

# SiVoDiM

---

## Sintesi Vocale per Dispositivi Mobili



### Verbale esterno 18/03/2016

Versione	1.0.0
Redattori	Gino Zaidan
Verificatori	Alberto Andriolo Riccardo Rizzo
Responsabili	Enrico Chiara
Uso	Esterno
Lista di distribuzione	Stark Labs Prof. Tullio Vardanega, Prof. Riccardo Cardin

Verbale di incontro tra il gruppo Stark Labs e il Proponente Giulio Paci, MIVOQ s.r.l.  
per il progetto SiVoDiM

## Registro delle modifiche

Attività	Autori	Data	Versione
Accettazione	Enrico Chiara	19/03/2016	1.0.0
Verifica	Riccardo Rizzo	19/03/2016	0.2.0
Correzioni errori	Gino Zaidan	19/03/2016	0.1.1
Verifica	Alberto Andriolo	18/03/2016	0.1.0
Stesura del documento	Gino Zaidan	18/03/2016	0.0.2
Creazione struttura documento	Gino Zaidan	18/03/2016	0.0.1

## Indice

<b>1</b>	<b>Informazioni Generali</b>	<b>1</b>
1.1	Riferimenti . . . . .	1
1.1.1	Informativi . . . . .	1
<b>2</b>	<b>Ordine del giorno</b>	<b>2</b>
2.1	Approfondimento sul processo di campionamento . . . . .	2
2.2	Modalità di accesso al sistema di campionamento dall'applicazione . . . . .	2
2.3	Utilizzo del modello di campionamento . . . . .	2
2.4	Utilizzo dei linguaggi di markup per esprimere le emozioni . . . . .	2
2.5	Tecniche grafiche consigliate per migliorare l'accessibilità su dispositivi mobili	2
2.6	Istruzioni per installazione e uso . . . . .	3
<b>3</b>	<b>Risultato dell'incontro</b>	<b>4</b>

## 1 Informazioni Generali

- **Data:** 18/03/2016;
- **Ora:** 15:00;
- **Durata:** 120 minuti;
- **Luogo:** Aula 1BC50, Torre Archimede, Via Trieste, 63, 35121 Padova PD;
- **Partecipanti interni:** Stark Labs
  - Alberto Andriolo;
  - Enrico Chiara;
  - Federico Rossetto;
  - Francesco Bizzaro;
  - Gino Zaidan;
  - Riccardo Rizzo.
- **Partecipanti esterni:** MIVOQ s.r.l.
  - Dott. Giulio Paci.

### 1.1 Riferimenti

#### 1.1.1 Informativi

- *Glossario v1.0.0.*

## 2 Ordine del giorno

Di seguito sono trascritti gli argomenti che il gruppo ha trattato durante la riunione con il referente Giulio Paci di MIVOQ s.r.l.. Il *meeting* ha avuto carattere informale e si è svolto mediante una conferenza su Skype<sub>G</sub>.

### 2.1 Approfondimento sul processo di campionamento

Il campionamento consiste nella lettura da parte dell'utente di un certo numero di frasi che vengono registrate e inviate al server di MIVOQ. Una volta elaborati i dati ricevuti, il server crea un nuovo modello di voce che può essere scaricato o utilizzato direttamente dal servizio *web* dell'azienda. Il *file* scaricato è compresso all'interno di un pacchetto in formato *.jar<sub>G</sub>*, la cui dimensione si attesta mediamente attorno ai 3 MB<sub>G</sub>, e che contiene tutti i dati relativi alla voce dell'utente. Tali informazioni consistono in modelli statistici del timbro vocale, necessari per una corretta sintesi della voce dell'utente. L'intero processo può essere completato nel corso di più fasi. Pertanto si consente all'utente di interrompere il campionamento e di riprenderlo in un secondo momento, conservando i dati elaborati precedentemente all'interruzione.

### 2.2 Modalità di accesso al sistema di campionamento dall'applicazione

Il meccanismo che l'azienda proponente vuole implementare è basato su OAuth 2.0<sub>G</sub>. Per tale ragione, gli utenti dovranno consentire all'applicazione di avere accesso ai loro dati: ogni utente dovrà quindi essere in possesso di un *account*.

### 2.3 Utilizzo del modello di campionamento

Il modello della voce campionata rimane privato e utilizzabile solamente dall'utente. Il modello è anche installabile nel server, ma una volta installato questo diventa pubblico e accessibile a tutti gli utenti.

### 2.4 Utilizzo dei linguaggi di markup per esprimere le emozioni

Il gruppo ha richiesto quale linguaggio di markup utilizzare per l'espressione dei sentimenti: il referente Giulio Paci ha consigliato l'uso del linguaggio SSML<sub>G</sub>, meno potente rispetto a Emotionml<sub>G</sub>, ma ben supportato dal motore FA-TTS<sub>G</sub>.

### 2.5 Tecniche grafiche consigliate per migliorare l'accessibilità su dispositivi mobili

Durante una discussione nata a proposito di possibili interfacce da implementare su mobile, il referente Giulio Paci ha presentato al gruppo una pagina *web* provvisoria che MIVOQ s.r.l. sta sviluppando per migliorare la modifica e l'inserimento degli effetti da applicare a una voce. Il servizio è disponibile al seguente link:

<https://www.mivoq.it/ttsdemo>

L'adozione di *slider<sub>G</sub>* per la gestione delle impostazioni può essere considerata come una fra le possibilità per migliorare l'accessibilità lato utente. Il gruppo ha pertanto preso in considerazione l'opportunità di strutturare l'interfaccia su mobile con un'impostazione grafica simile.

## 2.6 Istruzioni per installazione e uso

Il referente Giulio Paci ha spiegato il procedimento necessario all'installazione del motore FA-TTS<sub>G</sub> su computer o macchina virtuale equipaggiate con sistema operativo Debian<sub>G</sub> o Ubuntu<sub>G</sub>. Ha inoltre presentato le due API<sub>G</sub> che il server Mivoq TTS HTTP implementa:

- La FA-TTS interface, ancora in fase di sviluppo;
- La original Mary TTS interface.

La documentazioni delle API<sub>G</sub> presentate è disponibile presso il seguente link

<http://fic2fatts.tts.mivoq.it/documentation.html>

Infine sono stati discussi i diversi parametri necessari a ottenere un effetto vocale.

### 3 Risultato dell'incontro

L'incontro ha fatto maggior chiarezza su gran parte delle perplessità sorte durante l'analisi dei requisiti. Inoltre sono stati discussi, assieme al Proponente, i vari casi d'uso individuati dal *team* prima del giorno dell'incontro. Infine, a seguito di nuove preferenze espresse dal Proponente, è stato deciso di ampliare l'applicazione ideata in partenza con due nuove funzionalità:

- Arricchimento degli sceneggiati attraverso l'inserimento di audio come sottofondi musicali ed effetti sonori;
- Esportazione dello sceneggiato in formato video, in modo da facilitarne la condivisione e la diffusione sulle piattaforme *social*.