Explicit Knowledge Database Interface Model System Based on Natural Language Processing Techniques and Immersive Technologies

Luis Alfaro[1], Claudia Rivera[2], José Herrera[3], Antonio Arroyo[4], Lucy Delgado[5], Elisa Castañeda[6]
Universidad Nacional de San Agustín de Arequipa, Arequipa, Perú[1,2,4,5,6]
Universidad Nacional Mayor de San Marcos[3]

***Abstract*—This work is focused on the proposal and development of an interface system model, based on natural language processing, immersive technologies and natural user interfaces, for the interaction with Explicit Knowledge databases. Five phases were proposed: The user testing characterization, the establishment of the state of the art and the theoretical foundation, the design and development of software, the system implementation and the functional tests and evaluation of the usability of the interface model. In order to establish the user testing characterization and the corresponding theoretical framework, the expert guide on Knowledge Management and Virtual Reality was followed, based on the approach of Usability and Computer Ergonomics compatible with the ISO 9241 standard. The traditional interfaces and the proposal in this work were evaluated for each of the metrics defined by the ISO 9241 standard, considering the dimensions of effectiveness, efficiency and satisfaction. The statistical test of "T-Student" established that there is enough evidence to confirm the existence of the following significant differences: Effectiveness is lower using the proposed interface model; efficiency and satisfaction is higher using the proposed interface model. Based on the conducted tests, it can be established that the proposed interface model is superior to the traditional interface in terms of the "Efficiency and Satisfaction" dimensions and inferior in terms of "Effectiveness." Consequently, it can be concluded that the scientific article exploration model using VR and NLP is superior to the traditional model.**

Keywords*—Knowledge management; explicit knowledge databases; natural language processing; natural user interfaces; Immersive technologies

I. Introduction

Knowledge Management (KM) is one of the emerging academic disciplines which responds to the growing demands of the information society [1]. It is the result of the confluence or intersection of the areas of business administration, communication technologies and information systems, and is defined as the process of creation, storage and exchange of organizational knowledge. With the existence of a wide variety of modern organizations, objectives and cultures, KM becomes an extremely interdisciplinary area of operation which develops its own objectives and professional approach.

The knowledge-based vision asserts that knowledge management can facilitate innovative practices by transforming knowledge assets into new products and services through a series of management processes and activities. It can promote the exchange and distribution of knowledge necessary for innovation activities, thereby stimulating the generation of new ideas and ultimately enhancing innovation performance.

For [2], an important activity of the organizations is the integration of individual knowledge into collective knowledge. Today, there is a need to properly manage tacit and explicit business knowledge [3], which, in the latter case, may be document-based, having a direct relationship with the productivity of the workforce and the efficiency of the organization. In many organizations, much of the business documents are not structured and do not have an adequate management system. These documents take many forms, such as office documents, images, reports, e-mails, drawings [4].

The contemporary organizations have a large number of documents which can be located in various containers or repositories. These documents can be lost, misplaced or even stored by someone without the knowledge of other collaborators. Since the files may have been stored on several separate computers, locating and retrieving them can be a complex task. To this end, Web-based applications can be equipped with the ability to manage various types of documents, such a business, process and research documents, which can be explored with immersive technology resources [5], as well as other types of documents in public and private organizations. Likewise, web-based applications are much more accessible and share some common characteristics. They are distributed, which means that storage and access takes place in different physical locations. Similarly, for [6], Scientific data repositories have a key role in science.

Tasks of exploration, search, and information collection are carried out using computer systems with traditional interfaces, which may include forms, menus, windows, and hyper medial or hypertextual elements. Typically, these tasks involve using the author's name, keywords, publication year, and/or digital files, along with some additional search parameters. The results are presented in a text list format [7].

The information indexed in a repository can grow rapidly, and when using traditional interfaces, there can be difficulties in exploring and selecting different types of documents. This can impact the search processes conducted by collaborators, making them tedious and resulting in unexpected outcomes.

The originality and motivation of this research lie in an attempt to contribute to the design and exploration of tools for optimizing the review tasks of explicit documents within private, public, and research-based organizational settings. These tools aim to facilitate the work of executives, collaborators, and researchers through novel forms of representation and interaction with textual information. The approach involves utilizing efficient and agile system models for searching and selecting from the extensive variety and diversity of explicit documents stored in organizational repositories.

In this paper, Section II, describes methodology, Section III describes the theoretical framework; Section IV explores and analyses immersive technologies; Section V explores and analyzes the works related to the development of interaction systems with Textual Web-based repositories using Virtual Reality (VR) and NLP; Section VI describes the proposed model; Section VII evaluates the model in relation to a traditional one; Section VIII evaluates the results. In conclusion, the Section IX presents the conclusions and recommendations for future work.

II. Methodology

For the development of the theoretical framework, the required elements for a systematic literature review and meta-analysis (PRISMA) were also utilized. Consequently, a search string was defined concerning the topics of interest, identifying relevant keywords to construct the search string. This search string was applied to recognized databases such as IEEE Xplore, Science Direct, Web of Science, and Scopus, considering that the studies were articles from journals and conferences published in the last 20 years. Duplicate studies were removed, those that did not contribute to the study were excluded, and applying inclusion and exclusion criteria, a total of 43 studies were used in the research.

Five phases were proposed for the development of the model, as described in Fig. 1.

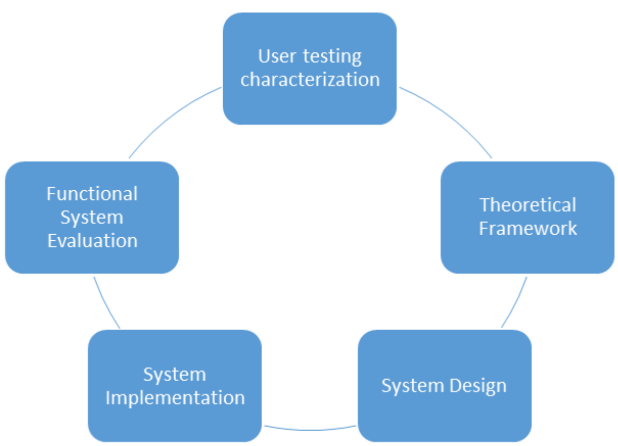


Fig. 1. Proposed methodology.

In the “Using Testing Characterization” phase, the theoretical framework was established, incorporating Knowledge Management, Web-Based Repositories, Natural Language Processing, and Immersive Reality along with their technological resources. This phase also involved the design and development of software, as well as the criteria required for evaluating the model, referred to as the usability interface.

In the “System Design” phase, the system was designed by establishing the background and reviewing related works. It was noted that there are various approaches to addressing and resolving the problem.

Moving on to the “System Implementation” phase, the different components of the system were implemented in the following steps: Physical environment setup for immersive experiences, Natural Interaction System Model, Implementation of Web Crawler for information retrieval from the model, Speech recognition and text-to-speech engine and Implementation of the syntactic analyser.

For the categorization of user tests, the expert guidelines in Knowledge Management (KM) and Virtual Reality (VR) were considered. These guidelines were based on the usability and ergonomics of computer science, aligning with the ISO 9241 standards. After configuring the interconnected devices and initializing the software modules, the evaluation of the model for exploring scientific articles using VR and NLP was conducted, alongside the traditional model of the scientific database ALICIA. Five users participated in the experiment, a number recommended by Nielsen [43] for evaluating interfaces, a criterion considered for obtaining the required dataset.

Participants were assigned the task of exploring scientific articles for a duration of 30 minutes using the interfaces of both models. During the experiment, users could request assistance from an expert who observed and documented the proceedings. Subsequently, the results were tabulated based on the metrics of ISO 9241, specifically efficiency, effectiveness, and satisfaction. These results are discussed in Section IX.

III. Theoretical Framework

In this section, a historical review and state-of-the-art analysis are conducted to explore the theoretical foundation required for the research.

A. Knowledge in the Organizations

The Knowledge Management System (KMS) is focused on information and enterprise social media, as well as on the connection and communication between employees. These two perspectives correspond to the categorizations of knowledge theory and KM. KM recognizes that virtually all jobs involve “knowledge work,” and employees are “knowledge workers” to some extent. Therefore, it places more emphasis on knowledge than on manual skills. Consequently, the creation, sharing, and use of knowledge are among the most important activities in nearly every organization.

Existing research provides the theoretical framework for carrying out these activities, and it is proposed that knowledge sharing can contribute to companies achieving innovation objectives, thereby improving organizational performance [8]. The positive impact of KM capabilities on organizational performance has also been studied [9], considering that KM contributes to a company’s innovation capacity [10].

Demuner-Flores and Nava-Rogel [11] emphasize a close relationship between knowledge and action, which also involves understanding and comprehension. The knowledge that an individual possesses is a product of their experience and incorporates the norms for evaluating novel contributions from their environment. For [12], knowledge involves framed experience, values, contextual information, and informed intuition.

It provides a framework for evaluating and incorporating new experiences and information. Knowledge originates and is applied in the minds of knowledge workers, but it can also be integrated into organizations through documents, repositories, practices, routines, and organizational standards.

KM is based on the concept of collecting, organizing, analyzing, and sharing knowledge in a way that employees can easily access and use it [13]. Therefore, a competitive organization relies, among other things, on the effectiveness of its workforce in creating new knowledge, sharing it within the organization, and using it to gain a competitive advantage.

Ultimately, KM aims for a positive impact on business value generated from knowledge. This is achieved through the creation of processes, culture, and technology in an environment where knowledge exchange is considered essential and occurs seamlessly. Thus, this research work can contribute to the exploration of textual documents in repositories of explicit web-based knowledge.

On the other hand, in the KMS, the following types of knowledge are defined:

1) Explicit Knowledge [14]: It classified as structured or unstructured, can be symbolically transferred, meaning it can be conveyed through linguistic or computational mediation, and can be documented and stored in databases. Some examples include written procedures, instructional manuals, lessons learned, best practices, as well as research results in the form of indexed scientific articles in repositories such as Web of Science, Scopus, Scielo, etc., or in academic data repositories of universities and other institutions, for the exploration, search, and collection of information [15]. Some databases, document types, and spreadsheets are examples of structured data and information, which are organized to allow for future retrieval. This work is focused on the exploration of emails, images, training courses, and audio and video selections, which are examples of unstructured knowledge because there is no clear way to reference them for retrieval.

2) Tacit Knowledge: It is knowledge hidden from the consciousness of the individual who possesses it [15]. It resides in the minds of human beings, and capturing or encoding it is not easy because it involves a complex process. Knowledge in its entirety resides in a tacit dimension. It is expressed through human actions, stories, as well as competencies, skills embedded in an individual's worldview, experiences, attitudes, evaluations, and perspectives that are often taken for granted. It can be observed through action and is less concrete compared to explicit knowledge. It can be described as a "tacit understanding" of something, which is difficult to document in writing or in databases. It is considered more valuable because it provides context for people, ideas, and experiences.

Managing both types of knowledge requires applying various methods and approaches, taking into consideration their unique characteristics. Some other types of knowledge in the context of corporate KM: (1) Individual knowledge, in which only one entity owns it; (2) Collective knowledge, whose transfer is focused on interaction and goal achievement.

Finally, the aim of KM's contribution to positive business value generated from knowledge and its management will be achieved by creating an environment that encompasses processes, culture, and technology for seamless knowledge sharing. Hence, the importance of this research work, which can contribute to the investigation and exploration of textual documents in web-based repositories of explicit knowledge.

B. Web Based Repositories

Information systems designed to preserve and organize commercial, scientific, and academic materials are referred to as repositories. They are used to support operations, research, and learning, while also ensuring access to information [16]. Institutional repositories also consist of interoperable web-based services, dedicated to disseminating the resources of the results and analysis of institutional, scientific, academic, or business operations of institutions, based on the listing of a specific set of data (metadata), so that these resources can be collected, catalogued, accessed, managed, disseminated and preserved. Respect for compatibility and interoperability standards makes it possible for the content of a repository to be more easily retrievable, not only for the institution itself, but also for the scientific community and society.

Scientific databases consist of components that include information related to library materials such as books and other types of documents, including journals and other scientific publications. These materials encompass scientific articles, research papers, conference proceedings, books, among others. Scientific databases are typically in electronic format and are accessible via the internet. They contain abstracts, bibliographic citations, references, and often the full text or links to the full text.

The government institution of the National System of Science and Technology and Technological Innovation of Perú (CONCYTEC), has subscriptions to various scientific databases, to which it provides free access and makes available to the scientific and academic communities and general population. The repository of scientific and technological information and innovation, called ALICIA, provides free access to intellectual production and research in science, technology and innovation, developed in the public sector and private institutions, with public funds.

The bibliographic review carried out in this part of the work, allows us to establish that the research in the design and development of friendly interfaces, from a unique and original perspective, resulting from the confluence of the use of methodologies and techniques, such as NLP, Natural Users Interfaces and VR, that facilitate the search and exploration in Explicit Knowledge database, can be very useful for the collaborators of the business organizations, researchers, academics, scientists and professionals in general, since it will not only facilitate the tasks that these carry out, but will also shorten the time spent, as well as diminish the costs, a fact that can undoubtedly contribute with the capacity of research.

C. Natural Language Processing (NLP)

NLP refers to the hardware and/or software elements of a computer system focused on analyzing or synthesizing spoken or written language [17]. NLP contributes to the retrieval of texts for user reading, which can vary in size from a paragraph to a book. Innovations in emerging technologies have made tasks such as storage, searches, and the retrieval of complete texts or parts of them from online documents relatively straightforward [18].

Many academic libraries have implemented various services to support users in their learning and research activities through chat. Chat serves as an important channel for accessing resources and services in university repositories. Repositories and libraries have accumulated a significant amount of data in the form of chat transcripts. Analyzing the content of these transcripts can help library officials understand user information needs, allocate library resources more efficiently, and improve the quality of chat reference services [3].

Conversational agents (chatbots) are programs that utilize natural language processing with a question-and-answer system, aiming to simulate intelligent dialogue with human interlocutors. This can be implemented through text messages or voice as a tool created to support the customer-company relationship, enabling virtual interaction through technology in the most human-like manner possible [19].

Recent applications of explicit database interfaces include two works by [20], who propose a Natural Language Query (NLQ) database query system (NADAQ) as an alternative solution. In their design, they introduce new models for translation, seamlessly merging deep learning with traditional techniques for database analysis and developing new techniques to allow the neural network to reject queries irrelevant to the content of the target database, and recommend that candidate queries be translated back into a tuned natural language. According to [21], the slow development of Natural Language Interfaces for Databases (NLIDB) stems from linguistic issues (such as language ambiguity) as well as domain portability. Brad et al. [21]assert that there is a demand for non-expert users to make queries to relational databases using their natural language instead of utilizing specified attribute values for the database domain. According to [22], the use of natural language instead of SQL-based queries enabled the development of NLIDB, a new type of processing that represents an approach to developing intelligent database systems aimed at improving the performance of flexible queries in databases.

Among the most important applications of NLP, we can mention human-machine interaction by voice, becoming one of the most successful applications with multiple uses, information retrieval and extraction, morphological, syntactic and semantic tagging, NLP for automatic responses to questions, text sentiment analysis, automatic text summarization, document classification, among others.

The review of the literature in NLP allows establishing the requirements of the users who demand the use of interfaces in their natural languages, making the research and development of research projects necessary, which contribute with the development of a model of interface systems for the interaction with Explicit Knowledge databases, which simplify and reduce the costs referred to the search and research tasks.

IV. Immersive Technologies

A. Virtual Reality (VR)

VR as defined by [23], is a system used to describe 3D environments generated by computers that can be explored and allow interaction with users. Users become part of a virtual world in which they are immersed and can interact with and manipulate objects, as well as perform various actions. VR is a simulated reality constructed using digital elements through computational systems. The construction and visualization of alternative reality require powerful hardware and software resources (e.g., immersion headsets, head-mounted displays, exoskeletons, power globes, 3D software, etc.) to make the creation of realistic immersive experiences in VR environments possible [24]. After many years of research and development, VR hardware and software are now available for use by the general public, researchers, and entrepreneurs.

VR can be classified into various categories, including virtual worlds, augmented reality (AR), mixed reality (MR), and extended reality (XR) [25]. Several researchers argue that VR and associated technologies contribute to stimulating creativity and enhancing various competencies. Furthermore, their presence has increased in the entertainment industry, generating interest in research areas such as education, knowledge management, psychology, engineering, architecture, among others. VR can be associated with Artificial Intelligence (AI) techniques to drive innovations in various fields, especially to enhance learning experiences by enabling direct experiences of real or imaginary phenomena [**?**]. As the design and development of VR and AI become more accessible to institutions, the levels of usage and dissemination could increase.

Liu et al. [27] argue that the essential elements characterizing VR are: Sensory immersion, through which users immerse themselves in a virtual world and can see their surroundings by using a Head Mounted Display. Interactive simulation involves recreating a virtual world and digitally representing physical objects through a computer, enabling implicit interaction that allows users a degree of control over their experiences using sensors and input devices like joysticks, keyboards, head-mounted displays, among others.

Immersive VR applications, where the user's perception is mediated entirely by visual and audio devices used in the virtual world. In this case, external information must be entirely isolated for the experience to be fully immersive. This technology can be costly and has some drawbacks, such as lower-definition images, computer overload, and environmental issues associated with simulators [28].

VR allows users to feel as if they were inside a simulated virtual world. Psotka [29] asserts that this contributes to a significant emotional factor, which facilitates cognition and improves information retention. The user's range of movement can also be expanded, allowing for navigation (useful for virtual tours). Immersion can be achieved in various configurations, such as using a large screen and anaglyph glasses, or in CAVE environments with a room containing four walls and a projector for each. Head-mounted displays (HMDs) are also used to project images that will be viewed by the user.

In this research, a VR immersion headset is used, which is comfortable for immersion due to its minimal space requirements compared to other devices. It also provides a greater sense of presence that contributes to memory retention and various cognitive processes [30]. However, it's important to note that in some users, VR can cause motion sickness, which is considered a significant disadvantage.

B. Natural User Interfaces (NUIs)

NUIs study new forms of human-computer interaction [31]. They are characterized by interactions that users find natural, common, and familiar. The design includes the use of new devices that make this type of interaction possible, not limited to just the use of a mouse and keyboard. They allow for gesture-based, touch-based, transduction-based, or body movement-based interactions, as well as voice command communication, among others. Innovations in input peripherals demand changes in the way we interact with digital screens and traditional interfaces like the mouse and keyboard, which are replaced by touch-based and motion-based interfaces [32].

NUIs represent a revolution in the field of computing, not only because they replace existing and widely used traditional interfaces but also because they enable the creation of new types of applications and original and innovative forms of interaction. These can be applied in various fields such as engineering, management, marketing, library science, and more.

C. Natural Devices Interaction

For [33], the emergence of new devices and interaction paradigms enables new forms of human-computer interaction, allowing users to interact with systems using body movements or transduction. Users can operate systems through intuitive actions such as gestures, touches, haptics, or speech. This offers a significant opportunity to enhance user experiences.

This literature review and the establishment of the state of the art have made it possible to recognize the immense potential of resources and elements in immersive technologies. Their use in proposals for interfaces with explicit knowledge databases will enable the conception and design of original and innovative models.

V. Related Works of Interaction Model Systems based on Immersive Technologies

The lack of flexibility in information systems has led users to employ their own strategies to carry out their daily activities in alignment with business objectives. Users fulfill their functions, enhance their performance, and save resources, especially time [34]. With the aim of formalizing the intentions of users of enterprise information systems for the execution of their daily tasks in line with their job roles and to improve their performance, Khodabandelou et al. [35] propose a method to discover the strategies users employ to achieve their personal objectives, aligned with business objectives.

The method is based on a supervised machine learning algorithm developed using general business activities and specific business rules. According to the authors, the model could be used to determine user behavior and enhance their performance. An advantage of this method is that it provides the option to work with flat or structured files. Additionally, the method allows for the structuring of a knowledge base that could be generalized for various types of businesses.

From another perspective, there is a considerable number of applications for exploring explicit knowledge databases that are designed to interact with traditional resources such as scientific databases [36]. However, there are some difficulties associated with the traditional environment. In the literature, several problems that hinder the search for scientific articles are reported, such as:

1) Traditional Devices: mostly have limited display space. It is assumed that newer generation LCD monitors have improved features [37], and they are currently the most commonly used means. Additionally, the visibility of LCD screens is lower when using an anti-reflective filter, so it is recommended for use when there is glare, a risk of visual fatigue, or when the user is performing tasks that do not depend on the detection of stimuli with low RGB values [38].

2) RepoVis: adopts the metaphor of visualizing source code listings hanging on a wall from a distance [39]. The complete visual description is the central piece of the structure (folders, files, and code) at a given moment in a software project. Source code files are represented as boxes with rows of colors that represent one or more lines of code.

3) Al-Amawi et al. [40]: propose the creation of a database to store explicit knowledge derived from lectures at a university. This database resembles a cognitive memory that grows and evolves over time, serving as a reference and repository of knowledge generated in the educational process. This is part of an e-learning system that provides benefits to students.

In the literature review, various works with different approaches were found, some of which make proposals that can contribute to the development of this research project.

VI. Model System Development

Below, the processes used for developing the different elements for building the model of the web repository interaction system based on immersive technologies and natural user interfaces are described.

A. Physical Environment Setup for Immersive Experiences

The equipment is configured, and the first step can be seen in Fig. 2 [15]. The Leap Motion was connected to the Oculus Rift, and then both devices were connected to a laptop using USB and HDMI ports. The installation software used was the Unity3D graphics engine.

(a) Schema setup (b) Actual operation

Fig. 2. Immersive environment configuration [5]

B. Usability for the Development of the Virtual Interface

Some usability principles should be considered for the design of the model in order to facilitate its use and proper handling of errors in its usage. Taking these principles into account in the design according to the defined and identified target audience should contribute to maintaining product usability quality to achieve maximum user satisfaction. One of the most important criteria for evaluating the usability of a virtual environment is the sense of presence. VR environments have the advantage of generating simulations that can act on the imagination of the participant, creating the sensation of psychologically transporting him/her to another place, which can be real or imaginary. Immersion in a VR environment with a high degree of presence will be an experience close to reality and even more pleasant. Tasks with a 2D interface use interactive output devices such as: keyboard and mouse, and interact with 2D objects and various tools (Marsh, 1999). Interaction in a VR system is carried out in a different way, with resources whose ergonomic evaluation criteria of the interfaces they use, is different from what most users are used to, helmets, glasses, gloves, virtual rooms, to mention a few.

Jordan [41] established usability principles, describing how and why they influence usability. Klevjer (2006) argues that the sense of presence in an environment can be enhanced through the use of an Avatar, which is a digital character that can resemble a real person, an animal, or take on another appearance as required, considering the model's requirements and user preferences [15]. The image shown in Fig. 3 corresponds to the avatar developed for this work, a resource that can enable user engagement, allowing them to perceive it as an assistant with whom they can interact using voice commands. The avatar assumes basic functions for exploring explicit knowledge documents contained in web-based databases.

Fig. 3. Avatar of the model system.

C. Natural Interaction System Model

After the immersive environment was set up, different techniques for interaction in immersive environments were investigated and explored:

1) Direct Manipulation: Technique implemented using a virtual hand to accurately emulate human hand movements. The selection of an element involves holding an object in the same way as in the real world. In this technique, elements are required to possess solid and physical properties for better interaction. Fig. 4, shows the selection of a node through direct manipulation; which as can be seen, is similar to holding a physical object in the real world.

2) Gestural manipulation: Immersive VR elements can be selected with virtual hand gestures. Holding an object with the pinch metaphor involves performing the 'Pinch' gesture and releasing it by performing the 'Stop' gesture using the

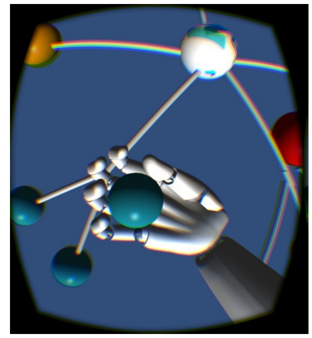
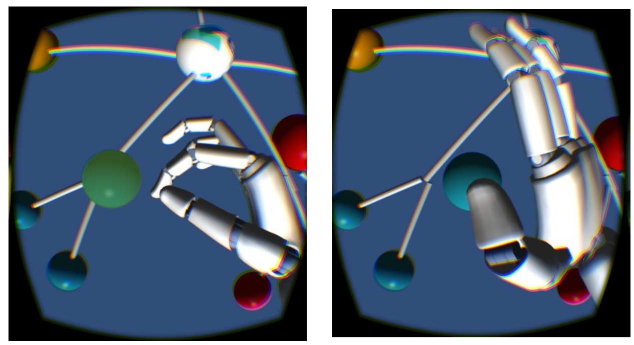


Fig. 4. Direct manipulation node selection.

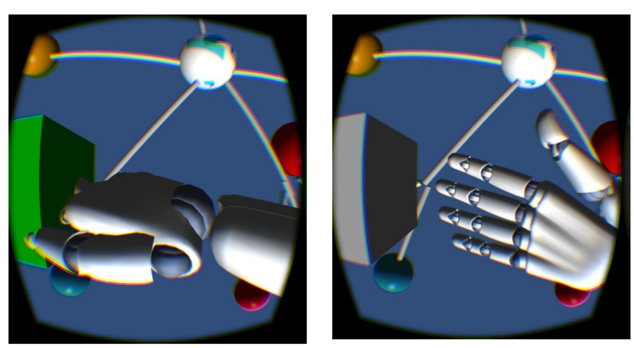


(a) Item Selecting (b) Element releasing

Fig. 5. Gesture of a node selection example.

open palm. In Fig. 5a, the node is selected by pinching, and in Fig. 5b, the palm is opened to stop selecting. This technique is not completely intuitive, because both gestures may mean something different to the user.

3) Artificial Manipulation: Winn [42], mentions the concepts of "Dimension", "Transduction" and "Reification", which allows the development of the corresponding metaphor, such as the crossing of objects. To leave the selection of the element, it is only necessary to go through it again. In Fig. 6a, a selected cube is visualized, changing color, and in Fig. 6b, the cube is visualized returning to its previous state.



(a) Item Selecting (b) Element releasing

Fig. 6. Artificial node manipulation example.

D. Implementation of Web Crawler for Information Retrieval

In order to carry out the tests, the interface model was connected to the ALICIA of Concytec scientific database and to enable the information retrieval, a program of the tracking type was developed and implemented, which allows automatic communication with the ALICIA and extracts the information required in the VR based interface through analysis of the source code of viewing concerning the respective visualization. The implementation involved investigating the pattern for performing the tracking and extraction for automated recovery.

E. Speech Recognition and Text-to-speech Engine

Intel Perceptual Computing features a library of raw audio-to-text conversion functions used to develop the recognition engine. To achieve accurate voice recognition, a certified wireless microphone was integrated into the prototype. A C# .Net library was developed to continuously capture user audio and then transmit it to Unity3D via Socket. A listening port was implemented to receive user voice commands.

F. Implementation of the Syntactic Analyzer

The Freeling library was used to perform the interpretation of the syntactic analysis of voice commands, taking into consideration the quality of the attributes of the functions for the treatment of the Spanish language. This process of syntactic analysis first involves breaking down the words in a sentence or phrase, to perform content analysis and add type labels: noun, verb, adverb, adjective, article, pronoun, etc. Fig. 7 shows an example of the analysis of a sentence, generating its respective syntactic tree.

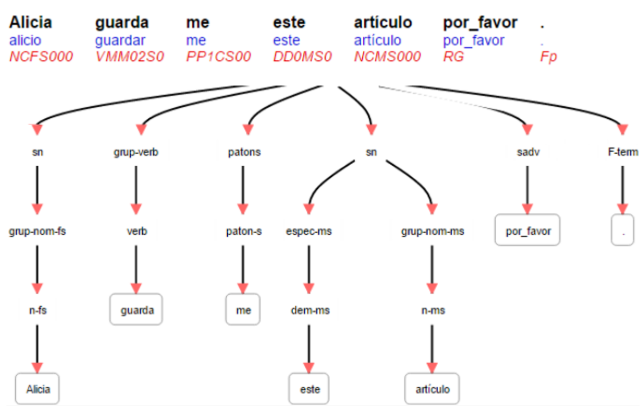


Fig. 7. Syntactic tree for a voice consultation generation.

The process diagram of the Proposed Model for document exploration in explicit knowledge databases, based on VR and NLP techniques, is presented in Fig. 8. The proposed model starts with a verbal query by a user who wants to search for a document in a web-based explicit knowledge database, then the speech recognition engine captures the speech signals and converts them into plain text. After the information is processed by the semantic analyzer, a syntactic tree is generated. Interaction and query commands are differentiated by the analyzer and then sent to the crawler, which establishes a connection with the virtual assistant to retrieve the information requested by the user. The results are visualized so that the user can begin

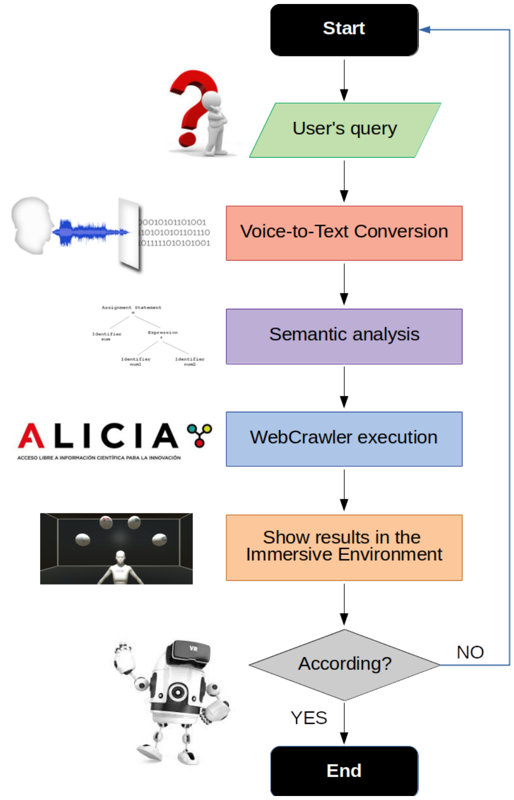


Fig. 8. Process diagram of the proposed model [5].

to interact. The process is concluded when the user agrees; otherwise, a new search is initiated.

VII. Evaluation of Proposed Model Regarding the Traditional

Usability is defined as the degree of satisfaction with which a product can be used by specific users to achieve specific goals, linked to effectiveness, efficiency, and satisfaction, within a particular context of use. The results of tests obtained using the proposed model in comparison to traditional models will be evaluated using the ISO 9241 standard, which pertains to interface usability. The dimensions to be evaluated include user-interface interaction, effectiveness, efficiency, and satisfaction.

A. Choice and Measurement of Variables

The following metrics are used:

1) Effectiveness: The accuracy achieved by users in relation to specified objectives also implies the absence of system errors and ease of learning and remembering. The metrics used include the number of important tasks executed, the number of relevant functions used, the number of tasks completed on the first attempt, the number of support requests, the number of accesses to help, and the number of functions learned.

2) Efficiency: This is the relationship of the resources employed (time, effort, etc.) with the accuracy and integrity with which the users achieve the specified objectives. The metrics used are: Time spent in the first attempt, time spent in relearning functions, productive time, time to learn characteristics, time to relearn, time spent in error correction.

3) Satisfaction: it is a subjective factor that takes into account the positive attitude towards the use of the product. The metrics used are: qualification of satisfaction with important features, qualification of ease of learning, qualification of error handling and the rate of voluntary system usage.

4) Preparation of Experiments with users: Once the devices are interconnected and the software modules are initiated, the evaluation of the document scanning model for explicit knowledge databases using the VR and NLP model, and the model with traditional interfaces of ALICIA (Web) were conducted. These evaluations were performed with five users, following Nielsen's [43] recommendations for interface evaluation. Each user was assigned the task of searching for scientific articles for thirty minutes using both interface models. During the experiment, users could request assistance from an expert, who was also responsible for documenting the experiment.

The first model evaluated was the exploration of scientific articles using the VR, NLP, and NUI-based model. In this model, the user immersed in the interface communicates with the article search mechanism through voice commands and can also interact with the elements being queried through touch. Fig. 9 shows a user utilizing the interface model, and Fig. 10 displays the information content presented to the user on the video monitor.

Fig. 9. Evaluation of the VR, NLP and NUI interface model.

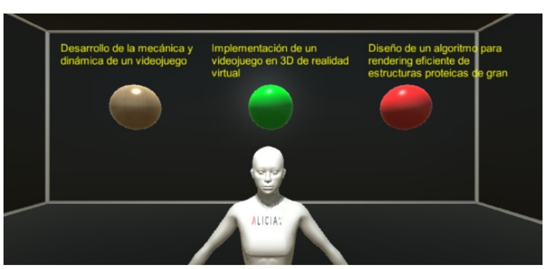


Fig. 10. View of the user with the model.

The second model to be evaluated is the traditional document exploration model, utilizing the Alicia scientific database interface. In this case, the user has a monitor for query visualization and uses the mouse and keyboard solely for interactions. Fig. 11 illustrates a user utilizing the traditional Alicia-Web interface, and what the user sees on the monitor is shown in Fig. 12.

Fig. 11. Evaluation of the traditional model of interface - Alicia Web.

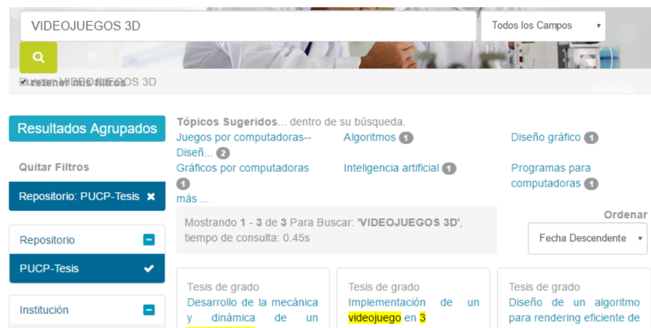


Fig. 12. View of the monitor with the information shown by the traditional interface.

VIII. RESULTS

The observations made on users using both types of interfaces allowed for the assessment of the effectiveness of each interface using the metrics outlined in ISO 9241. Effectiveness is defined as the achievement of goals or task completion, and it is represented by the number of goals accomplished. Various metrics count the goals achieved, tasks completed, functions used, etc. The results are tabulated in Table I.

The data from the results of using both interfaces for evaluating their effectiveness according to ISO 9241 are presented. Effectiveness is defined as the optimal use of the necessary resources to achieve a goal, where time is an important factor for the user of a software interface. The longer the time taken by the user to complete tasks using the interface, the lower the effectiveness. Conversely, if the time is shorter, the interface is considered more effective. The results are shown in Table I.

The satisfaction evaluation for each interface, considering the metrics defined by ISO 9241, takes into account that it is the degree or level of conformity of the user; it is their personal appreciation when evaluating the software interface. In order to measure software satisfaction in different aspects, the user must choose from five assessment alternatives: (1) Not at all satisfied, (2) Not very satisfied, (3) Fairly satisfied, (4) Very satisfied, (5) Completely satisfied. After each alternative of satisfaction there is a percentage of valuation corresponding to each alternative of satisfaction; these are respectively: 0.2, 0.4, 0.6, 0.8, 1.0. By averaging the levels of satisfaction, we can ultimately obtain the percentage of overall satisfaction required

TABLE I. COMPARISON OF EFFECTIVENESS USING THE TWO INTERFACES [15]

Metrics: Effectiveness Measure - ISO 9241	**Proposed model (Mean Quantity)**	**Web (Mean Quantity)**
Tasks performed	2.0	4.2
Functions used	3.8	7.8
Completed successfully Tasks at the first attempt	1.2	2.2
Important tasks performed	2.2	2.8
Calls for support	3.4	0.6
Access to help	0.4	0.2
Metrics: Measure of efficiency - ISO 9241	**Proposed model (Mean Quantity)**	**Web (Mean Cuantity)**
First attempt time	35.6	47.3
Time to relearn functions	31.2	16.6
Productive time	54.0	78.6
Learn characteristics time	16.6	31.2
Relearn characteristics time	54.0	78.6
Error correction time	55.8	54.6
Metrics: Measure of satisfaction - ISO 9241	**Proposed model (Mean %)**	**Web (Mean %)**
Qualification of satisfaction	0.72	0.32
Qualification of ease of learning	0.64	0.44
Error treatment qualification	0.24	0.28
Voluntary product use rate	0.92	0.64

in each metric.

IX. DISCUSSION

After processing the data obtained in the evaluation of the interfaces for each of the metrics following the ISO 9241 standards, allowing the assessment of the system in the dimensions of effectiveness, efficiency, and satisfaction, the “T-Student” statistical test was applied. This test revealed sufficient evidence that significant differences exist in the comparison of the evaluated dimensions:

- The proposed VR and NLP-based model showed low effectiveness compared to the traditional model of the scientific database ALICIA.
- The VR and NLP-based model exhibited higher efficiency compared to the traditional ALICIA database.
- Users expressed greater satisfaction with the VR and NLP-based model.
- The efficiency of the VR and NLP-based model was higher compared to the traditional ALICIA model.

The results obtained from the empirical analysis, based on user experiences with the VR and NLP-based model, generated considerable interest among users. They perceive it as a novel approach to interacting with explicit documents in institutional repositories, a proposal with significant development potential. Users faced initial difficulty with both types of models. Specifically, users who had never used ALICIA only utilized basic functions and not the tools and features available in the system. In contrast, users utilizing the VR and NLP-based interface initially struggled to adapt to the three-dimensional nature of immersion headsets. However, after a few minutes of adaptation, they expressed a desire to explore all available options and functionalities.

For exploration and document retrieval tasks, users employing the VR and NLP-based interface completed these tasks in less time than those using traditional web-based interfaces. It is noteworthy that a significant issue with VR immersion headsets is user discomfort.

X. CONCLUSION

An interface system model for explicit knowledge databases has been proposed and developed, based on natural language, natural users interface and immersive technologies processing techniques, which were successfully tested and found to be successful.

The scientific database ALICIA, was used for the evaluation of the proposed traditional model, using the ISO 9241 standard, considering Effectiveness, Efficiency and Satisfaction. The statistical test T-Student was applied to evaluate the data obtained by means of inferential statistics. The results allow to establish that the proposed model based on VR, NLP and NUI is superior to the traditional one in the dimensions of “Efficiency and Satisfaction” and is inferior in the dimension of “Effectiveness”. Therefore, it is concluded that the model of exploration of scientific articles with VR, NLP and NUI is better than the traditional model.

It is important to point out that currently, users of large databases have traditional interfaces based mainly on the mouse and keyboard. In the opinion of the authors, a greater effort should be made to research new interface models that are more efficient and effective and that have ergonomic and usability features to facilitate user interaction with the systems.

ACKNOWLEDGMENT

The authors would like to thank both the Universidad Nacional de San Agustín (UNSA), Arequipa – Perú, and the Universidad Nacional Mayor de San Marcos (UNMSM), Lima - Perú.

REFERENCES

[1] N. Dmytrenko, *Knowledge management in Erasmus + mobility projects teams*. (Master Thesis), University of Gothemburg, Sweden, 2015.

[2] M. Franco and F. Di Virgilio and L. Di Pietro, *Management of Group Knowledge and the Role of E-WOM for Business Organizations*. In Knowledge Management and Competitive Advantage: Issues and Potential Solutions, IGI Global, 2014, pp. 71-89.

[3] Y. Wang, *Using Machine Learning and Natural Language Processing to Analyze Library Chat Reference Transcripts*. Information Technology and Libraries, vol. 41, no. 3, 2022.

[4] G. Vladova and A. Ullrich and J. Bahrs and B. Bender, *Digitalisation and Enterprise Knowledge (net)Working*. Proceedings of the 51st Hawaii International Conference on System Sciences, 2018.

[5] R. Linares and J. Herrera and L. Alfaro, *AliciaVR: Exploration of scientific articles in an immersive virtual environment with natural user interfaces*. 2016 IEEE Ecuador Technical Chapters Meeting (ETCM), Guayaquil, 2016.

[6] M. Assante and L. Candela and D. Castelli and A. Tani, *Are Scientific Data Repositories Coping with Research Data Publishing?*. Data Science Journal, vol. 15, no. 6, 2016.

[7] T. Sakai and B. Flanagan and J. Zeng and T. Nakatoh and S. Hirokawa, *Search Engine Focused on Multiple Features of Scientific Articles*. 2012 IIAI International Conference on Advanced Applied Informatics, Fukuoka, Japan, 2012, pp. 214-217.

[8] S. Singh and S. Gupta and D. Busso and S. Kamboj, *Top management knowledge value, knowledge sharing practices, open innovation and organizational performance*. Journal of Business Research, vol. 128, pp. 788–798, 2021.

[9] A. Attia and I. Eldin, *Organizational learning, knowledge management capability and supply chain management practices in the Saudi food industry*. Journal of Knowledge Management, vol. 22, pp. 1217–1242, 2018.

[10] A. Mohamad and T. Ramayah and L. May, *Sustainable knowledge management and firm innovativeness: The contingent role of innovative culture*. Sustainability, vol. 12, no. 17, pp. 6910, 2020.

[11] M. Demuner-Flores and R. Nava-Rogel, *Gestión del Conocimiento al Interior de las Instituciones de Educación Superior*. GECONTEC: Revista Internacional de Gestión del Conocimiento y la Tecnología, vol. 6, no. 1, pp. 68-81, 2018.

[12] E. Hajric, *Knowledge Management System and Practices. A Theoretical and Practical Guide for Knowledge Management in Your Organization*. New York Press, 2018.

[13] C. Archer-Brown and J. Kietzmann, *Strategic knowledge management and enterprise social media*. Journal of Knowledge Management, vol. 22, no. 6, pp. 1288-1309, 2018, doi:10.1108/jkm-08-2017-0359

[14] P. Federico and M. Wagner and A. Rind and A. Amor-Amoros and S. Miksch and W. Aigner, *The Role of Explicit Knowledge: A Conceptual Model of Knowledge-Assisted Visual Analytics*. 2017 IEEE Conference on Visual Analytics Science and Technology (VAST), 2017.

[15] L. Alfaro and R. Linares and J. Herrera, *Scientific Articles Exploration System Model based in Immersive Virtual Reality and Natural Language Processing Techniques*. International Journal of Advanced Computer Science and Applications, vol. 9, no. 7, pp. 254-263, 2018, doi:10.14569/ijacsa.2018.090736

[16] E. Duperet and D. Pérez and M. Rodríguez and A. Ramírez and L. Montoya, *Importance of repositories for preserving and recovering information*. Medisan, vol. 19, no. 10, pp. 1283-1290, 2015.

[17] P. Jackson and I. Moulinier, *Natural Language Processing for Online Applications*. Text retrieval, extraction and categorization, vol. 5, 2007.

[18] D. Lewis and K. Jones, *Natural language processing for information retrieval*. Communications of the ACM, vol. 39, no. 1, pp. 92-101, 1996, doi:10.1145/234173.234210

[19] J. Gratch and J. Rickel and J. Cassell and E. Petajan, *Creating Interactive Virtual Humans: Some Assembly Required*. IEEE Intelligent Systems, vol. 17, no. 4, pp. 54-63, 2002.

[20] B. Xu and R. Cai and Z. Zhang and X. Yang and Z. Hao and Z. Li and Z. Liang, *NADAQ: Natural Language Database Querying Based on Deep Learning*. IEEE Access, vol. 7, pp. 35012-35017, 2019, doi:10.1109/access.2019.2904720

[21] F. Brad and R. Iacob and I. Hosu and T. Rebedea, *Data set for a Neural Natural Language Interface for Databases (NNLIDB)*. 8th International Joint Conference on Natural Language Processing, Taipei, Taiwan, 2017.

[22] N. Nihalani and S. Silakari and M. Motwani, *Natural language Interface for Database: A Brief review*. International Journal of Computer Science Issues, vol. 8, no. 2, 2011.

[23] A. Bhat and G. Bhagwat and J. Chavan, *A Survey on Virtual reality platform and its Applications*. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 4, no. 10, pp. 3775-3778, 2015.

[24] J. Martín-Gutiérrez and C. Mora and B. Añorbe-Díaz and A. González-Marrero, *Virtual technologies trends in education*. EURASIA Journal of Mathematics Science and Technology Education, vol. 13, no. 2, pp. 469-486, 2017.

[25] L. Alfaro and C. Rivera and J. Luna-Urquizo and E. Castañeda and J. Zuñiga-Cueva and M. Rivera-Chavez, *New Trends in e-Technologies and e-Learning*. 2021 IEEE World Conference on Engineering Education (EDUNINE), 2021.

[26] L. Alfaro and C. Rivera and J. Luna-Urquizo and S. Alfaro and F. Fialho, *Knowledge construction by immersion in virtual reality environments*. International Journal of Advanced Computer Sciences and Applications, vol. 10, no. 12, 2019, doi: 10.14569/ijacsa.2019.0101278

[27] D. Liu and K. Kumar and Y. Gao and T. Chang and R. Huang, *The Potentials and Trends of Virtual Reality in Education*. Springer Singapore, 2017, pp. 105–130.

[28] A. Alqahtani and L. Daghestani and L. Ibrahim, *Environments and system types of virtual reality technology in STEM: A survey*. International Journal of Advanced Computer Science and Applications (IJACSA), vol. 8, no. 6, pp. 77-89, 2017.

[29] J. Psotka, *Immersive training systems: Virtual reality and education and training*. Instructional Science, vol. 23, no. 5-6, pp. 405-431, 1995, doi:10.1007/bf00896880

[30] E. Lombardo and S. Agostinelli and M. Metge, *Could an interactive and total immersive device with hmd improve memory and give the presence sensation?*. Issues in Information Systems, vol. 14, no. 1, pp. 315-321, 2013.

[31] S. Mann, *Intelligent Image Processing*. John Wiley & Sons, Inc., 2001.

[32] A. Câmara, *Natural User Interfaces*. In Human-Computer Interaction – INTERACT 2011. Berlin: Springer, Berlin, Heidelberg, 2011.

[33] F. García-Peñalvo and L. Moreno, *Special issue on exploring new Natural User Experiences*. Universal Access in the Information Society, vol. 18, no. 1, pp. 1-2, 2019, doi:10.1007/s10209-017-0578-0

[34] O. Díaz-Rodríguez and M. Pérez, *Minería de intenciones a partir de una base del conocimiento y aprendizaje automático supervisado*. 3C TIC. Cuadernos de desarrollo aplicados a las TIC, vol. 10, no. 3, pp. 65-101, 2021.

[35] G. Khodabandelou and C. Hug and C. Salinesi, *A novel approach to process mining: Intentional process models discovery* International Conference on Research Challenges in Information Science, pp. 78-89, 2014, https://doi.org/10.1109/RCIS.2014.6861040

[36] C. Jiménez Noblejas and A. Perianes Rodríguez, *Recuperación y visualización de información en Web of Science y Scopus: una aproximación práctica*. Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información, vol. 28, no. 64, pp. 15-31, 2014, doi:10.1016/s0187-358x(14)70907-4

[37] M. Ghodrati and A. Morris and A. Price, *The (un) suitability of modern liquid crystal displays (LCDs) for vision research*. Frontiers in Psychology, vol. 6, no. 303, 2015, https://doi.org/10.3389/fpsyg.2015.00303

[38] A. Zunjic and L. Ristic and D. Milanovic, *Effects of screen filter on visibility of alphanumeric presentation on CRT and LCD monitors*. Work, vol. 41, pp. 3553-3559, 2012.

[39] J. Feiner and K. Andrews, *RepoVis: Visual Overviews and Full-Text Search in Software Repositories*. 2018 IEEE Working Conference on Software Visualization (VISSOFT), Madrid, Spain, 2018.

[40] A. Al-Amawi and S. Alsmarai and M. Maraqa, *Creating a Knowledge Database for Lectures of Faculty Members, Proposed E-Module for Isra University*. International Journal of Advanced Computer Science and Applications, vol. 6, no. 11, pp. 69-77, 2015.

[41] P. Jordan, *An Introduction To Usability*. CRC Press, 1998.

[42] W. Winn, *A conceptual basis for educational applications of virtual reality*. (Technical Report TR-93-9). Seattle, Washington: Human Interface Technology Laboratory, University of Washington.

[43] J. Nielsen, *Designing web usability: The practice of simplicity*. New Riders Publishing, 1999.