Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Arpiagrama: Supporting Visually Impaired Users in Creating and Understanding UML Diagrams

Pablo Nunes Vargas

pablo.vargas@unir.br

Universidade Federal Fluminense

Daniela Gorski Trevisan

Universidade Federal Fluminense

Research Article

Keywords: Visual impairment, Assistive technology, Non-textual representation, Multimodal interaction, UML diagrams

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11322708/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Arpiagrama: Supporting Visually Impaired Users in Creating and Understanding UML Diagrams

P.N. Vargas[1,2*] and D.G. Trevisan[1*]

[1*]Institute of Computing, Fluminense Federal University, Av. Gal. Milton Tavares de Souza, s/n, São Domingos, Niterói, 24210-346, Rio de Janeiro, Brazil.
[2]Academic Department of Computer Science, Federal University of Rondônia, BR-364, Km 9.5, Porto Velho, 76801-059, Rondônia, Brazil.

*Corresponding author(s). E-mail(s): pablo.vargas@unir.br; daniela@ic.uff.br;

Abstract

Purpose: Accessing and creating diagrams remain challenging for people with Visual Impairments (VI) and other stakeholders because most diagramming tools rely on visual and spatial interactions. This paper presents Arpiagrama, a multimodal assistive technology designed to support the creation and comprehension of Unified Modeling Language (UML) use case diagrams through 3D pieces, tactile exploration, computer vision, audio and visual feedback, and contextual descriptions.
Methods: Arpiagrama was evaluated in a mixed-methods study with ten participants: six students with VI and four sighted teachers experienced in educational contexts involving people with VI. Participants completed diagram creation and comprehension tasks, followed by the System Usability Scale (SUS), supplementary questions and observational analysis.
Results: All participants correctly constructed the expected diagram structure, while eight of ten also completed the final save action, resulting in an 80% full-task completion rate. The comprehension task achieved a 100% completion rate, and the mean SUS score was 91.0. Reflexive thematic analysis identified four themes: from abstraction to concrete multimodal representation, spatial construction strategies, interaction reliability, and configurable accessibility. Participants adopted different spatial arrangements while preserving the expected elements and relationships and reported challenges related to response time, auditory information, and external assistance.
Conclusion: The findings suggest that accessible diagramming systems should support creation and comprehension through distinct interaction mechanisms, accommodate semantically equivalent spatial arrangements, provide configurable multimodal feedback, and keep accessibility requirements open to revision. These results provide initial evidence that Arpiagrama can support accessible creation and comprehension of UML use case diagrams.

Keywords: Visual impairment, Assistive technology, Non-textual representation, Multimodal interaction, UML diagrams

1 Introduction

Non-textual representations play an important role in teaching and learning processes in fields such as computing, mathematics, and engineering because they make it possible to organize concepts, represent relationships, and communicate structures visually [26, 56]. In computing

education, for example, representations based on UML diagrams support the specification, analysis, and communication of structural and behavioral aspects of software systems [6, 34, 42]. However, access to non-textual representations still constitutes a significant barrier for people with VI [23, 24, 30, 50, 61].

Diagramming tools generally use graphical interfaces in which users perform actions within a two-dimensional workspace, relying predominantly on visual resources. For UML diagrams, some studies have explored tools designed for people with VI that involve the use of textual structures, often dependent on screen readers [12, 27, 46, 53]. Although screen readers are essential for interacting with digital interfaces, such as menus, buttons, and textual content, their sequential presentation does not always preserve the spatial and structural relationships inherent in non-textual representations. For people with VI, these structures and relationships require complementary mechanisms for exploration, navigation, and feedback, such as tactile graphics, audio-tactile interfaces, or other elements capable of supporting interaction with non-textual representations [4, 13, 17, 30, 40, 61].

Studies in the field of assistive technologies emphasize the importance of involving users throughout the design process because their experiences, needs, and contexts of use contribute to the development of solutions that are more appropriate, acceptable, and aligned with actual interaction practices [20, 25, 29, 35, 43, 44]. Some studies focused on people with VI mention the use of collaborative design approaches that seek to align the development of solutions more closely with users' actual needs [2, 3, 5, 18, 19, 31, 37–40, 47, 48, 60]. One approach that involves users from the beginning of the process is design thinking, a methodology that emphasizes empathy with users and rapid iteration [10], which may be promising for developing assistive solutions because of its collaborative nature.

Previous approaches have presented solutions based primarily on conversion to text, audio-tactile exploration, and creation through textual syntaxes. In general, these approaches support specific stages of interaction by addressing either the reading or the creation of representations separately. Therefore, the complete construction and comprehension of non-textual representations remain challenging [11, 33, 36], particularly for diagrammatic representations, for which the literature still offers few contributions [14, 24, 49]. In the context of UML diagrams, solutions that combine physical manipulation, tactile exploration, recognition of the structure during assembly, and generation of contextual descriptions within the same interaction flow remain underexplored.

To address this gap, the Arpiagrama tabletop system was developed as an assistive technology with a multimodal interface designed to support people with VI in creating and comprehending UML use case diagrams. Arpiagrama underwent a design thinking process involving teachers, students with total VI, and students with Low Vision (LV), with this article focusing on the testing and feedback phase. The system combines tactile exploration, audio feedback, adjustable visual representation, accessibility features, and artificial intelligence-supported generation of contextual descriptions. Through this process, the manipulation of the pieces is converted into a diagram representation that can be examined, described, saved, and shared, enabling multimodal comprehension of the resulting structures and relationships.

This study aims to investigate the extent to which Arpiagrama supports people with VI in creating and comprehending UML use case diagrams, considering task performance, perceived usability, and the interaction strategies adopted by the participants. To guide the investigation, the following research questions were formulated: To what extent does Arpiagrama enable people with VI to create and comprehend UML use case diagrams? How do participants perceive the usability and usefulness of Arpiagrama? What design implications does Arpiagrama offer for the development of assistive technologies involving diagrams?

This work makes two main contributions: the first is the development of a tabletop system that integrates 3D pieces, computer vision, and multimodal feedback to support the creation and comprehension of UML use case diagrams; the second is its mixed-methods evaluation with ten participants, relating task performance, perceived usability and usefulness, and design implications for the development of assistive technologies.

2 Related Work

This section discusses related work along three complementary lines relevant to the development and evaluation of Arpiagrama: accessibility of non-textual representations; assistive solutions for diagrams; and collaborative approaches involving people with VI in assistive technology. The first line examines recent studies on accessibility mechanisms for non-textual representations in general; the second presents assistive approaches specifically designed for diagrams and their contributions; and the third focuses on studies that applied user-centered design, co-design, or participatory design in the context of assistive technologies involving non-textual representations for people with VI.

2.1 Accessibility of non-textual representations

Among the digital approaches designed to provide access to non-textual representations, Thompson et al. [48] developed Chart Reader, a web-based tool that supports users with VI in reading and exploring three types of charts: single-series line charts, multi-series line charts, and stacked bar charts. The solution organizes chart components into an accessible structure, allowing navigation through regions such as axes, data points, filters, and contextual information, while combining other resources to support different comprehension strategies. In addition to this approach, Alam et al. [1] presented SeeChart, an interactive tool designed to improve the accessibility of data visualizations on web pages by generating an accessible natural-language version of charts. The design principles of these tools seek to make web-based charts accessible to users with VI through a range of accessibility mechanisms intended to promote their autonomy.

Another approach, Chart4Blind [32], consists of an intelligent interface designed to convert line charts into accessible formats, such as SVG or CSV, and generate alternative text. Maćkowski et al. [28] investigated the multimodal presentation of interactive audio-tactile charts for people with VI through a solution that combines a printed tactile image placed over a tablet screen capable of identifying the element being touched and providing contextual descriptions. The study also examined important parameters for tactile and auditory perception, including shape recognition, appropriate element width, intervals between touches, and description duration. Another recent tool is ChartA11y, proposed by Zhang et al. [60], which investigated accessible touch experiences for 2D visualizations on smartphones. The solution combines semantic navigation and direct touch mapping to support users with VI in comprehending line, bar, and scatter charts.

Ramôa and Müller [38] investigated a dynamic audio-tactile interface for exploring line charts on tactile graphic readers, specifically the Tactonom Reader. The proposed approach, called Melodic Line Tracing Exploration, uses melodies played by different instruments and variations in pitch to guide line tracing, support comparisons between series, and facilitate the identification of maximum and minimum points. Vargas and Trevisan [50] discuss barriers encountered by students with VI and teachers who work with people with VI when accessing charts, diagrams, and other images, using several instruments, including questionnaires, observations, and interviews. Their findings revealed different barriers for teachers and students with VI, highlighting the lack of public policies, the complexity of non-textual representations, and the need for solutions that integrate tactile resources, digital resources, AI, and collaborative environments.

2.2 Assistive solutions for diagrams

For diagrams, solutions such as Touch-and-Audio-based Diagram Access (TADA) demonstrate the importance of supporting the exploration of nodes, connections, and relationships through multimodal strategies, including touch, speech, sounds, and navigation between elements [61]. Another study that investigated accessible approaches to diagrams and representations based on relationships between elements was conducted by Yang et al. [58], who compared four tactile forms of presenting network data, including a braille list, an adjacency matrix, and node-link diagrams using two layouts. The results showed that node-link diagrams were preferred and more effective for tasks such as following paths and identifying clusters.

The studies by Cross et al. [14] and Wimer et al. [55] explored ways of making diagrams accessible by transforming their semantic structures into non-visual representations. Cross et al. [14] proposed an online system capable of defining specific diagrams and converting their semantics into text, with the possibility of speech output, to represent elements and relationships in a manner comprehensible to people with VI. Complementarily, Wimer et al. [55] presented Arboretum, a web-based system that generates accessible representations of data structures in tabular, navigable, and tactile formats while preserving the structural properties required to understand the representation.

In the study by Davis et al. [16], the authors developed TangibleCircuits, an interactive educational tool based on 3D-printed models and audio feedback to support people with VI in comprehending circuit diagrams. The system translates Fritzing diagrams into 3D tactile models and audio interfaces, allowing users to touch components and receive information about their names, positions, connections, and implementations. The authors observed that combining tangible interaction and audio feedback supported the comprehension of geometric, spatial, and structural information about circuits. Although focused on electronics education, the study is relevant because it demonstrates the potential of multimodal tangible interfaces applied to a type of diagram.

In computing, some studies have sought to make UML diagrams accessible to people with VI. Wildhaber et al. [53] developed a mobile solution for creating and editing class diagrams by combining an iPad application, iOS accessibility features, and a physical 3D-printed grid. The evaluation indicated a preference for the prototype in tasks involving the reading of classes and relationships, particularly among users without experience with textual languages, although more complex editing tasks still favored users experienced with PlantUML[1]. [12] reported the use of PlantUML in a systems analysis and design course, enabling a student with VI to create, edit, share, and collaborate on activities involving UML diagrams.

[1]PlantUML is a text-based tool that facilitates the creation of a wide range of diagrams, available at: ¡ https://plantuml.com/ ¿

In the same context, Serin and Romeo [46] proposed Latitude, an enriched text-based approach capable of generating HTML documents and accessible SVG diagrams by incorporating semantic descriptions to support screen-reader access to UML class diagrams. Verde et al. [51] presented Py2UML, a screen-reader-accessible web tool that enables the creation of UML class diagrams from Python code and seeks to support students with VI in constructing diagrams through an interface compatible with the NonVisual Desktop Access (NVDA) screen reader. In turn, Woodring et al. [57] explored an auditory approach to presenting relationships in UML class diagrams through sonification, using different auditory stimuli to communicate types of relationships between classes and reduce reliance on visual representations. These studies reinforce that access to diagrams by people with VI depends on preserving spatial and structural relationships, as well as providing complementary exploration and navigation mechanisms.

2.3 Collaborative approaches with people with VI

In the development of assistive technologies, user participation throughout the design process is essential for understanding specific needs, difficulties, preferences, interaction strategies, and application contexts. This perspective is reinforced by Ortiz-Escobar et al. [35], who emphasize the importance of involving users from the initial requirements-gathering stages through prototyping, evaluation, and solution refinement. The conclusions of Vargas and Trevisan [50] highlight the importance of a user-centered approach for understanding the different perceptions of teachers who work with people with VI and students with VI when developing an assistive solution involving both groups. Different studies have adopted collaborative approaches, such as user-centered design, co-design, and participatory design, to align assistive solutions more closely with the actual needs of users with VI and other stakeholders.

Shi et al. [47] adopted a user-centered approach with teachers of students with VI, involving design workshops, modification of 3D models, proposition of interactive elements, and an iterative process with teachers to design,

deploy, and evaluate interactive 3D models in an educational context. Using participatory design, Zhang et al. [60] developed ChartA11y through an iterative process with two blind collaborators who participated as active members of the team, particularly in designing multimodal and multisensory interactions for accessing visualizations on smartphones. Another study discussed previously used a five-month iterative co-design process to develop Chart Reader [48]. Ten users with VI participated in separate individual sessions in which they explored prototypes, provided feedback, and influenced subsequent versions of the tool.

In the educational field, Metatla et al. [31] investigated voice interfaces in inclusive schools through co-design with students with VI, students without VI, and their teachers, using focus group discussions, bodystorming, inclusive workshops, Lego construction, low-fidelity prototyping, and Wizard-of-Oz evaluation. Complementarily, Fan et al. [19] applied co-design with a statistics teacher, a teacher specializing in VI, and students who were blind or had LV, organizing sessions focused on identifying needs, exploring tools, prototyping, and evaluating learning activities. These studies contribute to Arpiagrama by demonstrating the importance of involving students with different visual conditions and educational stakeholders familiar with the content, teaching context, and barriers encountered within the same process.

For diagrams, Zhao et al. [61] presented TADA, which was developed using a user-centered approach consisting of interviews with people who were blind or had LV, through a design process involving the authors, including an author with VI who was also a diagram user. In developing their audio-tactile interface, Ramôa and Müller [38] adopted a human-centered participatory design approach involving two blind audio engineers in weekly discussions and design iterations. The main difference of Arpiagrama lies in the collaborative design process adopted through design thinking and in the integration of multiple stakeholders into the same assistive solution from the beginning of the design process.

3 Arpiagrama: Design and Interaction

This section presents Arpiagrama, a multimodal tabletop assistive technology developed to help students with VI and teachers create, explore, and comprehend UML use case diagrams. Its design was guided by requirements derived from previous phases of the research, considering barriers identified by students with VI, teachers, and specialists in the use of non-textual representations. These requirements include the need for multimodal feedback, autonomy, physical stability of the pieces, low cost, pedagogical flexibility, interface customization, and the possibility of expanding the system to other types of diagrams.

Arpiagrama was developed using a user-centered design thinking approach organized, according to Riley [41], into the phases of empathy, definition and ideation, prototyping, testing, and feedback. This process was adopted at a national reference center for people with VI so that the solution would be based not only on technical decisions but also on the actual experiences of people with VI and teachers involved in teaching non-textual representations. Development occurred iteratively, with data collected during the different stages guiding changes to the pieces, element attachment, accessibility features, and forms of interaction with the table.

During the empathy phase, different techniques were combined to understand the problem of accessing non-textual representations without initially directing the investigation toward a specific type of representation, such as diagrams. First, the literature review made it possible to map existing assistive solutions, the sensory modalities employed, and gaps related to reading and creating non-textual representations for people with VI [49]. This review identified fewer studies on diagrams and flowcharts than on other types of non-textual representations, such as charts, maps, educational images, and geometric shapes, which contributed to defining diagrams and flowcharts as the target representation types of this study. Questionnaires were subsequently administered, observations were conducted during a class on use case diagrams, and interviews were held with students with VI and teachers [50]. These activities revealed challenges such as

reliance on assistance from sighted people, the limitations of tools based exclusively on screen readers, the complexity of spatial relationships, and the need for adjustable multimodal resources.

During the definition and ideation phase, the data obtained in the previous phase were initially synthesized through empathy maps and personas, which were the only techniques applied without the direct involvement of participants with VI. These instruments helped represent three profiles central to developing the solution: blind students, students with LV, and teachers who work with non-textual representations in educational contexts. Two brainstorming sessions were subsequently conducted with teachers and students with VI, using adaptations of the Crazy Eights technique [22] into Crazy Four, as well as time adjustments and the use of smartphones or computers with their accessibility features instead of paper and pencil.

During these sessions, participants proposed ideas for an assistive solution involving tangible elements, low cost, navigation in parts, collaboration between people with VI and sighted people, multimodality, artificial intelligence resources, and user autonomy. The discussions made it possible to transform abstract needs into concrete design possibilities, converging on a solution involving a device that would facilitate joint exploration through interactive and tactile elements, with the UML use case diagram defined as its initial focus. Prototyping was then conducted, incrementally producing 2 prototypes during three brainstorming sessions with students with VI and teachers from the reference institute for people with VI.

The first prototype, used in two of these sessions, employed simple materials that could be assembled quickly, including a magnetic board, magnets, tape, sticks, pre-shaped Medium-Density Fiberboard (MDF), and simulated system responses using the Wizard-of-Oz technique [54]. This prototype made it possible to assess the initial feasibility of creating a table, using 3D pieces, and establishing interactions for creating and reading use case diagrams. Based on participant feedback, improvements were identified concerning piece attachment, tactile and visual differentiation, element size, stability, instruction clarity, and system responsiveness. A second prototype was subsequently created using 3D-printed pieces, velcro, adjustments to the organization of the table, and new interaction resources, bringing the proposal closer to the functional version of Arpiagrama. This prototype was used in a final brainstorming session employing the Substitute, Combine, Adapt, Modify, Put to another use, Eliminate, Reverse (SCAMPER) technique [15].

This design process resulted in Arpiagrama, an interactive table that combines tactile exploration with audio and visual feedback and generates contextual descriptions using AI to support the creation and interpretation of use case diagrams. The final solution reflects the decisions accumulated throughout the previous phases of the *design thinking* process, seeking to translate stakeholders' expectations and needs into design decisions. The following subsections describe the characteristics of Arpiagrama, including information about the components used, element organization, selected architecture, accessibility features, and implementation details.

3.1 System overview

The solution was named Arpiagrama, inspired by the harpy eagle, which is known for its keen eyesight, combined with the word "grama", a suffix found in the Portuguese words for both diagram and flowchart. The premise of the Arpiagrama table is to address both creation and reading tasks involving non-textual representations in the form of use case diagrams. Arpiagrama can be understood as an interactive and hybrid assistive system composed of people, 3D-printed pieces, input, output, and support devices connected to the table, and a mobile application responsible for the operations, detections, interpretations, descriptions, and records of the diagrams. The people include everyone directly involved in interactions with the table, such as teachers, students, and volunteers associated with the external assistance resource.

Figure 1 shows Arpiagrama ready for use, with its components installed in a room at a national reference institute for people with VI. The investment was approximately US$ 700. Figure 1a shows the components of Arpiagrama: a black rack-style table measuring 90 × 55 × 74.5 cm; 3D-printed pieces stored in containers with braille labels; a smartphone with wired video output; an articulated tabletop smartphone mount; a white background board measuring 60 × 40 cm with black adhesive velcro strips; a portable monitor with

a 15.6" touchscreen, Full HD 1080p resolution, and a 60 Hz refresh rate; an articulated tabletop mount for the portable monitor; a Bluetooth speaker; a wireless lapel microphone compatible with Android devices; a wired keyboard and mouse; and a 4-port USB hub for connecting the microphone, keyboard, and mouse to the portable monitor.

Overall, Arpiagrama operates in a collaborative environment in which physical objects, digital devices, and accessibility resources work together to support both students with VI and teachers. Interactions occur when users create or explore a diagram by manipulating 3D-printed pieces. The system then interprets and supports the spatial organization constructed by the users, returning information through audio and visual feedback while also enabling tactile exploration of this organization. This combination seeks to promote the autonomy of students with VI and encourage collaboration between teachers and students. It also facilitates the generation and sharing of diagrams in PDF and/or PNG, which are file types commonly used in academic contexts, as illustrated in Figure 1b.

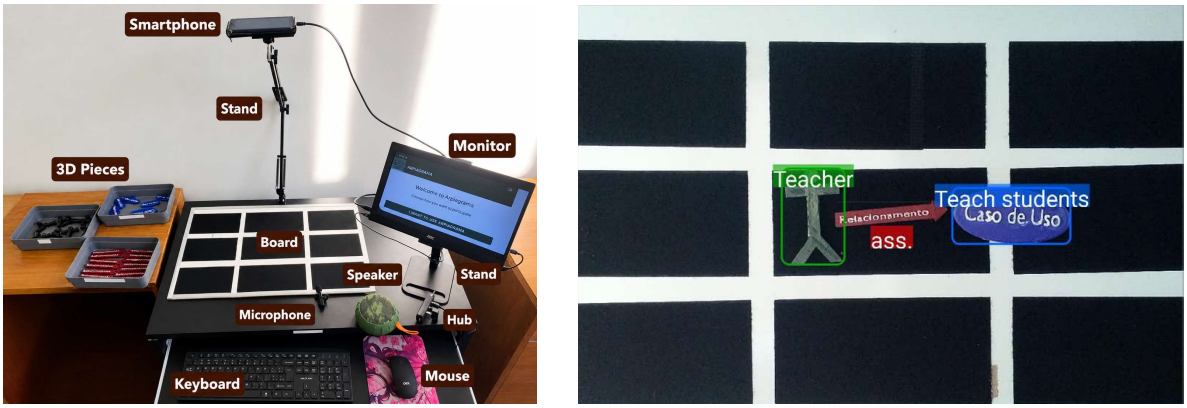
3.2 Layered architecture

The Arpiagrama architecture can be understood as a layered architecture composed of five main layers: physical, operational, acquisition, visual processing, and interpretation. Figure 2 presents this organization, in which most layers were implemented on the smartphone through the developed mobile application, except for the physical layer, to which the smartphone itself also belongs. This organization allows the physical manipulation of the 3D pieces to be captured, processed, and converted into an accessible digital representation capable of providing users with audio and visual feedback.

The first layer is the **physical layer**, consisting of the table, connected devices, whiteboard, 3×3 grid with braille labels, velcro strips, and 3D-printed pieces representing the elements of a use case diagram. This layer enables physical manipulation of the elements, supporting tactile exploration and spatial organization of the diagram and contributing to an understanding of the relationships among its elements. The grid formed by the whiteboard and velcro strips consists of 3 rows (A, B, and C) and 3 columns (1, 2, and 3). It serves as the manipulation and detection area, facilitating the location, orientation, and positioning of the pieces. For users with LV, contrast was created between the black velcro and the whiteboard. For blind users, tactile differentiation was provided between the rough texture of the velcro and the smooth texture of the board. braille labels were also placed immediately above the velcro to identify the cell represented by the strips (e.g.: A1, A2, A3, B1, B2, B3, C1, C2, and C3).

The 3D-printed pieces represent actors, use cases, and relationships and were produced in different sizes to support spatial organization and size preferences. Large and small pieces of each type were used, providing greater flexibility when creating diagrams. Regarding their dimensions, the use case pieces measured $11 \times 5 \times 0.6$ cm and $6.5 \times 3 \times 0.4$ cm; the actors measured $5 \times 9 \times 1$ cm and $2.5 \times 4.5 \times 0.4$ cm; and the relationships measured $11 \times 2 \times 0.6$ cm and $5 \times 1 \times 0.4$ cm. The evolution of these pieces incorporated adjustments resulting from the previously reported brainstorming sessions, particularly regarding stability, colors, tactile perception, and limitations related to the detection area and the pieces. An important observation is that the pieces had previously been printed with braille alongside conventional writing to indicate the type of each piece. However, after extensive discussion with participants during one of the brainstorming sessions, the braille was removed from the pieces because it occupied too much space. Furthermore, the shapes of the pieces facilitated their recognition.

The second layer is the **operational layer**, which encompasses the interface, interaction, control, service, and system accessibility components and mediates the flow among the other layers. This layer provides access to several resources, including contextual descriptions, data persistence, diagram visualization on the monitor, keyboard shortcuts, zoom, font size adjustment, a dark theme, audio speed control, and TalkBack activation. This layer is responsible for making the diagram accessible and customizable for different user profiles, including blind students, students with LV, and teachers. It also connects and controls the other layers, coordinating system actions according to user interactions and allowing diagrams to be saved or shared in formats such as PNG and PDF.



(a) (b)

Fig. 1: Arpiagrama setup with table, 3D pieces, equipment, and supporting materials. In (a), the table is ready for use shortly before the evaluation began. In (b), a test image in PNG format of a use case diagram generated by the table.

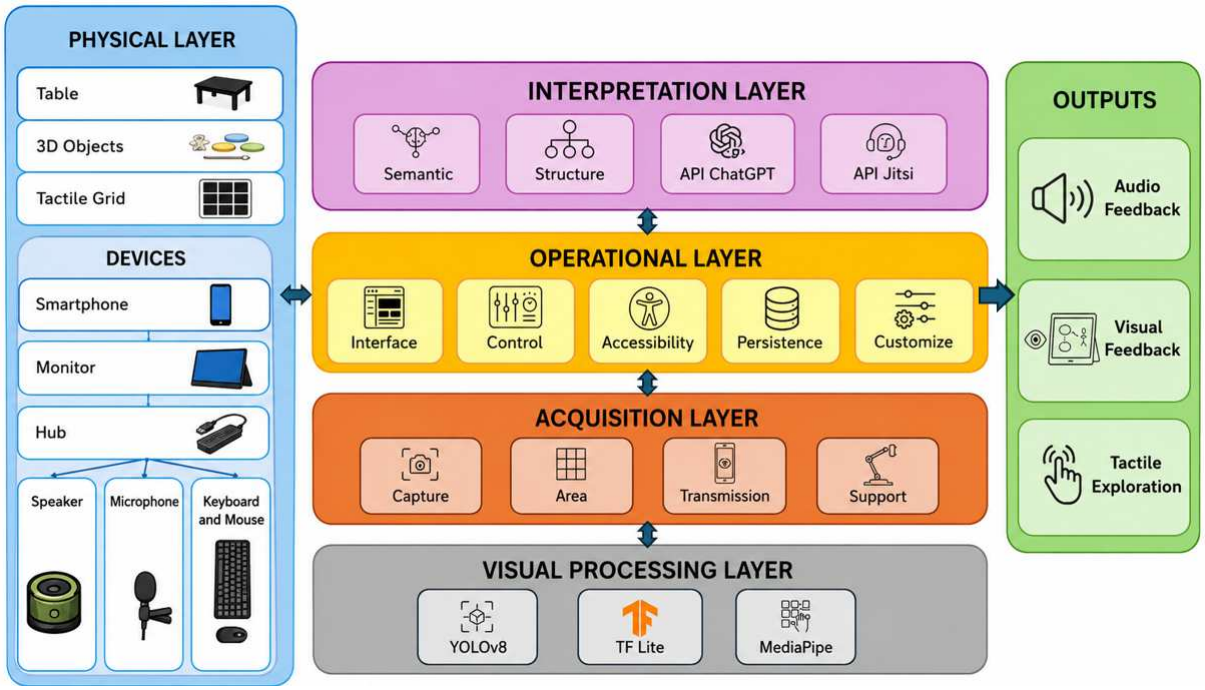


Fig. 2: Layered architecture of Arpiagrama

The third layer is the **acquisition layer**, consisting primarily of a smartphone secured to a 360º articulated mount at an approximate distance of 40 cm from the table. The device camera captures a top view of the diagram assembly area, recording the positions of the pieces on the board. The smartphone also serves as the system's main processing unit, centralizing image capture, execution, and communication with the other devices. During the evaluation, the smartphone transmitted its image to the portable monitor through a cable, allowing the visual representation of the

interaction to be followed, which may be useful for both teachers and students with LV. The important requirement is transmission compatibility among the devices (smartphone, monitor, speaker, microphone, keyboard, and mouse), allowing other forms of communication to be adopted, such as screen mirroring or Bluetooth.

The fourth layer is the **visual processing layer**, which processes the captured images and detects the 3D pieces within the board's detection area. In this layer, a model trained with YOLOv8[2] and exported to TensorFlow Lite[3] identifies actors, use cases, and relationships. To reduce false detections, the application considers a piece definitively detected only after it remains stable for approximately three seconds, after which it generates a bounding box and a textual indication of the element type. A piece is also not removed immediately: the system monitors its absence for approximately five seconds and evaluates possible occlusions caused by the user's hand with the support of MediaPipe[4].

The fifth layer is the **interpretation layer**, which transforms the individual detections of 3D pieces into a conceptual representation of the diagram, capable of interpreting specific elements rather than only the diagram as a whole. Thus, the system moves beyond working solely with objects detected in the image and organizes this information as semantic elements of a use case diagram, considering their types, definitions, redefinitions, and possible connections. For the structural recognition of the diagram, the ChatGPT API, model 5.2, was used. The image containing the diagram structure was sent to the API together with a prompt defining the output format, such as general descriptions of the diagram or specific descriptions of elements indicated by the user. Another interpretation-related resource was external assistance through Jitsi, which was still in the testing phase. Users were not required to use this resource, which relied on support from a sighted volunteer located separately from the participant.

[2]YOLOv8: a computer vision architecture applied to object detection, segmentation, and classification tasks, with support for deployment on different devices.

[3]TensorFlow Lite: an optimized framework for running machine learning models on mobile and embedded devices.

[4]MediaPipe: detects key hand landmarks in images or video streams, providing their coordinates and identifying whether each detected hand is left or right.

3.3 Accessibility and customization resources

Arpiagrama's accessibility resources were defined to support different forms of interaction with the non-textual representation, considering blind students, students with LV, and teachers who supervise or mediate educational activities. The solution includes three main features: contextual description, external assistance, and saving. In addition to being available through buttons in the interface, the table's main resources were associated with keyboard shortcuts: F1 for contextual description, F2 for external assistance, and F3 for saving. This choice sought to make access to the resources faster and more predictable, particularly for users familiar with keyboard navigation.

The F1 shortcut activates the diagram's contextual description, a resource designed to support comprehension of the structure constructed on the table by generating either a general description of the diagram or a specific description of an element. During prototyping, the activation of the available description modes underwent adjustments. Initially, a description of a specific element was expected to be activated after the user held a finger over a piece for several seconds. However, functional tests indicated unintentional activations when users were simply exploring the pieces by touch and kept their hands in the detection area for some time. Therefore, activation was restricted to the F1 shortcut or the corresponding interface button, while the software, supported by MediaPipe, began determining whether the description should be general or focused on an indicated element.

The F2 shortcut activates the external assistance resource, allowing users to request remote support through Jitsi to confirm interpretations, evaluate AI-generated descriptions, ask questions, or receive assistance while creating and reading the diagram. This resource requires Jitsi to be installed on both the Arpiagrama smartphone and the volunteer's device, as well as an active network of users and volunteers inspired by Crowdsourcing approaches[5]. The F3 shortcut allows users to save or share the diagram in formats such as PNG and

[5]Crowdsourcing: the practice of obtaining ideas, services, or solutions through the collaboration of multiple people, generally over the Internet.

PDF, accompanied by a textual description generated by ChatGPT. This prevents the exported artifact from being limited to a visual image and supports its subsequent use in educational activities, reviews, assessments, or sharing with teachers and classmates.

In addition to these shortcuts, the interface provides visual and audio customization resources. For people with LV, *zoom in* and *zoom out* were included, allowing the diagram view on the portable monitor to be enlarged or reduced. Font size adjustment and a dark theme were also added to improve legibility and reduce visual discomfort. For blind users, audio speed control allows the speech generated by the system to be adapted to the user's profile and experience, since more experienced users may prefer faster speech, whereas users who are still learning or working with more complex diagrams may choose a slower speed. The implementation of a custom TalkBack feature was also considered, allowing interface elements to be accessed similarly to a screen reader. Finally, all audio feedback generated by Arpiagrama can be interrupted using the ESC key, preventing auditory fatigue and giving users greater control during interaction.

3.4 Implementation details

The Arpiagrama software was developed as an application for Android devices, considering the ease of running it on a smartphone and integrating camera, audio, touch, accessibility, and external service communication resources. Choosing Android development provided greater portability and control over smartphone resources, with the smartphone serving as the solution's central processing unit. The Android Studio IDE was used, initially in the Narwhal version and subsequently updated to Panda 3, with the Android 15 SDK. Java was used as the primary programming language, with JDK 17 and Gradle for build automation, dependency management, and project organization. AndroidX libraries were also used to structure the application, manage screen lifecycles, and integrate native system resources, along with Material Design-based components to create part of the visual interface for users with LV and sighted users.

One of the central development stages involved creating a computer vision model to detect, on smartphones, the 3D pieces representing the elements of a use case diagram. For this purpose, images of the pieces were captured with and without the 3×3 board, including the elements together and separately, under different lighting conditions and from different angles. Roboflow[6] was used to assist with annotating the dataset, which initially contained 475 images: 316 for training, 82 for validation, and 77 for testing. During preprocessing, auto-orientation and resizing to 640×640 pixels were applied, along with data augmentation involving cropping and variations in zoom, saturation, brightness, and exposure. After this process, the resulting dataset contained 1183 images: 1062 for training, 81 for validation, and 40 for testing. The dataset was then exported in YOLOv8 format and used in Google Colab for training, validation of the best weights, and export to TensorFlow Lite in *float16*, including tests with confidence filters and *Non-Maximum Suppression* (NMS).

In addition to the YOLOv8 model exported to TensorFlow Lite, MediaPipe was used to detect hands and index fingers positioned over specific pieces, adding control mechanisms to the visual interpretation process. This resource made it possible to identify user interactions within the assembly area, associate pointing gestures with physical diagram elements, and prevent pieces from being incorrectly removed when occluded by the user's hands. Thus, the software receives camera images, identifies the pieces, records the detected structure, and produces visual and audio outputs. Room[7] was used for local data persistence, allowing structured information to be stored on the Android device, including user data, settings, and possible descriptions associated with the created representations.

Communication with external services was implemented using the OkHttp library, which structures HTTP requests, sends data, receives responses, and provides basic handling of communication between the application and the

[6] Roboflow: a platform for developing computer vision applications that provides resources for dataset organization, image annotation, model training, evaluation, and deployment for object detection and recognition.

[7] Room: provides an abstraction layer over SQLite, making data access and persistence on Android easier and safer.

APIs used. Among these integrations, the OpenAI/ChatGPT API[8] was used with model 5.2 to generate contextual descriptions of diagrams based on an image and a predefined prompt. When the system detects an index finger pointing toward a piece, information about the nearest element is sent, generating a description focused on that element and its connections. When no finger is detected, a general description is generated covering all elements and relationships in the assembled diagram.

Finally, native Android speech resources, particularly TextToSpeech and SpeechRecognizer, were incorporated to support audio output and voice input during interactions with the table. TextToSpeech converts textual information into audio feedback, allowing people with VI to receive system descriptions, instructions, and responses through audio. SpeechRecognizer enables the recognition of voice commands or input, expanding the available forms of interaction and reducing dependence on visual controls. The application is available on GitHub[9] and integrates Android development in Java, computer vision, local storage, authentication, communication with external services, automatic description generation, and audio and visual accessibility features.

4 User Study

This section presents the evaluation of Arpiagrama for the creation and comprehension of UML use case diagrams, corresponding to the final testing and feedback phase of the design thinking process adopted throughout the research. The evaluation was conducted as a mixed-methods user study, combining practical tasks with measures of task completion rate and completion time, as well as the System Usability Scale (SUS), supplementary closed-ended questions, and interviews guided by open-ended questions. The objective was to assess participants' performance in the proposed tasks and examine how they used the table's resources, the difficulties they encountered, and how the solution supported task completion. The following subsections describe the participants, tasks, study procedure, data analysis methods, and ethical considerations.

4.1 Participants

The evaluation involved ten adult participants (N=10), identified as P1 through P10 and presented in Table 1, along with their profile information. Six participants were students with VI, and four were sighted teachers with experience teaching or using diagrams in educational contexts involving people with VI. Among the students, two had low vision, two had congenital total visual impairment, and two had late-acquired total visual impairment. This composition enabled Arpiagrama to be evaluated from different user perspectives: students who interacted directly with the solution based on their experiences with VI and teachers who contributed a pedagogical perspective on teaching non-textual representations.

The participants comprised 3 females and 7 males and represented different educational levels, age groups, and levels of prior experience with diagrams, assessed on a scale ranging from none to excellent. Among the students, 2 participants had technical education and 4 had higher education, with experience levels ranging from little to reasonable and good familiarity with diagrams. Among the teachers, 3 participants held master's degrees and 1 held a doctoral degree, generally reporting good or excellent experience with diagrams. This diversity was considered important for observing the use of the table by people with different educational backgrounds and levels of experience with diagrams.

4.2 Tasks

The evaluation consisted of two main tasks: one involving the creation of a use case diagram and another involving its comprehension. Both representations were adapted from an academic textbook widely used in software engineering courses. The adaptations involved simplifying the names of the elements in both the creation and comprehension diagrams, as well as removing one actor and its associated use cases from the comprehension diagram. Figure 3 shows the use case diagrams employed in the evaluation tasks. While the first task evaluated diagram creation, the second

[8] OpenAI API: enables the integration of artificial intelligence models into applications, providing resources for processing and generating text, images, speech, and other multimodal tasks.

[9] Arpiagrama project, available at: ¡ https://github.com/pablonv/Arpiagrama ¿

Table 1: Participant Characteristics

ID	Gender	Education	Age	Type	Vision type	Diagram exp.
P1	Female	Higher education	18–29	Student	Low vision	Good
P2	Male	Technical education	30–39	Student	Low vision	Reasonable
P3	Male	Higher education	18–29	Student	Congenital blindness	Good
P4	Female	Technical education	30–39	Student	Congenital blindness	Good
P5	Male	Higher education	50–59	Student	Acquired blindness	Little
P6	Male	Higher education	30–39	Student	Acquired blindness	Little
P7	Male	Master's degree	30–39	Teacher	Not visually impaired	Excellent
P8	Male	Master's degree	40–49	Teacher	Not visually impaired	Excellent
P9	Female	PhD	40–49	Teacher	Not visually impaired	Good
P10	Male	Master's degree	50–59	Teacher	Not visually impaired	Excellent

focused on reading, exploring, and comprehending the previously constructed representation.

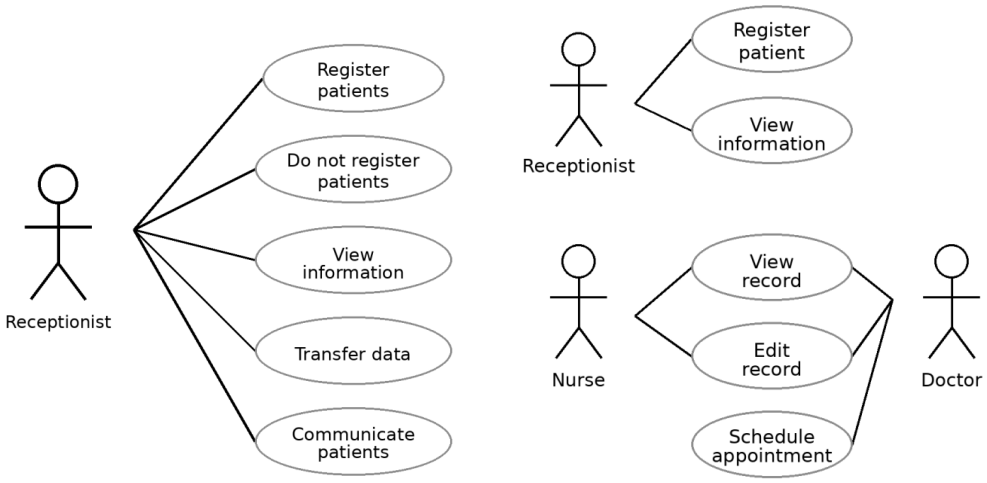
In the creation task, participants received an audio description of the use case diagram (Figure 3a) to be constructed and could request the description again as many times as they wished and in whatever order they considered necessary while performing the activity. At the end, they were required to save the created diagram in PDF or PNG format. The task had a maximum duration of 20 minutes and was used to assess participants' ability to generate a use case diagram and complete a save action, which is common in academic contexts and when using AT. If a participant correctly completed the diagram but did not remember to save it, the completion time was recorded at the moment of the last movement made with the pieces, but the task was considered incomplete.

The task was considered successfully completed when the diagram was constructed as expected, comprising 23 correct actions according to the description and representation provided: 11 connected elements, 11 completed definitions, and one save action. Layout differences were permitted, provided that the relationships among the elements were preserved, the assigned names were semantically similar to those expected, and the diagram was saved. Missing or incorrect pieces were counted according to the number of missing and/or incorrect elements. Failure to save the diagram was counted as one missing action, whereas extra elements corresponded to the number of inserted pieces that should not have been present in the representation.

The second task involved comprehending a use case diagram previously assembled on the table, using a representation (Figure 3b) that was slightly more complex than the one used in the first task, because the creation task entails a higher cognitive load [59]. Participants could explore the representation exclusively through Arpiagrama and were required to respond to five true-or-false statements concerning the diagram's elements and relationships (e.g., The physician is connected to the schedule appointment use case). This task also had a maximum duration of 20 minutes. The use of external assistance, which was still in the testing phase and involved one external volunteer, was optional. The task was considered successfully completed when all responses were correct. Incorrect true-or-false responses and unanswered statements were counted as errors if the time limit expired before the participant had responded to one or more of the statements.

4.3 Procedure

Before the full evaluation was conducted, two pilot studies were carried out: one with a sighted participant wearing a blindfold to test the interaction flow without visual access, and another with a participant with low vision. The pilot studies aimed to assess participant fatigue, refine the questions, evaluate the suitability of the tasks, and confirm the adequacy of the equipment used in the study. Adjustments were made to the questions, the training provided, and the developed resources to better accommodate the needs of participants with VI. The first pilot study, involving



(a) Diagram used in the creation task. (b) Diagram used in the comprehension task.

Fig. 3: Examples of UML use case diagrams used in the Arpiagrama evaluation.

the blindfolded participant, was conducted at Fluminense Federal University (UFF), whereas the second took place at the Benjamin Constant Institute (IBC), a national reference center for people with VI.

Participants were invited in person and through instant messaging applications. After their initial acceptance, the study sessions were scheduled. All sessions took place at the IBC in March and April 2026 and lasted an average of 1 hour and 17 minutes per participant. Before the tasks were performed, the researcher introduced the study, explained the overall evaluation procedure, and provided an electronic form containing the Informed Consent Form (ICF), the profile questionnaire, and the initial instructions. After agreeing to the ICF, which authorized audio and/or video recording, participants completed approximately 15 minutes of prior training with Arpiagrama to become familiar with the main components of the table, the 3D pieces, the available resources, and the available shortcuts.

Participants then individually performed the two evaluation tasks. Completion times and task completion rates were recorded for both tasks, along with observations made during the interactions. As previously mentioned, the first task involved creating a diagram, and the second involved comprehending a previously assembled diagram. Participants could request a 15-minute break before or after either task. After the first task, participants temporarily left the room so that the table could be prepared for the second task, which took approximately five minutes on average. Finally, participants completed the SUS questionnaire and answered open- and closed-ended questions about their experience with Arpiagrama.

4.4 Data analysis

The data analysis combined quantitative and qualitative methods to evaluate Arpiagrama as an assistive solution for working with use case diagrams by people with VI and other stakeholders. For the tasks, completion time and task completion rate were analyzed quantitatively, whereas observations of the interactions and the diagrams created were analyzed qualitatively. In the creation task, completion time and the expected, correct, extra, and incorrect/missing attributes relative to the requested diagram were quantitatively recorded. In the comprehension task, completion time and the expected, correct, and incorrect/missing attributes were recorded. These data made it possible to assess participants' performance in creating and comprehending the diagrams, as well as the completion rate achieved in each task.

Perceived usability was analyzed using the SUS [9], which consists of ten statements rated on a five-point scale ranging from strongly disagree

to strongly agree. The final score was represented on a scale from 0 to 100, enabling a quantitative interpretation of the overall perceived usability of Arpiagrama, as well as specific qualitative discussions of the questions for which responses varied most among participants. In addition to the SUS, supplementary closed-ended questions were used to quantitatively assess specific aspects of the user experience. These questions included participants' comparison of Arpiagrama with other assistive technologies, as well as their perceived usefulness and ease of use of the contextual description and external assistance features. The responses to these items were summarized using frequency distributions, allowing the analysis of how participants perceived the usability and usefulness of these specific resources. These questions were provided in the same form used for the SUS.

For the qualitative analysis, Reflexive Thematic Analysis (RTA), inspired by Braun and Clarke [8], was applied to an integrated corpus comprising responses recorded in the evaluation form, photographic records, researcher observations, and audio transcripts of the interviews with the ten participants. The open-ended questions were: What did you like most about the Arpiagrama table? Why?; What most hindered or bothered you while using the table?; and If you could change or add something to the table today, what would it be? The analysis adopted a predominantly inductive, semantic, and experiential orientation and followed the phases of familiarization, initial coding, theme construction and review, definition and naming, and production of the analytical narrative.

Initial coding was conducted by the first author through iterative readings of the transcripts, open-ended responses, and observational records. Coding was organized in a traceability spreadsheet and a codebook that was revised throughout the process rather than constituting a fixed scheme defined a priori. The final thematic structure was subsequently reviewed and discussed among the authors to assess its internal coherence, conceptual distinctiveness, and support within the corpus. This review was not intended to calculate inter-coder agreement but served as a reflexive stage for discussing and refining the interpretation. Photographic records were used as complementary qualitative data to describe observable spatial characteristics of the diagrams and relate them to the interview transcripts and observational records.

4.5 Ethical considerations

Participation in the study was voluntary and conditional upon agreement with the ICF, which was made available through Google Forms. Participants could choose the authorized form of recording, including participation with both audio and video recording or with audio recording only. Confidentiality was preserved through the use of pseudonyms, identifying participants as P1 through P10, without disclosing their real names in the records. The study was registered on Plataforma Brasil under number 76931024.9.0000.5243 and approved by the Research Ethics Committees (RECs) of UFF and IBC. The latter was the co-participating institution responsible for providing access to most of the target population of this study. In addition to the pilot studies conducted to assess the aforementioned suitability, measures were adopted to reduce the risk of identifying participants in visual materials, including anonymization or the application of filters to images in which people appeared.

The open-ended responses and audio recordings were analyzed in an integrated manner, preserving the meaning of the participants' statements and making only minimal grammatical adjustments or removing repetitions when necessary. Participants could also provide their final feedback either on site or remotely to reduce possible biases associated with the researcher's presence during the evaluation. Before any use of the artificial intelligence assistance tool, the data were anonymized by removing information that could identify the participants. The interpretations, analytical decisions, final definition of the themes, and validation of the results remained the responsibility of the researchers.

5 Results

This section presents the results of the Arpiagrama evaluation based on three main aspects: performance in the creation and comprehension tasks, perceived usability and usefulness, and the RTA findings. The results include quantitative

measures, such as completion time, task completion rate, and SUS score, as well as qualitative data from observations made during the study, responses to the administered form, and transcripts of the participant interviews. Overall, Arpiagrama received a positive evaluation, with most participants completing the tasks within the established time limit and achieving high completion rates in both tasks. Furthermore, all participants obtained SUS scores above the reference value of 68 points, which is commonly used as a benchmark in SUS evaluations [45]. The RTA resulted in four themes: from abstraction to concrete multimodal representation, spatial construction strategies, interaction reliability, and configurable accessibility.

5.1 Task performance

Overall, most participants completed both proposed tasks with high completion rates and within the established time limit, demonstrating Arpiagrama's potential to support both tasks. Table 2 presents participant performance in the two tasks, considering execution time (in minutes:seconds) and the other recorded attributes. In the creation task, participants were required to construct a use case diagram from a provided description, involving 23 expected actions: 11 connected elements, 11 definitions, and 1 save action. Of the 10 participants, 8 performed all actions correctly, whereas P3 and P6 each had only one missing action. Participant P3 constructed the diagram accurately but finished exactly at the 20-minute limit, leaving no time to perform the save action. Participant P6 also accurately completed the creation task in slightly more than 13 minutes but did not remember to save the diagram image.

No extra elements were recorded in the creation task for any participant. Overall, 228 of the 230 expected actions were completed correctly, corresponding to an overall action completion rate of 99.1%, whereas the full-task completion rate was 80% (8/10 participants). Regarding completion time, the creation task had a mean duration of 12min43s, ranging from 8min19s to 20min00s, noting that participants were not required to complete the task as quickly as possible. P8, a sighted teacher who chose to perform the tasks while blindfolded, completed the creation task in 13min33s. The longest times were recorded by students P3 and P4, both of whom had congenital total VI, whereas the shortest time was recorded by P9. Another noteworthy observation concerns the second-shortest time, which was recorded by P1, one of the students with low vision, at 09min12s.

In the comprehension task, all participants correctly answered the five true-or-false statements, totaling 50 correct responses out of 50 expected responses; thus, the task completion rate was 100%. The mean completion time was 6min24s, ranging from 0min59s to 13min07s. Participants P5 and P6, both students with late-acquired total VI, used the external assistance feature during this task and recorded the longest completion times, although they answered all statements correctly. Among the teachers, P8, who performed the tasks while blindfolded, again recorded the longest completion time, although it was close to that of P7, who recorded the second-longest time among the teachers, with a difference of only 21 seconds. Once again, P9, a teacher, completed the comprehension task in the shortest time. Among the students, P4 recorded the shortest time, at 06min13s, followed by P1 and P3, at 06min50s and 06min59s, respectively.

A comparison of the two tasks showed that creation required approximately twice the mean completion time of comprehension and resulted in a lower full-task completion rate. However, these differences should be interpreted cautiously because the tasks involved distinct operations and completion criteria. Although the comprehension diagram was structurally more complex, participants interacted with an already assembled representation and focused primarily on exploring and interpreting its elements and relationships. In contrast, the creation task required participants to interpret instructions, select and position pieces, define elements, establish connections, monitor the evolving structure, and save the final diagram. Therefore, the longer completion times observed during creation appear to reflect the greater number and variety of operations required rather than necessarily indicating greater difficulty. Overall, participants with different profiles were able to complete both types of activities using Arpiagrama.

Table 2: Task Performance in the Arpiagrama Evaluation

ID	Time (min:s)		Creation task				Comprehension task		
	Creation	Comprehension	Expected	Correct	Extra	Missing	Expected	Correct	Missing
P1	09:12	06:50	23	23	0	0	5	5	0
P2	12:24	07:59	23	23	0	0	5	5	0
P3	20:00	06:59	23	22	0	1	5	5	0
P4	15:42	06:13	23	23	0	0	5	5	0
P5	14:10	13:07	23	23	0	0	5	5	0
P6	13:09	12:24	23	22	0	1	5	5	0
P7	10:35	03:12	23	23	0	0	5	5	0
P8	13:33	03:33	23	23	0	0	5	5	0
P9	08:19	00:59	23	23	0	0	5	5	0
P10	10:04	02:44	23	23	0	0	5	5	0

5.2 Perceived usability

The perceived usability of Arpiagrama was evaluated using the SUS and several closed-ended questions concerning the experience compared with other AT, as well as the usefulness and ease of use of the available features. Figure 4 presents the distribution of responses to the positive and negative SUS items. For the negative items, most responses were concentrated in the disagreement options, indicating a favorable evaluation of Arpiagrama. For the positive items, the concentration of responses in the agreement options reinforced this favorable perception. The mean score was 91.0 points, with a median of 91.25, a standard deviation of 6.69, and a range from 77.5 to 97.5 points. These results indicate an overall positive perception of the solution, with all participants assigning scores above 68 points, a commonly used benchmark value. The highest SUS scores were recorded by students P2, P3, and P5, all of whom obtained 97.5 points. Only 2 teachers scored below 90 points: P7 and P9, with 77.5 and 82.5 points, respectively. Participants P1, P6, and P10 obtained 90 points, whereas P8 obtained 92.5 points and P4 obtained 95 points.

The analysis of the SUS items revealed particularly positive aspects related to ease of use, integration of the table's functions, and confidence during interaction, as 9 of the 10 participants strongly agreed with the items representing these aspects. Furthermore, all participants strongly disagreed with the statement that the table was confusing to use, while 9 of 10 disagreed that the table was unnecessarily complicated. These responses indicate that most participants perceived the table as relatively uncomplicated and easy to use. Other positive SUS items also showed a high concentration of strongly agree responses (8 of 10 participants), particularly those related to frequent use of Arpiagrama and how quickly participants could learn to use it.

The item with the greatest variation concerned the need for assistance from someone with technical knowledge. However, it is important to note that most participants disagreed with this statement and that those who agreed were teachers. Thus, the students with VI generally disagreed that such assistance was necessary, whereas the teachers' responses suggest that this perception should be investigated more thoroughly. Only one other negative SUS item received a single agreement response, concerning the need to learn before using the Arpiagrama table. This response was provided by a student with late-acquired total VI and little experience with diagrams. Although the mean SUS score was high, the responses to the individual items show that some participants still perceived a need for training or support during use.

The supplementary closed-ended questions also indicated positive perceptions of the table's features and enabled comparisons between Arpiagrama and other ATs used in similar contexts. Compared with other assistive technologies (e.g., screen readers, PlantUML, and BeMyEyes), 7 participants rated their experience with Arpiagrama as "much better", 2 as "better", and 1 as "the same". The latter response was provided by P5, who had little experience with diagrams. Figure 5 summarizes participants' ratings of the usefulness and ease of use of the contextual description and external assistance features. The contextual description was rated as "extremely useful" by 7

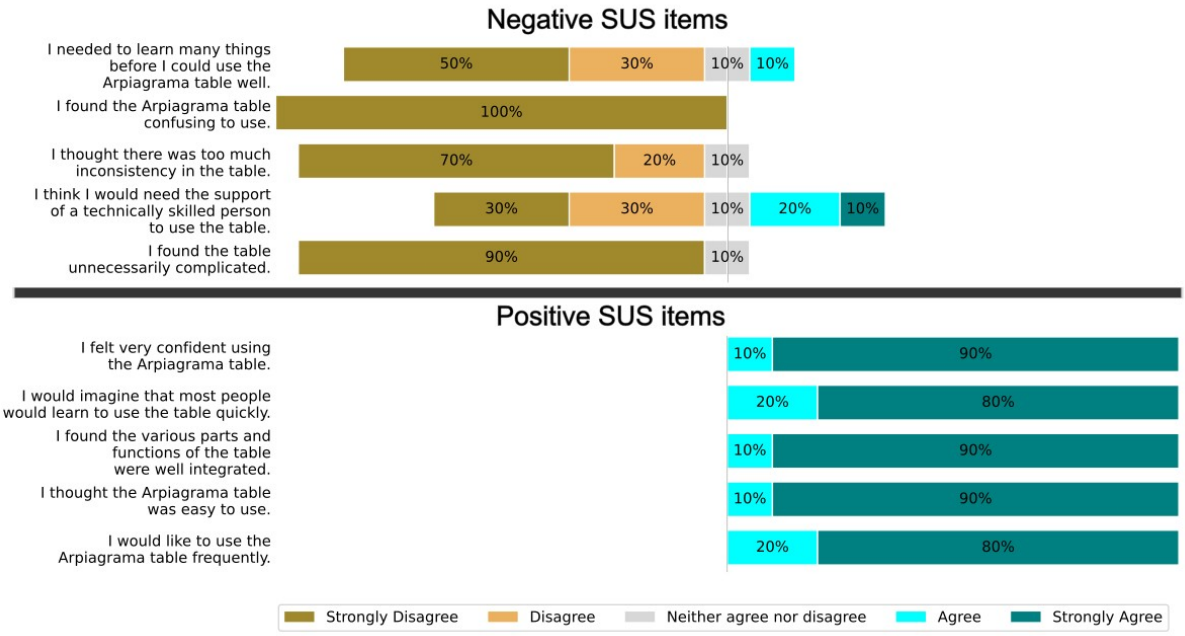


Fig. 4: Distribution of responses to the positive and negative SUS items in the Arpiagrama evaluation (N=10).

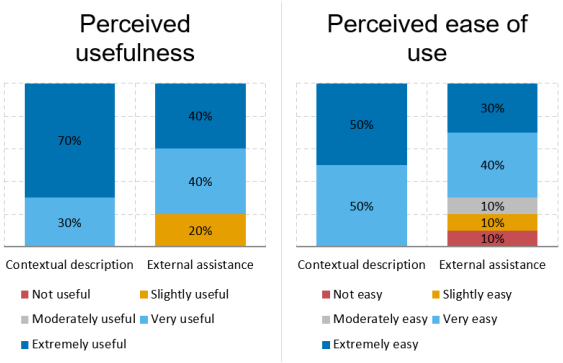


Fig. 5: Participants' ratings of the usefulness and ease of use of the contextual description and external assistance features.

participants and "very useful" by 3. All participants also considered this feature easy to use: 5 rated it as "extremely easy" and 5 as "very easy".

External assistance was positively evaluated by most participants, although the responses showed greater variation than those for the other questions. It should also be noted that this feature was not mandatory, although all participants activated it at least once. Regarding usefulness, 4 participants rated it as "extremely useful", 4 as "very useful", and 2 as "slightly useful", namely P4 and P9. Regarding the ease of use of external assistance, 3 considered it "extremely easy", 4 "very easy", 1 "reasonably easy", 1 "slightly easy", and 1 "not at all easy". The negative ratings were provided by P1 and P4, respectively. It is noteworthy that participants P5 and P6, who used external assistance during the comprehension task, also evaluated it positively. These results indicate that the table's features were recognized as important complementary support, although they still require adjustments in light of the negative evaluations reported by some participants, particularly regarding external assistance.

5.3 Thematic findings

The RTA integrated the interview transcripts, open-ended responses from the evaluation form, researcher observations, and photographic records collected during the study with the ten participants. The analysis resulted in four related themes: from abstraction to concrete multimodal representation, highlighting the tangible

and multimodal characteristics of the table; spatial construction strategies, reflecting the different arrangements used to construct semantically equivalent diagrams; interaction reliability, concerning limitations related to fluidity, predictability, and robustness; and configurable accessibility, encompassing participants' preferences and suggestions regarding personalization, alternative forms of interaction, and future expansion.

5.3.1 From abstraction to concrete multimodal representation

The first theme brings together accounts concerning the combined use of 3D pieces, tactile exploration, and the audio and visual feedback provided by Arpiagrama. The central idea of this theme is to understand how exploration through 3D pieces, complemented by audio and visual feedback and supported by additional features, transformed the use case diagram into something manipulable, comprehensible, and pedagogically valuable. Participants therefore highlighted not only the 3D pieces but also the integration of touch, spatial organization, audio, attachment, and, for participants with low vision and sighted participants, aspects of the visual interface.

Participant P1 valued the "interactivity with the materials", as well as the ease and simplicity of attaching and manipulating the pieces representing the use case diagram elements on the table using velcro. Complementing this theme, P2 highlighted the possibility of "exploring by touch", showing that manipulating the 3D pieces was important during the interactions. Based on his experience with low vision, he also stated that touch and residual vision complemented each other. The complementarity between modalities was reinforced by P3, who considered the "tactile and auditory feedback" important for enriching the construction process and enabling people with total VI or low vision to form a mental representation of the diagram. Similarly, P4 stated that the table made it possible to obtain a "tactile sense" of how the diagram was structured, meaning that tactile exploration contributed to the formation of a mental representation. P5 also valued tactile exploration for identifying the diagram elements, whereas P6 reported better comprehension through the type of interaction used on the table than through "typing" and "writing".

Among the teachers, P7 highlighted how the system "detects the element" of the use case diagram, valuing the computational recognition and the methods used to name the pieces. P8, a sighted teacher who performed the tasks while blindfolded, emphasized the importance of the pieces being "tactile", having distinguishable shapes, and being complemented by the proposed feedback, as expressed in the following statement: "because it also has audio. [...] it is easy to identify...I felt the stickman. That was easy for me". P9, the fastest participant overall, highlighted that the audio description facilitated comprehension of the elements and constituted a relevant teaching resource for diagram instruction. Finally, P10 stated that "nothing [...] hindered" his use of the table, emphasizing its pedagogical potential in light of the reported difficulty of demonstrating this type of diagram. Participants also highlighted its potential to contribute to teaching and learning this type of diagram and to increase the inclusion of students with different profiles.

Thus, Arpiagrama functioned as a hybrid solution capable of making use case diagrams explorable through different sensory modalities, using the manipulation of 3D-printed diagram elements to support the creation and comprehension of an accessible non-textual representation. The shapes and raised features of the pieces supported the recognition of diagram elements, the velcro kept these elements fixed during tactile exploration, and the system identified and named the components. Audio provided instructions, confirmations, and descriptions, primarily supporting people with total VI, whereas the visual interface supported participants with low vision and teachers.

5.3.2 Spatial construction strategies

The second theme concerns how Arpiagrama guided diagram construction without imposing a single layout, allowing spatially distinct yet semantically equivalent structures to be constructed using different strategies. The photographic records captured what the indirect observations had indicated, as participants used different approaches to organize the elements of the same diagram during their interactions. Figure 6 illustrates the different spatial structures adopted

by participants during the use case diagram creation task. These different structures showed that the same semantic diagram could be constructed using distinct organizational strategies on the tactile grid. Participants P1, P2, P5, and P6 produced similar spatial arrangements, with the actor placed near the central region, the use cases distributed around it, and the relationships positioned in different directions, resulting in a predominantly radial configuration, as shown in Figure 6a and Figure 6d.

Participant P3 was the only one to use the lower-right corner of the table, placing the actor at the edge and forming a highly condensed fan-shaped arrangement with the remaining relationships and use cases, as shown in Figure 6b. Participants P4, P7, P8, and P10 constructed the diagram's spatial structure using the first column of the grid as a reference for the actor element, as shown in Figure 6c and Figure 6e, arranging the other elements with more spacing and in a fan-shaped configuration. P9 used the middle column as a reference for the actor element, grouping the other elements in a fan-shaped and slightly more condensed arrangement, as shown in Figure 6f. P8 adopted a different strategy from the other participants, who generally began the construction with the actor and then sequentially placed a relationship and a use case, or a use case and a relationship. Although P8 also began with the actor, he preferred to place all the use cases first and the relationships afterward, as shown in Figure 6e.

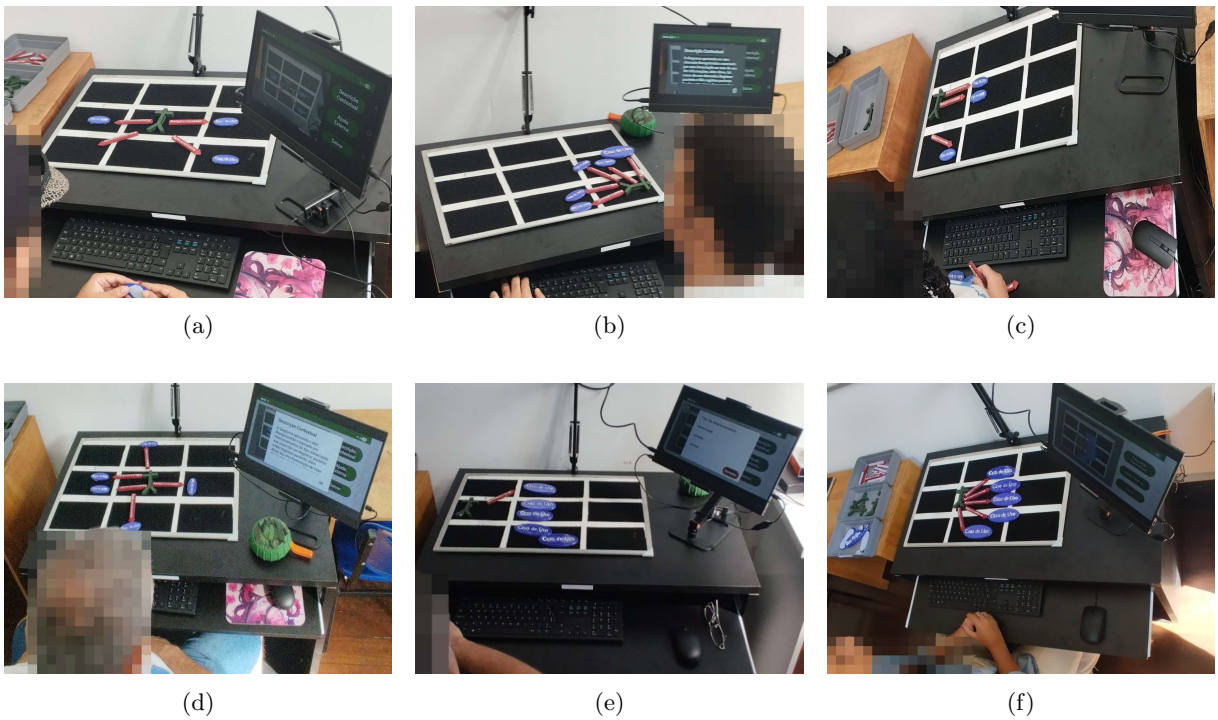
These differences suggest that participants appropriated the interaction area according to their own strategies for reaching, grouping, and representing elements. However, the effects of these spatial strategies on task completion time or efficiency were not directly evaluated in this study. The interview transcripts also contributed to this theme. P1 stated that the grid divisions helped maintain orientation while positioning the pieces. P2 highlighted the possibility of directing the elements wherever desired, allowing users to diversify the spatial structure of the same diagram. Meanwhile, P5 emphasized the spatial orientation provided by the tactile grid, which helped locate the diagram elements and influenced his exploration strategy.

Participant P7 valued the possibility of personalization, regardless of the "point" or "position" at which the pieces were placed, providing users with flexibility when structuring the diagram. P9 stated that the contextual description feature "greatly facilitates understanding of the arrangement of the elements", demonstrating that the feature provided spatial guidance concerning the diagram and the locations of its elements. Thus, the system should recognize and describe semantically equivalent diagrams even when they differ in the position, distance, orientation, and density of the pieces. Overlap or organization alerts may be provided as optional support, but the arrangement and semantics chosen by the user should be preserved as part of their strategy.

5.3.3 Interaction reliability

Although the overall evaluation of Arpiagrama was positive, negative usability aspects related to the speed, stability, and predictability of the system's responses were identified, resulting in the third theme. Participants P1, P3, P6, P7, and P8 reported waiting or delays at different moments while interacting with the Arpiagrama table, which affected the fluidity of the interactions. P1 mentioned "that delay", referring to the delay in activating external assistance, and had also indicated in the closed-ended questions that this feature was slightly difficult to use. For P3, the problem concerned the "delay in the feedback" during interactions, suggesting that the system's response time should be reduced. It should be noted that a piece-detection period begins when the pieces remain in a stable position before semantic definition is activated.

P6 described the delay more specifically by stating that it was necessary to "wait a little because of the AI's slowness", highlighting the moment when the contextual description was used, despite rating this feature as "extremely useful" and "very easy" in the closed-ended questions. P7 complained about the "time it [...] takes to register", associated with the time required to open the microphone when a piece is defined through this mechanism. He observed that, as a sighted participant receiving visual feedback, the process was "faster". Although P8 was uncertain about the moment at which a "certain delay" occurred, he stated that he would change the "response time".



(a) (b) (c)

(d) (e) (f)

Fig. 6: Examples of different spatial arrangements adopted by participants during the use case diagram creation task with Arpiagrama. In (a), the participant with low vision created the diagram with the actor positioned centrally and the other elements arranged around it; in (b), the participant with congenital total VI used the edge of the detection area to construct the diagram; in (c) and (d), both participants had late-acquired total VI but chose different ways to construct the diagram; in (e) and (f), both participants were sighted teachers who attempted to follow the visual structure shown in Figure 3a, but arranged the elements in different columns.

Furthermore, although P4 stated that nothing hindered her, she subsequently mentioned a "lack of familiarity" with the Arpiagrama table, demonstrating the need for familiarization and adaptation. The same participant provided negative responses regarding the external assistance feature, rating it as "not at all easy" and "slightly useful". P5 highlighted the verbosity of the messages, stating that the voice provided "too much information [...] at the same time". This participant also reported that the external assistance feature occasionally froze, despite rating it as "very easy" and "very useful" in the closed-ended questions. P5 also reported occasional problems with attaching the pieces, stating: "[...] sometimes they do not stick properly. But that is because the velcro will be used a lot". Therefore, actions need to be made more predictable and features more reliable, and the maintenance and attachment of the pieces should also be examined.

Participant P7 identified insufficient contrast in the focus indicator when navigating with the TAB key, even after activating the dark theme, indicating the need to increase contrast to improve visibility. P9 was uncertain about when to use voice, the keyboard, touch, or the buttons available in the interface, indicating that, even when different interaction modalities were supported, it was not always clear which one could be used. Furthermore, the same participant stated that "when there is a need to ask questions, it is because it (the piece) is not self-explanatory", referring to the absence of braille on the pieces and demonstrating the importance of providing complementary methods for identifying the elements. The accounts revealed challenges primarily related to response time, indication of the system state,

and the operation of external assistance. These aspects should guide future changes to the table, particularly regarding reduced latency and clearer communication of actions.

5.3.4 Configurable accessibility

The fourth theme brings together the opportunities identified for improving accessibility. These opportunities do not indicate a single configuration but rather alternatives that allow users to adapt the solution to different profiles and contexts. Audio speed exemplifies this theme, given the importance of enabling the customization of this mechanism, as noted by participants such as P3, P6, and P8. These participants also recognized that people with different profiles may have different speed preferences. It should be noted that audio speed could be configured during the study. Therefore, the solution should provide user profiles, independent speed controls, and options to increase or decrease verbosity, potentially using new shortcuts, such as F5 to reduce the speed and F6 to increase it.

An alternative version of the Arpiagrama table emerged from P1's suggestion of a portable "mini-table" the size of a notebook, whereas P2 proposed a version of the solution in which the main devices would be a webcam and a desktop computer. P4 identified a challenge related to expanding the solution, expressed as follows: "[...] space is something that helps you, but later it also hinders you", referring to increasing the size of the detection area and relating it to the desired diagram size. P7 suggested modifying the "relationship" piece, stating that it would be useful to "change the color of the tip" to represent different types of relationships. P10, in turn, proposed expanding the solution to entity-relationship diagrams, in line with other suggestions involving different UML diagrams and representations used in education.

The creation of alternative braille pieces was also discussed by teachers P8, P9, and P10, particularly for pieces representing relationships and use cases. The proposed optional sets of braille pieces would not include conventional writing, containing only braille and optimizing the available space on the piece. It should be noted that the brainstorming conducted during earlier phases of the design process included pieces printed with both conventional writing and braille, but not pieces containing only braille. The participants' suggestions showed that a single configuration may not equally meet the needs of all table user profiles. Alternatives were mentioned regarding the presence of braille, the size of the construction area, audio speed, and the amount of information presented.

6 Discussion and Design Implications

The results observed with Arpiagrama align with previous evidence that multimodal interfaces can support access to non-textual representations, including Tactile Graphics with a Voice [4], TangibleCircuits [16], TADA [61], and other audio-tactile interfaces [28]. These studies indicate that tactile resources can support the perception of shapes, sizes, spatial arrangements, and structural relationships, while audio can complement tactile exploration by providing names, definitions, connections, and contextual descriptions that may be difficult to convey through touch alone. In addition to focusing on UML use case diagrams, Arpiagrama is a tabletop assistive technology designed to support interaction among users with different visual profiles. Although the evaluation tasks were performed individually, the results also highlighted the participants' ability to interact with the system independently.

The creation task required approximately twice the mean time of the comprehension task. However, this result does not directly indicate greater difficulty because the activities differed in their structure, required operations, and completion criteria. During creation, participants had to interpret instructions; locate, define, and position elements; establish connections; monitor the structure; and save the diagram. During comprehension, they began with a completed diagram and focused on exploring and interpreting it. Zhang and Fiorella [59] indicate that creating representations may require more time and cognitive effort than comprehending them. Similarly, Wildhaber et al. [53] observed longer times for creation and editing tasks than for reading tasks, although participants preferred the evaluated prototype for most of the questions.

As a design implication, diagram creation and comprehension for people with VI should

be treated as related but operationally distinct modes, balancing time, accuracy, and user satisfaction. During creation, the system should provide confirmation of insertions, identification of elements that have not yet been defined, methods for undoing and redoing actions, and mechanisms for verifying the structure, which are also part of the comprehension activity. During comprehension, priority should be given to tactile exploration among the diagram elements, contextual descriptions, and external assistance features. Furthermore, both modes should provide interaction alternatives that, together with the aspects mentioned above, can reduce cognitive load without fragmenting the experience, because the activities remain related within the same system even though they involve distinct forms of operation.

The photographic records in Figure 6 showed that the same diagram was constructed using different strategies, resulting in distinct spatial arrangements while preserving the same semantic structure. The tactile grid provided references for locating and positioning elements without imposing a single spatial layout. Variations were observed in the region occupied, the distance between elements (density), orientation, and piece size. The importance of spatial references is also evident in SETUP09, which uses a 3×3 grid with directions, cell references, and points to support people with VI in exploration and diagram creation [21]. Furthermore, an accessible approach to teaching UML concluded that the ability to control the spatial organization of a diagram is important for diagram creation by both sighted users and people with VI [12].

Thus, the Arpiagrama results indicate that different users employ various construction and exploration strategies and that the semantics of the elements, connections, and spatial references should be preserved. This perspective is similar to the Arboretum approach, which prioritizes preserving the structural properties of diagrams while providing different forms of representation for the same diagram [55]. Regarding size, the study by Watanabe and Sato [52], which used 3D pieces of varying complexity, found that different sizes generated distinct perceptions of the pieces' architecture. In network representations, different organizational forms offered advantages depending on the task performed, with node-link diagrams being particularly effective for following paths and identifying clusters [58].

Something similar occurs in use case diagrams, which use relationship elements to establish connections among the other diagram elements, making it essential to "navigate" through these elements. Accordingly, one of Arpiagrama's strengths was tactile exploration, which allowed participants to move their hands across the diagram elements. Furthermore, the pieces had different sizes, and clusters could be identified with the support of contextual descriptions or external assistance. As a design implication, assistive systems should recognize different arrangements and element sizes as equivalent when the semantics of the elements and connections are preserved. The interface may provide tactile cells, lines, coordinates, and regions as references without requiring predetermined positions.

Despite the positive performance indicated by the quantitative data, the open-ended responses revealed challenges related to Arpiagrama's usability, demonstrating the contribution of the mixed methods used in this research. The quantitative measures indicated acceptance and task success, whereas the qualitative data identified specific issues that could affect user autonomy and experience. In the SUS analysis, the items with less positive indicators were related to ease of learning and the need for support. The supplementary open- and closed-ended responses broadened this interpretation by revealing difficulties primarily related to interaction response time, the amount and speed of information, and the operation of external assistance. Participants noted differences between voice speed and the total time required to complete an operation, as Arpiagrama provided voice-speed controls on the home screen.

Therefore, increasing the voice speed during interaction did not, by itself, eliminate the waiting time associated with detection, processing, or microphone activation, all of which occur repeatedly during use. One possible approach is illustrated by SeeChart, which allows users to adjust the speed and length of descriptions and to listen to or repeat content sentence by sentence [1]. Chart Reader similarly combines hierarchical navigation, different levels of information granularity, verbal descriptions, and sonification to support user exploration [48]. These studies suggest that no single accessibility feature

is sufficient to accommodate the different exploration strategies observed during the study. As a design implication, assistive solutions should provide multiple accessibility mechanisms that can be configured throughout the interaction. By default, they should present essential information first, provide additional details on demand, and use auditory cues to communicate system states such as detection, processing, confirmation, error, and saving.

Other preferences related to the presence of braille on the pieces, the size of the detection surface, portability, and devices reinforced that a single accessibility configuration does not equally serve all users and contexts. Some teachers suggested alternative sets of pieces with braille identification, particularly to support students with VI during the initial stages of learning and when recognizing diagram elements. Participants also discussed that a larger surface could accommodate more complex diagrams but would increase the tactile exploration distance and introduce technical challenges related to reach, piece location, and computational recognition.

The need for alternatives in Arpiagrama is consistent with Bornschein et al. [7], who, although they did not address the same type of non-textual representation, did not identify a universally preferred input modality for image creation. Pandey et al. [36] showed that, although audio labels were faster and easier to create, braille supported flexible exploration strategies for tactile graphics. In addition to the accessibility mechanisms already discussed, the solution should also enable control over other mechanisms generally used by people with low vision, as suggested by Alcaraz-Martínez et al. [2] regarding adjustments to contrast, legibility, resizing, and navigation.

Participants expressed different preferences regarding piece size, the presence of braille, and the amount of information. Therefore, these aspects should be configurable in future versions of Arpiagrama. Based on these findings, future versions could operationalize personalization through persistent profiles (concise, standard, and detailed) containing settings for message speed and level of detail, state sounds, configurable contrast, adjustable font size, on-screen zoom, and preferred input method. The concise profile would present only essential information, prioritizing speed and a smaller amount of verbal feedback. The standard profile would balance speed, amount of information, and level of detail. The detailed profile would provide more comprehensive descriptions, instructions, and confirmations, supporting novice users or more complex tasks.

The results also reinforce the importance of using a collaborative design process, showing that the involvement of the target population from the initial phases of the process, through various instruments, contributed to generating an accessible and positively evaluated solution. Other technologies, such as Chart Reader [48], ChartA11y [60], and TADA [61], were developed through iterative processes involving people with VI, allowing user strategies and preferences to influence successive increments. In Arpiagrama, features such as velcro, 3D pieces, the tactile grid, audio feedback, and customization resulted from previous cycles involving students and teachers. It is important to highlight the renewed discussion of braille, demonstrating that accessibility requirements remain open to revision throughout the process. As an implication, design decisions should record the participants, tasks, and contexts that supported them while preserving opportunities for alternatives and reassessment in later phases.

In summary, the results indicate four design implications for tabletop assistive technologies for diagrams. First, creation and comprehension should be treated as related but operationally distinct modes, providing specific mechanisms for constructing, confirming, and verifying the structure, as well as for exploring and interpreting the diagram. Second, the system should preserve semantic equivalence among different spatial arrangements, providing orientation references without imposing a single structure. Third, system states should be perceptible through clear and configurable feedback, with essential information presented first, additional details provided on demand, and distinct cues for the main actions. Finally, accessibility features should remain configurable and open to revision throughout the design process, with users continuing to be involved in subsequent stages of evaluating, revising, and refining the assistive technology.

7 Limitations and Future Work

The evaluation conducted with Arpiagrama presented limitations related to the sample, which consisted of 10 participants from a single institution with different visual conditions and profiles, preventing statistical comparisons between groups. Furthermore, the creation and comprehension tasks were performed in a single session after training and in the presence of one of the researchers, which may not represent independent and continued use of the solution. The two tasks also differed in the operations required, diagram complexity, and completion criteria.

A possible ceiling effect in the tasks should be investigated because the tasks may have been sufficiently simple to limit variation among participants or may have provided excessive time for completion. The table has limitations related to the detection area, which directly affects the size of the representations it can accommodate and poses a challenge for complex diagrams containing many elements. Other limitations are related to computer-vision-based piece detection because it depends on ambient lighting conditions. Furthermore, the technologies used to create the table depend on an Internet connection, paid AI services, and a crowdsourcing structure to maintain external assistance.

Future work should involve larger samples and multiple institutions, longitudinal studies in authentic educational activities, and experimental comparisons with alternatives such as PlantUML, static tactile graphics, and textual descriptions. Future studies should also consider counterbalanced task orders, pieces with braille identification, deeper investigation of spatial strategies, and faster and more robust detection mechanisms. Suggestions to expand the system to entity-relationship diagrams, class diagrams, and flowcharts indicate pedagogical potential and possibilities for future work that the Arpiagrama table could accommodate. However, this expansion will require consideration of the specific properties of each representation, such as arrow direction, relationship types, groupings, and definition methods.

8 Conclusion

This work presented Arpiagrama, a multimodal assistive technology developed to support the creation and comprehension of UML use case diagrams by people with VI. In the mixed-methods evaluation with ten participants, 8 of the 10 participants fully completed the creation task, corresponding to a full-task completion rate of 80%, with 228 of the 230 expected actions completed correctly. In the comprehension task, all participants successfully completed the activity by correctly responding to the five proposed statements. The mean SUS score was 91.0. These results provide initial evidence that integrating 3D pieces, tactile exploration, audio and visual feedback, and contextual descriptions can support both diagram creation and comprehension. In addition to the quantitative results, observations and interviews revealed different interaction strategies and spatial arrangements adopted by participants while using the table.

The 3D pieces and velcro grid supported the location and organization of diagram elements, while audio provided element identification and descriptions, and the visual interface supported sighted participants and those with low vision. In the evaluated context, the grid provided orientation references without imposing a single spatial arrangement, while different arrangements preserved the expected elements and relationships. At the same time, latency, the amount of auditory information, and difficulties with external assistance showed that effective interaction with Arpiagrama also depends on fluid, predictable, and configurable feedback.

During the study, the combination of tangible elements, spatial references, audio feedback, and visual information supported different ways of exploring and interacting with the diagrams. Participants' preferences regarding braille, piece size, and audio feedback further reinforce the importance of providing configurable options and user-specific alternatives. Arpiagrama and its design process may inform future assistive technologies that seek to support both the creation and comprehension of non-textual representations through accessible and collaborative interaction.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used ChatGPT (OpenAI) to support language revision, data organization, and the presentation of selected tables, figures, and qualitative findings. After using the tool, the authors reviewed and edited all outputs as needed and take full responsibility for the content of the publication.

Statements and Declarations

Ethics approval and consent to participate. Participation in the study was voluntary and conditional upon agreement with the Informed Consent Form (ICF). The study was registered on Plataforma Brasil under number 76931024.9.0000.5243 and approved by the Research Ethics Committees (RECs) of Fluminense Federal University (UFF) and the Benjamin Constant Institute (IBC). Participants could choose the authorized form of audio and/or video recording, and confidentiality was preserved through pseudonyms P1–P10.

Funding Declaration

This research received no specific funding from public, private, or not-for-profit funding agencies.

References

[1] Alam MZI, Islam S, Hoque E (2023) Seechart: Enabling accessible visualizations through interactive natural language interface for people with visual impairments. In: Proceedings of the 28th International Conference on Intelligent User Interfaces. Association for Computing Machinery, New York, NY, USA, IUI '23, p 46–64, https://doi.org/10.1145/3581641.3584099, URL https://doi.org/10.1145/3581641.3584099

[2] Alcaraz-Martínez R, Ribera M, Adeva-Fillol A, et al (2025) Enhancing statistical chart accessibility for people with low vision: Insights from a user test. Universal Access in the Information Society 24(4):2981–3019. https://doi.org/10.1007/s10209-024-01111-4, URL https://doi.org/10.1007/s10209-024-01111-4

[3] Amblès A, Leclet-Groux D, Potelle A (2024) Co-design of an Interactive 3D Objects Visualization Application: A User-Centered Approach for the Visually Impaired. International Journal On Advances in Software volume 17(numbers 3 and 4):226 to 238. URL https://hal.science/hal-04910988

[4] Baker CM, Milne LR, Drapeau R, et al (2016) Tactile graphics with a voice. ACM Trans Access Comput 8(1). https://doi.org/10.1145/2854005, URL https://doi.org/10.1145/2854005

[5] Bateman A, Zhao OK, Bajcsy AV, et al (2018) A user-centered design and analysis of an electrostatic haptic touchscreen system for students with visual impairments. International Journal of Human-Computer Studies 109:102–111. https://doi.org/10.1016/j.ijhcs.2017.09.004, URL https://doi.org/10.1016/j.ijhcs.2017.09.004

[6] Booch G, Rumbaugh J, Jacobson I (2005) The Unified Modeling Language User Guide. Addison-Wesley object technology series, Addison-Wesley, URL https://books.google.com.br/books?id=BqFQAAAAMAAJ

[7] Bornschein J, Bornschein D, Weber G (2018) Comparing computer-based drawing methods for blind people with real-time tactile feedback. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '18, p 1–13, https://doi.org/10.1145/3173574.3173689, URL https://doi.org/10.1145/3173574.3173689

[8] Braun V, Clarke V (2021) Thematic Analysis: A Practical Guide. SAGE Publications, URL https://books.google.com.br/books?id=mToqEAAAQBAJ

[9] Brooke J (1996) Sus: A "quick and dirty" usability scale. In: Usability Evaluation in

Industry. Taylor & Francis, p 189–194

[10] Brown T (2008) Design thinking. Harvard business review 86(6):84–92

[11] Butler M, Holloway LM, Reinders S, et al (2021) Technology developments in touch-based accessible graphics: A systematic review of research 2010-2020. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '21, https://doi.org/10.1145/3411764.3445207, URL https://doi.org/10.1145/3411764.3445207

[12] Carruthers S, Thomas A, Kaufman-Willis L, et al (2023) Growing an accessible and inclusive systems design course with plantuml. In: Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1. Association for Computing Machinery, New York, NY, USA, SIGCSE 2023, p 249–255, https://doi.org/10.1145/3545945.3569786, URL https://doi.org/10.1145/3545945.3569786

[13] Chase EDZ, Siu AF, Boadi-Agyemang A, et al (2020) Pantoguide: A haptic and audio guidance system to support tactile graphics exploration. In: Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility. Association for Computing Machinery, New York, NY, USA, ASSETS '20, https://doi.org/10.1145/3373625.3418023, URL https://doi.org/10.1145/3373625.3418023

[14] Cross C, Cetinkaya D, Dogan H (2020) Transforming diagrams' semantics to text for visually impaired. In: Design, User Experience, and Usability. Interaction Design: 9th International Conference, DUXU 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I. Springer-Verlag, Berlin, Heidelberg, p 339–350, https://doi.org/10.1007/978-3-030-49713-2_24, URL https://doi.org/10.1007/978-3-030-49713-2_24

[15] Dam RF, Teo YS (2024) Scamper: How to use the best ideation methods. https://www.interaction-design.org/literature/article/learn-how-to-use-the-best-ideation-methods-scamper, accessed: 2025-11-11

[16] Davis JU, Wu TY, Shi B, et al (2020) Tangiblecircuits: An interactive 3d printed circuit education tool for people with visual impairments. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '20, p 1–13, https://doi.org/10.1145/3313831.3376513, URL https://doi.org/10.1145/3313831.3376513

[17] Edman P (1992) Tactile Graphics. American Foundation for the Blind, URL https://books.google.com.br/books?id=C7vq4-Th71AC

[18] Engel C, Weber G (2017) Improve the accessibility of tactile charts. In: Bernhaupt R, Dalvi G, Joshi A, et al (eds) Human-Computer Interaction - INTERACT 2017. Springer International Publishing, Cham, pp 187–195, https://doi.org/10.1007/978-3-319-67744-6_12

[19] Fan D, Tomassetti O, Mouallem A, et al (2025) Promoting comprehension and engagement in introductory data and statistics for blind and low-vision students: A co-design study. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '25, https://doi.org/10.1145/3706598.3713333, URL https://doi.org/10.1145/3706598.3713333

[20] de Faria Borges LCL, Bertholdo APO, Filgueiras LVL, et al (2025) Motivation of persons with cerebral palsy through participatory design: where there is a will, there is a way. Universal Access in the Information Society 24(2):1221–1237. https://doi.org/10.1007/s10209-024-01131-0, URL https://doi.org/10.1007/s10209-024-01131-0

[21] Fernando S, Ohene-Djan J (2022) A computer aided drawing system evaluation with

early and late blind users. Assistive Technology 34(4):392–401. https://doi.org/10.1080/10400435.2020.1829745, URL https://doi.org/10.1080/10400435.2020.1829745, pMID: 33544663

[22] Goldman S, Zielezinski MB (2021) Design Thinking for Every Classroom: A Practical Guide for Educators, 1st edn. Routledge, https://doi.org/10.4324/9780429273421, URL https://doi.org/10.4324/9780429273421

[23] Gupta R (2022) Perceiving sequences and layouts through touch. In: Seifi H, Kappers AML, Schneider O, et al (eds) Haptics: Science, Technology, Applications. Springer International Publishing, Cham, pp 283–291, https://doi.org/10.1007/978-3-031-06249-0_32, URL https://doi.org/10.1007/978-3-031-06249-0_32

[24] Kim NW, Joyner SC, Riegelhuth A, et al (2021) Accessible visualization: Design space, opportunities, and challenges. Computer Graphics Forum 40(3):173–188. https://doi.org/10.1111/cgf.14298, URL https://doi.org/10.1111/cgf.14298

[25] Koontz A, Duvall J, Johnson R, et al (2022) Nothing about us without us: engaging at users in at research. Assistive Technology 34(5):499–500. https://doi.org/10.1080/10400435.2022.2117524, URL https://doi.org/10.1080/10400435.2022.2117524, pMID: 36170692

[26] Larkin JH, Simon HA (1987) Why a diagram is (sometimes) worth ten thousand words. Cognitive Science 11(1):65–100. https://doi.org/10.1111/j.1551-6708.1987.tb00863.x, URL https://doi.org/10.1111/j.1551-6708.1987.tb00863.x

[27] Luque L, Brandão L, Tori R, et al (2015) On the inclusion of blind people in uml e-learning activities. Revista Brasileira de Informática na Educação 23:18. https://doi.org/10.5753/rbie.2015.23.02.18

[28] Maćkowski M, Brzoza P, Kawulok M, et al (2023) Multimodal presentation of interactive audio-tactile graphics supporting the perception of visual information by blind people. ACM Trans Multimedia Comput Commun Appl 19(5s). https://doi.org/10.1145/3586076, URL https://doi.org/10.1145/3586076

[29] Mallin SSV, de Carvalho HG (2015) Assistive technology and user-centered design: Emotion as element for innovation. Procedia Manufacturing 3:5570–5578. https://doi.org/10.1016/j.promfg.2015.07.738, URL https://doi.org/10.1016/j.promfg.2015.07.738, 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015

[30] Melfi G, Müller K, Schwarz T, et al (2020) Understanding what you feel: A mobile audio-tactile system for graphics used at schools with students with visual impairment. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '20, p 1–12, https://doi.org/10.1145/3313831.3376508, URL https://doi.org/10.1145/3313831.3376508

[31] Metatla O, Oldfield A, Ahmed T, et al (2019) Voice user interfaces in schools: Co-designing for inclusion with visually-impaired and sighted pupils. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '19, p 1–15, https://doi.org/10.1145/3290605.3300608, URL https://doi.org/10.1145/3290605.3300608

[32] Moured O, Baumgarten-Egemole M, Müller K, et al (2024) Chart4blind: An intelligent interface for chart accessibility conversion. In: Proceedings of the 29th International Conference on Intelligent User Interfaces. Association for Computing Machinery, New York, NY, USA, IUI '24, p 504–514, https://doi.org/10.1145/3640543.3645175, URL https://doi.org/10.1145/3640543.3645175

[33] Mukhiddinov M, Kim SY (2021) A systematic literature review on the automatic creation of tactile graphics for the blind and

visually impaired. Processes 9(10). https://doi.org/10.3390/pr9101726, URL https://www.mdpi.com/2227-9717/9/10/1726

[34] Object Management Group (2017) OMG Unified Modeling Language (OMG UML), Version 2.5.1. Specification formal/2017-12-05, Object Management Group, URL https://www.omg.org/spec/UML/, oMG Document Number: formal/2017-12-05

[35] Ortiz-Escobar LM, Chavarria MA, Schönenberger K, et al (2023) Assessing the implementation of user-centred design standards on assistive technology for persons with visual impairments: a systematic review. Frontiers in Rehabilitation Sciences Volume 4 - 2023. https://doi.org/10.3389/fresc.2023.1238158, URL https://www.frontiersin.org/journals/rehabilitation-sciences/articles/10.3389/fresc.2023.1238158

[36] Pandey M, Subramonyam H, Sasia B, et al (2020) Explore, create, annotate: Designing digital drawing tools with visually impaired people. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '20, p 1–12, https://doi.org/10.1145/3313831.3376349, URL https://doi.org/10.1145/3313831.3376349

[37] Prescher D, Bornschein J, Weber G (2017) Consistency of a tactile pattern set. ACM Trans Access Comput 10(2). https://doi.org/10.1145/3053723, URL https://doi.org/10.1145/3053723

[38] Ramôa G, Müller K (2025) Leveraging dynamic audio-tactile uis to assist visually impaired users in exploring line charts through tactile graphic readers. In: Proceedings of the 27th International ACM SIGACCESS Conference on Computers and Accessibility. Association for Computing Machinery, New York, NY, USA, ASSETS '25, https://doi.org/10.1145/3663547.3746398, URL https://doi.org/10.1145/3663547.3746398

[39] Ramôa G, Moured O, Schwarz T, et al (2023) Enabling people with blindness to distinguish lines of mathematical charts with audio-tactile graphic readers. In: Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments. Association for Computing Machinery, New York, NY, USA, PETRA '23, p 384–391, https://doi.org/10.1145/3594806.3594818, URL https://doi.org/10.1145/3594806.3594818

[40] Ramôa G, Schmidt V, König P (2022) Developing dynamic audio navigation uis to pinpoint elements in tactile graphics. Multimodal Technologies and Interaction 6(12). https://doi.org/10.3390/mti6120113, URL https://www.mdpi.com/2414-4088/6/12/113

[41] Riley S (2024) Mindful Design: A Survival Guide for Responsible Product Designers (Design Thinking), 2nd edn. Apress, https://doi.org/10.1007/979-8-8688-0143-3, URL https://link.springer.com/book/10.1007/979-8-8688-0143-3

[42] Rumbaugh J, Booch G, Jacobson I (2010) The Unified Modeling Language Reference Manual. Addison-Wesley object technology series, Addison-Wesley, URL https://books.google.mu/books?id=05EVnwEACAAJ

[43] Santos A, Silveira Z (2020) At-d8sign: methodology to support development of assistive devices focused on user-centered design and 3d technologies. Journal of the Brazilian Society of Mechanical Sciences and Engineering 42. https://doi.org/10.1007/s40430-020-02347-w

[44] Santos A, Silveira Z (2021) Design for assistive technology oriented to design methodology: a systematic review on user-centered design and 3d printing approaches. Journal of the Brazilian Society of Mechanical Sciences and Engineering 43. https://doi.org/10.1007/s40430-021-03184-1

[45] Sauro J (2011) A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. Measuring Usability LLC, URL https://books.google.com.br/books?id=BL0kKQEACAAJ

[46] Serin F, Romeo K (2022) Drawing and understanding diagrams: An accessible approach dedicated to blind people. In: Universal Access in Human-Computer Interaction. User and Context Diversity: 16th International Conference, UAHCI 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26 – July 1, 2022, Proceedings, Part II. Springer-Verlag, Berlin, Heidelberg, p 94–109, https://doi.org/10.1007/978-3-031-05039-8_7, URL https://doi.org/10.1007/978-3-031-05039-8_7

[47] Shi L, Lawson H, Zhang Z, et al (2019) Designing interactive 3d printed models with teachers of the visually impaired. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '19, p 1–14, https://doi.org/10.1145/3290605.3300427, URL https://doi.org/10.1145/3290605.3300427

[48] Thompson JR, Martinez JJ, Sarikaya A, et al (2023) Chart reader: Accessible visualization experiences designed with screen reader users. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '23, https://doi.org/10.1145/3544548.3581186, URL https://doi.org/10.1145/3544548.3581186

[49] Vargas PN, Trevisan DG (2024) Revisiting visual accessibility with non-textual content: challenges and solutions for human-computer interaction. In: Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, IHC '24, https://doi.org/10.1145/3702038.3702105, URL https://doi.org/10.1145/3702038.3702105

[50] Vargas PN, Trevisan DG (2025) Seeing what we can't see in the non-textual representation from the educational context. In: Human-Computer Interaction – INTERACT 2025: 20th IFIP TC 13 International Conference, Belo Horizonte, Brazil, September 8–12, 2025, Proceedings, Part III. Springer-Verlag, Berlin, Heidelberg, p 333–355, https://doi.org/10.1007/978-3-032-05005-2_18, URL https://doi.org/10.1007/978-3-032-05005-2_18

[51] Verde A, Alves L, Fraga L, et al (2023) Py2uml: a tool for visually impaired students to build uml diagrams from python coding. In: Proceedings of the XXXVII Brazilian Symposium on Software Engineering. Association for Computing Machinery, New York, NY, USA, SBES '23, p 491–496, https://doi.org/10.1145/3613372.3613418, URL https://doi.org/10.1145/3613372.3613418

[52] Watanabe T, Sato K (2019) Suitable size of 3d printing architecture models for tactile exploration. Journal on Technology & Persons with Disabilities 7:45–53

[53] Wildhaber F, Salloum N, Gygli M, et al (2020) Self-directed creation and editing of uml class diagrams on mobile devices for visually impaired people. In: 2020 IEEE Tenth International Model-Driven Requirements Engineering (MoDRE). IEEE, pp 49–57, https://doi.org/10.1109/MoDRE51215.2020.00012

[54] Wilson J, Rosenberg D (1988) Chapter 39 - rapid prototyping for user interface design. In: HELANDER M (ed) Handbook of Human-Computer Interaction. North-Holland, Amsterdam, p 859–875, https://doi.org/10.1016/B978-0-444-70536-5.50044-0, URL https://doi.org/10.1016/B978-0-444-70536-5.50044-0

[55] Wimer BL, Kanchi R, Frierson K, et al (2026) Nonvisual support for understanding and reasoning about data structures. In: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '26, https://doi.org/10.1145/3772318.3791656, URL https://doi.org/10.1145/3772318.3791656

[56] Wong K, Sun D (2006) On evaluating the layout of uml diagrams for program comprehension. Software Quality Journal 14(3):233–259. https://doi.org/10.1007/s11219-006-9218-2, URL https://doi.org/10.1007/s11219-006-9

218-2

[57] Woodring I, Owen C, Islam S (2024) A method for presenting uml class diagrams with audio for blind and visually impaired students. In: Proceedings of the 17th International Conference on PErvasive Technologies Related to Assistive Environments. Association for Computing Machinery, New York, NY, USA, PETRA '24, p 15–20, https://doi.org/10.1145/3652037.3652056, URL https://doi.org/10.1145/3652037.3652056

[58] Yang Y, Marriott K, Butler M, et al (2020) Tactile presentation of network data: Text, matrix or diagram? In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, CHI '20, p 1–12, https://doi.org/10.1145/3313831.3376367, URL https://doi.org/10.1145/3313831.3376367

[59] Zhang Q, Fiorella L (2021) Learning by drawing: When is it worth the time and effort? Contemporary Educational Psychology 66:101990. https://doi.org/10.1016/j.cedpsych.2021.101990, URL https://doi.org/10.1016/j.cedpsych.2021.101990

[60] Zhang Z, Thompson JR, Shah A, et al (2024) Charta11y: Designing accessible touch experiences of visualizations with blind smartphone users. In: Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility. Association for Computing Machinery, New York, NY, USA, ASSETS '24, https://doi.org/10.1145/3663548.3675611, URL https://doi.org/10.1145/3663548.3675611

[61] Zhao Y, Nacenta MA, Sukhai MA, et al (2024) Tada: Making node-link diagrams accessible to blind and low-vision people. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM, CHI '24, p 1–20, https://doi.org/10.1145/3613904.3642222, URL http://dx.doi.org/10.1145/3613904.3642222