**Department of Computer Engineering**

**Vishwakarma Institute of Information Technology, Pune**

**Generating Images from Textual Description**

Husain Shaikh

Shubham Shende

Omkar Shelar

Omkar Waghmare

Project Guide – Prof. Mrs.Tiwaskar

**Table of contents**

1. **Requirement Analysis**
   1. Problem Statement
   2. Who will use this proposed system?
   3. Benefits
2. **Functional Specifications**
   1. System Overview
   2. Scenarios
   3. Details of scenarios
3. **External Interface Specification**
   1. User interfaces
   2. Hardware interfaces
   3. Database
4. **Technical Specifications**
   1. Programming languages
   2. Versions of different components
   3. IDE to be used
   4. Performance Constraints
5. **Requirement Analysis**
   1. **Problem Statement**

To generate an image given a textual description.

* 1. **Who will use the proposed system**

Any person who wants a visualisation of a given detailed text description.

* 1. **Benefits**
* The system would be able to generate images that do not actually exist.
* Abstract scenarios can be visualised.
* The creativity of humans can be seen iteratively through the program.

1. **Functional Specifications**
   1. **System Overview**

This system is going to be a web based application wherein a dyslexic student will enter a virtual world. This world is where the student can choose from multiple courses related to core fundamentals, take quizzes and tests. It will also provide a interactive visual learning in the form of 3D object generations as well as human avatar interaction.

* 1. **Scenarios**

The user would input a text string describing a scenery, this statement would be of the type”It’s Summer, the sun is shining through the leaves in the forest”. The system would then be able to generate an image which would match the description. So as to offer variations, a set of 16 images will be created for each text description. Each being unique when compared to each other

.

* 1. **Details of scenarios**

Consider another query, ”A desert with a palm tree ”. The problem that could occur is if the dataset that we have access to, does not have instances of “Palm Trees “to learn from.

Here the image generated would not be accurate as the system would not know what a “Palm Tree “is.

**3. External Interface Specification**

* 1. **User interfaces**

The User Interface will be an online environment, wherein the user would input his desired textual query. The system would parse through the query, and generate viable images for it.

* 1. **Database**

We would be storing metadata of images within our database. This would contain the image URL,IDs and labels. For each query the images matching the keywords will be retrieved and stored within the database..

**4.Technical Specifications**

* 1. **Programming languages**
* Backend – Python
* Database - MySQL
  1. **Versions of different components**
* Python – 3.7
* MySQL – 5.7
* Pandas v 0.23.4
* NumPy 1.15.1
* SciPy 1.1.0
  1. **IDE to be used**
* PyCharm
  1. **Performance Constraints**

The major constraint is the computing power required to generate the images. As each label requires approximately 200-300 images to learn how to generate a desirable image.

The next constraint is on how the keywords will be selected, The more keywords we accept ,will increase the possibilities to generate images but significantly increase the processing power required. Limited query processing is present in the application as the objects will be only scenery related instances. The query will not provide a suitable response if it contains words which are not present in the dictionary of the keywords.