**Abstract of Individual Research Component.**

# **Categorizing the text book content according to the topic/ chapter and storing in the database**

[A.M.M Mafaris IT13001162]

The objective of this component is to store the lecture contents content into the database according to the chapters. After choosing a subject to store in the database, have to identify the domain of ontology. There are two rules to follow this. That’s are Explicit mention rule and Implicit lexicon match rule. Then information will be storing into the ontology database. When storing the lecture contents to the ontology we have to write ontologies to get the stored content accordingly. These ontologies can be static because the lecture contents are not going to be changed any more. The final outcome of this component is a database which contains the lecture contents separated with chapters.

# **Filtering and identifying the content where the questions can be generated.**

[A.S.M Nibras IT14058424]

The objective of this sub component, is, extract elementary sentences from the complex sentences with the intention to generate more accurate questions. Initially information will be retrieved from the ontology. Information is retried as a paragraph. This will be tokenizing as sentences. Then tokenized sentences will be in the process of Sentence Simplification, where complex sentences will be passed to the BLLIP parser also known as the Charniak-Johnson parser or Brown Reranking Parser to construct a syntactic tree representation from the bracketed representation of the parsed sentence. We use the depth first algorithm, to read the tree nodes and leaves, which help us construct the elementary sentences, were we maintain if the phrases to be joined are sequentially correct with the respect of the sentence syntactical structure. At the end of this process elementary or simple sentences will be taken out from the complex paragraphs.

# **Identifying key phrases and words that should be used to generate appropriate questions.**

[I.S.M Arham IT14121852]

This is the section where the key phrases and words from elementary sentences will be identified to generate the questions such as noun phrase, verb phrase, noun, verb, adjectives, preposition etc. For each elementary sentence by classifying them into two classifiers named as Fine Classifier and Coarse Classifier to make candidate fine and coarse classes, where fine classes are reduced to coarse classes determined by the class hierarchy such as “Human”, “Entity”, “Location”, “Time”, “Count” etc. using Part-of-Speech(POS) Tagging concept. The part of speech tagging is used to identify and differentiate each word from other words in the sentence. Then the identified key phrases and words with their fine and coarse classes will be store in database.

# **Forming proper questions using the identified key phrases and words from the contexts.**

[M.F.F Mohamed IT14033506]

The objective of this component is to form proper questions based on the subject, object, preposition and verb for each elementary sentences. The coarse classes will be checked according to the word-to-word interaction rules to produce the type of questions that can be generated while considering the verb tense. Based on the coarse classification, we consider the relationship between the words in the sentence. For example, if a sentence has the structure: “Human Verb Human”, it will be classified as “whom and who” question types. If it is followed by a preposition that represents time, then we add the “When” question type to its classification. The rule check will produce the type of questions that can be generated while considering the verb tense.