# CHAPTER TWO

# LITERATURE REVIEW

## 2.1 Introduction

This chapter consist of previous project work carried out as related to the project, it sets the theoretical framework or base for the project and it also gives a brief explanation of the various terms pertaining to the research project.

## 2.2 Quick Response Code

QR (Quick Response) codes have gained significant popularity in recent years due to their versatility and ease of use. They are two-dimensional barcodes that can store a large amount of information, including text, URLs, contact details, and other data, in a compact format. QR codes consist of black modules arranged on a white background, forming a square or rectangular shape. These codes can be scanned and decoded using mobile devices with camera functionality and QR code scanning applications (Wachirawutthichai et al., 2022).

QR codes offer several advantages that make them suitable for various applications. They are easily readable by smartphones and can be scanned quickly, providing a seamless user experience. The error correction capabilities of QR codes allow for accurate decoding even if the code is partially damaged or distorted. Additionally, QR codes can store more information compared to traditional barcodes, making them ideal for storing complex data structures or URLs. Researchers and practitioners have explored the applications of QR codes in diverse fields. For instance, in marketing, QR codes are utilized to provide quick access to product information, promotions, or loyalty programs. They are also employed in ticketing systems, allowing users to easily access electronic tickets for events, transportation, or attractions. Furthermore, QR codes have found use in document verification, enabling the authentication and validation of various certificates, licenses, and identification documents (Yang et al., 2023).

Recent studies have focused on enhancing QR code technology to address specific requirements and challenges. For example, researchers have proposed techniques to improve the security of QR codes, such as incorporating encryption mechanisms to protect sensitive information encoded within the codes (Hsu et al., 2022). Other studies have investigated methods to enhance the visual aesthetics of QR codes by incorporating design elements or patterns while preserving their scannability (Tang et al., 2021).

### 2.2.1 Document verification

Document verification is a vast field such that there is bank type of documents, governmental type of documents, transactions type of document, educational certificates type of document and many more other kinds. Each of the domain and types can be treated differently and the content vary tremendously. For example, transactions can contain number in tabular form while educational certification may contain only textual information presented in paragraphs. Due to the vast differences in types of documents and how they are presented the research will focus on digital verification of paper-based graduation certificates.

Verification is the process of determining or confirming that someone (or something) is original. Documents Verification on the other hands can be define in various ways such as the researchers Warasart and Kuacharoen, (2012), defines document verification as the process of proving the correctness or authenticity of a document by using a proven method or technique. While the researchers Osman and Omar (2016), defines it as the process of ensuring that documents received from holder are genuine and that the holder is the rightful owner. Verification is the evidence that establishes or confirms the accuracy or truth of something while verifying is the act to prove the truth of, as by evidence or testimony.

A certificate verification is the act to prove that a certificate rightly and legally belongs to an organization or an individual or both. It is a computerized means of verifying someone’s claim of certificate- ship from an institution. Online Certificate Verification system improves the speed, quality of service of certificate authentication, globalization of markets, and cuts down cost (Nwachukwu & Igbajar, 2015). Educational establishments try to combat fraud and forgery in several ways, however, most of the methods are time-consuming because they are manual, partly automated or involve human to human interaction (Osman & Omar, 2016).

The main aim of document verification is the ability to trace the origins of a document to a specific person, the device that produced it or the place where it was produced (Srushti, Sanket, Aman, and Tyagraj, 2014). Forgeries pose a huge threat to the integrity of documents, with significant dangers in terms of authentication and trust. It is therefore important to protect the integrity of a document in order to prevent problems arising from the modification of a document by intruders (Srushti, Sanket, Aman, and Tyagraj, 2014). According to the research conducted by Nwachukwu and Igbajar (2015), all documents or credentials that are printed are potentially subject to counterfeiting and forgery. Forgery can cause a lot of damage when it comes to trust and authenticity (Hampo, 2011).

There is a high market for forgery as well as opportunity with low cost, high quality results available (Warasart, & Kuacharoen, 2012). Researchers have also found several significant problem areas when it comes to document verification. For instance, the technologies that are put forth to stop or prevent forgery do not seem to be moving as fast as the evolution of the forging techniques (Singhal, & Pavithr, 2015). With respect to academic documents, further authentication problems include the variations from one school to the next, which causes consistency issues that can be taken advantage of, especially in international situations (Boukar, Yusuf & Muslu, 2017).

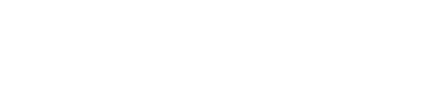
There are two basic document categories that are considered in document verification literature; digital based documents and the traditional paper or printed document. The research in this case deals with certificates. Almost all documents can be handled in a digital manner, except for the certificate. The reason for this exception is that all digital documents are easy to forge without leaving any clues (Tint & Win, 2014). Furthermore, the prevalence of forged certificates results from the increased global demand for higher education, which exceeds the university capacity of the world (Boukar, Yusuf & Muslu, 2017).

According to the research conducted by Tint and Win (2014), there are two main types of forgery, type 1 and type 2. Type 1 forgery is when some part of the original document is changed in order to benefit someone who was not benefitted by the original document. In this case, the base substance, normally the paper or plastic card, remains legal and valid, but the information that is contained therein is forged. The second, type 2 forgery is when both the base substance and the information contained therein is fake. However, it is often very difficult to tell whether it is real or fake because the base substance and the style of the document normally look authentic (Tint & Win, 2014). The researchers of the research Tint and Win (2014), outlined the characteristics of the classic unforgeable document. They also outlined three principles of the unforgeable document as follows;

1. The forged document normally has some difference from an authentic original document in some way
2. The detection of the forgery can happen without reference to the authentic original document
3. There is a concrete verification method that does not necessarily involve communication with an authentication bureau

### 2.2.2 Document verification workflow

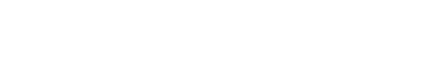
There are three entities that must be present to accomplish the process of document verification which are the issuer, the owner and the verifier. The issuer represents the entity that issues the document such as an educational institution or business organization or even a charity organization. The owner represents the person who owns the document. The verifier represents the employer/third parity that verifies the document. Based on that, the document verification workflow can be diagrammatically represented in Fig. 2.1



**Issuer**

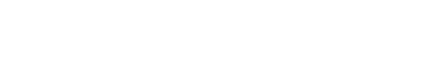
Such as organization, Educational

Institution, etc.



**Owner**

Such as applicant, student, etc.



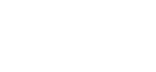
**Verifier**

Such as employee, third parity, etc.



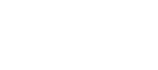
Issues the

document



Owns the

document



Verify the

document

Fig. 2.1. Document Verification Workflow.

### 2.2.3 Types of documents

Documents can be categorized to two categories which are paper based documents and digital based document (Tint & Win, 2014). Paper based document contains characters, digits, tables, etc. Its digital version or digital document is a computer file. Digital document is designed to produce visual information on the computer monitor (Tint & Win, 2014). Forgery of documents has increased jeopardizing the integrity of both the document holder and the organization that issued the document. The forgery of document is classified into two types which are 1) altering part of an authentic document that is original and 2) producing a new fake certificate with false information (Tint & Win, 2014). Forgery of document has become easier than the past mainly because of the technological advancements. For example, scanning and printing hardware are much more advanced than they use to and are not as expensive add to that the editing software that are widely spread and constantly being updated and enhanced. Unfortunately, as document forgery has become easier the increase of fake document has also increased. The latter is due to the lack in advancements in securing as well as verifying the paper-based documents (Osman & Omar, 2016). In other words, documents securing and verification are not advancing as fast as the tools that enable forgery are. For that, the document verification became an important task; it is the process of ensuring that documents presented by prospective employees are genuine and that the holder is the rightful owner.

**2.2.4 Paper based document**

This section will describe and detail on the first type of documents mentioned earlier and is the paper-based documents. Its importance is described and also how they are verified.

The paper-based documents are still widely used. There are many types of paper-based document such as graduation certificates, birth certificates, etc. The information inside the paper-based documents are subject to threats like forgery; despite measures taken to protect them attacks still happen. Authors Boukar, Yusuf and Muslu (2017), attributes that to the lack of verification. There are many cases where documents where forged throughout the globe. For example, one that happened in New Delhi, where five people obtained loans and cheated the banks using fake documents (Osman & Omar, 2016). Another example is one that happened in Bagdad, an investigation of 20,000 government employees by Iraqi's parliament showed that some employees have used forged educational certificates and fake diplomas to get their jobs. The issue extended in that those employees that used fake certificates became senior officials in the government (Srushti, Sanket, Aman, and Tyagraj, 2014). Forgery of documents can happen in any discipline or line of work. In U.S. for example, The National Health Care AntiFraud Association projected that United States of America lost 3% to 10% of total healthcare cost to fraud (GeeksforGeeks, 2018). Another example of forgery that happened in an area that involved the medical discipline is in Malaysia. The mainstream newspaper reported that a statement given by the Congress of Unions of Employees in the Public and Civil Services (CUEPACS) stated that more than 45,000 or 3% of 1.5 million government’s staff in Malaysia forged medical certificate as a reason of absence from work to do part-time jobs. Another discipline that was impacted is Education.

With that has been mentioned document verification is important to overcome many issues that could even do with life and death. Imagine a doctor forging his way into a medical school. Or a politician forging his way to power. As a result, many could be harmed of such a behavior. Document verification of a paper-based document has to be efficient to allow of seamless verification.

**2.2.5 Graduation certificate**

A university is an example of an organization that creates so many documents for their students. It issues a certificate and academic transcript for each of its graduates. The certificate contains information that certifies a person has graduated from a certain specialization and obtained results as stipulated in the certificate. The certificate can then be used for job hunting or pursuing academics or any other purpose. The graduation certificate issued by the universities/institutions is one of the important documents for the graduate. It is a proof of graduate’s qualification and can be used anywhere. Every year millions of students graduate from colleges and Universities, and their numbers are growing. Institutions issue certificates to those who have successfully completed the requirements of graduation. A graduation certificate is still in the form of a paper-based document because, as of yet, an electronic document cannot effectively replace a physical certificate (Smartsheet, 2019). With the rise of graduates and advancements in printing and photocopying technologies, came the rise of fake certificates as well threatening the integrity of both the certificate holder and the university that has issued the certificate (Abolaji, 2017).

The graduation certificate has to be verified to ensure that its content is true and also to ensure that the issued certificate comes from a real source (Osman and Omar, 2016). Fake certificates can be created easily and the quality of a fake certificate can now be as good as the original. The certificates of many prominent universities have been forged and these forgeries are very difficult to detect. Educational establishments try to combat fraud and forgery in several ways Hampo (2011); however, most of the methods are time consuming because they are manual and involve human interaction. A lot of the time is spent in either reaching out to the university to verify a certificate or in awaiting a reply from the university that the certificate is valid and true. This process can be extremely laborious and expensive especially if a company needs to check the certificates of several hundreds of applicants. This adds to need of having a cost-effective fast solution to verify certificates.

**2.2.5.1 Importance of Graduate Certificate**

Graduate certificates are of great importance to land a job or pursue further education they are the proof that the holder possesses the necessary knowledge to take a given position or pursue education. If these certificates are forged the whole foundation could collapse such that the employee would hold a position is not entitled to and could ruin or bring down the company. If it is in an educational institution, it could mean many things of which a seat could be occupied by unworthy person instead of a worthy person.

**2.2.6 Paper-Based Certificate**

Paper based certificates are still widely spread mostly because it is considered more secure than the digital certificate (Srushti, Sanket, Aman, and Tyagraj, 2014), (Singhal an&d Pavithr. 2015). Paper based certificate have stamps and signatures on them which can reflect originality (Abolaji, 2017). Many entities require a stamp and a signature to accept a given document and graduation certificates are no different. However, the issue that arises is that the holder would be bound to providing the original copy every time the stamp and signatures are required. Another importance for paper-based certificate is that they are easy to note from and on; Say the manuscript; modules can easily be highlighted and marked. Allowing multiple reviewers to go through it and do the same.

Paper based certificates despite being widely used they can be damaging. The most important disadvantages are:

1. With paper-based certificates is risk of loss and damage. Paper based certificates can easily be lost especially now as it is easy to relocate between different places and countries.
2. Paper-based certificates is that they can be costly especially if changes are required on the document; for example, a faulty name was printed, more papers would have to be used and that extra cost for the entity issuing the certificate; this indirectly also effects the environment.
3. Paper based certificates can easily be damaged be it a wet hand or a fire in the building; Once the paper documents are damaged, they are usually hard to recover. The holder either has to travel to source to generate the same or if the same is not regenerated it is a loss.
4. Paper based certificates can eventually consume physical space.
5. Paper based certificates can be slow to retrieve.

Despite these drawbacks with Paper based certificates entities still use it.

**2.2.7 Digital based Certificate**

The graduate digital certificate is the certificate that is issued in a digital form. It usually issued through a secure certification and verification method (Tint & Win, 2014). It is mostly adopted in order to solve the management problems of paper-based certificate. However, one of the important reasons why digital certificates are widely adopted is that digital certificates provide a unique feature which is portability (Adams & Blandford, 2012); it is easy to transfer documents when they are digital.

Digital based certificates are considered environment friendly and can easily be organized without taking much space. The digital certificates in the simplest form is the easiest to forge without the need for special hardware (Adams & Blandford, 2012). Editing softwares are widely spread and changes to manuscripts and graduation certificates can easily be made. Digital certificates are easily generated and can be amended with ease. Despite the advantages the digital based certificates they are not widely spread as the paper-based certificates and are not the preferred method for many universities. Even if digital based certificates are issued paper-based certificates are still required and needed.

## 2.3 Review of related literatures

Since (1995), or what Tenopir calls the “post web world” (2003), libraries have been seen as in danger of “substitution” The web is becoming “a ubiquitous source of information” giving an “illusion of depth and comprehensiveness” that leads to a questioning of the value of libraries and their collections. This review will not speculate on these future roles, but will focus instead on the certainty of changing technology, increasingly digital information resources and societal shifts that have changed user expectations of online certificate verification system.

Several approaches have been made to verify certificates and clear the issue of certificate forgery, however, certificate verification method still prevalent today is a manual process, whereby, whoever wants to verify a certificate trips to the institution or send a written request.

In light of the above, Srushti et al. (2014) presented a certificate generation system to ensure an efficient certificate management using huge data and to provide mark sheets for credit-based grading system (CBGS) in a very user-friendly manner. In this system, the admin enters the marks of each student. That information will be stored in internal collection information database, percentage and grade is calculated manually. The system embedded the digital form in mark sheet using encrypted QR code, so that any unauthorized user cannot retrieve any information. However, the system is partly automated made it inefficient.

Hampo (2014), in his work adopted the Structured System Analysis and Design Methodology (SSADM) which emphasizes on completing a phase of the software development before proceeding to the next phase and also being able to go back to the other phases in a purely sequential manner. The model used for this project is the Rapid Application Development (RAD) model proposed by International Business Machine (IBM) in 1980 and introduced to software community by James Martins through his book Rapid Application Development. Unfortunately, it was not a web-based application but a desktop application software which made the system less valuable as compared to web applications.

Osman and Omar (2016), incorporated cryptography approach and cloud-based model to enhance the verification mechanism and thereby reduce the incidence of certificate forgeries and ensure that the security, validity and confidentiality of graduation certificates would be improved. By using the Cloud-based model, some of the factors that result in reduced operational efficiency in student services at universities can be addressed and this should have a positive impact on the quality of services provided by universities. However, since cloud infrastructures are owned and managed by service providers, the cost of implementation is also high. Thus, most institutions could not afford its implementation.

Yusuf, Boukar and Shamiluulu (2018), research work enabled an end-user to define certificate template and template format without the requisite of XML knowledge by clicking few buttons and typing from the system GUI, verifying the certificate and generating one or more certificate(s) simultaneously. In the system, students' details are imported into the system using an excel file, thus, making the system partly automated and inefficient. Singhal and Pavithr (2015), to prevent the circulation of fake degree certificates adopted the use of the QR Code and Smart Phone Application. A QR Code contains a digital signature over the data such as degree holder's name, enrollment number, roll number, total marks obtained etc. which will be signed by university authorities. To verify the digital signature a person needs to use a specific smartphone application which will scan the QR Code and authenticate the certificate. The system was able to combat certificate fraud by embedding the QR Code on the degree certificate and by introducing the smartphone application which will read the digital data from the QR Code. It enables the verification of the certificate without depending on the certificate issuing institution. This did not only improve the authenticity mechanism of a certificate at a much faster rate than manual verification but also prevents the creation of fake certificates through cost-effective.

Musee (2015), in his study Employed Agile Methodology approach and Unified Process modelling to develop a cloudbased prototype which is used as SaaS to provide certificate verification. The prototype allowed users to request to get the academic certificates verified by filling the name of institution, course name, year of graduation and the verification code. All these processes were carried out in private cloud and accessible online.

Boukar, Yusuf and Muslu (2017), adopted the use of Java DataBase Connectivity (JDBC) and MySQL connector jar file hence designed a web-based approach proposed to replace the traditional (manual) verification process by retrieving certificate data from institutions in JSON format and archiving them in a database from which verification can be made eliminating security threats and human error. An SQL query was executed to retrieve relevant information from the database. Results are parsed and presented in a JSON format using the GSON jar file and JSON library functions. However, the use of NoSQL features in MySQL became the major deficiency of their system as it slows down the system operation.

Tint and Win [10] to control fake certificates, considered the combination of Elliptic Curve Digital Signature Algorithm (ECDSA) and Secure Hash Algorithm-1 (SHA-1) algorithm which provides strong cryptographic strength and optimizes the computational speed as well as space. In this process, the input message from the user is hashed into a message digest. This digest code is encrypted into signature value using the ECDSA algorithm. The signature value is converted into barcode. The user input message and barcode are combined into electronic certificate. If a user is a new user, he/she must register first. This user needs to input his/her information and generate public/private key pair. This user information and private key will be used to create an electronic certificate. The system, however, lacked the certification authority (CA) between user and server for a trusted third-party system and to get a more secure client-server authentication system.

Warasart & Kuacharoen (2012), in their paper, implemented a paper-based document authentication in which a document can be verified with the use of a digital signature and QR code. This enables the verification of the documents without depending on any special institute such as the forensic science centre or accessing the database. The verification process can be done automatically if the Optical character recognition or optical character reader (OCR) is accurate. Otherwise, human inspection is required. The inspector can see the differences between the printed message and the message in the QR code. This semi-automated process is the major drawback of their system.

Nwachukwu & Igbajar (2015), considered the adoption of Top-Down structure (a modular approach) with Iterative model and designed an online certificate verification system that can be implemented as a standalone application or embedded in a school official website depending on how the institution decides to use it.

The system was based on an RDBMS for certificate storage though can automate the process of certificate creation and management but lacked partition tolerance i.e. horizontal scalability, Flexibility and above all Efficiency when data became very large.

## 2.3 Summary

Based on what has been presented in the previous sections, there are many techniques proposed for paper-based document verification. Most of these techniques require change in the process of certificate generation either by changing template, changing paper, changing printers, adding hardware or even adding extra information. This change may mean that the university or verifier need the proper knowledge to execute and run the proposed technique. This also mean that older certificates may not work with the new introduced techniques. To also add some proposed techniques, require a change that is not always easy or cheap like in creating a third body to verify certificates.

As reflected some techniques are mostly suitable for specific domain and document like signature extraction for bank cheques. Others were proposed based on specific environments and conditions like environments that assumes both send and receive are known to each other (Osman & Omar, 2016), (Henrieta, 2015).