



VOIce-based Community-cEntric mobile Services for social development

Grant Agreement Number 269954

**Deliverable No D1.5**  
**First iteration of High Level Architecture**

September 2011

---

**SEVENTH FRAMEWORK PROGRAMME**  
THEME ICT-2009.9.1 – International cooperation



## PROJECT DELIVERABLE REPORT

Project	
Grant Agreement number	2699542
Project acronym:	VOICES
Project title:	<i>VOIce-based Community-cEntric mobile Services for social development</i>
Funding Scheme:	<i>Collaborative Project</i>
Date of latest version of Annex I against which the assessment will be made:	<i>18 February 2011</i>
Document	
Deliverable number:	D1.5
Deliverable title	First iteration of High Level Architecture
Contractual Date of Delivery:	September 2011
Actual Date of Delivery:	30 September 2011
Editor (s):	
Author (s):	David Blaisonneau (FT), Filipe Cabral Pinto (PTIN), Stéphane Boyera (WF), Morgan Richomme (FT)
Reviewer (s):	FT, PTIN, TNO, WebFoundation
Work package no.:	WP1
Work package title:	Sustainable architecture
Work package leader:	TNO
Work package participants:	WebFoundation, FT, VUA, CRS4, ESMT, SE, PTIN
Distribution:	PU
Version/Revision:	1.0
Draft/Final:	Final
Total number of pages (including cover):	20
Keywords:	High level architecture

**CHANGE LOG**

<b>Reason for change</b>	<b>Revision</b>	<b>Author</b>	<b>Date</b>
First draft	0.2	David Blaisonneau	29/07/2011
Corrections	0.3	Filipe Cabral Pinto	03/08/2011
Add summary. Correction on WP5 requirements. Correction and adds in 'Technical needs description'	0.4	David Blaisonneau	02/09/2011
Add TTS for WP4 use cases	0.5	David Blaisonneau	12/09/2011
Add Crowd sourcing for WP3	0.6	David Blaisonneau	15/09/2011
First review	0.7	Etienne Barnard	20/09/2011
Add Performance and ASR openness recommendations. Add conclusion. Final review	1.0	David Blaisonneau	29/09/2011

## DISCLAIMER

This document contains description of the VOICES project work and findings.

The authors of this document have taken any and all available measures in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any responsibility for actions that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the VOICES consortium and can in no way be taken to reflect the views of the European Union.

The European Union is established in accordance with the Treaty on European Union (Maastricht). There are currently 27 Member States of the Union. It is based on the European Communities and the member states cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice and the Court of Auditors. (<http://europa.eu.int/>)



**VOICES is a project funded in part by the European Union.**

## Table of contents

Flexibility .....	8
Reachability .....	8
Accessibility .....	8
Openness and Re-usability .....	8
Long-term sustainability .....	8
WP4 – m-Health .....	9
Use case 1 – Analysis result collection .....	9
Use case 2 - Quiz.....	10
Use case 3 – Information letter .....	11
WP5 – m-Agri .....	12
Use case 1 – Market information system .....	12
Use case 2 - M-Event Organizer for Regreening events .....	13
Multi channel accessibility .....	16
Service hosting and execution .....	16
Service execution.....	16
File Hosting .....	16
Database .....	16
Digital Asset Management .....	16
Crowd sourcing.....	16
Voices tools .....	16
IVR.....	16
ASR.....	17
TTS .....	17
Developers needs.....	17
Administration needs.....	17
Mobile Applications.....	17
Platform capacities.....	17

## Diagrams table

diagram 1: WP4 - Use case 1 .....	10
diagram 2: WP4 - Use case 2 .....	11
diagram 3: WP4 - Use case 3 .....	12
diagram 4: WP5 - Use case 1 .....	13
diagram 5: WP3 crowd sourcing tool.....	15
diagram 6: High level architecture .....	19

## **SUMMARY**

The report presents the VOICES high level architecture based on requirements elaborated in the task 1.1-Use cases and requirements. This architecture is built to be flexible and dedicated to emerging regions.

## ARCHITECTURE MAIN GOALS

### **Flexibility**

The VOICES platform must be as flexible as possible. It means that each component can be replaced by another with same functionalities. This implies that all components must target open and documented interfaces. This also implies that the services developers can use third party services through internet or another telecommunication channel.

### **Reachability**

The VOICES platform shall provide multiple accesses to reach the services hosted on it. In the VOICES context, the targeted public is mainly composed with mobile phones users, radio listeners and most of them share internet access when it is available.

### **Accessibility**

The VOICES platform shall provide components to enhanced service accessibility to reach illiterate people and local languages speakers.

### **Openness and Re-usability**

To be used after VOICES project or by other persons, the VOICES platform shall be built as much as possible on open components, free softwares, open interfaces and open data structures. The VOICES platform shall also provide an access to all services sources and documentation.

### **Long-term sustainability**

The architecture must not prevent the local adoption and exploitation of the VOICES tools and methods beyond the project lifetime.



## PILOTS NEEDS

### WP4 – m-Health

#### Use case 1 – Analysis result collection

This use case is a daily data collection from users without an internet access. This collection can be done by two ways:

- ⤴ using a graphical user interface on a mobile phone, then this application send preformatted SMS to the platform.
- ⤴ using call from a mobile to a Interactive Voice Responder.

These datas are collected, analysed and stored on the platform. The results must be consultable by a supervisor on an internet interface.

Fields collected can have those formats:

- ⤴ dates
- ⤴ IDs
- ⤴ sets of possibilities
- ⤴ free fields (string or audio format)

The supervisor interface shall provide:

- ⤴ An history of all declarations sortable by date and or lab ID
- ⤴ A possibility to listen an audio field and write the field as a string
- ⤴ An exportable monthly report
- ⤴ A map with reported data
- ⤴ A function to send SMS or voices messages to users that forget to declare the lasts cases.
- ⤴ A interface to add/modify/remove labs, users...

This is a diagram of this use case:

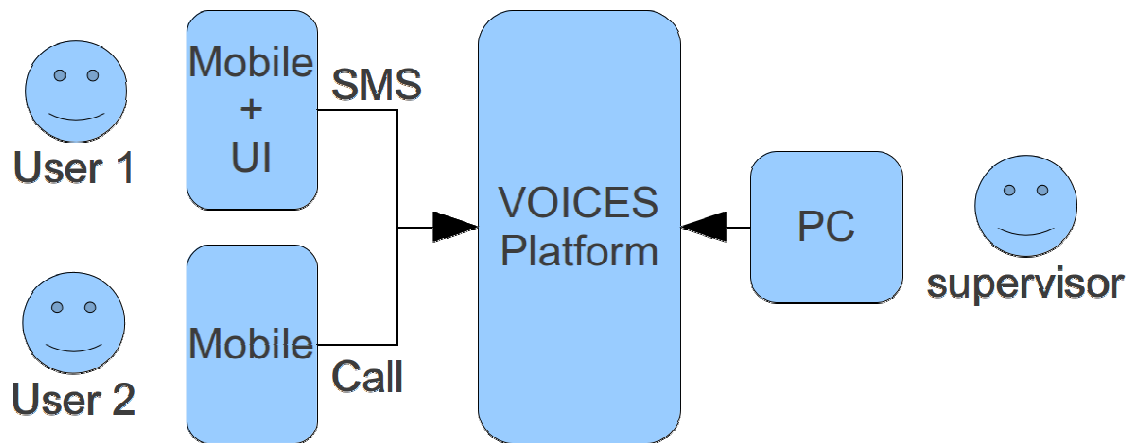


diagram 1: WP4 - Use case 1

The technical needs for this use case are:

- ^ mobile application
- ^ mobile/fixe originated calls to an interactive voice responder
- ^ mobile to platform SMS
- ^ platform to mobile SMS
- ^ web interface
- ^ database
- ^ French TTS engine

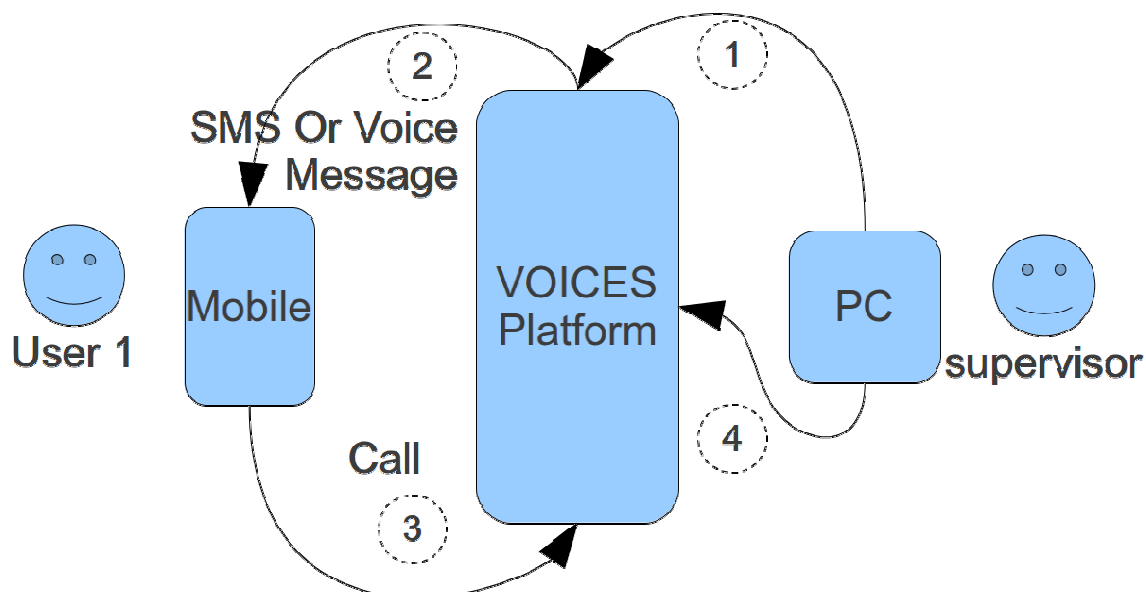
Needs of local number for countries: Senegal

### Use case 2 - Quiz

This use case is an anonymous vocal quiz for mobile users. A supervisor creates the quiz and then launches a massive SMS/call campaign to ask users to do the quiz.

This service runs in four steps:

1. The supervisor creates the quiz on an internet interface by writing the questions and the answers in text fields. The platform or the interface transforms those texts to audio phrases and the supervisor can listen and correct the spell.
2. When the supervisor validates the quiz, a mass notification campaign is launched. For a SMS notification, the SMS asks the user to call a specific number. For a call notification, the user can answer the quiz directly or call back latter.
3. The user answers to the quiz, and after it , he can have more information if he wished.
4. The supervisor can see statistics of the quiz.



*diagram 2: WP4 - Use case 2*

The technical needs for this use case are:

- ⤴ web interface
- ⤴ database
- ⤴ platform to mobile SMS
- ⤴ platform to phone calls
- ⤴ phone to platform calls
- ⤴ interactive voice responder
- ⤴ French TTS engine

Needs of local number for countries: Senegal

### **Use case 3 – Information letter**

This use case is a mass information campaign for mobile users. A supervisor creates the information letter then launches a massive SMS/call campaign to ask users to consult it.

This service runs in four steps:

1. The supervisor creates the information letter on an internet interface by writing the rubrics in text fields. The platform or the interface transforms those texts to audio phrases and the supervisor can listen and correct the spell.
2. When the supervisor validates the letter, a mass notification campaign is launched. For a SMS notification, the SMS asks the user to call a specific number. For a call notification, the user can listen directly or call back latter.
3. The user listen the letter and navigate through rubrics using DTMF or voice orders or consult old information letters
4. The supervisor can see statistics about the letter consultation.

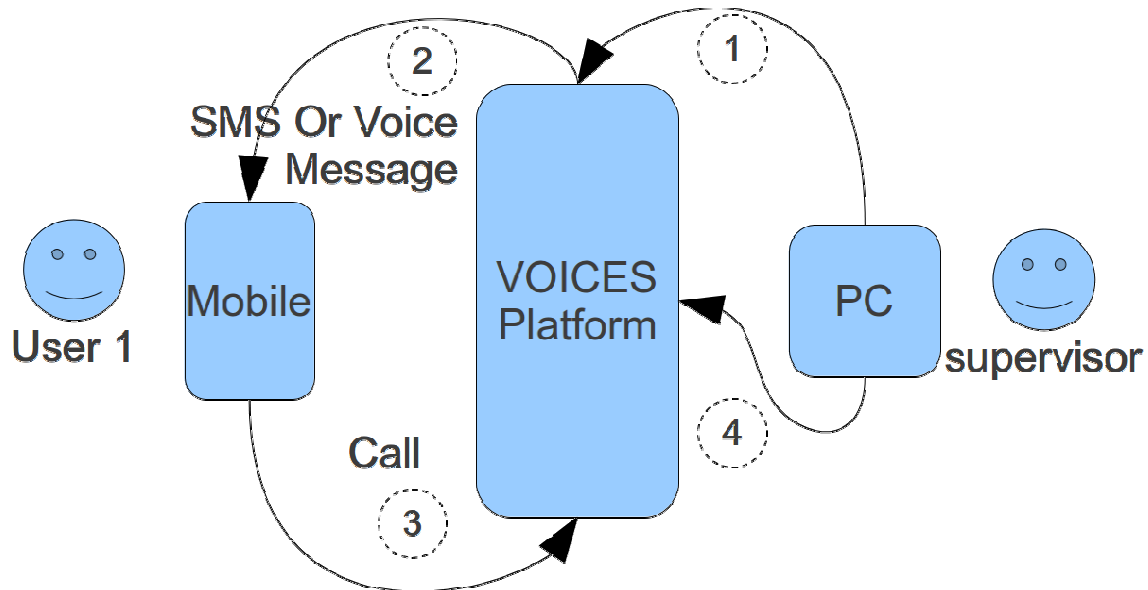


diagram 3: WP4 - Use case 3

The technical needs for this use case are:

- ⤴ web interface
- ⤴ database
- ⤴ platform to mobile SMS
- ⤴ platform to phone calls
- ⤴ phone to platform calls
- ⤴ interactive voice responder
- ⤴ French TTS engine

Needs of local number for countries: Senegal

## WP5 – m-Agri

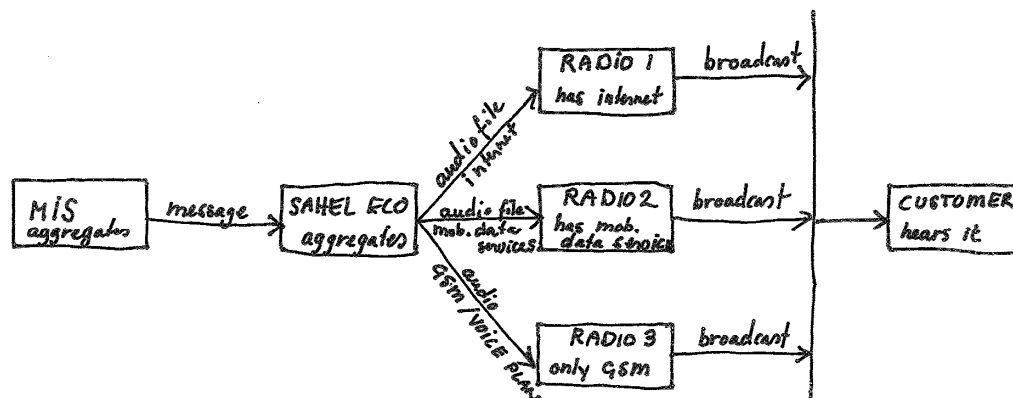
### Use case 1 – Market information system

Extract from W4RA\_UC\_V4.0 document:

This use case will focus on the MIS communiqué transfer from Sahel Eco to Radio.

This can be done through voice and phone, depending on the options available at the radio end. Specifically this can/will be:

- ⤴ Downloading the MIS communiqué AUDIO file and Broadcast:- Computer + Internet is available
- ⤴ Downloading the MIS communiqué AUDIO file directly to the phone and Broadcast:- Mobile Data Services available
- ⤴ Broadcast directly from the mobile phone at the Community Radio: GSM/Some voice plan with the Radio Station is available



### Sheabutter & honey usecase

#### diagram 4: WP5 - Use case 1

The technical needs for this use case are:

- ⤴ web interface
- ⤴ database
- ⤴ Text To Speech function
- ⤴ interactive voice responder (VoiceXML is asked by pilot developers)
- ⤴ massive SMS campaign

Needs of local number for countries: Mali

### Use case 2 - M-Event Organizer for Regreening events

This use case is an event notification and tracking system including the option of co-opting new members in the system.

This service runs in three steps:

1. Sahel Eco enter the event details on a web form
2. Sahel Eco broadcast the information via a voice broadcast message
  1. A subscribed user can answer the call directly
  2. A subscribed user can call latter to listen the message
  3. A non subscribed user can call to a given number to access the information triggering its automatically registration
3. The user can accept or reject the invitation, and leave a message

The technical needs for this use case are:

- ⤴ web interface
- ⤴ database
- ⤴ Text To Speech function
- ⤴ platform to phone calls
- ⤴ phone to platform calls

VOICES – 269954

[www.mvoices.eu](http://www.mvoices.eu)

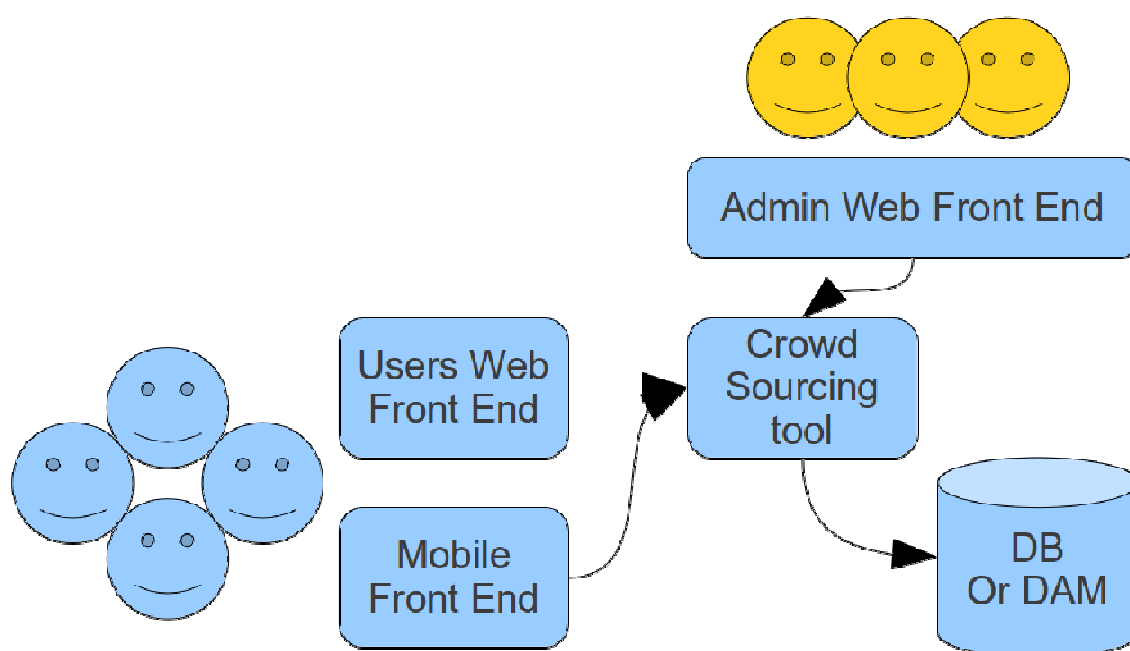
- ⤴ Interactive Voice Responder (VoiceXML is asked by pilot developers)
- ⤴ Audio file storage

Needs of local number for countries: Mali

## VOICE TOOLS NEEDS – WP3

Deliverable D3.1 provides a high-level overview of the approaches employed in current speech technology systems and discusses some of the resources needed to develop such systems in new languages. Those resources are mainly corpora of words, sentences and recordings, extended with acoustic, phonemic, syntactic and grammatical annotations.

To minimize the expertise and costs needed to develop such resources for the VOICES project and beyond, a crowdsourcing approach needs to be setup. This requires integrating to the project platform some new software components: a crowdsourcing framework and some dedicated interfaces to help workers to perform the tasks (i.e., entering sentences, annotating sentences, recordings, etc.)



*diagram 5: WP3 crowd sourcing tool*

The technical needs are:

- ^ web interface
- ^ database or DAM (Digital Asset Management)
- ^ Text To Speech function
- ^ phone to platform calls
- ^ Interactive Voice Responder (VoiceXML is asked by pilot developers)
- ^ Croud sourcing software
- ^ Optional:
  - o Platform to mobile SMS

Needs of local number for countries: Mali

## TECHNICAL NEEDS DESCRIPTION

The VOICES platform must provide those capacities:

### Multi channel accessibility

Data or services can be available from different media channels (in both way, from or to the platform):

- ^ fixed phone call
- ^ mobile call
- ^ SMS
- ^ Internet

The channels must be available for those countries:

- ^ Senegal
- ^ Mali

### Service hosting and execution

#### Service execution

The service must be executable on the platform. The service development language will depend on the components used, but open and well known development languages such as PHP are recommended.

#### File Hosting

All files needed for the services or created by them shall be hostable by the platform and sharable between the components (in the limit of a fair usage).

#### Database

A database must be available for each service. A relational database management system such as MySQL is recommended.

#### Digital Asset Management

A DAM is needed to manage the corpus of words, the sentences and the recordings from the crowd sourcing tool

#### Crowd sourcing

A crowd sourcing tool is needed for WP3 needs

### Voices tools

#### IVR

An Interactive Voice Responder function must be available for each services.

The IVR function can be piloted from several ways: VoiceXML, PHP Scripts or Asterisk language. The recommended one is to use the VoiceXML.



## **ASR**

An Automatic Speech Recognition function shall be available for services hosted by the platform. The languages packages shall be updated during the project.

The openness recommendation applies here for both the ASR engine and its compliance with open integration of new language. For this purpose, de-facto open standard such as HTK acoustic models shall be used.

## **TTS**

A Text To Speech function shall be available for services hosted by the platform. This TTS must propose different languages to fit the use cases needs. This function shall be compliant with Speech Synthesis Markup Language (SSML) and Pronunciation Lexicon Specification (PLS)

## **Developers needs**

Development tools such as versioning tools, logs or sandboxes shall be available for all developers.

## **Administration needs**

The platform must contains all tools needed to administrate, supervise, and manage developers needs.

## **Mobile Applications**

The Use-case 4.1 needs a mobile application to be installed on the mobile.

## **Platform capacities**

The platform shall have the following capacities:

- ^ 1 local phone number per use case
- ^ 1 SMS short code per country (with as many keywords as use case)
- ^ SIP lines (for development purpose)
- ^ Manage concurrent calls (> 5)
- ^ Telecom capacities may be limited by quotas (to avoid unwanted usage, not to limit the normal usage of the service)
- ^ Web hosting capacities may not be limited (in regard to a fair usage)

## TECHNICAL REQUIREMENTS LIST

Multi channel accessibility	Analogic and mobile calls - Senegal	R.1.1
	SMS - Senegal	R.1.2
	Internet - Senegal	R.1.3
	Analogic and mobile calls - Mali	R.1.4
	SMS - Mali	R.1.5
	Internet - Mali	R.1.6
Service hosting	Service execution	R.2.1
	File hosting	R.2.2
	Database	R.2.3
	DAM	R.2.4
	Crowd sourcing tool	R.2.5
Voices tools	IVR	R.3.1
	ASR	R.3.2
	TTS	R.3.3
Developpers needs		R.4.1
Administration needs		R.5.1
Mobile Applications		R.6.1
Platform capacities		R.7.1

## HIGH LEVEL ARCHITECTURE OVERVIEW

The architecture of the VOICES platform should be like this:

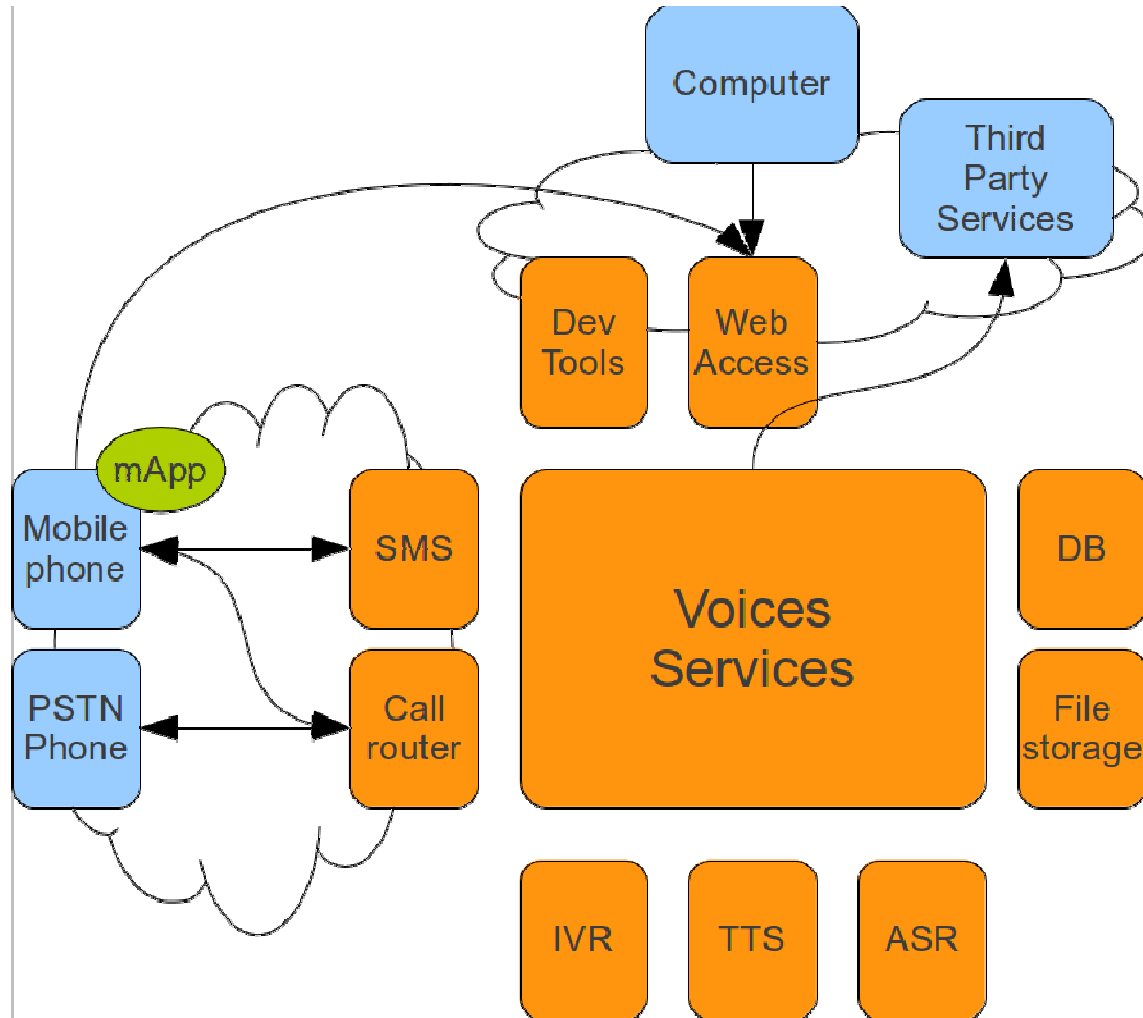


diagram 6: High level architecture

mApp = mobile application

## Conclusion

This report has described the high level architecture of the platform selected for the Voices project. The technical needs of each Voices use case have been described. Finally, this document has also detailed the technical requirements for the Voices platform.