REPUBLIC OF TURKEY

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Electrical and Computer Engineering

# AUTOMATING ORGANIZATIONAL INTEROPERABILITY IN E-GOVERNMENT SERVICES USING BLOCKCHAIN TECHNOLOGY

**Sarah Kareem Raheem AL-SAEDI**

Master's Thesis

Supervisor

Asst. Prof. Dr. Sefer KURNAZ

Istanbul, 2022

# AUTOMATING ORGANIZATIONAL INTEROPERABILITY IN E-GOVERNMENT SERVICES USING BLOCKCHAIN TECHNOLOGY

**Sarah Kareem Raheem AL-SAEDI**

Electrical and Computer Engineering

Master's Thesis

ALTINBAŞ UNIVERSITY

2022

The thesis titled AUTOMATING ORGANIZATIONAL INTEROPERABILITY IN E-GOVERNMENT SERVICES USING BLOCKCHAIN TECHNOLOGY prepared by SARAH KAREEM RAHEEM AL-SAEDI and submitted on 17/08/2022 has been **accepted unanimously** for the degree of Master's Thesis in Electrical and Computer Engineering.

________________________________

Asst. Prof. Dr. Sefer KURNAZ

Supervisor

Thesis Defense Jury Members:

| | | |
|---|---|---|
| Asst. Prof. Dr. Sefer KURNAZ | Faculty of Engineering and Architecture,<br>Altinbas University | ________________ |
| Asst. Prof. Dr. Oguz KARAN | Faculty of Engineering and Architecture,<br>Altinbas University | ________________ |
| Asst. Prof. Dr. Serdar KARGIN | Faculty of Engineering and Architecture,<br>Beykent University | ________________ |

I hereby declare that this thesis meets all format and submission requirements of a Master's thesis

Submission date of the thesis to the Graduate Education Institute: ____/____/____

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Sarah Kareem Raheem AL-SAEDI

# DEDICATION

First and foremost, I dedicate this thesis to the Almighty God, thank you for your guidance and for giving us the will, power of the mind, protection, support, and healthy life, I am grateful for all of this and ask God to enable me to serve the humanity through this achievement.

I have to thank and give very special gratitude to my beloved parents, without their love, words of encouragement, and support, this thesis would not have been made possible, they are my constant source of inspiration.

I also dedicate this work to the best persons I ever had, my husband , his family and my little angel my daughter " Deema ". thank you for being there for me throughout the entire program to earn the Master's degree. thank you for your endless support, continual care, and attention. Thank you for being the persons who always stand beside me.

My dear brothers and sisters, thank you for all the things you have done for me. I know that you are proud of me for doing this accomplishment. You are truly the best brothers and sisters that a person could ever have.

I also express my heartfelt gratitude to my worthy and respected supervisor Dr. Sefer KURNAZ, Assistant professor, Computer Engineering Department, for his valuable guidance in performing this thesis under his efficient supervision, motivation, enlightening, and collaboration.

## ABSTRACT

## AUTOMATING ORGANIZATIONAL INTEROPERABILITY IN E-GOVERNMENT SERVICES USING BLOCKCHAIN TECHNOLOGY

AL-SAEDI, Sarah Kareem Raheem,

M.S., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Asst. Prof. Dr. Sefer KURNAZ

Date: August / 2022

Pages: 53

The research is concentrated on the development of a smart system for the preservation of government records in smart E-Government benchmark repository. In literature, many existing preservation techniques and models have been discussed and presented with their detailed comparison as preservation rate of digital data growth has increased. Many western countries have already upgraded their paper-based systems to the smart system for preservation of government records including Turkey. In methodology, we used python language for the implementation of smart system application with E-Government benchmark to preserve archived paper-based records, it likewise delivers a roadmap for the innovative arithmetical environment which is trained using the automation based blockchain.. The blockchain model proposes construction that knows how to be utilized to relate household tasks and commitments in the interior of a confined structure. Major archives (i.e. open access, closed, restricted, proprietary) are creating certain assistance for safeguarding numerical objects and files in the E-Government benchmark dataset. Records accession, normalization, and transformation have also been performed on the records during the process of preservation to clean the records and format conversion. As a result, we achieved good preservation of most of the records from 2015 to 2025, the goal is to preserve most of the records it belongs to digital format. The distribution of records preservation for major archives in terms of paper-based and digitally preserved records from the year 2015 to 2025. The preservation of government records was recorded as very low in 2015 as much of the record archival was paper-based in all major repositories of the government. However, year by year the smart system tends to preserve the records in digital format from the old paper-based format, and by 2025 large number of records are being converted to the digital format with a whopping accuracy of 98.68% for all the records where the system was trained on 80% of data and tested on rest of the 20% of data. Having such a smart system can be very helpful for preserving the records of the government, and useful for understanding how records are being preserved and their functionality.

**Keywords:** Smart system, e-government, modeling, blockchain, interoperability.

# TABLE OF CONTENTS

**Pages**

# LIST OF TABLES

**Pages**

# LIST OF FIGURES

**Pages**

# ABBREVIATIONS

| | | |
|---|---|---|
| SIP | : | sustainable information packages |
| OAIS | : | Open Archive Information System |
| ISO | : | International Standards Organization |

# 1. INTRODUCTION

In this thesis, robust smart system and recognition is performed which is an important part of vision-based form recognition for government records preservation. It is worth to say that the records are vital forms for human being to communicate and carried out many important tasks on a daily basis [1]. Record and other government parts tracking from government records has revolutionized not only the scientific world but for entertainment like gaming industry and animation creation. It also took its place in traditional hard work on factory automation, virtual reality, rehabilitation purposes & disability support, performance measurement and many more [2]. Now, there are approaches where user can have unique hardware attached to its government which can have good analysis of the joints and geometry of some government parts. But it is neither natural nor comfortable. In addition, it is relatively expensive to wear specific hardware designed for certain tasks. Therefore, vision-based applications are preferred over those approaches for simplicity, cost effectiveness and natural interface experience [3].

In this research work, the main motive is to develop a smart system for the government for the purpose of preserving the existing paper-based records to the government records. As described in [4], experts are aware of the dangers of maintaining a digital environment for an extended period of time. It is the purpose of this project to discover new methods of informing academics about digital preservation efforts in order to enhance the long-term viability of both research data and government documents. As an approach to digitization, the study focused on actions in a government setting that result in the construction of sustainability everything now for three specific research digital libraries (electricity, finance, science, and technology) [5]. The Open Archive Information paradigm is based on paper-based information bundles. International Standards Organization standard OAIS is a framework for the emerging digital world in which paper-based data will be normalized and transformed into government records.

Almost every civilization on the earth has adopted digital formats for mediated communication and knowledge, as described in [6]. Digital transformation's consequences have been researched from various angles, but it may take generations to completely understand its influence. We proposed a new digital model over four years of governmental records from 2015 to 2019, It is possible, as described in [7], to accidentally delete priceless data sources from our collective knowledge repository by mistake due to the exponential rise of digital data. This can lead to both remarkable chances for finding and the accidental deletion of precious data sources. Despite rapid advancements in data storage technology (hardware and software), individual and organizational behavior have lagged far behind [8]. While cloud storage has become easier, data protection has become more difficult because of the task's scale and the fact that it is done in an unseen, abstract manner.

People needs to use mouse, keypads and remotes to control and manipulate machines. This interface between devices and humans can certainly be improved in coming years. Several researches have been going on to modify existing tools and components which com- bine human and machine interface [9]. It can be thought of by giving commands to a computer or robot with speech and/or some forms. This is the way we humans conduct communication among ourselves and we can definitely extend this way of communications to machines as well [10]. This task is the small attempt in that direction. We would like it to understand human actions and commands with forms.

## 1.1 ANALYSIS OF THIS SUBJECT AREA

There are many related issues of record form recognition. The most widely recognized ones are:

a. Government Record Localization: Means to decide the picture position of each single location of record.

b. Government Record Identification: Plans to recognize the nearness and area of highlights, for example, lines, paragraphs, figures, sections and forms.

c. Government Record Recognizable Proof: Thinks about an info picture against a database and reports a match, assuming any.

d. Government Record Verification Issue: Checks the case of the normalization of record in an information picture.

e. Government Record Gauge Issue: Gauge the area and conceivable direction of a record in a picture grouping.

f. Government Record Form Articulation Acknowledgment: Concerns recognizing the full of states (To check if the lines on records changes over time in different records) of different organizations.

Obviously, smart system is the initial phase in any mechanized framework which takes care of the above issues.

## 1.2 REVIEW OF ANALOGUES

Smart system and recognition for movement of records using different machine learning techniques as it is most important task as mentioned in [11]. It can also be considered as in terms of human mimicking in this huge task for making the robot suitable for industry work. The idea is that, the robot should observe human record movements and make other easy to understand at some point what happen around its surrounding as mentioned in [12]. That's why accurate and robust real time record movement detection is priority task for this project which can be the first step forward to proceed. Machine learning will eventually work side by side with humans in industry area where it has to understand record form and be able to repeat at some point of time as mentioned in [13]. It must have a very good insight about where humans' records are with respect to its world coordinates and also it must be able to detect an object for being operated upon. In this task, we are limiting our approach to human records tracking without any records in record as given in [14]. It would be simpler approach where record movements are being observed by machine learning and with form or trajectory detection algorithm developed in the later part of this thesis would help to recognize record to understand that movement. Record motion tracking algorithm locates records on frame by frame basis and continuously do so as they move in 3D space as mentioned in [16]. Most tracking algorithms consider object's shape size and color in to account. But once tracking is missed or lost, algorithms fails to find records and again needs to reinitialize [17]. Skin mask is an important feature for smart system which will be the indicator of records and face in the scene.

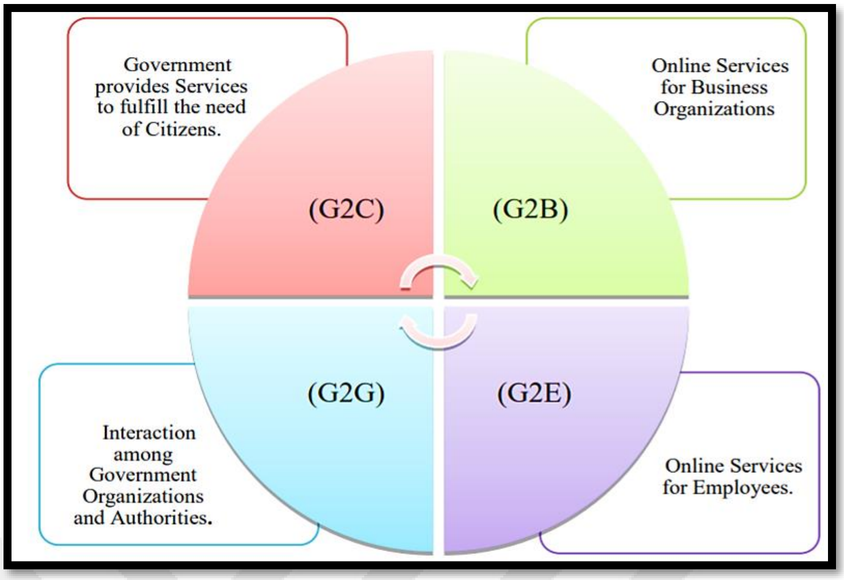


**Figure 1.1:** An example of four forms. Each form represents a label of every record that belongs to the smart system based on blockchain [17].

Without clear skin mask, depth is not helpful to track records in this case. Although false skin segmentation vs tracking efficiency needs to be balanced perfectly; by putting strict thresholds would limit false detection but at the same time will have less information to process as mentioned in [18]. Hence efficiency of tracking algorithm will suffer. On the other record, if we keep thresholds too big then the noise and false skin indicator will increase error in detection.

Record motion tracking algorithm locates records on frame by frame basis and continuously do so as they move in record space. Most tracking algorithms consider object's shape size and color in to account [19]. But once tracking is missed or lost, algorithms fails to find records and again needs to reinitialize. Skin mask is an important feature for smart system which will be the indicator of records and face in the scene [20]. Without clear skin mask, depth is not helpful to track records in this case. Although false skin segmentation vs tracking efficiency needs to be balanced perfectly; by putting strict thresholds would limit false detection but at the same time will have less information to process. Hence efficiency of tracking algorithm will suffer. On the other record, if we keep thresholds too big then the noise and false skin indicator will increase error in detection [21].

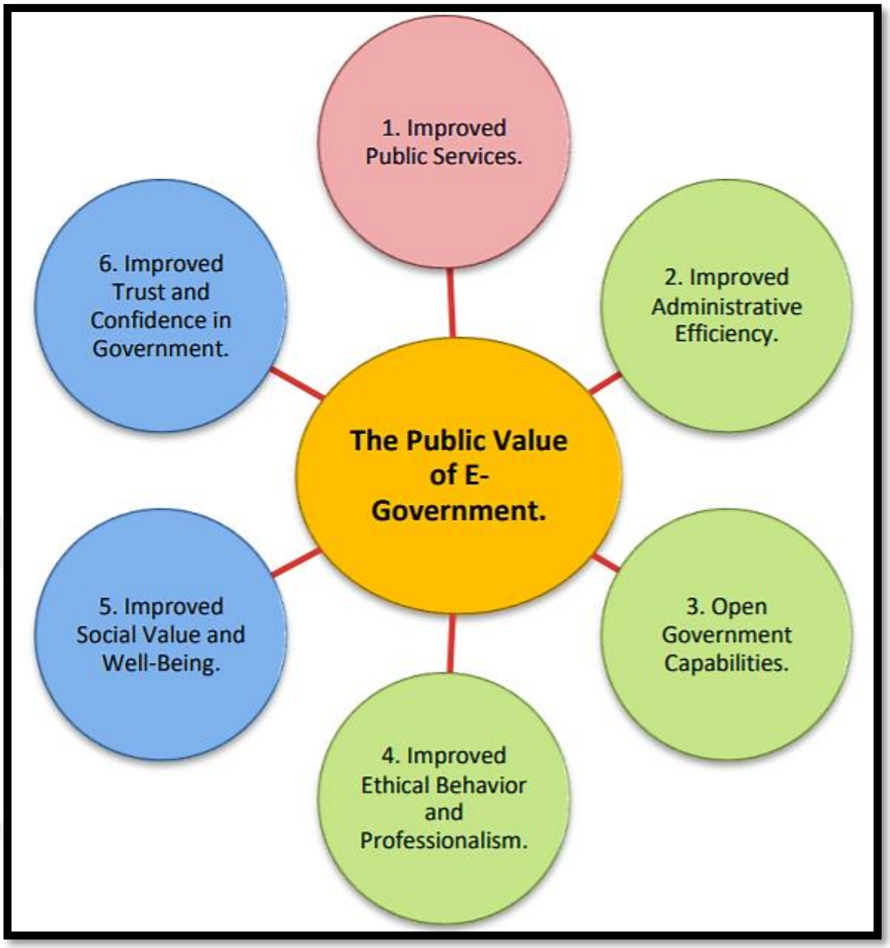


**Figure 1.2:** All basic services that should be provided by E-Government [21].

There are several attempts made for real time robust record tracking. This problem is quite old and many approaches has been developed to detect and understand record movement with many different devices [22]. Earlier people used to wear gloves and some special hard- ware to detect record postilion and used to interact with the world. The record camera was one of the simple solutions for complex hardware problem where people can extract skin information from color records with different chrominance models [23].

State of the art thesis include many record-based identification and tracking approaches. Representation with modeling, specifying key features for detection, contours tracking, or by using classification techniques such as neural network and binary tree classification [24]. Although machine learning algorithms and classifiers needs robust training with lots of data only then they are able to segment record from records. This approach needs lots of processing and hence computationally expensive [25]. The more traditional method where we use record form depth information and try to find record from those records combined; is an ideal approach for this task.

## 1.3 ANALYSIS OF BASIC TECHNOLOGICAL SOLUTION

Our system is designed to maximize user experience. Moreover, our system differs from existing systems in the following ways.

i. Just records. Many existing systems require users to wear gloves or markers to be capture by the camera. We feel this limits the usability of the application and aim to keep the prediction of the system entirely marker free.

ii. Real-time. Our system should run smoothly on average modern computers with a dedicated graphics card. The system should also recognize record form at a high frame rate, nearing real-time speeds. Our desired frame rate that we want achieved a frame rate for record storage. In our design and implementation, one driving goal is to squeeze every millisecond as possible.

iii. No calibration. Once trained, our system should require no further calibration before working in a new environment.

iv. Robust and accurate. Our system should have an accurate estimation of where the user's records are and what forms they use with low false positive. Moreover, the system should be insensitive to various background, user's location, camera position and other noise.

v. Arbitrary forms. Our system should be able to easily incorporate new types of forms. With sufficient training, our system should support many complex forms such as those seen in the sign language.

## 1.4 SUMMARY

i. It is important to distinguish between data processing and data preservation. According to [26], "permanent storage" is achieved through long-term data storage in a secure mechanism or model, such as the disk drive. Storage, on the other hand, requires a "place," such as a hard drive. When it comes to preserving data, today's technology relies primarily on the ability to store more information on smaller and smaller devices. It's like having parts confined in a bottle out at sea as devices and hard disks come and go. It is possible to avoid "digital forgetfulness," but only when the data are carefully adhered to. Media obsolescence and other threats to digital data preservation can be addressed through automation, but human interaction is required to assess, describe and prioritize digital artifacts. To preserve its accessibility, content, and meaning, "digital libraries" are required. Intangible assets, such as digitally stored information, can be amassed at a low cost and are becoming increasingly so as technology advances.

ii. A method for collecting records in real time. Based on machine learning, we build and implement a complete real-time record form recognition system. The system uses a web socket server to broadcast the location and form of the data, which can be used by other applications.

iii. A low-cost method of producing accurate labeled data for record form recognition. Using aligned depth records and government documents, we can quickly generate labeled data. The creators of [28] create training samples using advanced computer graphics. We found this to be too expensive, and developers can easily create their own unique form using coloured gloves. Another benefit of our method is that it employs real-world raw depth recordings as training samples. These capture realistic noise like shadows and hardware noise in a natural way. Simulating these chaotic effects with computer generated visuals, as done in [29], is extremely difficult. Color gloves are not required for end users; they are primarily used in training.

iv. A computational insight into a data model for preservation of records. To our knowledge, there does not appear to be any literature comparing digital from a computational standpoint. Rather of simply providing experimental accuracy as is done in most machine learning work, we provide an in-depth complexity analysis of the two methods.

v. The system's extensive experimental evaluations. By carefully exploring a broad array of parameters, we undertake lengthy trials to evaluate the effectiveness of the random forest classifier. The results are intriguing, as they lead to a better understanding of random forests.

# 2. RELATED WORK

## 2.1 DEFINITION OF REQUIREMENTS

In [30] when the author formalized and implemented the model for preservation of government records, The model's credibility, reliability, intimacy, and self-awareness would all help to build institutional trust by giving people the impression than people are in charge (not self-serving). A major difficulty will be establishing systems of trust and delineating roles at all levels of government within a single nation. Additional issues arise when these models are applied outside of national borders. They've also been keen to preserve their personal digital assets and traditions to society at large, particularly in countries with more advanced technology infrastructures. England, France, New Zealand and Australia were among the countries that acknowledged the importance of digital preservation for future generations [31]. A number of early-adopter countries are drafting guidelines to help their country deal with the issue. Everyone agrees that technology alone cannot provide a solution. The OAIS is one of many international (ISO) standards aimed at resolving issues that have arisen as the globe has become increasingly interconnected. Focusing on the average best validation error, two statistics seem to correlate with its absolute improvement. On one record, the improvement seems to shrink with the number of classes and thus the number of original samples in the training set. As the number of selected classes is increased from seven to ten, the difference becomes smaller and even negative. However, even though relatively small, the classification accuracy is higher when the data augmentation is used for twenty classes as well as the whole dataset.

On the other record, the data augmentation becomes more effective for subsets with larger maximum sequence length. While there was a similar amount of original training samples in the three five-class subsets, the maximum sequence length strongly varied. The largest absolute improvement was achieved for the classes where the maximum sequence length was the highest [32]. While the maximum length was equal in the experiments with seven, eight, nine and ten classes, where a shrinking improvement was observable with increasing training set size, the maximum sequence lengths are significantly larger when the experiments are repeated on the classes. In these cases, the data augmentation has positive effects again. As the maximum number of frames gets clipped to smaller numbers for the created sequences, the synthetic samples become less helpful for improving the classification accuracy on the whole dataset. The classification scores are all smaller than the original baseline when the synthetic sequences have a maximum length of only 100, 200 or 500 frames.

These high-level observations do not allow definite conclusions and it seems likely that the effectiveness of the data augmentation depends on many other factors, foremost the shape, type and variations of the selected classes as mentioned in [33]. That is why, it is difficult to determine a definite limit for which the augmentation method starts to fail. However, it is intuitive that the data augmentation is more advantageous on smaller sets which can be more easily enriched with new information. While the visual quality of the synthetic samples is negatively affected by the smaller training set, the generated sequences appear to carry enough information to have a beneficial impact. Moreover, a classifier's proneness to overfitting rises with decreasing amount of training samples, increasing the need for any regularizing input.

The model's credibility, reliability, intimacy, and self-awareness would all help to build institutional trust by giving people the impression than people are in charge (not self-serving). A major difficulty will be establishing systems of trust and delineating roles at all levels of government within a single nation. Additional issues arise when these models are applied outside of national borders. They've also been keen to preserve their personal digital assets and traditions to society at large, particularly in countries with more advanced technology infrastructures. England, France, New Zealand and Australia were among the countries that acknowledged the importance of digital preservation for future generations [31]. A number of early-adopter countries are drafting guidelines to help their country deal with the issue. Everyone agrees that technology alone cannot provide a solution. The OAIS is one of many international (ISO) standards aimed at resolving issues that have arisen as the globe has become increasingly interconnected.

## 2.2 FUNCTIONAL REQUIREMENTS

There are both personal and collective obligations in the digital sustainable practices for data, in which an info package results from a complex set of behaviors, decisions, rules and practices. Success in a systemic method is determined by how behaviors are shared, according to [36]. Online, people's views aren't as tied to geography. Group resemblance can be found across political boundaries and cultures, and the market encourages it. There are three ways to make a choice according to diffusion theory: individual, group, and government. [37] The Bureau is enforcing an institutional innovation judgment on researchers who can act within their specialty to produce a collective choice led by collaborative societal pressure despite working at an institution where academic freedom dictates that the conditions be voluntary. A company's digital preservation team has a variety of roles and responsibilities. The institution's digital assets have a strategic and long-term value that must be understood by senior officials. The choices they make have a ripple effect throughout the organization. The systems themselves require the attention of technologists. People who work in archives serve as both stewards and role models for those who wish to follow in their footsteps. These individuals are expected to be well-versed in both technology and law and policy in order to effectively manage the archives   According to [39], their social and professional behavior should be impacted by their sector, discipline, and associates. All archives must be administered by their authors, who are fully responsible for both local and distant system administration (at least initially). In order to guarantee that information recorded now will be accessible four years from now in government, it should be explained how these factors work together.

In [40] author calculates depth histogram and thresholds depth record based on that. They assume that the records are in front of the camera all the time and the rest of human government is considered as part of background. They calculate distance of records and other government parts from accumulated histogram by analyzing rapid increase in the histogram slops as described in [41]. They assume that the records are in front of camera and if slope is greater than certain value then this region is selected as candidate ROI. They use color information to segment other records at the same depth level, thus they turn to skin color detection. This method is highly restrictive for record movement, since it requires records to be in front of camera all the times.

It deals solely with depth records and finds record and its movement as described in [42]. They propose classifier that combines a boosting and cascade structure. The features for training are depth differences at the stage of detection and learning. They also implement depth region grow and Depth Adaptive Mean Shift. Although they claim successful results the tracking algorithm can be quite slow because of region growing method.

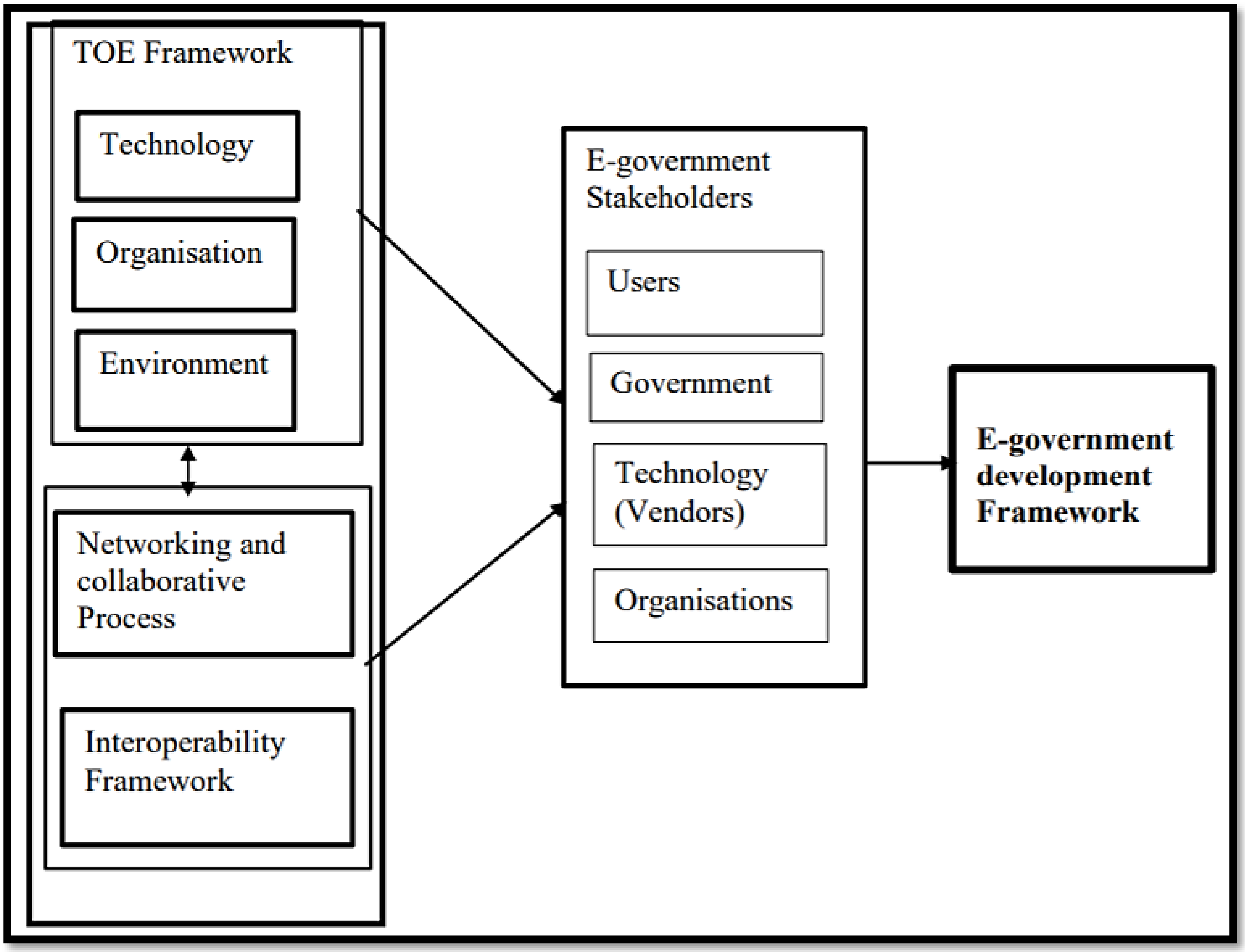


**Figure 2.1:** Record form recognition using the meta data from smart system based on blockchain. Source [43]

Research faculty need to be made aware of the importance of digital preservation in their work, and this case investigates the best ways to make that knowledge available to them. Researchers sought a way to promote improved digital storage at the university with communication with the awareness that faculty are significantly influenced within their field [44]. Departmental communications, rather than college or university-level ones, should be more effective because faculty are examined, funded, & tenured at the departmental level. According to this survey, researchers in each of the three departments were asked to describe how they handle their individual digital data. The type of research is the researcher [45]. Digital preservation procedures are described using a dataset and a blockchain activity theory model. As stated in [46], the preservation and management of government records is at the heart of the digitalization concept.

The three primary collections of government are being preserved in this initiative, whose scope extends beyond the technologies we use to access it. Decisions about the preservation of digital artifacts are made by their creators, the systems that manage them, the techniques that preserve them, and policy & funding for them are all based on these decisions. Digital artifacts have a monetary value that can only be determined by the researchers who created them. It is imperative that preservation procedures change to meet the society's challenges of the digital age. Choosing what data to keep has become much more difficult as a result of the ephemerality of the digital world. Record digitization relies on the elements of the DOI theory, which has been quantified through digital system analysis.

Specifically for the government's digital plan and activity clusters, this research has been designed. Regulations, resources, instruments, and cultural norms all fall within the purview of Government Activity Theory (GAT). According to [47], digital preservation procedures have been introduced and disseminated throughout the government because of digitalization. All digital libraries are compared to government's primary digital libraries in the endeavor to preserve records digitally.

## 2.3 NON-FUNCTIONAL REQUIREMENTS

Digital sustainability can be improved by understanding researcher processes in relation to digital self-archiving habits. Having access to various skills and resources can help them gain a better understanding of their own experiences. The conversation on campus should be improved through leadership support and guidance, which should lead to more collaborative solutions. Academic disciplines have a significant impact on how researchers conduct themselves. Researchers rely heavily on input from their peers in the area to improve their research methods. Activities like submitting funding applications to the finance or science ministries affect how researchers think, feel, and act. When it comes to a researcher's job, they are held to the same standard as any other employee at a university or college. To negotiate and learn best practices, researchers, technologists, and administrators should engage in a discourse on campus. Motivation and leadership are the keys to achieving this on a college campus. The structure, policies, and practices needed to deal with this problem are currently being developed in both the public and commercial sectors. To ensure the long-term viability of archival data, we turned to datasets. It provides a framework for identifying roles and responsibilities in a constrained system. As mentioned in [49], library services are developing specific preservation guidelines for digital artifacts. Employers, administrators, technicians, and organizations must take into account an all-inclusive dbms mandated by both the financial and scientific and technological sectors. Finance and science ministry mandates are key events in the digital preservation of humanity's wisdom.

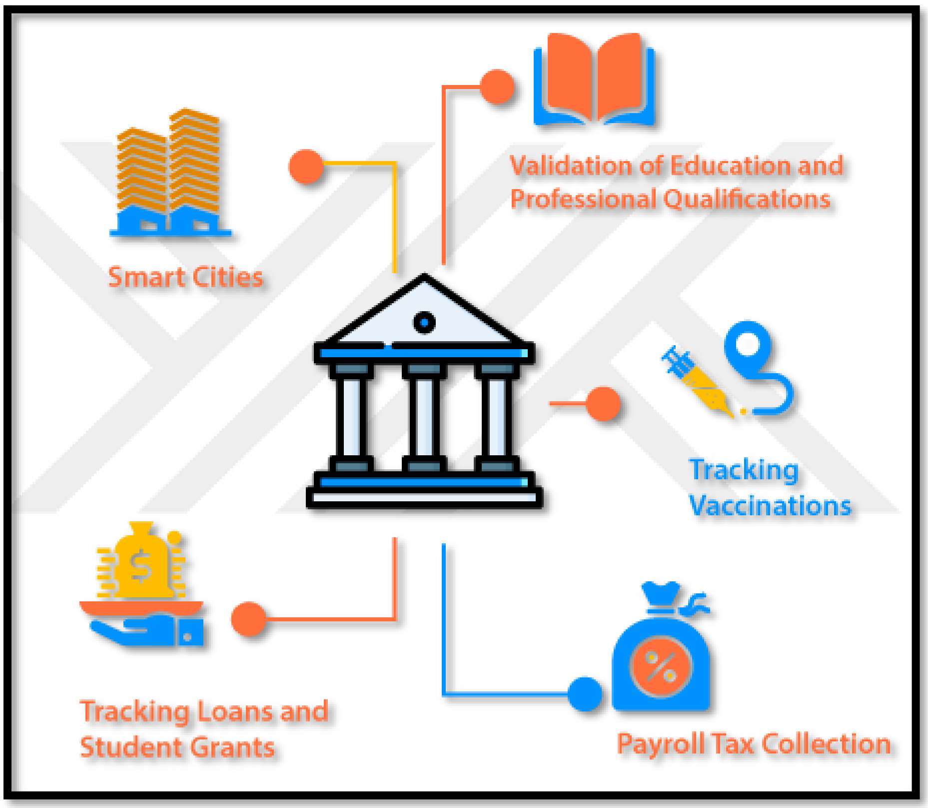


**Figure 2.2:** Blockchain Solutions for Government Records [49].

According to the findings of the research, the most advanced method for preserving records has short-sighted attitudes. While many researchers are disappointed with their own abilities, few are concerned about preserving their work for future generations. They claim that unless it aids them in finishing their research, the pressure to publish undermines their efforts to alter their behavior. Providing legislation, resources, and encouraging attitudes to inculcate long-term digital habits, as outlined in [50], is the responsibility of the government. The ability to establish mechanisms that assist their research faculty is a necessity for administrators. To ensure that their data is accessible in the future, scientists who seek for these awards should confer with experts in their institution and their field. Data source workshops and classes, for example, give the community another means of exchanging ideas and reiterating key messages.

# 3. METHODOLOGY

Even while the public records have employed ever deeper networks, intelligent technology techniques will benefit greatly from the flexibility of the blockchain algorithm. It is possible to ascertain more complex functions with more data models because of their larger representational capability. In addition, the field's rapid evolution brings new methodologies and structures on a regular basis, expanding the designer's toolkit. It is reasonable to say that the connection should be maintained as basic as possible, given that science has always favored simplicity. The CNN technique has been used in a study that questioned the use of complex activation functions in deep learning. Recurrent layers with the right stride and filtration size parameters should be used instead.

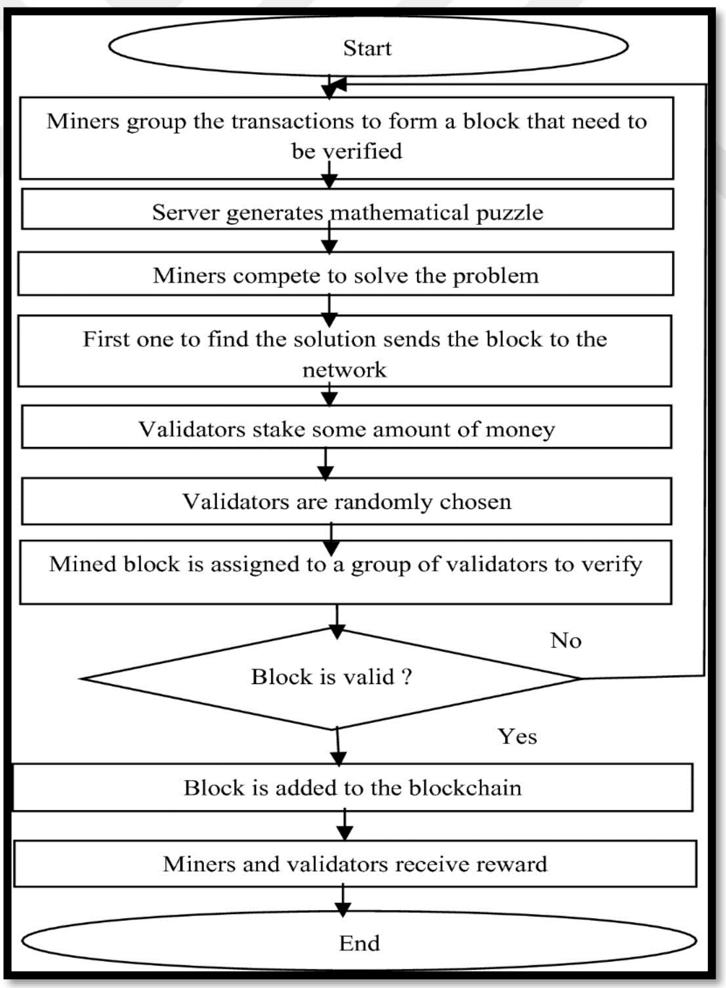


**Figure 3.1:** Proposed system depicted by flowchart using blockchain.

## 3.1 ARCHITECTURE OF THE PROPOSED SOLUTION

The blockchain is in charge of feature extraction in this architecture, with the characteristics to which they will reply defined through a learning process of the variable input connections. Lower-level CNN models will extract simple characteristics like specific line and edge orientations, and higher-level blockchain models would extract more global information (e.g. parts of learned patterns). The blockchain model receives inputs from preceding levels' blockchain models via fixed links and is responsible for robust pattern detection (i.e. decreasing sensitivity to a deformation or location shift of a pattern). Each CNN model receives inputs from a number of blockchain models ('the input window') in one cell plane, i.e. blockchains created as part of this research and whose networks are similar. It expresses the behavior of the models mathematically as convolutions and subsampling, respectively. The network, on the other hand, distinguishes itself from cognition in its training process. Whereas neo-cognition uses a layer-by-layer, unsupervised training procedure for the cell layers and supervised training for the output layer, blockchain uses back-propagation to find a global minimum across all parameters. The essential concepts that resemble biology and make blockchains so fusion for record smoothing. models that extract the same feature from multiple spatial locations If at least one of the models provides an input, the models will fire. As a result, the feature will be detected even if the input features' locations change, making the system less sensitive. However, the models' behavior can also be read from a different perspective. The input windows for the various models are extremely close together, thus models can be thought of as spatially blurring the excitatory signals they get from the models. The input window signals, which are models that perceive the same feature at slightly different locations, are averaged to generate this spatial pooling. Furthermore, a tiny inhibitory area generally frames the excitatory cell input window. This allows the models to identify a continuous line from the line's endpoint and the network to recognize features even after they have been blurred. As a result of the models' fixed input windows, the size of the models blockchain panel. This work gave rise to blockchains, which have a network design that is comparable to that of this study. It expresses the behavior of the models mathematically as convolutions and subsampling, respectively. The network, on the other hand, distinguishes itself from cognition in its training process. Whereas neo-cognition uses a layer-by-layer, unsupervised training procedure for the cell layers and supervised training for the output layer, blockchain uses back-propagation to find a global minimum across all parameters. The essential concepts that imitate biology and make blockchains suitable for record smoothing.

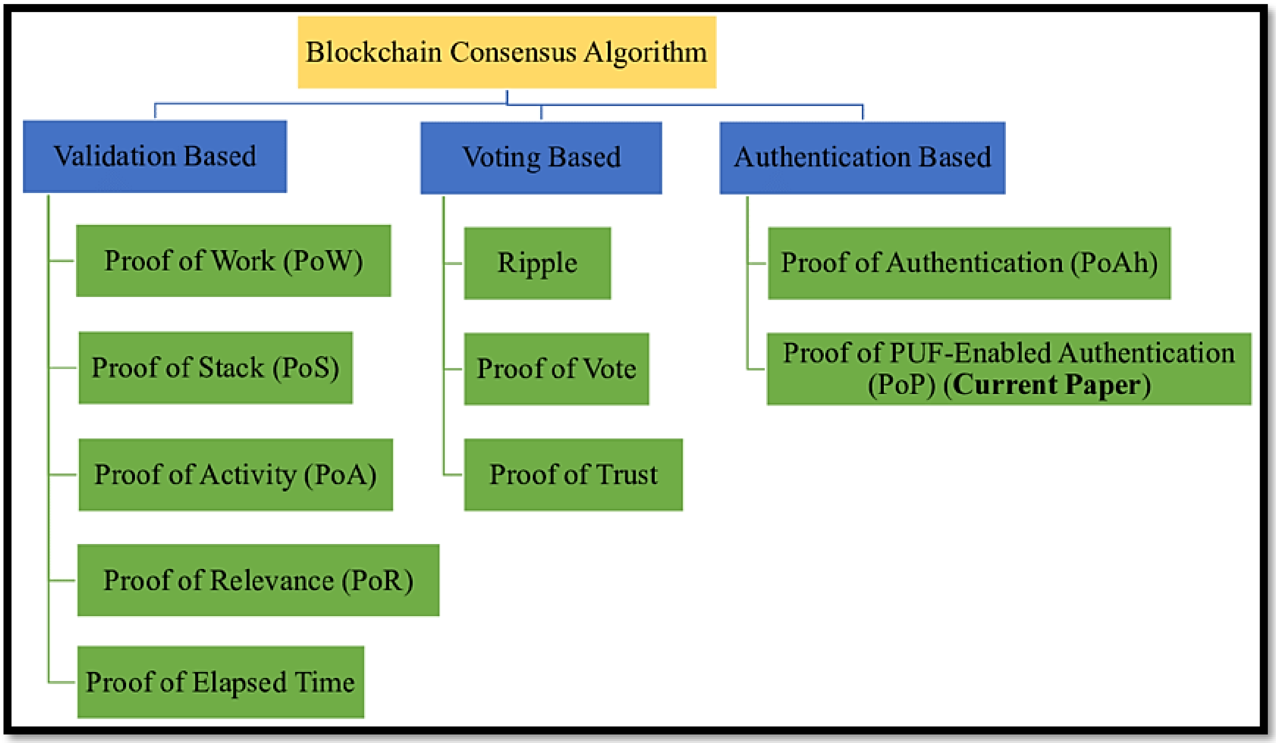


**Figure 3.2:** Illustration of a blockchain algorithms.

Due to the high density of research in record illumination estimation, multi-exposure records often appears clumped. Thus, The major goal of this study is to find a patch-wise solution for illumination estimation with human intervention using blockchains that are derived from this work and have a comparable network design. It expresses the behavior of the models mathematically as convolutions and subsampling, respectively. The network, on the other hand, distinguishes itself from cognition in its training process. Whereas neo-cognition uses a layer-by-layer, unsupervised training procedure for the cell layers and supervised training for the output layer, blockchain uses back-propagation to find a global minimum across all parameters. The essential concepts that resemble biology and make blockchains so fusion for record smoothing. Separating different instances of multi-exposure fusion records using a blockchain approach. Seeded segmentation, for example, is a traditional segmentation method. The goal of this research is to overcome these dominant problems in the patch-wise illumination estimation for multi-exposure government records. We will use a blockchain for multi-exposure government records. We will train the neural network on hand-annotated records and evaluate its performance with object-based metrics.

### 3.1.1 Key Concepts of Validation Based Algorithm

Convolutional layer as records contains highly correlated 2D information (generally in three RGB-channels), the idea of local connections has been exploited by creating units in a convolutional layer which receive weighted input from local patches in the previous layer, referred to as their receptive fields (RF's). By relying on local RF's, this introduces sparse connections, rather than the whole-scale record statistics passed on to units and causing high-dimensional problems in fully connected feed-forward networks. The weight vector of a unit is also known as filter bank and through training the weights will be updated to filter for specific features, giving the units the behavior of filter detectors. The weighted inputs and additive bias are then passed through the activation function of said unit, where ReLU's have portrayed much faster learning in deep networks compared to other activation functions.

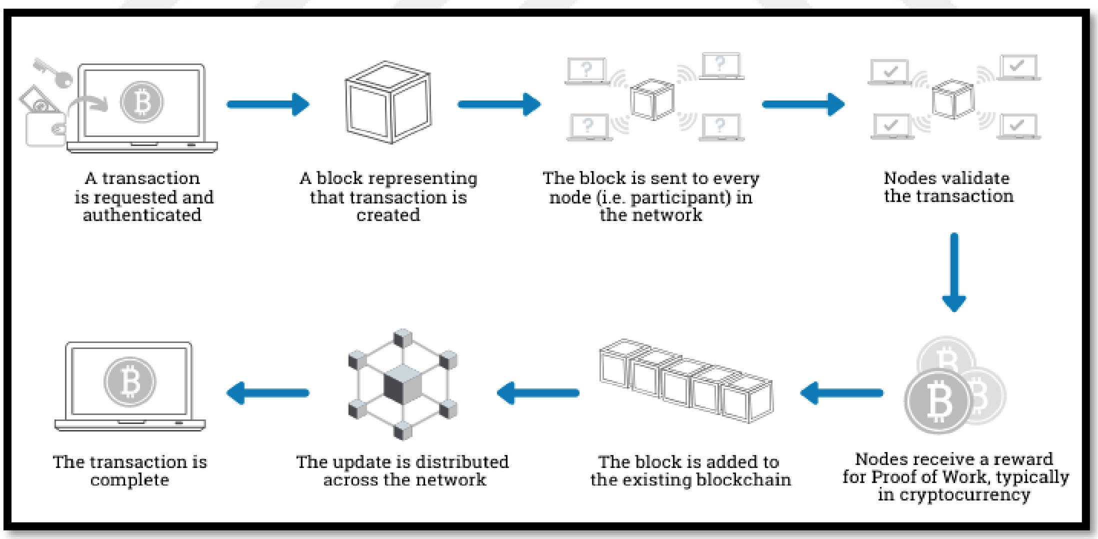


**Figure 3.3:** Basic building blocks of proof of work based blockchain.

Furthermore, recordings frequently depict the same qualities in multiple geographical places. As a result, units with varied receptive fields will have the same weight vectors to extract similar features, which is known as parameter sharing and decreases the problem's dimensionality. On a structural level, the outputs of units with the same weight vector are organized in feature maps.

Multiple feature maps (with different weight vectors) in each blockchain layer enable scanning for different features in each local record patch. The optimal number of feature maps needed to capture the record dynamics tends to increase with the complexity of the input records.

### 3.1.2 Blockchain Layer

The feature maps in the convolutional layer serve as input for the next stage, being a pooling layer. Given that the relative location of features with respect to each other contains significantly more information than their precise pixel-location, further stressed by the notion that small shifts in the precise feature locations are to be expected in different records, the idea of local pooling of features to reduce spatial resolution is implemented at this stage. This results in a network invariant to small shifts in the records and the stages of convolutional/pooling layers equip the network to handle variable size inputs.

Pooling units will scan the feature maps with (non-)overlapping p1 x p2 receptive fields and output a single value. Typical choices for pooling units are subsampling, where the output of a unit in the $j^{th}$ layer is the average of units in the receptive field, supplemented with additive bias b and passed through an activation function and max pooling.

### 3.1.3 Deep Hierarchies in Blockchain

Finally, the use of many layers represents the different stages in the simplified model of the ventral stream. Features at higher-level stages results from combinations of lower-level features, thus enabling deep structures to capture the compositional hierarchy of natural records.

### 3.1.4 Training and Validation

A hybrid model of the blockchain algorithm is used in the implementation. It was out that the visualization method was useful in understanding the predictions, even though the outcomes were unexpected. Initially, it was hoped that the blockchain systems would be able to detect even the tiniest of symptoms of skin infection risk. As a result, they were able to identify a wider range of possible patterns that could predict the likelihood of fractures. This is, of course, the polar opposite of what I just said. A new type of clinical prediction ability can be achieved by combining automation and imaging approaches, as demonstrated in these studies. Despite this, their full potential is still to be realized due to a lack of data and the large range of architectural options available to them. Regardless of the model, it is also possible that data analysis alone will not be able to produce the best forecast of fracture risk. An improved prediction algorithm that incorporates a wider range of input data could help smooth the edges of the record. In the studies, using hybrid robotic techniques yielded the best results for record edge smoothening prediction. The results show that hybrid automation has the claimed advantages, such as quicker training times and the ability to handle huge data.

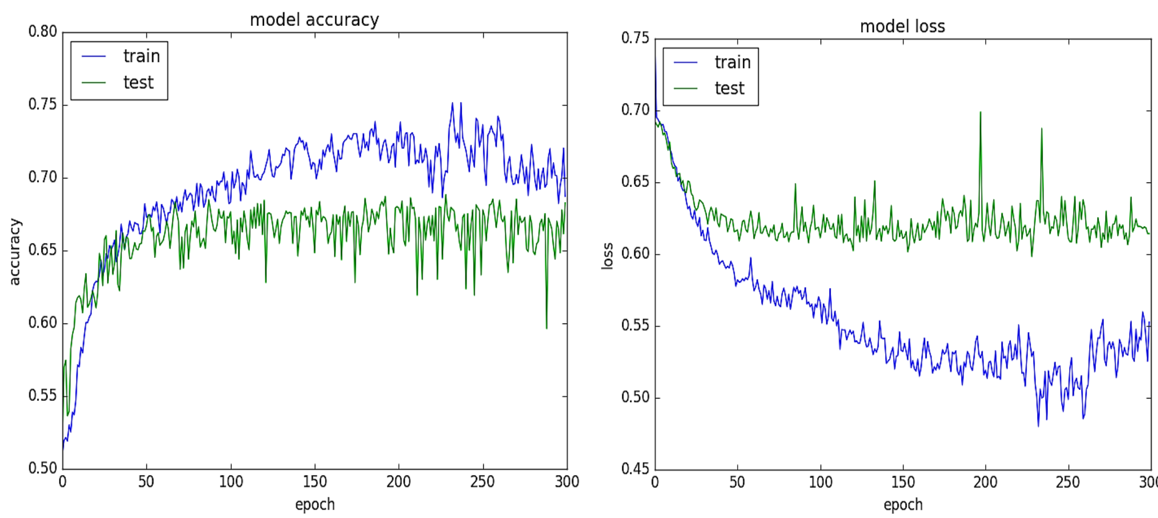


**Figure 3.4:** The blockchain model accuracy (accuracy vs epochs) model loss (loss vs epochs).

Given the current hype of automation, optimizations of blockchain architectures and designs is in hot pursuit. This section will therefore be devoted to some of the powerful blockchain models that have been developed in the past few years and will be used throughout this work. Specific, revolutionary design-elements that have assisted these improvements are marked throughout this discourse.

### 3.1.5 E-Government Validation

An interpretable machine can teach humans how to make better choices through cross-validation, which is employed in AI-based systems. Transparency can be improved by using graphical representations to locate the records that are most critical to the model's predictions. Imagining dot impact on forecast score can further emphasize the textures the model observes. Guided backpropagation is another visualization method. Record edge smoothening prognosis is illustrated by showing the discovered features. Neurons whose output had a favorable effect are only taken into account, while those whose output had a negative effect are set to 0. The blockchain model's predictions can be explained visually using a variety of visualization techniques. It is possible to increase the training set in order to better cover the issue space. The term "cross validation" refers to a technique in which samples are subjected to random transformations, such as rotational and flipping, as well as changes in property values, such as brightness and contrast.

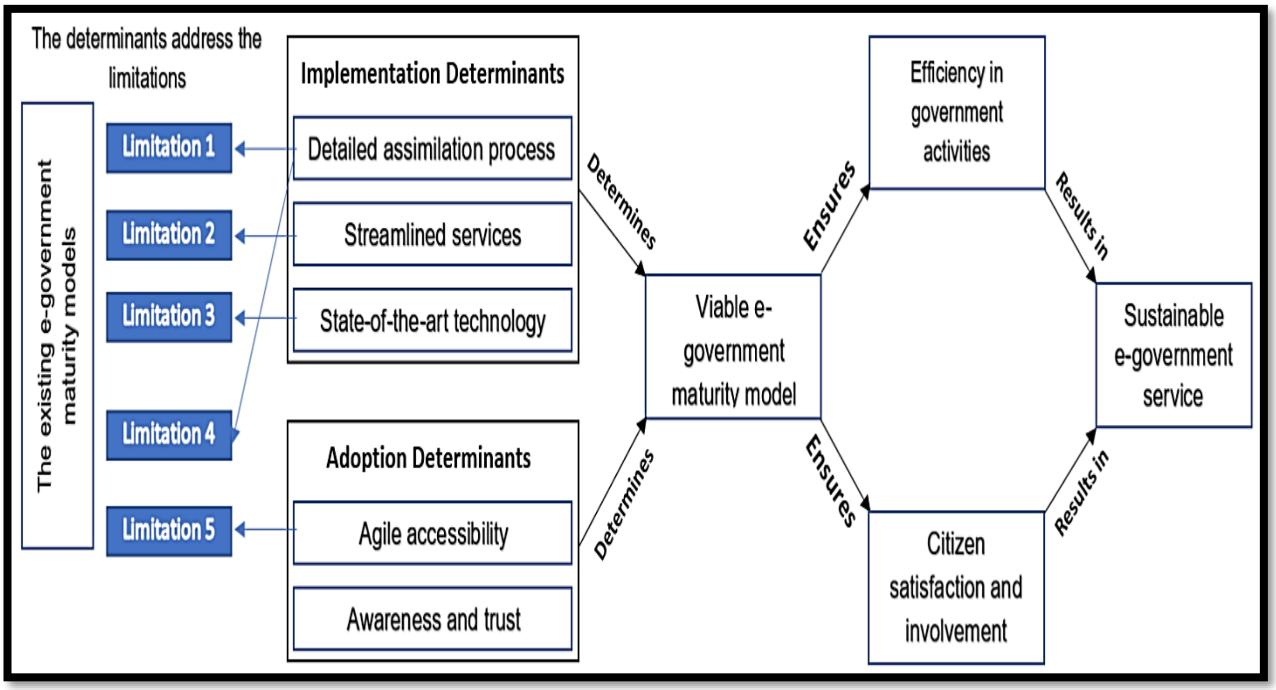


**Figure 3.5:** An example of services being provided by E-government being derived by blockchain.

### 3.1.6 Data Evaluation

Record analysis may appear to be a straightforward matter of loading the network with basic record data because automation techniques require numerical data as input. In the end, records are just grids of integers that represent pixel intensities. For simple recognition jobs with small, normalized records, this strategy works effectively. The issue is a lack of ability to expand. The amount of calculation required skyrockets as the size ( number of layers of an input increase.

**Table 3.1:** The number of epochs per training instances for whole dataset.

| Number of Epochs | Loss | Accuracy | Validation Loss | Validation Accuracy |
|---|---|---|---|---|
| 25 | 0.3301 | 0.8492 | 0.7911 | 0.5360 |
| 50 | 0.2418 | 0.9047 | 0.7874 | 0.5360 |
| 75 | 0.1891 | 0.9350 | 0.7896 | 0.5360 |
| 100 | 0.1702 | 0.9426 | 0.7456 | 0.5379 |
| 125 | 0.1384 | 0.9606 | 0.4709 | 0.7803 |
| 150 | 0.1246 | 0.9625 | 0.3569 | 0.8220 |
| 175 | 0.1020 | 0.9768 | 0.3543 | 0.8295 |
| 200 | 0.0809 | 0.9829 | 0.3598 | 0.8371 |
| 250 | 0.0747 | 0.9853 | 0.3662 | 0.8295 |
| 300 | 0.0678 | 0.9867 | 0.3709 | 0.8314 |

## 3.2 ALGORITHM FOR SOLVING THE PROBLEM

Smoothing approaches, filter status and topography as well as the features of the filter and other categorization variables can all cause variation. For this reason, the task of distinguishing between distinct field record filters can become unclear due to local variances. This makes it challenging to design robust record computational methods that are able to recognize all types of field examples without becoming overfitted to the specific environment in which they are being used. Although changes within a single scene are evident, satellite records can reveal even greater variability within a single category of data. Varied filter zones, different growing seasons, and various climatic and illumination conditions are all examples of this. Using extra and diverse ground truth and fuse data information (e.g. from several research regions in parallel) could also increase general model correctness, resilience, and transferability, especially with the growing public provision of high quality fused record datasets. Overfitting can be mitigated by adding additional data augmentation such as randomised scaling, deformities, color offset or jitter, or the insertion of record noise to the data.

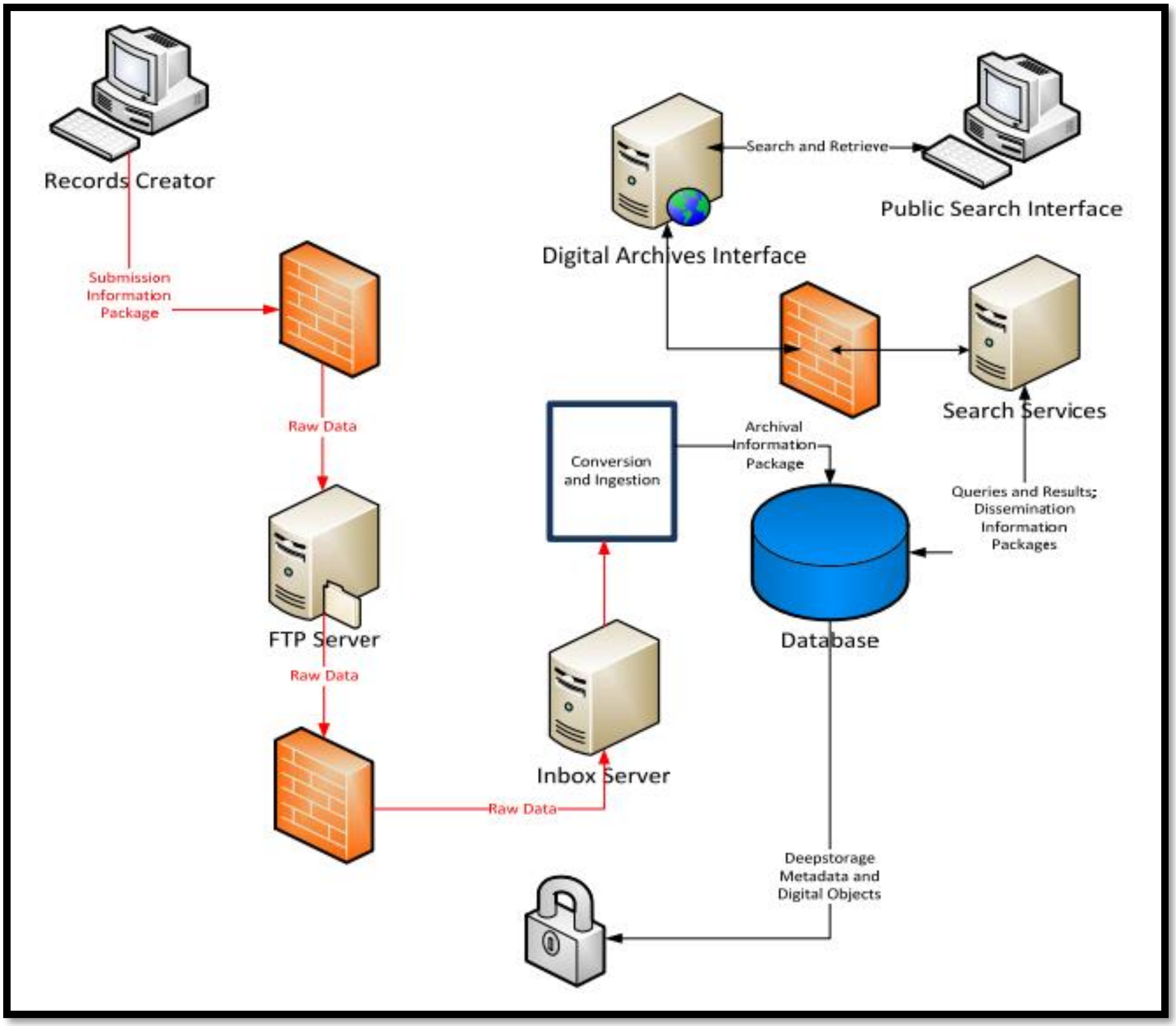


**Figure 3.6:** Functional overview of preservation system based on blockchain.

The government records will be used as a guideline in the following conversation to explore the most poignant results acquired using blockchains, as well as some astonishing new components. The top-5 error rate on the dataset of the nets is shown to offer an indication of the performance of these models. It's worth noting that the presented nets, training, and testing schemes include a slew of other hyper-parameters that were tweaked to get these results (e.g., a 7-network ensemble of blockchain-architectures).

Compiling the delineated features of government documents with boundary smoothing using the blockchain method and classifying records of the same class is what this research focuses on.

Semantic segmentation meets edge detection (which predicts the boundary and class for a large number of records but does not segment them) in feature extraction . A single instance of the blockchain algorithm on a field may be found using semantic segmentation if it was entirely surrounded by sections of distinct record classes.

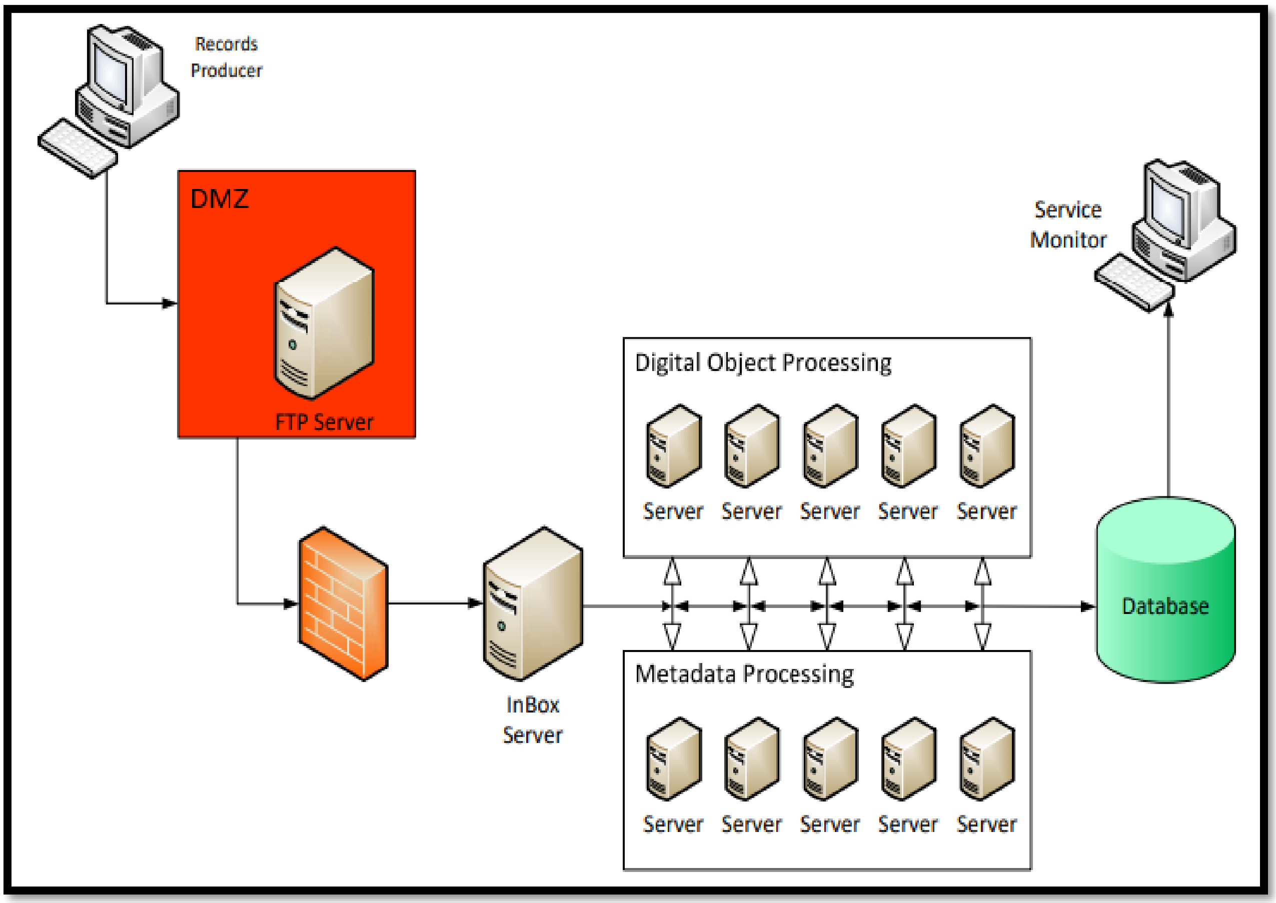


**Figure 3.7:** The process of government record object creation along with database.

These files are mainly government records derived through human field boundary tracing. Manual tracing of record edge smoothening, on the other hand, is highly reliant on the recorder and extra data used. It's also a very subjective task, with mistakes and ambiguities inevitable based on the operator's priorities. Existing property parcel information, which may be indistinguishable based on the recorder alone, can be added to the ground truth dataset. Furthermore, due to unreported land use changes during the growing season, a single field record edge smoothening can potentially reveal numerous subfields. Given these constraints, even a very powerful automated record processing algorithm can only become as good as the ground truth data but will never be in perfect alignment with reality. Although the model can generalize within certain limits (local generalization), its transferability is mainly limited by the variability of the training data. However, filter field are often directly adjacent to each other and have touching boundaries. In this case, edge segmentation would yield the record edge smoothening. Same filter instance segmentation is able to distinguish between these connected or overlapping instances of the same class. Automation has not reached mainstream adoption in the remote sensing community yet.

### 3.2.1 Preservation Of Government Record

The government record is basically a buffer of intensities per pixel. For this reason, the record at this stage is very dark and not suitable for a computer vision application. When the record is fusion, each pixel is mapped to a three-channel RGB representation with red, green and blue channels. The CNN algorithm determines the intensities of the other channels, based on information from neighboring pixels. This is a loss stage, as information in terms of bits per pixel is being lost. The government records have a huge tone, which is due to the Bayer mosaic pattern. In the Bayer mosaic, green is the dominant color as the color arrays in the record sensor are placed in rows of green-blue and red-green. As can be seen, color transformation corrects for that hue tone. The effect of gamma correction is not directly visible when compared to the output of the color transformation stage, but adds to the highlights and black tone appearance of the final result. The effect of the tone mapping stage is clearly visible, as it makes the record pleasant to the human vision, by fixing the brightness of different parts of the record. Furthermore, the blockchain’s training typically required many GPUs and was a time-consuming operation (up to 3 weeks). As a result, it is evident that training identical, high-performance blockchain designs on real-world data is nearly impossible in most circumstances. The blockchains that were taught to recognize the records dataset have a great ability to extract very distinct properties from natural records, making them applicable for other datasets. This is supported by the idea that the lower levels of these architectures extract general features from the records and become increasingly more task-specific deeper in the network. From this the idea of transfer learning sprung forth, which aims to kick start the process of the random weight initialization by using the pre-trained weights of deep blockchain’s to facilitate the training process.

### 3.2.2 Key Notions Concerning Government Records for Storage

The blockchain algorithm learning entails training a source network on a source dataset for a source task, transferring the learned knowledge (e.g. features or weights) to a new network for a different learning task and/or on a different target domain. Transfer learning is used with the objective of increasing performance on the target task by exploiting the trained knowledge of the source task. It can either serve as a 'Kick-starter' for the process improving initial performance, increase final performance or speed up the training of the target task. Three main types of transfer learning can be distinguished with respect to similarities in the source/target domains/tasks: (1) inductive transfer learning with the goal of inducing a generalizable, predictive model for the target task from a different source task; (2) transfer learning where a model should be derived in the target domain for the target task lacking labeled data, based on the source task containing a substantial amount of labeled data; and (3) unsupervised transfer learning which targets unsupervised learning in the target domain based on a different source task.

For purposes of record classification, the approach of inductive transfer learning has gained momentum in the past years. Given the large training time and requirements on labeled data, architectures are pre-trained on the records (source domain) for 1000-class classification purposes (source task) and knowledge is transferred to a labeled dataset of choice (target domain) for e.g. classification as target task. Thus, it primarily exploits the learned feature-representations in the source domain and aims to transfer these to the target domain with the aim of improving classification performance whilst making use of deep, data-greedy blockchain architectures.

The transferability of different levels of features was discussed in this study, with results confirming that the deepest layers are the most task-specific, and that employing characteristics from higher levels in the model yields superior outcomes. This was validated by the research, which also suggested that combining traits and pooling them together could boost performance. We added to the evidence by documenting higher performances for blockchain's initiate with pre-trained weights and fine-tuned for the task at hand, rather than through random weight initialization and training from scratch. Furthermore, they investigated the transferability of specific feature layers and reported co-adaptation of intermediate feature layers, where splitting between frozen and trainable layers at this level might result in a drop in performance.

### 3.2.3 List Of Development Issues

i. The trials show that data storage and automation techniques can be combined to provide predicting abilities that are superior to present clinical procedures.

ii. The issue of limited data and a large range of architectural options; their actual potential in this context has yet to be discovered.

iii. In the future, regardless of the blockchain model, it is possible that data analysis alone will not be able to deliver the best fracture risk prediction.

iv. The blockchain technique cannot be used with the objective of increasing performance on the target task by exploiting the trained knowledge of the source task.

v. The ground truth dataset can be complemented by existing property parcel information that could be indistinguishable based on the recorder alone.

## 3.3 DESCRIPTION OF DATA

The data was obtained from the e-Government Benchmark dataset, which is an open-source resource. The physical appearance of records can vary depending on a variety of physical and human geographical conditions, both on the ground and in government recorders. The cultivated species, plant status, topography, soil qualities, climatic conditions, cultivation methods, and other human impacts are all potential sources of diversity. The task of accurately distinguishing between many field records can become ambiguous due to such local variances. This makes it challenging to develop robust algorithms that can recognize a wide range of field cases without being overly reliant on the picture. Satellite recordings can reveal much stronger intra-class variances than variations within a single scene. This holds true for data collected from various agricultural regions, during various times and seasons, and under various meteorological and lighting circumstances. In addition, given the growing public availability of high-quality geo datasets, using more and more heterogeneous ground truth and satellite record data (e.g. from several study regions in parallel) could increase overall model correctness, resilience, and transferability. Additional data augmentation (e.g. random scaling, distortions, color offset or jitter, record noise introduction) may also help prevent overfitting and boost robustness to diverse geographic settings, as well as meteorological and lightning conditions. The data reference model is given in this link: https://digital-agenda-data.eu/datasets/e-gov

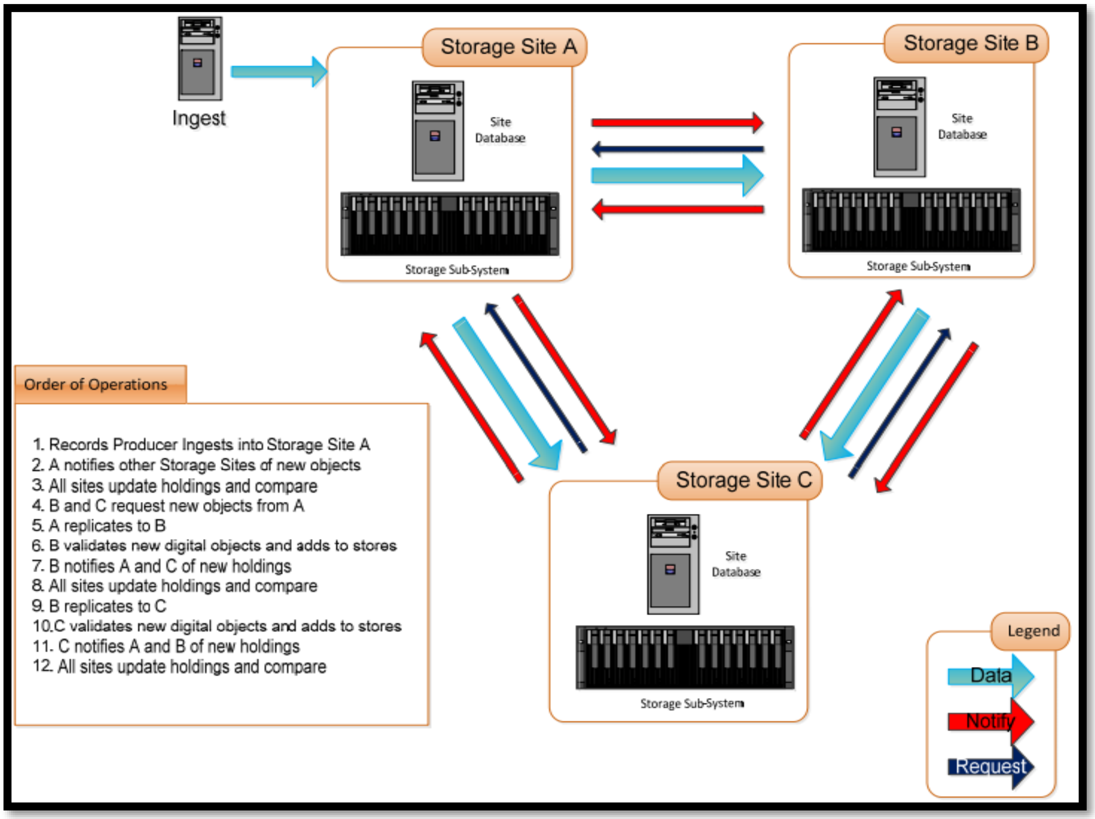


**Figure 3.8:** The final storage site replication process followed by record preservation.

# 4. RESULTS

The primary goal of this study is to create a smart structure for the government that will preserve the government's existing paper-based data. For a long time, the dangers of a digital world have been well-known to experts. It is the purpose of this project to discover new methods of informing academics about digital preservation efforts to enhance the long-term viability of their study results and government documents. As a result of this research, sustainable information packages (SIPs) were created for three residential research digital libraries, i.e. electricity and money for digitalization as the CNN method to digitization. The Access Archive Information (OAIS) paradigm is based on paper-based information bundles. World Customs Organization (ISO) standard OAIS is a framework for the emerging digital world in which paper-based data will be normalized and transformed into government records.

A digital format has been adopted by practically every civilization on the planet for mediated communications and knowledge. Digital transformation's consequences have been researched from various angles, but it may take generations to completely understand its influence. For four years, from 2015 to 2025, we introduced a new computer version to handle government records, and it has enhanced both the multitude of remarkable potentials for revelation and the take the chance of inadvertently removing priceless data resources from the repository of collective human awareness with a keystroke. Despite the rapid advancement of data storage technologies (hardware and software), individual and organizational behavior have been significantly slower to adjust in the new environment. While information management has become easier, data protection has become more difficult because of the task's scale and the fact that it is done in an unseen, abstract manner. Even while government records have historically used more complex networks, the architectural flexibility of the Blockchain algorithm and enhanced automation techniques will allow for greater efficiency. Classification approaches like neural networks and binary trees can also be used to represent data utilizing representations such as modeling, designating key detection features, or tracking contours. Because ml algorithms and filters require extensive training with a large amount of data before they can be able to segment records from each other, this is not a problem.

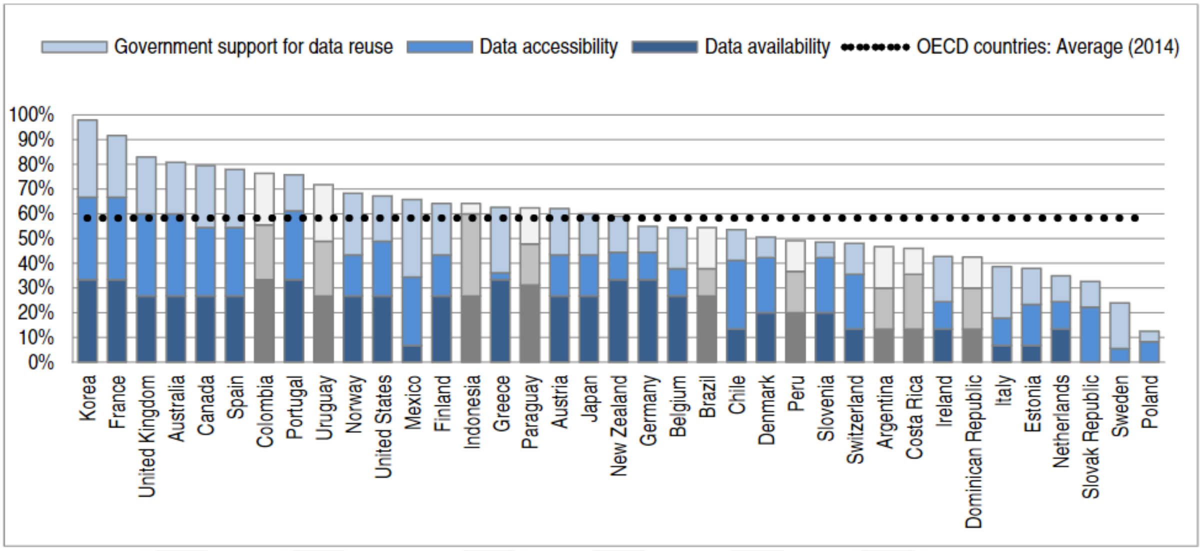


**Figure 4.1:** The prediction of e-government services as well as the minimum threshold of services being evaluated for different countries in terms of government support, data accessibility and data availability from database from percentage on x-axis with countries on the y-axis.

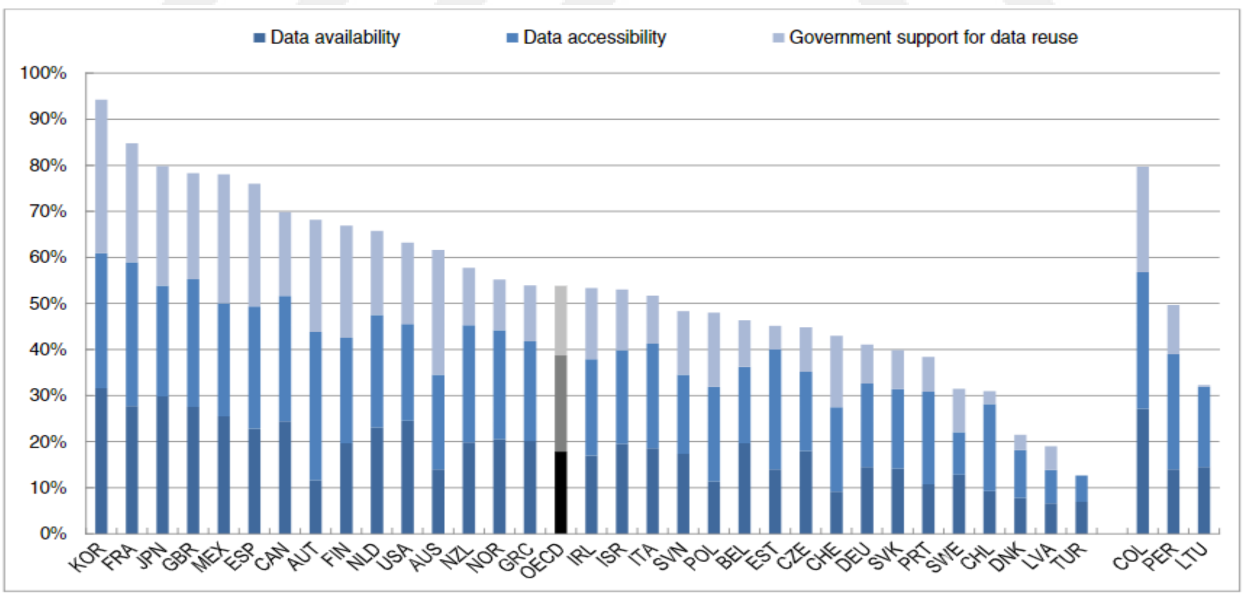


**Figure 4.2:** The prediction of e-government services without minimum services threshold being evaluated for different countries in terms of government support, data accessibility and data availability from database from percentage on x-axis with countries on the y-axis.

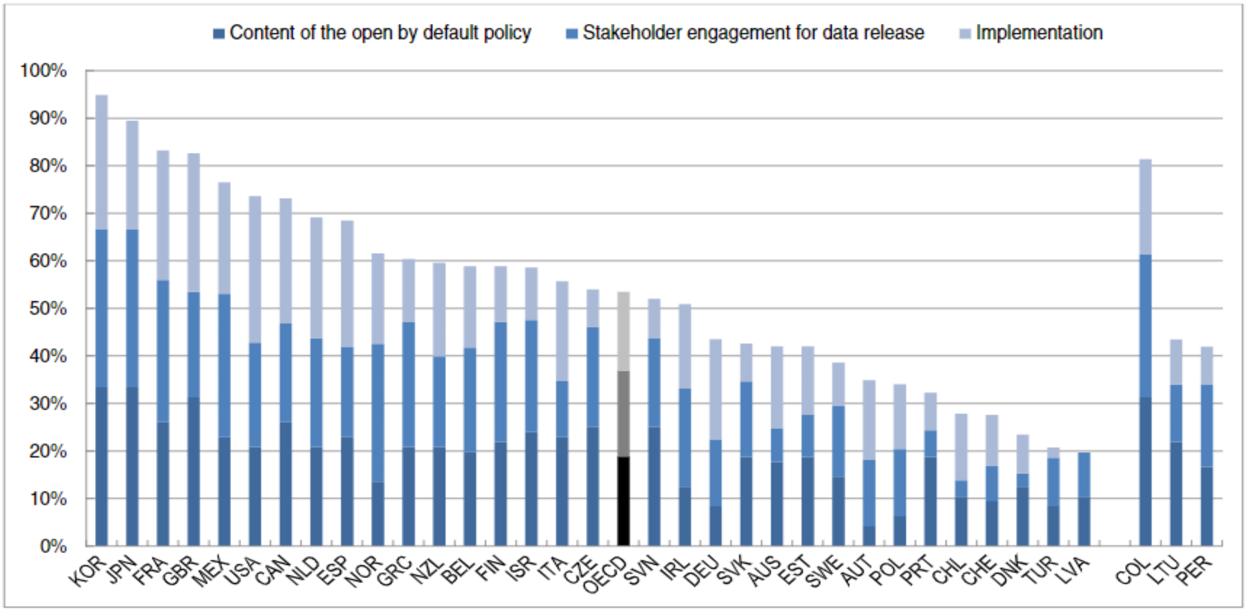


**Figure 4.3:** The prediction of e-government services being evaluated for different countries in terms of government restricted default policy, stakeholder engagement for data release and implementation of blockchain based on database where percentage is given on x-axis with countries on the y-axis.

For three distinct digital libraries, we rely on decisions made by the people who created the artifacts, the systems that manage them, the techniques that keep them running, and policies that provide the funding to keep them preserved. Digital artifacts have a monetary value that can only be determined by the researchers who created them. Preservation practices should then evolve to meet the needs of society because the internet world will rely on the preservation of our data. As the digital world becomes more fleeting, we've had to rethink how we save data. Using digital system analysis, DOI elements can be quantified for record digitalization.

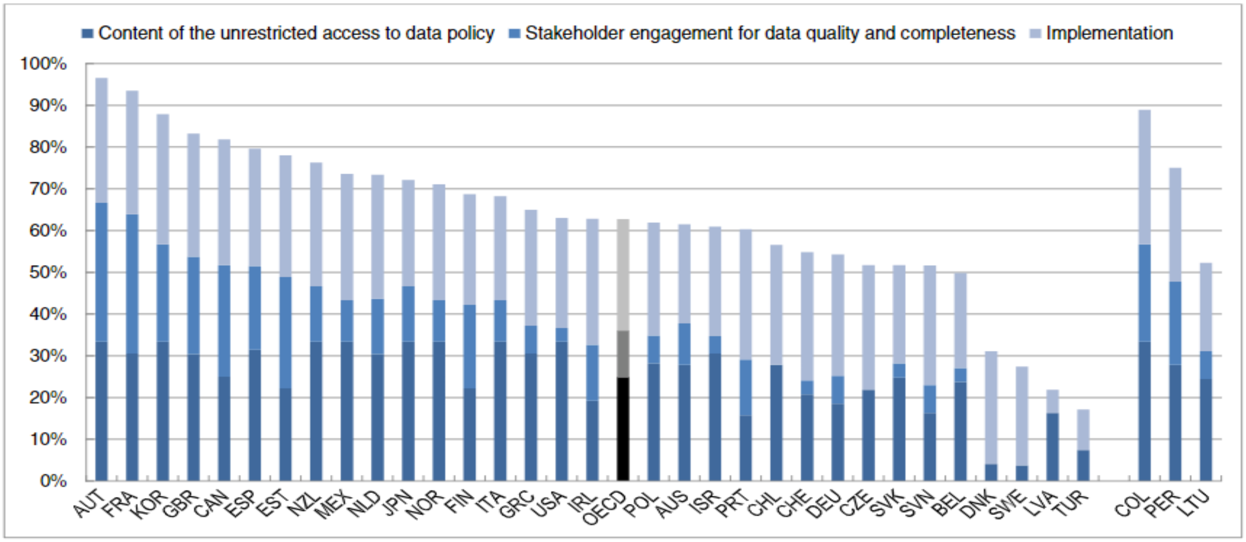


**Figure 4.4:** The prediction of e-government services being evaluated for different countries in terms of government unrestricted data policy, stakeholder engagement for data quality and implementation of blockchain based on database where percentage is given on x-axis with countries on the y-axis.

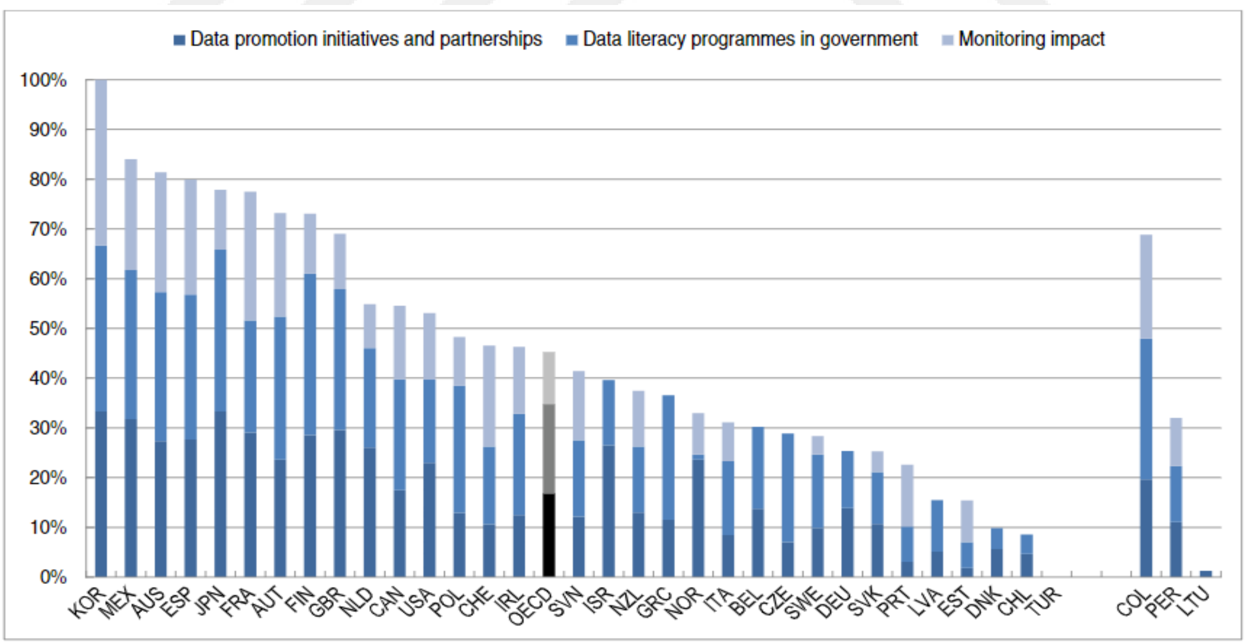


**Figure 4.5:** The prediction of e-government services being evaluated for different countries in terms of government data promotion partnership, data literacy programs in government and monitoring impact of blockchain based on database where percentage is given on x-axis with countries on the y-axis.

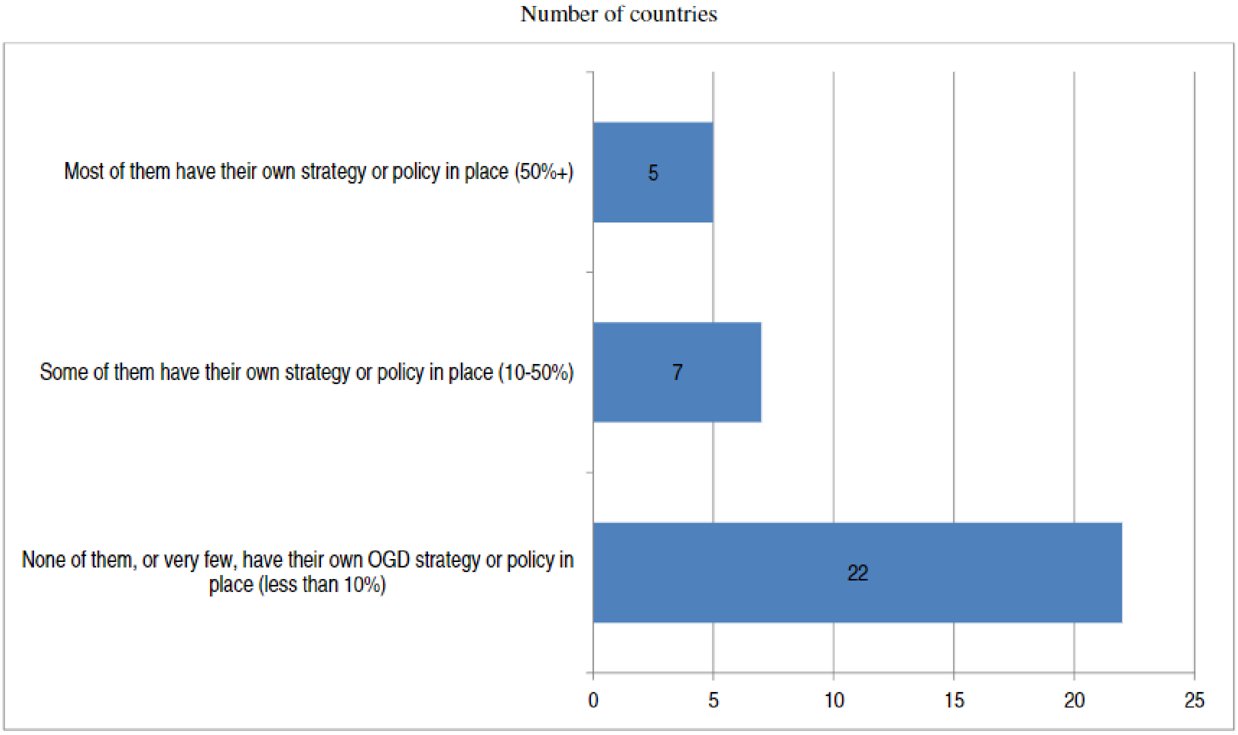


**Figure 4.6:** The discovery of government services as well as the policy and strategy based on the dataset from year 2015 until 2025 being evaluated by proof-of-work blockchain.

It was one of the goals of this research to find a way to influence robust digital restoration at the organization through correspondence with the knowledge that educators are greatly influenced within their field. Messages tailored to the department level are more likely to be successful than messages from the undergraduate and graduate level because faculty are assessed, supported, and gainfully employed at the department level. The scientists in the four components were asked to describe how they manage their someone else's digital data in order to better understand how the units interact with one another. The unit of analysis is the individual researcher. There are many interrelated processes in the preservation of digital process that can be represented by using a dataset and Activity theory. Digitalization's central tenet is the safekeeping and management of government records. The scope of this project, which involves digital preservation for three of the government's most important digital libraries, far exceeds the capabilities of the tools we use to access them.

Number of countries

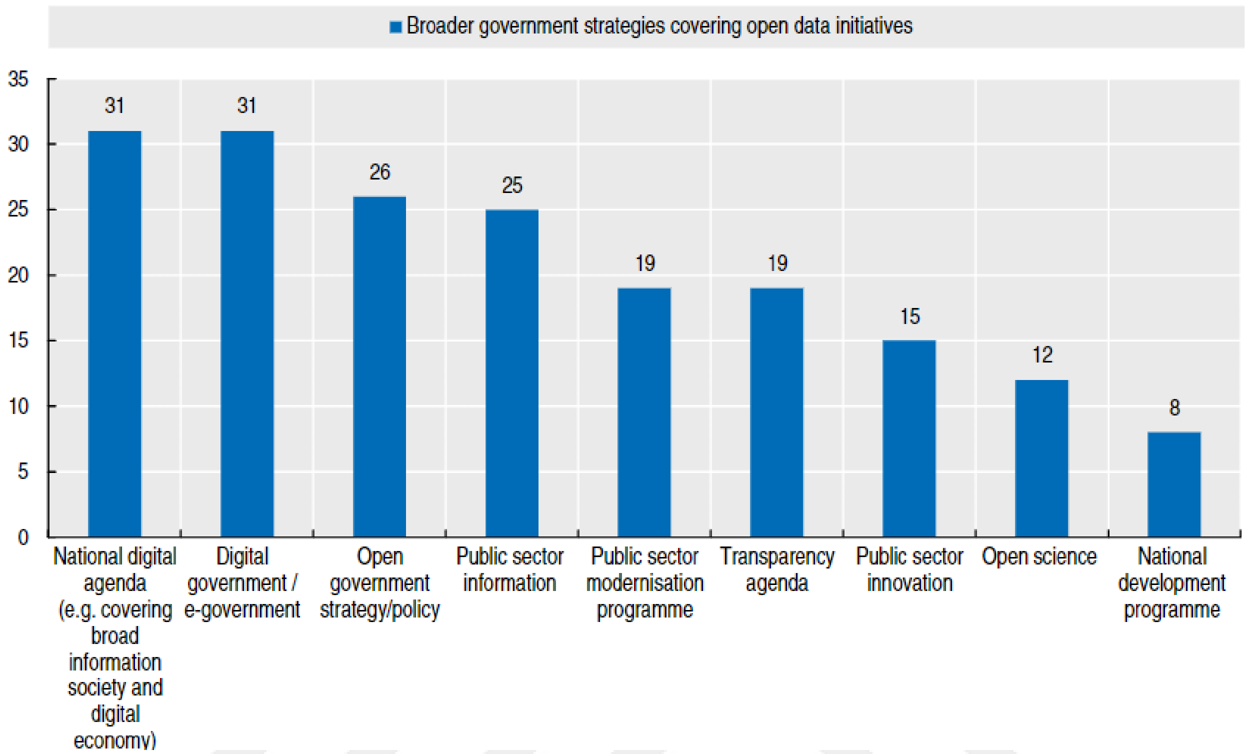


**Figure 4.7:** The broader government strategies covering open data initiative for different countries where we are limited only by the software that we have at our disposal to access it from dataset from year 2015 until 2025 for number of countries on x-axis with services on the y-axis.

Given these constraints, even a very powerful automated record processing algorithm can only become as good as the ground truth data but will never be in perfect alignment with reality. Although the model can generalize within certain limits (local generalization), its transferability is mainly limited by the variability of the training data. However, filter field are often directly adjacent to each other and have touching boundaries.

Number of countries

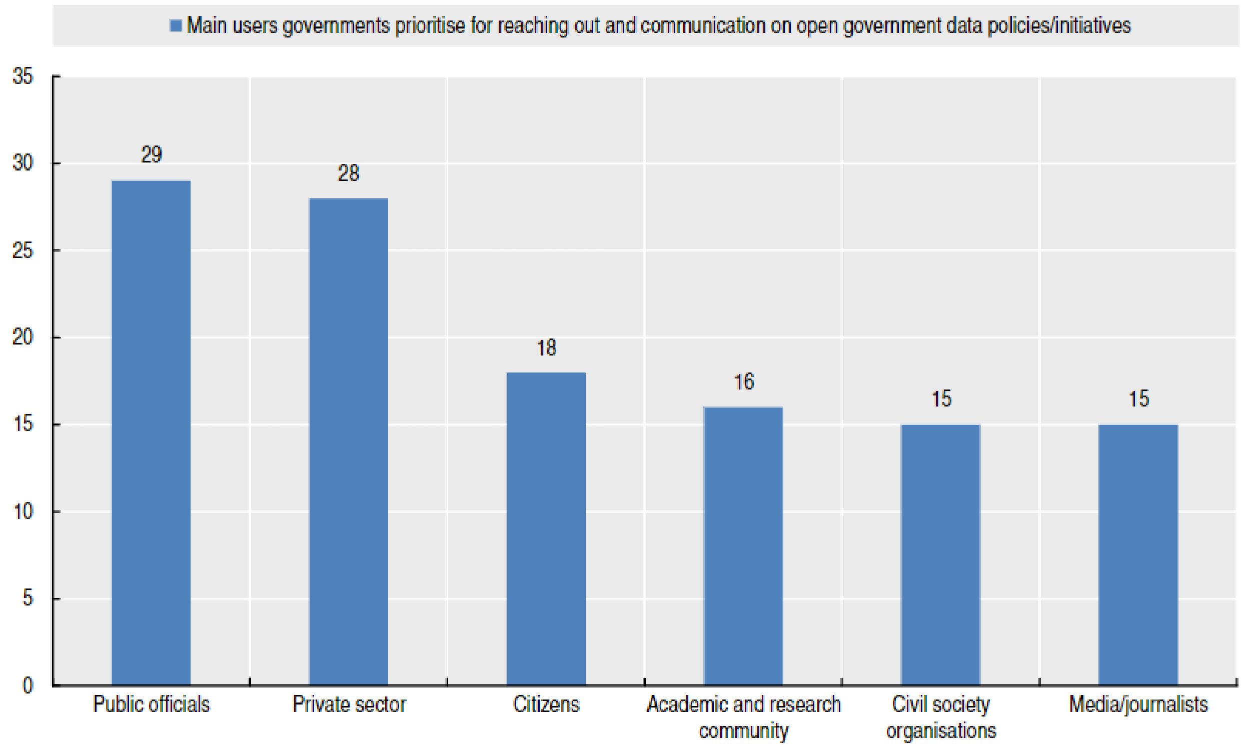


**Figure 4.8:** The main user's government priorities for reaching out and communication on open government data policies for different countries where we are limited only by the software that we have at our disposal to access it from dataset from year 2015 until 2025 for number of countries on x-axis with services on the y-axis.

# 5. DISCUSSION

Government record preservation and management is at the core concept of digitalization. It is being done for the 3 key government digital libraries, whose scope extends beyond the technology we use to access them. For 3 distinct digital libraries, we rely on decisions made by the people who created the artifacts, the systems that manage them, the techniques that keep them running, and policies that provide the funding to keep them preserved. Digital artifacts have a monetary value that can only be determined by the researchers who created them. Process is carefully must evolve to meet the needs of society because the digitalization will depend on the preservation of our data. As the digital world becomes more fleeting, we've had to rethink how we save data. For the digitalization of records, DOI is a well-established theory with elements that can be quantified through digital system analysis. State of the art thesis include many record-based identification and tracking approaches. Representation with modeling, specifying key features for detection, contours tracking, or by using classification techniques such as neural network and binary tree classification. Although machine learning algorithms and classifiers needs robust training with lots of data only then they are able to segment record from records. This approach needs lots of processing and hence computationally expensive. The more traditional method where we use record form depth information and try to find record from those records combined; is an ideal approach for this task. The E-Government dataset is one of several international (ISO) standards that attempt to resolve problems in an increasingly connected world. Focusing on the average best validation error, two statistics seem to correlate with its absolute improvement. On one record, the improvement seems to shrink with the number of classes and thus the number of original samples in the training set. As the number of selected classes is increased from seven to ten, the difference becomes smaller and even negative. However, even though relatively small, the classification accuracy is higher when the data augmentation is used for twenty classes as well as the whole dataset. On the other record, the data augmentation becomes more effective for subsets with larger maximum sequence length. While there was a similar amount of original training samples in the three five-class subsets, the maximum sequence length strongly varied.

The research has been designed solely for the government digitalization plan and activity clusters with respect to digital preservation activity. Government record activity theory provides a framework that includes rules, resources, tools and culture for the preservation of records. Both can explain the introduction and diffusion of digital preservation practices throughout the government by digitalization. The general comparison between the major archives of repositories with all journals for the efforts of records preservation digitally.

**Table 5.1:** The comparison for the preservation of records in digital form among existing digital libraries.

| ARTICLE | TECHNIQUE | DIGITAL LIBRARIES |
|---|---|---|
| [51] | Trustworthy Digital Object (TDO) | National Digital Information Infrastructure Preservation Plan (NDIIPP) |
| [52] | Byte Replication and Encapsulation | Information Communication Technology (ICT) |
| Proposed | Validation based Proof-of-Work Blockchain | E-Government Benchmark |

# 6. CONCLUSION

## 6.1 CONCLUSION

In this work, digital preservation is being performed for the three major digital libraries at governmental level, The scope of the project is not limited by the technology we use to move it from publication to digital format preservation. This void led directly towards the preservation of digital information. With this predictive algorithm, the files can be found in any computer database where they are stored, making it possible to recover paper-based data. Many government regulations now require that documents be stored in digital form for reporting purposes, as many other prosperous nations have updated their system using cryptocurrency. The majority of the records will be very well preserved from 2019 to 2023 as a result of our efforts., the goal is to preserve which most of the records it belongs to digital format. The distribution of records preservation for major archives in term of paper based and digitally preserved records from year 2015 to 2025. The preservation of government records was recorded very low in 2015 as much of the record archival were paper-based in all major repositories of the government. However, year-by-year the smart system tends preserve the records in digital format from old paper-based format and till 2025 large amount of records are being converted to the digital format with a whopping accuracy of 98.68% for all the records where the system was trained on 80% of data and tested on rest of the 20% of data. The 3 major facilities of government are being preserved in this work, whose scope extends beyond the technology we use to access it. Three different digital libraries' artifacts, management systems, maintenance techniques, and policy and funding are all based on decisions made by the people who created them. Scientists are the ones who generate the data; therefore, they bear the greatest responsibility for determining what digital artifacts are worth. It is imperative that preservation practices evolve to meet the societal needs of the digital age.

## 6.2 FUTURE WORK

The E-Government benchmark dataset is single of several international (ISO) standards that effort to resolve questions in an progressively more linked globe. Focusing on the average best validation error, two statistics seem to correlate with its absolute improvement.

i. In future, the improvement will shrink with the number of classes and thus the number of original samples in the training set. As the number of selected classes is increased from seven to ten, the difference becomes smaller and even negative.

ii. However, even though relatively small, the classification accuracy is higher when the data augmentation is used for twenty classes as well as the whole dataset.

iii. In future, the research will aim to perform data augmentation becomes more effective for subsets with larger maximum sequence length. While there was a similar amount of original training samples in the three five-class subsets, the maximum sequence length strongly varied.

iv. The largest absolute improvement was achieved for the classes where the maximum sequence length was the highest and it would be determined in future.

v. In future, the maximum length was equal in the experiments with seven, eight, nine and ten classes, where a shrinking improvement was observable with increasing training set size, the maximum sequence lengths are significantly larger when the experiments are repeated on the classes. In these cases, the data augmentation has positive effects again.

vi. As the maximum number of frames gets clipped to smaller numbers for the created sequences, the synthetic samples become less helpful for improving the classification accuracy on the whole dataset.

# REFERENCES


[1] D. Halkias, M. Neubert, P. W. Thurman, and N. Harkiolakis, “The Advantage of the Multiple Case Study Design for Management Researchers,” The Multiple Case Study Design, pp. 22–26, Feb. 2022

[2] P. Baxter and S. Jack, “Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers,” The Qualitative Report, Jan. 2015.

[3] H. Paxson, “Participant-Observation and Interviewing Techniques,” Food Culture, pp. 92–100, Feb. 2017.

[4] T. K. Bikson and S. A. Law, “Electronic information media and records management methods: A survey of practices in UN organizations,” The Information Society, vol. 9, no. 2, pp. 125–144, Apr. 1993.

[5] B. Chachage and P. Ngulube, “Management of business records in Tanzania: An exploratory case study of selected companies,” SA Journal of Information Management, vol. 8, no. 3, Dec. 2006.

[6] D. Wickman, “Recordkeeping legislation and its impacts: the PARBICA Recordkeeping for Good Governance Toolkit,” Comma, vol. 2011, no. 1, pp. 51–59, Jan. 2011.

[7] H. Demirtel and Ö. G. Bayram, “Efficiency of Electronic Records Management Systems: Turkey and Example of Ministry of Development,” Procedia - Social and Behavioral Sciences, vol. 147, pp. 189–196, Aug. 2014.

[8] L. Duranti, “Concepts and principles for the management of electronic records, or records management theory is archival diplomatics,” Records Management Journal, vol. 20, no. 1, pp. 78–95, Mar. 2010.

[9] F. Hsu, T. Chen, and S. Wang, “Efficiency and satisfaction of electronic records management systems in e-government in Taiwan,” The Electronic Library, vol. 27, no. 3, pp. 461–473, Jun. 2009.

[10] Harvard's Records Management Services. (2015). Controlling Costs and Promoting Efficiency. Available at: http://library.harvard.edu/sites/default/files/ControllingCostsAndPromotingEfficiency. pdf [Accessed 21 May 2021].

[11] H. Zinner Henriksen and K. Viborg Andersen, “Electronic records management systems

implementation in the Pakistani local government," Records Management Journal, vol. 18, no. 1, pp. 40–52, Feb. 2008.

[12] R. Cox, "Empty Temples: Challenges for Modern Government Archives and Records Management," Records & Information Management Report, vol. 22, no. 8, pp. 1–13, Oct. 2006.

[13] Johnston, G.P. and Bowen, D.V. (2015), The benefits of electronic records management systems: a general review of published and some unpublished cases. Records Management Journal, 15 (3), pp. 131 - 140.

[14] Keith, G., (2015). Implementing an electronic records management system. Records Management Journal, 15 (2). pp. 80 - 85 Available at: http://dx.doi.org/10.1108/09565690510614229 [Accessed 10 May 2021].

[15] Ki-Young, J., Phillips, D. T. (2015). Operational efficiency and effectiveness measurement. International Journal of Operations & Production Management, 21 (11), pp.1404 - 1416. Available at: http://dx.doi.org/10.1108/EUM0000000006223 [Accessed 12 May 2021].

[16] Mäkinen, S. and Henttonen,P. (2011),Motivations for records management in mobile work", Records Management Journal, 21 (3). pp. 188 - 204 Available at:http://dx.doi.org/10.1108/09565691111186867 [Accessed 13 May 2021].

[17] Meijer, A. (2016). Accountability in an information age: Opportunities and risks for records management. Archival Science.1(4), pp. 361-372. Available at: http://link.springer.com/article/10.1007%2FBF02438902 [Accessed 13 May 2021].

[18] Mokhtar, U.A. and Yusof, Z.M. (2019), Electronic records management in the Malaysian public sector: the existence of policy, Records Management Journal, 19 (3) pp. 231 -244 Available at: http://dx.doi.org/10.1108/09565690910999201 [Accessed 20 May 2021]

[19] Myers, M. D., (2017). Qualitative Research in Information Systems. MIS Quarterly 21(2), pp.241-242. [Online] Available at: [Accessed 10 May 2021].

[20] Myers, M.D. and Newman, M. (2007). The qualitative interview in IS research: Examining the craft. Information and Organization. 17 (1). Available at: doi: 10.1016/j.infoandorg.2006.11.001 [Accessed 31 May 2021].

[21] M. Bianchi, "Guide to Resource Management Records for Farms: Farm Water Quality Planning Series," Sep. 2002

[22] Patience Coetzer, X. (2020). The status of records management at the University of Zululand.

Master Thesis. University of Zululand. Available at: http://uzspace.uzulu.ac.za/bitstream/handle/10530/1219/THE+STATUS+OF+RECORDS+MANAGEMENT+AT+THE+UNIVERSITY+OF+ZULULAND.pdf?sequence= 1 [Accessed 1 May 2021].

[23] Porter-Roth,B. (2016). Applying Electronic Records Management in the Document Management Environment: an Integrated Approach. [Pdf] Xerox. Available at: http://docushare.xerox.com/pdf/docushare_RM_whitepaper.pdf [Accessed 20 May 2021]

[24] Quan-Haase, A. (2013). Theoretical Perspectives on Technology. Don Mills, Ontario: Oxford University Press. Richmond, H. (n.d.). Why Records Management? The Top Ten Business Reasons. pp.1-4. Available at: http://condar.ca/CONDAR%20Articles/article1-%20ten%20reasons.pdf [Accessed 11 May 2021].

[25] Reed, B. (2017),"Electronic records management in Australia", Records Management Journal, 7 (3). pp. 191 - 204. Available at: http://dx.doi.org/10.1108/eb027111 [Accessed 21 May 2021].

[26] Shaw,A. and Shaw, D.T. (2016). Electronic Records Management Criteria and Information Security. Proceedings of 7th Australian Information Warfare and Security Conference. Available at: http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1018&context=isw [Accessed 10 May 2021].

[27] Sprehe, T. J. (2015). The positive benefits of electronic records management in the context of enterprise content management. Government Information Quarterly.22(2).pp.297–303.Available at: http://www.sciencedirect.com/science/article/pii/S0740624X05000195S0740624X05000195 [Accessed 3 May 2021].

[28] Tambudzai Nengomasha , C. (2009). A study of electronic records management in the Namibian public service in the context of e-government. Ph.D. University of Namibia. Available at: http://wwwisis.unam.na/theses/nengomasha2009.pdf [Accessed 10 May 2021].

[29] Turner, D. W. (2010). Qualitative Interview Design: A Practical Guide for Novice Investigators. The Qualitative Report .5 (3). pp. 754-760Available at: http://www.nova.edu/ssss/QR/QR15-3/qid.pdf [Accessed 22 May 2021].

[30] Umi Asma' Mokhtar and Zawiyah Mohammad Yusof, (2009). Electronic records management in the Malaysian public sector: the existence of policy. Records Management

Journal, 19 (3). pp. 231 - 244. Available at: http://dx.doi.org/10.1108/09565690910999201 [Accessed 10 May 2021].

[31] R. O. Zerbe, "Ethical benefit–cost analysis as art and science: ten rules for benefit–cost analysis," Principles and Standards for Benefit–Cost Analysis, pp. 264–293.

[32] Yampolskiy, R. V. (2011). Efficiency Theory: A Unifying Theory for Information, Computation and Intelligence. [Online] Available at: http://arxiv.org/abs/1112.2127222 [Accessed 10 March 2021].

[33] Financial Reporting Council, The UK Corporate Governance Code. London: The Financial Reporting Council Limited, 2012. Available at: https://www.frc.org.uk/OurWork/Publications/Corporate-Governance/UK-Corporate-Governance-Code-September2012.pdf; accessed on 15th May 2021.

[34] Gollins, Tim, "Parsimonious preservation: Preventing pointless processes!", Online Information, 2009.

[35] Gough, Roger (ed.), "An End to Sofa Government: Better working of Prime Minister and Cabinet", Conservative Democracy Task Force, 2007. Available at http://image.guardian.co.uk/sysfiles/Politics/documents/2007/03/27/DemocracyTaskForce.pdf; accessed on 20th May 2021.

[36] S. A. Buchanan, C. Stratton, Y. Sun, and A. Chaudhary, "Survey research on tasks and competencies to inform records management education," Records Management Journal, vol. 27, no. 1, pp. 2–18, Mar. 2017.

[37] International Records Management Trust (IRMT), "Consultancy Projects". Available at: http://irmt.org/consultancy-services/consultancy-projects-2; accessed on 20th May 2021.

[38] Jisc infoNet, "Guidance notes for the Records Management Maturity Model." Available at: www.jiscinfonet.ac.uk/records-management/measuring-impact/maturity-model; accessed on 25th May 2021.

[39] Lemieux, Victoria, "The records-risk nexus: exploring the relationship between records and risk", Records Management Journal, Vol. 20 No. 2, 2010, pp. 199-216.

[40] V. L. Lemieux, "One Step Forward, Two Steps Backward? Does E-Government Make Governments in Developing Countries More Transparent and Accountable?," Jan. 2016.

[41] McKemmish, Sue, "Traces: Document, record, archive, archives", in McKemmish, Sue et al. (eds.), Archives: Recordkeeping in Society. Wagga Wagga: Charles Sturt University, 2005,

pp. 1-20

[42] Lipchak, Andrew, "Information management to support evidence-based governance in the electronic age ", a Public Policy Forum Discussion Paper, November 2002.

[43] McLeod, Julie et al, "A Critical Assessment of Records Management Capacity and Compliance Toolkits", Final Research Project Report. Newcastle: Northumbria University, 2006.

[44] McLeod, Julie, Sue Childs and Rachel Hardiman, "AC+ erm – Accelerating positive change in electronic records management", Final Research Project Report. Newcastle: Northumbria University, 2010.

[45] McLeod, Julie, Sue Childs and Susan Heaford, "A Critical Assessment of Records Management Capacity and Compliance Toolkits", Final Research Project Report. Newcastle: Northumbria University, 2006.

[46] National Archives, "About the National Archives – National Archives History". Washington: National Archives and Records Administration.

[47] National Records of Scotland, "Implementation of the Public Records (Scotland) Act". Edinburgh: Public Records (Scotland) Act 2011 Implementation Team, 2012.

[48] Perkins, Eric, "Records Management Training: Introduction to Records Management". Chester: Virginia Information Technology Agency (VITA), 2010.

[49] Share, Felex. "Education Records Management in Shambles." The Herald, 13th January, 2012.

[50] University College London, "University College London Records Management Policy". London: University College London Library Services, 2011.

[51] Gladney, Henry. (2006). Principles for Digital Preservation. Commun. ACM. 49. 111-116. 10.1145/1113034.1113038.

[52] MATLALA, Emanuel. Long-term preservation of digital records at the University of KwaZulu-Natal archives. Journal of the South African Society of Archivists, 2019, 52: 95-109..