

会議における聴覚障害者補助 コンポーネントの開発

目次

1	はじめに.....	3
1.1	概要.....	3
1.1	各機能の説明.....	3
1.3	開発環境.....	4
2	ソフトウェア	5
2.1	NAME（氏名入力コンポーネント）	5
2.2	WebScraping（ウェブスクレイピングコンポーネント）	6
2.3	JudgeTheElement（音声認識結果作成コンポーネント）	7
2.4	TEXTtoSUBTITLE（字幕作成コンポーネント）	8
2.5	TEXTtoTEXTFILE（議事録作成コンポーネント）	9
2.5	KeywordMatchConfirmation（文字列比較コンポーネント）	10
2.5	KeyOutput（スライド切り替えコンポーネント）	11
3	本コンポーネント群の利用手順.....	12
3.1	Python パッケージと Webdriver のインストール.....	13
3.3	コンポーネントの接続	15

4	システムの利用	16
4.1	システムの利用	16
5	参考文献.....	17

はじめに

1.1 概要

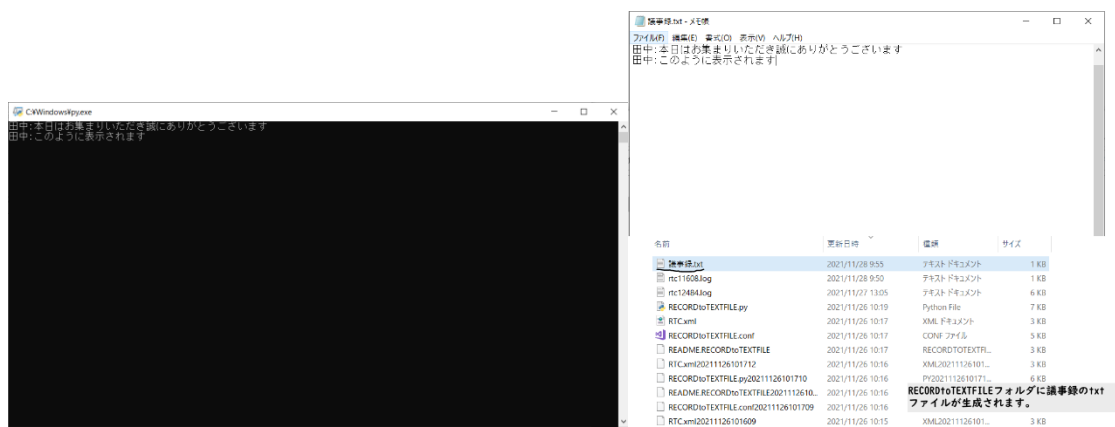
我々は,音声認識を利用した議事録作成システムとパワーポイント操作システムを作成した。このシステムを利用することで、会議中などに字幕を表示したり、音声でパワーポイントのスライドを切り替えることができる。

また、本システムの音声認識システムは、自作の音声認識 web サイトから web スクレイピングをして音声認識をしている。

1.2 各システムの説明

このシステムには、字幕機能、議事録作成機能、スライド音声切り替え機能の 3 つの機能がある。以下の表に各機能の詳細と使用しているコンポーネントを示す。

字幕機能	NAME,WebScraping1.2,JudgeTheElement,RECORDtoSUBTITLE の 5 つのコンポーネントを利用して、議事録の体裁に直した音声認識結果を RECORDtoSUBTITLE のコンソール上にリアルタイムで表示します。
議事録作成機能	NAME,WebScraping1.2,JudgeTheElement,RECORDtoTEXTFILE の 5 つのコンポーネントを利用して、議事録を作成するシステム。以下の画像の様なテキストファイルが RECORDtoTEXTFILE フォルダ内 に生成される。
スライド音声切り替え機能	WebScraping1.2,JudgeTheElement,KeywordMatchConfirmation,KeyOutput の 5 つのコンポーネントを利用して、特定のワードを言ったときに、パワーポイントにスライド変更のキー入力を送るシステム。



字幕機能(RECORDtoSUBTITLE.py コンソールに表示) ↑

議事録作成機能 ↑

1.2 開発環境

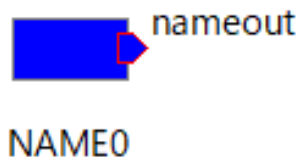
本コンポーネントの開発環境を下記の表に示す。

OS	Windows 10
開発環境	Visual Studio 2017
RT ミドルウェア	OpenRTM-aist-1.2.0-RELEASE(Python 版)
Python	Python 3.7.2

ソフトウェア

2.1 NAME（氏名入力コンポーネント）

- ・説明
アクティベートされた時に、氏名を入力するコンポーネント
- ・画像



- ・データポート
以下の表にデータポートの説明を示す。

データポート	ポート名	データ型	説明
Outport	nameout	TimedWString	議事録で作成する氏名を送信する

2.2 WebScraping（ウェブスクレイピングコンポーネント）

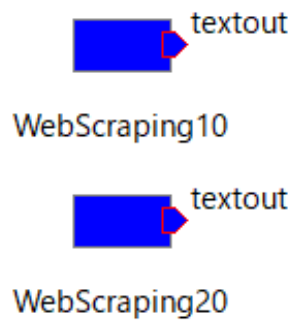
- ・説明

アクティベートされた時に、指定された URL を開き、指定された要素をスクレイピングするコンポーネント。

- ・補足説明

スクレイピングする URL と要素は、コンフィグレーション変数を書き換えることで変更することが可能。

- ・画像



- ・データポート

以下の表にデータポートの説明を示す。

データポート	ポート名	データ型	説明
Outport	textout	TimedWString	スクレイピングした要素を送信する。

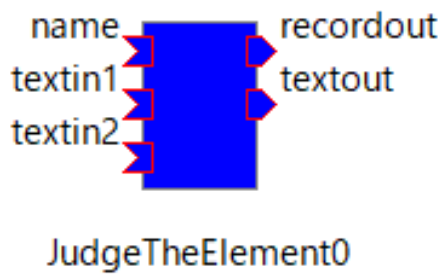
- ・依存ライブラリ

selenium

Webdriber

2.3 JudgeTheElement（音声認識結果作成コンポーネント）

- ・説明
- namein と、textin から受け取ったデータを、音声認識のみのデータを textout、議事録の体裁に整えたデータを recordout から送信する。
- ・画像



・データポート

以下の表にデータポートの説明を示す。

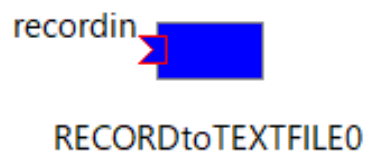
データポート	ポート名	データ型	説明
Inport	namein	TimedWString	Name コンポーネントから送られたデータを受信する。
Inport	textin1	TimedWString	WebScraping1 コンポーネントから送られたデータを受信する。
Inport	textin2	TimedWString	WebScraping2 コンポーネントから送られたデータを受信する。
Outport	recordout	TimedWString	受け取った音声認識結果と氏名を組み合わせ、議事録の体裁にし、送信する。
Outport	testout	TimedWString	受け取った音声認識結果を送信する。

2.4 TEXTtoSUBTITLE（字幕作成コンポーネント）

- ・説明

recordin から送られてきたデータを受信して、Python 上で文字を表示する。

- ・画像



- ・データポート

以下の表にデータポートの説明を示す。

データポート	ポート名	データ型	説明
Inport	recordin	TimedWString	JudgeTheElement コンポーネントから議事録の体裁の音声認識結果を受信する。

2.5 TEXTtoTEXTFILE（議事録作成コンポーネント）

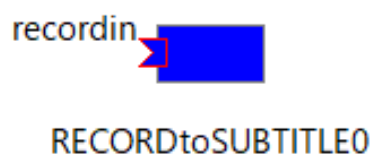
- ・説明

recordin から送られてきたデータを受信して、ディアクティベートされた時に議事録をテキストファイルとして作成する。

- ・補足説明

テキストファイル名は、コンフィグレーション変数を書き換えることで変更が可能。

- ・画像



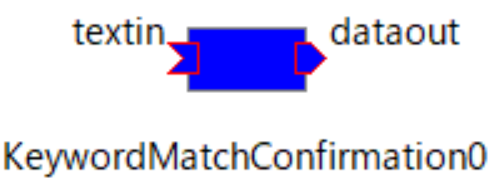
- ・データポート

以下の方にデータポートの説明を示す。

データポート	ポート名	データ型	説明
Inport	recordin	TimedWString	JudgeTheElement コンポーネントから議事録の体裁の音声認識結果を受信する。

2.6 KeywordMatchConfirmation（文字列比較コンポーネント）

- ・説明
アクティベートされた時に、スライドを変更する時のワードを入力する。textin から受け取った音声認識結果がスライドを変更するワードと一致した場合は true、一致しなかった場合は false を dataout から送信する。
- ・画像



- ・データポート
以下の方にデータポートの説明を示す。

データポート	ポート名	データ型	説明
Inport	textin	TimedWString	JudgeTheElement コンポーネントから音声認識結果を受信する。
Outport	dataout	TimedWString	KeyOutput コンポーネントに成否判定結果を送信する。

2.7 KeyOutput (スライド切り替えコンポーネント)

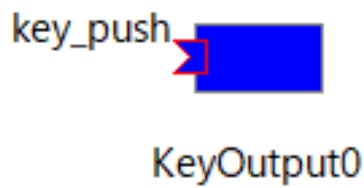
- ・説明

datain から受信したデータが true か false かを判断し、true だった場合はパワーポイントに n キーを文字入力する。

- ・補足説明

入力する文字はコンフィグレーション変数を書き換えることで変更可能

- ・画像



- ・データポート

以下の方にデータポートの説明を示す。

データポート	ポート名	データ型	説明
Inport	textin	TimedWString	KeywordMatchConfirmation コンポーネントからデータ を受信する。

- ・依存ライブラリ

PyAutoGUI

本コンポーネント群の利用手順

3.1 Python パッケージと Webdriver のインストール

- ・「Webdriver」

このシステムは、今回は、「Webdriver」と、後述の Python パッケージ「selenium」を利用した web スクレイピングで音声認識をしている。

- ・ Python パッケージ「Selenium」

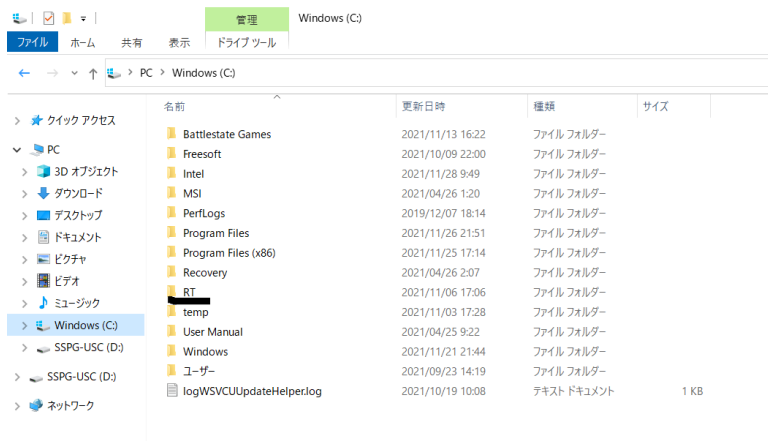
このシステムでは、前述の「Webdriver」と、Python パッケージ「Selenium」を利用して web スクレイピングをしている。

- ・ Python パッケージ「PyAutoGUI」

このシステムのスライド自動切り替え機能は、Python パッケージ「PyAutoGUI」を利用している。

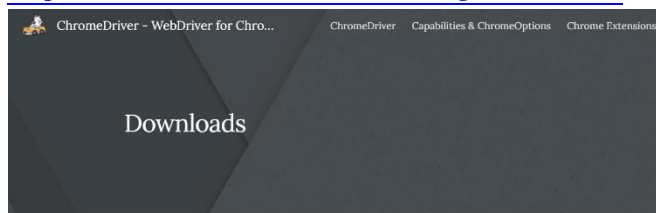
- ・「Webdriver」のインストール手順

[1]Cドライブ直下に、「RT」という名前のフォルダを作成する。



[2]以下の URL から、chrome のバージョンに対応したファイルをインストールする。

<https://chromedriver.chromium.org/downloads>



Current Releases

- If you are using Chrome version 97, please download [ChromeDriver 97.0.4692.20](#)
- If you are using Chrome version 96, please download [ChromeDriver 96.0.4664.45](#)
- If you are using Chrome version 95, please download [ChromeDriver 95.0.4635.69](#)

• For older version of Chrome, please see below for the version of ChromeDriver that supports it.

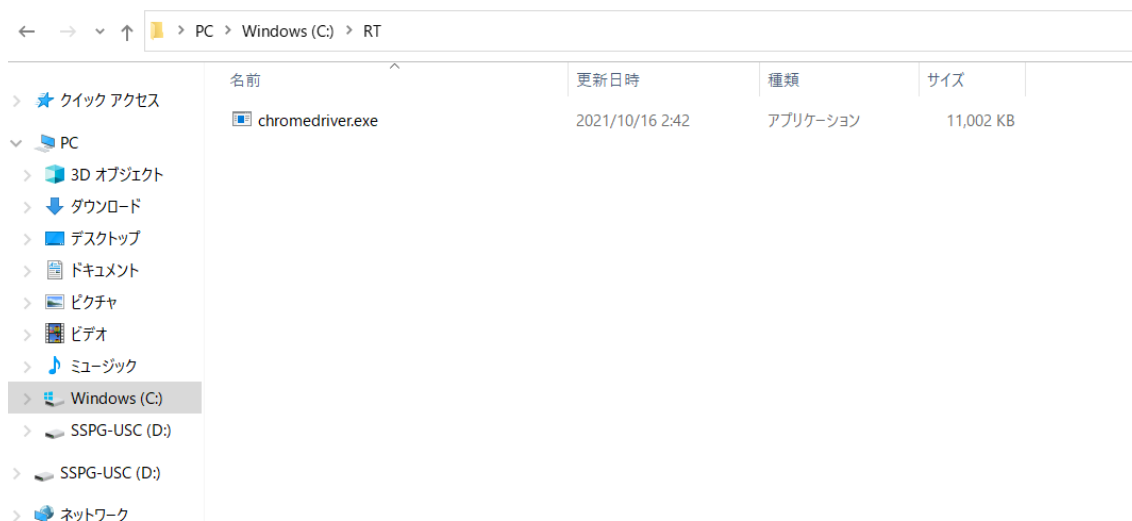
If you are using Chrome from Dev or Canary channel, please following instructions on the [ChromeDriver Canary](#) page.

For more information on selecting the right version of ChromeDriver, please see the [Version Selection](#) page.

[ChromeDriver 97.0.4692.20](#)

ここから、使用しているChromeのバージョンに最も近いバージョンを選択し、ダウンロードします。

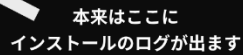
[3] インストールした ZIP ファイルを解凍し、中に入っている chromedriver.exe を [1] で作った RT フォルダに移動させる。



[1] コマンドプロンプトを起動する

[2]以下のようなコマンドを実行する

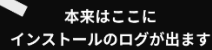
```
pip install selenium
```



[1] コマンドプロンプトを起動する

[2] 以下のようなコマンドを実行する。

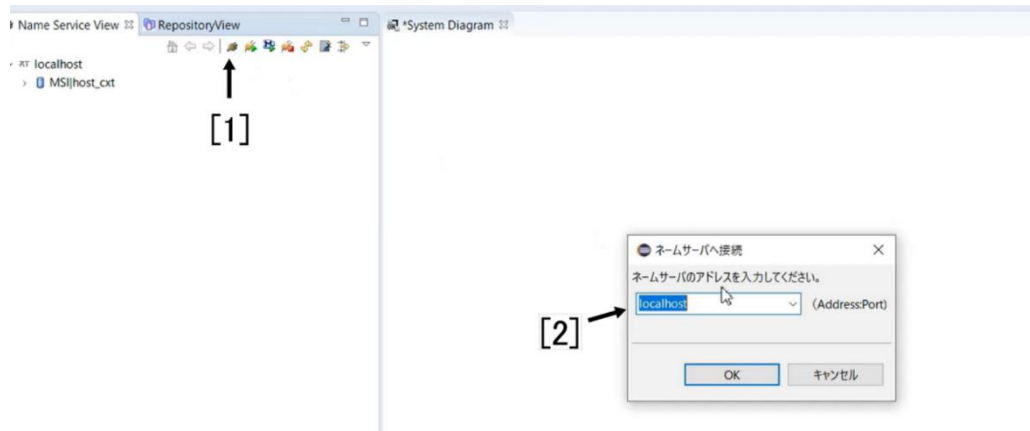
```
pip install pyautogui
```



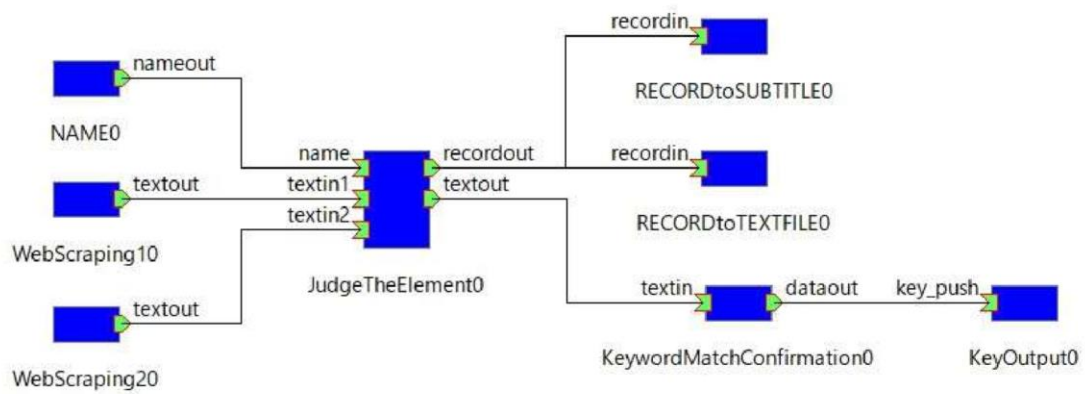
3.2 コンポーネントの接続

[1] ネームサーバを起動する。

[2] 複数人で利用する際は、以下の図のようにネームサーバーを接続する。



[3] 以下の図のようにコンポーネントを接続する。



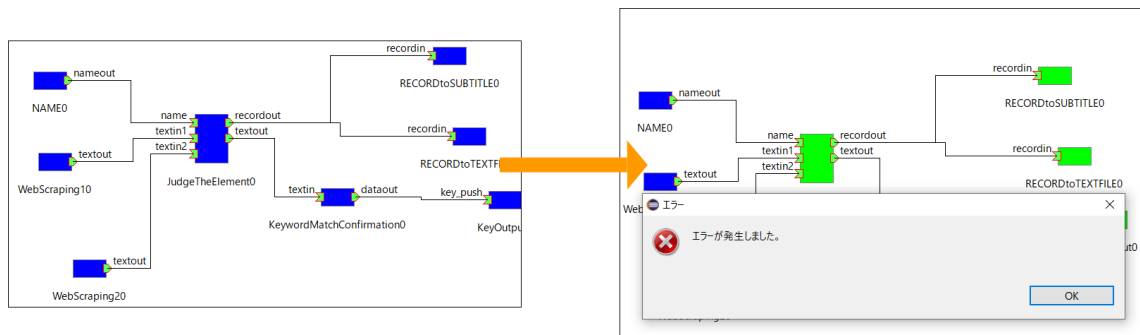
システムの利用

4.1 システムの利用

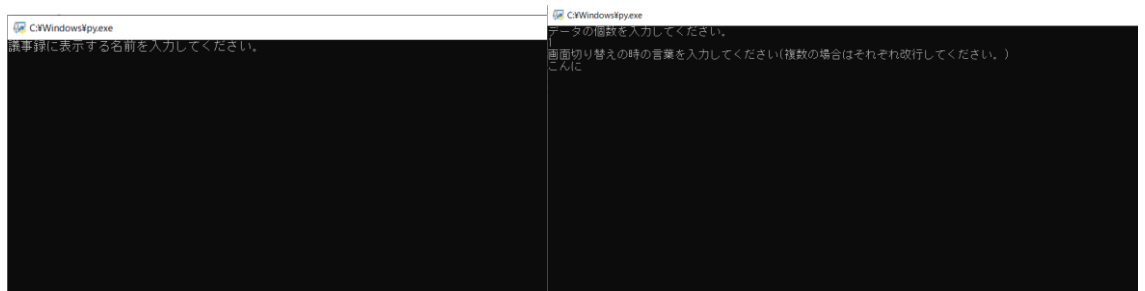
このシステムの利用方法を示す。

[1]以下のように接続したコンポーネントをアクティベートする。

*アクティベート時にエラーが出ますが、ウェブスクレイピング用のブラウザの起動時のロードの為です。OK を押すと正常に作動します。



[2]NAME.py と KeywordMatchConfirmation.py にそれぞれ氏名とスライド切り替えワードを入力する。



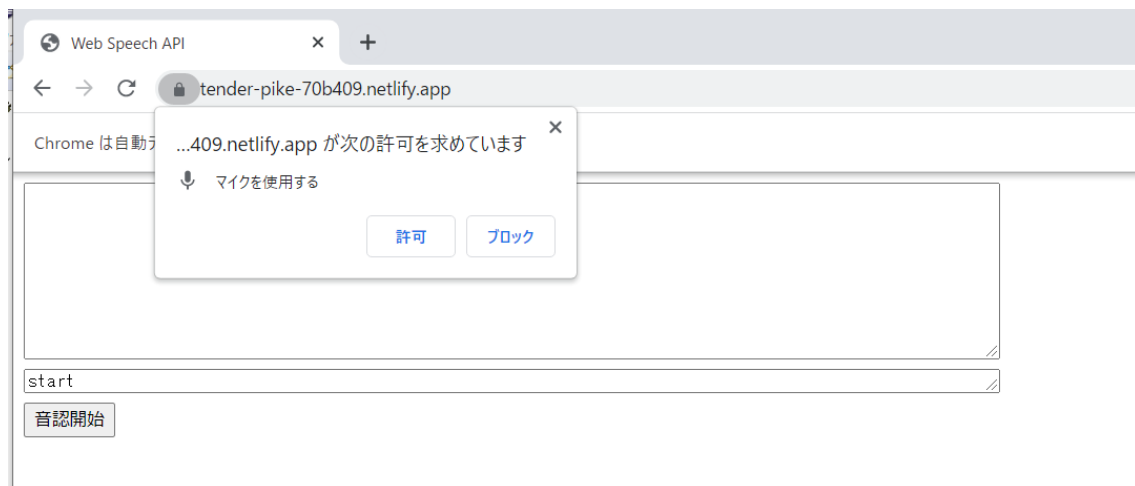
↑ NAME.py

↑ KeywordMatchConfirmation.py

[3]使用するパワーポイントのスライドを開始する。

[4]開かれた音声認識ウェブサイトのマイクのアクセスを許可する。

ここで正常に作動していれば、何かしゃべるとウェブサイトに音声認識結果が表示される。



[5]パワーポイントのスライドのウィンドウを最前列に表示する。

参考文献

久保田浩平: ゴミ捨て動作により映像が 変化するディスプレイ付き ゴミ箱の開発