**Chapter 3** **METHODOLOGY**

This chapter presents the method of research used, sampling technique, instrument, data gathering procedure, statistical treatment of the data and the system architecture of the software.

**Research Design**

The researcher used the descriptive survey method of research which they applied in this study. The purpose of this study is to know the percentage or the total assessment of the PRC services, the evaluation of the benefits of continuing education, and the level of acceptance of the proposed system.

Calmorin (2002) explains descriptive method as one that concentrates on the present condition where the purpose was to find a new truth. Descriptive studies are valuable in providing facts in which specific judgments may be based. This study is primarily concerned with on the assessment of the professional in current benefits of continuing education, and about the current services provided by the PRC.

The researcher gathers the data using the research instruments to acquire the necessary numerical data and statistics that is based on the respondents’ nature of answers.

**Sources of Data**

The population used in this study comes from the selected different universities around Metro Manila area. The population selected are licensed professionals that are regulated by the PRC.

The researcher used purposive sampling to accumulate the data needed for the said study. The researcher selects the professionals that are licensed and those who are continuing their education to answer the given questionnaire. The researcher also used the datasets that contains the list of seminars obtained from the PRC.

**Research Instrument**

The researcher used the questionnaire as the primary instrument in conducting the study which includes the different certain questions. The questionnaire is designed to get the percentage or the total assessment of the PRC services, the evaluation of the benefits of continuing education, and the level of acceptance of the proposed system.

The questionnaire answers the issues and questions of the statement of the problem number 2 and 3, and by computing the result, the formulated result answers the statement of the problem number 1.

The researcher used the following software: (a) xampp, using mysql for databasing and phpMyAdmin for local-hosting in times of development; (b) sublime text editor for programming the web system; and (c) android studio for android application development. The researcher used some computer hardware with these following specifications: (a) laptops with Intel i5 3rd generation to 7th gen processors; (b) equipped with 4gb ram; (c) an android device for testing the android application.

The researcher used the Likert Scale or Point Scale to determine the respondent’s response in a particular statement.

### Table 1

**Likert Scale for Level of Agreement**

|  |  |
| --- | --- |
| **Numeric Value** | **Response** |
| 5 | Strongly Agree |
| 4 | Agree |
| 3 | Partially Agree |
| 2 | Disagree |
| 1 | Strongly Agree |

### Table 2

**Corresponding Remark for Likert’s Scale**

|  |  |
| --- | --- |
| **Score** | **Corresponding Remark** |
| 4.01-5.00 | Excellent |
| 3.01-4.00 | Good |
| 2.01-3.00 | Satisfactory |
| 1.01-2.00 | Fair |
| 0.01-1.00 | Not Good |

**Data Generation Procedure**

The researcher accumulated the local and foreign related literatures and studies mostly through online web research.

The researcher asked the selected respondents for their availability and level of willingness in terms of answering the survey. The respondents come from different selected universities around Metro Manila area, and the researcher has two types of approach, through giving them a conventional paper questionnaire and by making them answer through online survey, by using google forms.

The respondents answered the questions that describes themselves as an individual through demographic profiling, their evaluation on their current experience in dealing with PRC’s services, about acquiring their CPD units, and about their current status as a professional here in the Philippines.

Also, the testing of the system was done online, by providing a hosted link of the Green Light prototype for web users, particularly the PRC administrator and the CPD providers, and the testing of the mobile application by the selected professionals is done using a smartphone.

**Respondents’ Profile**

The respondents’ professional profile and level of agreement in implementing an online CPD application system tool is determined by the evaluation of the following tables, namely: Table 3, 4 and 5.

### Table 3

**Length of time as a Professional Identification Cardholder**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Frequency** | **Percent** | **Rank** |
| < 1 year | 20 | 20% | 3 |
| 1-3 years | 36 | 36% | 1 |
| 4-6 years | 27 | 27% | 2 |
| 7-10 years | 12 | 12% | 4 |
| more than 10 years | 5 | 5% | 5 |
| Total | 100 | 100% |  |

Based on the accumulated data from the respondents, the highest computed percentage is 36, ranked as number one, which are professionals working around 1-3 years, while the lowest computed percentage is 5, ranked as number five, which are those professionals that are working more than years.

### Table 4

**In favor of upgrading knowledge and skills as a prerequisite to**

**license renewal**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Frequency** | **Percent** | **Rank** |
| Yes | 66 | 66% | 1 |
| No | 34 | 34% | 2 |
| Total | 100 | 100% |  |

Based on the accumulated data from the respondents, the highest computed percentage is 66, ranked as number one, which are professionals that are in favor, while the lowest computed percentage is 34, ranked as number two, which are those professionals that are not in favor.

### Table 5

**In favor of renewing their license online if given a chance**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Frequency** | **Percent** | **Rank** |
| Yes | 83 | 83% | 1 |
| No | 17 | 17% | 2 |
| Total | 100 | 100% |  |

Based on the accumulated data from the respondents, the highest computed percentage is 83, ranked as number one, which are those in favour of online renewing license, while the lowest computed percentage is 17, which are those who’s not in favour of online renewal licensing.

**Ethical Considerations**

The following was observed in gathering the data:

1. The researcher produced multiple copies of conventional paper questionnaires and by making online surveys through google forms, to distribute to the respondents.
2. In this research, the researcher did not force the respondents to answer the survey questionnaire, considering their willingness to participate in the research
3. The researcher keeps the respondents’ personal information confidential.
4. The researcher explains to the respondent the objectives and the details they need to know before answering the questionnaire.
5. The researcher did not harm or abuse the respondents both physically and psychologically, in conducting of the research.
6. The researcher was conducted the surveying in a calm and friendly manner to ease the tense and to establish rapport with the respondents.

## **Data Case Analysis**

The following statistical tools was used to give meaning to the data to be gathered:

1. **Frequency and Percentage** was utilized to determine the demographic profile of the respondents of the study.

Wherein: P = percentage / f = frequency

n = total number of the respondents

**F =**  Wherein: n= total number of responses

N= total number of respondents

1. **Weighted Mean** was used to determine the level of agreement of the respondents on the benefits of continuing education thru CPD law implementation. It was also used to get the mean average of the level of acceptance of system’s capabilities, in terms of functionality and design, in accordance to the respondents’ answers.

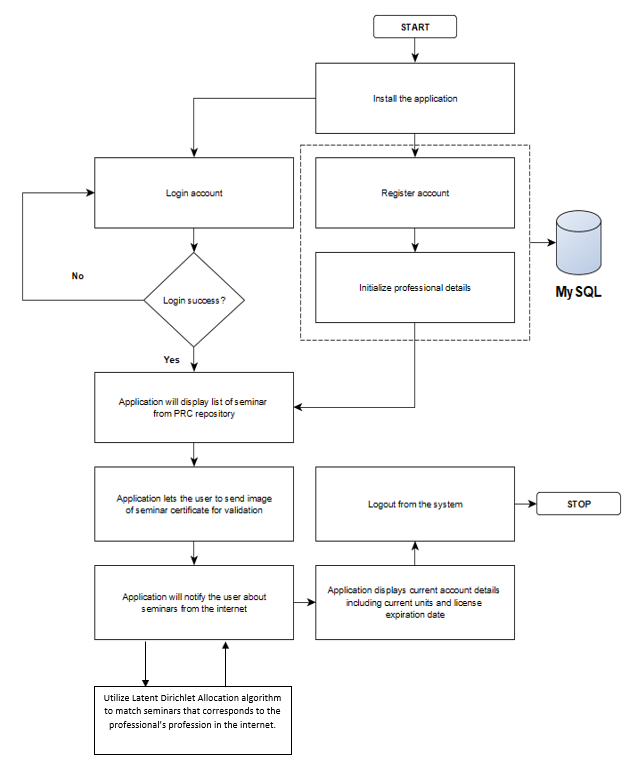
Wherein: f = frequency

x = corresponding rank of the verbal Interpretation

n = total number of the respondents

## **Process Flow Diagram**

Figure 4. Process Flow Diagram of the Developed Mobile Application



The professional will install the mobile application, after the installation user can either log in to an existing account or register for a new one. Upon registration, the system will require the user to initialize account details which includes the current units the professional's name, profession, current number of units and the starting date of the professional's license. After registration the user will automatically login the system and redirect to the home page.

At the home page, the application will display a list of seminars based on the user's profession which are stored in the repository, the user can now choose from these available seminars. The application also provides an interface where the user can upload an image of the certificate, he/she received from the seminar which will be saved in the database. Also, the application will notify the user on seminars available from the internet which may be national or international. The utilization of Latent Dirichlet Allocation is also linked in the process for matching the seminars that corresponds on the professional’s profession in the internet.

Lastly, the user can view his/her current accumulated points and the expiration date of the license to be able to track his/her progress and then the use can logout of the system which ends the process.

## **System Architecture**

Figure 5. Green Light System Architecture

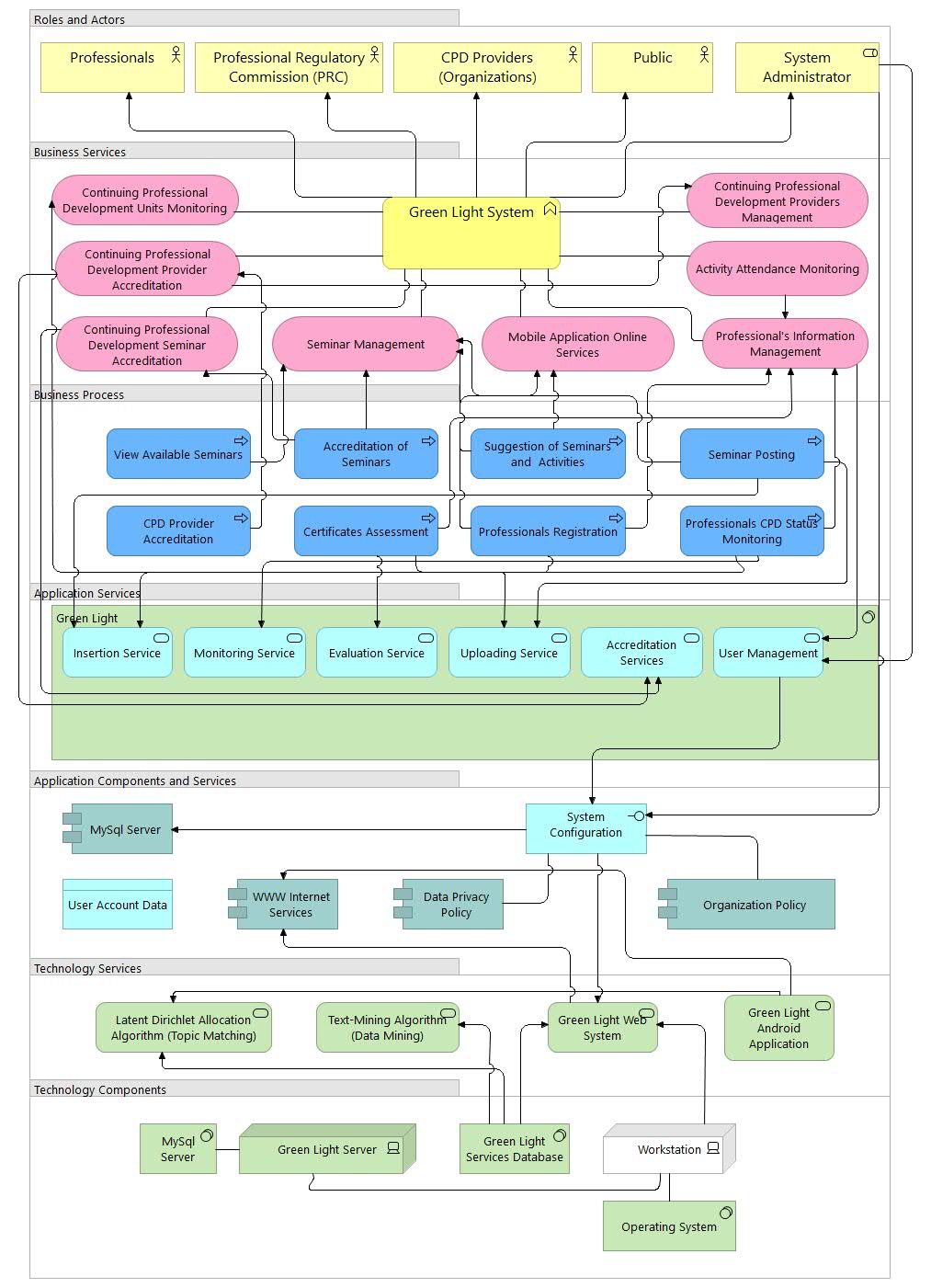


Figure 5 shows the components of the developed continuing professional development application system – Green Light. It has 7 classifications consist of the roles and actors, business services, business process, application services, application components and services, technology services and technology components. There are 4 roles and actors in the system, these are the professionals, professional regulation commission (PRC), the organizations, the public, and the system administrator.

In the business services, Green Light includes Professional’s Information Management, Continuing Professional Development (CPD) Seminar Management, CPD Providers Accreditation, Seminar Accreditation, Mobile Application Online Services, Activity Attendance Monitoring, CPD Unit Tracking and Professional’s Information Monitoring.

The business process of Green Light System Architecture contains 8 activities such as the viewing, posting, accreditation and suggestions of seminars, accreditation of CPD providers, certificate assessment, and professionals’ registration and CPD status tracking. For the Application Services, it has 6 modules named as Insertion Service, Uploading Service, Monitoring Services, Evaluation Services, Accreditation Service and the User Management.

The system used MySQL Server for storing the data of the Professionals, the www internet services for seminar mining and matching, and includes a System Configuration to update the Data Privacy Act and Professional Regulation Commission’s Policy.

The system used technology services such as mobile application, text-mining and topic-matching Algorithm and Web system. There are 4 technology components involved in the system, these are MySQL Server, Green light Server, Green light Services Database and the Workstation for the operating system.

## **Features and Complexity**

This section explains the main features and the producible reports of the system, together with its intelligence capability that will determine and decipher the overall complexity of the system architecture.

### Table 6

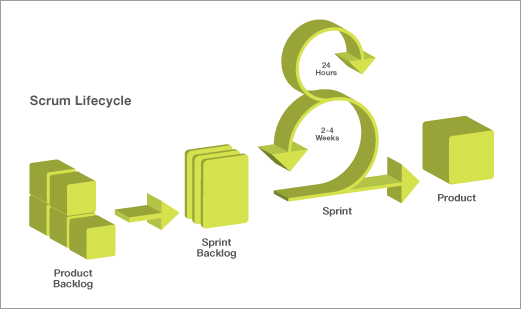
**Green Light Application System Summary of Technical Capabilities**

|  |  |  |
| --- | --- | --- |
| **Main Features** | **Reports** | **Intelligence** |
| 1. Administrator’s Dashboard - statistical analysis in that helps the administrator to evaluate the performance and outputs of the mobile application. 2. Light-Weight Design - the application installed on the mobile devices are literally light-weight, consuming less space and has an optimized application startup. The overall design will also be minimal, that will help the users to navigate more with ease. 3. Data Mining from big data - feature will enable the users to find all available CPD programs around the world, and with the use of the algorithm, they can find referrals of CPD programs that relates to their current profession, thus, finding CPD programs that will satisfy their required number of units will be hassle-free and less time-consuming. 4. Viewing of CPD Units - will let the users (professionals) to monitor their CPD units, so they will know when to apply for license application/renewal, or, if they still needed to gain more units, meaning they will utilize the data mining feature to find seminars that they are interested with, or related with, depending on their profession. | 1. List of Attendance - The attendance of seminars is recorded by the organization; it can be filtered and printed. 2. List of Seminars 3. Pending Seminars – list of pending seminars requested by the organizations; it can be filtered and printed. 4. Revision Seminars – list of for revision seminars required by the PRC; it can be filtered and printed. 5. Approved Seminars – list of approved seminars approved by the PRC that are available for public viewing; it can be filtered and printed. 6. List of CPD Providers - These are organizations accredited by the PRC that provides seminars with CPD units; it can be filtered and printed. 7. List of Professionals – These are the licensed professionals registered in the PRC; it can be filtered and printed. 8. Users of the System – these are the users of the system: the professionals, providers, and the PRC administration; it can be filtered and printed. | 1. Data Migration – the users of the systems can use the excel sheet import capability for fast and easy recording. The following can be uploaded through excel import: 2. The organization can upload a particular seminar attendance. 3. The admin can setup the providers records. 4. The admin can setup the current seminars available. 5. Data Mining – The users of the system can find seminars with the help of LDA algorithm. The following discusses the process of mining: 6. The local seminars are automatically recommended to a particular professional, based on his/her profession. 7. The capability of meta-data search helps the application to find the matching of seminars to a profession accurately and fast. 8. The use LDA algorithm in terms of text modeling comparison is effective in big-data mining from the internet. It serves as the search engine for non-registered seminars around the world. |

Table 6 discusses the technical capabilities of the system. The features are discussed based on the proposed features on the chapter 1, and the reports are formulated based on the features. The system’s intelligence is articulated from the features and reports it produces, thus, the researcher hypothesized that the system can provide quality service to its users, after analyzing the complexity of the system architecture.

## **System Development**

Figure 6. Agile Scrum Development Life Cycle

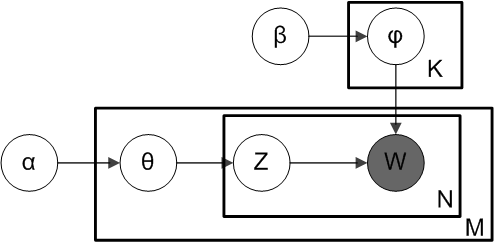


The figure 6 shows the agile scrum development life cycle which was used in the development of the proposed system. Agile scrum development is an iterative and incremental framework for managing software/product development. It is a full strategy where development team work as a unit to reach a common goal in a short period of time.

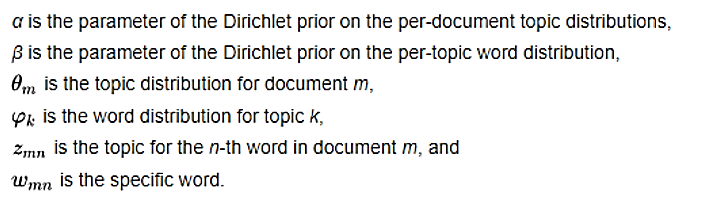
The researcher utilized the use of this software methodology to deliver a quality prototype with a minimum amount of time, equipped with a fast-paced development environment and quality assurance, that helped the researcher to use the prototype as an instrument for survey.

## **Simulation Strategy**

Figure 7. Plate Notation for LDA with Dirichlet distributed topic-word distributions



This plate notation is often used to represent probalistic graphical models, the dependencies among the many variables can be captured succinctly. The boxes are “plates” representing replicates, which are repeated entities. The outer plate represents documents, while the inner plate represents the repeated word positions in a given document, each of which position is associated with a choice of topic and word. M denotes the number of words in a document. The other variable names are defined as follows:

Figure 8. Latent Dirichlet Allocation Plate Notation Symbol Translation

Data was simulated by using topic modeling in text mining through Latent Dirichlet Allocation. Latent Dirichlet Allocation is best suited to categorized text in a document to a specific topic.

The algorithm was integrated and used in the system prototype to determine and simulate big data of different professions with different CPD programs and seminars offered by the PRC to get the require units to be able to renew the professional license.

# Table 7

**Topic Modeling of 44 Licensed Professions in to different seminars in the Philippines using 3 Different Algorithms in MATLAB**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Results of Prediction** | | **Processing Time** | **Accuracy** |
| **Latent Dirichlet Allocation** | D1  D2  D3  D4  D5  D6  D7 | 68 Items  50 Items  1 Item  1 Item  10 Items  99 Items  73 Items | 2s | 60.89% |
| **Latent Semantics Analysis** | D1  D2  D3  D4  D5  D6  D7 | 21 Items  73 Items  2 Items  3 Items  1 Item  1 Item  46 Items | 5s | 58.43% |
| **Explicit Semantics Analysis** | D1  D2  D3  D4  D5  D6  D7 | 1 Item  3 Items  5 Items  1 Item  1 Item  1 Item  30 Items | 4s | 56.21% |

The interpretation of the table 7 are as follows:

The Latent Dirichlet Allocation algorithm is tested in 7 documents, and for the result of the simulation, the result of prediction is totaled in 302 items, thus, the selection of suggestions is many. The result is shown after 2 seconds, and the measured accuracy is 60.89%.

The Latent Semantics Analysis algorithm is also tested in 7 documents, and for the result of the simulation, the result of prediction is totaled in 147 items, thus, the selection of suggestions is close to appropriate. The result is shown after 5 seconds, and the measured accuracy is 58.43%.

The Explicit Semantics Analysis algorithm is also tested in 7 documents, and for the result of the simulation, the result of prediction is totaled in 42 items, meaning, the number of results is not that many, thus, the selection of suggestions is not enough. The result is shown after 4 seconds, and the measured accuracy is 56.21%.