CHAPTER 2
THEORETICAL FOUNDATIONS

This chapter will contain detailed explanation on the theory used in composing this thesis. This chapter will begin with basic theories and will be continued by some theories that are related with the scopes of this project.

2.1. The Internet

The internet is a global network that connects all networks all over the world. This global network enables people from different places or areas to share the information. Unlike the other sharing information tools such as radio and television, the internet is not centralized distribution system based. This ability of the internet enables each computer to become a central of information sharing. Unlike radio and television that are based on one way communication system, the internet enables two way communication that well known as request and respond method.

The information that can be accessed using the internet normally consists of interconnected hypertext and other resources of the World Wide Web (WWW). Sharing information through the internet can be done through so many types of interfaces such as web browsers, chat client, electronic mail, and so many others.

The information that is transferred across the internet will be partitioned into smaller piece of data called packets. The receiving client computers will integrate these small pieces into the original document. Packet itself consists of two sections which are the header and the payload. The header will contain the control information which consists of the source and destination address, the error detection codes, and the sequencing information. The payload section will contain the information itself.

2.1.1. Web Pages

The World Wide Web (WWW) is conjured of information that is presented as “web pages” or HTML (hypertext mark-up language). World Wide Web is part of internet which use transport through the internet from client/server architecture to handle every type of digital information such as text, hypermedia, graphic, and audio [13]. These web pages can be displayed using the web browsers. These browsers recognize the HTML language that constructs the web pages and display the HTML code on the window of the browser in different way. HTML is used to define structure of a web page [21]. HTML is the most common text-based tagging language from creating documents and setting up hypertext links between documents on web [19].

The communication through the internet is based on “request and respond” method and it is applied in web pages system. The web pages are requested from the server and this server called as the

web server. The one who request for the web pages is the client. The server served the requested file web pages using the Hypertext Transfer Protocol (HTTP).

2.1.2. Hypertext Transfer Protocol (HTTP)

Hypertext Transfer Protocol (HTTP) is a protocol that is used for data transfer in the World Wide Web (WWW). This protocol has generic behavior so it can be used for any type of documents. This protocol is used for “request and respond” between the server and the client. When the client make a request, it will creates Transmission Control Protocol (TCP) connection to a specific port (default port:80). HTTP server listens on that port wait for the client to send a request message. Upon receiving the request, the server reports status of the line.

2.1.3. TCP/IP

Transmission Protocol/Internet Protocol (TCP/IP) is a standard for data communication that is used for the information exchange through the internet. Both TCP and IP cannot stand alone because they depend each other. TCP/IP describes the rules how the communication should be over the network. TCP/IP is usually illustrated as having four abstraction layers which are:

- **Link Layer**

  This layer is the lowest section layer of the internet protocols. The link layer is responsible for the packets movement between the internet layer interfaces of two different hosts which have the same link.

- **Internet Layer**

  This layer is responsible for routing and encapsulation of the network data packet into IP packets.

- **Transport Layer**

  The Transport Layer is responsible for providing the service connections between applications together through the use of ports.

- **Application Layer**

  This layer is responsible in giving access for the application to the TCP/IP.

2.1.4. Web Server

The term Web Server can be interpreted into two understandings: The computer program that accepts HTTP Requests from web clients and serves HTTP response or the computer that runs the

program [17]. Web server is a computer that acts as a server for a website. This server is located in the internet service provider (ISP) in order to connect to the internet in full 24 hours. Web server is a core in the internet organization which is complex today. Every time a browser is connected to a website, no matter it is on the internet or intranet, it means that the browser is connected with the web server. The server listen to the network for the request and give the respond to the client by carrying a type or types of data.

Web server is a computer that provides service to internet. This server is called as a host. To enable host a website that can is accessible through the internet, we have to assign addresses which is provided by the web server.

2.1.5. Web Browser

Web browser is software that is easy to use for showing a webpage, to accessing a web, and any other type of resources from the internet [14]. A browser recognizes graphic information, audio, and video. There are so many browsers exist in the world such as Internet Explorer, Mozilla Firefox, Opera, Safari, and so many others. In the development of this thesis, the team has selected Internet Explorer and Mozilla Firefox as the standard of the development process.

2.2.System Analysis and Design

2.2.1. The System Development Life Cycle

The systems development life cycle (SDLC) is the process of understanding how an information system (IS) can support business needs, designing the system, building it, and delivering it to users [1]. Because the SDLC describe the overall list of processes and sub-processes required, it is essential for establishing a system project plan. It helps in establishing a system project plan because it gives the overall list of processes and sub-processes required for developing a system. The phases in SDLC are consisting of project planning phase, analysis phase, design phase, implementation phase, and a support phase [18].

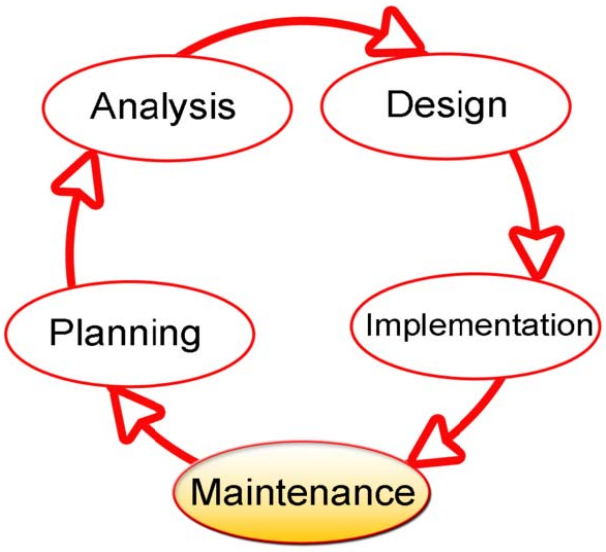

figure 1: System Development Life Cycle [2]

SDLC has four fundamental phases which are:

- **Planning Phase**

  The planning phase is a process of understanding why the current system should be built or the reason why the new system is needed. The project team will also determine the way the project executed.

- **Analysis Phase**

  During the analysis phase, the analyst determines where the problem is in order to fix the system. This phase requires the analyst to break down the structure of the system into smaller structure and draws diagrams to analyze the current situation.

- **Design Phase**

  This phase describe the way the system will operate including the hardware, software, and network infrastructure. It contains in detail the screen layouts, diagrams, and any other documentation.

- **Implementation Phase**

  The final step of the SDLC is the implementation phase. This phase will be executed after the design that

has been created during the design phase is accepted by the client of the project. This phase will get the most attention because this is the most expensive part among all the steps. This step will turn every theory into practice.

2.2.2. Use Case Analysis and Diagrams

Use case analysis is a technique that determines the requirements that are needed in a system. Use case describes how actors or users trigger an event which is needed to be responded by the system. Use case help the system developer to understand the situation and help make the design simpler.

When developing use cases, the developer should collaborate with the actors that will have interaction with the system to gather the information needed. This collaboration can be done by using requirement gathering technique such as interviews and observations.

Designing a system from the user's perspective and point of view is a primary goal of the use case analysis. Use case analysis is also finding out the way the system will be used, the roles of the actors in the system, and also the system response to actions done actors.

A use case diagram contains describe in general all activities that can be done by actors. They are useful for presentations to management and/or project stakeholders, but for actual development you will find that use cases provide significantly more value because they describe "the meat" of the actual requirements [22]. This activity determines what requirements should be included during the development of the system. There are 3 elements of the use case diagram which are:

- **Actors**

figure 2: the actors in use case diagram [3]

Actors are someone or something outside the system that interacts with the system by doing some sort of actions. The picture above, as shown on figure 2, is the way actors are drawn in a use case diagram. The actors will be linked to one or many use cases in a use case diagram means that they can do some several action to the system.

- **Use Case**

figure 3: the use cases in use case diagram [3]

Figure 3 shows us how use cases are drawn in a use case diagram. A use case is a single task that can be done by every actors that have connection or linked with it in a use case diagrams. Before declaring a use case, a developer should discover what actions are exist in the business function. The developers do not create the use case but it is more like they discover it through the business function.

- **System Boundary**

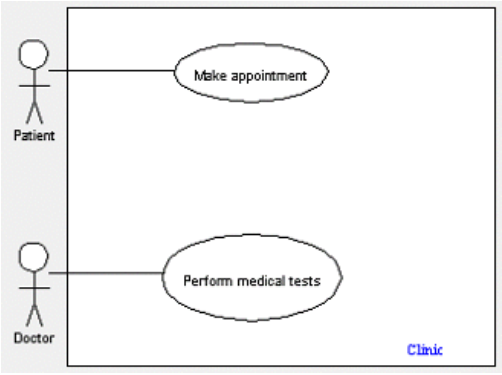

figure 4: the system boundary in use case diagram [3]

Figure 4 shows us a system with some functionality that can be done by several actors. The system is impossible to have infinite numbers of function so the system boundary of use cases determines the limitation of the system. This boundary is drawn in rectangle and contains use cases inside.

2.2.3. Data Flow Diagrams

Data flow diagrams (DFD) describe how the system converts input data into valuable information. It shows how data travels through the system but does not explain the program logic or processing steps. A set of DFDs provides a logical model that shows what he system does, not how it does it [20].

2.2.3.1. Elements of Data Flow Diagrams

Data flow diagram describes the flow of data in a system which is drawn in notations that helps the developer understand the logic of the system. Data flow diagrams enable the developer having clear visualization of the flow of the data in very detail. This requires the developer break down the system into smaller system to view specific data flow in the system. There are four symbols that represent the data flow diagrams which are:

- **The Process**

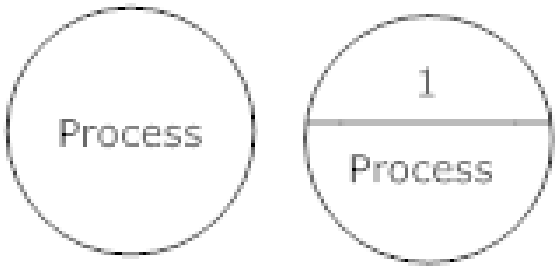

figure 5: the process in data flow diagrams [4]

Process is an activity which is performed for some particular business reason. It will transform the input data and produce an output data.

- **Data Store**

figure 6: the data store in data flow diagrams [4]

Data store is a group of data that is stored in an address. Sometimes, they are referred as files.

- **Data Flow**

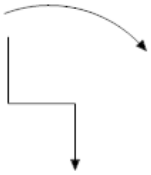

figure 7: the Data Flow in Data Flow Diagrams [4]

In data flow diagram, the data flow is the line that connects data store, process, and the external entity. Data flow itself is a single piece of data or a logical set of some pieces of information.

- **External Entity**

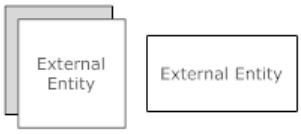

figure 8 the External Entity in Data Flow Diagrams [4]

As the explanation above, data flow diagram describes a single pieces of system from the big system so a big system may have several data flow diagrams of its smaller systems. The output of one data flow diagram may be use for another data flow diagram which means on

different smaller system. To draw this, there is external entity notation in data flow diagrams. External entity is a system or person that is outside the system but interacts with it.

2.2.3.2. Context Diagrams

The context diagrams describe the way the computer system interacts with the environment, mainly focus on the external entities. Context diagram explains the data flow from the external entities which describe the whole big system. This diagram is the highest level among all the data flow diagrams which only contain a process. This diagram does not have data store notations and looks very simple.

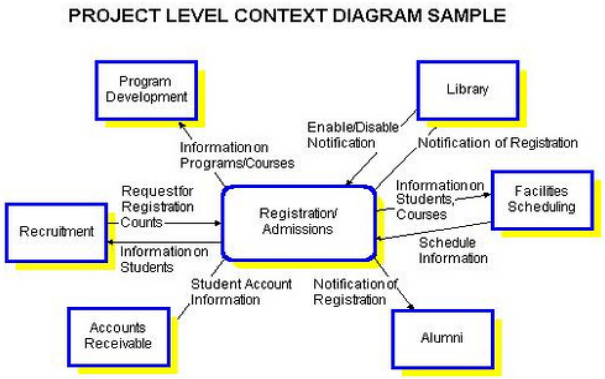

figure 9 Context Diagrams [5]

2.3.Distributed System

Distributed system is a collection of independent computers that appears to its users as a single coherent system [6]. The goals of the distributed systems are to making the resources accessible and to achieve the distribution transparency. To making the resources accessible means that the user of the system can access the resources on different computers or remote computers easily and they can share the information in a controlled and efficiency way. In this project, the practice of this ability of the distributed system is the communication between smaller systems [example: departments in an organization].

To achieve the distribution transparency does not mean that all the system is visible to users but it means that the system transparent to user as a single system although the reality that the system interacts with other systems. There are so many types of transparency which are:

- **Access Transparency**

  Access transparency means that the user of the system cannot identify the differences in representation of the data and the way the resources accessed. For an example that two websites that has different architecture, let say that the first website is using Oracle and the other one is using MySQL. This information in some ways should be hidden from the users and this is that called as access transparency.

- **Location Transparency**

  Location transparency means that the user of the system unable to locate where is the location of the accessed resources. For an example that the user of the ATM machine unable to know from which office's database the data they accessed.

- **Migration Transparency**

  Migration transparency means that the user is unable to know whether a resource moved to other location. Migration transparency means that the system has ability to move the resources without affect how the resources can be accessed.

- **Relocation Transparency**

  Relocation transparency is a stronger condition than the migration transparency. It is when the resources can be relocated when it is being accessed by the user without the user noticing anything.

- **Replication Transparency**

  Replication transparency is when several copies of the resources exist in the system without let the user knows it. To achieve this transparency, the system should named the resources with the same name. The system should support the location transparency before achieving the replication transparency.

- **Concurrency Transparency**

  Concurrency transparency means that the users are unable to know if the resources they use are shared. For example if there are two person are accessing a resource at the same time, both of them is unable to know that someone is also accessing that resource at that time.

- **Failure Transparency**

  Failure transparency is when there is failure in the system, the user can still executing the resources without noticing that there is failure in the system. The user also should not notice that the system is currently recovering from the failures.

- **Security Transparency**

  Security transparency means that the system can hide the security technique that protects the system from the user. This transparency should have the minimum level of user intervention in order to avoid the security failures.

2.4.Database Management System

Database management system consists of two elements which are the database and the application programs. Database is a collection of data which are interrelated. The applications used to access, update, manipulate, and manage that data. These applications commonly categorized as the

management system. The definition of the database management system itself is a set of applications that manipulates the data from the database. Database management system is implemented everywhere in every sector today such as banking, airlines, universities, telecommunication, and so many others. The goal of using database management system is to achieve the convenient and the efficiency in retrieving and storing data. Database management systems enable to remove some major disadvantages that occur in file system such as:

- **Data redundancy and inconsistency**

  Data redundancy means that the same data exist in or duplicated in several places. Redundancy will cause the inconsistency of the system.

- **Difficulty in accessing data**

  File system is unable to give data from a request which is not programmed in the system. Example if bank officers want to retrieve data from several customer which are live in a specific state, file system is impossible to give data which has this criteria but it can give all data and let the officers find it out manually and it means time consume.

- **Data isolation**

  Data might be in different files and formats so it is difficult to write new applications programs.

- **Integrity problems**

  Data needs to satisfy some constraints and it is difficult to use file processing approach to change constraints of the data.

- **Concurrent access anomalies [multiple users]**

  When there are multiple users access same data, it can be result with inaccurate data. Example: when there are 2 users withdraw money (say $50 and $100 respectively) from an account which has $1,000 in it, if the program runs concurrently, it can be result in $950 or $900. Both of them may read $1000 at the same time and the result is depending which data writes the last.

2.4.1. Data Abstraction

One purpose of the database system is to provide users with abstract view of the data. This means that the user unable to find out how the data stored and maintained. Considering that not all people is well computer trained, the developer of database management systems hide the complexity from the users through several levels of abstraction:

- **Physical level**

  Physical level defines the way the data stored. This level is the lowest level of abstraction and describes the complexity of low level data structure in detail.

- **Logical level**

  Logical level describes the type of data which is stored and its relationships. Normally, the logical level is where the database administrator is.

- **View level**

  View level is the highest level of the abstraction. This level defines the section or address in the database for particular users.

2.4.2. Entity Relationship Diagram (ERD)

Object-based data model use concepts as entities, attributes, and relationships. This model describes data at the view and conceptual level including with its constraints. One of well known object-based logical model concept is the entity relationship diagram (ERD). These are the elements that build the ERD:

- **Entity**

  Entity is an object which is the place where the data is stored.

- **Attributes**

  Attributes defines the characteristic and properties of an entity.

- **Relationships**

  Relationships show the way the entity shares their information in the database structure.

- **Links**

  Links connect the entities, attributes, and relationships.

There are 2 kinds of relationships in an entity relationship diagram

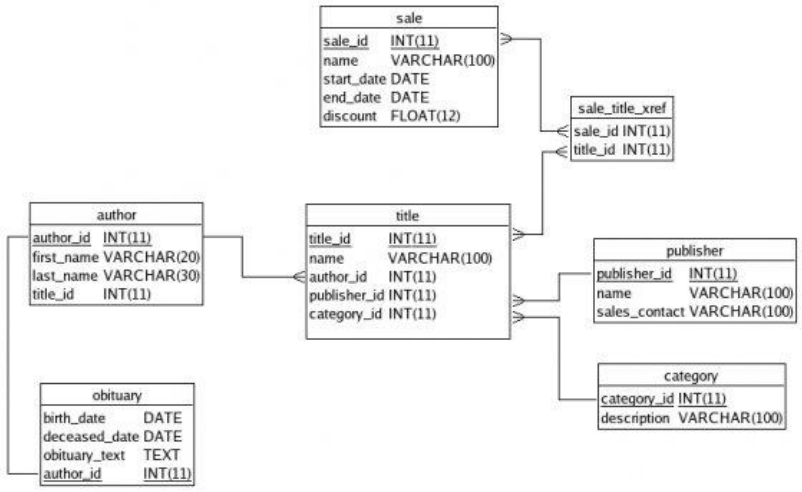

figure 10 simple ERD diagram [11]

2.5.Security

In information system, security is the condition where the information is well protected from any violation of the policy that is applied on particular information systems. The way to secure the information is only one which is to protect the information. There are 6 [9] security functions that are needed to be protected which are:

- **Identification and Authentication**

  This function protect from unauthorized user accessing the system. There are so many ways to do this function in a web system such as using secured login using password and terminal access.

- **Access Control**

  Access control defines the boundaries and limitations of the information that can be accessed by actors

- **Accountability**

  The traceability of actions performed on a system to a specific system entity (user, process, and device). For example, the use of unique user identification and authentication supports accountability; the use of shared user IDs and passwords destroys accountability [7].

- **Object Reuse**

  When an object is initially assigned, allocated, or reallocated to a subject, that object must not contain any data that the subject is not authorized to access [8].

- **Accuracy**

  Accuracy function guarantee that the information resources remain protected from corrupted data.

- **Reliability of Service**

  In order to achieve the reliability of service, the information should be available when the authorized users want to access it.

2.6.PHP

PHP is a server side scripting language which is designed especially for networks. In HTML pages, PHP script can be inserted and will be executed every time users visit the pages [12]. PHP code will be translated on the server and generate the view on the HTML pages. PHP is widely used for dynamic web programming but it is possible also used for static web programming. PHP have some advantages compared with other web programming language such as:

- PHP is a scripting language which is not do a compilation in the execution

- Web servers that support PHP scripting language can be easily found such as IIS and apache with easy configuration.
- PHP is a open source language which can be used in every type of operating system such as Linux, Unix, and Windows and can be executed in runtime through console including system command
- Large amount of database interfaces like MySQL, MS SQL, Oracle, Informix, and PostgreSQL etc. [15]

2.7.MySQL

MySQL is native database from PHP. As a database server, MySQL runs on the server side from an application client server. One of the greatest advantages from MySQL than other SQL based DBMS is no price and operation cost. MySQL is a freeware. The operation of MySQL is based on SQL statements such as SELECT, DELETE, CREATE, DROP, and so many others.

MySQL is a proven and cost effective database solution that lowers 90% of the cost for database software infrastructure. As stated by Sun Microsystems, Inc, These are the way MySQL cut down the cost of the operation [16]:

- Reducing database licensing cost by over 90%
- Cutting system downtime by 60%
- Lowering hardware expenditure by 70%

- Reducing administration, engineering, and support cost by up to 50%

2.8. Cellular Phone

Cellular telephone, sometimes called mobile telephone, is a type of short-wave analog or digital telecommunication in which a subscriber has a wireless connection from a mobile telephone to a relatively nearby transmitter [25]. Cellular phone enables the user to do some services that available in it such as SMS (Short Message Service), radio, television, audio player, videophone, even internet connection.

2.9.SMS (Short Message Service)

SMS (Short Message Service) is a technology that enables to send and receive message through cellular phone [24]. SMS basically is a supporting service for the cellular phone user. The SMS service is very popular today which is not only applicable on GSM but also with CDMA, TDMA, even with the fixed phone.

2.10. SMS Gateway

SMS Gateway is a bridge for data (receive and send) transport of the SMS [24]. SMS gateway enables to connect the cellular phone with the computer which ease the SMS usage. SMS gateway normally use for voting system today. The data sent will be saved into the destination database and the developer of the system can use create scripts to manipulate the data. The thesis team is going to implement the SMS gateway for course management purpose.