# Introduction

## Say It Smart

Say It Smart is a Unified CVP technology that handles the breakdown of formatted data into an array of audio files played one after the other to render the data in a manner understandable by a caller. While many Text To Speech (TTS) engines can perform a similar function, the power of Say It Smart is that it can handle the playback using pre-recorded audio. Each Say It Smart type lists the audio files required to fully render all the formatted data it can handle. The user need only record these files according to the guidelines specified below and Say It Smart does the rest.

Each Say It Smart type is handled by a separate plug-in deployed on Cisco Unified Call Studio (Call Studio) and Cisco Unified CVP VXML Server (VXML Server). Unified CVP includes many common types such as dates and times. Developers can produce their own plug-ins to either extend Unified CVP Say it Smart plug-in functionality, or introduce new types.

**Note**  The grammer logic supplied with the out-of-the-box plug-in follows English grammer logic only. To achieve logic for other languages, you must develop your own plug-in.

The following defines the characteristics a Say It Smart plug-in requires:

- **Type** – A Say It Smart plug-in is associated with a single type that defines on a high level what kind of data can be handled by the plug-in. Numbers, dates, or currency values are examples of types.

- **Input Format** – A Say It Smart plug-in can have from one to many input formats that define how the data appears when it is sent to the plug-in. These formats may reflect different ways that type can be represented. For example, a date may appear in MMDDYYYY format or YYYYMMDD.

- **Output Format** – A Say It Smart plug-in can have from one to many output formats that define how to express the data passed to the plug-in. Output formats are dependent on input formats, once an input format is changed, the output formats available also change. Output formats can encapsulate differences in expression, such as reading back a value with pauses. They can also reflect language differences or even preferences in how to tailor the output. For example, a time may have an output format that reads 12:00 as *noon* or another that reads back the time in Spanish.

- **Fileset** – A Say It Smart plug-in can have from one to many filesets that list all the audio files required to render a particular output format. Filesets are dependent on output format, once an output format changes, the filesets available also change. Different filesets represent different combinations of files that will render the same data in the specified output format. The most common use of filesets is to use different groups of files to render the data so it sounds better by using more files, or using fewer files but with a more robotic sound. Another use for filesets would be to provide a different gender or playback speed. For example, a fileset may be introduced that reads back a number slowly for those applications where the audience requires it.

- **Audio Files** – Say It Smart plug-ins return a list of audio files needed to render the data in the manner specified by the above criteria. The application designer is required to record all the audio files specified by the fileset(s) they intend on using, name the audio files appropriately, and place them in a centrally servable location. Some criteria on audio files are:

  - All audio files must be given names listed in the specification (with the appropriate audio type extension). All Unified CVP Say It Smart plug-ins use filenames in lowercase and are named such that they can exist on any computing platform without naming issues (the names do not include spaces or unusual punctuation). Any naming inconsistencies will cause Unified CVP Say It Smart plug-ins to use TTS for those files.

  - All audio files for a Say It Smart format must be of a single audio type. Mixing WAV and VOX files, for example, is not possible.

  - Not all files listed need to be recorded. If the user is fairly sure some files will never be encountered, they can be left off. Unified CVP Say It Smart plug-ins use TTS as a backup so if a missing audio file is requested, it will be read as TTS. This may be a bit disconcerting to the caller but does not cause any issues for the application. For example, the Unified CVP Number Say It Smart plug-in can handle numbers up to 999 trillion and the user may know that their application will not handle numbers larger than ten thousand so may choose not to record *million*, *billion*, or *trillion*.

  - Many of the Unified CVP Say It Smart plug-ins use filesets whose contents include audio files specified by the Unified CVP Number Say It Smart plug-in. Recording the audio files to support Number will greatly reduce the number of files needed for other types.

  - All audio files for a particular plug-in must be stored within the same directory. Unified CVP Say It Smart plug-ins require the audio files used by the plug-in to reside in a single directory, though custom plug-ins can require subdirectories of this root directory.

  - Audio files must be placed in a location made accessible via an HTTP request from the voice browser. Unlike the Unified CVP software itself, serving audio files does not require an application server, they can be served by any web server such as IIS or Apache.

**Note** For types, input formats, output formats, and filesets, a plug-in defines a name for each as well as a display name. The display name is used for readability purposes and is what Call Studio shows when a new Say It Smart audio item is configured. The actual name is used by VXML Server and the developer when they build dynamic voice element configurations.

The Say It Smart plug-ins requiring the use of a pause produce VoiceXML using the `<break>` tag. Some voice browsers do not support this tag so Say It Smart playback normally including pauses on these browsers would hear no pauses.

This document presents full specifications for all Unified CVP Say It Smart plug-in types, including all input formats, output formats, filesets, and audio files required. The display names of these are also provided.