

SiVoDiM

Sintesi Vocale per Dispositivi Mobili



Verbale esterno 09/03/2016

Versione	1.0.0
Redattori	Francesco Bizzaro
Verificatori	Alberto Andriolo Riccardo Rizzo
Responsabili	Enrico Chiara
Uso	Esterno
Lista di distribuzione	Stark Labs Prof. Tullio Vardanega, Prof. Riccardo Cardin Giulio Paci, MIVOQ s.r.l.

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

Registro delle modifiche

Attività	Autori	Data	Versione
Accettazione	Enrico Chiara	10/03/2016	1.0.0
Verifica	Alberto Andriolo	10/03/2016	0.2.0
Correzione errori	Francesco Bizzaro	10/03/2016	0.1.1
Verifica	Riccardo Rizzo	10/03/2016	0.1.0
Stesura del documento	Francesco Bizzaro	09/03/2016	0.0.2
Creazione struttura documento	Francesco Bizzaro	09/03/2016	0.0.1

Indice

1	Informazioni Generali	1
1.1	Riferimenti	1
1.1.1	Informativi	1
2	Ordine del giorno	2
2.1	Precisazione sullo scopo del progetto	2
2.2	Motivazioni e vantaggi dello sviluppo su dispositivi mobile	2
2.3	Problematiche riguardanti la connessione al servizio remoto	2
2.4	Integrazione del modulo di sintesi con il sistema	2
2.5	Implementazione separata delle varie componenti	2
2.6	Tempo di campionamento della voce	3
2.7	Strumenti di lavoro offerti da MIVOQ	3
3	Risultato dell'incontro	4

1 Informazioni Generali

- **Data:** 09/03/2016;
- **Ora:** 15:00;
- **Durata:** 90 minuti;
- **Luogo:** MIVOQ s.r.l., sede di Padova;
- **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.

2.1 Precisazione sullo scopo del progetto

Il capitolato è rimasto volutamente vago, per non imporre alcun limite sul tipo di applicazione da realizzare. L'unico requisito primario è l'utilizzo del motore di sintesi, con lo scopo ultimo di metterne in risalto le potenzialità. Tuttavia è fortemente desiderabile che l'applicazione si dimostri utile per qualche utilizzo reale, in linea con gli esempi suggeriti nel capitolato.

2.2 Motivazioni e vantaggi dello sviluppo su dispositivi mobile

La scelta è stata dettata dal desiderio del Proponente di entrare in tale settore. Le motivazioni date, durante la riunione, sono le seguenti:

- Incentivare la diffusione dell'applicazione;
- Utilità nell'avere l'applicativo su un dispositivo *mobile*, il quale risulta facilmente trasportabile per eventuali dimostrazioni della tecnologia FA-TTS_G.

2.3 Problematiche riguardanti la connessione al servizio remoto

Le problematiche emerse possono essere di vario genere:

- La connessione può cadere;
- I tempi di risposta della connessione potrebbero non essere ottimali:
 - I *file* da trasferire sono codificati in formato WAV_G e pertanto bisogna tenere conto del loro peso;
 - L'elaborazione effettuata dal *server* necessita di un certo tempo d'attesa.

Per poter arginare il problema del ritardo, si potrebbero adottare strategie di *caching*_G, conoscendo con anticipo i possibili testi da elaborare.

2.4 Integrazione del modulo di sintesi con il sistema

Il modulo *software* deve essere capace di integrare il servizio remoto di MIVOQ con il sistema *mobile* nel quale viene eseguita l'applicazione. Per risolvere la possibile assenza di connessione, che renderebbe impossibile lo sfruttamento del motore di sintesi, devono essere implementati meccanismi di *fallback*_G, utilizzando per esempio le voci presenti nel sistema stesso. In questo modo è possibile effettuare la sintesi con le voci di sistema e consentire all'applicazione di funzionare correttamente. Tale applicazione deve permettere all'utente di aggiungere voci in modo dinamico e di poterle salvare con determinati parametri all'interno del sistema.

2.5 Implementazione separata delle varie componenti

L'applicazione deve essere formata da quattro parti:

- Un modulo per la sintesi, a sé stante, che corrisponde all'implementazione del motore FA-TTS_G di MIVOQ;

- Un'applicazione per la configurazione, che deve essere in grado di interagire con il modulo per la sintesi al fine di modificarne determinati parametri. Ad esempio si vuole fornire la possibilità di aggiungere nuove voci, assieme a nuovi *preset_G* di effetti associabili alle voci;
- Una libreria_G che faciliti l'utilizzo delle funzionalità aggiuntive, permettendo il riuso della libreria stessa;
- Un'applicazione innovativa che sfrutti le suddette componenti e che dimostri in modo chiaro le potenzialità offerte dal motore FA-TTS_G.

Malgrado sia preferibile l'approccio modulare appena descritto, è stato specificato che è possibile sviluppare un'unica applicazione contenente tutte le funzionalità sopra elencate.

2.6 Tempo di campionamento della voce

Il processo di campionamento richiede circa 45 minuti per essere completato. Tale attesa è determinata dalla necessità di raccogliere ed elaborare un elevato numero di dati, tant'è che solitamente occorrono più di 100 frasi per ottenere un buon risultato. Il microfono incorporato negli *smartphone_G* odierni è adatto per ottenere un campionamento di qualità soddisfacente. Considerando l'elevato tempo di attesa, è di estrema importanza assistere al meglio l'utente durante il processo di registrazione.

2.7 Strumenti di lavoro offerti da MIVOQ

MIVOQ fornisce il motore di sintesi assieme alle voci predefinite, materiale che può essere scaricato direttamente dal sito dell'azienda. Su richiesta del gruppo sarà possibile avere accesso ad un sito interno di MIVOQ per campionare la propria voce e poterla riutilizzare all'interno dell'applicazione da realizzare. Non viene dato alcun vincolo sugli strumenti di sviluppo utili per implementare le applicazioni richieste.

3 Risultato dell'incontro

L'incontro ha chiarito alcuni dei dubbi sorti fra i membri del gruppo in seguito alla presentazione del capitolato e ad un suo studio iniziale. Come conseguenza delle direttive date e delle preferenze espresse dal proponente, è stato deciso di implementare un'applicazione per dispositivi *mobile* in grado di strutturare, gestire e registrare interi sceneggiati associando a ciascun personaggio una voce personalizzata.