

# INTELLIT Version 2.0 – Advancements and Evaluation of an Enhanced Platform for the Modern Romanian Literature

**Alin-Gabriel Dita**

National University of Science  
and Technology

POLITEHNICA Bucharest

313 Splaiul Independentei,  
Bucharest, Romania  
alin\_gabriel.dita@stud.acs.u  
pb.ro

**Mihaita Stroe**

The “G. Călinescu” Institute of  
Literary History and Theory,

Romanian Academy

Calea 13 Septembrie, Bucharest,  
Romania st.mihaita@gmail.com

**Laurentiu-Marian Neagu**

National University of Science  
and Technology POLITEHNICA

Bucharest

313 Splaiul Independentei,  
Bucharest,  
Romanialaurentiu.neagu@upb.ro

**Mihai Dascalu**

National University of Science and Technology  
POLITEHNICA Bucharest

313 Splaiul Independentei, Bucharest, Romania  
mihai.dascalu@upb.ro

**Lucian Chisu**

The “G. Călinescu” Institute of Literary History and  
Theory, Romanian Academy

Calea 13 Septembrie, Bucharest, Romania  
institutulcalinescu@gmail.com

## ABSTRACT

Creating digital solutions for Romanian literature is essential to preserve Romanian cultural identity. The development of the INTELLIT platform (version 1.0) represented a significant first step in the digitalization process of Romanian literature while also providing a high-quality educational resource. Overall, the platform aims to present information in a clear and user-friendly manner for those interested in enriching their knowledge of Romanian literature. INTELLIT is based on two essential corpora published by the “G. Călinescu” Institute of Literary History of the Romanian Academy: the General Dictionary of Romanian Literature (DGLR) and the Chronology of Romanian Literary Life (CVLR). Additionally, the platform features work by 32 Romanian writers from the “Fundamental Works” Collection. This study introduces the updated version (2.0) of INTELLIT that considers an improved architecture with a rewritten backend that indexes information in an Elasticsearch server, enhanced workflows including semantic search based on sentence embeddings, and an optimized UI/UX in the form of a responsive web platform. A study on this version showed that users had a pleasant experience, selected the most attractive features of the INTELLIT platform, and provided suggestions for future developments.

### Author Keywords

Web application; User Experience; Romanian literature; INTELLIT; DGLR; CVLR, semantic search.

### ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

DOI: 10.37789/ icusi.2024.26

## INTRODUCTION

The INTELLIT platform aims to address the lack of valuable and accessible online resources to support the learning process and facilitate the systematic and efficient study of Romanian literature. As digitalization is one of the essential pillars of progress, contributing to the well-being of a nation and increasing global competitiveness among education and research systems, converting the General Dictionary of Romanian Literature (DGLR) and the Chronology of Romanian Literary Life (CVLR) volumes into a digital format was a necessity [1].

DGLR contains information about all writers and their works in the Romanian language, as well as concepts, movements, and literary institutions. There are no geographical limits to the publications' origins as long as they add value to Romanian literature. The dictionary was printed in two editions, the first comprising seven volumes released between 2004 and 2009. The second edition of DGLR was published in two stages: the first four volumes were released between 2016 and 2018, and the last four were issued in 2019. Through its eight volumes, the second edition of DGLR expanded the dictionary's information with a substantial number of writers and publications and corrected and updated the data from the first edition [2; 3].

CVLR addresses the issue of recording Romanian publications across time to highlight the importance of writers and observe the evolution of Romanian literature in terms of contemporary events and mentalities. CVLR considers different periods in Romanian history, starting from World War II in 1944, documenting writings from the communist era until 1989, and the post-communist period

<sup>1</sup> <https://intellit.ro/>

(1990-2000). A new CVLR project is underway for 2000-2010 [4].

Given the growing trend toward online educational resources derived from a large volume of information transposed from physical resources, data analysis, and categorization were important aspects in developing the updated architecture of the INTELLIT platform. With efficiency as a key factor, users needed to enjoy an optimal experience. Intrinsically, DGLR and CVLR have separate sections within the INTELLIT platform, but a deeper categorization of the data was required to increase usability. As such, CVLR data can be displayed according to the selected year and month. At the same time, information from DGLR can be arranged according to different parameters such as writer, publication, concept, or literary genre. To reduce search time, the INTELLIT interface includes search and data sorting functionalities based on the selected parameters.

Our main contributions are as follows:

- Introducing INTELLIT Version 2.0 with enhanced workflows and optimized UI/UX; The platform is freely available online, and all data is available for research purposes upon request;
- Performing user validations to assess user experience and the utility of this specialized platform for Romanian literature.

This paper is structured as follows. The second section introduces a state-of-the-art detailing the initial version of the INTELLIT platform (Version 1.0) and corresponding experiments. Next, it showcases the updated version, focusing on aspects such as ease of use, coherence of functionalities, organization, and presentation of information. Results from the performed user validations are presented afterward, as well as areas for improvement for future versions of the INTELLIT platform.

### STATE OF THE ART

The digitization of specialized literary resources has an increasing interest. However, no such platforms besides INTELLIT version 1.0 are available for Romanian literature, while there is only one fully free alternative<sup>2</sup> for English literature that includes a search engine, thematic channels, events, essays, and the complete texts of an encyclopedia [5], and Larousse<sup>3</sup> for the French literature.

Initially, as part of the INTELLIT project, researchers from the Romanian Academy provided a dataset from DGLR in Adobe InDesign format, along with a series of Microsoft Word documents containing details about the lives and works of canonical writers. These files were converted into HTML to enable the data to be processed, parsed, and stored. The selected method involved indexing the data on an Elasticsearch server optimized for information retrieval tasks. Consequently, meaningful information about writers, such as their names, birth years, places of birth, biographies, lists of works, and more, was indexed in Elasticsearch [2].

This open-source search engine was justified since Elasticsearch can function as a distributed search engine with real-time analysis capabilities and can be scaled to manage large volumes of structured and unstructured data [6].

As a result, experiments and analyses were conducted on demographic data such as the birthplaces of all writers, the publication periods of each canonical writer, the number of writings per writer per year, and the places where canonical writers traveled [2]. Thus, the utility of digitizing DGLR data was argued, and publishing it on the INTELLIT platform enabled interested individuals to access it for documentation or analysis.

Regarding CVLR, limited research relies on digitization to quickly capture subjects of interest in literature at any given time and subsequently suggest relevant periods for different searches. Similar to DGLR, CVLR was also stored in Elasticsearch. An experiment on CVLR by Neagu et al. [4] identified contemporary themes for each period. The intended outcome was for users to search for various events in Romanian literary history based on keywords. Additionally, the system could recommend specific periods during which writings of similar interest were published.

These experiments marked the initial steps in adapting Romanian literature to the digital era by introducing Natural Language Processing techniques and applying them within the literary domain. Despite the advancements made by the INTELLIT platform in digitalizing and categorizing Romanian literature, a critical gap remains in the availability and accessibility of reliable, high-quality online resources for Romanian literature. Existing platforms are often incomplete or poorly organized, lacking advanced search functionalities and comprehensive coverage of Romanian literary works. Furthermore, the integration of semantic search and user-friendly interfaces in the context of Romanian literature still needs to be explored. This paper aims to address these gaps by introducing INTELLIT Version 2.0, which enhances the platform's architecture with improved search capabilities and a more intuitive user interface, making Romanian literary resources more accessible and useful for educational and research purposes.

### INTELLIT VERSION 2.0

Given the purpose of creating an educational website with material as structured as possible for an intuitive user experience, the second version of the INTELLIT platform was divided into five major sections, which are then broken into individual pages, as required. Thus, the main pages are: "HOME" - the landing page; "DGLR" - the page dedicated to the content of the General Dictionary of Romanian Literature; "CVLR" - the page dedicated to the content of the volumes in the Chronology of Romanian Literary Life; "WORKS" - containing volumes free of copyright restrictions written by the most important Romanian classic writers; and "ABOUT US" - providing information on the institutions that collaborated in the INTELLIT project.

### Architecture

The architecture of the updated platform is available in

<sup>2</sup> <https://www.britannica.com/art/English-literature>

<sup>3</sup> <https://www.larousse.fr/>

Figure 1. INTELLIT version 2.0 includes a web front-end platform developed in React.js version 184 and NextJS version 13.5.65 to display the processed information and facilitate user interaction. The backend was implemented using the ExpressJS framework<sup>6</sup>, which communicates with the Elasticsearch server where the data is indexed, and with a FastAPI server<sup>7</sup> written in Python, which enables semantic search through the descriptions of publications from DGLR. Both the web client and web server serve static and dynamic content. Static content refers to content that is stored directly on the client or on the server, such as images, or documents. Dynamic content refers to content which is retrieved dynamically from a data source, such as the Elasticsearch database.

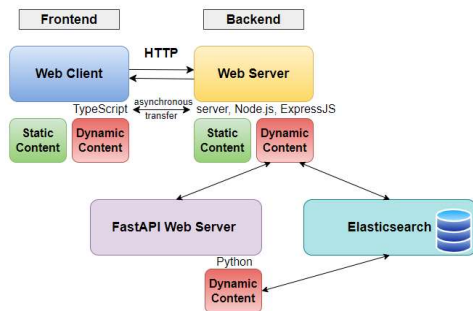


Figure 1. INTELLIT Version 2.0 Architecture.

In addition to the backend written in ExpressJS framework, the implementation requires the Cross-Origin Resource Sharing (CORS) [7] middleware to manage access and resource-sharing properties and the "path" module to handle file paths, given the numerous images and PDFs stored directly on the backend. Additionally, a controller was implemented to manage operations related to the Elasticsearch database – reading, writing, and updating data. This middleware controller adds a layer of security to the operations and manages data encapsulation and errors.

**UI Graphical Design**

The homepage has a straightforward, minimalist style. This page supports users in exploring the platform's other sections. The navigation bar is located at the top of the page. The page's content is organized into three sections, each displaying the title of the digital work when the user hovers over each representative image. Subsequently, a consistent layout is applied throughout the entire platform.

Since the General Dictionary of Romanian Literature (DGLR) is quite complex with 9,572 individual entries, transposing it into an online platform was challenging. First, the functionalities were divided into the basic elements of Romanian literature that DGLR addressed and, simultaneously, into the data categories indexed in

Elasticsearch. Thus, on the main page dedicated to DGLR, 9 categories were introduced, namely: a) Writers; b) Publications; c) Concepts and Currents; d) Literary and Folkloric Genres; e) Religious Texts; f) Writings with Unclear Authorship; g) Institutions, Societies, and Associations; h) Literary and Folkloric Genres; i) Other Literature in the Romanian Space. Out of the total 9,572 literary entries, the most represented was the literary publications (1,801 entries), followed by poets (1,575 entries), and prose writers (1,523 entries). On the lower end, the least represented subcategory was the religious texts (4 entries), followed by other works of literature in the Romanian space (5 entries), and versifiers (7 entries).

The most elaborated page is the one devoted to writers. To ensure rapid searches, the writers are categorized into 15 groups. Once a category is selected, information on all writers is pulled from Elasticsearch. Moreover, users can select all writers by the first initial of their last name or search for a writer by name on this page. Once the selection is made by an initial, an overlay is rendered as a blue circle to indicate the number of matching results (i.e., the number of writers in that category whose names start with the selected letter) – see Figure 2.

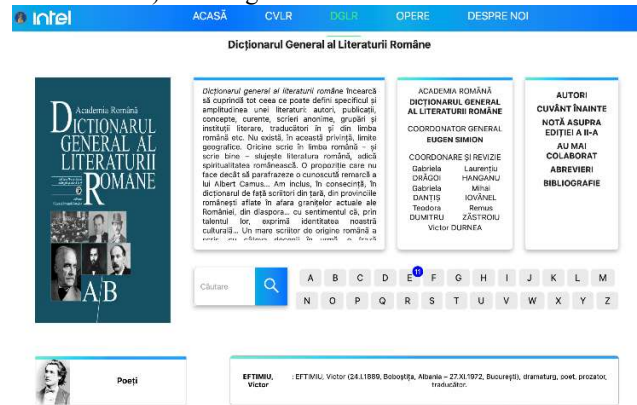


Figure 2. DGLR – Selection of all authors starting with “E”.

A separate page is displayed upon choosing a writer. This page presents the writer's life, works, and important references. Furthermore, there is a group of canonical writers for whom more in-depth information is accessible. The canonical writer's page comprises two parts, as can be seen in Figure 3. The first section (left and central part of the screen) includes their biography, works, quotes, critical reviews, timeline, writings, bibliography, and images capturing important moments in their lives. The second part is exclusive to canonical writers and consists of a card with a portrait image of the writer, their birth and death details, and a list of notable works.

<sup>4</sup> <https://react.dev>  
<sup>5</sup> <https://nextjs.org/>  
<sup>6</sup> <https://expressjs.com/>  
<sup>7</sup> <https://fastapi.tiangolo.com/>



Figure 3. DGLR – Description of Mihai Eminescu.

Except for the writers’ page, the other pages within DGLR have data arranged in the form of cards containing the titles of the respective entities. These pages use a dynamic parameter from the route to generate content (i.e., to make requests to Elasticsearch for data). This implementation encourages flexibility and code reuse. On these pages, the user can apply a filter by the initial letter of the entity's name and perform a semantic search, which is presented in a subsequent section. When an entity is selected, a specific page opens, displaying the text representing the description or the corresponding content from DGLR.



Figure 4. DGLR – Description of the modernism literary trend.

The layout chosen for the Chronology of Romanian Literary Life page follows a simplified navigation (see Figure 5). Underneath the heading, three separate sections indicate the years during which the volumes of CVLR were released: 1944 - 1964, 1965 - 1979, and 1990 - 2000. Every part contains links to corresponding PDF files like "Foreword" and "Edition Note," enabling users to access in-depth information on those time frames. Users can select the desired year between 1944 and 2000 below the cover of the corresponding CVLR year in the sidebar. The highlighted year helps quickly identify the year of the event within the navigation context. The initial selection is made for 1944 and the first month with available data (i.e., August). Selecting a year alters the literary events, the CVLR cover image, the listed year, and the link to the corresponding "Note on the Year" PDF document. Beneath the search bar, the page shows the months of each selected year as tabs. Users can move between these tabs to see literary events happening each month, offering seamless and quick navigation.

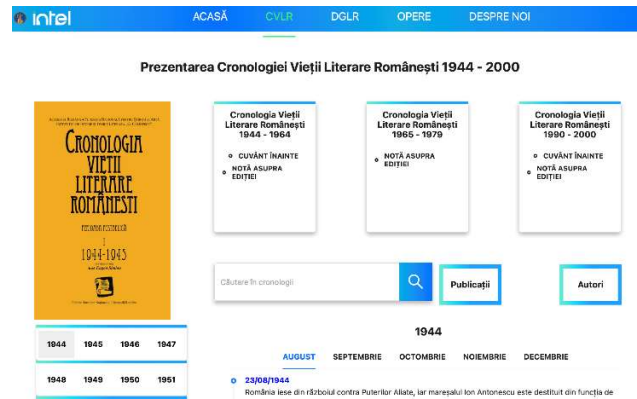


Figure 5. CVLR – Landing page.

Additionally, there are the "Publications" and "Writers" filters included, that open related models displaying the appropriate publications and writers. The information in these pop-up windows is specialized and changes based on the chosen year (see Figure 6).



Figure 6. CVLR – Selection of authors for a given year.

When choosing a publication or writer from the designated year, their name is entered into the search bar to enable sorting by the selected criteria. This feature helps users easily view all publications and writers mentioned in the chosen year and swiftly access information from CVLR in a highly efficient and straightforward manner.

A “WORKS” section is included on the main page and the sections for DGLR and CVLR, highlighting canonical writers and their key works free of copyright restrictions from classical Romanian literature, as in can be seen in Figure 7. The 32 writer cards are organized in a grid layout, making navigating and visually recognizing them easier. Access is provided to all available works in PDF format.



Figure 7. Fundamental works – Main interface.

Semantic Search

Semantic search has been implemented across all data categories in DGLR, except writers, to boost search accuracy

and enhance user experience. This functionality enables users to find the desired information, even if they make small typing mistakes or are unsure of the precise title or name of the literary work they seek.

Semantic recommendations are performed based on cosine similarity between sentence embeddings computed on the user's search text and precomputed on all text descriptions in Elasticsearch. The pre-trained all-MiniLM-L6-v2 model<sup>8</sup> was selected due to its balanced performance between speed and search quality. This model generates embedding vectors with 384 dimensions. Furthermore, it has been pre-trained in over 50 languages spoken globally, including Romanian, enabling its successful integration within the INTELLIT platform.

The initial stage required a substantial decrease in the data processed for the search to be completed. Hence, the data was filtered in Elasticsearch based on its previously established categories. After this initial filtering process, a vector containing the names and descriptions of literary elements from Elasticsearch was created. Semantic search relies on full-text descriptions; nevertheless, keeping the publication names and links to their page for later display on the front end is also necessary.

Besides considering the initial embedder, which provides similarity scores between the user's search text and the text descriptions, a cross-encoder evaluates the scores, leading to a ranked list of similarities, greater than a pre-defined threshold value. Tensor vectors were pre-computed and stored on the web server for all the 1,801 publications.

The semantic search functionality leverages the all-MiniLM-L6-v2 model, a pre-trained model known for balancing speed and accuracy. This model was selected because it generates 384-dimensional embeddings suitable for multilingual contexts, including Romanian. When a user inputs a search query, sentence embeddings are computed for the query text and compared to pre-computed embeddings of all indexed descriptions using cosine similarity. This approach enables a more nuanced search, accommodating slight variations in user input and providing more relevant results. The cross-encoder then ranks the similarities, ensuring the most pertinent results are presented to the user.

## Results

A study involving 18 respondents was conducted in July 2024 to measure the utility and user experience of this second version of the specialized platform for Romanian literature. The respondents were from Bucharest and Craiova, mainly software engineers, high school teachers, and students. Most of them were already familiar with using online resources to learn or prepare materials. The participants used the platform and explored its functionalities. To ensure that users navigated through all sections of the platform before completing the questionnaire, they were provided with a list of tasks to accomplish on the

platform, such as in the “DGLR” section – to search for Nichita Stănescu's biography; in the “CVLR” section – to see which literary events took place on, for example, 7th September 1972; or in the “WORKS” section – to read a poem from Ion Barbu's volume "Joc secund".

Based on the study outcomes, INTELLIT version 2.0 is user-friendly, providing a pleasant experience for its users. The System Usability Scale (SUS) [8] was used to measure how effectively users can utilize the platform to meet their needs and objectives (see Table 1). This SUS score was derived from the calculation of UMUX-LITE (Usability Metric for User Experience LITE) scores based on the first two questions of the questionnaire. The resulting SUS score for the INTELLIT platform was 74.77%, which is considered good and above the standard average - 68%.

Previous user experience with online resources in Romanian literature indicated that these resources often require more complete and reliable information and could be better organized. Despite this, the perception of the INTELLIT version 2.0 was notably more favorable. Furthermore, the score was 4.72 on a 5-point Likert scale [9] when users were asked whether the INTELLIT platform could serve as an educational tool in Romanian language and literature courses in schools.

The platform's intuitive design was also reflected in the high score of 4.61 out of 5 when considering the ease of learning how to use it. This suggests that most users can navigate the platform without spending excessive time understanding its functionality. The platform performed well in terms of navigation and response times, with users rating the speed of navigation on INTELLIT at 4.72 out of 5. Additionally, the main functionalities enabling users to quickly obtain the desired information were rated favorably.

**Table 1.** INTELLIT User Experience Feedback  
(Mean and Standard Deviation values corresponding to qualities scored on a 5-point Likert scale).

UX Quality	M (SD)	UX Quality	M (SD)
Willingness to use	4.33 (0.66)	Fast navigation	<b>4.72 (0.55)</b>
Ease of use	4.38 (0.89)	Good mix of features	<b>4.50 (0.76)</b>
Ease of finding information	4.38 (0.67)	Clear and organized visual display	4.44 (0.89)

When developing the updated version, significant attention was given to the visual presentation of data, ensuring a user-friendly and aesthetically pleasing interface. This effort was validated by a score of 4.44 out of 5 from the study participants, indicating that data is displayed in a clear, readable, and well-organized manner. Users were particularly interested in the section dedicated to CVLR, appreciating the convenience of reducing searches through physical volumes to a few classifications by the year and month of literary events.

Regarding user-suggested improvements for the INTELLIT platform, the primary focus was on enhancing the search functionalities. Further improvements should focus on expanding the semantic search model across all data categories and platforms' pages. Additionally, exploring and

<sup>8</sup> <https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2>

implementing a higher-quality artificial intelligence model for improved results is being considered. An interesting feature to consider is also applying a similarity model on writer biographies, generating a top-five list of writers with the most similar biographies based on a similarity score.

Improvement suggestions from the survey include uploading audio and video recordings of Romanian writers, where available (such as poetry readings, interviews, and speeches). Furthermore, implementing role-based authentication is being considered to enable administrator users to add or edit content on the platform. This functionality will be accompanied by a user authentication flow and a user management dashboard moderated by the Romanian Academy personnel.

Despite the current advancements in INTELLIT Version 2.0, several limitations remain. The current semantic search model could benefit from expansion to cover all categories of data and improve accuracy. Additionally, incorporating multimedia content could further enrich the user experience. Addressing these limitations can enhance the platform's functionality and user engagement.

#### Conclusions

The INTELLIT concept stemmed from the desire to create an easy-to-use and maintained platform that not only displays information from targeted literary volumes but also serves as an informative source and educational resource for students studying Romanian literature. Given the scarcity of specialized platforms in this field, with most existing websites offering low-quality essays containing incorrect and incomplete information, the INTELLIT platform, with its comprehensive updates – to Version 2.0, adds value by providing verified and curated information, good separation of the modules, filtering and semantic search based on sentence embeddings. INTELLIT Version 2.0 was also evaluated by a group of both students and teachers, and the perception was positive in terms of speed and interaction, underlying fast navigation, a good mix of features, and a clear and organized visual display. Future work may be conducted to perform role-based authentication for managing the platform content or the ability to upload audio and video recordings for writers or publications.

#### ACKNOWLEDGMENT

This work was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS – UEFISCDI, project number PN-III-P4-PCE-2021-0300/ PCE 127, INTELLIV– “Soluții digitale inteligente pentru cercetarea și diseminarea literaturii române vechi și premoderne”.

#### REFERENCES

1. Rahmatullah, A. S., Mulyasa, E., Syahrani, S., Pongpalilu, F., & Putri, R. E. (2022). “Digital Era 4.0: The Contribution to Education and Student Psychology”. *Linguistics and Culture Review*, 6, 89-107.
2. Neagu, L.M., Toma, I., Dascălu, M., Trăușan-Matu, Ș., Hanganu, L., & Simion, E. (2020). “A Quantitative Analysis of Romanian Writers’ Demography Based on the General Dictionary of Romanian Literature”. In *Ludic, Codesign, and Tools Supporting Smart Learning Ecosystem and Technologies* 197, 253-261.
3. "Plan de Activitate 2021". Academia Română. Institutul de Istorie și Teorie Literară "G. Călinescu". [Online]: [https://acad.ro/bdar/rapInt2021/01lit/plan/1Inst\\_IstTeorieLiteraraGCalinescu\\_Plan.pdf](https://acad.ro/bdar/rapInt2021/01lit/plan/1Inst_IstTeorieLiteraraGCalinescu_Plan.pdf).
4. Neagu, L.M., Cotet, T.M., Dascălu, M., Trăușan-Matu, Ș., Chisu, L., Simion, E. (2020). "Semantic Recommendations and Topic Modeling Based on the Chronology of Romanian Literary Life".
5. "Britannica in the digital era," britannica.com, 2023. in "Encyclopædia Britannica". [Online]: <https://www.britannica.com/topic/Encyclopaedia-Britannica-English-language-reference-work/Britannica-in-the-digital-era>
6. Gormley, C., Tong, Z. (2015) "Elasticsearch: The Definitive Guide: A Distributed Real-Time Search and Analytics Engine". O'Reilly Media, Inc.
7. Chen, J., Jiang, J., Duan, H., Wan, T. Chen, S., Paxson, V. & Yang, M. (2018) “We still don’t have secure cross-domain requests: an empirical study of CORS”. In 27th USENIX Security Symposium, pp. 1079–1093.
8. Kaya, A., Ozturk, R., & Gumussoy, C. A. (2019) “Usability Measurement of Mobile Applications with System Usability Scale ( SUS)”. *Industrial engineering in the big data era. Lecture notes in management and industrial engineering*, Springer, Cham, pp. 389-400.
9. Joshi, A., Kale, S., Chandel, S. and Pal, D. (2015), “Likert scale: explored and explained”, *British Journal of Applied Science and Technology*, Science domain International, vol. 7, no. 4, pp. 396-40.