



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México

**Centro Nacional de Investigación
y Desarrollo Tecnológico**

Tesis de Maestría

**Representación visual de datos multimodales de
comportamiento para generar analítica visual del
proceso de evaluación de la experiencia de usuario**

presentada por

Ing. Tlaloc Humberto Vergara Morales

como requisito para la obtención del grado de
Maestra en Ciencias de la Computación

Director de tesis

Dr. Juan Gabriel González Serna

Codirector de tesis

Dr. Nimrod González Franco

Cuernavaca, Morelos, México. Diciembre de 2025.



Centro Nacional de Investigación y Desarrollo tecnológico
Departamento de Ciencias Computacionales

Cuernavaca, Mor., 18/noviembre/2025

OFICIO No. DCC/337/2025

Asunto: Aceptación de documento de tesis
CENIDET-AC-004-M14-OFICIO

CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial de **TLALOC HUMBERTO VERGARA MORALES** con número de control **M23CE088**, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado **"Representación visual de datos multimodales de comportamiento para generar analítica visual del proceso de evaluación de la experiencia de usuario"** y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

ATENTAMENTE

Excelencia en Educación Tecnológica®

"Conocimiento y Tecnología al Servicio de México"

Dr. Juan Gabriel González Serna
Director de tesis

Dr. Nimrod González Franco
Codirector de tesis

Dr. Noé Alejandro Castro Sánchez
Revisor 1

Dra. Olivia Graciela Fragozo Díaz
Revisor 2

C.c.p. Depto. Servicios Escolares.
Expediente / Estudiante



2025
Año de
La Mujer
Indígena

Interior Internado Palmira S/N, Col. Palmira,

C. P. 62490, Cuernavaca, Morelos Tel. 01 (777) 3627770, ext. 3201,

e-mail: dcc_cenidet@tecnm.mx tecnm.mx | cenidet.tecnm.mx

cenidet
Centro Nacional de Investigación
y Desarrollo Tecnológico





Centro Nacional de Investigación y Desarrollo tecnológico
Subdirección Académica

Cuernavaca Mor, 20/noviembre/2025

Oficio No. SAC/269/2025

Asunto: Autorización de impresión de tesis

TLALOC HUMBERTO VERGARA MORALES
CANDIDATO AL GRADO DE MAESTRO
EN CIENCIAS DE LA COMPUTACIÓN
P R E S E N T E

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado **"Representación visual de datos multimodales de comportamiento para generar analítica visual del proceso de evaluación de la experiencia de usuario"**, ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

ATENTAMENTE

Excelencia en Educación Tecnológica®

"Conocimiento y Tecnología al Servicio de México"



CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO

c.c.p. Departamento de Ciencias Computacionales
Departamento de Servicios Escolares

CMAZ/lmz



2025
Año de
La Mujer
Indígena

Interior Internado Palmira S/N, Col. Palmira,
C. P. 62490, Cuernavaca, Morelos Tel. 01 (777) 3627770, ext. 4104,
e-mail: acad_cenidet@tecnm.mx tecnm.mx | cenidet.tecnm.mx

cenidet
Centro Nacional de Investigación
y Desarrollo Tecnológico



Dedicatoria

A mi madre, Olga, por apoyarme tanto económicamente como sentimentalmente. La persona que me crió y forjó realmente como soy ahora. A mi hermano, Yair, por estar siempre ahí cuando lo necesitaba. En parte, este también es su logro de ambos, porque no se habría logrado sin su sacrificio, paciencia, ayuda y amor por mí.

A todos mis amigos que estuvieron ahí para mí, incluso si solo fue para charlar, tanto para mi círculo más cercano: Laura, Mónica, Olivetto, Luis, Jacinto.

A mi círculo de amigos más lejanos, pero que aun así sentí su apoyo: Alexa, Hazael y Leonardo.

A mi asesor que ha sido tan comprensible y paciente conmigo, Dr. Gabriel González Serna.

A mis profesores de la carrera que me dieron su conocimiento y ayuda durante mi estancia en Cenidet: Dr. Nimrod, Dr. Max, Dr. Dante, Dr. Noé.

A la generación 2023-2025 del departamento de Ciencias de la Computación, tanto a mis compañeros de la misma línea como a otros que conocí durante mi estancia y también al personal de oficinas que me ayudó a resolver todos mis problemas con paciencia e interés.

A mis profesores de mis anteriores instituciones, Upemor, Cetis No.99, Ignacio Manuel Altamirano, que han dejado una gota de su conocimiento conmigo y que hoy utilizo en parte para lograr este proyecto.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por su apoyo económico; sin este, no hubiera sido posible el desarrollo de esta tesis.

Agradezco al Tecnológico Nacional de México, campus Centro Nacional de Investigación y Desarrollo Tecnológico (TecNM/CENIDET), y a todo su personal por darme la oportunidad de realizar mis estudios de maestría en Ciencias de la Computación.

Agradezco a mi director de tesis, el Dr. Juan Gabriel González Serna, por ser mi asesor, guía y apoyo durante todo el periodo de mi estancia en el TecNM/CENIDET. A mi codirector, el Dr. Nimrod González Franco, y a mis revisores, el Dr. Noé Alejandro Castro Sánchez y la Dra. Olivia Graciela Fragosó Díaz, quienes objetivamente aportaron en este trabajo, me compartieron de su valioso tiempo para observaciones, consejos y sugerencias

Agradezco a mi madre y a mi hermano por permanecer a mi lado y brindarme su apoyo y amor incondicional a lo largo de todo este nuevo proceso en mi vida.

Desde el fondo de mi corazón, muchas gracias a todos los que me ayudaron a lograr esta tesis.

Resumen

En el proceso de desarrollo de software, la fase de pruebas es esencial para validar la funcionalidad de un prototipo o de una versión temprana del producto. En esta etapa se busca no solo identificar fallas técnicas, sino también obtener retroalimentación directa de los usuarios, ya que su experiencia de interacción aporta información valiosa para mejorar el diseño. Existen métodos simples para recabar opiniones, como cuestionarios o encuestas, pero los enfoques más ricos suelen implicar la observación y el registro detallado de la interacción mediante videos, audio y otros datos. El problema de estos métodos es que generan grandes volúmenes de información que deben revisarse varias veces, lo que convierte el análisis en una tarea lenta y exigente para los evaluadores. Aunque se han propuesto diferentes herramientas para apoyar este proceso, muchas presentan limitaciones, como la sobrecarga de datos sin una organización clara o la falta de precisión en la representación de eventos específicos, lo que dificulta la interpretación y disminuye su utilidad práctica.

Frente a este panorama, este proyecto de investigación se enfocó en diseñar y desarrollar un reproductor de video que facilitara la evaluación de la experiencia de usuario mediante la integración de datos multimodales en una única interfaz. El sistema permite superponer información directamente sobre los videos de interacción, mostrando con claridad aspectos como el seguimiento ocular, las expresiones faciales y la verbalización del usuario. Además, incorpora recursos de análisis visual que ayudan a identificar trayectorias de la mirada, zonas de interés en la interfaz y patrones de atención, representados tanto en porcentajes como en mapas de calor. Información visual sobre comentarios positivos, negativos y neutrales. Presentación de información sobre el análisis facial, como la detección de emociones, unidades de acción y estados cognitivos. La propuesta busca simplificar la labor de los evaluadores al organizar los datos de manera intuitiva y, al mismo tiempo, evitar la saturación de información que dificulte el análisis.

Durante la etapa de validación se realizaron 13 pruebas completas, en las que se comprobó la calidad de los datos obtenidos y la efectividad del reproductor en la tarea de revisión. Los resultados mostraron que el sistema reduce, en promedio, en un 69.08% el tiempo necesario para extraer observaciones relevantes, lo que representa un avance significativo en eficiencia. Este ahorro de tiempo no solo facilita el trabajo de los evaluadores, sino que también permite analizar una mayor cantidad de material sin sacrificar la profundidad de las observaciones. Asimismo, se comprobó que la interfaz del reproductor resulta práctica y no interfiere con la visualización de los datos, lo que contribuye a una revisión más rápida de todos los videos de la experiencia de usuario.

En conjunto, los hallazgos respaldan el valor práctico y técnico de esta herramienta, que se consolida como un apoyo sólido para la evaluación de prototipos digitales. Su capacidad para integrar de manera organizada diferentes datos visuales y facilitar la interpretación de información compleja lo convierte en una herramienta útil para identificar y corregir errores de diseño en software. Con ello, no solo se optimiza la revisión de la experiencia de usuario, sino que también se promueve un ciclo de desarrollo más ágil y eficiente, con beneficios tanto para los equipos de diseño como para los usuarios finales.

Palabras clave: Analítica visual, experiencia de usuario, representación visual, interfaz de usuario, interacción humano-computadora.

Abstract

In the software development process, the testing phase is essential for validating the functionality of a prototype or early version of the product. This stage aims not only to identify technical failures but also to obtain direct user feedback, as their interaction experience provides valuable information for improving the design. There are simple methods for gathering opinions, such as questionnaires or surveys, but richer approaches usually involve observing and recording interactions in detail through video, audio, and other data. The problem with these methods is that they generate large volumes of information that must be reviewed multiple times, making the analysis a slow and demanding task for evaluators. Although different tools have been proposed to support this process, many of them have limitations, such as data overload without clear organization or insufficient precision in representing specific events, which complicate interpretation and reduce their practical usefulness.

Given this situation, this research project focused on designing and developing a video player that facilitates user experience evaluation by integrating multimodal data into a single interface. The system allows overlaying information directly onto interaction videos, clearly displaying aspects such as eye tracking, facial expressions, and the user's verbalizations. In addition, it incorporates visual analysis resources that help identify gaze trajectories, interface areas of interest, and attention patterns, represented both as percentages and heatmaps; visual information about positive, negative, and neutral comments; and facial analysis data, including emotion detection, action units, and cognitive states. The proposal seeks to simplify the evaluators' work by organizing the data intuitively while avoiding information overload that would hinder analysis.

During the validation stage, 13 complete tests were conducted to verify the quality of the data obtained and the effectiveness of the video player in the review task. The results showed that the system reduced the time required to extract relevant observations by an average of 69.08%, resulting in a significant improvement in efficiency. This time savings not only makes the evaluators' work easier, but also allows for the analysis of a greater amount of material without sacrificing the depth of observations. Likewise, it was confirmed that the player's interface is practical and does not interfere with data visualization, contributing to faster review across all user experience videos.

Overall, the findings support the practical and technical value of this tool, which stands as a solid aid for evaluating digital prototypes. Its ability to integrate diverse visual data in an organized manner and to facilitate the interpretation of complex information makes it a valuable resource for identifying and correcting software design errors. In doing so, it not only optimizes user experience review but also promotes a more agile, efficient development cycle, benefiting both design teams and end users.

Keywords: Visual analytics, user experience, visual representation, user interface, human-computer interaction.

Índice

Dedicatoria.....	III
Agradecimientos	IV
Capítulo 1: Introducción	1
1.1 Antecedentes	3
1.2 Planteamiento del problema.....	4
1.3 Objetivos.....	6
1.4 Justificación	6
1.5 Alcances y limitaciones	7
Capítulo 2: Marco teórico	8
2.1 Experiencia de usuario.....	9
2.2 Datos multimodales	9
2.3 Analítica visual (Visual Analytic)	10
2.4 Datos de comportamiento	11
Capítulo 3: Trabajos relacionados	14
3.1 Trabajos relacionados con seguimiento ocular	16
3.2 Trabajos relacionados con expresión facial	17
3.3 Trabajos relacionados con el protocolo Think-Aloud o minería de opinión	18
Capítulo 4: Propuesta de solución	25
4.1 Etapa 1: Análisis de trabajos relacionados	26
4.2 Etapa 2: Diseñar reproductor que implemente analítica visual	26
4.3 Etapa 3: Implementar cambios en el reproductor de UXLab	27
4.4 Etapa 4: Plan de pruebas	27
4.5 Etapa 5: Análisis de resultados	27
4.6 Etapa 6: Correcciones y adecuaciones.....	27
4.7 Etapa 7: Entrega de resultados.....	27
Capítulo 5: Recolección de datos.....	28
5.1 Naturaleza de los datos del repositorio	29
5.2 Grabaciones de las pruebas.....	29
5.3 Datos esperados	33
Capítulo 6: Experimentación y resultados	35
6.1 Métodos de evaluación	36
6.2 Pruebas funcionales	37
6.3 Comparativa de tiempos total vs tiempo segmento	52

6.4	Calidad del sistema y carga mental de usuarios	54
6.5	Formulario de evaluación heurística	56
Capítulo 7: Conclusiones		70
7.1	Trabajos futuros	72
Referencias.....		73
ANEXO		76
Anexo A		76
Anexo B		76
Anexo C		76
Anexo D.....		76

Índice de figuras

Figura 1. Arquitectura de UXLab. Fuente: UXLab Arquitectura (González, 2019). ...	4
Figura 2. Representación del problema	6
Figura 3. Esquema del funcionamiento del seguimiento ocular con sensor infrarrojo (Fouilloux, 2018).	11
Figura 4. Análisis multimodal de expresión facial y postura (Toribio, 2023)	12
Figura 5. Representación gráfica de una sesión de Think Aloud Protocol.	13
Figura 6. Representación gráfica de la minería de opiniones. (Villena, 2015).	13
Figura 7. Metodología de solución.	26
Figura 8. Parte del modelo de procedimiento para la obtención de datos.	30
Figura 9. Distribución de los participantes por género.	34
Figura 10. Distribución de los participantes según su rango de edad.	34
Figura 11. Interfaz del reproductor de UXLab original.	37
Figura 12. Interfaz del reproductor UXLab actual que se divide en 5 segmentos importantes: (1) Los botones de menú y de activación de la información visual. (2) Recuadro donde se reproduce el video de la webcam, que puede moverse y cambiar de tamaño. (3) La reproducción del escritorio. (4) La segmentación del video según los comentarios del usuario. (5) Segmentación del video según el cambio de contexto.	38
Figura 13. Representación de la situación en la que se observan los puntos de visión.	40
Figura 14. Prueba en la que se utiliza el mapa de calor para ver en qué parte se presta atención.	41
Figura 15. Prueba en la que se utilizan mapas de rutas para identificar la ruta de visión del usuario.	42
Figura 16. Prueba en la que se utilizan las áreas de atención.	43
Figura 17. Prueba donde se visualizan las emociones del usuario durante la prueba.	44
Figura 18. Prueba en la que se utilizan las unidades de acción.	45
Figura 19. Prueba en la que se utiliza la información de análisis cognitivo.	46
Figura 20. Segmentación del video según los comentarios y la transcripción.	47
Figura 21. Representación visual de cómo Auditok segmenta el audio.	48
Figura 22. Prueba en la que se utilizan varias representaciones visuales a la vez sin ningún problema.	51
Figura 23. Gráfica de cantidad de comentarios de los usuarios.	56
Figura 24. Primera vista del reproductor UXLab.	58
Figura 25. Botones tanto de las representaciones visuales como de las funcionalidades.	60
Figura 26. Comparación del reproductor actual de UXLab con un reproductor de video convencional.	62
Figura 27. Mensaje de advertencia.	64

Índice de Tablas

Tabla 1. Proporción de artículos seleccionados por año.....	15
Tabla 2. Comparación de trabajos sobre analítica visual.....	23
Tabla 3. Recursos que se usaron en las pruebas.	32
Tabla 4. Comparación de tiempos de revisión.....	52
Tabla 5. Tiempos de la segmentación de las primeras pruebas (Sin estándar de calidad).	53
Tabla 6. Evaluación del formulario PSSUQ.....	54
Tabla 7. Resultados del formulario NASA-TLX de la tarea 1.	54
Tabla 8. Resultados del formulario NASA-TLX de la tarea 2.	55
Tabla 9. Resultados del conteo de comentarios positivos, negativos y neutrales.....	55
Tabla 10. Visibilidad del estado del sistema.....	57
Tabla 11. Relación entre el sistema y el mundo real.	59
Tabla 12. Control de usuario y libertad.	60
Tabla 13. Consistencia y estándares.	61
Tabla 14. Prevención de errores.	63
Tabla 15. Reconocimiento en lugar de recordar.	64
Tabla 16. Flexibilidad y diseño minimalista.....	66
Tabla 17. Diseño estético y minimalista.....	66
Tabla 18. Ayudar a los usuarios a reconocer, diagnosticar y recuperar errores.	67
Tabla 19. Ayuda y documentación.	68
Tabla 20. Tabla de puntuación de la evaluación heurística.	69
Tabla 21. Puntuación heurística.....	69

Capítulo 1: Introducción

Durante el desarrollo de software, normalmente existe una fase de pruebas en la que el producto se evalúa mediante un prototipo o una versión Alpha/Beta. Estas versiones suelen ser revisadas primero por los propios desarrolladores para comprobar su funcionamiento. Sin embargo, en ciertos casos resulta necesario obtener retroalimentación directa del público objetivo, lo cual puede lograrse de diversas maneras. Algunos métodos son sencillos y comunes, tales como aplicar cuestionarios tras el uso del programa o utilizar encuestas de satisfacción que permiten calificarlo mediante métricas simples, como estrellas o emojis. Otros enfoques, en cambio, son más sofisticados, observar la interacción del usuario con el software en un entorno controlado y tomar notas, o incluso registrar esa interacción mediante cámaras y micrófonos para un análisis más detallado. Asimismo, es importante señalar que la información más detallada y relevante suele obtenerse mediante métodos sofisticados. En el caso de las grabaciones de la interacción, no solo es posible analizar cómo el usuario utiliza el software y cuáles son sus observaciones, sino también extraer datos pasivos que pueden procesarse posteriormente. Entre estos datos se incluyen sus expresiones faciales, la ubicación de la mirada en la interfaz mediante cámaras especializadas, la transcripción de sus comentarios y la identificación de si estos reflejan una valoración positiva o negativa, entre otros aspectos. No obstante, este tipo de análisis presenta una desventaja: la revisión de todo el material requiere muchas horas para ser realmente exhaustiva. Los revisores no solo deben ver cada video, cuyo tiempo puede variar entre un mínimo de cinco minutos y un máximo indefinido, aunque lo habitual es que dure alrededor de treinta minutos, sino también registrar todas las observaciones de los usuarios y añadir las suyas propias al detectar otras áreas de oportunidad a partir de los datos disponibles. Esto implica que cada video deba revisarse varias veces, retrocediendo y avanzando según sea necesario para analizar con detalle la información requerida. Además, este esfuerzo corresponde únicamente a un solo video; en la práctica, suelen recopilarse varios para obtener la perspectiva de diferentes usuarios. Esto genera la necesidad de desarrollar una herramienta que facilite la tarea y reduzca el tiempo necesario para revisar los videos.

En la literatura consultada se identifican diversos trabajos que proponen distintas iteraciones de reproductores diseñados para presentar al evaluador los datos necesarios. Sin embargo, esta investigación identificó dos problemas principales en dichos desarrollos. La primera es la sobrecarga de información, ya que tiende a presentar todos los datos de manera desorganizada, sin un control de calidad que facilite su comprensión. Un ejemplo claro es el trabajo de da Silva Franco et al. (2019) con su sistema Uxmood, en el que la información se presenta de manera caótica y con filtros limitados únicamente a las emociones. Esta falta de organización y categorización de los datos puede generar confusión y dificultar la labor del evaluador, cuando el objetivo principal debería ser, precisamente, simplificar y facilitar el análisis.

El segundo es la falta de precisión en la presentación de la información en pantalla, donde la información que ofrecen las herramientas visuales del reproductor resulta poco clara o poco intuitiva para el evaluador. En el trabajo de Soure et al. (2021) sobre la interfaz CoUX, si bien la propuesta resulta visualmente atractiva al incorporar marcas en la barra de reproducción para señalar la presencia de comentarios, no se especifica con exactitud cuándo ocurren dichos eventos. En su lugar, el sistema agrupa los comentarios en “capítulos” o segmentos del video, lo que reduce la precisión y puede dificultar un análisis puntual de la información.

Por tanto, esta investigación tiene como objetivo desarrollar un reproductor de video capaz de representar visualmente datos multimodales de comportamiento y, de este modo, generar una analítica visual del proceso de evaluación de la experiencia de usuario. Todo ello orientado a apoyar la toma de decisiones en el desarrollo de productos digitales, permitiendo identificar y corregir errores de diseño en etapas tempranas del proceso.

1.1 Antecedentes

Máquina de aprendizaje para el reconocimiento de emociones mediante datos fisiológicos multimodales (Morales Morante, J., 2021)

En esta tesis de maestría se desarrolló una metodología y un conjunto de algoritmos basados en el aprendizaje automático para el reconocimiento de emociones básicas. Dichos algoritmos utilizan como entrada datos fisiológicos multimodales provenientes del electroencefalograma (EEG), del electrocardiograma (ECG) y de la respuesta galvánica de la piel (GSR) para su evaluación. Estos clasificadores se incorporaron a la plataforma UXLab para apoyar la evaluación de la experiencia de usuario.

Algoritmo para analizar el movimiento ocular del usuario para identificar la correlación con estados mentales (Lagunés Ramírez, D., 2020)

En esta tesis de maestría se desarrolló un método para analizar el movimiento ocular de los usuarios durante el proceso de evaluación UX, con el fin de obtener datos relevantes y generar mapas de calor, trayectorias visuales y áreas de interés correlacionadas con el estado cognitivo durante la interacción con un producto digital. Este método se incorporó como un módulo de visualización en la plataforma de evaluación de la experiencia de usuario (UXLab).

Método para la detección de emociones mediante el análisis de expresión facial en videos de evaluación de la experiencia de usuario (Hernández Pérez, L., 2022)

En esta tesis de maestría se implementó un algoritmo de aprendizaje automático capaz de reconocer emociones a partir del análisis de las expresiones faciales durante la evaluación de la UX. El método se basó en el sistema EMFACS, propuesto por Ekman, que permite describir distintas emociones. El modelo fue entrenado con un repositorio propio y evaluado en el conjunto de datos CK+. Como resultado, se alcanzó una precisión del 84.52 % en el reconocimiento de cuatro emociones: felicidad, enojo, sorpresa y neutral.

Método para integrar y sincronizar datos EEG y multimedia para su aplicación en el proceso de evaluación de la experiencia del usuario (Fouilloux Quiroz, D., 2018)

UXLab es una plataforma de software diseñada para la captura, sincronización y visualización de datos multimodales provenientes de distintas fuentes, como el video del rostro del usuario, la captura de pantalla, señales fisiológicas (EEG, ECG y GSR), el seguimiento ocular y el audio. Esta plataforma fue desarrollada como parte de la tesis

“Método para integrar y sincronizar datos EEG y multimedia para su aplicación en el proceso de evaluación de la experiencia del usuario” (Fouilloux, 2018).

La grabación de estos datos multimodales se utiliza para analizar y evaluar la experiencia del usuario, en particular para identificar estados emocionales. Sin embargo, en las primeras versiones de UXLab, este análisis no se realizaba automáticamente, sino que lo hacía completamente un experto. Como resultado de esta tesis, se integró un módulo de análisis emocional que toma como entrada las señales fisiológicas grabadas y devuelve dos valores: uno de valencia y otro de activación. Con esta información, es posible rastrear las emociones y crear un recorrido de los distintos estados emocionales experimentados por un usuario a lo largo de una prueba UX.

En la figura 1 se muestra la integración de las funciones de UXLab y las etapas asociadas a cada una.

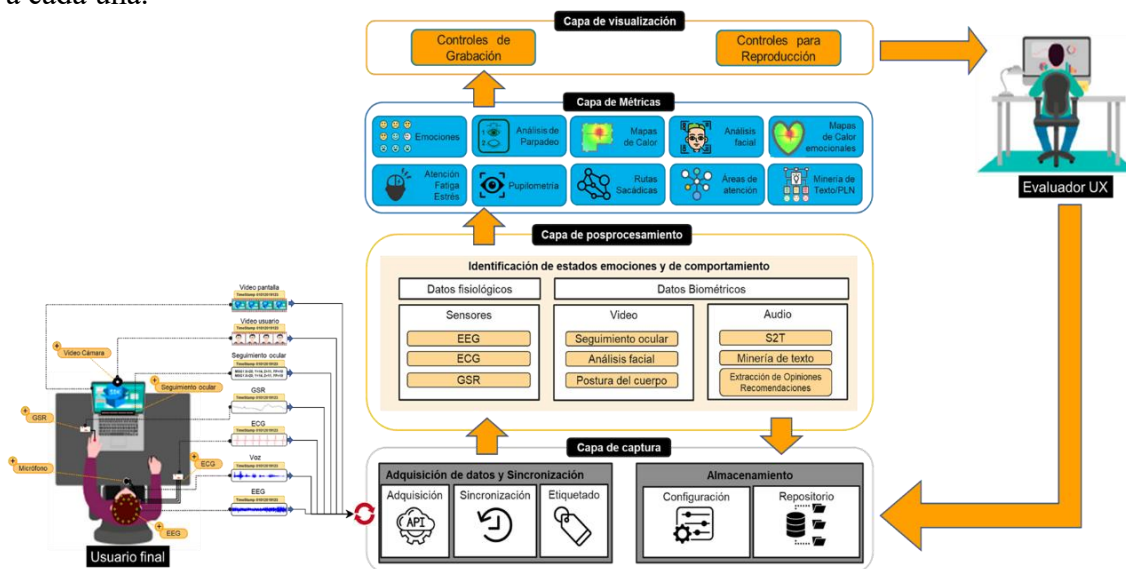


Figura 1. Arquitectura de UXLab. Fuente: UXLab Arquitectura (González, 2019).

1.2 Planteamiento del problema

La evaluación de la Experiencia de Usuario (UX) ha evolucionado al integrar múltiples datos que permiten una comprensión más profunda y holística de la interacción entre el usuario y la interfaz. Este enfoque multimodal combina datos de diversas fuentes, como el seguimiento ocular, la minería de opiniones obtenidas mediante el protocolo Think-Aloud y las grabaciones en video de la interacción. Si bien esta estrategia permite realizar una evaluación más completa, también plantea importantes desafíos computacionales que se describen a continuación.

- **Integración de datos multimodales:** La primera problemática surge en la etapa inicial de integración de datos de distintas modalidades. Cada fuente posee características propias, con formatos y escalas de medición distintos, lo que dificulta la creación de un marco unificado para su fusión. Al reunir todos estos datos, surge la duda sobre cómo integrarlos en un mismo ecosistema para representarlo de forma clara. Esto puede abordarse mediante la fusión de datos para generar información unificada, una tarea especialmente compleja, o bien

mediante una integración basada en las representaciones visuales de los datos obtenidos, que resultan más fáciles de crear e interpretar, considerando el tiempo de desarrollo disponible. La calidad de los datos aporta una dificultad adicional por considerar. Es necesario asegurar la claridad del audio, la buena resolución del video y la fiabilidad de los datos de seguimiento ocular, no solo para contar con información útil y representativa, sino también para garantizar que los algoritmos que generan las representaciones visuales funcionen correctamente.

- **Visualización coherente:** La representación visual de datos multimodales constituye un desafío adicional. Crear herramientas que muestren la información de forma ordenada, sin abrumar al usuario final, exige una atención cuidadosa al diseño visual y a su distribución en la interfaz. Una problemática común en otros trabajos es la ausencia de una propuesta clara sobre cómo distribuir los datos visuales, lo que dificulta su uso. Además, la falta de orden en las herramientas visuales puede llevar a que no se utilicen en absoluto o, peor aún, a que generen confusión al estorbar.
- **Privacidad:** La opinión, especialmente en la actualidad, es una herramienta muy poderosa que puede resultar incómoda para muchas personas cuando deja de ser un simple texto y se le asocia un rostro o un nombre, en particular cuando se trata de críticas. Por ello, mantener la información en el mayor anonimato posible resulta esencial. Así pues, resulta fundamental mantener la privacidad de los usuarios al recopilar y gestionar la información derivada de dichas pruebas.

La problemática que se presenta en este proyecto de investigación busca resolver está relacionado con la creación de analítica visual; el enfoque que combina técnicas de análisis de datos con herramientas de visualización para ayudar a las personas a entender información compleja, un campo que representa un desafío desde la perspectiva computacional, principalmente porque muchos de los datos aún no pueden transformarse en reportes completamente automatizados. Todavía se requiere intervención humana para una revisión exhaustiva, aunque este proceso puede optimizarse mediante herramientas que permitan una evaluación más rápida y, al mismo tiempo, completa. Superar estas problemáticas demanda un enfoque interdisciplinario que abarque el diseño de la experiencia de usuario, la estadística, la visualización de datos y la ética computacional. A medida que avance esta investigación, la capacidad de representar y visualizar datos multimodales de forma ordenada y clara se convertirá en un elemento clave para comprender y mejorar la revisión de los videos de la experiencia de los usuarios. Y también, para mejorar las futuras iteraciones de esta misma investigación, así como la implementación de otras y mejores analíticas visuales o incluso de herramientas que automatizan parte del proceso.

Para una interpretación visual, véase la figura 2 a continuación, que resume y grafica la problemática a resolver en este trabajo.

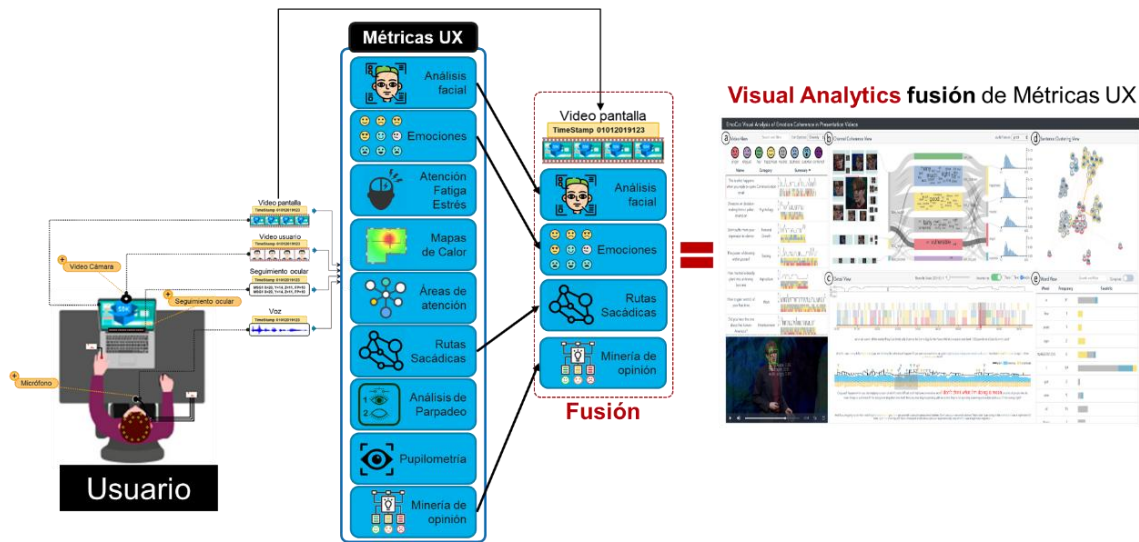


Figura 2. Representación del problema

1.3 Objetivos

1.3.1 Objetivo general

Diseñar, desarrollar y evaluar un reproductor de video que integre analítica visual con datos multimodales del proceso de evaluación de la experiencia de usuario, para apoyar la toma de decisiones, permitir la identificación y corrección de errores de diseño en el desarrollo de software.

1.3.2 Objetivos específicos

- Modificar el reproductor de UXLab para lograr una representación más clara y enfocada del video de la interacción del usuario.
- Implementar una analítica visual que integre la información de seguimiento ocular con el video de la interacción. Por ejemplo: mapas de calor.
- Implementar una analítica visual que integre la información del análisis facial con el video de la interacción. Por ejemplo: unidades de acción del rostro.
- Incorporar una analítica visual que segmente el video para localizar comentarios positivos, negativos y neutros en el reproductor.
- Evaluar un prototipo con el reproductor para validar su utilidad en la elaboración de reportes de experiencia de usuario.

1.4 Justificación

Los informes sobre el proceso de evaluación de la experiencia de usuario o de usabilidad han cobrado gran relevancia en los últimos años. Con el desarrollo continuo del software, se vuelve necesaria la participación de los usuarios que lo utilizan, en muchos casos, a diario. Resulta especialmente conveniente implementar modificaciones en etapas tempranas del desarrollo, cuando son más manejables en un entorno controlado, en comparación con los ajustes que deben hacerse una vez que el producto ya está en el mercado. La elección de la metodología para elaborar los

informes de evaluación de experiencia del usuario es variada: puede ir desde simples cuestionarios en los que el usuario califica y describe algún problema o comentario hasta reportes más robustos que recopilan datos de comportamiento, como seguimiento ocular, análisis facial y comentarios en audio. Sin duda, este último enfoque, en el que no solo se consideran las observaciones del usuario, sino también el análisis de su interacción con el software, resulta el más completo, aunque también el que demanda más horas de trabajo humano para procesar la información. De acuerdo con la revisión del estado del arte, muchas de estas herramientas se centran en el desarrollo de algoritmos para la obtención de datos, pero no en su implementación orientada a la analítica visual o a la representación gráfica de la información, aspectos que facilitan al experto la detección de problemas de diseño en un producto de software. Solo unas pocas propuestas se enfocan en este aspecto, limitándose al uso de una o dos fuentes de información, lo que reduce la posibilidad de identificar con precisión las problemáticas que se evidencian en una evaluación más completa (Soure et al., 2021; Zeng et al., 2020). Además, incluso aquellas propuestas que sí incorporan analítica visual presentan limitaciones importantes: las representaciones suelen carecer de orden y coherencia en sus interfaces, y muestran falta de claridad en aspectos específicos de la información visual. En otras palabras, se prioriza el desarrollo de métodos algorítmicos para obtener los datos, mientras que la presentación de la información y el diseño de las interfaces quedan relegados.

En este trabajo se plantea, por tanto, una oportunidad de investigación: diseñar y desarrollar un reproductor que integre analítica visual con datos multimodales, con un enfoque centrado en la forma en que la información se representa, sin dejar de lado el modo en que se obtiene. El objetivo de integrar la analítica visual en el proyecto es reducir significativamente el tiempo que un experto dedica a elaborar informes de evaluación.

1.5 Alcances y limitaciones

A continuación, se presentan los alcances y limitaciones del proyecto, determinados por el tiempo y los recursos disponibles durante la investigación.

1.5.1 Alcances del proyecto

- Implementar analítica visual en el reproductor, tomando en cuenta los datos de seguimiento ocular, los análisis faciales y el audio de comentarios.
- Desarrollar una forma visual de segmentar la barra de reproducción de acuerdo con los comentarios de los usuarios.
- Integrar la analítica visual en el reproductor de UXLab, de modo que se enfoque en el video de la interacción del usuario.

1.5.2 Limitaciones del proyecto

- No se considerarán datos recolectados por sensores fisiológicos
- No se desarrollarán algoritmos para identificar emociones a partir de datos de seguimiento ocular, análisis facial o audios de comentarios.
- No se desarrollarán nuevos algoritmos para obtener datos adicionales de las fuentes de seguimiento ocular, de análisis facial o de audio de los comentarios.

Capítulo 2:

Marco teórico

En esta sección se describen los fundamentos teóricos utilizados en el tema de investigación para una mejor comprensión de este.

2.1 Experiencia de usuario

La Experiencia de Usuario (UX, por sus siglas en inglés) se refiere a la percepción global y emocional que un usuario tiene al interactuar con un producto, sistema o servicio. Norman (2013) resalta su importancia al afirmar que “el diseño no es solo cómo se ve y se siente un producto, sino también cómo funciona”.

Asimismo, la norma ISO 9241-210 define la UX como las percepciones y respuestas de los usuarios al utilizar un producto, sistema o servicio. Esta norma establece un marco para el diseño centrado en el usuario, destacando la importancia de comprender sus necesidades y capacidades para ofrecer interacciones efectivas. Además, subraya la participación activa de los usuarios a lo largo de todo el proceso de diseño, desde la conceptualización hasta la evaluación, con el propósito de crear interfaces intuitivas y funcionales.

2.2 Datos multimodales

Los datos multimodales se refieren a la información recopilada a través de múltiples fuentes o modalidades, que pueden incluir datos digitales, físicos, fisiológicos, psicométricos y ambientales. Este tipo de información se emplea para obtener una comprensión más completa y precisa de un fenómeno en particular. En el contexto del aprendizaje multimodal (MMLA), por ejemplo, los datos multimodales se utilizan para lograr una visión más detallada y precisa de los procesos de aprendizaje (Mu et al., 2020). Algunas de sus características principales son:

- **Diversidad de Fuentes:** Los datos multimodales pueden provenir de una variedad de fuentes, como registros de interacción, encuestas, entrevistas, grabaciones de sesiones de usuarios, análisis de emociones faciales, seguimiento ocular, entre otros.
 - **Complementariedad:** Cada modalidad de datos aporta una perspectiva única. Por ejemplo, mientras que los datos cuantitativos pueden proporcionar herramientas visuales objetivas, los datos cualitativos pueden ofrecer información detallada sobre las experiencias subjetivas de los usuarios.
 - **Interconexión:** Los datos de diferentes modalidades están interrelacionados para proporcionar una imagen más completa. La combinación de datos cuantitativos y cualitativos, por ejemplo, puede revelar patrones y contextos que podrían pasar desapercibidos al analizar cada modalidad por separado.
- Ejemplos de Modalidades de Datos Multimodales:
- Datos cuantitativos: datos numéricos, tiempos de interacción, tasas de éxito, etc.
 - Datos cualitativos: retroalimentación verbal, comentarios de los usuarios, opiniones subjetivas.
 - Datos visuales: grabaciones de sesiones de usuarios, mapas de calor, análisis de eye-tracking.
 - Datos emocionales: análisis de expresiones faciales, respuestas fisiológicas para evaluar las emociones del usuario.

- Datos de interacción y expresión facial: datos relacionados con los toques, gestos y movimientos del usuario.

El uso de datos multimodales en la evaluación de la UX permite a los profesionales obtener una comprensión más holística y precisa de cómo los usuarios interactúan con un producto, proceso o servicio. Esto no solo facilita la identificación de áreas de mejora, sino que también aporta información valiosa para la toma de decisiones centradas en el usuario y el diseño iterativo.

2.3 Analítica visual (Visual Analytic)

La analítica visual es un campo de investigación que combina técnicas de visualización y análisis de datos para permitir a los usuarios obtener información mediante la interacción con las visualizaciones. El objetivo principal de la analítica visual es facilitar la generación de conocimiento guiada por el usuario (Christino et al., 2022).

De acuerdo con Keim et al. (2008), la analítica visual se define como “el uso de datos interactivos para facilitar el descubrimiento de conceptos, patrones y anomalías en datos complejos”. Esta definición enfatiza la importancia de la interacción humano-computadora en el proceso analítico, donde la visualización de datos se combina con técnicas analíticas para generar conocimiento procesable.

En el contexto de la evaluación de la Experiencia de Usuario (UX), las técnicas de Visual Analytics se refieren al uso de métodos y herramientas de visualización de datos para analizar y comprender cómo los usuarios interactúan con un producto de software. Esta aproximación integra la visualización de datos con el análisis de la UX para proporcionar información significativa y accionable sobre el rendimiento del sistema y la satisfacción del usuario.

Sus principales características son:

- Visualización de Datos UX: Emplea representaciones visuales de datos relacionados con la experiencia del usuario, como mapas de calor, diagramas de embudo, gráficos de flujo de usuarios, entre otros, para facilitar la comprensión de patrones y tendencias.
- Interactividad: Proporciona herramientas interactivas que permiten a los profesionales de UX explorar visualmente datos complejos y realizar análisis para identificar áreas de mejora o patrones de comportamiento.
- Integración de Datos Multimodales: Combina datos de diferentes modalidades, como datos cuantitativos, comentarios cualitativos y grabaciones de sesiones de usuario, en visualizaciones unificadas para obtener una visión completa de la experiencia del usuario.
- Monitorización Continua: Facilita la monitorización continua de la experiencia del usuario a lo largo del tiempo, permitiendo a los diseñadores y profesionales de UX realizar ajustes iterativos en función de los cambios en el comportamiento y las preferencias del usuario.
- Identificación de Puntos de Fricción: Permite identificar visualmente los puntos de fricción o las áreas problemáticas en la interfaz de usuario, lo que ayuda a priorizar mejoras y optimizaciones.

Aplicaciones en la evaluación UX:

- **Análisis del recorrido del usuario:** Utiliza visualizaciones para seguir el recorrido del usuario a través de la interfaz e identificar dónde los usuarios tienden a tener dificultades o a abandonar el proceso.
- **Evaluación de la efectividad de las funcionalidades:** Visualiza la analítica clave sobre la eficacia de las funciones y características del producto para comprender su impacto en la experiencia del usuario.
- **Visualización de datos emocionales:** Incorpora herramientas de visualización para representar los datos emocionales recopilados durante la interacción del usuario, como reacciones faciales o comentarios emocionales, y así comprender mejor la satisfacción del usuario.
- **Comparación de variantes de diseño:** Permite comparar visualmente diferentes variantes para evaluar su impacto en la experiencia del usuario y tomar decisiones informadas sobre el diseño final.
- **Seguimiento de datos clave de rendimiento:** Utiliza visualizaciones para seguir datos clave de rendimiento, como tiempos de carga, tasas de conversión y de abandono, y proporciona una visión rápida del rendimiento general del producto.

2.4 Datos de comportamiento

2.4.1 Seguimiento ocular

El seguimiento ocular es una tecnología que permite comprender la atención visual del usuario. Existen dos tipos principales de sensores para registrar el movimiento ocular:

1. Los que utilizan cámaras de video para analizar el movimiento de los ojos, y
2. Los que emplean tecnología infrarroja para procesar la reflexión de la luz en la pupila del usuario, véase en la figura 3 un ejemplo visual.

En este último caso, la reflexión de la luz sobre la pupila es capturada por un sensor que calcula la posición en coordenadas (x, y), lo que representa el punto de la mirada del usuario en el plano espacial de la pantalla. El seguimiento ocular permite identificar en qué punto mira el usuario en un momento determinado, cuánto tiempo fija la atención en un elemento y la trayectoria que siguen sus ojos. En el campo de la UX, esta tecnología contribuye a comprender de manera más completa la percepción del usuario (Romano Bergstrom & Schall, 2014).

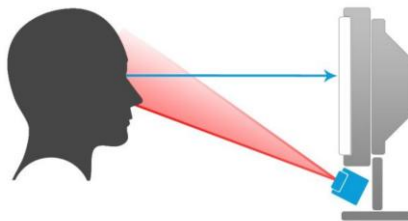


Figura 3. Esquema del funcionamiento del seguimiento ocular con sensor infrarrojo (Fouilloux, 2018).

2.4.2 Análisis facial

La detección de puntos de referencia faciales (King, 2015) se define como la tarea de identificar puntos clave del rostro y, posteriormente, rastrearlos, y es resistente a deformaciones faciales, tanto rígidas como no rígidas, producidas por movimientos de la cabeza y expresiones. Estos puntos se utilizan para localizar y representar regiones destacadas de la cara, como los ojos, las cejas, la nariz, la mandíbula y la boca. Esta técnica ha sido empleada en diversas aplicaciones, entre ellas la alineación facial, la estimación de la postura de la cabeza, el intercambio de rostros, la detección del parpadeo, la detección de la somnolencia y el reconocimiento de expresiones faciales (King, 2015).

En este contexto, la detección de puntos de referencia faciales requiere identificar las estructuras faciales relevantes mediante métodos de predicción de la forma. El proceso implica dos pasos principales:

1. Localización del rostro en la imagen.
2. Identificación de estructuras faciales clave, como:
 - Nariz
 - Mandíbula
 - Ojos izquierdo y derecho
 - Cejas izquierda y derecha
 - Boca

La propuesta se orienta a identificar unidades de acción (AU, por sus siglas en inglés) planteadas por Baltrušaitis et al. (2016), con el fin de detectar estados cognitivos como el interés o el aburrimiento. Para la distracción, se propone emplear las AU relacionadas con los movimientos de la cabeza y los ojos. El método de extracción de características y el algoritmo de clasificación serían los mismos que se utilizan para la detección de emociones, realizando las adecuaciones necesarias para las nuevas características, véase en la figura 4 un ejemplo visual.

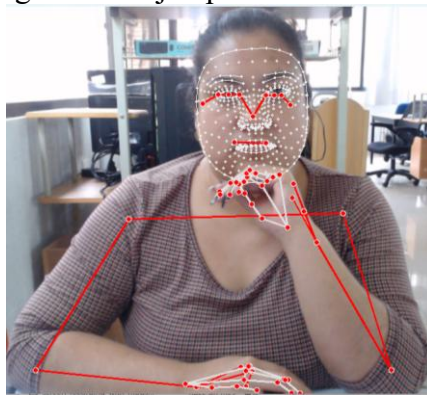


Figura 4. Análisis multimodal de expresión facial y postura (Toribio, 2023)

2.4.3 Protocolo Think Aloud

El protocolo de pensar en voz alta, también conocido como Think-Aloud Protocol, es una técnica que consiste en pedir a los usuarios que expresen en voz alta sus pensamientos mientras realizan una tarea a través de una interfaz digital, en la figura 5 se puede observar un ejemplo. La verbalización de estas percepciones se graba en audio, el cual posteriormente se transcribe automáticamente mediante STT para su

¿Ese comentario fue positivo, negativo o neutral?

UX

Think-Aloud Protocol

random content's comments, comments

Chris D. Random Person's Website

Permanent content

Comments

Voy a registrarme utilizando el prototipo del sistema web.

No entiendo cómo...

Es fácil de usar el ...

Me parece adecuado ...

La minería de opinión, también conocida como análisis de sentimientos, es un campo de estudio que se centra en identificar, extraer y cuantificar la información subjetiva en los textos, con el propósito de comprender las actitudes, emociones y opiniones expresadas por los autores (Liu, 2012). Esta disciplina combina técnicas de procesamiento de lenguaje natural, estadística y aprendizaje automático para detectar y clasificar opiniones en distintas fuentes, como redes sociales, foros, reseñas de productos y noticias. El proceso incluye varias etapas: recopilación de datos, preprocesamiento de textos, identificación de entidades y aspectos, detección de polaridad y agregación de resultados; véase la figura 6 como un ejemplo de la aplicación de la minería de opinión, una sopa de letras que indica que palabra está más presente utilizando el tamaño de las letras. La minería de opiniones se aplica en múltiples ámbitos, como el monitoreo de la reputación de las marcas, el análisis de la satisfacción del cliente, la predicción de tendencias de mercado y la toma de decisiones políticas. Su capacidad de analizar grandes volúmenes de datos textuales de manera eficiente la convierte en una herramienta valiosa para las organizaciones que buscan comprender mejor las percepciones del público.



Capítulo 3:

Trabajos relacionados

En esta sección se presentan los resultados de la revisión sistemática de la literatura (RSL) realizada según la metodología propuesta por García (2017). Esta revisión incluye las investigaciones más relevantes sobre el uso de la analítica visual para representar datos multimodales en el proceso de evaluación de la experiencia de usuario. La RSL abarca trabajos publicados entre 2015 y 2023. Se seleccionó un total de 20 estudios: 1 indexado en Scopus, 13 en JCR y 6 en congresos. La selección se basó en dos criterios principales:

1. Que presenten propuestas de analítica visual capaces de mostrar datos útiles para la evaluación de la experiencia de usuario, por ejemplo: segmentación en la barra de video o representación visual de estados emocionales.
2. Que usen técnicas de obtención de información relacionadas con investigaciones previas en las plataformas QUXBOX o UXLab, tales como el Think-Aloud Protocol, el seguimiento ocular y el reconocimiento facial, incluyendo, además, el uso de minería de texto.

En la primera parte, en el apartado 3.1, se incluyen trabajos que emplean analítica visual basada en datos obtenidos mediante seguimiento ocular. Dado que el uso de esta técnica en los análisis de experiencia de usuario no es tan común, resulta pertinente destacar estos estudios de manera especial.

En la segunda parte, todo el segmento 3.2, se agrupan los trabajos que presentan analítica visual enfocada en el reconocimiento facial y en el análisis de videos de usuarios durante la interacción con aplicaciones o servicios.

En la tercera parte, la final de este capítulo, el subcapítulo 3.3, se describen investigaciones que utilizan analítica visual aplicada exclusivamente al protocolo Think-Aloud y/o a la minería de texto, ya sea proveniente de transcripciones de audio, de retroalimentación u de otras fuentes escritas. Además, se incluyen estudios que, aunque se centran más en la representación de emociones o posturas que en la evaluación directa de la experiencia de usuario, resultan relevantes para esta tesis debido a la naturaleza del análisis, al explorar las emociones de los usuarios durante la interacción con el sistema. En esta revisión se excluyeron los trabajos relacionados con la detección de emociones mediante sensores fisiológicos (EEG, ECG y GSR), así como aquellos que abordan la regulación emocional a partir de recursos tecnológicos no vinculados a la evaluación de la experiencia de usuario (UX). La distribución por año se presenta en la tabla 1.

Los trabajos se obtuvieron mediante búsquedas en Google Scholar, IEEE Xplore, ACM, MDPI y ResearchGate. La selección se llevó a cabo en tres etapas: primero, considerando los títulos con mayor concordancia; después, evaluando la afinidad de los resúmenes con el tema de estudio; y finalmente, verificando la pertinencia de los resultados reportados.

Tabla 1. Proporción de artículos seleccionados por año.

Año	Número de artículos	Proporción (%)
2023	1	5%
2022	1	5%
2021	3	15%

2020	2	10%
2019	2	10%
2018	3	15%
2017	4	20%
2016	3	15%
2015	1	5%
Total	20	

3.1 Trabajos relacionados con seguimiento ocular

VETA: Visual eye-tracking analytics for the exploration of gaze patterns and behaviours (Goodwin, S., Prouzeau, A., Whitelock-Jones, R., Hurter, C., Lawrence, L., Afzal, U., & Dwyer, T., 2022)

En este trabajo se presenta VETA, una herramienta que incorpora analítica visual para el análisis de datos de seguimiento ocular, lo que permite explorar patrones y comportamientos de la mirada del usuario. VETA emplea técnicas de aprendizaje automático para identificar problemas de usabilidad y presentarlos a los profesionales de UX como apoyo en su proceso de análisis. La herramienta facilita tanto una visión general de los patrones de exploración visual como un análisis detallado de áreas específicas de la mirada. Además, se llevó a cabo un estudio con 30 profesionales de UX para evaluar la eficacia de VETA frente a otros métodos.

Este trabajo se vincula con la presente investigación, ya que constituye un ejemplo de analítica visual aplicada de manera útil en el ámbito de la experiencia de usuario.

VA2: A Visual Analytics Approach for Evaluating Visual Analytics Applications (Blascheck, T., John, M., Kurzhals, K., Koch, S., & Ertl, T., 2017)

En este trabajo se presenta VA2, una herramienta que incorpora analítica visual para evaluar aplicaciones de análisis visual. VA2 utiliza protocolos de pensamiento en voz alta, registros de interacción y seguimiento ocular para analizar las estrategias visuales de los usuarios. La herramienta ofrece un entorno de visualización altamente interactivo que permite mostrar y analizar de manera integrada los datos registrados mediante el protocolo de pensar en voz alta, así como la interacción y el movimiento ocular. Los algoritmos de búsqueda automática de patrones facilitan una exploración eficiente y apoyan el razonamiento para identificar patrones comunes de interacción entre la mirada y el pensamiento de los participantes. Uno de los principales problemas en los videos de experiencia de usuario que emplean los expertos para elaborar sus reportes es la ausencia de un marcador que indique con precisión el momento y el lugar en que se presenta un problema con el producto digital. En VA2, los problemas de usabilidad se describen de manera sencilla y accesible para los evaluadores.

M2Lens: Visualizing and explaining multimodal models for sentiment analysis (X. Wang, J. He, Z. Jin, M. Yang, Y. Wang, H. Qu., 2021)

En este trabajo se presenta M2Lens, un sistema interactivo que emplea la analítica visual para representar y explicar modelos multimodales en el análisis de sentimientos. M2Lens proporciona explicaciones sobre las interacciones intra e intermodales a nivel global, de subconjunto y local. En particular, resume la influencia de tres tipos de interacción típicos (dominancia, complemento y conflicto) en las predicciones del modelo. Además, identifica características multimodales frecuentes e influyentes y permite explorar de manera multifacética los comportamientos del modelo a partir de las modalidades de lenguaje, acústica y visual.

A Multimodal Deep Log-Based User Experience (UX) Platform for UX Evaluation (Hussain, J., Khan, W. A., Hur, T., Bilal, H. S. M., Bang, J., Hassan, A. U., Afzal, M., & Lee, S., 2018)

En este trabajo se propone una plataforma para evaluar la experiencia de usuario (UX) mediante un método de triangulación. La plataforma integra mediciones observacionales y fisiológicas, además de las tradicionales. Con ello, se busca reducir el sesgo subjetivo y validar las percepciones de los usuarios, registradas a través de distintos sensores, objetivando así la naturaleza subjetiva de la evaluación en UX. Asimismo, ofrece soporte plug-and-play para diversos dispositivos y capacidades de análisis avanzadas que permiten obtener información sobre la UX, considerando múltiples participantes. Los resultados muestran que la metodología propuesta supera a los enfoques existentes en la adaptación de interfaces de usuario, al incorporar el contexto y la experiencia de los usuarios. Por su relevancia y utilidad en la elaboración de reportes de experiencia de usuario, este trabajo fue considerado inicialmente como el principal referente en nuestra investigación y, aun con un panorama más amplio, continúa siendo una base sólida sobre la cual plantear nuestra propuesta.

3.2 Trabajos relacionados con expresión facial

uxSense: Supporting user experience analysis with visualization and computer vision (Batch, A., Ji, Y., Fan, M., Zhao, J., & Elmqvist, N., 2023)

En este trabajo se presenta uxSense, una herramienta que incorpora analítica visual y emplea métodos de aprendizaje automático para extraer el comportamiento del usuario a partir de grabaciones de audio y video tratadas como flujos de datos paralelos con marcas de tiempo. La implementación se basa en el reconocimiento de patrones, la visión por computadora, el procesamiento del lenguaje natural y el aprendizaje automático para identificar el sentimiento del usuario, así como sus acciones, posturas, palabras y otras características de las grabaciones. Estos flujos se representan como líneas de tiempo paralelas en una interfaz web, lo que permite al investigador buscar, filtrar y anotar datos en función del tiempo y el espacio. La herramienta destaca tres cualidades clave que deseamos replicar en este proyecto: la visualización de la pantalla

de la aplicación o del servicio, el video del rostro del usuario y la analítica visual de los datos obtenidos.

EmoCo: Visual Analysis of Emotion Coherence in Presentation Videos (Zeng, H., Wang, X., Wu, A., Wang, Y., Li, Q., Endert, A., & Qu, H., 2019)

En este trabajo se presenta EmoCo, una herramienta que incorpora analítica visual para facilitar el análisis eficiente de la coherencia emocional en videos de presentaciones. EmoCo permite a los usuarios obtener una visión general rápida de la coherencia emocional y de su evolución temporal, además de ofrecer una exploración detallada y comparaciones tanto a nivel de oración como de palabra. El sistema se evaluó mediante dos escenarios de uso: videos de TED Talks y entrevistas con dos expertos del dominio. Los resultados demuestran la efectividad de EmoCo para obtener información sobre la coherencia emocional en las presentaciones. La interfaz de EmoCo sirve como referente para el diseño del reproductor, retomando las tres cualidades clave mencionadas en el trabajo anterior, pero con mejoras en el apartado visual y en la implementación de la analítica visual.

Towards developing a tool for UX evaluation using facial expression (Munim, K. M., Islam, I., Khatun, M., Karim, M. M., & Islam, M. N., 2017)

El trabajo propone una herramienta web para evaluar la experiencia de usuario (UX) mediante la medición de las expresiones faciales. La herramienta es capaz de detectar el género y distintas emociones, como el compromiso, la valencia, el desprecio, la sorpresa, la ira, la tristeza, el asco, el miedo y la alegría. Este estudio parte de la premisa de que las expresiones faciales son un componente esencial de las manifestaciones físicas y pueden emplearse como datos precisos de satisfacción. La herramienta planteada ofrece una nueva forma de evaluar la UX, complementando los métodos tradicionales basados en cuestionarios o encuestas, que pueden introducir sesgos y limitar la precisión de los datos sobre los sentimientos de los usuarios objetivo. Se trata, nuevamente, de una representación gráfica relevante que abre la oportunidad de mejorar su apariencia e incorporar mejores enfoques de analítica visual en el diseño del reproductor.

3.3 Trabajos relacionados con el protocolo Think-Aloud o minería de opinión

CoUX: collaborative visual analysis of think-aloud usability test videos for digital interfaces (Soure, E. J., Kuang, E., Fan, M., & Zhao, J., 2021)

El trabajo presenta CoUX, una herramienta de analítica visual colaborativa para evaluar la usabilidad de interfaces digitales. CoUX facilita la identificación, anotación y discusión de problemas de usabilidad mediante técnicas de aprendizaje automático para extraer características acústicas, textuales y visuales de los videos de las sesiones de prueba. Además, CoUX permite la colaboración entre evaluadores de UX para registrar, comentar y consolidar los problemas detectados. Un estudio de usuario

demostró que CoUX resulta útil y efectivo tanto para identificar problemas como para fomentar el trabajo colaborativo. Se trata de una de las interfaces más depuradas de todas, y se aspira a seguir las buenas prácticas aplicadas en este trabajo para el desarrollo del reproductor de este proyecto.

UXmood—A Sentiment Analysis and Information Visualization Tool to Support the Evaluation of Usability and User Experience (Franco, R. Y. S., Lima, R. S. D. A. D., Paixão, R. D. M., Santos, C. G. R. D., & Meiguins, B. S., 2019)

En este trabajo se presenta UXmood, una herramienta web para la evaluación de la experiencia de usuario y la usabilidad que incorpora analítica visual. Combina datos de video, audio, registros de interacción y seguimiento ocular en un panel configurable. Funciona como un reproductor multimedia que permite a los evaluadores revisar el proceso de interacción del usuario y aplica técnicas de análisis de sentimientos para obtener información sobre la experiencia del usuario. La motivación principal para desarrollar UXmood fue apoyar el análisis conjunto de usabilidad y experiencia de usuario, sincronizar diferentes tipos de datos en un mismo panel y posibilitar el análisis de interacciones desde cualquier dispositivo mediante un navegador web. Al ser el trabajo en el que más enfoques de analítica visual se aplicaron dentro del alcance de nuestra investigación, UXmood constituye un buen precedente para identificar qué prácticas conviene adoptar y cuáles conviene evitar en el reproductor de este proyecto.

VisTA: Integrating Machine Intelligence with Visualization to Support the Investigation of Think-Aloud Sessions (M. Fan, K. Wu, J. Zhao, Y. Li, W. Wei, and K. N. Truon., 2020)

En este trabajo se presenta VisTA, una herramienta de analítica visual inteligente (Intelligent Visual Analytics) diseñada para asistir a los profesionales en la evaluación de la experiencia de usuario (UX) y de la usabilidad de las interfaces. VisTA emplea el protocolo de pensamiento en voz alta (Think-Aloud Protocol) y técnicas de aprendizaje automático para identificar problemas de usabilidad. Los profesionales de UX pueden analizar dinámicamente las sesiones mediante una visualización en línea de tiempo de las predicciones y de las características de entrada del modelo de aprendizaje automático. Un estudio demostró que los profesionales de UX identificaron más problemas al utilizar VisTA que con un enfoque de referencia. Al ser una de las primeras iteraciones en el desarrollo de herramientas web para el análisis de la experiencia de usuario, propuesta por Fan y Zhao, constituye un buen contraste respecto al desarrollo lógico que puede seguir una herramienta de esta naturaleza, línea que también adoptará el reproductor de esta investigación.

StanceVis Prime: visual analysis of sentiment and stance in social media texts (Kucher, K., Martins, R.M., Paradis, C., 2020)

En este trabajo se presenta StanceVis Prime, una plataforma que incorpora analítica visual para el análisis de sentimientos y posturas a partir de datos textuales temporales de diversas fuentes de redes sociales. StanceVis Prime procesa documentos

provenientes de múltiples transmisiones de texto, aplica clasificación de sentimientos y posturas, y ofrece a los usuarios finales tanto una visión general de las series de datos resultantes como una vista detallada para la lectura cercana y el examen de la salida de los clasificadores. Los escenarios de uso previstos para StanceVis Prime incluyen la monitorización de redes sociales y la investigación sociolingüística. Aunque StanceVis se aleja del ámbito tradicional de la “experiencia de usuario”, no deja de estar vinculado a la naturaleza de la minería de textos. Considerando esto, este trabajo demuestra que es posible emplear enfoques de analítica visual no centrados únicamente en las emociones y, aun así, constituir una herramienta útil para la evaluación de la experiencia de usuario.

TargetVue: Visual Analysis of Anomalous User Behaviors in Online Communication Systems (Cao, N., Shi, C., Lin, S., Lu, J., Lin, Y.-R., & Lin, C.-Y., 2016)

En este trabajo se presenta TargetVue, una herramienta que incorpora analítica visual para identificar comportamientos anómalos de los usuarios en sistemas de comunicación en línea. TargetVue emplea técnicas de aprendizaje automático para detectar incidencias y presentarlas a los profesionales de UX como apoyo en su análisis. La herramienta ofrece tanto una visión general de los patrones como un análisis detallado de la exploración local de la mirada. Además, se llevó a cabo un estudio con 30 profesionales de UX para evaluar la eficacia de TargetVue frente a otros métodos. Siguiendo la línea de dejar de lado el enfoque exclusivo en las emociones, puede parecer que la analítica visual presentada en este trabajo no resulta directamente útil para el análisis de la experiencia de usuario. Sin embargo, puede adaptarse si se toma en cuenta el texto, ya sea proveniente de feedback o de transcripciones del método Think-Aloud, lo que ofrece un nuevo uso para este desarrollo.

Exploring Emotions in Online Movie Reviews for Online Browsing (Bader, N., Mokryn, O., & Lanir, J., 2017)

En este trabajo se presenta un análisis orientado a determinar si las emociones expresadas en las reseñas reflejan las emociones evocadas durante la visualización de una película. Utilizando un conjunto de datos con millones de reseñas en línea de más de 9000 películas estrenadas en IMDb entre 1972 y 2015, los autores muestran que es posible extraer las emociones evocadas por una película a partir de sus reseñas y construir una “firma emocional” tanto para cada película como para cada género. Este estudio constituye un primer paso hacia un sistema de interfaz de usuario basado en emociones que permitiría a los usuarios navegar por las películas según las emociones que estas evocan. Si bien el enfoque de analítica visual en este trabajo no se relaciona directamente con el tema de la tesis, la forma en que se representan las emociones puede extrapolarse fácilmente a otros contextos. En particular, podría adaptarse al ámbito de la experiencia de usuario, utilizando la visualización de emociones para representar los resultados de experimentos y enriquecer el proceso de análisis en UX.

An Intelligent Interface for Organizing Online Opinions on Controversial Topics (Gao, M., Do, H. J., & Fu, W. T., 2017)

En este trabajo se presenta una interfaz inteligente para organizar opiniones en línea sobre temas controvertidos mediante análisis visual. La interfaz fue diseñada para permitir la visualización interactiva y la categorización de publicaciones originales sobre un tema controvertido, así como la captación de las reacciones y emociones de los participantes desde distintas posturas. Los autores evaluaron la interfaz a partir de publicaciones en Reddit sobre los candidatos presidenciales de EU. Los resultados demuestran que la herramienta puede mitigar la exposición selectiva y ayudar a los usuarios a acceder a un espectro más amplio de opiniones que el de la interfaz tradicional de Reddit. Aunque se desarrolla en un contexto completamente distinto, el uso de la representación de posturas puede adaptarse fácilmente. Ya sea para dar a conocer opiniones sobre una candidatura presidencial o para reflejar la postura de los usuarios respecto a un programa o sistema, este enfoque es útil y aplicable, con los cambios necesarios.

SentiCompass: Interactive visualization for exploring and comparing the sentiments of time-varying Twitter data (Wang, F. Y., Sallaberry, A., Klein, K., Takatsuka, M., & Roche, M., 2015)

En este trabajo se presenta SentiCompass, una herramienta basada en analítica visual para explorar y comparar los sentimientos en datos de Twitter que varían con el tiempo. El diseño de visualización de SentiCompass combina un modelo psicológico bidimensional del afecto (es decir, la emoción) con una representación en forma de túnel temporal. Para ilustrar su funcionamiento, se llevaron a cabo dos estudios de caso. Se trata de una propuesta clara, fácil de comprender y, sobre todo, muy útil para representar gráficamente las emociones de un gran número de personas a lo largo de un periodo determinado.

PyPlutchik: Visualising and comparing emotion-annotated corpora (Semeraro A, Vilella S, Ruffo G., 2021)

En este trabajo se presenta PyPlutchik, una biblioteca de Python diseñada específicamente para la visualización de emociones en textos o en corpus. PyPlutchik genera la flor de Plutchik, en la que cada pétalo se dimensiona según la presencia de una emoción en el corpus y, además, representa tres grados de intensidad para cada una de ellas. De manera destacada, PyPlutchik permite mostrar también diadas primarias, secundarias, terciarias y opuestas de forma compacta e intuitiva. Se trata de otro buen ejemplo de cómo llevar a cabo la representación gráfica de emociones y sentimientos, además de ofrecer instrucciones claras para utilizar la biblioteca en otros proyectos.

From Opinion Lexicons to Sentiment Classification of Tweets and Vice Versa: A Transfer Learning Approach (Bravo-Marquez, F., Frank, E., & Pfahringer, B., 2016)

En este trabajo se propone un modelo simple para transferir etiquetas de sentimiento de palabras a tweets y viceversa. El modelo representa tanto los tweets como las palabras mediante vectores de características que comparten el mismo espacio. Los

tweets se describen utilizando características estándar de procesamiento de lenguaje natural, como unigramas y etiquetas de parte del habla. Las palabras, por su parte, se representan promediando los vectores de los tweets en los que aparecen.

El enfoque se evaluó en dos problemas de aprendizaje por transferencia:

1. Entrenar un clasificador de polaridad a nivel de tweet a partir de un léxico de polaridad.
2. Inducir un léxico de polaridad a partir de una colección de tweets anotados con polaridad.

La cantidad de opiniones expresadas sobre un tema resulta especialmente relevante, ya sea en publicaciones de Twitter o en evaluaciones de usuarios sobre un sistema o servicio.

MultiStream: A Multiresolution Streamgraph Approach to Explore Hierarchical Time Series (Cuenca, E., Sallaberry, A., Ying Wang, F., & Poncet, P., 2018)

En este trabajo se presenta MultiStream, una plataforma visual diseñada para facilitar el análisis de múltiples series de tiempo. MultiStream emplea un enfoque de Streamgraph con múltiples resoluciones para explorar series temporales jerárquicas. Este enfoque permite analizar y comparar la evolución de las series en distintos niveles de granularidad mediante componentes visuales interactivos. Las series de tiempo múltiples se organizan en una estructura jerárquica, agrupando las individuales según su proximidad. Un ejemplo de analítica visual fuera de lo estándar, que se adapta perfectamente para mostrar cambios a lo largo del tiempo en torno a un tema general o para representar las emociones de un usuario en un periodo determinado.

ConceptVector: Text Visual Analytics via Interactive Lexicon Building Using Word Embedding (D. Park, S. Kim, J. Lee, J. Choo, N. Diakopoulos, and N. Elmqvist., 2018)

En este trabajo se presenta ConceptVector, una herramienta de analítica visual que guía al usuario en la construcción de conceptos y, posteriormente, los aplica para analizar documentos. Los conceptos se generan a partir de un pequeño conjunto de términos semilla mediante técnicas de incrustación de palabras. No obstante, la aplicación ingenua de estas técnicas puede generar falsos positivos debido a la polisemia del lenguaje humano. Para mitigar este problema, ConceptVector ofrece un análisis detallado de los documentos mediante estudios de caso con conjuntos de datos reales. Con el fin de apoyar la elaboración de conceptos, se introduce un modelo de concepto bipolar y se añade soporte para gestionar las palabras irrelevantes. Su relevancia radica en que, utilizada adecuadamente con el texto obtenido a partir del feedback de los usuarios o de transcripciones del Protocolo Think-Aloud, puede desempeñar un papel significativo en el análisis de la experiencia de usuario.

Which Feature is Unusable? Detecting Usability and User Experience Issues from User Reviews (Bakui, E., & Guzman, E., 2017)

En este trabajo se propone un enfoque para detectar automáticamente, a partir de las opiniones de los usuarios, las fortalezas y debilidades de la experiencia de usuario (UX) de un software. El método utiliza un algoritmo de colocación para extraer características, análisis de sentimientos mediante minería de opiniones en los textos para determinar la satisfacción de los usuarios respecto a cada característica, y técnicas de aprendizaje automático para identificar los problemas específicos de UX que afectan a la aplicación. Además, se presentan dos visualizaciones de los resultados. La evaluación inicial del enfoque, comparada con el juicio humano, arrojó resultados mixtos. Este trabajo resulta relevante porque la usabilidad y la experiencia de usuario tienen un impacto decisivo en la calidad y el éxito de un software.

Puede verse en la tabla 2 una comparación de los trabajos investigados, identificándolos por el tipo de dato que utilizan: seguimiento ocular, expresión facial, Think Aloud, texto o multimodal. Y además el tipo de detección que identifica si la detección de los datos se obtiene datos de emociones, posturas como estar de acuerdo o en desacuerdo u otro que no califican en esta distinción, pero pueden ser útiles.

Tabla 2. Comparación de trabajos sobre analítica visual.

#	Trabajo	Año	Autor(es)	Tipo de dato					Tipo de detección		Otros
				Seguimiento ocular	Expresión facial	Think Aloud	Texto	Multimodal	Emoción	Postura	
	CoUX: collaborative visual analysis of think-aloud usability test videos for digital interfaces	2021	Soure, E. J., Kuang, E., Fan, M., & Zhao, J.			✓			✓		
2	UXmood—A Sentiment Analysis and Information Visualization Tool to Support the Evaluation of Usability...	2019	Franco, R. Y. S., Lima, R. S. D. A. D., Paixão, R. D. M., Santos, C. G. R. D., & Meiguins, B. S.			✓			✓		
3	VisTA: Integrating Machine Intelligence with Visualization to Support the Investigation of Think-Aloud Sessions	2020	M. Fan, K. Wu, J. Zhao, Y. Li, W. Wei and K. N. Truong			✓		✓	✓		
4	uxSense: Supporting user experience analysis with visualization and computer vision	2023	Batch, A., Ji, Y., Fan, M., Zhao, J., & Elmqvist, N.		✓	✓		✓	✓		
5	VETA: Visual eye-tracking analytics for the exploration of gaze patterns and behaviours	2022	Goodwin, S., Prouzeau, A., Whitelock-Jones, R., Hurter, C., Lawrence, L., Afzal, U., & Dwyer, T.	✓							✓
6	EmoCo: Visual Analysis of Emotion Coherence in Presentation Videos	2019	Zeng, H., Wang, X., Wu, A., Wang, Y., Li, Q., Endert, A., & Qu, H.		✓	✓		✓	✓		

7	VA2: A Visual Analytics Approach for Evaluating Visual Analytics Applications	2016	Blascheck, T., John, M., Kurzhals, K., Koch, S., & Ertl, T.	✓		✓		✓			✓
8	StanceVis Prime: visual analysis of sentiment and stance in social media texts	2020	Kucher, K., Martins, R.M., Paradis, C. et al.				✓			✓	
9	TargetVue: Visual Analysis of Anomalous User Behaviors in Online Communication Systems	2016	Cao, N., Shi, C., Lin, S., Lu, J., Lin, Y.-R., & Lin, C.-Y.				✓			✓	
10	Exploring Emotions in Online Movie Reviews for Online Browsing	2017	Bader, N., Mokryn, O., & Lanir, J.				✓		✓		
11	An Intelligent Interface for Organizing Online Opinions on Controversial Topics	2017	Gao, M., Do, H. J., & Fu, W. T.				✓			✓	
12	SentiCompass: Interactive visualization for exploring and comparing the sentiments of time-varying Twitter data	2015	Wang, F. Y., Sallaberry, A., Klein, K., Takatsuka, M., & Roche, M.				✓		✓		
13	PyPlutchik: Visualising and comparing emotion-annotated corpora	2021	Semeraro A, Vilella S, Ruffo G				✓		✓		
14	From Opinion Lexicons to Sentiment Classification of Tweets and Vice Versa: A Transfer Learning Approach	2016	Bravo-Marquez, F., Frank, E., & Pfahringer, B.				✓			✓	
15	M2Lens: Visualizing and explaining multimodal models for sentiment analysis	2021	X. Wang, J. He, Z. Jin, M. Yang, Y. Wang, H. Qu	✓	✓	✓	✓	✓	✓		
16	MultiStream: A Multiresolution Streamgraph Approach to Explore Hierarchical Time Series	2018	Cuenca, E., Sallaberry, A., Ying Wang, F., & Poncelet, P.				✓		✓		
17	ConceptVector: Text Visual Analytics via Interactive Lexicon Building Using Word Embedding	2018	D. Park, S. Kim, J. Lee, J. Choo, N. Diakopoulos and N. Elmqvist				✓		✓		
18	A Multimodal Deep Log-Based User Experience (UX) Platform for UX Evaluation	2018	Hussain, J., Khan, W. A., Hur, T., Bilal, H. S. M., Bang, J., Hassan, A. U., Afzal, M., & Lee, S.	✓	✓	✓	✓		✓		
19	Which Feature is Unusable? Detecting Usability and User Experience Issues from User Reviews	2017	E. Bakiu and E. Guzman				✓				✓
20	Towards developing a tool for UX evaluation using facial expression	2017	K. M. Munim, I. Islam, M. Khatun, M. M. Karim and M. N. Islam		✓				✓		

Capítulo 4:

Propuesta de solución

La Figura 7 presenta la metodología para resolver el problema abordado en este trabajo de tesis. Esta se compone de 7 fases, las cuales se describen a continuación:

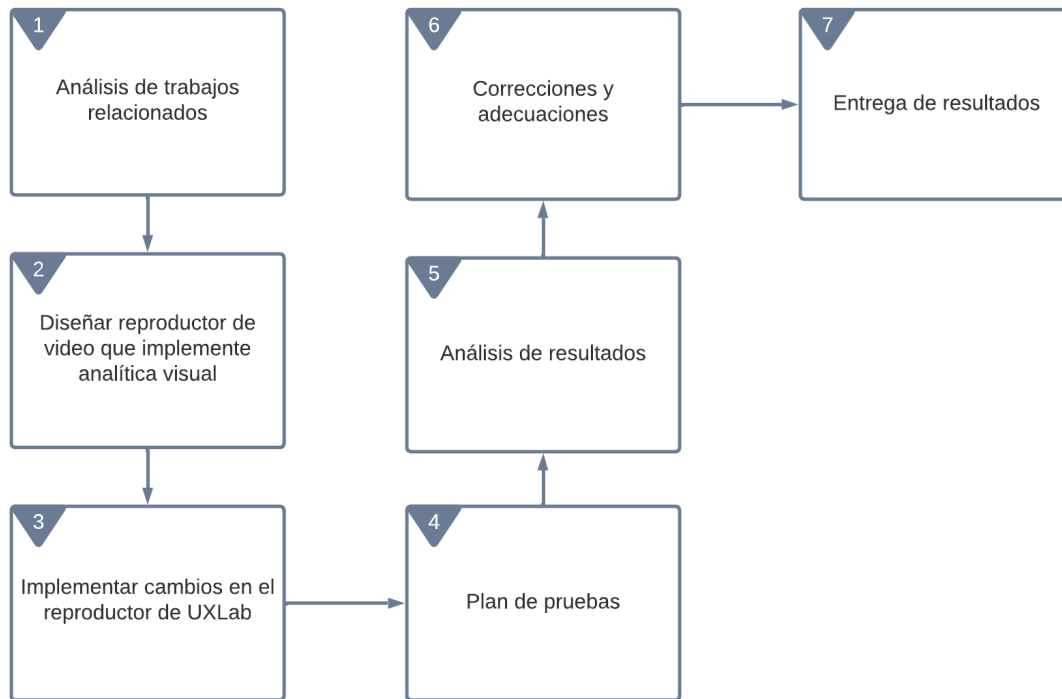


Figura 7. Metodología de solución.

4.1 Etapa 1: Análisis de trabajos relacionados

En esta etapa se llevó a cabo el análisis de trabajos relacionados, con el fin de identificar estudios previos vinculados al tema de investigación o que resultaran útiles para la solución planteada en esta tesis. La complejidad de esta fase radica en la selección de aquellos trabajos que utilicen algún enfoque de analítica visual para presentar información de seguimiento ocular; como en los trabajos de VETA (Goodwin, 2021) o VA2 (Blascheck, 2017), reconocimiento facial, como en EmoCo (Zeng, 2019), así como el uso de información derivada de técnicas de think-aloud, minería de texto o representación de emociones y sentimientos como en CoUX y UXmood (Soure, 2021; Franco, 2019).

De esta forma, se consideran tanto trabajos directamente relacionados con el desarrollo de herramientas para el análisis de la experiencia de usuario como aquellos que, mediante enfoques de analítica visual, presentan información de calidad que puede ser recontextualizada y empleada para la solución de la problemática abordada en esta tesis. Puede verse con mayor profundidad en el capítulo 3.

4.2 Etapa 2: Diseñar reproductor que implemente analítica visual

En esta etapa se diseñó la interfaz del reproductor implementando la analítica visual de los datos multimodales donde se tuvo ciertos criterios para su diseño:

1. Un enfoque central en el video sobre la interacción del usuario.

2. Colocar el video de la cara del usuario en una esquina.
3. Colocar botones para activar o desactivar la analítica visual implementada en el reproductor.
4. Implementar una analítica visual que segmente la barra de video en partes donde el usuario esté comentando

4.3 Etapa 3: Implementar cambios en el reproductor de UXLab

En esta etapa, se implementó todo el diseño realizado en la anterior sobre el antiguo reproductor para actualizarlo y mejorarlo durante el proceso. Además, se implementaron otros aspectos más interactivos en la interfaz para mantener la comunicación entre el reproductor y los evaluadores que lo utilizan.

4.4 Etapa 4: Plan de pruebas

Esta etapa se enfocó en la prueba del reproductor mediante el análisis de experiencias de usuario utilizando videos y datos de evaluación previamente generados en la plataforma UXLab, o bien con nuevos datos, en caso de que los anteriores no se ajusten correctamente a los parámetros de la plataforma desarrollada.

4.5 Etapa 5: Análisis de resultados

Etapa en la que se revisaron los resultados de la evaluación del prototipo y de la analítica visual desarrollada, mediante la elaboración de gráficas comparativas de la información. Estos resultados se contrastarán con los de evaluaciones previas de la experiencia de usuario o con los de otros sistemas enfocados en la misma tarea.

4.6 Etapa 6: Correcciones y adecuaciones

En esta etapa, se identificaron los errores generados durante la implementación de la plataforma web mediante la analítica visual. Estos pueden abarcar desde problemas visuales menores en la interfaz de usuario, dificultades en la visualización de datos en la analítica visual, inconvenientes en la entrega de reportes o en la presentación de la información de la evaluación de la experiencia de usuario, hasta cualquier otro tipo de error que pueda surgir durante el desarrollo de esta tesis.

4.7 Etapa 7: Entrega de resultados

Etapa final en la que se documentaron todos los procesos pendientes y se incorporaron las correcciones derivadas de la etapa anterior, con el fin de preparar la versión definitiva para su entrega.

Capítulo 5:

Recolección de datos

Para llevar a cabo una evaluación integral de la interfaz junto con la analítica visual desarrollada, fue indispensable realizar pruebas de experiencia de usuario en condiciones controladas, registrando todos los datos necesarios para su posterior procesamiento y análisis. Estos datos permitieron obtener las herramientas visuales requeridas, que luego se visualizaron en el reproductor de la plataforma UXLab, con el propósito de construir un repositorio de información confiable y alineado con los objetivos del estudio.

5.1 Naturaleza de los datos del repositorio

La naturaleza específica del proceso de grabación y reproducción de videos en el contexto de las evaluaciones de la experiencia de usuario imponen una serie de restricciones importantes. Por un lado, se requiere contar con datos generados mediante hardware especializado, como cámaras de seguimiento ocular compatibles con el sistema de recolección. Por otro lado, los videos y los datos asociados deben corresponder estrictamente a pruebas reales de UX, en las que las observaciones del comportamiento del usuario resultan especialmente relevantes.

Ante estas condiciones, utilizar un repositorio de datos externo implicaría enfrentar diversos desafíos. No solo supondría una tarea extremadamente laboriosa y costosa en términos de tiempo, sino que tampoco garantizaría que los datos obtenidos cumplieran con los criterios necesarios para una evaluación adecuada. Por ejemplo, es fundamental que las pruebas se hayan desarrollado bajo metodologías reconocidas, como el protocolo Think-Aloud, y que los dispositivos utilizados, como el modelo específico de la cámara de seguimiento ocular, sean plenamente compatibles con la infraestructura técnica de UXLab. Asimismo, se requiere que cada grabación esté orientada a una sesión de evaluación de la experiencia de usuario, ya sea estructurada mediante un guion o realizada como una prueba libre, siempre dentro del marco metodológico que rige estas prácticas. A esto se suma la necesidad de obtener un audio con poco ruido de fondo, a fin de no afectar los resultados de los modelos de transcripción de audio a texto empleados durante las pruebas.

Por estas razones, se optó por realizar las grabaciones y la recopilación de datos de manera autónoma y personalizada. Esta decisión permitió asegurar que los datos cumplieran con los estándares de calidad esperados, eliminando la incertidumbre sobre su integridad o relevancia. Al mantener el control total del proceso, se evitó el riesgo de obtener información incompleta, mal calibrada o irrelevante para los objetivos de la evaluación, lo que garantizó un entorno óptimo para validar la funcionalidad y la eficacia de la analítica visual integrada en el software UXLab.

5.2 Grabaciones de las pruebas

Las pruebas se planificaron cuidadosamente a partir de cuatro documentos clave, diseñados para guiar de manera estructurada cada etapa del proceso: antes, durante y después de la grabación. Estos archivos fueron elaborados con base en la experiencia

adquirida y en las observaciones realizadas durante una sesión previa de recolección de datos, en la que se detectaron diversos errores operativos y metodológicos.

A partir de dicha experiencia se identificaron aspectos susceptibles de mejora, lo que permitió optimizar el protocolo para esta nueva fase de pruebas, como prueba véase la figura 8 como un apoyo visual de este proceso.

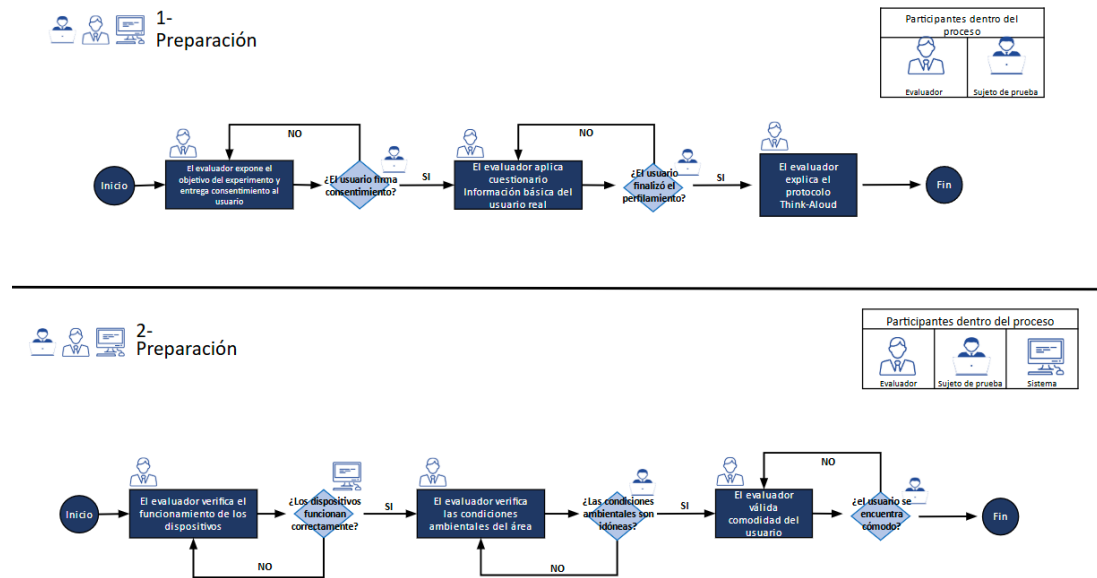


Figura 8. Parte del modelo de procedimiento para la obtención de datos.

De los cuatro documentos elaborados, uno estaba en formato impreso y los otros tres, digitales. El documento impreso consistía en una carta de consentimiento informado, entregada a todos los participantes antes de iniciar la prueba. En este documento se explicaba con claridad el propósito de la evaluación, los procedimientos a seguir, los tipos de datos a recolectar (incluidos video, audio y seguimiento ocular, entre otros) y el uso previsto de dicha información con fines académicos y de investigación.

Asimismo, se detallaban las tareas que el usuario debía realizar durante la sesión de prueba y se incluían cláusulas que garantizaban la confidencialidad de los datos personales. El consentimiento informado concluía con la firma del participante, lo que validaba su autorización expresa para utilizar los datos en el análisis posterior. Esta etapa resultó esencial, no solo desde el punto de vista ético, sino también para garantizar la transparencia del proceso y la legitimidad en el uso de la información obtenida.

El archivo de la carta de consentimiento puede consultarse en el Anexo A para corroborar toda la información presentada de manera general antes de iniciar la prueba.

En cuanto a los documentos digitales empleados durante la preparación de las pruebas, se trabajó con tres archivos fundamentales que abordaban distintos aspectos del proceso:

1. El procedimiento para el usuario.
2. La guía para el facilitador responsable de la prueba.

3. La gestión de acceso al prototipo utilizado para la evaluación.

El primero de estos documentos fue el “Procedimiento para el Usuario”, un archivo que describía cómo debía actuar el participante antes, durante y después de la prueba. Aunque no se entregaba directamente a los usuarios desde el inicio, servía como guía interna para los facilitadores, quienes ofrecían a los participantes un resumen verbal claro de las instrucciones. En caso de que surgiera alguna duda respecto al rol o al comportamiento del usuario durante la sesión, este documento podría presentarse como material de apoyo. En él se especificaban los parámetros esperados de la grabación, se aclaraban detalles sobre cómo debía desarrollarse la interacción y, además, se incluían ejemplos prácticos para facilitar la comprensión de los objetivos planteados.

El segundo documento, titulado “Guía para el Facilitador”, estaba dirigido exclusivamente al personal responsable de conducir y registrar las pruebas. Este archivo establecía, paso a paso, las acciones a seguir para asegurar tanto la correcta ejecución de la prueba como la calidad de los datos obtenidos. Se detallaban aspectos clave como la supervisión del cumplimiento del protocolo Think-Aloud, la verificación del funcionamiento de los dispositivos de seguimiento ocular, el monitoreo de los sistemas de análisis facial y la calidad del audio grabado. El objetivo principal de esta guía era garantizar la estandarización del proceso y evitar la pérdida de información crítica, promoviendo la obtención de datos consistentes, completos y útiles para el análisis posterior mediante analítica visual.

Por último, se incluyó un archivo para almacenar las credenciales de acceso utilizadas durante las pruebas. Este documento contiene los datos de inicio de sesión necesarios para acceder al prototipo del sistema en evaluación. Debido a ciertas limitaciones técnicas del prototipo, como restricciones en la gestión de múltiples usuarios, se optó por utilizar una única cuenta para todas las sesiones. Esta decisión permitió evitar modificaciones innecesarias que podrían retrasar el proceso o comprometer la estabilidad del entorno de prueba.

En conjunto, estos tres documentos digitales fueron fundamentales para estructurar y ejecutar de manera coherente y profesional todas las sesiones de prueba. Su uso permitió mantener el control sobre cada aspecto del proceso, minimizar errores operativos y asegurar que los datos recolectados respondieran fielmente a los objetivos establecidos en la evaluación de la experiencia de usuario.

Estos documentos pueden consultarse en los apartados de anexos: Anexo B, correspondiente al documento de Procedimiento para usuarios; Anexo C, correspondiente al archivo de Guía para facilitadores; y Anexo D, donde se incluyen los datos de inicio de sesión.

Con el fin de garantizar la calidad y consistencia de los datos obtenidos durante las pruebas, se decidió realizar todas las sesiones en un mismo espacio físico: un aula acondicionada específicamente para este propósito. Esta elección respondió a la necesidad de mantener condiciones controladas, especialmente en cuanto a la calidad del sonido capturado, la iluminación ambiental y la estabilidad de la imagen de la cámara web. Al reducir al mínimo las variables externas, se aseguró un entorno homogéneo en todas las grabaciones, lo cual resultó esencial para preservar la consistencia de las pruebas. Las especificaciones técnicas del equipo de cómputo y de los sensores especializados utilizados para recolectar datos sobre el comportamiento de los participantes se presentan en la tabla 3.

Tabla 3. Recursos que se usaron en las pruebas.

Recurso	Equipo
Plataforma donde se grabó (Laptop)	Lenovo LOQ 83JC0000US
Seguimiento ocular	GP3 HD Eye Tracker
Webcam	Logitech Streamcam Streaming USB-C
Micrófono	Hollyland Lark M2 Combo

El procedimiento general de grabación se estructuró en una serie de pasos claramente definidos:

1. Montaje del equipo técnico: Se preparaba el aula con todos los elementos necesarios para la grabación. Esto incluía la instalación de un monitor, un teclado, un ratón, una cámara web, una cámara de seguimiento ocular, un micrófono y el sistema de grabación de la plataforma UXLab. Se verificaba que todo el equipo estuviera debidamente configurado y en funcionamiento antes de iniciar cualquier sesión.
2. Invitación y consentimiento del usuario: Se le preguntó al participante si estaba dispuesto a colaborar en la prueba. En caso afirmativo, se le ofrecía una explicación general de los objetivos del estudio y, posteriormente, se le entregaba la carta de consentimiento informado para que conociera con claridad las actividades a realizar durante la sesión.
3. Aplicación de formulario demográfico: Como parte de las políticas de calidad y para apoyar las evaluaciones posteriores, se recopilaba información demográfica general mediante un formulario. Se omitía el nombre para preservar el anonimato de los participantes.
4. Explicación del protocolo y condiciones previas: Una vez firmado el consentimiento, se explican instrucciones detalladas sobre el rol del participante durante la prueba. Se explicaba el protocolo Think-Aloud y se le indicaba que debía verbalizar todas sus opiniones, pensamientos y percepciones mientras interactuaba con la interfaz. También se le pedía que colocara su teléfono en silencio y cerrara el aula para mantener el audio con el menor ruido de fondo posible.
5. Inicio de la grabación: Cuando el participante confirmaba haber comprendido las instrucciones, se iniciaba la grabación y se comenzaban las tareas asignadas.
6. Desarrollo de la prueba sin intervención directa: Durante la sesión se procuraba no intervenir para no sesgar el comportamiento del participante. No obstante, si transcurría un periodo prolongado sin comentarios verbales, el facilitador podía intervenir con una pregunta relacionada con la interfaz para reactivar la verbalización de los pensamientos.
7. Finalización de la prueba y recolección de datos posevaluación: Una vez completadas todas las actividades, se detenía la grabación. Mientras la plataforma UXLab procesaba automáticamente los datos recopilados, se solicitaba al participante que completara dos cuestionarios de evaluación: el

PSSUQ (Post-Study Usability Questionnaire) y el NASA-TLX (Task Load Index), con el fin de obtener métricas subjetivas de usabilidad y carga cognitiva.

Gracias al riguroso seguimiento de este protocolo y a la verificación constante de los criterios de calidad antes mencionados, así como de los que se encuentran en el documento para el facilitador de la prueba del Anexo C, se recopiló un total de 13 pruebas completas antes, durante y después de cada sesión. Todas presentaron una calidad adecuada en los datos esperados, incluyendo audio, video, seguimiento ocular, expresión facial y verbalización, lo que permitió contar con un conjunto de datos sólido y confiable para el análisis posterior mediante la analítica visual de la plataforma UXLab.

5.3 Datos esperados

Una vez garantizada la calidad de los datos durante la grabación, se estableció una estructura clara para los archivos que debían generarse al finalizar cada prueba de experiencia de usuario. En condiciones ideales, se esperaba que cada sesión generara cinco archivos, almacenados en la carpeta correspondiente al participante.

Los archivos eran los siguientes:

1. Video del escritorio (.mp4): grabación completa de la pantalla del usuario durante la prueba, que capture todas las acciones realizadas en la interfaz del sistema evaluado.
2. Video de la cámara web (.mp4): grabación del rostro del usuario para analizar expresiones faciales, movimientos y comportamiento no verbal durante la interacción.
3. Archivo de audio (.wav): pista de audio independiente, generalmente de mayor calidad, que recoge la voz del usuario siguiendo el protocolo Think-Aloud, lo que facilita el análisis posterior del contenido verbal.
4. Archivo de seguimiento ocular (.csv): denominado Gaze, contiene los datos recolectados por la cámara de eye-tracking, incluyendo coordenadas de fijación, movimientos oculares y tiempos exactos. Este archivo es clave para reconstruir el recorrido visual del usuario en la interfaz.
5. Archivo de eventos técnicos (.txt), también conocido como events, almacena información técnica detallada proporcionada por la cámara de seguimiento ocular, como los tiempos de inicio y fin de la sesión, el modelo del dispositivo utilizado, la resolución de captura y otros metadatos relevantes para la validación y sincronización de datos.

Estos cinco archivos constituyen la base esencial para el procesamiento posterior. A partir de ellos se extraen las representaciones visuales necesarias para alimentar los módulos de analítica visual desarrollados en la plataforma UXLab, lo que permite una visualización rica, precisa y estructurada del comportamiento del usuario durante la prueba.

Gracias al uso del formulario, se obtuvieron datos demográficos de 13 usuarios. En la figura 9 se muestra la distribución por género; en la figura 10, el rango de edad.

Porción de participantes de acuerdo a su género

■ Hombre ■ Mujer

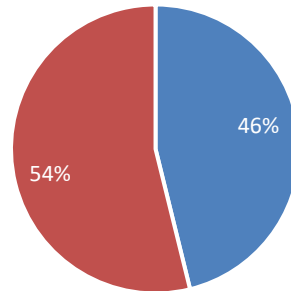


Figura 9. Distribución de los participantes por género.

Porción de participantes de acuerdo a su edad

■ 21 a 29 años ■ 30 a 39 años ■ 40 a 49 años

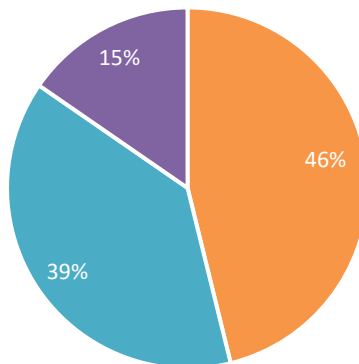


Figura 10. Distribución de los participantes según su rango de edad.

Capítulo 6: Experimentación y resultados

En esta sección se describe el proceso de experimentación realizado para evaluar el desempeño de la analítica visual desarrollada, así como los resultados obtenidos tras la aplicación de las pruebas finales. El objetivo fue validar tanto la funcionalidad de la herramienta como su utilidad práctica en el análisis de datos de la experiencia de usuario.

6.1 Métodos de evaluación

Debido a la naturaleza específica de una interfaz de usuario, el enfoque de evaluación difiere considerablemente del que se aplicaría en modelos de clasificación, donde es posible utilizar métricas cuantitativas tradicionales como Accuracy, Precision, Recall o F1-Score. En el caso de las interfaces, el rendimiento no puede medirse únicamente en términos de aciertos o errores, sino que requiere un enfoque más cualitativo y contextual.

Por esta razón, se optó por aplicar un método de evaluación basado en dos enfoques complementarios.

En primer lugar, se incorporó un análisis comparativo basado en el tiempo de revisión del contenido audiovisual, con el objetivo de medir la eficiencia de la herramienta en el proceso de evaluación. Para ello, se contrastaron dos escenarios:

1. El tiempo requerido para revisar el video completo de la prueba, sin ningún tipo de filtrado ni ayuda visual.
2. El tiempo necesario para revisar únicamente los segmentos en los que el usuario realizó comentarios verbales relevantes, extraídos mediante el protocolo Think-Aloud y visualizados en la analítica visual de UXLab.

En segundo lugar, se utilizaron los resultados de los cuestionarios PSSUQ y NASA-TLX para evaluar tanto la calidad percibida como la carga mental percibida por los propios usuarios durante la prueba. De esta forma, fue posible obtener datos cuantitativos sobre el desempeño del reproductor y, al mismo tiempo, datos cualitativos percibidos por los participantes.

6.1.1 Comparativa de tiempos completos vs por segmentación por comentarios

Esta técnica de evaluación permite estimar la eficiencia en el análisis de videos de experiencia de usuario al comparar el tiempo total necesario para revisar todo el contenido con el tiempo requerido para observar únicamente los segmentos en los que el usuario realiza comentarios verbales. La idea central es reducir la duración del proceso sin perder información relevante para la interpretación.

En su forma más básica, se mide el tiempo total del video original (tiempo de visualización completo) y se compara con la suma de los tiempos de los fragmentos que contienen comentarios (tiempo segmentado). La diferencia entre ambos valores permite estimar el ahorro temporal logrado mediante la segmentación. Además de mejorar la eficiencia, este enfoque facilita la focalización en los momentos clave de la interacción, reduciendo la carga cognitiva del evaluador al eliminar los períodos de silencio o de interacción pasiva sin aportes verbales significativos. La Fórmula 1 describe el cálculo del porcentaje de reducción temporal:

$$Reducción (\%) = \left[\frac{(Tiempo\ total - Tiempo\ segmentos)}{Tiempo\ total} \right] \times 100 \quad (1)$$

Esta métrica no solo ayuda a cuantificar la ganancia de tiempo, sino que también permite justificar el uso de herramientas de segmentación automática como apoyo en el análisis de pruebas de usabilidad.

6.2 Pruebas funcionales

Esta sección tiene como objetivo explicar en detalle cómo se implementaron, cómo funcionan y cómo deben interpretarse los distintos módulos de analítica visual integrados en la interfaz del reproductor UXLab. Se describe qué datos se utilizan, cómo se comportan durante la reproducción del video y cómo los evaluadores deben interpretarlos al analizar una sesión de experiencia de usuario.

6.2.1 Cambio principal de interfaz

Antes de detallar las analíticas visuales implementadas, se presenta el cambio general en la interfaz del reproductor de UXLab. En la figura 11 se muestra la interfaz original antes de las modificaciones, mientras que en la figura 12 se observa la versión modificada, con ajustes en la disposición de los elementos, mejoras en la organización visual y diversas adaptaciones orientadas a una mayor usabilidad.

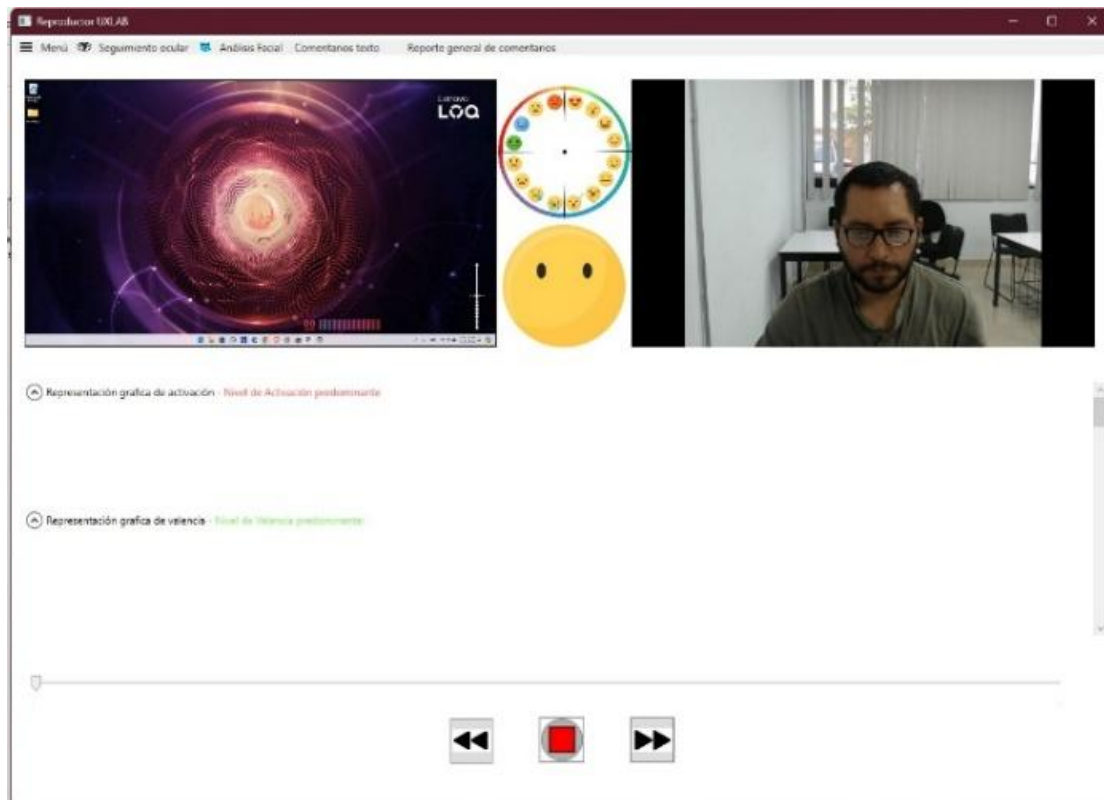


Figura 11. Interfaz del reproductor de UXLab original.

Como puede observarse en la figura 12, la interfaz fue diseñada con una disposición que prioriza el espacio destinado a la visualización de gráficas, respondiendo a la necesidad previamente identificada de mostrar datos fisiológicos, tal como ocurría en iteraciones anteriores del sistema. Esta decisión de diseño redujo parcialmente el protagonismo del área de reproducción de video, con el fin de otorgar mayor visibilidad a las analíticas visuales generadas durante las pruebas de experiencia de usuario. Para facilitar el acceso y la visualización de distintos tipos de datos, como la analítica visual de seguimiento ocular y el análisis facial, se incorporó un menú superior en la interfaz. Este menú permitía al evaluador seleccionar la métrica que deseaba analizar y, al hacerlo, reemplazaba el video original por otro preprocesado que mostraba la visualización correspondiente a dicha métrica.

Si bien este enfoque resultaba funcional, presentaba limitaciones importantes en términos de flexibilidad, ya que solo permitía observar una métrica a la vez, lo que dificultaba el análisis cruzado entre distintos tipos de datos. Estas restricciones fueron uno de los principales motivos para rediseñar tanto el sistema de procesamiento como la interfaz del reproductor en esta nueva iteración, lo que ahora permite una visualización más rica, dinámica y simultánea de múltiples representaciones visuales del mismo video base.

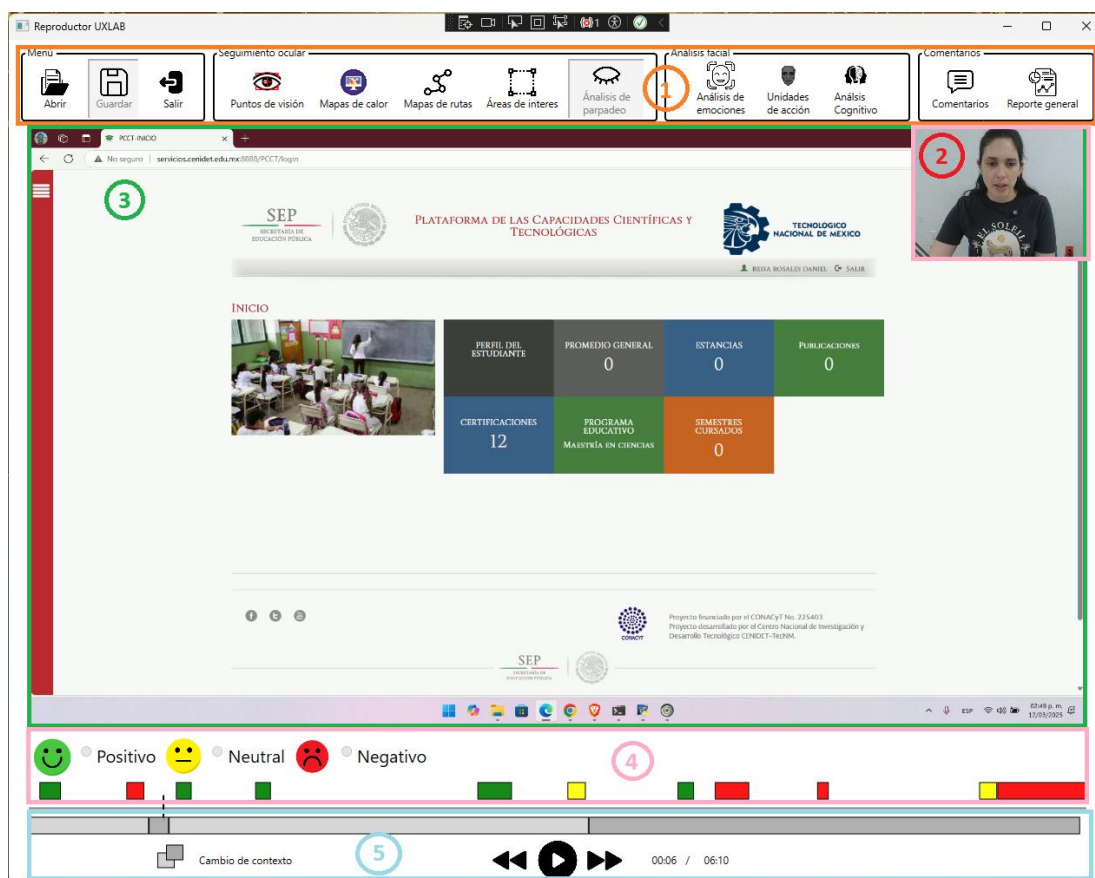


Figura 12. Interfaz del reproductor UXLab actual que se divide en 5 segmentos importantes: (1) Los botones de menú y de activación de la información visual. (2) Recuadro donde se reproduce el video de la webcam, que puede moverse y cambiar de tamaño. (3) La reproducción del escritorio. (4) La segmentación del video según los comentarios del usuario. (5) Segmentación del video según el cambio de contexto.

En la iteración actual de la interfaz se llevó a cabo una reestructuración significativa en la forma de presentar los datos, orientando el diseño hacia una visualización más clara, directa y centrada en la interacción del usuario con el prototipo. Esta nueva versión de la interfaz se organiza en cinco segmentos principales que pueden verse en la figura 12, cada uno con una función específica que facilita el análisis.

(1) En la parte superior se encuentran los botones de menú y los controles para activar distintas herramientas de analítica visual, lo que permite superponer diversas capas de información sin saturar la pantalla. (2) La cámara del usuario ahora se presenta como un recuadro flotante, movable y redimensionable, ubicado de forma estratégica para no obstruir la visualización del escritorio. (3) El área principal está ocupada por la reproducción del video del escritorio, priorizando la observación directa de la interacción del usuario con el sistema evaluado. (4) Justo debajo del video se incluye una barra de segmentación basada en los comentarios verbales del usuario, lo que permite identificar momentos clave relacionados con percepciones o problemas expresados durante la prueba. (5) Finalmente, se incorpora una segunda barra de segmentación que muestra los cambios de contexto durante la tarea, lo que facilita la navegación y el análisis de las transiciones entre las diferentes etapas de la interacción.

Este nuevo enfoque permite una experiencia de análisis más fluida y centrada en el comportamiento del usuario. Deja de lado las métricas fisiológicas, ya en desuso en UXLab, y favorece el uso de datos como el seguimiento ocular, los comentarios verbales y las expresiones faciales, lo que permite una evaluación más precisa y contextualizada.

6.2.2 Punto de visión

Objetivo:

Presentar una visualización sencilla del seguimiento ocular que permita identificar con precisión la ubicación de las fijaciones del usuario durante su interacción con la interfaz, incluso cuando sus comentarios verbales no especifican con claridad el componente al que se refieren.

Observaciones:

Durante la interacción, la herramienta genera puntos rojos que indican la posición exacta de cada fijación. En la figura 13 se aprecia una concentración de estos puntos en áreas específicas de la pantalla mientras el usuario emite un comentario negativo. Aunque sus palabras no señalan explícitamente el elemento de la interfaz al que alude, la distribución de las fijaciones permite inferir con claridad el componente observado en ese instante.

Resultado:

La visualización del punto de fijación demuestra ser una herramienta efectiva para complementar y aclarar los comentarios verbales del usuario. Permite identificar con precisión el elemento de la UI asociado al comentario, incluso en situaciones de ambigüedad o falta de claridad en la verbalización, lo que contribuye a elaborar reportes más completos y precisos sobre los problemas de usabilidad detectados.

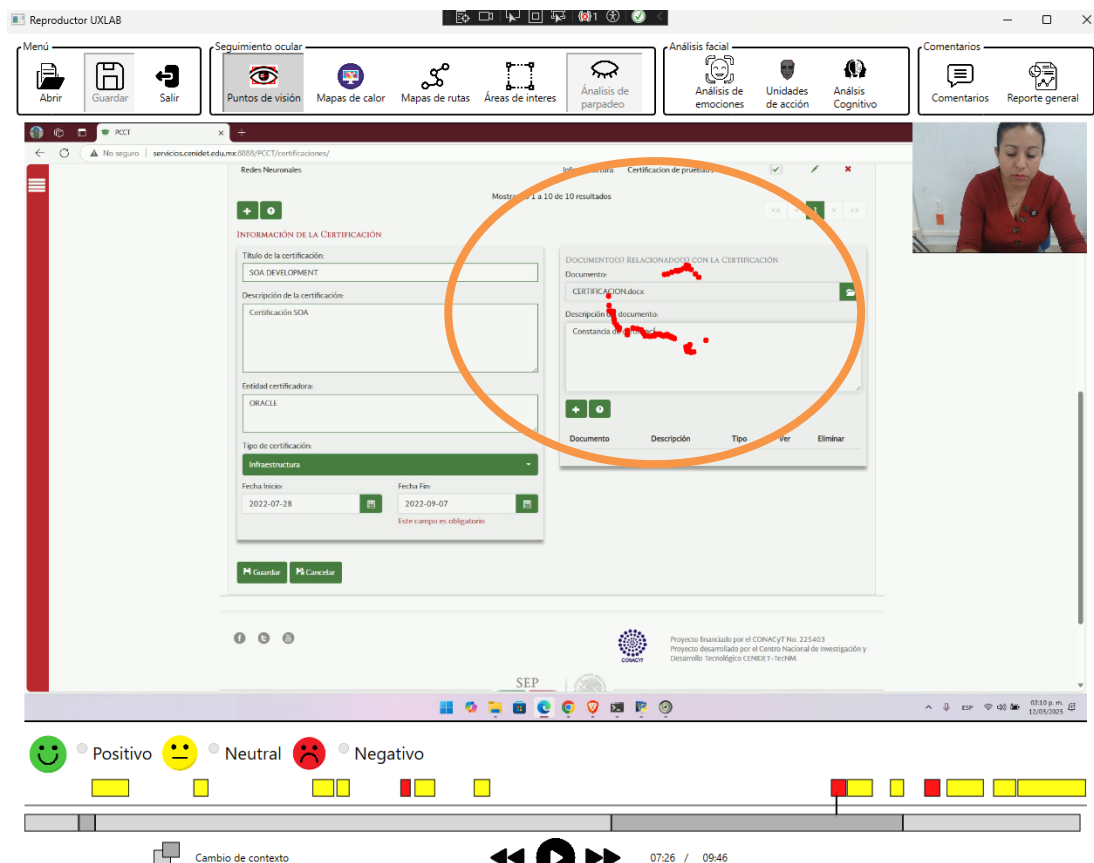


Figura 13. Representación de la situación en la que se observan los puntos de visión.

6.2.3 Mapa de calor

Objetivo:

Representar el uso del mapa de calor como una herramienta visual que permita visualizar la intensidad de la atención del usuario en distintas áreas de la interfaz durante la interacción.

Observaciones:

El mapa de calor representa de forma gradual las zonas en las que el usuario fijó la mirada con mayor frecuencia o durante más tiempo, mediante variaciones de color que indican el nivel de atención. Esta visualización permite identificar áreas de alta y baja concentración visual a lo largo de periodos específicos, incluso cuando no se cuenta con comentarios verbales que orienten la interpretación. Además, resulta útil en estudios comparativos, ya que facilita observar diferencias en los patrones de atención entre distintos grupos de usuarios. Por ejemplo, puede evidenciar qué elementos atraen más la mirada en personas jóvenes que en adultos, así como los momentos en los que estas diferencias ocurren durante la prueba, véase la figura 14 una captura de esta representación visual.

Resultado:

El uso del mapa de calor proporciona una representación visual clara y dinámica que facilita la detección de posibles problemas de diseño o de usabilidad. Asimismo, aporta información relevante para el análisis de patrones de comportamiento visual en

diversos perfiles de usuarios, lo que constituye un recurso valioso tanto para evaluaciones puntuales como para estudios comparativos más amplios.

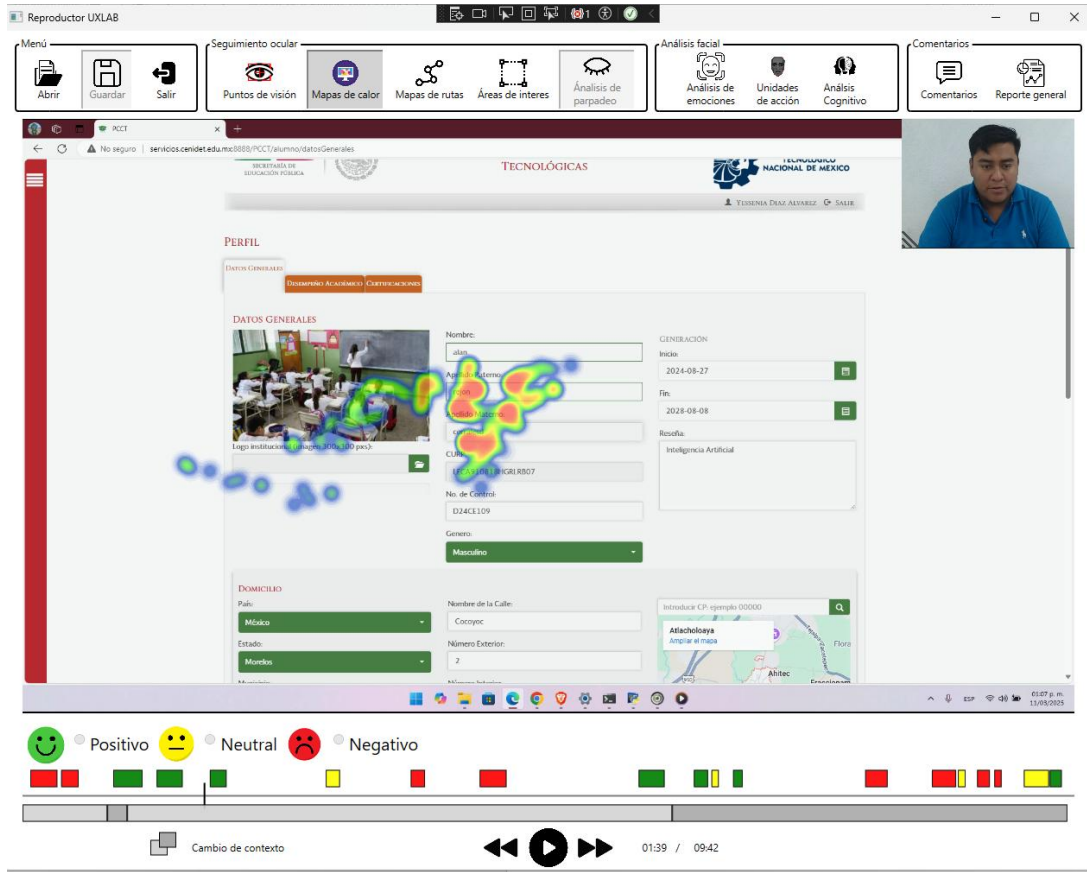


Figura 14. Prueba en la que se utiliza el mapa de calor para ver en qué parte se presta atención.

6.2.4 Mapas de rutas

Objetivo:

Utilizar el mapa de rutas o el scanpath como herramienta para visualizar secuencialmente el recorrido de la mirada del usuario durante la interacción con la interfaz.

Observaciones:

El scanpath conecta, mediante líneas, las fijaciones realizadas por el usuario, mostrando el orden y la dirección del recorrido visual. Esta representación permite analizar la distribución de la atención a lo largo del tiempo, identificar los elementos observados y comprender la secuencia y la fluidez del proceso de exploración. Asimismo, facilita la detección de patrones de navegación, de saltos de atención o de distracciones durante la prueba. Esta información es especialmente útil para entender el flujo cognitivo del usuario, ya que evidencia cómo procesa la información y toma decisiones visuales. También permite comparar usuarios o distintas versiones de un diseño, revelando diferencias en las estrategias de exploración que podrían no ser visibles con otras representaciones visuales, véase la figura 15 un ejemplo.

Resultado:

El mapa de rutas ofrece una visión dinámica y estructurada del comportamiento visual del usuario, lo que contribuye a detectar problemas de usabilidad, evaluar la eficacia del diseño y proponer mejoras basadas en el análisis del movimiento ocular en escenarios reales de uso.

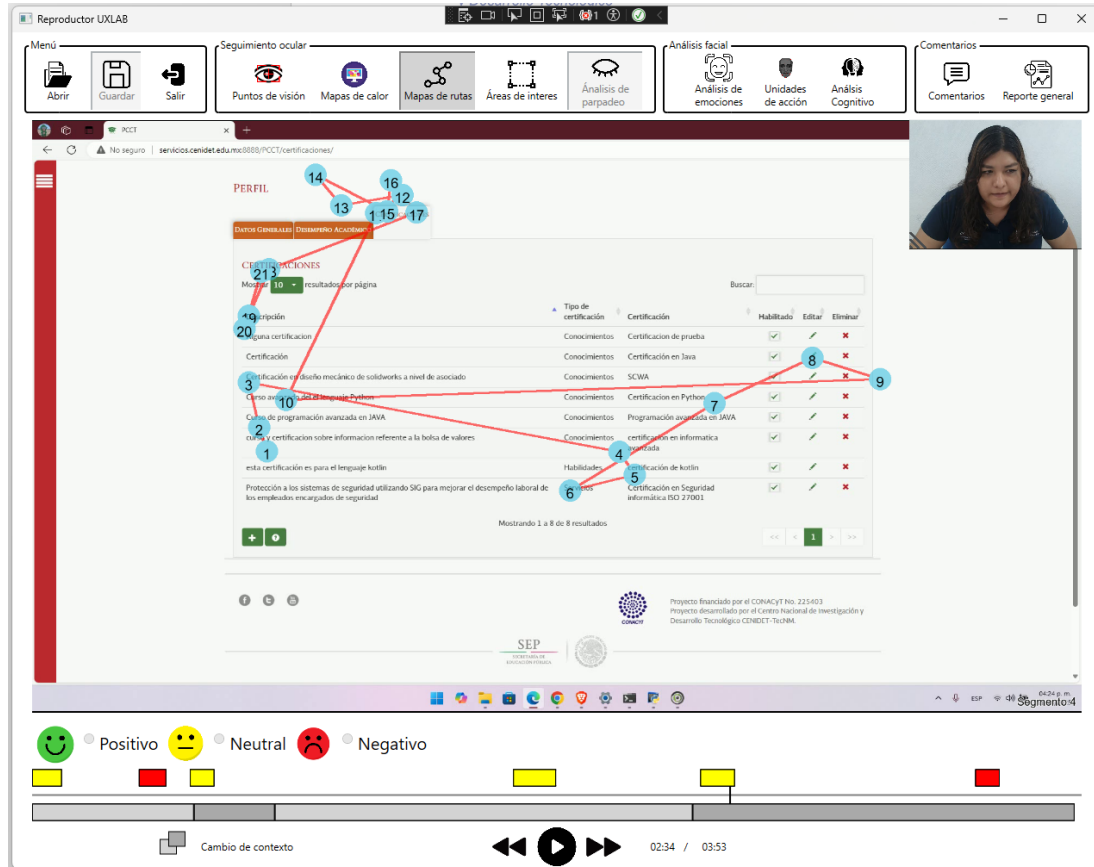


Figura 15. Prueba en la que se utilizan mapas de rutas para identificar la ruta de visión del usuario.

6.2.5 Áreas de atención

Objetivo:

Usar la herramienta de áreas de atención para identificar las zonas de la interfaz que resultaron visualmente relevantes para el usuario en segmentos específicos del video.

Observaciones:

Esta representación visual agrupa las áreas de atención en bloques de análisis denominados contextos, donde cada cambio de contexto corresponde a un periodo definido del video. A diferencia de representaciones continuas o acumulativas, esta organización permite observar con claridad y precisión qué regiones fueron relevantes en cada intervalo. Este enfoque ayuda a analizar momentos clave de la interacción sin perder la referencia al flujo general del uso, y facilita además la comparación entre distintos segmentos del video. La agrupación por contexto reduce el “ruido” visual que podría surgir de una visualización continua, optimizando la interpretación de patrones de atención y permitiendo identificar cambios en la relevancia de los elementos de la interfaz a lo largo del tiempo, véase un ejemplo visual en la figura 16.

Resultado:

Las áreas de atención segmentadas permiten un análisis más enfocado, estructurado y contextualizado del comportamiento visual del usuario. Esta representación visual revela con precisión qué partes de la interfaz fueron importantes en momentos específicos del video, lo que apoya la identificación de patrones de atención y la evaluación de posibles problemas de diseño o de usabilidad.

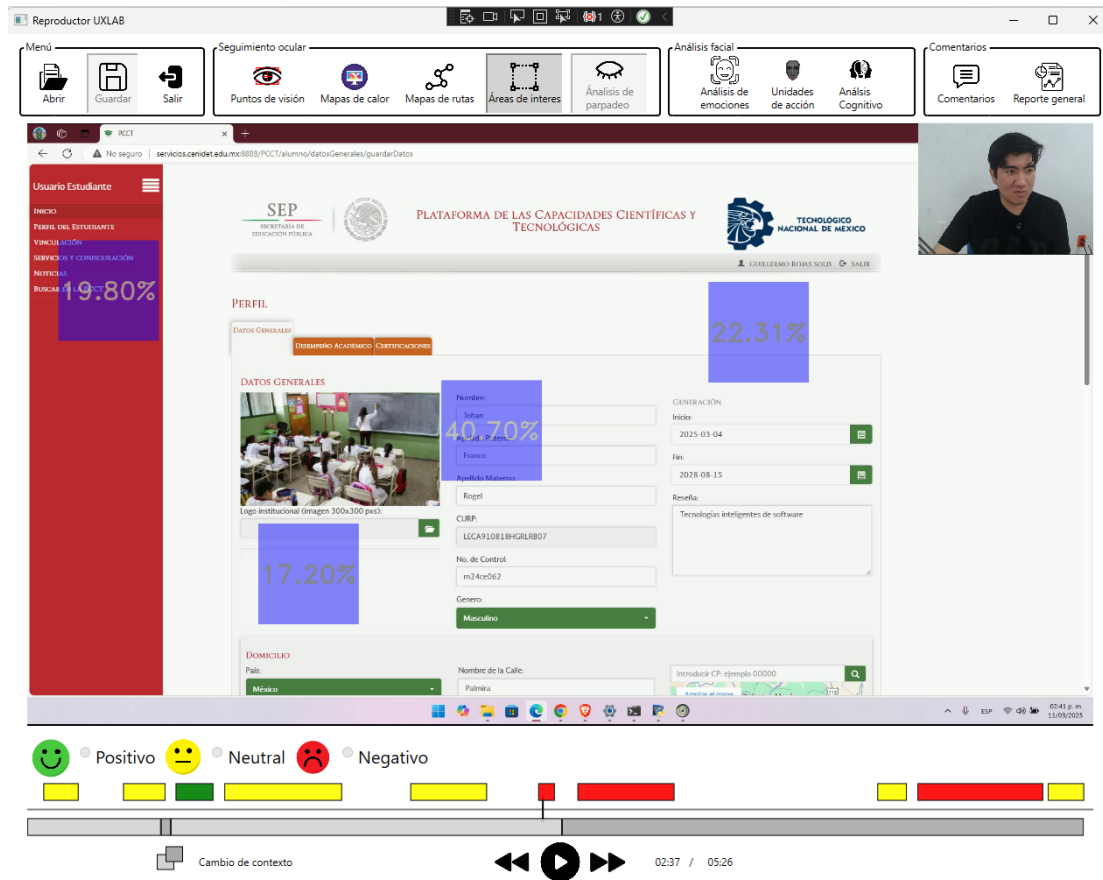


Figura 16. Prueba en la que se utilizan las áreas de atención.

6.2.6 Análisis de emociones

Objetivo:

Implementar el análisis de emociones faciales como forma de identificar los estados emocionales del usuario durante la interacción, aportando información no verbal que complementa el análisis basado en audio y comentarios.

Observaciones:

El análisis emocional interpreta expresiones faciales asociadas a emociones como la felicidad, el enojo, la sorpresa o la frustración. Esta herramienta es especialmente útil cuando el usuario no expresa verbalmente su experiencia, pero sus gestos reflejan emociones relevantes. Por ejemplo, un participante puede mostrar frustración al completar un formulario sin emitir ninguna queja. En estos casos, los comentarios del protocolo Think-Aloud no detectan el malestar, pero el reconocimiento emocional sí identifica expresiones de molestia o de confusión. En la figura 17 se observa un caso de este tipo: el usuario muestra signos claros de frustración mientras intenta completar

un formulario, lo que sugiere una problemática posiblemente relacionada con la falta de claridad en los datos requeridos por los campos del formulario.

Resultado:

El análisis de las emociones faciales aporta una capa adicional de evidencia no verbal que enriquece la evaluación de la experiencia de usuario. Permite detectar problemáticas que no se manifiestan explícitamente en el discurso del participante y orienta la identificación de aspectos del diseño que requieren mejoras, fortaleciendo así la interpretación global del comportamiento del usuario.

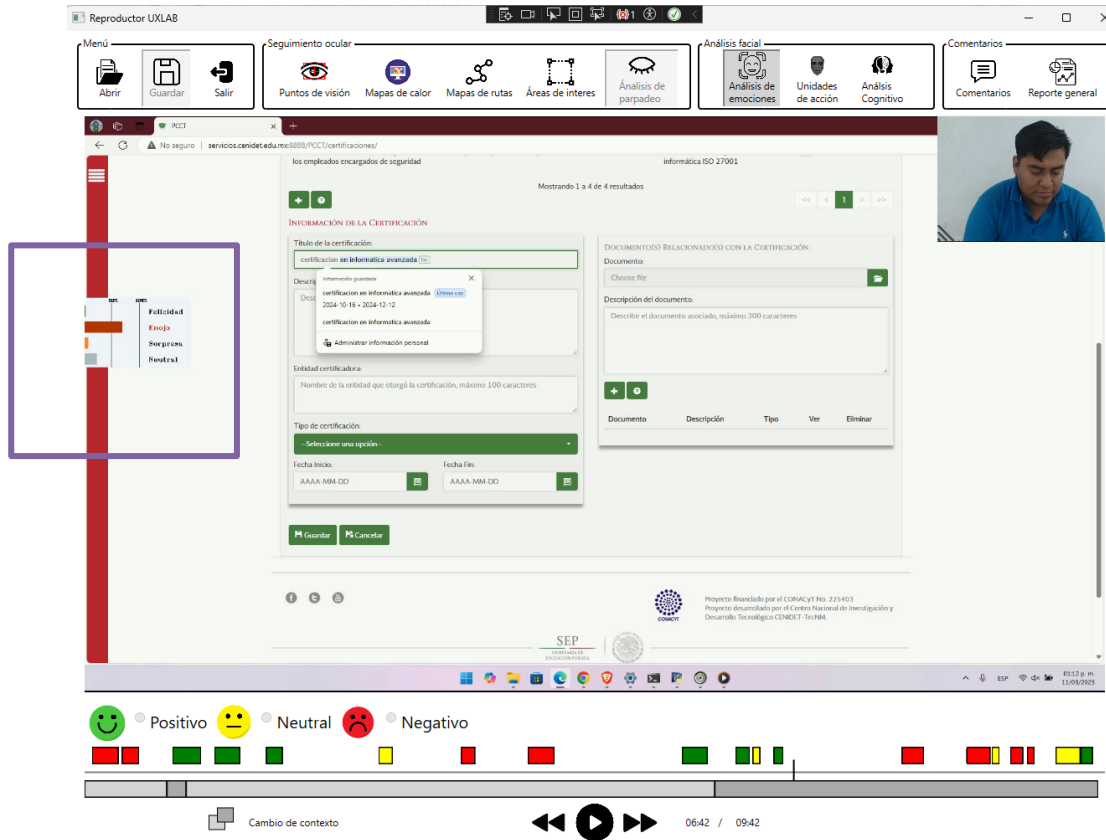


Figura 17. Prueba donde se visualizan las emociones del usuario durante la prueba.

6.2.7 Unidades de acción

Objetivo:

Adaptar la información de unidades de acción (AU) para identificar movimientos musculares específicos del rostro que permiten analizar con mayor detalle las expresiones faciales del usuario durante la interacción.

Observaciones:

La información AU descompone las expresiones faciales en movimientos musculares concretos, como el levantamiento de cejas, el fruncimiento del ceño o la contracción de los labios. A diferencia del análisis emocional, que clasifica emociones generales, las unidades de acción permiten detectar variaciones sutiles que no siempre corresponden a emociones estándar, pero que evidencian reacciones cognitivas o tensiones específicas frente a la interfaz. Por ejemplo, una elevación repentina de las cejas (AU1 y AU2) puede indicar sorpresa o confusión ante un cambio inesperado,

incluso sin manifestarse como una emoción completa. Esta representación visual es especialmente útil cuando el usuario mantiene un rostro aparentemente neutro, pero realiza microexpresiones breves que indican atención, esfuerzo cognitivo o incomodidad. En la figura 18 se muestra un caso de este tipo: ante un menú con múltiples opciones ambiguas, el usuario presenta una combinación de unidades de acción asociadas al esfuerzo mental, como la contracción del entrecejo (AU4) y el entrecerrar de los ojos (AU7). Además, la interfaz del reproductor permite mover y redimensionar la ventana de la cámara, lo que facilita al evaluador ampliar la visualización del rostro para analizar con mayor precisión las unidades de acción o, si se requiere, aplicar analítica visual posterior, como el análisis cognitivo.

Resultado:

El análisis mediante unidades de acción ofrece una visión más detallada y objetiva de las reacciones faciales del usuario, lo que permite identificar momentos críticos que podrían pasar desapercibidos en evaluaciones basadas únicamente en emociones o comentarios verbales. Esta herramienta enriquece la interpretación de la experiencia de uso al revelar señales de fricción, esfuerzo cognitivo o tensión, lo que apoya una evaluación más precisa del comportamiento del usuario frente a la interfaz.

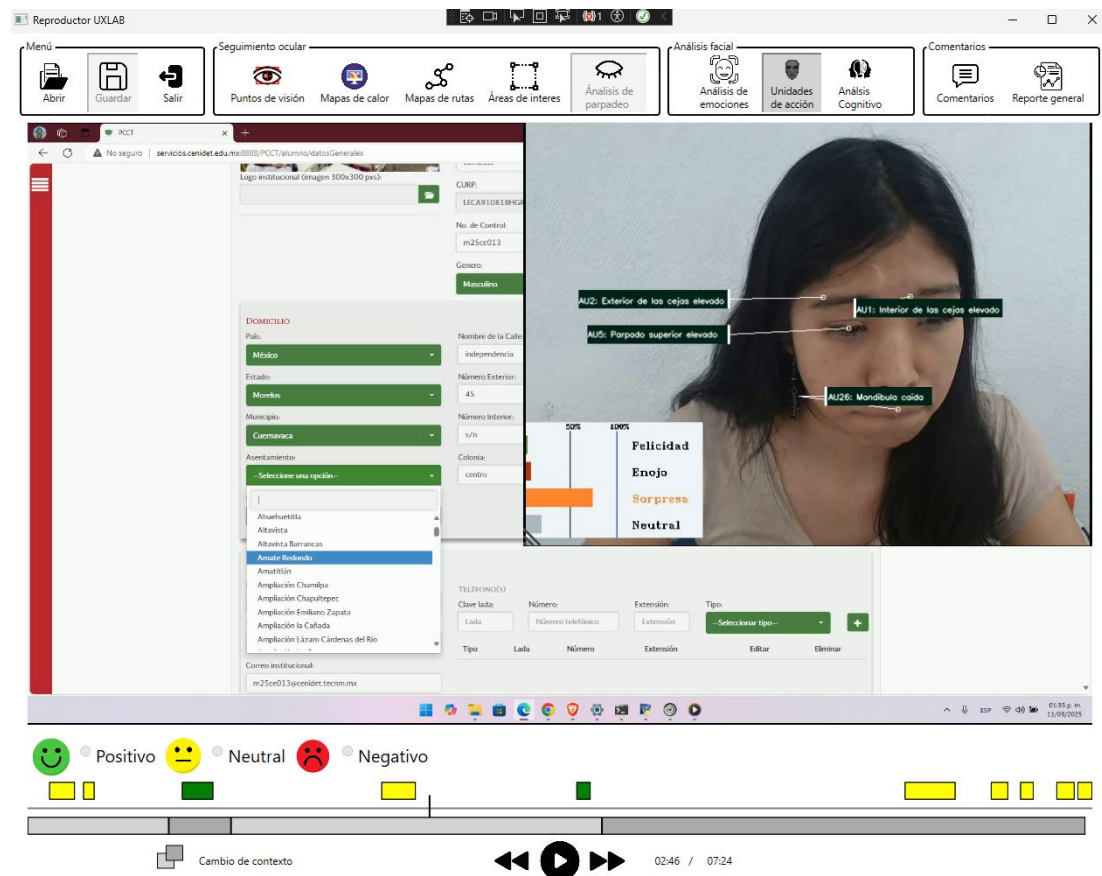


Figura 18. Prueba en la que se utilizan las unidades de acción.

6.2.8 Análisis cognitivo

Objetivo:

Implementar el análisis del estado cognitivo para identificar el nivel de carga mental o de esfuerzo que experimenta el usuario durante su interacción con la interfaz.

Observaciones:

El análisis del estado cognitivo se basa en indicadores como las expresiones faciales, los microgestos, la duración de las fijaciones visuales y el ritmo de parpadeo, lo que permite inferir si el usuario está concentrado, sobrecargado, distraído o relajado. A diferencia de herramientas centradas únicamente en las emociones o en los movimientos musculares, esta herramienta busca relacionar señales fisiológicas con procesos mentales involucrados en la toma de decisiones y la resolución de tareas. Por ejemplo, una expresión neutra sostenida, combinada con una reducción del parpadeo y un enfoque visual prolongado, puede indicar una alta concentración, mientras que movimientos oculares erráticos, microgestos o gestos repetitivos pueden sugerir fatiga o confusión. En la figura 19 se muestra un caso en el que un formulario complejo provoca una elevada carga cognitiva, evidenciada por una mirada fija prolongada y gestos corporales asociados al esfuerzo mental. Esta herramienta visual resulta especialmente útil para identificar momentos de dificultad cognitiva que no se manifiestan ni emocional ni verbalmente.

Resultado:

El análisis del estado cognitivo ofrece una comprensión más profunda de los procesos mentales del usuario durante la interacción, lo que permite detectar puntos de fricción, niveles de complejidad excesivos y momentos de sobrecarga. Esta información es clave para mejorar los flujos de navegación, ajustar la dificultad de las tareas y optimizar el diseño de la interfaz desde una perspectiva cognitiva, complementando de manera efectiva otras herramientas visuales y emocionales.

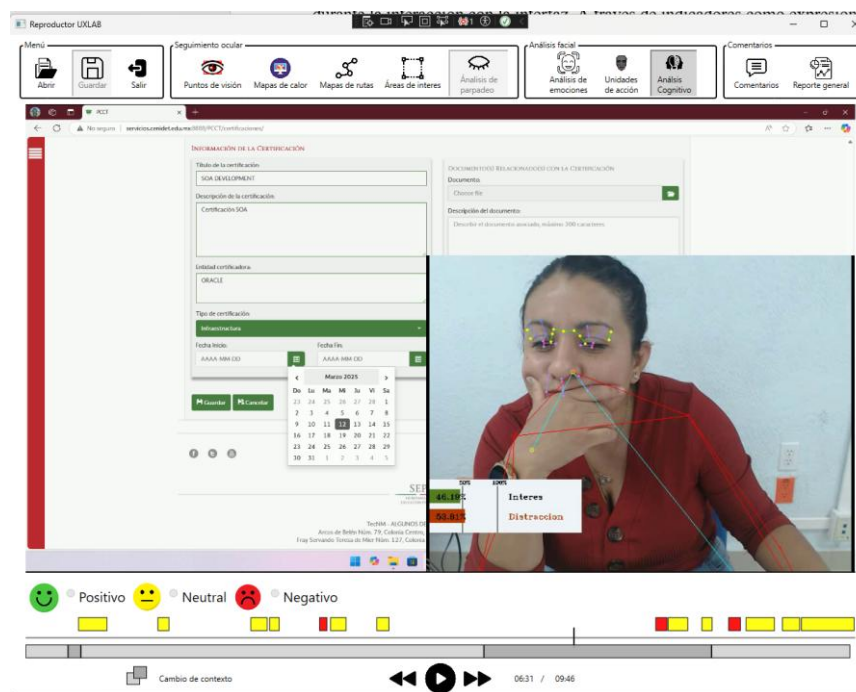


Figura 19. Prueba en la que se utiliza la información de análisis cognitivo.

6.2.9 Segmentación y etiquetado por comentarios

Objetivo:

Representar la segmentación del video basada en los comentarios del usuario mediante el Think-Aloud Protocol, incorporada al módulo reproductor de UXLab como una herramienta de analítica visual interactiva.

Observaciones:

La segmentación consiste en registrar el tiempo completo de cada intervención verbal del participante y generar un rectángulo que represente visualmente dicho segmento. El color del rectángulo indica la polaridad del comentario: verde para positivo, rojo para negativo y amarillo para neutral. Al hacer clic con el botón izquierdo sobre un segmento, el evaluador puede saltar directamente al momento del video en el que se realizó el comentario y verificar su contexto auditivo y la transcripción correspondiente. Con un clic derecho es posible consultar la transcripción en detalle. Si el etiquetado fue asignado de forma incorrecta, el evaluador puede modificarlo seleccionando una nueva categoría a través de los controles de emociones y guardar el ajuste con el botón correspondiente. Esta funcionalidad combina la analítica visual con la interacción directa, permitiendo revisar y corregir los datos procesados durante la evaluación, véase la figura 20 un ejemplo de cómo se ve.

Resultado:

La segmentación basada en comentarios proporciona una herramienta visual clara e interactiva que facilita el análisis y la validación de la información verbal emitida por el usuario. Permite identificar rápidamente los momentos relevantes del video, revisar el etiquetado emocional y corregirlo cuando sea necesario, lo que fortalece la precisión del análisis y mejora la eficiencia del proceso de evaluación de la experiencia de usuario.

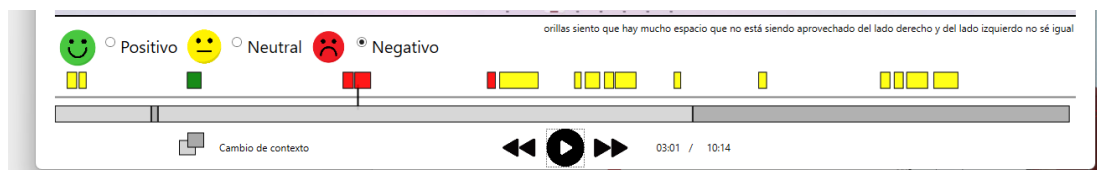


Figura 20. Segmentación del video según los comentarios y la transcripción.

Para implementar esta segmentación, se diseñó un algoritmo que utiliza tres librerías principales: una para la segmentación del audio, otra para la transcripción del audio a texto y una tercera para el etiquetado de los segmentos. A continuación, se detalla cuáles son estas librerías y cómo fueron utilizadas:

6.2.9.1 Segmentación de audio

Para la segmentación de audio se utilizó la librería Auditok, versión 0.3.0. Esta librería emplea el método Split, que permite dividir un audio en fragmentos al detectar aumentos en la energía de la voz. En otras palabras, identifica los momentos en los que el usuario pasa de un silencio a una intervención oral. Véase la figura 21 para un ejemplo visual de este proceso.

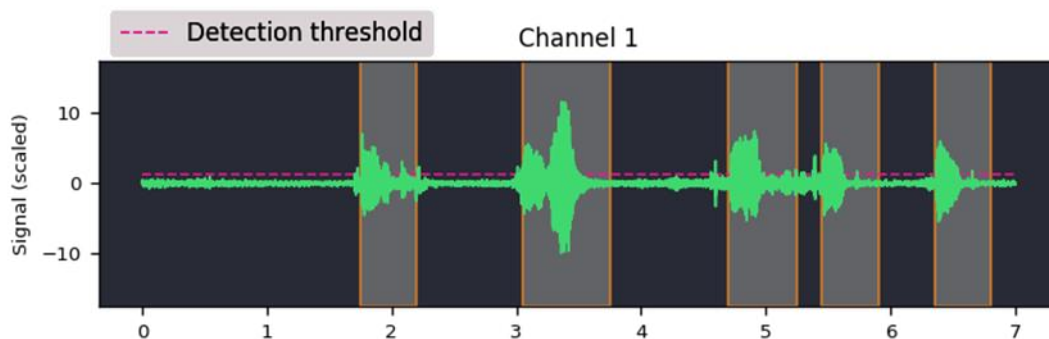


Figura 21. Representación visual de cómo Audiotok segmenta el audio.

Para configurar el método Split se requieren cinco parámetros que permiten ajustar la segmentación de acuerdo con las necesidades del análisis:

1. Archivo de audio: el archivo de entrada sobre el que se realizará la segmentación.
2. Duración mínima: tiempo mínimo (en segundos) que debe tener un segmento para ser considerado válido.
3. Duración máxima: tiempo máximo (en segundos) que puede alcanzar un segmento; si se supera este límite, el segmento se divide.
4. Silencio máximo: periodo de silencio permitido en un segmento. Este parámetro se aplica cuando el audio ya cumple con la duración mínima, extendiendo el segmento hasta el límite de silencio definido. Por ejemplo, si un segmento dura 5 segundos, con una duración mínima de 3 segundos y un límite de 3 segundos de silencio, el segmento abarcará los 5 segundos originales más otros 3 segundos de espera, por si el usuario continúa comentando.
5. Energy_threshold: umbral mínimo de energía que debe superar un fragmento de audio para ser considerado sonido válido o significativo. Este valor se ajusta según la calidad del audio y el nivel de ruido de fondo.

En la práctica, los parámetros que ofrecieron mejores resultados en la segmentación fueron: duración mínima de 4 segundos, duración máxima de 999 segundos (para no limitar el tamaño del segmento), silencio máximo de 3 segundos y un energy_threshold de 57. Adicionalmente, mediante el método save, es posible guardar cada segmento detectado en archivos independientes con extensión .wav, que contienen únicamente el fragmento correspondiente.

6.2.9.2 Transcripción de audio a texto

Para la tarea de transcribir el audio a texto se empleó la misma librería utilizada en iteraciones anteriores de UXLab: Whisper. Su buen desempeño ya había sido validado en la transcripción de audio a texto en trabajos anteriores de la plataforma, y tras revisar las alternativas disponibles no se encontró una mejora significativa en otras librerías. Por esta razón, se decidió mantener su uso en esta etapa.

La transcripción se realizó con el modelo “large”, especificando el parámetro de idioma como “es” para procesar audios en español. Como insumo, se utilizaron los fragmentos de audio generados mediante el método save de la librería AudioKit, lo que permitió integrar ambas herramientas de forma directa.

6.2.9.3 Etiquetado de segmentos

Por último, el etiquetado de los segmentos se realizó mediante la librería PySentiment, configurada con el modelo de sentimientos (task = "sentiment") y el idioma español (lang = "es"). De esta forma, fue posible obtener un puntaje de polaridad para cada comentario, clasificándolo como positivo, negativo o neutral. Todas estas librerías se integraron en un único archivo de Python, que posteriormente generaba un archivo de salida en formato .txt. En este archivo se registraba, para cada segmento:

- el número de segmento,
- el tiempo inicial (en segundos),
- la etiqueta detectada,
- el puntaje de las tres etiquetas obtenidas en el modelo,
- una variable de control inicializada en 0 para indicar si los datos del segmento habían sido modificados,
- la transcripción del audio a texto, y
- el tiempo final del segmento (en segundos).

En el Algoritmo 1 se presenta el pseudocódigo del proceso empleado para segmentar el audio.

Algoritmo 1. Segmentación de audio.

Entrada: archivo de audio.wav de la prueba de experiencia de usuario

Salida: carpeta llamada regions con archivos .wav de las segmentaciones y archivo datosAudio.txt con los datos de segmentación.

1. Dividir el archivo de audio en múltiples regiones usando un detector de silencio, configurado con una duración mínima, una duración máxima, un silencio máximo y un umbral de energía.
 2. Informar al usuario de que se está cargando el modelo de análisis.
 3. Inicializar un analizador de sentimiento en español.
 4. Cargar un modelo de transcripción automática (ASR) preentrenado.
 5. Crear un directorio para almacenar las regiones segmentadas.
 6. Abrir un archivo de texto para registrar los resultados del análisis.
 7. Para cada región de audio segmentada:
 - a. Asignar un nombre único al archivo de la región.
 - b. Guardar la región como archivo de audio si no existe previamente.
 - c. Transcribir el contenido de la región utilizando el modelo de ASR.
 - d. Analizar el texto transcrito para identificar el sentimiento.
 - e. Escribir en el archivo de texto:
 - i. el índice de la región,
 - ii. el tiempo de inicio,
 - iii. el sentimiento detectado,
 - iv. las probabilidades del análisis,
 - v. un valor fijo (cero),
 - vi. el texto transcrito,
 - vii. y el tiempo de fin.
 8. Informar al usuario de que el procesamiento ha finalizado.
-

6.2.10 Segmentación por cambio de contexto

Objetivo:

Identificar los momentos en que el usuario cambia de una ventana o vista a otra dentro del video, segmentando la interacción por contextos para asegurar que cada información de seguimiento ocular se analice únicamente dentro de su vista correspondiente.

Observaciones:

La analítica visual, situada en la parte inferior del reproductor, muestra rectángulos que representan los segmentos del video asociados a cada ventana utilizada por el usuario. Estos rectángulos señalan los instantes en que ocurre un cambio de contexto, dividiendo el video según las estampas de inicio y de fin de cada vista. Aunque la herramienta es sencilla, resulta altamente útil, ya que limita la información de seguimiento ocular a los datos de una misma vista, evitando interferencias o sesgos que surgirían si no se consideraran los cambios de contexto. Para su funcionamiento, el reproductor requiere un archivo denominado contexto.txt que debe contener un 0 como primer valor, seguido de los tiempos (en segundos) en que ocurre cada cambio de contexto y finalizar con la duración total del video. Esto se realiza automáticamente solo tomando en cuenta el inicio y el final del video, pero se puede modificar manualmente para complementar los datos.

Resultado:

La segmentación por cambio de contexto permite analizar con precisión la interacción del usuario en cada ventana del sistema, garantizando que los datos visuales se interpreten de manera coherente con la vista activa, véase la figura 20. Además, establece una base sólida para futuras mejoras del sistema, en las que se espera integrar un algoritmo que detecte automáticamente los cambios de contexto de manera óptima.

6.2.11 Sobreposición de herramientas e información

Objetivo:

Permitir la visualización simultánea de múltiples analíticas visuales en el reproductor UXLab mediante la generación y superposición de imágenes con fondo transparente, superando la limitación de las versiones anteriores, en las que solo era posible mostrar una información visual a la vez.

Observaciones:

En versiones previas del reproductor, cada analítica visual se procesaba directamente sobre el video base, lo que generaba un archivo único por métrica. Esto impedía combinar varias herramientas visuales simultáneamente, ya que cada una quedaba incrustada de forma estática en su respectivo video. Para resolver esta restricción, se realizaron dos modificaciones principales:

1. Reestructuración de los métodos de procesamiento.

Se reutilizó la lógica existente que generaba las analíticas visuales, pero en lugar de producir un video con la métrica incrustada, se generaron series de imágenes con fondo transparente que contienen únicamente la herramienta visual. Estas imágenes funcionan como capas independientes que pueden superponerse al video original. Para asegurar la fluidez y la sincronización, se definió una frecuencia de 24 imágenes por segundo.

2. Adaptación del reproductor para la superposición dinámica.

El reproductor fue modificado para renderizar imágenes superpuestas en tiempo real en lugar de reproducir un video con las analíticas integradas. Este cambio hizo posible combinar múltiples capas visuales, como mapas de calor, áreas de atención, fijaciones o rutas de mirada, de manera simultánea y eficiente.

Gracias a esta modificación, el sistema ahora ofrece una experiencia de análisis más flexible, permitiendo que las diferentes fuentes de datos visuales interactúen entre sí durante la reproducción. Además, la combinación de métricas amplía la capacidad del evaluador para identificar patrones y comprender mejor el comportamiento del usuario. Por ejemplo:

- La segmentación del audio permite localizar un comentario relevante.
- Las fijaciones muestran exactamente dónde miraba el usuario durante dicho comentario.
- El mapa de calor permite identificar zonas con mayor actividad visual.
- El scanpath ayuda a detectar si el usuario navega de manera errática o poco eficiente.
- Las áreas de atención revelan qué regiones de la interfaz son más relevantes.
- Las unidades de acción permiten comprender reacciones cognitivas o tensiones.
- El análisis emocional identifica frustración, estrés o confusión.
- El análisis cognitivo detecta la sobrecarga mental o la pérdida de atención.

Resultado:

La implementación de la sobreposición de imágenes transparentes permitió superar la limitación de visualizar una sola analítica a la vez y habilitó la combinación simultánea de múltiples métricas visuales. Esto ofrece una visión más completa e integrada del comportamiento del usuario, permitiendo identificar con mayor precisión los problemas de interacción, comprender el estado emocional y cognitivo asociado, y priorizar adecuadamente los hallazgos en el proceso de mejora del prototipo. El resultado final es un sistema de análisis más potente, flexible y eficaz para la evaluación de experiencias de usuario, véase la figura 22.



Figura 22. Prueba en la que se utilizan varias representaciones visuales a la vez sin ningún problema.

6.3 Comparativa de tiempos total vs tiempo segmento

Debido a la necesidad de contar con una evaluación cuantitativa que refleje mejor los resultados obtenidos, se compararon dos escenarios: el tiempo requerido para revisar los videos en su totalidad y el tiempo necesario para revisar únicamente los segmentos correspondientes a los comentarios del usuario, véase esos tiempos en la tabla 4.

Objetivo:

Cuantificar la reducción del tiempo de análisis al comparar el tiempo requerido para revisar los videos completos con el tiempo necesario para revisar únicamente los segmentos asociados a los comentarios del usuario, a fin de obtener una métrica precisa del ahorro temporal.

Tabla 4. Comparación de tiempos de revisión.

#No.	Tiempo total del video	Tiempo con segmentos	Reducción en tiempo
1	482 segundos	208 segundos	56.84 %
2	729 segundos	180 segundos	75.30 %
3	884 segundos	306 segundos	65.38 %
4	614 segundos	147 segundos	76.05 %
5	326 segundos	188 segundos	42.33 %
6	582 segundos	189 segundos	67.52 %
7	444 segundos	96 segundos	78.37 %
8	575 segundos	222 segundos	61.39 %
9	233 segundos	40 segundos	82.83 %
10	325 segundos	129 segundos	60.30 %
11	330 segundos	15 segundos	95.45 %
12	361 segundos	102 segundos	71.74 %
13	586 segundos	184 segundos	68.60 %
Promedio	498 segundos	154 segundos	69.07 %

Retomando la fórmula que comentamos en su momento para calcular la reducción del tiempo:

$$Reducción (\%) = \left[\frac{(Tiempo\ total - Tiempo\ segmentos)}{Tiempo\ total} \right] \times 100 \quad (2)$$

Aplicando la ecuación 2 da como resultado:

$$Reducción (\%) = \left[\frac{(8.18 - 2.34)}{8.18} \right] \times 100$$

Transformador todo a segundos para tener un mejor cálculo da como resultado:

$$Reducción (\%) = \left[\frac{(498 - 154)}{498} \right] \times 100 = 69.07\%$$

Observaciones:

Así pues, el resultado obtenido es comparable al porcentaje de reducción del tiempo reportado en pruebas anteriores, con la diferencia de que dichas pruebas carecían de un proceso que garantizara la calidad de los datos. Esto ocasionó que elementos relevantes,

como el audio o los registros de seguimiento ocular, no se almacenaran correctamente o presentaran errores, lo que derivó en una representación deficiente de las herramientas visuales. Los tiempos recopilados en esas pruebas pueden consultarse en la tabla 5; si bien no cumplen plenamente con los criterios de calidad vigentes, sirven como referencia para el proceso de segmentación.

En esta etapa, gracias a una mejor planificación, a la optimización de la duración de las pruebas y a la reducción de la cantidad total de datos por sesión, la disminución del tiempo obtenido es menos pronunciada que en las pruebas anteriores; no obstante, sigue siendo representativa para evaluar la eficiencia del proceso de análisis segmentado.

Tabla 5. Tiempos de la segmentación de las primeras pruebas (Sin estándar de calidad).

Id	Tiempo total del video	Tiempo total de todos los segmentos	Reducción en tiempo
1	22.28 minutos	10.12 minutos	54%
2	22.41 minutos	11.15 minutos	50%
3	15.36 minutos	6.50 minutos	58%
4	22.51 minutos	9.23 minutos	59%
5	19.09 minutos	9.57 minutos	50%
6	12.07 minutos	5.42 minutos	55%
7	22.35 minutos	11.25 minutos	50%
8	20.31 minutos	8.05 minutos	60%
9	12.24 minutos	6.22 minutos	49%
10	15.56 minutos	7.12 minutos	54%
11	13.14 minutos	5.33 minutos	59%
12	18.55 minutos	9.26 minutos	50%
13	15.31 minutos	6.37 minutos	58%
14	16.38 minutos	8.11 minutos	50%
15	10.24 minutos	4.44 minutos	57%
16	15.19 minutos	7.41 minutos	51%
17	9.26 minutos	4.18 minutos	55%
18	16.33 minutos	8.12 minutos	50%
19	20.12 minutos	10.24 minutos	49%
20	5.43 minutos	2.30 minutos	58%
21	14.37 minutos	7.28 minutos	49%
22	10.09 minutos	7.11 minutos	30%
23	12.49 minutos	6.36 minutos	49%
24	12.49 minutos	5.43 minutos	57%
25	10.14 minutos	5.27 minutos	48%
26	7.03 minutos	4.29 minutos	39%
27	7.34 minutos	4.12 minutos	44%
28	19.26 minutos	10.45 minutos	46%
29	16.10 minutos	7.27 minutos	55%
30	8.06 minutos	4.54 minutos	44%
	14.73 minutos /	7.08 minutos / promedio	52%

Resultados:

En términos generales, los resultados muestran que continuar siendo una estrategia eficaz consiste en evaluar únicamente los segmentos de tiempo en los que el usuario emite comentarios relevantes. Esta metodología permite a los evaluadores concentrarse directamente en los momentos clave de la interacción, reduciendo el tiempo invertido sin comprometer la profundidad del análisis. Comparación con cuestionarios y resultados, comentarios de usuarios.

6.4 Calidad del sistema y carga mental de usuarios

La última forma de evaluar la implementación consiste en comparar los resultados obtenidos en los formularios aplicados al finalizar la prueba de experiencia de usuario con las observaciones derivadas de la segmentación y del etiquetado del audio.

En el análisis del cuestionario PSSUQ, que mide la calidad del sistema, la información y la interfaz, se obtuvieron resultados mayoritariamente positivos, con una puntuación promedio de 80 puntos en sus tres dimensiones. De este formulario puede concluirse que la percepción de los usuarios respecto al prototipo fue favorable. Véase la tabla 6 para el resumen de los valores obtenidos.

Tabla 6. Evaluación del formulario PSSUQ.

Factores	Preguntas	Score General	Evaluación
Calidad el sistema (SYS)	(P1-P6)	5.65	81
Calidad de la información (INFO)	(P7-P12)	5.79	83
Calidad de la interfaz (INTER)	(P13-P16)	5.81	83
CALIDAD GENERAL (OVERALL)	(P1-P16)	5.75	82

En el análisis de los cuestionarios NASA-TLX, que miden el esfuerzo realizado por los usuarios, así como su rendimiento y nivel de frustración, se encontró que ambas tareas requirieron un esfuerzo medio. En palabras de la propia evaluación: “Se requirieron concentración y cierto trabajo”. Los resultados detallados se observan en las tablas 7 y 8, correspondientes a la evaluación de las tareas 1 y 2, respectivamente.

Tabla 7. Resultados del formulario NASA-TLX de la tarea 1.

Dimensión	Media	Interpretación Cualitativa
Exigencia Mental	5.31	La tarea requirió atención y algo de concentración.
Exigencia Física	3.85	Casi ningún esfuerzo físico. Interacción fluida.
Exigencia Temporal	4.31	Se requirieron concentración y cierto esfuerzo.
Esfuerzo	5.38	Se requirieron concentración y cierto esfuerzo.
Rendimiento	7.62	Se completó la tarea, aunque con esfuerzo.
Frustración	5.38	Algo de incomodidad o confusión.

Tabla 8. Resultados del formulario NASA-TLX de la tarea 2.

Dimensión	Media	Interpretación Cualitativa
Exigencia Mental	4.77	La tarea requirió atención y algo de concentración.
Exigencia Física	4.15	Interacción física moderada, algo repetitiva.
Exigencia Temporal	5.08	Algo de prisa o de espera, pero manejable.
Esfuerzo	4.08	Se requirieron concentración y cierto esfuerzo.
Rendimiento	5.69	Se completó la tarea, aunque con esfuerzo.
Frustración	4.85	Algo de incomodidad o confusión.

Al comparar estos resultados con las segmentaciones obtenidas a partir de los videos, puede afirmarse que existe una relación consistente entre los comentarios registrados y los eventos observados en las representaciones visuales. En las 13 pruebas realizadas, se identificó un total de 44 comentarios positivos, 33 negativos y 99 neutrales. Estos datos pueden consultarse en la tabla 9 y la figura 23, donde se contabilizan los comentarios encontrados en cada prueba de usuario. Los resultados sugieren que, si bien la interfaz del prototipo no fue percibida como perfecta, sí se evidenciaron errores y áreas de mejora señaladas por los usuarios. Un caso recurrente fue la confusión y/o molestia generada por el proceso de inicio de sesión. La mayoría de los participantes no encontraron el diseño confuso en sí, sino que manifestaron inconformidad por el tiempo excesivo de carga necesario para acceder a la página principal del sistema. Este tipo de hallazgos demuestra que las representaciones visuales generadas guardan una correspondencia adecuada con lo detectado en los formularios. Cabe señalar que, aunque se registró un número considerable de comentarios etiquetados como neutrales, esto no debe interpretarse como una limitación del sistema, sino como una oportunidad para el evaluador, ya que muchas de estas observaciones pueden no ser fácilmente clasificables mediante un modelo automático.

En este sentido, se reafirma la intención con la que se desarrolló el sistema: servir como herramienta de apoyo, no de reemplazo, para el trabajo del evaluador. La plataforma facilita el análisis, organiza la información y permite enfocarse rápidamente en los momentos clave, pero el juicio experto del evaluador es el que determina qué aspectos deben cambiarse, mejorarse o analizarse con mayor profundidad.

Tabla 9. Resultados del conteo de comentarios positivos, negativos y neutrales.

Usuario	POS	NEG	NEU	TOTAL
1	4	3	12	19
2	9	0	5	14
3	3	2	13	18
4	1	2	13	16
5	1	3	6	10
6	7	8	4	19
7	2	0	8	10

8	10	0	8	18
9	0	2	4	6
10	1	6	12	19
11	1	0	1	2
12	5	4	2	11
13	0	3	11	14
Promedio	3	3	8	14

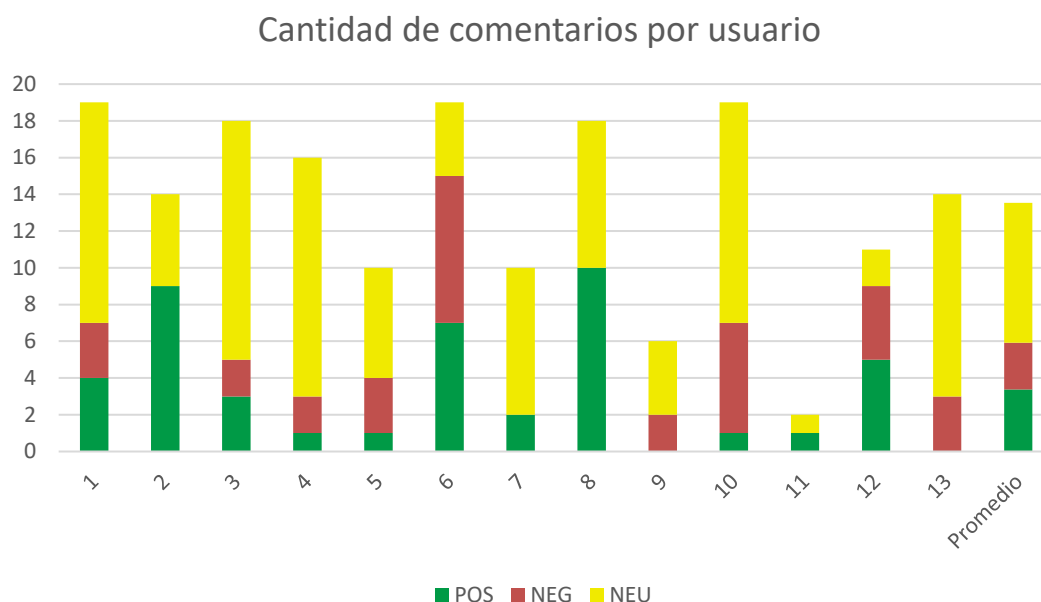


Figura 23. Gráfica de cantidad de comentarios de los usuarios.

6.5 Formulario de evaluación heurística

A continuación, se presentan todas las preguntas incluidas en el formulario de evaluación heurística, basado en el propuesto por Pierotti, D. (1995).

Objetivo:

Este instrumento se utilizó para analizar la interfaz desarrollada, complementándolo con capturas de pantalla que contextualizan visualmente cada heurística evaluada. Dichas imágenes facilitan la comprensión de los aspectos analizados y aportan evidencia concreta sobre el comportamiento de la herramienta en situaciones reales.

Es verdad que esta evaluación originalmente está enfocada más en las interfaces de las aplicaciones de venta, como menús, formularios y botones. Pero se realizó como un añadido para validar la usabilidad del reproductor, sin necesidad de ayuda externa de otros usuarios.

Observaciones:

De esta manera, se ofrece una representación visual que refuerza los hallazgos obtenidos mediante el cuestionario heurístico y contribuye a la validez del análisis. El formato empleado corresponde al checklist PEH_UAR.

6.5.1 Visibilidad del estado del sistema

En la primera tabla se aborda el apartado de visibilidad del estado del sistema, que establece que la aplicación debe mantener informados a los usuarios en todo momento sobre lo que ocurre. Para ello, se requiere proporcionar mensajes de retroalimentación pertinentes y en un plazo razonable. De esta manera, el usuario puede comprender con claridad qué sucede en la aplicación, véase en la tabla 10.

Tabla 10. Visibilidad del estado del sistema.

#	Lista de preguntas para evaluar	Sí	No	N/A	Comentarios
1.1	¿Cada pantalla comienza con un título o encabezado que describe los contenidos de la pantalla?	X			La pantalla principal donde se ve el video no cuenta como tal con un título, pero si encabezados de los apartados de las representaciones visuales. En las vistas de los comentarios y reporte general tienen tanto título como encabezados.
1.2	¿El esquema de diseño de los íconos y su estética es consistente en todo el sistema?	X			Se buscó que todos los iconos no salieran de la temática gris del reproductor actual.
1.3	Los menús de instrucciones, puntos de entrada de datos y mensajes de error, ¿aparecen en forma ordenada dentro de la pantalla?	X			
1.4	Si se utilizan ventanas emergentes (pop-up) para mostrar mensajes de error, ¿permiten estas ventanas que el usuario detecte el error que se menciona?	X			Se muestra el error que ocurre; por ejemplo, se indica cómo cargar el video si se trata de reproducir sin ningún video cargado previamente.
1.5	¿Hay algún tipo de retroalimentación (feedback) para cada acción u operación?	X			Sí, cuando se pasa por encima de los botones y se les da clic.
1.6	Después de que el usuario completa una acción o un grupo de acciones, ¿la retroalimentación (feedback) del sistema indica qué acciones pueden realizarse?	X			Sí, por ejemplo, el botón de guardado solo se activa cuando se carguen los datos que se pueden guardar.
1.7	¿El sistema provee algún tipo de retroalimentación (feedback) visual en menús o campos de texto que indiquen las opciones que pueden seleccionarse?	X			Hay un feedback de qué tipo de métrica se puede seleccionar y feedback visual sobre datos visuales en la línea de reproducción,
1.8	Si hay menús o campos de texto en donde pueden seleccionarse múltiples opciones, ¿el sistema provee algún tipo de retroalimentación visual que indique cuáles son las opciones que ya han sido seleccionadas?			X	No hay algo como selección múltiple, pero sí, al seleccionar la etiqueta de los comentarios, se actualiza para indicar el cambio de la etiqueta.
1.9	Luego de que el usuario complete una o un grupo de acciones, ¿El feedback del sistema indica que el siguiente grupo de acciones puede comenzarse?			X	Como no hay algo necesario que, por consecuencia, requiera una 3.ª acción, no se considera.
1.10	Si existen demoras mayores de 15 segundos en las respuestas del sistema, ¿el usuario es informado del progreso en la espera de la respuesta?			X	No hay demoras muy tardías en el sistema.
1.11	¿El sistema provee algún tipo de feedback visual en menús o cajas de diálogo que indique en cuál de las posibles opciones se halla posicionado el curso o el puntero del ratón?	X			Sí, se ve claro cuando el ratón está encima de algo con lo que puede interactuar.
1.12	¿Los tiempos de respuesta del sistema son adecuados al proceso cognitivo del usuario?	X			Cualquier otra acción, además de la carga del video, es inmediata.

1.13	¿Hay retroalimentación visual cuando los objetos son seleccionados o movidos?	X		Quando se selecciona alguna de las opciones, el botón con el que se activa adquiere un efecto de presión, dejando claro que está seleccionado.
1.14	¿Está claramente indicado el estado actual de un ícono, botón o vínculo?	X		
1.15	¿Hay indicaciones difíciles de entender dentro del sistema?	X		No hay indicaciones difíciles de entender.
1.16	¿Los tiempos de respuesta son apropiados para la tarea?	X		El tiempo de carga del video es óptimo; no es inmediato, pero no tarda un tiempo que pueda resultar incómodo.
1.17	Tareas simples y frecuentes: menos de 1 segundo	X		
1.18	Tareas comunes: 2 a 4 segundos	X		
1.19	Tareas complejas: 8 a 12 segundos	X		
1.20	Se requiere continuidad de pensamiento y la información debe recordarse a lo largo de varias respuestas: menos de dos segundos.		X	Como la reproducción del video y las representaciones visuales son directas, no aplica la continuidad.
1.21	Los altos niveles de concentración no son necesarios y no es necesario recordar la información: de dos a quince segundos.	X		
1.22	¿El sistema o aplicación proporciona visibilidad, es decir, al mirar, puede el usuario decir el estado del sistema y las alternativas de acción?	X		
1.23	¿Los menús, botones, vínculos o elementos virtuales hacen obvio qué elemento ha sido seleccionado?	X		
1.24	¿Los menús, botones, vínculos o elementos virtuales hacen obvio si la selección es posible?	X		
1.25	Si los usuarios deben navegar dentro del sistema, aplicación o ambiente virtual, ¿Este utiliza etiquetas que indiquen dónde se encuentra, mapas de menú o marcadores de lugar como ayuda de navegación?		X	Como no hay navegación profunda más que abrir otra ventana, no aplica.

Retomando algunos de los comentarios en este primer segmento, se procuró que la parte visual favoreciera un entendimiento claro, incluso para quienes no son expertos en la evaluación de la experiencia de usuario. El objetivo fue mantener una comprensión general de la utilidad y la función de cada elemento de la interfaz, de modo que pudiera interpretarse fácilmente a primera vista. En la figura 24 se aprecia que la mayoría de los componentes de la interfaz pueden comprenderse sin dificultad al primer vistazo.

Asimismo, se buscó dejar claro qué opciones están disponibles y cuáles no. Por ejemplo, los botones de Guardar y Análisis de parpadeo muestran un feedback visual que indica que, por el momento, no están habilitados.

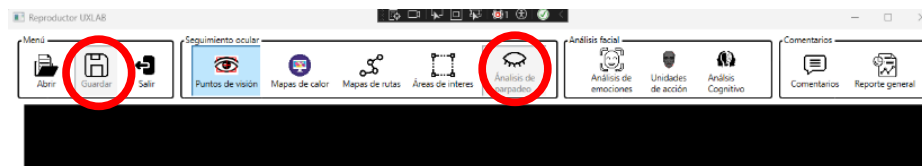


Figura 24. Primera vista del reproductor UXLab.

6.5.2 Relación entre el sistema y el mundo real

En el siguiente apartado se centra más en que el prototipo debe hablar el idioma de los usuarios, con palabras, frases y conceptos familiares, en lugar de términos técnicos. Tomando el mundo real para que la información aparezca en un orden natural y lógico, véase en la tabla 11.

Tabla 11. Relación entre el sistema y el mundo real.

#	Lista de preguntas para evaluar	Si	No	N/A	Comentarios
2.1	¿Los iconos, imágenes o elementos virtuales son concretos y familiares?	X			Sí, presentan una normativa que cualquiera podría identificar
2.2	¿Se ordenan las opciones de menú de la manera más lógica, mostrando al usuario los nombres de los elementos más representativos?	X			Se colocan primero los botones de funcionalidad y, posteriormente, los botones de las opciones de datos.
2.3	Si hay una secuencia natural de opciones de menú, ¿se han utilizado?	X			Como en otras interfaces, se colocan las funcionalidades generales en la primera parte del lado izquierdo y, posteriormente, todos los demás.
2.4	¿Los campos relacionados e interdependientes aparecen en la misma pantalla?	X			
2.5	Si los elementos se usan como una señal visual, ¿coinciden con las ideas culturales de los usuarios?	X			
2.6	¿Los colores seleccionados corresponden a las expectativas comunes sobre los códigos de color?	X			Sí, una perspectiva gris/clara, dejando de lado el icono del mapa de calor; es necesario que salgan un poco de esta norma.
2.7	Cuando las indicaciones implican una acción necesaria, ¿las palabras en el mensaje son consistentes con esa acción?	X			
2.8	En las secciones de captura de datos, ¿las tareas descritas son familiares para los usuarios?			X	No hay una sección de llenado de dato, así que no aplica
2.9	¿Se proporcionan avisos dentro de la sección de captura de datos?			X	No hay una sección de llenado de dato, así que no aplica
2.10	Para las secciones de formularios y cuestionarios, ¿las preguntas, enunciados o instrucciones se formulan en un lenguaje claro y sencillo?			X	No hay una sección de llenado de dato, así que no aplica
2.11	¿Las elecciones de menú encajan lógicamente en categorías que han comprendido fácilmente los significados?	X			
2.12	¿Los títulos de menú son entendibles gramaticalmente?	X			
2.13	¿El lenguaje del sistema o aplicación emplea palabras coloquiales del usuario y evita tecnicismos?			X	Como es una herramienta específicamente creada para la evaluación de los datos en una experiencia de usuario, se usan tecnicismos.
2.14	¿Se han evitado textos poco comprensibles?	X			
2.15	¿Los campos numéricos de dinero colocan algún signo de moneda y un decimal para las entradas monetarias?			X	No hay una sección de llenado de dato, así que no aplica
2.16	¿Los campos numéricos solo aceptan números?			X	No hay una sección de llenado de dato, así que no aplica

Al revisar apartados más concretos, puede observarse una correspondencia entre lo mostrado en la interfaz y el mundo real, reflejada en los iconos seleccionados. Algunos de estos ya estaban presentes en iteraciones anteriores de UXLab, pero fue necesario complementarlos con otros que mantuvieran la misma estética en tonos grises y claros.

No obstante, surge un pequeño reto: al tratarse de una herramienta especializada, resulta inevitable el uso de ciertos tecnicismos, lo que implica dejar en segundo plano algunos aspectos heurísticos. Esto puede apreciarse en la fila completa de iconos de los botones, como se muestra en la figura 25.



Figura 25. Botones tanto de las representaciones visuales como de las funcionalidades.

6.5.3 Control de usuario y libertad

Para el siguiente apartado, se enfoca en que marque claramente "salida de emergencia" para abandonar el estado no deseado sin tener que pasar por un diálogo prolongado. Soporte deshacer y rehacer, véase en la tabla 12.

Tabla 12. Control de usuario y libertad.

#	Lista de preguntas para evaluar	Si	No	N/A	Comentarios
3.1	Las acciones que se requieren en una sección de captura de información, ¿son particularmente fáciles de recordar?			X	No hay un apartado de captura de datos, así que no aplica.
3.2	En sistemas o aplicaciones que muestran varias ventanas, elementos virtuales o marcadores, ¿es fácil para los usuarios visualizar la pantalla?	X			
3.3	En sistemas o aplicaciones que utilizan opciones de paginación, navegación o sitios de una sola página, ¿es fácil para los usuarios cambiar entre ventanas?	X			
3.4	Cuando la tarea de un usuario se completa, ¿el sistema espera una confirmación del usuario antes de procesarla?	X			En el cambio de comentario se tiene un feedback completo.
3.5	¿Existe una función de "deshacer" en secciones de captura de información, importar archivos, crear elementos o acciones?			X	No hay un apartado de captura de datos, así que no aplica.
3.6	¿Pueden los usuarios cancelar operaciones fuera de operación en curso?			X	No hay un apartado de captura de datos, así que no aplica.
3.7	¿Pueden los usuarios reducir el tiempo de captura de datos al copiar y pegar información, elementos, imágenes o texto dentro del sistema o aplicación?			X	No hay un apartado de captura de datos, así que no aplica.
3.8	Si las listas de menú son largas (más de siete elementos), ¿pueden los usuarios seleccionar un elemento moviendo el cursor, el dedo o el control?			X	No hay un apartado de captura de datos, así que no aplica.
3.9	¿Los menús son amplios (muchos elementos en un menú) en lugar de profundos (muchos niveles de menú)?	X			
3.10	Si el sistema tiene múltiples secciones, ¿Es fácil navegar por sus secciones?	X			
3.11	¿Pueden los usuarios avanzar y retroceder mediante un botón, vínculo o elemento virtual?			X	Como no es un sistema donde haya ventanas completas con un nivel de profundidad más de 2 ventanas a la vez, no aplica.
3.12	Si el sistema, aplicación o ambiente virtual deja a los usuarios avanzar y retroceder entre las interfaces de usuario sin perder			X	No es un sistema con este tipo de aplicaciones, así que no aplica.

	información (sesiones, creación de elementos, niveles, cookies)?				
3.13	Si el sistema o aplicación contiene formularios, ¿los usuarios pueden volver a las preguntas anteriores o pasar a preguntas posteriores?			X	No hay un apartado de captura de datos, así que no aplica.
3.14	¿Pueden los usuarios revertir fácilmente sus acciones?			X	No hay un apartado de captura de datos, así que no aplica.
3.15	¿Pueden los usuarios establecer sus propios valores predeterminados (preferencias) al utilizar el sistema o aplicación (crear usuario, sesiones, información personal)?			X	No hay un apartado de captura de datos, así que no aplica.

Esta sección tiene una naturaleza ligeramente distinta del resto del análisis heurístico, ya que su enfoque principal está en la captura de datos, una función que no resulta central en el reproductor UXLab, cuyo propósito principal es servir como visualizador más que como recolector de información.

6.5.4 Consistencia y estándares

En esta parte se centra en que los usuarios no deben preguntarse si diferentes palabras, situaciones o acciones significan lo mismo. Debe tener un estándar para los símbolos, textos o imágenes que se utilicen en el prototipo, el sistema, el sitio web o la aplicación. Seguir las convenciones de la plataforma, véase en la tabla 13.

Tabla 13. Consistencia y estándares.

#	Lista de preguntas para evaluar	Si	No	N/A	Comentarios
4.1	¿Se han seguido los estándares de formato de la industria o de la empresa de manera consistente en todo el sistema o aplicación?	X			Se siguió la misma intención de reproductores como el de VLC o incluso reproductores no convencionales como el reproductor de software de Gazepoint, donde se coloca la grabación principal centrada y con mayor importancia.
4.2	¿Se ha evitado un uso intensivo de textos con colores fosforescentes?	X			
4.3	¿Se tiene alineación de textos en el sistema o aplicación?	X			
4.4	¿Los iconos tienen alguna descripción (texto abajo del icono, al sobrepasar el cursor)?	X			
4.5	¿Hay señales visuales para identificar dónde se encuentra el usuario?	X			
4.6	¿Cada sección tiene un título?	X			
4.7	¿Se han establecido los estándares de la industria o de la compañía para el diseño de interfaz y se aplican de manera coherente en todo el sistema o aplicación?	X			Sí, se siguieron los estándares de los reproductores de video de darle mayor importancia al video principal, el del escritorio, y se colocó la grabación del usuario en una esquina, siguiendo la tendencia en video de colocar la cámara donde se ve la cara de los usuarios en una esquina.
4.8	Si "salir" es una opción de menú, ¿siempre aparece en la parte inferior de la lista?		X		Se colocó en otra ubicación para acompañar a los botones del menú.
4.9	¿Los títulos del menú están centrados o justificados?	X			
4.10	¿Las instrucciones aparecen en una ubicación uniforme en todas las pantallas?	X			
4.11	¿Las etiquetas de texto y los campos de texto se distinguen tipográficamente (letra de títulos, subtítulos, negritas, color)?	X			

4.12	¿Las etiquetas de texto son consistentes en todo el sistema o aplicación?	X		
4.13	¿Las etiquetas de texto cuentan con tamaños adecuados y colores que las distingan?	X		
4.14	¿Se han evitado las combinaciones de colores de croma alta y colores espectrales extremos?	X		
4.15	¿Se evitan colores saturados para texto, iconos, fondos o elementos virtuales?	X		
4.16	¿La información más importante se coloca al comienzo del aviso?	X		
4.17	¿La información de los avisos proporciona más información que el encabezado del aviso?	X		Se coloca la problemática o la solución del problema.
4.18	Para las secciones de formularios, ¿se enumeran las entradas de cada pregunta o ítem?		X	No se usan formularios, así que no aplica.
4.19	¿Los nombres de opciones de menú son consistentes, tanto dentro de cada menú como en todo el sistema o aplicación, en términos gramaticales y terminológicos?	X		
4.20	¿La estructura de los nombres de cada opción del menú coincide con los títulos de las secciones?	X		
4.21	Si el sistema o aplicación cuenta con varias secciones de formularios, ¿estas tienen el mismo título?		X	No se usan formularios, así que no aplica.
4.22	¿Se usan colores de alto valor y alto cromatismo para atraer la atención?		X	En este caso, no, se decidió mantener un tema gris/claro en toda la interfaz para mantener una visualización fuera de irritación, se debe usar por mucho tiempo.

A partir de las observaciones y peticiones realizadas durante las evaluaciones tutoriales del proyecto, se decidió adoptar un diseño más cercano al de un reproductor de video convencional. Esto implicó dejar de lado la disposición de datos en formatos grandes, como gráficas que incluso llegaban a ocupar más espacio que los propios videos en las iteraciones anteriores.

En su lugar, se optó por una propuesta más moderna, centrada en destacar el video principal y presentar las herramientas de manera sutil y poco invasiva. La figura 26 muestra una comparación directa entre el reproductor actual y un reproductor de video estándar.

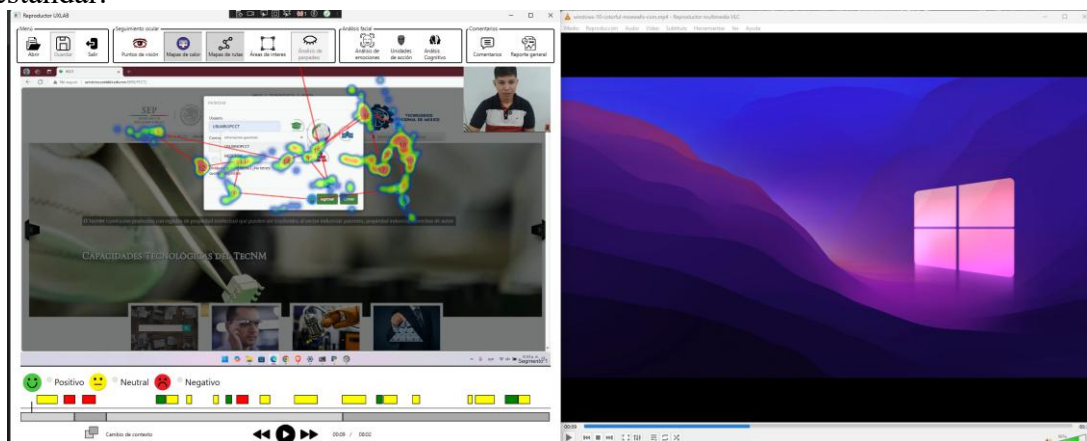


Figura 26. Comparación del reproductor actual de UXLab con un reproductor de video convencional.

6.5.5 Prevención de errores

Las aplicaciones deben contar con una buena arquitectura y una navegación fluida para evitar condiciones propensas a errores. Validar posibles errores y presentar a los usuarios una opción de confirmación antes de que se comprometan con la acción nos ayudará a evitar tener que utilizar tantos mensajes de error, véase en la tabla 14.

Tabla 14. Prevención de errores.

#	Lista de preguntas para evaluar	Si	No	N/A	Comentarios
5.1	¿Los mensajes de alerta se usan para señalar un error?	X			
5.2	¿Se formulan sugerencias de manera constructiva, sin críticas explícitas o implícitas al usuario?	X			
5.3	¿Las indicaciones implican que el usuario tiene el control?	X			
5.4	¿Son mensajes breves e inequívocos?	X			
5.5	¿Los mensajes de error están redactados para que el sistema, no el usuario, asuma la culpa?		X		No se han redactado de forma que se culpe al mismo sistema del error, pero tampoco para que se culpe al usuario.
5.6	Si se utilizan mensajes de error humorísticos, ¿son apropiados e inofensivos para la población de usuarios?			X	No se utilizan mensajes de error humorístico, por lo que no aplica.
5.7	¿Los mensajes de error son gramaticalmente correctos?	X			
5.8	¿Los mensajes de error evitan el uso de signos de exclamación?	X			
5.9	¿Los mensajes de error evitan el uso de palabras violentas u hostiles?	X			
5.10	¿Los mensajes de error evitan un tono antropomórfico?	X			
5.11	¿Todos los mensajes de error en el sistema usan un estilo gramatical, forma, terminología y abreviaturas consistentes?	X			
5.12	¿Los mensajes colocan a los usuarios en control del sistema?	X			
5.13	Si se detecta un error en un campo de texto, ¿el sistema coloca el cursor en ese campo o resalta el error?			X	No hay campos de texto, por lo que no aplica.
5.14	¿Los mensajes de error informan al usuario de la gravedad del error?	X			
5.15	¿Los mensajes de error sugieren la causa del problema?	X			
5.16	¿Los mensajes de error proporcionan información semántica y sintética apropiada?	X			
5.17	¿Los mensajes de error indican qué acción debe tomar el usuario para corregir el error?	X			
5.18	Si el sistema admite usuarios principiantes y expertos, ¿hay varios niveles de detalle de mensajes de error disponibles?	X			

Por su naturaleza como reproductor de datos, UXLab solo puede prevenir errores en la fase de carga de información, ya sea porque falte algún dato o se presente un error durante el proceso de carga. Un ejemplo de esta prevención puede verse en la figura 27, donde se muestra un mensaje de advertencia al intentar reproducir el video antes de cargar los datos a visualizar.

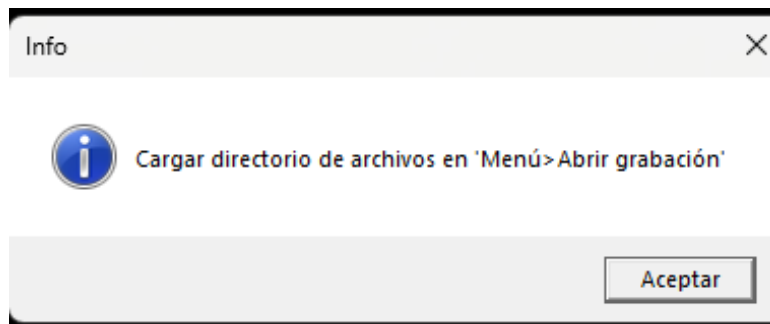


Figura 27. Mensaje de advertencia.

6.5.6 Reconocimiento en lugar de recordar

En este segmento se trata de minimizar la carga de memoria del usuario, haciendo visibles los objetos, acciones y opciones. El usuario no debe recordar información de una parte del diálogo a otra. Las instrucciones para el uso del sistema deben ser visibles o fáciles de recuperar cuando sea apropiado, véase en la tabla 15.

Tabla 15. Reconocimiento en lugar de recordar.

#	Lista de preguntas para evaluar	Si	No	N/A	Comentarios
6.1	Para las interfaces de pregunta y respuesta, ¿se utilizan señales visuales y espacios en blanco para distinguir preguntas, indicaciones, instrucciones y aportes del usuario?			X	No aplica en este caso.
6.2	¿Se visualiza en pantalla la captura de datos?			X	No hay captura de datos, por lo que no aplica.
6.3	¿Las etiquetas de texto de palabras múltiples se colocan horizontalmente (no apiladas verticalmente)?			X	No hay selección múltiple.
6.4	¿Se muestran todos los datos que un usuario necesita en cada paso de una secuencia de transacción?			X	Esto se refiere a los pasos relacionados con la captura de datos, por lo que no aplica.
6.5	¿Se colocan avisos, pistas y mensajes donde es probable que el usuario observe cierta sección de la pantalla?	X			
6.6	¿Las áreas de texto tienen espacio a su alrededor?	X			
6.7	¿Existe una distinción visual obvia entre "elegir uno" y "elegir muchos" en algún menú, opciones múltiples, respuesta o eliminar elementos?	X			
6.8	¿El sistema se vuelve oscuro en las secciones inactivas?		X		No se implementó esa parte dentro del sistema.
6.9	¿Se usa el espacio en blanco para crear simetría y dirigir la vista en la dirección apropiada?	X			
6.10	¿Los elementos se han agrupado en zonas lógicas y se han usado encabezados para distinguir entre secciones?	X			
6.11	¿Las secciones han estado separadas por espacios, líneas, colores, letras, títulos en negrita, líneas de reglas o áreas sombreadas?	X			
6.12	¿Las etiquetas de texto están cerca de las cajas de texto, pero separadas por al menos un espacio?	X			

6.13	¿Están los campos de entrada opcionales claramente marcados?	X		
6.14	¿Se utilizan divisiones para dividir secciones con mucha información?	X		
6.15	¿Se utilizan colores llamativos para atraer la atención del usuario?		X	Como parte de seguir el tema gris/claro, no se colocaron colores llamativos para atraer la atención.
6.16	¿Se utiliza algún color, sombra o línea en algún elemento que ha sido seleccionado?	X		
6.17	¿Se usan el tamaño, la negrita, el subrayado, el color, el sombreado o la tipografía para mostrar la cantidad relativa o la importancia de los diferentes elementos de la pantalla?		X	Se hace una distinción por la forma en que se agrupan, por lo que no.
6.18	¿Se ha usado el mismo color para agrupar elementos relacionados?	X		
6.19	¿La codificación de color es consistente en todo el sistema?	X		
6.20	¿Hay un buen contraste de color y brillo entre la imagen y los colores de fondo?	X		
6.21	¿Se han utilizado colores saturados, brillantes y luminosos para enfatizar los datos y se han utilizado colores más oscuros, más apagados y desaturados para restar importancia a los datos?	X		
6.22	¿Los elementos del menú inactivos están atenuados u omitidos?	X		
6.23	¿Hay valores predeterminados de selección de menú?	X		
6.24	Si el sistema tiene muchas opciones y algunas son restringidas, ¿los usuarios tienen forma de identificar que no pueden utilizar ciertas secciones?	X		
6.25	¿Ofrecen los menús una interacción con el usuario, es decir, hacer obvia donde la selección es posible?	X		
6.26	¿Hay señales visuales sobresalientes para identificar la ventana activa?	X		
6.27	¿Los formularios y los cuadros de diálogo indican cuándo los campos son opcionales?		X	No aplica en este caso.
6.28	En los formularios y cuadros de diálogo, ¿se muestran los campos dependientes solo cuando es necesario?	X		

Como parte de la misma interfaz, implementar poca necesidad de interacción y limitarse, en parte, a mostrar las representaciones visuales de los datos recopilados. El aspecto heurístico de mantener un diseño que se reconozca en vez de recordar no fue una prioridad y se puede ver en su evaluación.

6.5.7 Flexibilidad y diseño minimalista

Los atajos no vistos por el usuario novato a menudo pueden acelerar la interacción para el usuario experto de modo que el sistema pueda atender tanto a inexpertos como a usuarios experimentados. Permitir a los usuarios personalizar las acciones frecuentes es una manera de tener satisfechos a todo tipo de usuarios, véase en la tabla 16.

Tabla 16. Flexibilidad y diseño minimalista.

#	Lista de preguntas para evaluar	Sí	No	N/A	Comentarios
7.1	Si el sistema o aplicación admite tantos usuarios principiantes como expertos, ¿hay varios niveles de detalle de mensajes de error disponibles?	X			Dependiendo del error, puede indicar algo fácil como elegir los datos de forma correcta o indicar que un archivo en específico no está.
7.2	¿El sistema o aplicación permite a los principiantes usar una gramática sencilla y a los expertos usar una gramática posicional?	X			Al utilizar tanto lenguaje normal como técnico, se adapta a las necesidades y conocimientos de cada uno.
7.3	¿El sistema o aplicación permite a los usuarios ingresar información de manera sencilla y a los usuarios experimentados poder agregar parámetros si ellos lo desean al ingresar información?			X	No se toma en cuenta la parte de captura de datos, por lo que no aplica.
7.4	¿Los usuarios pueden guardar información sin complejidad?	X			Sí, solo con presionar el botón de guardar.
7.5	Si las listas de menú son cortas (siete elementos o menos), ¿pueden los usuarios seleccionar un elemento moviendo el cursor, el dedo o el control?			X	No hay dicho recurso, por lo que no aplica.
7.6	¿El sistema o aplicación usa atajos para ayudar a los usuarios expertos?		X		No hay atajos.
7.7	¿El sistema o aplicación tiene una buena navegación y un fácil entendimiento para los usuarios inexpertos?	X			Sí, al ser tan pocas las ventanas, se realizó con empeño para poder navegar con facilidad.
7.8	¿El sistema ofrece accesos directos "encontrar siguiente" y "encontrar anterior" para búsquedas?			X	No hay un sistema de brusquedad, por lo que no aplica.
7.9	En los sistemas o aplicaciones, ¿Los usuarios expertos tienen opciones para personalizar las tareas sin problemas?	X			Sí, la misma de mezclar las representaciones visuales haciendo una elección por el mismo.
7.10	¿Pueden los usuarios expertos omitir los cuadros de diálogo o algún atajo dentro de la aplicación o del sistema?			X	No hay como tal cuadros de diálogo, así que no aplica.

6.5.8 Diseño estético y minimalista

En la siguiente parte se centra en que los diálogos no deben contener información irrelevante o rara vez necesaria. Cada unidad adicional de información en un diálogo compite con las unidades de información relevantes y disminuye su visibilidad relativa, véase en la tabla 17.

Tabla 17. Diseño estético y minimalista.

#	Lista de preguntas para evaluar	Sí	No	N/A	Comentarios
8.1	¿Es solo (y toda) la información esencial para la toma de decisiones que se muestra en la pantalla?			X	No hay como tal una situación parecida a esa, por lo que no aplica.
8.2	¿Todos los iconos en un conjunto son visualmente y conceptualmente distintos?	X			
8.3	¿Se han usado objetos grandes, líneas en negrita y áreas simples para distinguir iconos?	X			
8.4	¿Cada icono se destaca de su fondo?		X		Solo unos pocos destacan
8.5	¿Los menús se apegan a la especificación siempre que sea posible?	X			
8.6	¿Los grupos significativos de artículos están separados por espacios en blanco?	X			

8.7	¿Cada interfaz de usuario tiene un título corto, simple, claro y distintivo?	X			
8.8	¿Las etiquetas de texto son breves, familiares y descriptivas?	X			
8.9	¿Los mensajes se expresan afirmativamente y utilizan la voz activa?	X			
8.10	¿Son breves los títulos de los menús, pero suficientes para comunicarse?	X			
8.11	¿Hay menús desplegables o desplegables dentro de los campos de entrada de datos que tienen muchas opciones de entrada bien definidas?			X	No hay campos de entrada, por lo que no aplica.

6.5.9 Ayudar a los usuarios a reconocer, diagnosticar y recuperar errores

Los mensajes de error deben expresarse en lenguaje sencillo (sin códigos, palabras técnicas o bloques de código) e indicar el problema y sugerir constructivamente una solución, véase en la tabla 18.

Tabla 18. Ayudar a los usuarios a reconocer, diagnosticar y recuperar errores.

#	Lista de preguntas para evaluar	Si	No	N/A	Comentarios
9.1	¿Se han usado etiquetas de texto para indicar la longitud del campo?			X	No hay campos de entrada, por lo que no aplica.
9.2	¿Las opciones de menú son lógicas, distintivas y mutuamente excluyentes?	X			
9.3	¿La información capturada es verificada al colocarla en los campos de texto?			X	No hay campos de entrada, por lo que no aplica.
9.4	Si el sistema muestra varias ventanas, la navegación entre ventanas es simple y visible	X			
9.5	¿Se utilizan validaciones antes de almacenar información?		X		
9.6	Si el sistema o aplicación cuenta con términos específicos, ¿se usan de manera uniforme y sencilla dentro de la interacción del sistema o aplicación?	X			
9.7	¿El sistema evita que los usuarios cometan errores siempre que sea posible?	X			En la carga de los datos, por ejemplo.
9.8	¿El sistema advierte a los usuarios si están a punto de cometer un error potencialmente grave?		X		No se implementó esa parte.
9.9	¿Los campos de texto y los cuadros de diálogo indican la cantidad de espacios de caracteres disponibles en un campo?			X	No hay campos de entrada, por lo que no aplica.
9.10	¿Los campos de texto y los cuadros de diálogo contienen valores predeterminados cuando corresponde?			X	No hay campos de entrada, por lo que no aplica.

6.5.10 Ayuda y documentación

El prototipo, sitio web, sistema o aplicación debe proporcionar ayuda al usuario y una sección de documentación que responda a cualquier duda sobre su funcionamiento. Debe tener información breve, fácil de buscar e interpretar, enfocada en las tareas del usuario, y presentada mediante una secuencia de pasos concretos para solucionar alguna situación, véase en la tabla 19.

Tabla 19. Ayuda y documentación.

#	Lista de preguntas para evaluar	Si	No	N/A	Comentarios
10.1	Si los usuarios están utilizando el sistema o aplicación con ayuda de un documento impreso, ¿Se explica mediante imágenes cómo utilizar el sistema o aplicación?	X			
10.2	¿Las instrucciones en línea son visualmente distintas?			X	Se están realizando cambios a la misma interfaz, por lo que es normal que lo reportado en documentos pasados no refleje el reproductor de hoy.
10.3	¿Las instrucciones siguen la secuencia de acciones del usuario?	X			
10.4	Si las opciones del sistema o aplicación son ambiguas, ¿Se proporciona información explicativa adicional cuando se selecciona un elemento?		X		No tendría mucho caso en un reproductor; eso fue lo que se pensó en su momento, pero puede que sea una buena modificación.
10.5	¿La explicación de la sección ayuda es sencilla y fácil de entender para cualquier tipo de usuario?		X		En este caso, no hay una sección “ayuda”, por lo que debería añadirse.
10.6	¿Existe una ayuda para utilizar “atajos”, ya sea a través de una referencia rápida en línea o solicitando?		X		
10.7	Existe una sección de ayuda que sea visible; ¿por ejemplo, un botón con el texto de AYUDA o un botón especial?		X		
10.8	¿La interfaz de la sección de ayuda (preguntas FAQ, comando de voz, contacto) es coherente con las interfaces del sistema o de la aplicación?			X	No hay dicha sección, por lo que no aplica.
10.9	Navegación: ¿la información es fácil de encontrar?	X			
10.10	Presentación: ¿El diseño visual está bien diseñado?	X			
10.11	Conversación: ¿Es la información precisa, completa y comprensible?	X			
10.12	¿La información es relevante?	X			
10.13	Objetivo orientado (¿Qué puedo hacer con este sistema o aplicación?)	X			
10.14	Descriptivo (¿Para qué sirve esto?)	X			
10.15	Procedural (¿Cómo hago esta tarea?)	X			
10.16	Interpretativo (¿Por qué sucedió eso?)	X			
10.17	Navegacional (¿Dónde estoy?)	X			
10.18	¿Hay ayuda contextual?		X		Falta agregar alguna ventana o apartado para aclarar y colocar explicaciones para los apartados que puedan no entenderse.
10.19	¿Puede el usuario cambiar el nivel de detalle disponible?		X		Está limitado a solo el nivel de detalle que tiene.
10.20	¿Pueden los usuarios cambiar fácilmente entre la ayuda y las demás secciones del sistema o de la aplicación?			X	No tiene una sección de ayuda dentro del reproductor. Por lo que no aplica.
10.21	¿Es fácil acceder a y regresar desde la sección de ayuda?			X	No tiene una sección de ayuda dentro del reproductor. Por lo que no aplica.
10.22	¿Pueden los usuarios reanudar el trabajo donde lo dejaron después de acceder a la ayuda?			X	Esta apartada está más relacionada con el aspecto de la captura de datos, por lo que no aplica en este caso.

6.5.11. Resultados de evaluación heurística final

Resultados:

Una vez realizada la evaluación heurística con los 10 formularios para obtener una puntuación y verificar si se cumplen o no las reglas heurísticas, se debe contabilizar el total de preguntas y el total de las respuestas que se respondieron de forma positiva o que no tuvieron una respuesta clara (N/A). Después, se podrá obtener el puntaje total de cada formulario. Véase en la tabla 20 la contabilización del total de preguntas en cada formulario (A) y del total de respuestas (B). Con los valores de la pregunta (C) se deben multiplicar por el número de respuestas (B) para obtener el total (D).

Tabla 20. Tabla de puntuación de la evaluación heurística.

# de regla	Total, de preguntas por regla (A)	Total, de respuestas positivas+N/A (B)	Valor por pregunta (C)	Total (D) B*C
1	25	25	4.0	100
2	16	15	6.2500	93.75
3	15	15	6.6664	99.996
4	22	21	4.5453	95.4513
5	18	17	5.5558	94.4486
6	28	25	3.5715	89.2875
7	10	9	10.0000	90
8	11	10	9.0911	90.911
9	10	8	10.0000	80
10	22	16	4.5455	72.728
			Total	906.5724

Total (C): 906.5724 /100 = 9.065724

Al sumar el total de cada formulario se obtiene un valor final, en este caso de 9.06. Después, gracias a la tabla de puntuación, se puede verificar si la evaluación cumple o no con las reglas heurísticas, véase en la tabla 21.

Tabla 21. Puntuación heurística.

Escala cuantitativa	Escala cualitativa
10 - 9	Cumple con las reglas heurísticas de forma excelente
7 - 8	Cumple con las reglas heurísticas de forma idónea
5 - 6	No cumple con todas las reglas heurísticas como se desea
≤ 4	No cumple con ninguna regla heurística de forma adecuada

Como se pudo observar a lo largo de la evaluación heurística, la interfaz obtuvo una puntuación final de 9.065724 (promedio), lo que refleja un cumplimiento excelente de los principios heurísticos establecidos. Esta calificación indica que la mayoría de los elementos evaluados presentan una implementación adecuada en términos de usabilidad, claridad, consistencia y diseño centrado en el usuario.

Dicho resultado respalda la calidad del diseño de la interfaz y demuestra que, desde una perspectiva experta, el sistema cumple con los lineamientos clave para ofrecer una experiencia de usuario sólida, intuitiva y funcional.

Capítulo 7: Conclusiones

El objetivo general establecido para este proyecto, “Diseñar, desarrollar y evaluar un reproductor de video que integre analítica visual con datos multimodales del proceso de evaluación de la experiencia de usuario, para apoyar la toma de decisiones, permitir la identificación y corrección de errores de diseño en el desarrollo de software”, se cumplió satisfactoriamente. Esto se evidencia en la correlación observada entre las respuestas de los cuestionarios posprueba y las observaciones obtenidas mediante el reproductor de video modificado de UXLab.

Además, el sistema demostró ser capaz de reducir entre un 52% y un 69% el tiempo de análisis al permitir que el evaluador revisara únicamente los segmentos de comentarios de los usuarios, en lugar de visualizar el video completo. De igual manera, se lograron mejoras significativas en relación con los dos principales problemas detectados en iteraciones previas dentro del estado del arte:

- Mayor precisión en los datos concretos, al ubicar con exactitud posibles problemas mediante la segmentación de comentarios y complementarlos con información de seguimiento ocular, lo que permite localizar con precisión el área de la interfaz en cuestión.
- Mejor organización de las herramientas visuales en el reproductor, priorizando el video de la interacción del usuario como elemento central, sin descuidar los demás datos relevantes recolectados durante la prueba.

Asimismo, se integraron opciones que brindan mayor control al evaluador, como la posibilidad de manipular el video del rostro del usuario y de activar simultáneamente una o varias analíticas visuales, lo que otorga flexibilidad y mejora la usabilidad del sistema.

De igual forma, los objetivos específicos definidos para la tesis fueron completados:

- Se modificó el reproductor de UXLab para lograr una representación más clara y enfocada del video de la interacción del usuario, reutilizando los algoritmos ya concretados en iteraciones anteriores y mejorándolos en gran medida.
- Se implementó una analítica visual que integra la información de seguimiento ocular con el video de la interacción, lo que permite utilizar más de una al mismo tiempo.
- Se implementó una analítica visual que integra la información del análisis facial con el video de la interacción, lo que permite utilizar más de una al mismo tiempo.
- Se incorporó una segmentación visual que dividiera el video para localizar comentarios positivos, negativos y neutros en el reproductor.
- Se evaluó un prototipo con el reproductor para validar su utilidad en la elaboración de reportes de experiencia de usuario.

Durante el desarrollo de la investigación, fue necesario realizar adecuaciones en los alcances y las limitaciones, e incluso en las modificaciones tanto del objetivo general como de los objetivos específicos. Esto se debió a un malentendido inicial entre los conceptos de “fusión visual” y “fusión de datos”. El segundo planteaba un nivel de complejidad considerablemente mayor que, debido al tiempo disponible y a la experiencia del investigador, no resultaba viable mantener dentro de los lineamientos del proyecto.

Así pues, la fusión visual se concretó mediante la superposición de las representaciones visuales en la interfaz del reproductor. Además, dado que gran parte del trabajo originalmente orientado a la fusión de datos tenía puntos de conexión con la perspectiva visual, la transición no supuso mayores dificultades. El principal reto consistió en un cambio de enfoque sobre cómo debía desarrollarse la investigación y en la incorporación de herramientas visuales en el reproductor.

La modificación en la forma de interactuar con el reproductor de UXLab representa un avance significativo para futuras iteraciones del sistema. Este rediseño abre la posibilidad de incorporar nuevas herramientas visuales, ajustar o mejorar las ya implementadas e incluso integrar aspectos automatizados para optimizar el proceso de generación de reportes. Además, este trabajo deja un precedente sobre cómo debería orientarse el diseño de este tipo de herramientas especializadas, colocando como foco principal el video de la interacción del usuario e integrando en torno a él el resto de las herramientas visuales. De esta manera, se asegura que puedan utilizarse libremente y de forma combinada según las necesidades y decisiones del evaluador.

7.1 Trabajos futuros

Como se ha planteado en diferentes partes de esta investigación, existen aspectos que, debido a limitaciones de tiempo, no fue posible desarrollar, pero que representan oportunidades valiosas para continuar el proyecto:

1. Ampliar las analíticas visuales

Explorar nuevas formas de análisis o mejorar las ya implementadas. Por ejemplo, utilizar el seguimiento ocular con una cámara facial convencional en lugar de una cámara especializada. Esto permitiría reducir costos y facilitar la aplicación en contextos más amplios.

2. Incorporar datos fusionados y su visualización

Aunque no se profundizó en este punto, se sabe que al combinar dos o más fuentes de información es posible obtener resultados más útiles para el evaluador, lo que le ayuda a identificar patrones con mayor rapidez y confiabilidad. Esta fusión podría implementarse tanto a nivel de datos como en representaciones visuales más integradas, superando lo logrado en esta investigación.

3. Integrar procesos de automatización en el reproductor

No se busca automatizar todo el proceso, pero sí ciertas tareas específicas que ahorren tiempo al evaluador. Por ejemplo, generar automáticamente un informe que recopile los comentarios negativos y los vincule con áreas concretas de la interfaz. Además, resulta necesario ampliar el repositorio de pruebas. Originalmente se contaba con 30 pruebas obtenidas con un prototipo anterior, pero al no existir un estándar de calidad en la recolección, la mayoría de esos datos resultaron incompletos o de baja calidad. Esto terminó afectando la generación de las representaciones visuales, por lo que gran parte de las grabaciones de las pruebas fue desechada. Así pues, al final solo se obtuvieron 13 pruebas, pero de buena calidad, lo que permite verificar con claridad la analítica visual implementada tanto para su evaluación como para su uso.

Referencias

- Armengol Castells, L. (2007). Los protocolos de pensamiento en voz alta como instrumento para analizar el proceso de escritura. *Revista Española de Lingüística Aplicada*, 20, 27–35.
- Bader, N., Mokryn, O., & Lanir, J. (2017). Exploring Emotions in Online Movie Reviews for Online Browsing. En *IUI 2017 – Companion of the 22nd International Conference on Intelligent User Interfaces* (pp. 35–38). Association for Computing Machinery. <https://doi.org/10.1145/3030024.3040982>
- Bakiu, E., & Guzman, E. (2017, septiembre). Which Feature is Unusable? Detecting Usability and User Experience Issues from User Reviews. En *Proceedings of the 2017 IEEE 25th International Requirements Engineering Conference Workshops (REW)* (pp. 182–187). <https://doi.org/10.1109/REW.2017.76>
- Baltrušaitis, T., Robinson, P., & Morency, L.-P. (2016). OpenFace: An open source facial behavior analysis toolkit. En *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1–10). IEEE. Disponible en <https://www.cl.cam.ac.uk/research/rainbow/projects/openface/wacv2016.pdf>
- Batch, A., Ji, Y., Fan, M., Zhao, J., & Elmqvist, N. (2024). uxSense: Supporting User Experience Analysis with Visualization and Computer Vision. *IEEE Transactions on Visualization and Computer Graphics*, 30(7), 3841–3856. <https://doi.org/10.1109/TVCG.2023.3241581>
- Blascheck, T., John, M., Kurzhals, K., Koch, S., & Ertl, T. (2016). VA²: A Visual Analytics Approach for Evaluating Visual Analytics Applications. *IEEE Transactions on Visualization and Computer Graphics*, 22(1), 61–70. <https://doi.org/10.1109/TVCG.2015.2467871>
- Bravo-Marquez, F., Frank, E., & Pfahringer, B. (2016). From Opinion Lexicons to Sentiment Classification of Tweets and Vice Versa: A Transfer Learning Approach. En *IEEE/WIC/ACM International Conference on Web Intelligence (WI)* (pp. 145–152). <https://doi.org/10.1109/WI.2016.29>
- Cao, N., Shi, C., Lin, S., Lu, J., Lin, Y.-R., & Lin, C.-Y. (2016). TargetVue: Visual Analysis of Anomalous User Behaviors in Online Communication Systems. *IEEE Transactions on Visualization and Computer Graphics*, 22(1), 280–289. <https://doi.org/10.1109/TVCG.2015.2467196>
- Christino, L., Rezaeipourfarsangi, S., Milios E. E., & Paulovich, F. V. (2022). A Theoretical Approach for Structuring and Analysing Knowledge Provenance for Visual Analytics.
- Cuenca, E., Sallaberry, A., Wang, F. Y., & Poncelet, P. (2018). MultiStream: A Multiresolution Streamgraph Approach to Explore Hierarchical Time Series. *IEEE Transactions on Visualization and Computer Graphics*, 24(12), 3160–3173. <https://doi.org/10.1109/TVCG.2018.2796591>
- da Silva Franco, R. Y., Santos do Amor Divino Lima, R., Monte Paixão, R. d., Resque dos Santos, C. G., & Serique Meiguins, B. (2019). UXmood—A Sentiment Analysis and Information Visualization Tool to Support the Evaluation of Usability and User Experience. *Information*, 10(12), 366. <https://doi.org/10.3390/info10120366>

- Fan, M., Wu, K., Zhao, J., Li, Y., Wei, W., & Truong, K. N. (2020). VisTA: Integrating Machine Intelligence with Visualization to Support the Investigation of Think-Aloud Sessions. *IEEE Transactions on Visualization and Computer Graphics*, 26(1), 343–352. <https://doi.org/10.1109/TVCG.2019.2934797>
- Fouilloux Quiroz, D. (2018). *Método para la integrar y sincronizar datos EEG y multimedia para su aplicación en el proceso de evaluación de la experiencia del usuario* (Tesis de maestría). CENIDET/TecNM.
- Gao, M., Do, H. J., & Fu, W. T. (2017). An Intelligent Interface for Organizing Online Opinions on Controversial Topics. En *Proceedings of the 22nd International Conference on Intelligent User Interfaces (IUI 2017)* (pp. 119–123). Association for Computing Machinery. <https://doi.org/10.1145/3025171.3025230>
- García Peñalvo, F. J. (2017, 16 de marzo). *Revisión sistemática de literatura en los Trabajos de Final de Máster y en las Tesis Doctorales*. Universidad de Salamanca. <https://repositorio.grial.eu/handle/grial/813>
- Goodwin, S., Prouzeau, A., Whitelock-Jones, R., Hurter, C., Lawrence, L., Afzal, U., & Dwyer, T. (2022). VETA: Visual Eye-Tracking Analytics for the Exploration of Gaze Patterns and Behaviours. *Visual Informatics*, 6(2), 1–13. <https://doi.org/10.1016/j.visinf.2022.02.004>
- Hernández Pérez, L. (2022). *Método para la detección de emociones mediante el análisis de expresiones facial en videos de evaluación de la experiencia de usuario* (Tesis de maestría). CENIDET/TecNM.
- Hussain, J., Khan, W. A., Hur, T., Bilal, H. S. M., Bang, J., Hassan, A. U., Afzal, M., & Lee, S. (2018). A Multimodal Deep Log-Based User Experience (UX) Platform for UX Evaluation. *Sensors*, 18(5), 1622. <https://doi.org/10.3390/s18051622>
- International Organization for Standardization. (1998). *ISO 9241-11: Guidance on usability*.
- International Organization for Standardization. (2010). *ISO 9241-210: Human-centred design for interactive systems*.
- Keim, D., Andrienko, G., Fekete, J. D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual Analytics: Definition, Process, and Challenges. En *Information Visualization: Human-Centered Issues and Perspectives (LNCS Vol. 4950)* (pp. 154–175). Springer. https://doi.org/10.1007/978-3-540-70956-5_7
- King, D. E. (2015). Max-margin object detection. *arXiv*. <https://arxiv.org/abs/1502.00046>
- Kucher, K., Martins, R. M., Paradis, C., & Kerren, A. (2020). StanceVis Prime: Visual Analysis of Sentiment and Stance in Social Media Texts. *Journal of Visualization*, 23(6), 1015–1034. <https://doi.org/10.1007/s12650-020-00684-5>
- Lagunes Ramírez, D. (2020). *Método para analizar el movimiento ocular...* (Tesis de maestría). CENIDET/TecNM.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool. <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
- Morales Morante, J. (2021). *Metodología para el preprocesamiento y clasificación de datos fisiológicos multimodales...* (Tesis de maestría). CENIDET/TecNM.
- Mu, S., Cui, M., & Huang, X. (2020). Multimodal Data Fusion in Learning Analytics: A Systematic Review. *Sensors*, 20(23), 1–27. <https://doi.org/10.3390/s20236856>

- Munim, K. M., Islam, I., Khatun, M., Karim, M. M., & Islam, M. N. (2017). Towards developing a tool for UX evaluation using facial expression. En *EICT 2017* (pp. 1–6). IEEE. <https://doi.org/10.1109/EICT.2017.8275227>
- Norman, D. A. (2013). *The Design of Everyday Things: Revised and Expanded Edition*. Basic Books.
- Park, D., Kim, S., Lee, J., Choo, J., Diakopoulos, N., & Elmqvist, N. (2018). ConceptVector: Text Visual Analytics via Interactive Lexicon Building Using Word Embedding. *IEEE TVCG*, 24(1), 361–370. <https://doi.org/10.1109/TVCG.2017.2744478>
- Romano Bergstrom, J., & Schall, A. (Eds.). (2014). *Eye Tracking in User Experience Design*. Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-408138-3.00001-4>
- Semeraro, A., Vilella, S., & Ruffo, G. (2021). PyPlutchik: Visualising and comparing emotion-annotated corpora. *PLOS ONE*, 16(9), e0256503. <https://doi.org/10.1371/journal.pone.0256503>
- Soure, E. J., Kuang, E., Fan, M., & Zhao, J. (2021). CoUX: collaborative visual analysis of think-aloud usability test videos for digital interfaces. *IEEE TVCG*, 28(1), 643–653. <https://doi.org/10.1109/TVCG.2021.3114822>
- Wang, F. Y., Sallaberry, A., Klein, K., Takatsuka, M., & Roche, M. (2015). SentiCompass: Interactive Visualization for Exploring and Comparing the Sentiments of Time-Varying Twitter Data. *PacificVis*, 129–133. <https://doi.org/10.1109/PacificVis.2015.7156368>
- Wang, X., He, J., Jin, Z., Yang, M., Wang, Y., & Qu, H. (2022). M2Lens: Visualizing and Explaining Multimodal Models for Sentiment Analysis. *IEEE TVCG*, 28(1), 802–812. <https://doi.org/10.1109/TVCG.2021.3114794>
- Zeng, H., Wang, X., Wu, A., Wang, Y., Li, Q., Endert, A., & Qu, H. (2020). EmoCo: Visual Analysis of Emotion Coherence in Presentation Videos. *IEEE TVCG*, 26(1), 927–937. <https://doi.org/10.1109/TVCG.2019.2934656>

ANEXO

Se recomienda descargar los documentos para su visualización real, debido a que se cambia un poco el formato y la forma por la transformación de los archivos Word al subirlos a Drive.

Anexo A

Anexo de la carta de consentimiento informado:

https://docs.google.com/document/d/166KnxU2ALx5gN_uJtxvBbaPaVU80a2qH/edit?usp=drive_link&oid=103913902180264131469&rtpof=true&sd=true

Anexo B

Anexo para el procedimiento para usuarios

https://docs.google.com/document/d/1zczLLcX5OLrDV41psL6xXu1sTAFTRlCv/edit?usp=drive_link&oid=103913902180264131469&rtpof=true&sd=true

Anexo C

Anexo para la guía para el facilitador

https://docs.google.com/document/d/1f2YHkkrKe9YTgupKAAhYv-4Smgcucn9f/edit?usp=drive_link&oid=103913902180264131469&rtpof=true&sd=true

Anexo D

Anexo del documento con los datos de las cuentas y datos de inicio de sesión

https://docs.google.com/document/d/1sB8cWPSNCEgB7GZzOudq-OSveBUstnRG/edit?usp=drive_link&oid=103913902180264131469&rtpof=true&sd=true