

## CHAPTER 3

### RESEARCH METHODOLOGY

#### 3.1 Introduction

Research methodology is a disciplinary way which researchers need to conduct their research. It explained detail processes on how researchers address the issue and present their outcome in line with the objectives from the data obtained throughout the study. This chapter start off with research design approach, in general, which discusses the research design framework that will be used during the research process and the prior/post activities on the research method chosen with proper justification. Then next section will outline the detail processes which includes data collections, conventional and proposed methods for extraction and profiling stage. Last section of this chapter shows the overall proposed research frameworks are discussed in details and the types of research methods suitable for this research area are explained. It also outlined how the outcome of the research will be obtained while the objective of the study is met.

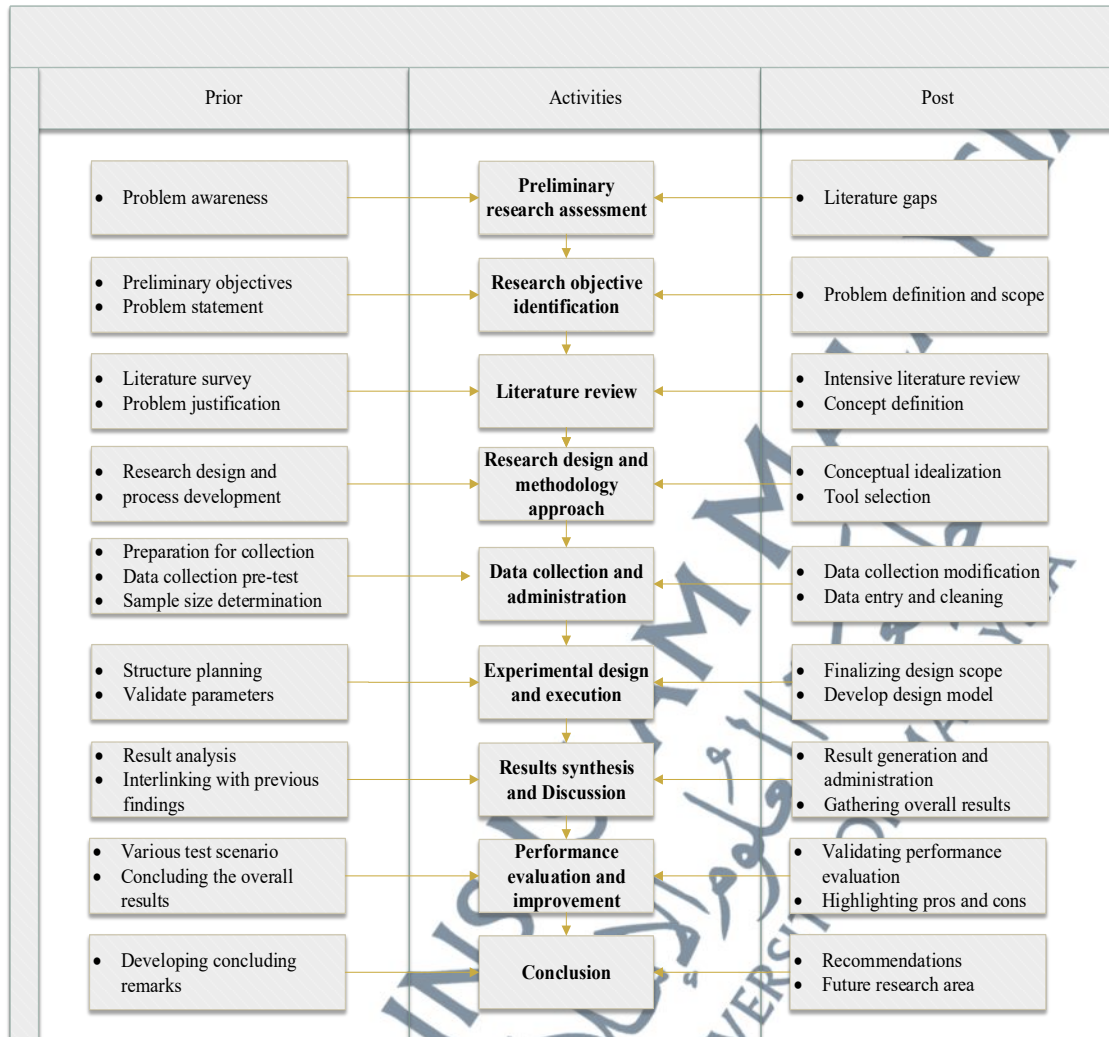
#### 3.2 Research Design Framework

The research design presented here is intended to provide an appropriate framework throughout the study. In general, an exploratory research approach will determine how relevant the information for a study will be obtained. Considering this study is exploratory research using complex speech processing method for profiling system, a wise decision in research design process is crucial based on the choice of methodologies to be made regarding research approach. To address the key research objectives, a preliminary research assessment and problem statement should be

initiated. Literature review should be done intensively prior the research design and methodology. Next stage is the data collection method of two datasets that must be cleaned and administered properly. The detail specification and sampling requirement will be described further in next section. Experimental design and execution are performed to obtain the intended result for synthesis and analysis prior the performance evaluation and modification of the system. The research framework is ended with conclusion and recommendation of the research area for future works. The fundamental research design framework and flow process, in general, are depicted in Figure 3.1. This study employed a mixed type of methods in spectral analysis for audio profiling system. Next section will describe further on the methodology and detailed framework applied throughout the study.

### **3.3 Research Method**

The methodology provides investigations in sequence that build information from one level to the next. In this study, the proposed profiling system known as Maqamat Profiling System (MPS) system should follow a specific methodology to be most effective. This methodology commonly uses a bottom-up approach. It starts at the most fundamental level of data and proceed to a higher level of structure over the data. This is to allow problems at lower levels are found and at the higher level it can be factored into the analysis. If a top-down approach is used, data inaccuracies at the lower levels may confuse the process and make it difficult to establish the true data rules. To make the higher levels of data profiling more successful these issues need to be corrected at each level before moving to a higher level.



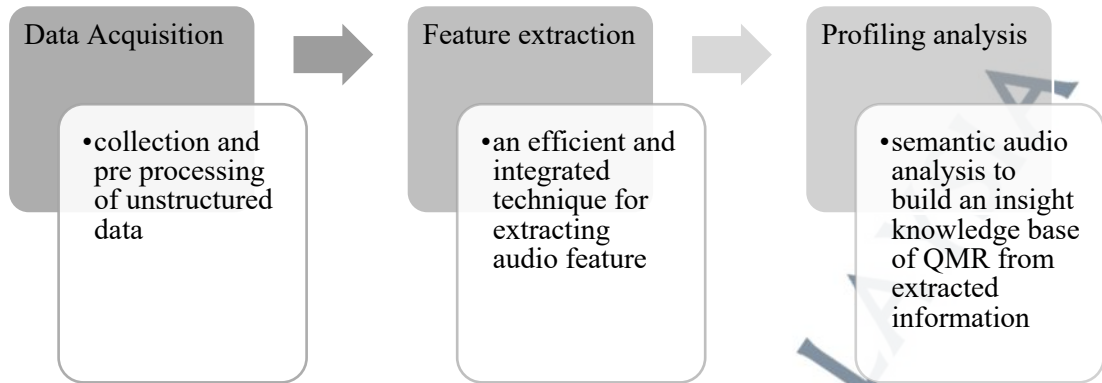
**Figure 3.1** The fundamental research design and flow process

Hence, the aim is to design a framework that consist of three main stages to improve the performance of the profiling system. Figure 3.2 shows the overall process framework presented and graphically displayed in this thesis. The overall framework consists of 3 stages as below:

1. Data Acquisition - Collection of raw data samples consisting of 242 audio recordings of QMR which grouped into to datasets, primary and secondary, as described further in section 3.4.

2. Ontological-based Audio Feature Extraction - Identify an adaptive and efficient approach for extracted features based on ontology that suitable for the development of MPS. As such, the methods need to be robust and efficient in term of distinguishing between the desired features and background noise or scene. Suggest an efficient enhanced and integrated technique that allow the extraction process shortened in term of time and high in accuracy. First part of feature extraction process which involved two existing AFE technique namely, conventional MFCC and featured enhancement approaches with warping DFT, as comparison. Three (3) spectral descriptors will be also utilized to detect the features prior developing the constructed knowledgebase in profiling stage. The mathematical model and algorithms have been described in section 2.3 and more details will be discussed in chapter 4 for the detailed process and chapter 5 for evaluation.
3. Profiling analysis - Find a robust technique for acoustical based profiling system using proposed MPS prior constructing an insight knowledge base of QMR from extracted information using semantic audio analysis which discussed detail in Chapter 6.

The software tools that will be utilized is MATLAB which requires additional toolbox namely DSP System Toolbox and Signal Processing Toolbox to run signal processing analysis. Figure 3.2 shows a diagram describing the work process of overall proposed framework for MPS development. Next section will cover the brief processes for each stage involved in the proposed framework.



**Figure 3.2** Flow diagram describing the work process of proposed framework for MPS development.

### 3.4 Data collection

As described in previous chapter, seven (7) maqamats are selected based on their audio signal in which contributing significant features and hence lead to optimum output of extracted features. These include *Maqamat Bayyati*, *Hijaz Rast*, *Jiharkah*, *Nahawand*, *Soba*, and *Sikah*. These maqamat also have been selected since they were mostly used among public reciters and experts.

There are 32 sets of pre-recorded audios which selected from 24 surah. Most of the surahs only includes certain verses, according to the selected theme. The *maqamat* recordings are based on selected surah in *Juzzu' Amma* as described further in later section. Both primary and secondary datasets are recorded in a recording studio in Faculty of Leadership and Management, Universiti Sains Islam Malaysia, Nilai, Negeri Sembilan, Malaysia. The audio is sampled with standard frequency sampling of 44100Hz for optimum audio quality. The datasets for instance, should be small to ensure the compatibility with the developed system.

Denote that the recordings should include all the designated requirement for QMR to be optimum state. These includes pitch, theme-content, and song features. To have all these features, the recordings should be captured with the duration of 20s – 30s per recording. As mentioned earlier, the data collections are divided into two datasets; primary and secondary, which based on the functionality of the sources. Both primary and secondary datasets will be used for development and validations during performance evaluation stage. All these pre-recorded audio files are considered as raw and unstructured data that has not yet been organized in knowledgebase for educational purposes. The next sections elaborate how the data were grouped and functionality of the datasets.

### 3.4.1 Primary Datasets

The primary data sources are selected based on the elements contained surah. Each of the verses is recited with all *tabaqat* types namely *qarar*, *nawa*, *jawab* and *jawabul jawab*. As described earlier in previous chapter, *tabaqat* represent the level of pitch in Arabic maqam. These will give the optimum energy for the *maqamat* to be recited. The primary data contained all *tabaqat* types due to its functionality for the purposes of training model in system development. All of the selected surahs is listed with recording duration as tabulated in Table 3.1 below.

**Table 3.1:** Primary datasets from selected surahs

No.	Surah	Tabaqat	Verse	Theme
1.	Al-Asr	All types	1-3	Greeting, command and assertiveness
2.	Al-Ikhlās	All types	1-4	Command and assertiveness
3.	Al-Falaq	All types	1-5	Command and assertiveness
4.	Al-Nas	All types	1-6	Narration

### 3.4.2 Secondary Datasets

The selection of secondary datasets is based similar criteria in primary dataset. This is to verify the MPS system development and validate the results. Table 3.2 show the secondary datasets form selected surahs.

**Table 3.2** Secondary datasets from selected surahs

No.	SURAH	Tabaqat	Verse	Theme
1.	Al-Naba'	All types	1-2	Warning, assertiveness
2.	Al-Naba'	All types	18-20	Assertiveness
3.	Al-Naba'	All types	38	Assertiveness
4.	Al-Naba'	All types	40	Assertiveness
5.	Al-Nazi'at	All types	6-8	Assertiveness
6.	Al-Nazi'at	All types	10-11	Assertiveness
7.	Al-Nazi'at	All types	12-15	Assertiveness
8.	Al-Nazi'at	All types	16-19	Assertiveness
9.	Al-Nazi'at	All types	21-24	Assertiveness
10.	Abasa	All types	11-12	Assertiveness
11.	Al-takwir	All types	1-6	Assertiveness
12.	Al-takwir	All types	25-27	Assertiveness
13.	Al-Infitar	All types	1-4	Assertiveness
14.	Al-Mutaffifin	All types	10-13	Assertiveness
16.	Al-Insyiqaq	All types	16-22	Assertiveness
17.	Al-Buruj	All types	17-19	Assertiveness
18.	Al-Tariq	All types	11-14	Assertiveness
19.	Al-Fajr	All types	15-17	Assertiveness
20.	Al-Balad	All types	17	Assertiveness
21.	Al-Shams	All types	13-14	Assertiveness
22.	Al-Layl	All types	14-16	Assertiveness
23.	Al-Duha	All types	9-11	Assertiveness
24.	Al-Alaq	All types	1-5	Assertiveness
25.	Al-Adiyat	All types	1-5	Assertiveness
26.	Al-Qari'ah	All types	1-5	Assertiveness
27.	Al-Fil	All types	3-5	Assertiveness
28.	Al-Ma'un	All types	1-7	Assertiveness

### 3.5 Audio Feature Extraction

In digital signal processing, the existing techniques for speech recognition and classification is highly dependable on the audio feature extraction results. Many extraction techniques such as MFCC, Linear Prediction Coefficients, Linear Prediction Cepstral Coefficients, Line Spectral Frequencies, Discrete Wavelet Transform and Perceptual Linear Prediction were discussed and evaluated (Sabur et.al, 2018). As mentioned earlier, the need for an enhanced algorithm is crucial to develop an efficient profiling system for feature extraction technique. In this study, MFCC technique is selected to be enhanced with warping function for complex audio signal. Further details will be elaborated in the next section.

#### 3.5.1 Spectral Descriptors

SD are widely used in machine and deep learning applications, and perceptual analysis. The spectral centroid represents the "center of gravity" of the spectrum. As been described in section 2.3.5, SD defines the equations used to determine the spectral features, cites common uses of each feature, and provides examples to gain intuition about what the spectral descriptors are describing. The QMR signal will be applied to the ontological based of SD wherein a certain number of spectral descriptors that will be extracted per each input frame. More elaborations and explanations on the works will be presented in the next chapter.

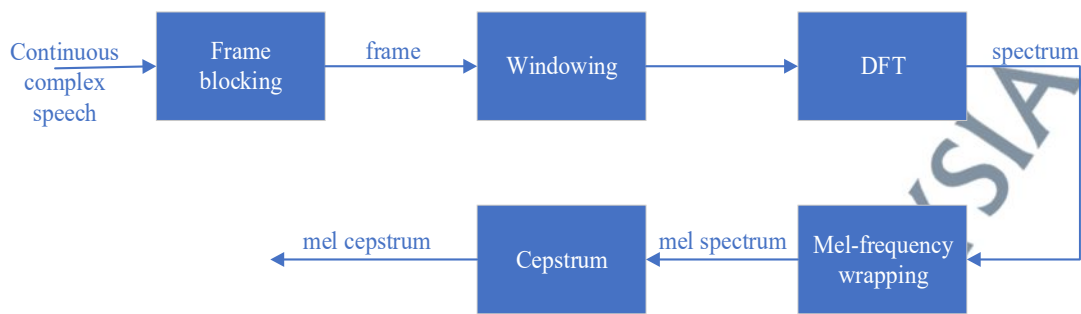
#### 3.5.2 Cepstral Coefficient analysis (CCA)

The objective of these analysis is to separate the speech into its source and system components without any a priori knowledge about source and/or system. As explain in section 2.3.2, the multiplication of the excitation and system component need

to be deconvolved to vocal tract components in the time domain (Giacobello et al., 2010). For this purpose, cepstral analysis is used for transforming the multiplied source and system components in the frequency domain to linear combination of the two components in the cepstral domain. The result of CCA will be used for comparison and performance evaluation.

### 3.5.3 MFCC Analysis

The conventional MFCC involved various step which includes windowing and framing the signal, applying the DFT, taking the log magnitude spectrum on a Mel-scale filterbank, followed by applying the inverse DCT as shown in Figure 3.3. In this study, MFCC is developed with frequency warping function as an adaptive feature extraction technique for complex speech profiling system. As mentioned in section 2.4.1, the enhancement of MFCC with frequency warping is design using HMM model. The first step is to compute a measure matrix for HMM dissimilarity measurement. Once the measured matrix is computed the clustering technique is applied to obtain the K clusters. The algorithm produces a sequence of clustering of decreasing number of clusters at each step. The clustering produced at each step results from the previous one by merging two clusters into one. Finally, isolating the sequences into K groups we fit each HMM to each cluster using all the observation sequences in cluster for HMM training.



**Figure 3.3** Conventional MFCC using HMM model

MFCC among the well-known method for recognition. As explained in section 2.3.3, the aim is to mimics the frequency response of the human hearing, so the coefficients are depend on filterbank energies from mel-frequency spacing. According to steps from HMM model, the initial step is to windowed using is Hamming window in overlapping steps. In these experiments, each window is set to time frame value,  $T_w=25$  ms with overlapping steps for shift frame,  $T_s = 10$ ms so this gives a total of 100 windows. The window width and overlap can be defined to optimize the visualizations. For each window, the log power spectrum is computed using a discrete Fourier transform (DFT). According to Reynold, the log spectral coefficients are perceptually weighted by a non-linear map of the frequency scale which known as Mel-scale. This is to emphasizes the mid-frequency bands that proportionate to their perceptual importance. The process is further transforming the Mel-weighted spectrum to coefficients value. Only the first 12 order of coefficients are selected while remaining higher order are discarded. This will transform the log energy into 13- dimensional feature vectors (12 MFCCs plus energy) at a 100 Hz rate.

### 3.5.4 W-DFT via MFCC

Generally, this method is performed by enhancing the conventional MFCC method. In this experiment, an alternative frequency warping strategy by replacing mel filterbank with direct warping function with uniformly warped DFT (W-DFT). The experiment is started again from the Hamming windowed frame but apply a warped DFT (WDFT) matrix to obtain a warped spectrum:

$$\tilde{\mathbf{X}}[k] = \tilde{\mathbf{F}}x[k] \quad (3.1)$$

where  $\tilde{\mathbf{X}}[k] \in \mathbb{C}^{N'}$  are the complex-valued warped spectrum and  $\tilde{\mathbf{F}} \in \mathbb{C}^{N' \times N'}$  are the W-DFT matrix. Its elements are given by:

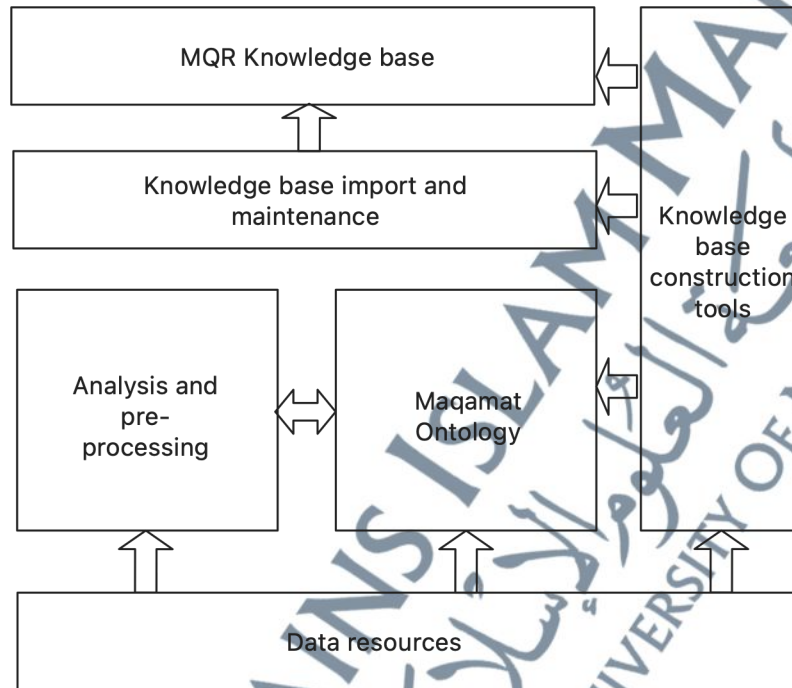
$$F_{k,n} = e^{-j\tilde{\omega}_k n} \quad (3.2)$$

where  $\tilde{\omega}_k = \frac{2\pi}{N'} \tilde{k}$ . The uniform value of  $\tilde{k}$  is replacing the mel-scale. The rest of the warped process undergoes exactly the same steps as conventional MFCC namely filterbank integration, log compression and DCT. Since the spectrum is already pre-warped, the triangular filters are now uniformly spaced in Hertz, rather than mel scale.

### 3.6 QMR Knowledgebase construction

This section provides the finale part of profiling stages. These involves in attribute tagging for rule matching based on audio features from AFE. The overall process of constructing QMR knowledge base is depicted in Figure 3.4. The data resources is performed pre-processing and analyzed using knowledge base construction tools and maqamat ontology. The ontological features are based on the peak detection of extracted

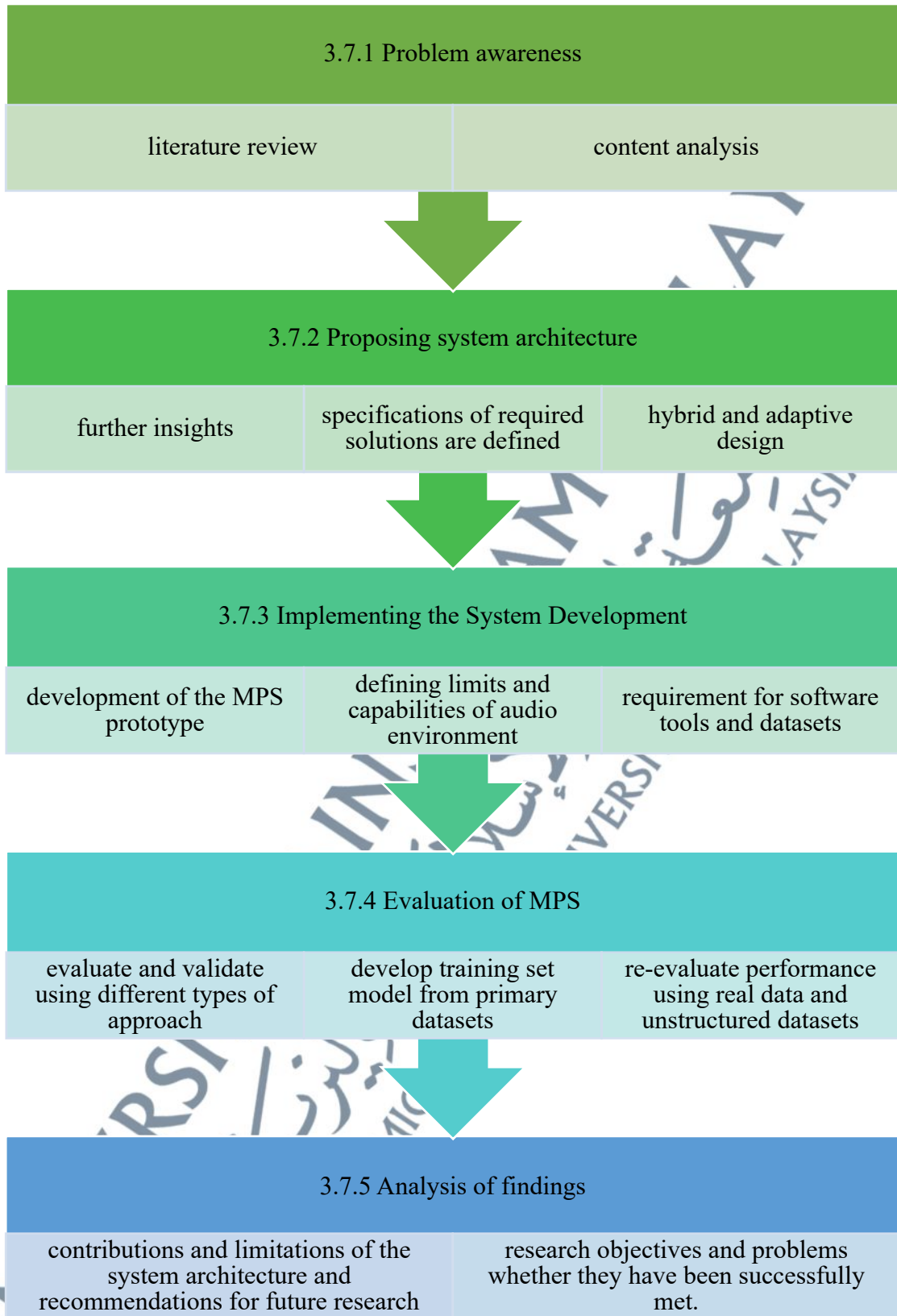
signals and signal matching. The construction of the knowledgebase is developed based on semantical features of the proposed ontology. The structured knowledgebase of MPS will be automatically generated based on the ontological features extracted in section 3.5.



**Figure 3.4** QMR Knowledgebase Construction Model

### 3.7 Overall research framework with outcome

The overall research framework can offer significant utilities for the research initiatives on QMR because it clarifies the attributes of the research environments, outlines the significant features and supports from the public reciters level to the QMR experts' level. Based on the fundamental research design in section 3.2, here are five (5) stages of the overall research framework implying the proposed MPS method together with the expected outcomes as shown in Figure 3.5.



**Figure 3.5** Overall research framework

### **3.7.1 Problem awareness**

The first stage of this method is literature review and content analysis. This is to understand and outline the problem statement, prototype's requirement, objectives and scope of the study focusing on QMR. During the cycles, several characteristics and engineering concept were furnished and combined with key insights to produce well planned direction for development. Through intensive literature review, more design tools of DSP and concept for speech processing such as AFE, spectral analysis and so on, are studied and compared.

### **3.7.2 Proposing the system architecture of MPS**

At this second stage, further insights into problem solving are required during initial analysis and design which aim to explore and evaluate potential solutions. Furthermore, the specifications of required solutions are defined. A hybrid and adaptive design combining two or more algorithms and computational techniques to solve complex and challenging problems. In this case, a complex speech processing for QMR issues. The proposed system intends to address drawback of QMR by integrating AFE, warping function and correlation techniques. Detailed descriptions on the algorithms of the proposed system are provided in Chapter 4 accordingly.

### **3.7.3 Implementing the System Development**

The third stage involves the development of the MPS prototype. This section describes the development toolkit (software and hardware resources) that will be used to develop and implement the proposed architecture in 3.4. The limits and capabilities of a sound profiling environment should be recognized. The major challenges that hinder sound profiling is the limitation and reliability of the captured data. Each

recording has different sound features depending on the hardware tools, reciters and sound features of the extracted data. The requirement for software tools and datasets has been elaborated in previous section 3.3. It starts with execution of the hybrid AFE with adaptive warping function. The overall execution of MPS system took place prior to the knowledge base construction. Then look at the sorts of global properties for having a robust system and discuss variations on the architecture which make validation of system more practical. Details on the development of the proposed system is described in Chapter 4 and 5.

#### **3.7.4 Evaluation of MPS**

This stage aims to prove that the implementation system satisfies the initial requirements. The evaluation is performed partly within submodule then followed by the overall system. To evaluate and validate the research outcome, different types of approach have been proposed such as mathematical proofs, logical reasoning, benchmarking, simulation, metrics, experimentation and demonstration (Vaishnavi and Kuechler, 2007). In this study, the evaluation methods used to validate MPS system is by developing the training set model from primary datasets using Classification Learner Toolbox from MATLAB. The model will be performed using further triangulation evaluation method based on unstructured secondary datasets and current real-world datasets. This method is used for both AFE and profiling performance evaluation. This is due to limited database exist in real world for all maqamat with studio quality standard.

### **3.7.5 Analysis of findings**

This section summarizes the work conducted in this research, the contributions and limitations of the system architecture and recommendations for future research. This section also determines the research objectives and problems whether they have been successfully met. The details of this section will be discussed in Chapter 7.

### **3.8 Summary**

Research methodology is the path through which researchers need to conduct their research. This chapter describes the research design framework that will be used during the research process and the prior/post activities on the research method chosen with proper justification. The research methodology includes data collections, conventional and proposed method, and profiling stage. It also shows the overall proposed research frameworks are discussed in details and the types of research methods suitable for this research area are explained.