

[2.6.3](#) Fragment identifiers

Identical to that of "application/xml" as described in [RFC 3023 \[RFC3023\], section 5](#).

[2.6.4](#) Macintosh File Type Code

"TEXT"

Froumentin

Expires November 2, 2005

[Page 5]

Internet-Draft

W3C Speech Interface Media Types

May 2005

[2.6.5](#) Person & email address to contact for further information

World Wide Web Consortium <web-human@w3.org>

[2.6.6](#) Intended usage

COMMON

[2.6.7](#) Change Controller

The Speech Interface Framework specifications set is a work product of the World Wide Web Consortium's Voice Browser Working Group. The W3C has change control over these specifications.

[3.](#) Registration of application/srgs

MIME media type name: application

MIME subtype names: srgs

Required parameters: none

Optional parameters: none

[3.1](#) Encoding considerations

The ABNF Form of SRGS follows the character encoding handling defined for XML: an ABNF grammar processor must accept both the UTF-8 and UTF-16 encodings of ISO/IEC 10646 and may support other character encodings.

[3.2](#) Interoperability considerations

There are no known interoperability issues.

[3.3](#) Published specifications

Speech Recognition Grammar Specification Version 1.0 [[SRGS](#)]

[3.4](#) Applications which use this media types

Various SRGS implementations use this media type.

[3.5](#) Security Considerations

Several instructions in SRGS may cause arbitrary URIs to be dereferenced. In this case, the security issues of [\[RFC2396\]](#), [section 7](#), should be considered.

In addition, because of the extensibility features of SRGS, it is possible that the registered media types may describe content that has security implications beyond those described here. However, if the processor follows only the normative semantics of the specifications, this content will be ignored. Only in the case where the processor recognizes and processes the additional content, or where further processing of that content is dispatched to other processors, would security issues potentially arise. And in that case, they would fall outside the domain of this registration document.

[3.6](#) Additional Information

[3.6.1](#) Magic numbers

The ABNF self-identifying header must be present in any legal stand-alone ABNF Form grammar document. The first character of an ABNF document must be the "#" symbol (x23) unless preceded by an optional XML 1.0 byte order mark. The ABNF byte order mark follows the XML definition and requirements. For example, documents encoded in UTF-16 must begin with the byte order mark. The optional byte order mark and required "#" symbol must be followed immediately by the exact string "ABNF" (x41 x42 x4d x46) or the appropriate equivalent for the document's encoding (e.g. for UTF-16 little-endian: x23 x00 x41 x00 x42 x00 x4d x00 x46 x00). If the byte order mark is absent on a grammar encoded in UTF-16 then the grammar processor should perform auto-detection of character encoding in a manner analogous to auto-detection of character encoding in XML. Next follows a single space character (x20) and the required version number which is "1.0" for this specification (x31 x2e x30).

[3.6.2](#) File extensions

.gram

[3.6.3](#) Macintosh File Type Code

"TEXT"

[3.6.4](#) Person & email address to contact for further information

World Wide Web Consortium <web-human@w3.org>

[3.6.5](#) Intended usage

COMMON

[3.6.6](#) Change Controller

The SRGS specification is a work product of the World Wide Web Consortium's Voice Browser Working Group. The W3C has change control over the SRGS specification.

[4.](#) IANA Considerations

This document calls for registration of six new MIME media types, according to the registrations in [Section 2](#) and [Section 3](#).

5. IPR Disclosure Acknowledgement

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

6. Normative References

- [CCXML] Auburn, RJ., Ed., "Voice Browser Call Control: CCXML Version 1.0, W3C Working Draft", January 2005, <<http://www.w3.org/TR/2005/WD-ccxml-20050111/>>.
- [PLS] Baggia, P., Ed., "Pronunciation Lexicon Specification (PLS) Version 1.0, W3C Working Draft", February 2005, <<http://w3.org/TR/2005/WD-pronunciation-lexicon-20050214/>>.
- [RFC2396] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax. IETF [RFC 2396](#)", August 1998, <<http://www.ietf.org/rfc/rfc2396.txt>>.
- [RFC3023] Murate, M., St.Laurent, S., and D. Kohn, "XML Media Types", January 2001, <<http://www.ietf.org/rfc/rfc3023.txt>>.
- [SRGS] Hunt, A., Ed. and S. McGlashan, Ed., "Speech Recognition Grammar Specification Version 1.0, W3C Recommendation", March 2004, <<http://www.w3.org/TR/2004/REC-speech-grammar-20040316/>>.
- [SSML] Burnett, D., Ed., Walker, M., Ed., and A. Hunt, Ed., "Speech Synthesis Markup Language (SSML) Version 1.0, W3C Recommendation", September 2004, <<http://www.w3.org/TR/2004/REC-speech-synthesis-20040907/>>.
- [VoiceXML2.0] McGlashan, S., Ed., "Voice Extensible Markup Language

(VoiceXML) Version 2.0, W3C Recommendation", March 2004,
<<http://www.w3.org/TR/2004/REC-voicexml20-20040316/>>.

[VoiceXML2.1]

Oshry, M., Ed., "Voice Extensible Markup Language
(VoiceXML) Version 2.1, W3C Working Draft", July 2004,
<<http://www.w3.org/TR/2004/WD-voicexml21-20040728/>>.

Froumentin

Expires November 2, 2005

[Page 10]

Internet-Draft

W3C Speech Interface Media Types

May 2005

Author's Address

Max Froumentin
World Wide Web Consortium

Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Copyright (C) The Internet Society (2005). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.