Journal of Infrastructure, Policy and Development 2024, 8(7), 3520.
https://doi.org/10.24294/jipd.v8i7.3520

Article

Prediction of accessibility testing using a generalized linear model for e-government

Ahmad Althunibat[1], Wael Alzyadat[1], Siti Sarah Maidin[2], Adnan Hnaif[1,*], Basem Alokush[1]

[1] Department of Software Engineering, Faculty of Science and IT, Al-Zaytoonah University of Jordan, Amman 130, Jordan
[2] Faculty of Data Science and Information Technology (FDSIT), INTI International University, Nilai 71800, Malaysia
*** Corresponding author:** Adnan Hnaif, adnan_hnaif@zuj.edu.jo

CITATION

Althunibat A, Alzyadat W, Maidin SS. (2024). Prediction of accessibility testing using a generalized linear model for e-government. Journal of Infrastructure, Policy and Development. 8(7): 3520. https://doi.org/10.24294/jipd.v8i7.3520

ARTICLE INFO

Received: 4 December 2023
Accepted: 5 January 2024
Available online: 30 July 2024

Abstract: Using the United Nations' Online Services Indicator (OSI) as a benchmark, the study analyzes Jordan's e-government performance trends from 2008 to 2022, revealing temporal variations and areas of discontent. The research incorporates diverse testing strategies, considering technological, organizational, and environmental factors, and aligns with global frameworks emphasizing usability, accessibility, and security. The proposed model unfolds in three stages: data collection, performing data operations, and target selection using the Generalized Linear Model (GLM). Leveraging web crawling techniques, the data collection process extracts structured information from the Jordanian e-government portal. Results demonstrate the model's efficacy in assessing accessibility and predicting web crawler behavior, providing valuable insights for policymakers and officials. This model serves as a practical tool for the enhancement of e-government services, addressing citizen concerns and improving overall service quality in Jordan and beyond.

Keywords: e-government; predict; accessibility; generalized linear model; accuracy; sustainable growth

1. Introduction

In 2001, the first rating of e-government was created. It was followed up every two years and achieved eleven rankings through 2022. There are many challenges for e-government including public values where citizens refuse inadequate information and process, the barriers in the sophisticated stage faced by citizens during using e-government conducts are lack of resources, quality, and delivery services online (Al-rawahna, 2018, p. 2; Althunibat et al., 2022, p. 6) that leads to the question how can the primary components of e-government be made reliable? We respond to this question by obtaining data from e-government, driving the analysis of the data, and evaluating the quality features that are reflected in the trustwor-thiness of citizens (Alhroob et al., 2020; Meiyanti et al., 2018; Norris et al., 2012). Obviously, United National (UN) proposed measurement called Online Services Indicator (OSI) which able to scale the e-government based on several factors, the survey report from UN the one of countries is Jordan's e-government that showed the OSI from 2008 the score of 0.605 degrees by OSI; in 2010 and 2012. As a consequence, the OSI score declined by −0.072 and −0.141 degrees, respectively. However, the scale increased by 0.127 degrees in 2014 before falling back to a score of 0.63 in 2016. 2018 had an increase in OSI of 0.037 degrees (Althunibat, 2014; Manoharan, 2018). Finally, the pandemic covid-19 outbreak in 2020 was a time of crisis for both the public and policymakers.

Notably, Jordan's e-government's score fluctuates negatively in terms of its ability and willingness to serve its citizens. To provide based on comparisons of find and identify content and indexing pages from Jordanian e-government applications, to derive classification collected data to recognize between validate services via concern, correct-ness, and completeness, and to encourage an understanding of why the OSI indicates citizens dissatisfied with a long-term vision. Moreover, applying a generalized linear model that anticipates well established skills for high quality services in response to citizens' requests, we examine the e-government system and suggest a test strategy.

The next subsection focused on the e-government test strategies emphasis to achieve quality services. The second subsection delves more into the investigation that has been done on e-government services as well as conceptual frameworks and models that have been tried and evaluated.

2. Literature overview

2.1. Test strategies of e-government applications

To prevent the failure of e-government services as a means of modernizing the public sector, the dependability, credibility, and trust that promotes institutions and research to expand greatly to realize the objectives of e-government testing (Althunibat et al., 2022; Twizeyimana and Andersson, 2019). For e-government, it is very crucial to comprehend the test strategies because they are important categories that establish the purpose, influencing factors, and constraints, the purpose category describes the result using measures such as performance and acceptability referred to as the main goal of e-government.

Influencing factors category consists of (1) technological, (2) organizational, and (3) environmental, which have an impact role to reduce inconsistencies and conflicts; constraint category adjusting relies on the influencing factor restrictions imposed demanded. An example, the organizational constraint is imposed by laws, standards, or regulations from policymakers. According to Forum (2019), argue for the information flow from e-government to citizens concern to challenge how the irrelevant sources specifically when use social media and researchers (Falco and Kleinhans, 2018) supported to consider to data flow for security perspective to assess the risk by proposed hierarchical model for Saudi e-government that integrate De Militarized Zone (DMZ) with block chain implemented through Yesser website to addressed vulnerability categories.

A study by Assiri et al. (2021) conducted on usability based on the World Wide Web Consortium (W3C) used optimization tools, accessibility, and security attributes for Kyrgyz Republic e-government which meagre of qualified specialists, particularly in the field of security into information aspect. On other hand, the accessibility evaluated was concerned by Ismailova (2017) and Moya et al. (2023), the infarction indicates that the Jordanian government website did not fulfil the minimal requirements for being considered accessible for any significant impairment groups.

The next five stages are described by the impediments to e-government that are ostensibly in technical and nontechnical areas, such as the several problems Iran's e-government encountered. First, there are no implementation instructions that appear

to have a negative impact on a technological level, particularly on architecture interoperability, data standards, legacy systems, and technical standards. The strategy level has con-ducted funding goals and objectives. As a result, privacy is not addressed at the policy level for citizens and industry, to summarize all level hurdles at the organizational level (AL-Naimat, 2013). To advance theoretical knowledge and assist government officials and agencies in better positioning their strategies to ensure the success of e-government projects by in-creasing the likelihood that the employed e-government services will be used, researchers (Zhang et al., 2014) investigated the proposed citizen centric model through literal review and related work. To design e-government systems where users are satisfied based on usability, accessibility, and security, it is important to comprehend how it is used and to discover elements that are influenced by supply and demand with adoption from citizens' points of view.

As indicated in **Table 1**, developing countries including Saudi Arabia, the Kyrgyz Republic, Jordan, Iran, and Egypt adopt e-government systems. It was variable in its quality. The accessibility measures cross-issue among those countries, for example, refers to literature reviews as a state of the art, quantity, quality, or systematic review, and they tried to improve the proposed guide by adopting different perspectives, such as tailoring the accessibility standards from W3C institutes to provide services for citizens with various knowledge and backgrounds to achieve the satisfying results. The focal point is to assess quality attributes for e-government online services or exchange of knowledge such as availability, security, usability, performance, and accessibility, whatever the digitalization channel (Mobile) and the future technology-driven growth. They require to lean towards validation and verification of services to contribute an iron triangle (Technology, Organizational, and environmental) parallel with cutting-edge technology as UAE Government launched a strategy entitled UAE Strategy for Artificial Intelligence (AI) covered nine sec-tors such as healthcare, education, technology, water, etc. One of the aims is an integrated smart digital system (Bigdeli and Cesare, 2011).

Table 1. Reviewed test strategies of e-government studies.

Ref	Consider to	e-gov	Category	Measures
Assiri et al., 2021	Data flow and process	Saudi	Technology adopted with blockchain	Security
Ismailova, 2015	Improve usability, accessibility, and security	Kyrgyz republic	Technology evaluation tool Organizational define priority	Usability, accessibility, and security
Abuaddous et al., 2019	Improve accessibility for impairment groups	Jordan	Technology of web content (information, web documents)	Accessibility
Bigdeli et al., 2011	Attempts to rank the barriers	Iran	Organizational (policy, vision, objectives) technology (data standard, architecture) environmental (citizens, business)	Recognizing the issues with e-government services
UAE National Strategy	Propose a citizen-centric e-government adoption model	Egypt	Technology (e-services) Organizational (developer, citizens)	Usability, accessibility, and security

Government services are changing because of technological advancements in ma-chine learning, AI, and big data. The details of how e-government services work in con-junction with cutting-edge technology are provided in the next sections.

2.2. E-government frameworks

The researcher Al-Mushayt (2019) and UAE National Strategy for Artificial Intelligence (2031, 2018), suggested a framework that uses Artificial Intelligence applications to accomplish two aims. Automate e-government services first, and then facilitate thinking about gulf e-government in the second services is the information resources, automated e-government services, and smart e-government platform architecture, the framework is separated into three stages. Using AI, Big Data, and the Internet of Things; an information resource proposes an architecture for the centralized management of e-government information resources, the big data services center oversees all big data policies and procedures (collection, storage, processing, and transmission) (Abdallah et al., 2022). Moreover, automating the role of services introduces several deep learning models, including those for Arabic sentiment categorization, handwritten digit identification, and handwritten letter recognition. Lastly, the development of a cutting-edge platform that enables a smart web for better e-government transactions and services may be facilitated by incorporating semantic web, multi-agent systems, autonomic computing, and AI techniques (Alhroob et al., 2020).

The Czech Republic had the highest value of artificial intelligence in Europe in 2021, with a value of 39.74, followed by Austria with a value of 36.66 and Greece with a value of 33.86. In 2021, the total value of artificial intelligence throughout all of Europe was 25.13 (Leogrande et al., 2022).

The authors (Alexopoulos et al., 2019) debate to question How Machine Learning is changing E-Government; answer the inter-section of three pathways big data, classification, and decision making; can be used to analyze big data, including government data, or even to generate new knowledge, while classification problems can be met. Despite the variety of methods for resolving categorization issues. Yet, the performance of machine learning models can be increased by experimenting with various strategies, enhancing their cur-rent services, but the actions are directed toward decision making between different do-mains.

The various frameworks/models from multiple countries are summarized sequentially in **Table 2** Greek, the first three rows, Estonians, Taiwan, and Turkey which consolidated them to the services of e-government. The citizens' acceptance of the services is measured by abundant indicators, such as accessibility that conducts the right services with the correct information with guidance or procedural. The FLOM framework's corner-stone for sharing information uses a sharing mechanism by partitioning data using ma-chine learning to transform Greek E-Government into share knowledge (Linders et al., 2018), as well X-Road data exchange for Estonians E-Governments to define the requirement of government 3.0 (Robles et al., 2019) alongside the COBRA purposed life cycle turkey e-government through validation and verification phase by the stages are test, refined, and validate for e-services to define the weakness, strength, opportunity, and threat. The benefits of COBRA help to be update and manipulate services based on citizens need and desire such as ERDSS framework (Osman et al., 2014) and UTAUT-based model (Almaiah et al, 2020) and Goodwin, 2012) and several research on e-government services, including proactive e-government, which is especially prevalent in e-Taiwan, and the e-GIF approach to

Table 2. Relatives e-government frameworks.

Frameworks of e-government	Purpose & focus	Conduct	Ref
FLOM	The machine learning approach enables collaborative learning among many parties and their data, whether private or open, by training models on such partitioned data without sharing it between parties throughout any stage of the training or inference process, thereby establishing shared knowledge.	Considering machine learning's use of the learning process for iterative model aggregation through parallel data partitioning for prediction and machine learning models analysis	Emanuel et al.
e-GIF approach	To enhance public administrations' electronic cooperation and communication Include three levels: • Systems level: Certificate Tools. Services process, Xml Schema, Web Services Repository, and System References • Standards level: Enterprise Architecture, Information Architecture, Computational Architecture, Standards (Data, Technical, and Authentication), and Web Portal. • Vision level: Vision, Goal, Coordination, marketing, management, guidelines	Specify the standards, specifications, and guidelines that should be followed when developing and deploying web-based front and back offices.	Charalabidis et al., 2008
Greek E-Gif ontology modeling	To be able to provide licenses in accordance with Directive 123/2006/EC, the ideas and relations that are utilized to arrange the data in the ERMIS Greek portal for public administration must be modeled with the relevant data. Focus on how linked data applications are implemented.	Services, document information systems, government entities, and web services are the layers that arrange the information for public administration.	Fragkou et al., 2014
X-Road data exchange	Estonians e-government for facilitating data exchange between public and private sectors using Open-Source Software. Define requirements for Government 3.0 solutions. Aim to data sharing infrastructure	Gather information (Public documents: Publicly available)	Robles et al., 2019
Flipping the service delivery model "Proactive e-Governance"	Network infrastructure. • Integration of "one window" services and online government services (e-Taiwan). • Citizen-centricity "putting citizens in the center". • Data-driven personalization. • Civil servants. • Range of channels (m-Taiwan & u-Taiwan)	To switch from a strategy of proactive, integrated service and information delivery along with data-driven personalization to passive and static information repositories.	Linders et al., 2018
COBRA	For Turkey e-government services regarding to: • Develop • Test • Refined • Validate Proposed A mathematical model to show the relationships between the constructs, associated variables, and user satisfaction	Cost, Benefit, Risk, and Opportunity, the primary goal is the success of e-services measured by key performance indicators (KPIs), cost-benefit for user perspective in Turkey, based on SWOT qualitative.	Osman et al., 2014

develop E-Government 3.0 and smart cities. A semantic web-based administration that adapts all its services to the requirements and preferences of each citizen in conjunction with blockchain, big data, and artificial intelligence technologies may be referred to as "E-Government 3.0". The study model is motivated in the next section utilizing three subsections of operations and data collection methods, followed by raw data to describe the correlations between attributes and the application of the GLM

algorithm as a machine learning approach. The results are then presented, followed by a discussion and recommendations.

3. Materials and methods

The evaluation of Jordanian e-government services, outlined in our proposed model, stems from the analysis of website pages. Despite the existence of the Government Website Standards (GWS_2019-Mode-v14 0-En) released by the Ministry of Digital Economy and Entrepreneurship of Jordan, focusing on five indicators—accessibility, usability, content and site architecture, responsive design, and security—our research specifically targets the assessment of the accessibility indicator. This involves scrutinizing all website pages within the Jordanian E-Government, ensuring citizens have equal access to information and services (Althunibat et al, 2024; Althunibat et al, 2021; Althunibat et al, 2011; Almaiah et al. 2020).

Our model, illustrated in **Figure 1**, delineates three integral stages: data collection, perform data, and select target (generalized linear model). Each stage encompasses distinct adoption methods and algorithms. **Figure 1** provides a visual representation of the model, where rectangles with a white background symbolize the main stages—data collection, performing data, and selecting a target (generalized linear model). The blue-rounded rectangles within each stage signify the methods and models employed, including the crawling method, index content, and predict model, with an emphasis on their intersection capabilities during the implementation phase utilizing an Object-Oriented Paradigm (OOP).

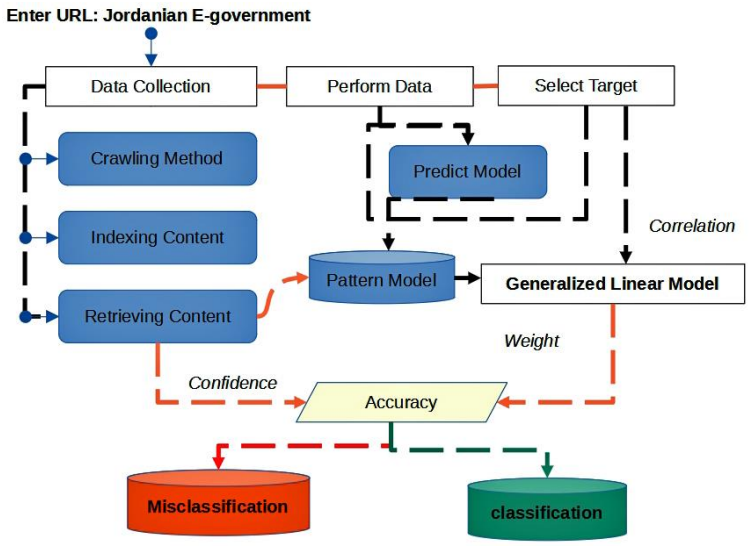


Figure 1. Proposed model to evaluate the services of Jordanian E-Government.

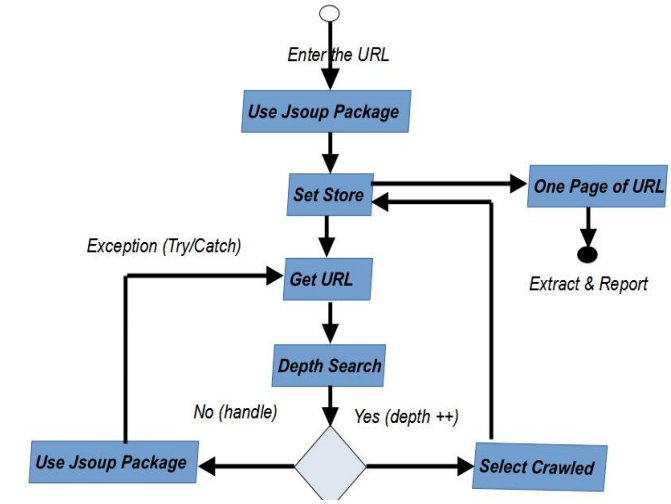
Furthermore, the yellow parallelogram shape represents a crucial model metric—accuracy—calculated to determine confidence and weight. Finally, the output, depicted as a magnetic disk, classifies outcomes into two categories: red for misclassification and green for classification. This comprehensive model strives to address accessibility concerns and equalize access to information and functionalities for all users on the Jordanian E-Government platform.

3.1. Data collection

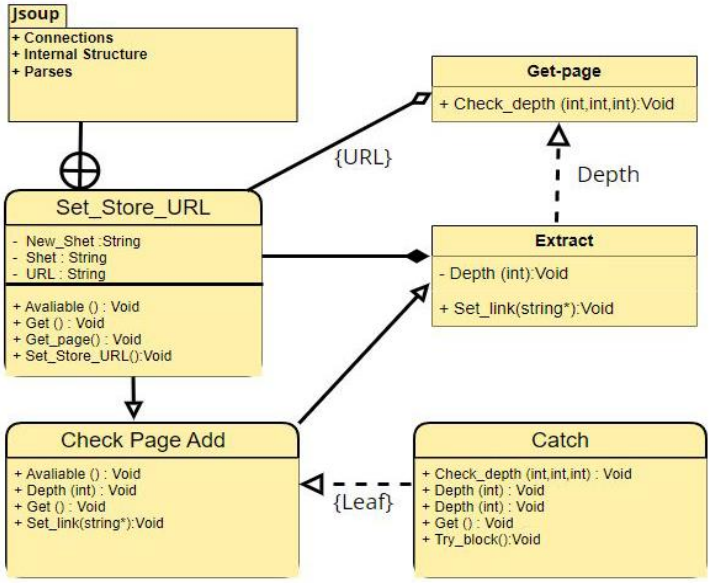
The process of gathering, extracting, and storing the content index of the URL (Jordanian E-Government) is known as the data collection stage. This dataset is maintained as shown later in Box 1. To gather structured and context sensitive data for the dataset, we use a web crawling approach.

Figure 2: (a) The flow chart diagram presents the main function for data collection and the process from enter the URL of Jordan E-Government to generated raw data; (b) the Class diagram shows the attributes, method for each class with visibility types where the relationship among these classes belongs to object-oriented design.

The declaration for operation from **Figure 2** is expressed in the code (presented in Box 1) by parsing Html into the documents fetching documents using library, set to storage data into two-dimension tables (rows and column) using jtable class.



(a)



(b)

Figure 2. (a) flow chart for data collection; **(b)** class diagram for data collection.

Box 1. Code of the declaration for operation.

```
//parsing HTML into A Document
Element content = doc.getElementById("content");
Elements links = content.getElementsByTag("a");
for (Element link : links) {
String linkHref = link.attr("href");
String linkText = link.text();
//Fetch
Document doc = Jsoup.connect("https://portal.jordan.gov.jo/wps/portal?lang=en/").get();log(doc.title());
Element newsHeadlines = doc.select("#mp-itn b a");
for (Element headline: newsHeadlines) {log("%s\n\t%s", headline.attr("title"), headline.absUrl("href"));}
//Set HTML to element
Element div = doc.select("div").first();
div.html("<p>lorem ipsum</p>");
div.prepend("<p>First</p>");
div.append("<p>Last</p>");
Element span = doc.select("span").first();
span.wrap("<li><a href='http://example.com/'></a></li>");
```

Requesting to write the data into a table using two dimensions after retrieving the URL (row and column). Parsing defines two elements. First element, that is content, is in-tended to read document with a specific id as a unique value, the second element is the link where define the tag for loop, define the link current document and links present an external resource to acquire the relationship between those attributes. In Fetch method in-clude create new connection then define element newsHeadlines contain the select docu-ment id of <a….> div tag, forward to validate the fetching data while monitoring and collecting by wrap method to be derived to set method. The storage data defined in the code (presented in Box 2) by set_store through the created two-dimension table using Jtable class.

Box 2. Code of storage data.

```
Public set_store{
table =new Jtable (new DefualtTableModel(new object []{},
new string []{"URL"})
Public boolean isCellEditable(int row, int column)
{returen false;});
Jpanel MatchesPanel1= new Japanel1();
MatchesPanel1.SetBorder(BorderFactory.creattitleBorder("Matches"));
MatchesPanel1.setLayout(new BorderLayout());
MatchesPanel1.add(new Jscrollpane(table), BorderLayout.CENTER);
```

The scope of the research in the portal entitled Jordanian E-Government; this portal has two languages, English and Arabic. The data collection selects the English language and represents the structure of data collection (shown in **Table 3**) as follows 15 attributes (columns) and 92,029 rows (records).

The pairs values of each attribute level and missing indicates the impact of the attribute on an overall dataset. Furthermore, assign the prediction accuracy through the classification error criterion as shown in the **Table 3**, the missing for these attributes pro-duce distorted class distribution by vertical action (column) to conduct those values. For example, the size bytes include 907 levels and the 1% missing value, which in simple observe the low negative influence indicated by 0.011% as well as the

link path and type attributes. Consequently, the middle negative influence approximate rate (between 0.2% and 0.8%) will conduct attributes are path type, link origin, follow, and link position. The high negative influence is covered by Anchor (3.57%), Target (236.17%), Alt Text (12%) and Rel (610%).

Table 3. Summary of Jordanian e-government dataset (reliable data).

Attributes	Description	Levels	Missing
Size Bytes	The set of header content length (KB)	907	1
Alt Text	The text is tag for image	83	99,707
Status Code	Consider to HTTP with responds code	7	1
Status	Like status code but only to header	13	5
Follow	Determine the direction of two paths are True or discretize	2	1
Rel	Indicate the relationship between URL elements	2	122,130
Path Type	References attribute of the link	5	1
Link Position	link located in the code	6	5
Link Origin	Recognize through HTML HTTP	3	1
Source	The root data	786	-
Destination	Indicate the interlink via URL	2731	-
Anchor	The text use it presents by Unicode	1164	41,582
Target	The page when open based on security and performance	5	118,086
Link Path	Link positions within page	3283	5
Type	Attribute types of links	8	1

The idle attributes are source and destination both are zero missing. To obtain pre-diction model, we need to solve the missing problem attributes by preprocess component, the preprocess wide board techniques based on related works, we adopted the stability technique (Al-zyadat et al., 2021) as expressed in the equation.

$$\text{Stability} = \sum i\ (\text{NonMissing})/N \quad (1)$$

where i present the value of attributes, which is 15 attributes of the dataset. N is the number of rows.

The ID ness is measuring the degree to which this attribute resembles an ID. The number of different values for the attribute (column) indicates a distinct value divided by the number of data rows (indexing) in **Table 4** to classify the attributes positive and negative effects and determine the target feature as the first action, this study chose the status attribute as the target for the entire dataset because it has the highest degree of stability (83.23 percent). Meanwhile, the status code produces similar results as the status code attribute, but it is avoided because the value of ID-ness is less than the 10% the target we chose. Furthermore, correlation is a statistical technique for defining the weight by value of the status attribute. Compared to the set attributes, the status attribute has a high positive significance and includes seven distinct indexes, which are used to fetch the dataset Ok, domain name system (DNS) lookup failed, Blocked by robots.txt, Connection Refused, Found, Found: Moved Temporarily, and Blocked by roboots.txt connection Refused DNS lookup failed found, by observation the from the status attribute consequence to distinct indexing ok (2446 repetitions where is 83.25%), DNS lookup failed (376 repetitions where is 12.80%), Blocked by robots.txt

(96 repetitions where is 3.27%), connection refused found (12 repetitions where is 0.41%), found (4 repetitions where is 0.14%), found: moved temporarily (3 repetitions where is 0.10%), and blocked by roboots.txt connection Refused DNS lookup failed found (1 repetitions where is 0.03%).

Table 4. Stability for selected attributes.

Attributes	Distinct value	ID ness	Stability
Size Bytes	45	1.53%	53.15%
Status	8	0.27%	83.23%
Status code	5	0.17%	83.23%
Type	7	0.24%	71.96%
Source	796	0.27.8%	0.48%
Destination	106	3.61%	44.4%

3.2. Perform data

The initial operation determines the weights for showing the influence for generating model to adjust the dataset through algorithms to achieve the prediction, the **Figure 3** below illustrates how define the pattern by independent and dependents variables.

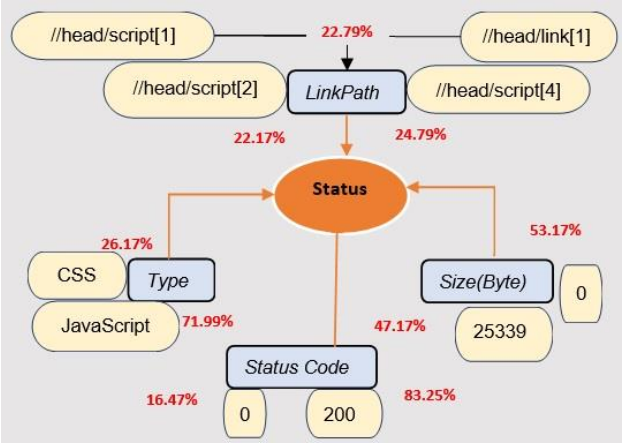


Figure 3. Weights from dataset define by target.

The weight of a dataset determined by correlation is classed as low dimension data (2D), with different weight outcomes for a single attribute. The relevant attribute for the target performs dataset is status, as shown in **Figure 3**, with a wide gap value in the correlation indexing. To address this issue, it is suggested that high dimensional (3D) data be used to cover all possible prediction data and to improve the accuracy scale for determine gap values of the indexing status attribute.

Based on the status attribute, the indexing rank categorizes four categories: status moved momentarily and found is between 10% and 14%, DNS lookup failed and blocked by roboots.txt is between 3.27 percent and 12.80 percent, connection Refused is 0.41 percent, and Ok is 83.25 percent. While the problem appeared to be about considering the target by using correlation weight, the justification to examine the different methods began by weight measure with validated the output, which led to

assigning a supervisor learning include linear classification, support vector machine, and dimensional classification.

3.3. Select target (generalized linear model)

The classification process is a pivotal step in this study, involving the careful selection of a model to effectively encompass and distinguish the entire dataset (Alrawashed et al., 2019; Alrawashed et al., 2021; Almarashde et al., 2014). This process comprises two key elements: first, building a model from statistical and label data, and second, validating the model's predictions on actual data. The fundamental objective of classification, to establish a specific label (predict) for a dataset, is crucial for obtaining reliable data based on recall and accuracy scales.

In our approach to classification, particular attention is directed towards the Generalized Linear Model (GLM). The GLM plays a central role in the classification process, working in tandem with the processes of recall and memory. It facilitates the detection of the intricate connection between response and prediction. To ensure the robustness of our model, a suitable training model is employed, incorporating cross-validation to enhance accuracy.

In our analytical framework, a traditional linear regression model is utilized to delve into data consistency. This model helps illustrate the variation in attributes through coefficient measures and standard deviation calculations, providing insights proportional to the mean (Zuniga et al., 2021).

Table 5. Coefficient for prediction model.

Attributes	**Status**			
	Index	**Blocked by robots.txt Connection Refused DNS lookup failed Found, Moved Temporarily Moved Permanently Moved Temporarily Not Found OK Object Moved Redirect See Other See other**	**Blocked by robots.txt**	**Ok**
Status Code	Continuous	0.843	0	−0.08079
	Missing	0	0	0
	302	0	0	−0.648
	200	0	−1.192	5.432
	0	0	1.422	−3.029
Link Path	String	0.843	0	−0.081
	Missing	0	−0.276	−0.615
	//head/script[5]	0	2.844	0
	//head/link[1]	0	−0.638	0
Type	CSS	0	−0.707	0.181
	JavaScript	0	1.908	0.268
	CSS,HTML Hreflang HTTP, Redirect Hyperlink Iframe Image JavaScript	0.843	0	−0.081

As depicted in **Table 5**, the concordance index assumes a crucial role in our correlation objectives. It serves to identify the distinction between observer and predictive expectations. One notable feature is its exclusive consideration of parameterization as a contradictory component. This consideration is essential to

obtain a simulation model that mitigates the risk of lacking discriminatory power (Alhroob et al., 2020).

It is imperative to note that the data for our study were sourced from the Jordanian e-government, as detailed in **Table 4**. The GLM is instrumental in this context for identifying risk factors, with a specific focus on accuracy and misclassification. This multifaceted approach ensures a nuanced understanding of the data and its implications for our overarching objectives.

The assessment of web page misclassification in our proposed model encompasses accuracy measures, incorporating detailed weights and coefficients as outlined in the table above, and adheres to Web Content Accessibility Guidelines (WCAG 2.1) standards. Our model validation involves experiments utilizing the WAVE Web Accessibility Evaluation tool and a similar tool named LERA, focusing on specific URLs:

1) First URL—Title: "تسجيل الوكالات التجارية" (Registration of Commercial Agencies)

2) Second URL—Title: "معاملات منح وتجديد الرخص العمومية المؤقتة والانشائية" (Granting and Renewing Temporary Public Drivers' Licenses and Construction Permits)

Our analysis employs four criteria to evaluate the critical impact on accessibility. The first criterion necessitates components to be enclosed by parent roles, possess correct attribute values, have readable text on buttons, meet WCAG 2 AA contrast ratio standards, and ensure links have recognizable text with unique id attribute values (Zuniga et al., 2021). Notably, the moderate criterion is absent.

The severity distribution for the first link path is categorized as follows:

- Critical criteria: 64%, detecting 16 issues, including information and relationship (tablist), name, value, and role (tab-*: tab-1, tab-2, etc.), and section 508 (#search_submit, #weatherIcon.newhokamtilogo > img, and .mart10).
- Serious criteria: 28%, detecting 7 issues related to control (keyboard), link purpose, and contrast.
- Minor criteria: 8%, detecting 2 issues parsing by id values of attributes.

(**Table 3** provides a summary of the experiments conducted to test the proposed model, minimizing data requirements for Jordanian e-government.)

In conducting this analysis, the Generalized Linear Model (GLM) plays a crucial role, enabling the construction of responses and predictions with accuracy, particularly in predicting the effect on content optimization for better design and accessibility. Amid the COVID-19 pandemic, the demand for reliable, practical, and current information from governments has surged. The Jordanian e-government vision of expanding digital government aims to provide residents in rural or underdeveloped regions with access to digital services, fostering digital inclusion and identification. The simplicity and straightforwardness of systems are crucial, especially for citizen activities such as searching, as indicated by the path column in **Table 5** (#search_submit).

When considering data attributes related to HTTP status codes, link paths, and types, the GLM can be leveraged to create a predictive model for inferring or forecasting the behavior of a web crawler or website interactions based on these

attributes.

For instance, by using HTTP status codes, link paths, and types as independent variables and 'Status' as the dependent variable, a GLM could predict the likelihood of different status outcomes based on these attributes. Each attribute's weight or coefficient in the model indicates its impact on predicting the 'Status' outcome, providing valuable insights into the significance and influence of each attribute in determining the 'Status' outcome.

This analytical approach equips researchers and practitioners with a quantitative understanding of the nuanced relationships between different attributes and their implications for predicting status outcomes. The interpretability of the assigned weights or coefficients enhances the model's transparency, facilitating a more comprehensive comprehension of the factors contributing to the predicted outcomes (see **Figure 4**).

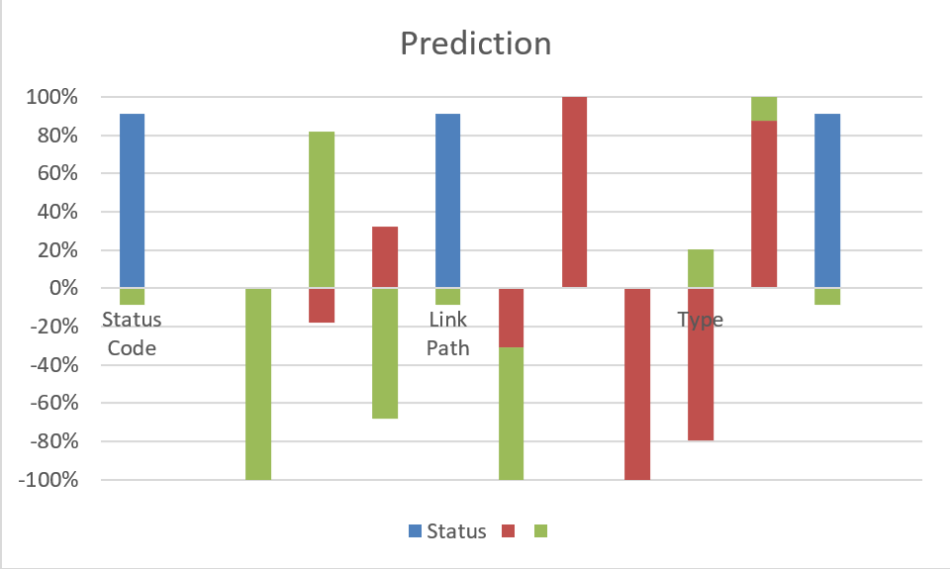


Figure 4. Prediction outcome.

4. Discussion

The study delves deeply into the intricate dynamics of missing data and its profound impact on the accuracy of the model. By meticulously examining how attributes and their interplay influence the development of a model pattern, we aim to provide a nuanced understanding of the challenges posed by missing data. Our findings notably highlight that the target attribute experiences a discernible influence, with an impact ranging from 3% to 12% due to missing data. This underscores the pivotal role of comprehensive and complete data in crafting a model that faithfully mirrors the intricacies inherent in the dataset.

In parallel, our investigation extends to a thorough examination of the Generalized Linear Model (GLM) within the e-government framework. Emphasizing its efficacy, we shed light on how the GLM adeptly recognizes risk indicators and categorizes data based on accuracy and recall. The study underscores the well-established utility of the GLM, particularly in the realm of black box testing, elucidating its crucial role in ensuring the seamless operation of e-government

processes. This underscores the robust and versatile capabilities of the GLM in the specific context of e-government applications.

Central to our study is the crucial role of accessibility requirements in evaluating e-government services. Here, the GLM emerges as a cornerstone tool, with a specific focus on its role in maintaining prediction accuracy in alignment with the stringent standards of the Web Content Accessibility Guidelines (WCAG) 2 AA criterion. This sheds light on the intricate interdependence between predictive precision and the rigorous accessibility standards that govern the landscape of e-government services.

As a final note, our study recognizes and underscores the vast potential inherent in big data management. By addressing various content types, from unstructured to semi-structured and structured data, big data emerges as a transformative force. This recognition not only contributes to the existing body of knowledge but also paves the way for additional studies and innovative e-government applications. In essence, our study provides a thorough and detailed exploration of missing data implications, GLM efficacy, accessibility considerations, and the expansive possibilities presented by big data in the realm of e-government.

5. Conclusions

This study underscores the multifaceted impact of missing data on both accuracy and the intricate relationships between attributes in generating a model pattern for the dataset. The observed missing data, constituting a noteworthy percentage range of (3% to 12%) of the total target attribute, highlights the significance of addressing data imbalances. In this context, the interrogator's role, particularly in the context of weight measurement, emerges as a pivotal factor in delineating attribute behaviors. This nuanced understanding serves as the backbone for developing a well-founded ranking model, crucial for determining accuracy scores.

To address these complexities, our approach involves leveraging a weighted dataset and conducting horizontal and vertical analyses. This strategy aims to recognize different types and attributes of relevance by calculating distance scores. The experimental results reveal a substantial relationship that becomes evident through this meticulous analysis. Additionally, there is a discernible potential to apply these methodologies to handle big data, particularly concerning specific characteristics, as indicated in previous studies (Alhroob et al., 2020; Althunibat et al., 2021).

The evaluation of the generalized linear model applied to e-government URLs, incorporating weights and coefficients to determine accessibility criteria, yields promising outcomes. Testing according to the WCAG 2 AA accessibility standard demonstrates that the generalized linear model precisely aligns with the expected accuracy results. This underscores the model's pivotal role in black box testing, ensuring the seamless execution of e-government processes while maintaining accuracy.

Looking ahead, the focus of our upcoming work will pivot towards digital transformation, leveraging cutting-edge technologies already integrated into the digital economy and smart cities. In the realm of Artificial Intelligence (AI), we anticipate employing classification techniques and natural language processing. These techniques will play a crucial role in confirming the accuracy of information,

especially considering the bilingual nature (Arabic and English) of the Jordanian e-government. However, the vast landscape of the big data paradigm presents an extensive opportunity for managing diverse content types, including structured, semi-structured, and unstructured content. This expansive potential opens avenues for future research and application development in the e-government domain.

Author contributions: Conceptualization, WA and AA; methodology, WA; software, AA; validation, AA, WA, and SSM; formal analysis, WA; investigation, AA and BA; resources, AA; data curation, AA; writing—original draft preparation, WA; writing—review and editing, AA and AH; visualization, AA; supervision, AA and AH; project administration, AA; funding acquisition, WA. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: The authors would like to thank Al-Zaytoonah University of Jordan for supporting this research.

Conflict of interest: The authors declare no conflict of interest.

References

Abdallah, M., Hammad, A., & AlZyadat, W. (2022). Towards a Data Collection Quality Model for Big Data Applications. Lecture Notes in Business Information Processing, 103–108. https://doi.org/10.1007/978-3-031-04216-4_11

Angelo, L., Nicola, M., Gabriele, C., & Alessandro, M. (2022). e-government in Europe. A Machine Learning Approach. Available online: https://mpra.ub.uni-muenchen.de/112242/ (accessed on 17 July 2022).

Abuaddous, H. Y. (2019). The Accessibility Diagnosis on Jordan E-Government Website. International Journal of Science and Applied Information Technology, 8(6), 45–48. https://doi.org/10.30534/ijsait/2019/05862019

Alexopoulos, C., Lachana, Z., Androutsopoulou, A., et al. (2019). How Machine Learning is Changing e-government. In: Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance. https://doi.org/10.1145/3326365.3326412

Alhroob, A., Alzyadat, W., Tareq Imam, A., & M. Jaradat, G. (2020). The Genetic Algorithm and Binary Search Technique in the Program Path Coverage for Improving Software Testing Using Big Data. Intelligent Automation & Soft Computing, 26(4), 725–733. https://doi.org/10.32604/iasc.2020.010106

Alhroob, A., Alzyadat, W., Almukahel, I., & Altarawneh, H. (2020). Missing Data Prediction using Correlation Genetic Algorithm and SVM Approach. International Journal of Advanced Computer Science and Applications, 11(2). https://doi.org/10.14569/ijacsa.2020.0110288

Alhroob, A. M., Alzyadat, W. J., Almukahel, I. H., & Jaradat, G. M. (2020). Adaptive Fuzzy Map Approach for Accruing Velocity of Big Data Relies on Fireflies Algorithm for Decentralized Decision Making. IEEE Access, 8, 21401–21410. https://doi.org/10.1109/access.2020.2969204

Al-Mushayt, O. S. (2019). Automating E-Government Services With Artificial Intelligence. IEEE Access, 7, 146821–146829. https://doi.org/10.1109/access.2019.2946204

AL-Naimat, A. M. (2013). The Critical Success Factors for E-Government Implementation in Jordan. In: Proceedings of the 4th International Conference on Computing and Informatics, ICOCI 2013; Sarawak, Malaysia. pp. 391–398.

Al Olama, O. S. (2018). UAE National Strategy for Artificial Intelligence 2031. Available online: https://ai.gov.ae/wp-content/uploads/2021/07/UAE-National-Strategy-for-Artificial-Intelligence-2031.pdf (accessed on 12 December 2023).

Althunibat, A., Abdallah, M., Amin Almaiah, M., et al. (2022). An Acceptance Model of Using Mobile-government Services (AMGS). Computer Modeling in Engineering & Sciences, 131(2), 865–880. https://doi.org/10.32604/cmes.2022.019075

Althunibat, A., Alrawashdeh, T. A., & Muhairat, M. (2014). The Acceptance of Using M-government Services in Jordan. In: Proceedings of the 2014 11th International Conference on Information Technology: New Generations. https://doi.org/10.1109/itng.2014.65

Althunibat, A., Binsawad, M., Almaiah, M. A., et al. (2021). Sustainable Applications of Smart-Government Services: A Model to Understand Smart-Government Adoption. Sustainability, 13(6), 3028. https://doi.org/10.3390/su13063028

Althunibat, A., AlNuhait, H., Almanasra, S., et al. (2024). Culture and law enforcement influence on m-government adoption: An exploratory study. Journal of Infrastructure, Policy and Development, 8(5), 3353.

Althunibat, A., Alokush, B., Tarabieh, S. M., & Dawood, R. (2021). Mobile Government and Digital Economy Relationship and Challenges. International Journal of Advances in Soft Computing & Its Applications, 13(1).

Al Thunibat, A., Zin, N. A. M., & Sahari, N. (2011). Mobile government user requirements model. Journal of E-governance, 34(2), 104–111.

Almarashde, I., Althunibat, A., & Fazidah El, N. (2014). Developing a Mobile Portal Prototype for E-government Services. Journal of Applied Sciences, 14(8), 791–797.

Alrawashdeh, T., ElQirem, F., Althunibat, A., & Alsoub, R. (2021). A prioritization approach for regression test cases based on a revised genetic algorithm. Information Technology and Control, 50(3), 443–457.

Alrawashed, T. A., Almomani, A., Althunibat, A., & Tamimi, A. (2019). An automated approach to generate test cases from use case description model. Computer Modeling in Engineering & Sciences, 119(3), 409–425.

Almaiah, M., Al-Khasawneh, A., Althunibat, A., & Khawatreh, S. (2020). Mobile government adoption model based on combining GAM and UTAUT to explain factors according to adoption of mobile government services. International Journal of Interactive Mobile Technologies (iJIM), 14(03), 199–225. https://doi.org/10.3991/ijim.v14i03.11264

Alzyadat, W., Muhairat, M., Alhroob, A., & Rawashdeh, T. (2022). A Recruitment Big Data Approach to interplay of the Target Drugs. International Journal of Advances in Soft Computing and Its Applications, 14(1), 02–13. https://doi.org/10.15849/zujijasaca.220328.01

Assiri, H., Nanda, P., & Mohanty, M. (2021). A novel e-government Framework using Blockchain. Journal of Information Assurance & Cybersecurity, 1–14. https://doi.org/10.5171/2021.164568

Awwad, M. S., & Omoush, K. S. A. (2012). Governance of information technology-business relationship quality and performance outcomes. Electronic Government, an International Journal, 9(4), 350. https://doi.org/10.1504/eg.2012.049724

Bigdeli, A. Z., & de Cesare, S. (2011). Barriers to e-government Service Delivery in Developing Countries: The Case of Iran. Emerging Themes in Information Systems and Organization Studies, 307–320. https://doi.org/10.1007/978-3-7908-2739-2_24

Cabrera-Moya, D. R. R., Vasudevan, H., & Prieto-Rodriguez, G. (2023). Kash training models: increasing levels of commitment and organizational effectiveness. Business: Theory and Practice, 24(1), 239–249. https://doi.org/10.3846/btp.2023.17480

Charalabidis, Y., Lampathaki, F., Sarantis, D., et al. (2008). The Greek Electronic Government Interoperability Framework: Standards and Infrastructures for One-Stop Service Provision. In: Proceedings of the 2008 Panhellenic Conference on Informatics. https://doi.org/10.1109/pci.2008.37

Falco, E., Kleinhans, R. (2018). Beyond Information Sharing. A Typology of Government Challenges and Requirements for Two-Way Social Media Communication with Citizen. The Electronic Journal of e-government, 19(1), 32–45.

Forum, J. S. (2019). E-Government in Jordan a Guide for Policy Maker, 2019. Available online: https://jsf.org/sites/default/files/EN%20E-Government%20Report%20.pdf (accessed on 7 October 2022).

Fragkou, P., Galiotou, E., & Matsakas, M. (2014). Enriching the e-GIF Ontology for an Improved Application of Linking Data Technologies to Greek Open Government Data. Procedia—Social and Behavioral Sciences, 147, 167–174. https://doi.org/10.1016/j.sbspro.2014.07.141

Guberović, E., Alexopoulos, C., Bosnić, I., & Čavrak, I. (2022). Framework for Federated Learning Open Models in e-government Applications. Interdisciplinary Description of Complex Systems, 20(2), 162–178. https://doi.org/10.7906/indecs.20.2.8

Ismailova, R. (2015). Web site accessibility, usability and security: a survey of government web sites in Kyrgyz Republic. Universal Access in the Information Society, 16(1), 257–264. https://doi.org/10.1007/s10209-015-0446-8

Leogrande, A., Magaletti, N., Cosoli, G., & Massaro, A. (2022). E-Government in Europe. A Machine Learning Approach. Preprints, 2022, 2022020359. https://doi.org/10.20944/preprints202202.0359.v1

Linders, D., Liao, C. Z. P., & Wang, C. M. (2018). Proactive e-Governance: Flipping the service delivery model from pull to push in Taiwan. Government Information Quarterly, 35(4), S68–S76. https://doi.org/10.1016/j.giq.2015.08.004

Manoharan, A. P., & Ingrams, A. (2018). Conceptualizing E-Government from Local Government Perspectives. State and Local Government Review, 50(1), 56–66. https://doi.org/10.1177/0160323x18763964

Meiyanti, R., Utomo, B., Sensuse, D. I., & Wahyuni, R. (2018). e-government Challenges in Developing Countries: A Literature Review. In: Proceedings of the 2018 6th International Conference on Cyber and IT Service Management (CITSM). https://doi.org/10.1109/citsm.2018.8674245

Ministry of Digital Economy and Entrepreneurship, Amman (2019). Available online: https://www.modee.gov.jo/ebv4.0/root_storage/ar/eb_list_page/gws_2019-modee-v16_0-en.pdf (accessed on 25 May 2022).

Mujali Al-rawahna, A. S., Chen, S. C., & Hung, C. W. (2018). The Barriers of E-Government Success : An Empirical Study from Jordan. International Journal of Managing Public Sector Information and Communication Technologies, 9(2), 01–18. https://doi.org/10.5121/ijmpict.2018.9201

Norris, D. F., & Reddick, C. G. (2012). Local E-Government in the United States: Transformation or Incremental Change? Public Administration Review, 73(1), 165–175. https://doi.org/10.1111/j.1540-6210.2012.02647.x

Osman, I. H., Anouze, A. L., Irani, Z., et al. (2014). COBRA framework to evaluate e-government services: A citizen-centric perspective. Government Information Quarterly, 31(2), 243–256. https://doi.org/10.1016/j.giq.2013.10.009

Robles, G., Gamalielsson, J., & Lundell, B. (2019). Setting Up Government 3.0 Solutions Based on Open Source Software: The Case of X-Road. Electronic Government, 69–81. https://doi.org/10.1007/978-3-030-27325-5_6

Twizeyimana, J. D., & Andersson, A. (2019). The public value of E-Government—A literature review. Government Information Quarterly, 36(2), 167–178. https://doi.org/10.1016/j.giq.2019.01.001

Zhang, H., Xu, X., & Xiao, J. (2014). Diffusion of e-government: A literature review and directions for future directions. Government Information Quarterly, 31(4), 631–636. https://doi.org/10.1016/j.giq.2013.10.013

Zuniga, F., Kozubowski, T. J., & Panorska, A. K. (2021). A generalized linear model for multivariate events. Journal of Computational and Applied Mathematics, 398, 113655. https://doi.org/10.1016/j.cam.2021.11365