

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE HÁBITAT, INFRAESTRUCTURA Y
CREATIVIDAD

CARRERA DE INGENIERÍA EN SISTEMAS DE INFORMACIÓN



Trabajo de Titulación:

DESARROLLO DE UN PROTOTIPO FUNCIONAL DE UNA
EXTENSIÓN DE NAVEGADOR PARA INTERPRETAR POLÍTICAS DE
PRIVACIDAD MEDIANTE IA GENERATIVA

AUTOR:

Patricio Sebastián Martínez Flores

DIRECTOR:

Henry Nelson Roa Marin, PhD

QUITO DM, ENERO DE 2026

DEDICATORIA

Este trabajo está dedicado a mi familia, por el apoyo constante, la estabilidad y el acompañamiento brindado a lo largo de estos años de formación académica.

También está dedicado a Jhoel, cuyo apoyo hizo que este proceso resultara más llevadero y humano.

AGRADECIMIENTOS

Agradezco a la Pontificia Universidad Católica del Ecuador por la formación académica brindada a lo largo de mi carrera universitaria.

Expreso un especial agradecimiento a mis docentes, por la calidad de la enseñanza impartida, la cual ha contribuido de manera significativa a mi desarrollo académico y profesional.

Finalmente, agradezco a mi familia por el apoyo y la motivación brindados a lo largo de este proceso académico.

Tabla de contenido

| | |
|--|----|
| DEDICATORIA | 2 |
| AGRADECIMIENTOS | 3 |
| ÍNDICE DE FIGURAS | 6 |
| ÍNDICE DE TABLAS | 6 |
| 1. INTRODUCCIÓN | 7 |
| 1.1. JUSTIFICACIÓN..... | 7 |
| 1.2. PLANTEAMIENTO DEL PROBLEMA..... | 8 |
| 1.3. OBJETIVOS..... | 8 |
| 1.3.1. OBJETIVO GENERAL | 8 |
| 1.3.2. OBJETIVOS ESPECÍFICOS: | 9 |
| 1.4. ALCANCE | 9 |
| 1.5. MARCO TEÓRICO Y CONCEPTUAL..... | 10 |
| 1.5.1. ANTECEDENTES O MARCO REFERENCIAL | 10 |
| 1.5.2. MARCO TEÓRICO | 10 |
| 1.5.3. MARCO CONCEPTUAL | 11 |
| 2. REVISIÓN DE LITERATURA | 13 |
| 2.1. POLÍTICAS DE PRIVACIDAD Y CONSENTIMIENTO INFORMADO | 13 |
| 2.2. PROCESAMIENTO DE LENGUAJE NATURAL APLICADO A TEXTOS LEGALES..... | 15 |
| 2.3. INTELIGENCIA ARTIFICIAL GENERATIVA PARA RESUMEN Y EXPLICACIÓN..... | 17 |
| 2.4. METODOLOGÍA DE DESARROLLO DEL PROTOTIPO | 18 |
| 2.5. UX ÉTICA Y SISTEMAS DE APOYO A DECISIONES | 20 |
| 2.6. CONSIDERACIONES ÉTICAS Y LIMITACIONES | 21 |
| 3. DISEÑO Y DESARROLLO DE LA EXTENSIÓN DE NAVEGADOR | 24 |
| 3.1. ENFOQUE DE DISEÑO DEL SISTEMA..... | 24 |
| 3.2. REQUERIMIENTOS DEL SISTEMA | 25 |
| 3.2.1. REQUERIMIENTOS FUNCIONALES | 25 |
| 3.2.2. REQUERIMIENTOS NO FUNCIONALES..... | 26 |
| 3.3. CASOS DE USO DEL SISTEMA | 26 |
| 3.4. ARQUITECTURA GENERAL DEL SISTEMA..... | 28 |
| 3.5. COMPONENTES DE LA EXTENSIÓN DE NAVEGADOR..... | 31 |
| 3.5.1. CONTENT SCRIPT Y DETECCIÓN DE MOMENTOS DE CONSENTIMIENTO | 31 |

| | | |
|---------------|---|----|
| 3.5.2. | INTERFAZ DE USUARIO Y PRESENTACIÓN DE RESULTADOS | 33 |
| 3.5.3. | BACKGROUND SCRIPT Y GESTIÓN DE ESTADO | 37 |
| 3.6. | BACKEND Y PROCESAMIENTO DE LENGUAJE NATURAL | 38 |
| 3.6.1. | ARQUITECTURA DEL BACKEND | 38 |
| 3.6.2. | PIPELINE DE PROCESAMIENTO DE LENGUAJE NATURAL | 39 |
| 3.7. | MANEJO DE NAVEGACIÓN, CANCELACIONES Y CONSISTENCIA | 42 |
| 3.8. | CONSIDERACIONES ÉTICAS Y LIMITACIONES TÉCNICAS | 44 |
| 3.9. | RESUMEN DEL CAPÍTULO | 45 |
| 4. | EVALUACIÓN DEL PROTOTIPO | 47 |
| 4.1. | ENFOQUE DE EVALUACIÓN | 47 |
| 4.2. | DISEÑO DE LA EVALUACIÓN | 48 |
| 4.3. | CRITERIOS E INDICADORES DE EVALUACIÓN | 49 |
| 4.4. | RESULTADOS DE LA EVALUACIÓN | 50 |
| 4.4.1. | FASE 1: EVALUACIÓN DEL BACKEND Y DEL PIPELINE DE PROCESAMIENTO | 51 |
| 4.5. | ANÁLISIS Y DISCUSIÓN DE RESULTADOS | 55 |
| 4.6. | RESUMEN DEL CAPÍTULO | 57 |
| 5. | CONCLUSIONES Y RECOMENDACIONES | 59 |
| 5.1. | CONCLUSIONES | 59 |
| 5.2. | APORTES DEL PROYECTO | 60 |
| 5.3. | LIMITACIONES DEL PROYECTO | 61 |
| 5.4. | RECOMENDACIONES | 62 |
| 5.5. | CIERRE FINAL DEL TRABAJO | 64 |
| 6. | BIBLIOGRAFÍA | 65 |
| 7. | ANEXOS | 67 |
| 7.1. | ANEXO A: GUÍA DE PREGUNTAS UTILIZADAS EN LA EVALUACIÓN DEL PROTOTIPO | 67 |
| 7.2. | ANEXO B: SÍNTESIS DE RESULTADOS CUALITATIVOS DE LA EVALUACIÓN | 68 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| <i>Figura 1:</i> Diagrama de casos de uso del sistema | 27 |
| <i>Figura 2:</i> Vista de la Arquitectura del sistema..... | 29 |
| Figura 3: Overlay inyectado en la pagina web | 33 |
| Figura 4: Vista Inactiva del Popup | 34 |
| Figura 5: Vista del sistema en ejecución | 35 |
| Figura 6: Interfaz de Resultados del Sistema | 36 |
| Figura 7: Implementación del pipeline de procesamiento de lenguaje natural en el backend. | |
| Figura 8: Ejecución del pipeline de procesamiento de lenguaje natural y registro de etapas en consola. | 42 |

ÍNDICE DE TABLAS

| | |
|--|----|
| Tabla 1: Síntesis de resultados de evaluación del backend y pipeline | 69 |
| Tabla 2: Síntesis de resultados de evaluación de la interfaz y el overlay..... | 69 |
| Tabla 3: Síntesis de resultados de evaluación del sistema integrado | 70 |

1. INTRODUCCIÓN

1.1. JUSTIFICACIÓN

En la actualidad, cada usuario de internet se enfrenta diariamente a un desafío constante: comprender y aceptar políticas de privacidad extensas, técnicas y poco transparentes. Estas políticas, que deberían garantizar el uso responsable de la información personal, suelen estar redactadas en un lenguaje jurídico complejo que desincentiva su lectura y comprensión. Como resultado, millones de personas aceptan términos sin tener claridad sobre cómo se recopilan, procesan y comparten sus datos. Este fenómeno genera una problemática real y transversal que afecta a cualquier persona con acceso a servicios digitales, desde redes sociales hasta plataformas de comercio electrónico.

La relevancia de este tema radica en que el manejo inadecuado o la falta de comprensión de estas políticas abre la puerta a vulneraciones de derechos fundamentales, como la privacidad y la seguridad de la información. En un contexto en el que los datos se han convertido en uno de los recursos más valiosos, el consentimiento informado de los usuarios deja de ser una formalidad legal para convertirse en una necesidad ética y social. No obstante, el escenario actual refleja una clara asimetría entre las empresas que diseñan estas políticas y los individuos que las aceptan.

El presente proyecto busca contribuir a la reducción de esta brecha mediante el desarrollo de una extensión de navegador que utilice inteligencia artificial generativa para interpretar y resumir de forma clara y accesible las políticas de privacidad. De esta manera, se pretende brindar a los usuarios una herramienta que les permita tomar decisiones más informadas sobre el uso de sus datos. La propuesta se centra en un enfoque ético, al priorizar la transparencia y la usabilidad, y responde a la necesidad urgente de fomentar la responsabilidad digital en un entorno marcado por la sobreexposición de información personal.

El proyecto es factible debido a la disponibilidad de modelos de lenguaje natural avanzados, frameworks para extensiones web y metodologías de diseño centradas en el usuario.

Además, se alinea con tendencias actuales en materia de regulación de datos y ética en la inteligencia artificial, y aporta un prototipo funcional que, aunque limitado en alcance, busca demostrar la viabilidad de soluciones tecnológicas orientadas a proteger la privacidad de todos los usuarios de internet.

1.2. PLANTEAMIENTO DEL PROBLEMA

El uso masivo de servicios digitales ha convertido las políticas de privacidad en un elemento central para la protección de los datos personales. Sin embargo, en la práctica, estas políticas representan un obstáculo más que una garantía para los usuarios. La mayoría de los documentos se presentan en largos textos legales, redactados en un lenguaje técnico y poco accesible, lo que provoca que los usuarios acepten los términos sin leerlos ni entenderlos. Este comportamiento, ampliamente extendido, genera un consentimiento más formal que real, debilitando la transparencia y el control efectivo sobre la información personal.

La problemática se agrava en un entorno donde las empresas recopilan y procesan grandes volúmenes de datos, a menudo con fines que exceden lo que los usuarios imaginan. El desconocimiento de cláusulas sensibles —como la cesión de datos a terceros, el almacenamiento transfronterizo o el uso con fines publicitarios— pone en riesgo derechos fundamentales, incrementando la asimetría entre quienes diseñan estas políticas y quienes deben aceptarlas.

En este contexto, surge la necesidad de herramientas que acerquen estos textos a los usuarios de manera clara y comprensible. El desafío radica en transformar la complejidad legal en información útil y transparente, fomentando un consentimiento informado y real, y contribuyendo al uso responsable de los datos en internet.

1.3. OBJETIVOS

1.3.1. OBJETIVO GENERAL

Desarrollar un prototipo funcional de extensión de navegador que interprete y resuma políticas de privacidad mediante inteligencia artificial generativa, facilitando la comprensión de los usuarios y promoviendo el uso responsable de su información personal en internet.

1.3.2. OBJETIVOS ESPECÍFICOS:

1. Diseñar la arquitectura de una extensión de navegador que permita extraer y procesar automáticamente las políticas de privacidad de las páginas web.
2. Implementar un modelo de lenguaje generativo para generar resúmenes claros y accesibles de las políticas analizadas.
3. Evaluar la usabilidad y la comprensión del prototipo mediante pruebas con usuarios, identificando fortalezas y limitaciones en su aplicación práctica.

1.4. ALCANCE

El proyecto aborda el diseño y desarrollo de un prototipo funcional de extensión de navegador que utiliza inteligencia artificial generativa para interpretar y resumir políticas de privacidad. Su finalidad es mejorar la comprensión del usuario sobre cómo se gestionan sus datos personales en la web, promoviendo decisiones más informadas y un uso responsable de la información en entornos digitales. El prototipo debe extraer automáticamente políticas de privacidad de páginas web, procesarlas con modelos de lenguaje natural y generar resúmenes claros, acompañados de alertas sobre cláusulas sensibles. Además, se realizarán pruebas de usabilidad con usuarios no expertos para validar la comprensión, pertinencia y utilidad del sistema. El trabajo no contempla la creación de un producto comercial ni la implementación de mecanismos legales de validación. Tampoco incluye la cobertura de todas las variantes de políticas de privacidad existentes, lo que limita su aplicación a un entorno experimental diseñado para evaluar la viabilidad técnica y la experiencia de usuario del prototipo.

1.5. MARCO TEÓRICO Y CONCEPTUAL

1.5.1. ANTECEDENTES O MARCO REFERENCIAL

En los últimos años, distintos estudios han abordado el uso de técnicas de procesamiento de lenguaje natural (NLP) para analizar y simplificar políticas de privacidad. Investigaciones como “Natural Language Processing of Privacy Policies: A Survey” examinan cómo los modelos de lenguaje permiten extraer información relevante, resumir cláusulas y detectar riesgos asociados al tratamiento de datos personales (Harkous et al., 2020). De manera similar, el estudio “Beyond the Text: Analysis of Privacy Statements through Syntactic and Semantic Role Labeling” propone identificar los elementos semánticos más importantes en los textos legales, evidenciando las dificultades de comprensión que enfrentan los usuarios comunes (Ravichander et al., 2020). También, trabajos como “How to Keep Text Private? A Systematic Review of Deep Learning Methods for Privacy-Preserving Natural Language Processing” analizan métodos de aprendizaje profundo que protegen la confidencialidad durante el procesamiento textual (Klymenko & Yampolskiy, 2022).

Estos antecedentes reflejan una creciente preocupación por la accesibilidad de la información y la transparencia en el tratamiento de datos personales. Sin embargo, la mayoría de las soluciones se concentran en el ámbito académico y no llegan a los usuarios finales. Persiste, por tanto, la necesidad de herramientas aplicadas que acerquen estos avances tecnológicos a la vida cotidiana, lo que justifica el desarrollo de un prototipo accesible mediante una extensión de navegador.

1.5.2. MARCO TEÓRICO

El marco teórico de este proyecto se fundamenta en la convergencia entre la inteligencia artificial generativa, el procesamiento del lenguaje natural y los principios éticos de la privacidad digital. Los modelos de lenguaje como Ollama, basados en arquitecturas transformer, han demostrado una alta capacidad para comprender textos complejos y generar

resúmenes coherentes sin perder el sentido contextual. Este tipo de tecnologías constituye la base técnica para automatizar la interpretación de documentos extensos, como las políticas de privacidad.

De la misma manera, el procesamiento de lenguaje natural aplicado a textos legales permite segmentar y clasificar cláusulas según su relevancia o riesgo potencial. Estas técnicas, combinadas con enfoques de aprendizaje supervisado y semántico, facilitan la identificación de secciones sensibles relacionadas con el uso o la cesión de datos personales.

En el plano conceptual, este proyecto también se apoya en la teoría de la *integridad contextual* propuesta por Helen Nissenbaum, que entiende la privacidad como el flujo apropiado de información según las normas sociales y el contexto en el que se produce. Este enfoque respalda la idea de que los usuarios deben contar con información clara sobre cómo y para qué se utilizan sus datos. Por último, se consideran los principios de la privacidad diferencial, orientados a garantizar que los sistemas basados en IA no vulneren la confidencialidad de la información procesada durante su funcionamiento.

1.5.3. MARCO CONCEPTUAL

Para la presente investigación, se establecen los conceptos fundamentales que orientan el desarrollo del proyecto. Una extensión de navegador se define como un módulo complementario que amplía las funciones de un navegador web y permite interactuar con el contenido de los sitios visitados. En este caso, la extensión se utilizará para extraer, procesar y resumir las políticas de privacidad directamente desde las páginas consultadas.

La inteligencia artificial generativa hace referencia a un conjunto de modelos capaces de producir texto coherente y contextualizado a partir de datos de entrada mediante técnicas de aprendizaje profundo. Estos modelos, al aplicarse al procesamiento de lenguaje natural, permiten transformar grandes volúmenes de texto en información accesible y comprensible.

Una política de privacidad se entiende como un documento legal que explica los métodos mediante los cuales una organización recopila, almacena y utiliza los datos de sus usuarios. Dentro de estas políticas, las cláusulas sensibles son aquellas que implican riesgos potenciales para la privacidad, como la transferencia de información a terceros o el uso de datos con fines comerciales.

Finalmente, la privacidad diferencial es un enfoque que protege la identidad de los individuos mediante la adición de ruido estadístico a los datos, evitando que la información personal pueda ser reidentificada.

2. REVISIÓN DE LITERATURA

2.1. POLÍTICAS DE PRIVACIDAD Y CONSENTIMIENTO INFORMADO

En los entornos digitales, uno de los principales mecanismos con los que las organizaciones informan a sus usuarios sobre la recopilación, uso, almacenamiento y transferencia de sus datos personales es el aviso legal. Desde una perspectiva legal, estos documentos establecen las condiciones por las cuales se realiza el tratamiento de la información, además, técnicamente están escritos y dirigidos al usuario para que consienta sobre estas prácticas. Sin embargo, en la práctica, su efectividad como un instrumento de comunicación dirigido a un público general resulta limitada, especialmente para usuarios no técnicos.

Una de las principales dificultades asociadas con las políticas de privacidad y documentos similares radica en su extensión y complejidad. Los documentos, por diseño, están redactados en un lenguaje jurídico especializado. Tienen una estructura densa y referencias normativas que dificultan su lectura y comprensión. Como consecuencia, una gran parte de los usuarios acepta estas políticas de privacidad sin comprender plenamente su contenido, lo que convierte al consentimiento en una formalidad más que en el resultado de una decisión verdaderamente informada. Este fenómeno ha sido documentado en estudios empíricos recientes sobre privacidad digital y comportamiento de los usuarios, los cuales evidencian que incluso políticas de privacidad diseñadas para ser simples, prominentes y fácilmente accesibles pueden ser mal comprendidas por una proporción significativa de los usuarios (Korunovska et al., 2020). En este contexto, se observa una brecha considerable entre la intención normativa de las políticas de privacidad y su impacto real en la protección efectiva de los usuarios.

El concepto de consentimiento informado implica que el usuario comprende de manera clara los datos que se recopilan, la finalidad de su recolección, cuánto tiempo se almacenan y/o si se comparten con terceros. Sin embargo, en el contexto de la web actual, el consentimiento real rara vez se materializa de forma efectiva. En su lugar, prevalece un modelo en el que el usuario se ve presionado a aceptar rápidamente para acceder a un servicio.

A esta problemática se le suma el factor del momento en el que las políticas de privacidad son presentadas. Usualmente, aparecen en forma de banners, ventanas emergentes o enlaces secundarios durante la navegación en la página. Para ese momento, el usuario ya se encuentra enfocado en otra tarea dentro de la página, como acceder a un contenido o completar un registro. Entonces, la aceptación se convierte en un paso obligatorio para continuar, más que en una oportunidad real de reflexión. El problema, por tanto, no se limita únicamente al contenido de las políticas, sino también a cuándo se entregan y a la forma en que se integran dentro de la experiencia de usuario. Estudios recientes demuestran que el diseño y la arquitectura de elección de los banners de consentimiento influyen significativamente en las decisiones de los usuarios, generando aceptación automática incluso cuando esta no refleja sus preferencias reales de privacidad. (Bielova et al., 2024)

Para entender esta dinámica, resulta fundamental distinguir entre consentimiento formal e informado. Mientras que el consentimiento formal se otorga mediante la simple aceptación de un documento, el consentimiento informado requiere que el usuario comprenda y anticipe las consecuencias derivadas del tratamiento de sus datos personales. En la práctica, el consentimiento otorgado en muchos servicios digitales se aproxima más al primer tipo, ya que los usuarios tienden a aceptar las condiciones de forma rutinaria y con una conciencia limitada sobre las protecciones legales que se les ofrecen, priorizando la continuidad del uso del servicio sobre la evaluación del contenido presentado (Ahrari, 2024). Esta situación debilita la autonomía del usuario y limita su capacidad de tomar decisiones conscientes sobre su privacidad.

En este escenario, surge la necesidad de mecanismos que faciliten la interpretación de las políticas de privacidad. El desafío consiste en transformar información legal compleja en contenido comprensible y contextualizado, respetando la autonomía del individuo y evitando prácticas coercitivas. Esta necesidad sienta las bases para el uso de herramientas de apoyo que actúen como intermediarias informativas, orientadas a mejorar la comprensión en el momento adecuado, sin reemplazar el rol legal de las políticas ni la responsabilidad final del usuario.

Este proyecto se ha desarrollado dentro de este enfoque, al considerar las políticas de privacidad no solo como documentos legales, sino también como elementos de comunicación que deben ser interpretables por los usuarios a quienes van dirigidas. Desde esta perspectiva, la mejora del consentimiento informado no depende exclusivamente de

cambios normativos, sino también de soluciones tecnológicas que prioricen la claridad, el contexto y la experiencia del usuario, estableciendo así el fundamento conceptual para la incorporación de técnicas de procesamiento de lenguaje natural y sistemas de asistencia basados en inteligencia artificial.

2.2. PROCESAMIENTO DE LENGUAJE NATURAL APLICADO A TEXTOS LEGALES

El NLP (Natural Language Processing), procesamiento de lenguaje natural, constituye el área de la inteligencia artificial orientada a analizar textos del lenguaje humano. En los últimos años, su aplicación se ha extendido a especializaciones aisladas, como el análisis de documentos legales. En este contexto, el NLP se presenta como una herramienta potencial para apoyar la estructuración, organización e interpretación preliminar de textos jurídicos extensos, como las políticas de privacidad. (Katz et al., 2023)

Los textos legales presentan características propias que los diferencian de otros tipos de documentos y que suponen desafíos para el análisis automático. En primer lugar, suelen estar redactados en un lenguaje formal y técnico, con oraciones extensas y estructuras sintácticas complejas. Además, emplean términos con significados específicos dentro del ámbito jurídico, cuyo sentido puede variar según el contexto. Esta complejidad no responde a errores de redacción, sino a una necesidad deliberada de precisión y cobertura legal, lo que dificulta su procesamiento mediante técnicas lingüísticas generales. (Ariai, Mackenzie, & Demartini, 2025)

Otro desafío relevante radica en la ambigüedad presente en muchos textos legales. A diferencia de otros temas, donde la ambigüedad se considera un defecto, en el ámbito jurídico puede utilizarse como un mecanismo para permitir interpretaciones flexibles o para abarcar múltiples escenarios futuros. Para los sistemas de NLP, esta ambigüedad representa una limitación importante, ya que los modelos tienden a buscar una interpretación dominante del texto, lo cual puede no reflejar la complejidad real de las disposiciones legales. (Katz et al., 2023; Ariai et al., 2025)

A pesar de estas dificultades, diversas aplicaciones del procesamiento de lenguaje natural han demostrado ser útiles en el análisis de documentos legales. Entre las más comunes se

encuentran la clasificación de fragmentos según su contenido, la extracción de información relevante y la identificación de patrones recurrentes. Estas técnicas permiten transformar documentos largos y densos en representaciones más estructuradas, facilitando su exploración y análisis por parte de sistemas informáticos y usuarios humanos. (Ariai, Mackenzie, & Demartini, 2025)

En el contexto específico de las políticas de privacidad, se han desarrollado enfoques de resumen automático que buscan condensar el contenido de estos documentos en versiones más breves y accesibles. Sin embargo, estos resúmenes deben entenderse como aproximaciones informativas y no como interpretaciones legales profundas, debido a que el significado normativo de los textos originales no puede ser reemplazado completamente mediante procesos automáticos. (Ariai et al., 2025)

Las limitaciones del procesamiento del lenguaje natural en el ámbito legal son un aspecto central para considerar. Los modelos de NLP dependen de los datos con los que han sido entrenados y pueden reproducir sesgos o errores en dichos conjuntos. Además, carecen de una comprensión real del marco legal subyacente, lo que impide que evalúen la validez, la legalidad o la aplicabilidad de una cláusula específica. En consecuencia, el uso de NLP en este dominio requiere una limitación clara de su alcance y la supervisión humana.

Desde este punto de vista, el procesamiento de lenguaje natural debe verse como una herramienta de apoyo para la interpretación de información legal, y no como un sustituto de un análisis jurídico profesional. Su valor reside en su capacidad para reducir la carga cognitiva asociada a la lectura de documentos extensos y para destacar elementos relevantes que puedan orientar al usuario en la toma de decisiones informadas. Esta aproximación resulta especialmente pertinente en sistemas diseñados para asistir a usuarios no técnicos, donde la claridad y la contextualización de la información son prioritarias.

En resumen, el NLP aplicado a textos legales ofrece oportunidades para mejorar el acceso y la comprensión de documentos complejos, siempre que se conozcan explícitamente sus límites. Esta visión permite integrar técnicas de procesamiento automático del lenguaje dentro de sistemas de apoyo informativo, estableciendo una base teórica adecuada para la incorporación de modelos más avanzados de inteligencia artificial y para el desarrollo de soluciones orientadas a la transparencia y la experiencia del usuario.

2.3. INTELIGENCIA ARTIFICIAL GENERATIVA PARA RESUMEN Y EXPLICACIÓN

La inteligencia artificial generativa se ha estado observando como una evolución del procesamiento de lenguaje natural, orientada no solo al análisis de textos, sino también a la producción automática de contenido en lenguaje humano. Este enfoque se apoya en técnicas de generación de lenguaje natural (Natural Language Generation, NLG), cuyo objetivo es transformar representaciones internas de información en textos comprensibles para los usuarios. A diferencia de otros métodos de NLP centrados en tareas descriptivas, los sistemas generativos permiten producir resúmenes, reformulaciones y explicaciones adaptadas al contexto, manteniendo coherencia lingüística y fluidez discursiva. (Dong et al., 2022)

Uno de los usos más relevantes de la inteligencia artificial generativa es el resumen automático de documentos extensos. En este ámbito, la literatura distingue entre enfoques extractivos, que seleccionan fragmentos del texto original, y enfoques abstractivos, que generan nuevas formulaciones que condensan las ideas principales del documento. Estudios recientes han demostrado que los modelos basados en arquitecturas transformer son particularmente eficaces para el resumen de textos largos, especialmente cuando combinan técnicas extractivas preliminares con generación abstractiva posterior (Pilault et al., 2020). Esta combinación permite reducir la complejidad del texto de entrada y mejorar la coherencia del resultado generado.

En cuanto a documentos legales, como las políticas de privacidad, el resumen generativo se plantea como una herramienta para reducir la carga cognitiva asociada a la lectura de textos extensos y densos. Sin embargo, estos resúmenes no deben interpretarse como sustitutos del documento original, sino como representaciones informativas que facilitan una comprensión preliminar. Tal como señala la literatura, los sistemas que resumen texto no poseen una comprensión semántica profunda del contenido normativo, sino que generan texto en función de patrones aprendidos durante el entrenamiento del modelo (Pilault et al., 2020).

Además del resumen, los modelos generativos permiten elaborar explicaciones en lenguaje natural, reformulando conceptos técnicos en términos más accesibles para usuarios no especializados. Esta capacidad resulta especialmente valiosa en sistemas de apoyo a la toma de decisiones, donde el objetivo no es reemplazar el juicio humano, sino proporcionar

información contextualizada que oriente al usuario. No obstante, estas explicaciones deben entenderse como aproximaciones y no como interpretaciones expertas del contenido (Gatt & Kraemer, 2018).

El uso de modelos generativos también introduce riesgos significativos, siendo uno de los más relevantes las alucinaciones. Este término describe la generación de información que, aunque lingüísticamente plausible, no se encuentra respaldada por el texto original ni por fuentes verificables. Este comportamiento es inherente a los modelos de lenguaje de gran escala, que producen respuestas basadas en la continuación más probable de una secuencia de palabras, sin contar con mecanismos internos de validación factual (Azamfirei et al., 2023).

En dominios sensibles como el legal, las alucinaciones representan una limitación crítica. La literatura advierte que los modelos generativos no deben emplearse como autoridades interpretativas ni como sustitutos del análisis humano, ya que pueden producir explicaciones incorrectas o incompletas con apariencia de veracidad. Por esta razón, se enfatiza la necesidad de delimitar claramente el rol de la inteligencia artificial, incorporando supervisión humana y estableciendo expectativas realistas sobre su alcance (Azamfirei et al., 2023).

En esta perspectiva, la inteligencia artificial generativa debe concebirse como una herramienta de apoyo informativo, orientada a facilitar la comprensión de textos complejos sin reemplazar la lectura ni la responsabilidad del usuario. Su valor reside en la capacidad de resumir información y presentar explicaciones, siempre que se sepan explícitamente sus limitaciones. Esta visión fundamenta el uso de modelos generativos dentro de sistemas de asistencia, y establece el marco teórico para su aplicación responsable en el análisis de políticas de privacidad.

2.4. METODOLOGÍA DE DESARROLLO DEL PROTOTIPO

El desarrollo del prototipo se llevó a cabo mediante un enfoque secuencial incremental con validación continua, seleccionado en función de las características y objetivos del proyecto. Desde un inicio, el trabajo contó con un objetivo claramente definido y un alcance delimitado, orientado a la construcción de un prototipo funcional de una extensión de

navegador destinada a apoyar la comprensión de políticas de privacidad. Durante el desarrollo, el alcance del proyecto no sufrió modificaciones sustanciales, lo que permitió estructurar el trabajo de manera planificada y progresiva.

El proceso de desarrollo se organizó en fases secuenciales claramente identificables: análisis del problema, diseño del sistema, implementación, pruebas y cierre del prototipo. No obstante, a diferencia de un modelo secuencial rígido, la implementación se realizó de forma incremental, desarrollando el sistema por módulos funcionales. Cada módulo fue implementado y probado de manera independiente antes de avanzar al siguiente, permitiendo detectar y corregir errores de forma temprana y garantizar la estabilidad del sistema en cada etapa.

La fase de análisis se centró en la identificación del problema de investigación, la delimitación del alcance del sistema y la definición de los requisitos funcionales y no funcionales del prototipo. En esta etapa se establecieron los principios éticos y de experiencia de usuario que guiarían el desarrollo, tales como la no coerción, la transparencia y el respeto a la autonomía del usuario.

Posteriormente, en la fase de diseño, se definió la estructura general del sistema, el rol de cada componente y la forma en que interactuarían entre sí. Este diseño previo permitió orientar la implementación de manera coherente y evitar cambios estructurales en las etapas posteriores del desarrollo.

La implementación se realizó de manera modular, desarrollando progresivamente los distintos componentes del sistema. Cada módulo fue sometido a pruebas funcionales inmediatas con el objetivo de verificar su correcto funcionamiento antes de integrarlos con el resto del sistema. Este enfoque de validación continua permitió asegurar la consistencia del prototipo y reducir la propagación de errores entre componentes.

Las pruebas no se limitaron a una etapa final, sino que se integraron a lo largo del proceso de desarrollo. Además de las pruebas técnicas, se incorporaron observaciones puntuales y retroalimentación informal proporcionada por el docente tutor, las cuales permitieron realizar ajustes específicos sin alterar el alcance ni la estructura general del proyecto.

Finalmente, la fase de cierre del prototipo consistió en la consolidación del sistema, la verificación de su estabilidad y la preparación para su evaluación. Este enfoque

metodológico resultó adecuado para este proyecto de carácter académico, cuyo objetivo principal era obtener un prototipo funcional y coherente.

Entre las principales ventajas del enfoque adoptado se encuentra la claridad en la organización del proceso, el control del alcance y la posibilidad de validar cada componente de forma temprana. No obstante, una de sus limitaciones es la menor flexibilidad para incorporar cambios significativos en etapas avanzadas del desarrollo, lo cual se asumió conscientemente dado el carácter delimitado del proyecto.

2.5. UX ÉTICA Y SISTEMAS DE APOYO A DECISIONES

En el diseño de sistemas interactivos que presentan información relevante para la toma de decisiones del usuario, resulta fundamental diferenciar entre mecanismos de asistencia informativa y estrategias de persuasión o manipulación. En este contexto, los sistemas de apoyo a la toma de decisiones (Decision Support Systems, DSS) se caracterizan por ofrecer información estructurada con el objetivo de facilitar la comprensión del usuario, sin sustituir el juicio humano ni imponer una acción específica.

Diversos estudios en el ámbito del diseño centrado en el ser humano destacan que los sistemas de apoyo bien diseñados deben mantener al usuario como agente principal del proceso de decisión, promoviendo un enfoque human-in-the-loop. Bajo esta visión, la tecnología actúa como un facilitador cognitivo que reduce la carga informativa y resalta elementos relevantes, mientras que la responsabilidad final de la decisión permanece en el usuario. Este enfoque resulta especialmente pertinente en sistemas basados en inteligencia artificial, donde la automatización completa podría generar dependencia o desplazamiento de la autonomía humana (Sutton et al., 2020).

Desde una perspectiva ética, el diseño de sistemas que influyen en la toma de decisiones requiere principios claros de transparencia y delimitación funcional. La literatura sobre el diseño ético de sistemas de toma de decisiones basados en inteligencia artificial subraya que estos sistemas deben comunicar de forma comprensible cómo se genera la información que presentan, cuáles son sus limitaciones y qué alcance tiene su función dentro de la experiencia del usuario. La ausencia de estas consideraciones puede conducir a una sobreconfianza en el

sistema o a interpretaciones erróneas de su rol, especialmente en usuarios no técnicos (Floridi et al., 2022).

Un aspecto crítico en el diseño de experiencias de usuario relacionadas con decisiones sensibles es la distinción entre persuasión legítima y manipulación. Investigaciones en el ámbito de la interacción humano-computador han evidenciado que ciertas prácticas de diseño, aun cuando buscan mejorar la experiencia de usuario, pueden derivar en patrones manipulativos si reducen la capacidad de resistencia del usuario o explotan vulnerabilidades cognitivas. Cosas como ocultar de información relevante, la presentación falsa de opciones puede comprometer la autonomía del usuario, incluso cuando no existe una intención explícita de manipulación por parte del diseñador (Gray et al., 2023).

En contraste, una UX ética prioriza la autonomía del usuario a través de la presentación equilibrada de la información, la posibilidad de rechazo y la ausencia de coerción. La transparencia no se limita únicamente al contenido mostrado, sino también al momento y al contexto en que se presenta. En este sentido, el concepto de resistencia resulta central. Un diseño ético debe permitir que el usuario pueda ignorar, posponer o rechazar la información presentada sin penalización ni bloqueo funcional.

En el contexto de las políticas de privacidad, estos principios adquieren una relevancia particular. Dado que se trata de documentos legales complejos, el objetivo de un sistema de apoyo no debe ser dirigir al usuario hacia la aceptación o el rechazo, sino facilitar la comprensión de los elementos clave que afectan su privacidad. Desde esta perspectiva, la experiencia de usuario debe diseñarse como una herramienta informativa, que acompañe al usuario en el momento adecuado, sin sustituir la lectura del documento original ni asumir un rol normativo (Floridi et al., 2022).

Por lo tanto, una UX ética en sistemas de apoyo informativo se fundamenta en la claridad, la contextualización y el respeto a la autonomía del usuario. Estos principios constituyen la base conceptual para el diseño de soluciones tecnológicas orientadas a mejorar el consentimiento informado.

2.6. CONSIDERACIONES ÉTICAS Y LIMITACIONES

El desarrollo de sistemas asistidos por inteligencia artificial en un contexto sensible, como el análisis de políticas de privacidad, exige una delimitación clara de su alcance y de las responsabilidades asociadas a su uso. En este proyecto, la inteligencia artificial se concibe explícitamente como una herramienta de apoyo informativo, y no como un sustituto del análisis legal ni de la lectura del documento original. En consecuencia, el sistema no pretende interpretar normativamente las políticas de privacidad ni emitir juicios legales sobre su contenido.

Una de las principales consideraciones éticas de este enfoque es la preservación de la autonomía del usuario. El sistema no impone decisiones, no bloquea la navegación ni condiciona el acceso a los servicios en internet. Su función se limita a presentar información sintetizada y contextualizada que facilite la comprensión de los elementos más relevantes de una política de privacidad, manteniendo siempre la responsabilidad final de la decisión en el usuario.

Asimismo, se reconoce la posibilidad de falsos positivos en la detección y categorización de información. Dado que el análisis se basa en técnicas automáticas de procesamiento del lenguaje natural e inteligencia artificial generativa, es posible que ciertos fragmentos sean clasificados de manera imprecisa o que se destaquen elementos que no resulten críticos desde una perspectiva legal. Este riesgo se asume como un compromiso consciente, priorizando la visibilidad temprana de posibles riesgos.

En este sentido, el diseño del sistema prioriza la claridad y la comprensibilidad por encima de la exhaustividad técnica. El objetivo no es cubrir todas las cláusulas ni reproducir el contenido completo de la política de privacidad, sino ofrecer una síntesis accesible que permita al usuario identificar rápidamente los aspectos que pueden afectar su privacidad. Esta decisión responde a la evidencia presentada en capítulos anteriores sobre las limitaciones de los usuarios para procesar documentos extensos y altamente técnicos.

Otra limitación relevante de este sistema radica en la naturaleza de los modelos de lenguaje utilizados. Aunque se implementan mecanismos para reducir errores y mantener la coherencia de los resultados, no es posible garantizar la ausencia total de imprecisiones o simplificaciones excesivas. Por esta razón, los resultados generados deben interpretarse como orientativos y no como descripciones definitivas del contenido legal.

En conjunto, estas consideraciones éticas y limitaciones definen el marco de responsabilidad del sistema propuesto y garantizan que su uso se mantenga dentro de un enfoque de apoyo al consentimiento informado, respetando la autonomía del usuario y reconociendo las capacidades y restricciones inherentes a las tecnologías empleadas.

3. DISEÑO Y DESARROLLO DE LA EXTENSIÓN DE NAVEGADOR

3.1. ENFOQUE DE DISEÑO DEL SISTEMA

El diseño del sistema se envisionó como un asistente informativo, cuyo propósito principal es apoyar al usuario en la comprensión de políticas de privacidad en el momento en que se solicita su consentimiento. La idea central del sistema consiste en activarse cuando un sitio web presenta mensajes del tipo “al aceptar, aceptas nuestra política de privacidad o cookies”, ofreciendo al usuario información sintetizada sobre el contenido de dichos documentos antes de que tome una decisión.

El sistema fue diseñado teniendo como usuario objetivo al usuario común y casual, caracterizado por no poseer conocimientos técnicos o legales avanzados, pero mostrar interés o curiosidad por comprender qué ocurre con sus datos personales. Este perfil de usuario suele priorizar la rapidez y la facilidad de uso durante la navegación web, lo que implica que cualquier herramienta de apoyo debe integrarse de manera discreta y no intrusiva dentro de la experiencia de uso.

A partir de este perfil, se definieron principios de diseño orientados a preservar la autonomía del usuario y minimizar la interferencia con su navegación. Entre estos principios se incluyen la no coerción, evitando forzar decisiones o bloquear el acceso a contenidos; la transparencia, presentando información clara y comprensible sobre el contenido de las políticas; y la claridad, priorizando la presentación de evidencia relevante de manera estructurada y accesible.

Asimismo, el diseño del sistema evita asumir un rol autoritativo o normativo. La extensión no recomienda explícitamente aceptar o rechazar una política, ni emite juicios legales sobre su contenido. En su lugar, actúa como una capa informativa complementaria que presenta resúmenes y explicaciones en lenguaje natural, permitiendo que el usuario tome una decisión informada de manera autónoma. Como parte de la presentación de resultados, el sistema incorpora un indicador visual de riesgo expresado en niveles cualitativos (alto, medio, bajo). Este indicador tiene una función exclusivamente informativa y busca sintetizar la severidad de las prácticas detectadas para facilitar una comprensión rápida del contenido analizado. En

ningún caso representa una evaluación legal ni implica una recomendación explícita sobre la aceptación o rechazo de la política de privacidad, manteniendo la responsabilidad final de la decisión en el usuario.

Durante el diseño se asumieron ciertas consecuencias de forma consciente. En particular, se aceptó la posibilidad de falsos positivos en la detección y clasificación de información relevante, priorizando la visibilidad temprana de posibles riesgos sobre un análisis legal completo. De igual forma, se optó por privilegiar la experiencia de usuario y la claridad de la información frente a un análisis técnico exhaustivo, considerando que este último podría resultar contraproducente para el usuario objetivo.

En conjunto, el enfoque de diseño del sistema busca equilibrar la utilidad informativa con el respeto por la experiencia de usuario, estableciendo una base coherente con los principios éticos y conceptuales desarrollados en el Capítulo dos.

3.2. REQUERIMIENTOS DEL SISTEMA

3.2.1. REQUERIMIENTOS FUNCIONALES

El sistema debe cumplir con los siguientes requerimientos funcionales, definidos a partir de los objetivos del proyecto y del enfoque de diseño planteado:

1. Detectar automáticamente la presencia de mensajes de consentimiento relacionados con políticas de privacidad o uso de cookies durante la navegación web.
2. Extraer el texto completo de la política de privacidad asociada al sitio web, ya sea desde la página actual o desde enlaces relevantes.
3. Enviar el contenido extraído al backend para su procesamiento mediante modelos de procesamiento de lenguaje natural.
4. Generar un resumen explicativo del contenido de la política de privacidad en lenguaje natural, orientado a usuarios no técnicos.
5. Clasificar la información relevante de la política en categorías temáticas (por ejemplo, recopilación de datos, uso de datos, entrega a terceros, derechos del usuario).

6. Presentar los resultados del análisis mediante un popup informativo integrado en el navegador, incluyendo el resumen general y las categorías
7. Permitir al usuario confirmar la comprensión de la información o cancelar el análisis sin bloquear la navegación web.
8. Gestionar correctamente estados de cancelación, inactividad o cambio de página para evitar el cruce de información.

3.2.2. REQUERIMIENTOS NO FUNCIONALES

Además de los requerimientos funcionales, el sistema debe cumplir con una serie de requerimientos no funcionales que garantizan su calidad y alineación con los principios éticos del proyecto:

1. El sistema debe ser intuitivo y comprensible para usuarios no técnicos, evitando interfaces complejas o sobrecarga de información.
2. La extensión no debe bloquear la navegación ni interferir de forma agresiva con la experiencia del usuario.
3. La información presentada debe ser clara, comprensible y explicativa, evitando juicios normativos o recomendaciones explícitas.
4. El sistema no debe almacenar datos personales del usuario ni registrar información sensible durante el análisis.
5. La arquitectura debe permitir la sustitución o mejora de los modelos de lenguaje sin requerir cambios estructurales en la extensión.
6. El sistema debe preservar la autonomía del usuario, evitando mecanismos coercitivos o persuasivos en la presentación de la información.

3.3. CASOS DE USO DEL SISTEMA

Con el fin de describir las interacciones entre los actores y el sistema, se definieron los principales casos de uso del prototipo. Estos casos de uso permiten representar el

comportamiento funcional del sistema desde la perspectiva del usuario y de los sistemas externos con los que interactúa.

Dado que el sistema implementa una arquitectura distribuida, el análisis de casos de uso se centra en la extensión de navegador como componente principal, mientras que el backend de análisis se muestra como un actor externo que provee el procesamiento del lenguaje natural. De esta manera, se delimita claramente el alcance de la extensión y sus responsabilidades dentro del sistema completo.

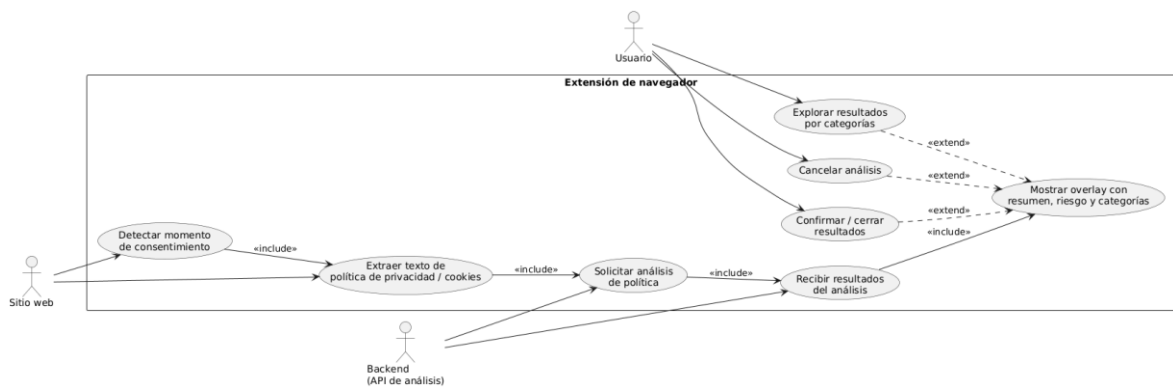


Figura 1: Diagrama de casos de uso del sistema

El diagrama representa los principales casos de uso de la extensión de navegador, incluyendo la detección de momentos de consentimiento, la extracción y análisis de políticas de privacidad, y la presentación de resultados al usuario mediante un overlay informativo. Asimismo, se modelan las interacciones opcionales del usuario con los resultados y la comunicación con el backend de análisis.

El flujo principal del sistema inicia cuando, durante la navegación web, se detecta un posible momento de consentimiento asociado a políticas de privacidad o uso de cookies. En ese contexto, la extensión presenta un overlay de confirmación, para informar al usuario que se ha identificado una situación relevante y solicitar su autorización para realizar el análisis. Este paso muestra el enfoque de diseño no coercivo del sistema, asegurando que el procesamiento de la política solo se haga con el consentimiento del usuario.

Si el usuario acepta el análisis, el sistema procede a extraer el texto de la política de privacidad asociada al sitio web y lo envía al backend para su procesamiento. Una vez recibido el resultado del análisis, la extensión muestra un segundo overlay informativo que presenta un resumen general, un indicador cualitativo de riesgo y las categorías temáticas identificadas en el documento. Desde este punto, el usuario puede explorar la información

de manera detallada, revisar categorías específicas o cerrar la visualización para continuar con su navegación.

En caso de que el usuario rechace el análisis desde el overlay de confirmación, el flujo se detiene y no se realiza ningún procesamiento adicional. Este comportamiento garantiza que el sistema actúe únicamente como un asistente informativo bajo control del usuario, evitando cualquier análisis no solicitado.

Finalmente, los casos de uso definidos se corresponden directamente con los requerimientos funcionales y no funcionales descritos en la sección anterior, garantizando la coherencia entre el diseño del sistema y su implementación. Esta alineación facilita la validación del prototipo.

3.4. ARQUITECTURA GENERAL DEL SISTEMA

El prototipo desarrollado se basa en una arquitectura modular de tipo cliente–servidor, diseñada para separar claramente las responsabilidades de detección, interacción con el usuario y procesamiento de información. Esta separación permite integrar la extensión de navegador de manera ligera y no intrusiva dentro de la experiencia de navegación, mientras que las tareas de análisis de texto y generación de resultados se delegan a un backend especializado.

De forma general, el sistema se compone de cinco elementos principales: la página web visitada por el usuario, una capa de detección y activación implementada mediante scripts inyectados en el navegador, la interfaz de usuario de la extensión, el servicio de fondo de la extensión y un backend local de procesamiento de lenguaje natural. Cada uno de estos componentes cumple un rol específico dentro del flujo del sistema y se comunica de forma controlada con los demás.

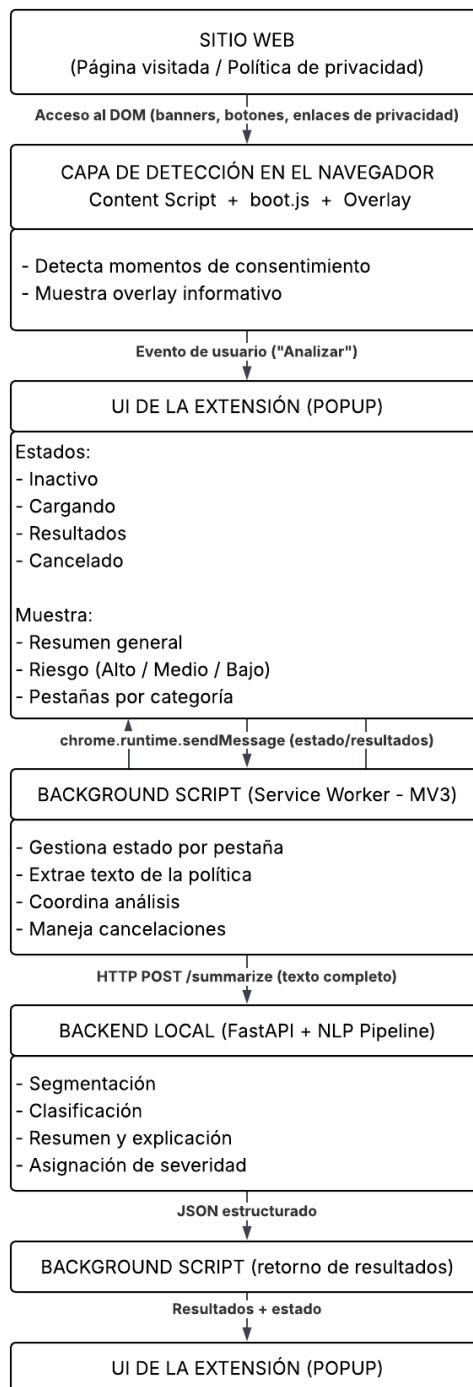


Figura 2: Vista de la Arquitectura del sistema

Presenta una vista general de la arquitectura del sistema y del flujo principal de información entre sus componentes.

La interacción inicia en la página web visitada por el usuario, donde el content script de la extensión accede al modelo de objetos del documento (DOM) con el objetivo de identificar posibles momentos de consentimiento asociados a políticas de privacidad o cookies. Esta lógica de detección se apoya en un script de arranque independiente (boot.js), encargado de evaluar el contexto de la página y desplegar un overlay informativo cuando se detecta una situación relevante. Este overlay actúa como un punto de entrada para el usuario, sin bloquear la navegación ni interferir de manera coercitiva con el contenido del sitio.

La interfaz principal de la extensión se implementa mediante un popup, desde el cual el usuario puede iniciar el análisis y visualizar los resultados generados. La interfaz gestiona distintos estados de interacción, como inactivo, cargando, resultados y cancelación, y se encarga de presentar la información de forma estructurada, incluyendo un resumen general, un indicador de riesgo y pestañas por categoría temática. El popup constituye el principal canal de comunicación entre el usuario y el sistema, manteniendo el control explícito sobre la ejecución del análisis.

El background script, ejecutado como un service worker bajo el estándar Manifest V3, cumple el rol de coordinador del sistema. Este componente mantiene el estado asociado a cada pestaña del navegador, gestiona el ciclo de vida del análisis y coordina la extracción del texto de la política de privacidad previamente localizada. Asimismo, es responsable de establecer la comunicación con el backend local y de gestionar la cancelación de procesos cuando el usuario navega, cambia de pestaña o interrumpe la operación.

El backend del sistema se implementa como un servicio local basado en FastAPI, exponiendo un endpoint principal (/summarize) que recibe el texto completo de la política de privacidad y devuelve un resultado estructurado en formato JSON. En este componente se ejecuta el pipeline de procesamiento de lenguaje natural, encargado de segmentar el contenido, clasificarlo por categorías y generar resúmenes explicativos en lenguaje natural. Para el alcance de este proyecto, el backend se ejecuta en un entorno local, lo cual resulta suficiente para un prototipo funcional; no obstante, su despliegue en infraestructura en la nube se identifica como una posible línea de trabajo futuro orientada a mejorar la portabilidad y escalabilidad del sistema.

El flujo general de información del sistema sigue una secuencia bien definida: detección del momento de consentimiento en la página, intervención informativa mediante el overlay,

inicio explícito del análisis por parte del usuario, extracción del texto de la política coordinada por el background script, envío del texto al backend para su procesamiento y presentación de los resultados en la interfaz de usuario. De forma complementaria, el sistema incorpora mecanismos de control de consistencia mediante la cancelación del análisis en curso cuando se detectan cambios en el contexto de navegación, evitando así la presentación de resultados desfasados o incorrectos.

En conjunto, esta arquitectura general permite materializar el enfoque asistivo del prototipo, garantizando una integración respetuosa con la experiencia de usuario, una clara separación de responsabilidades entre componentes y un flujo de información coherente con los principios éticos y de diseño establecidos en capítulos anteriores.

3.5. COMPONENTES DE LA EXTENSIÓN DE NAVEGADOR

3.5.1. CONTENT SCRIPT Y DETECCIÓN DE MOMENTOS DE CONSENTIMIENTO

El content script constituye el punto de entrada del sistema dentro de la página web visitada por el usuario. Su función principal es operar sobre el contenido del documento cargado en el navegador, permitiendo acceder al DOM con el fin de identificar elementos relacionados con políticas de privacidad y consentimiento. Este componente se ejecuta de forma automática al cargarse la página y actúa como la primera capa de análisis del contexto, sin requerir interacción previa del usuario.

La detección de momentos de consentimiento se basa en la identificación de patrones comunes utilizados por los sitios web para solicitar la aceptación de políticas de privacidad o cookies. Entre estos patrones se incluyen banners visibles, botones de aceptación, enlaces a documentos legales y textos que sugieren la recopilación o tratamiento de datos personales. Para ello, el content script analiza la estructura del DOM y busca elementos que coincidan con estos criterios, considerando tanto su contenido textual como su ubicación y visibilidad dentro de la página.

Con el objetivo de reducir falsos positivos y evitar activaciones innecesarias del sistema, la detección no se realiza de forma binaria. En su lugar, se implementa un sistema de

puntuación (scoring), mediante el cual distintos indicadores aportan valores parciales a una puntuación acumulada. Este enfoque permite ponderar múltiples señales —como la presencia de palabras clave, enlaces legales o botones de aceptación— antes de determinar si existe un momento de consentimiento relevante. Solo cuando la puntuación supera un umbral predefinido, el sistema considera que el contexto justifica la intervención informativa.

Una vez identificado un posible momento de consentimiento, el content script delega la presentación del overlay informativo a un script de arranque independiente (boot.js). Esta separación permite mantener la lógica de detección desacoplada de la lógica de presentación, permitiendo la modularidad y facilitando el mantenimiento del sistema. El overlay cumple una función informativa y actúa como un punto de entrada para el usuario, señalando la existencia de información relevante sin interrumpir ni bloquear la navegación normal del sitio.

El diseño del overlay se alinea con los principios éticos y de experiencia de usuario definidos en capítulos anteriores. En particular, se optó deliberadamente por un enfoque no bloqueante, evitando prácticas coercitivas como la desactivación del contenido, la superposición completa de la página o la exigencia de una acción obligatoria. El usuario conserva en todo momento el control sobre su navegación y puede decidir si interactúa o no con el sistema.

El timing de la activación del overlay constituye un aspecto central del diseño. La intervención ocurre únicamente cuando el sistema detecta un contexto asociado a la solicitud de consentimiento, es decir, en el momento en que la información resulta potencialmente relevante para la toma de decisiones del usuario. De este modo, el sistema evita mostrar información fuera de contexto o en etapas tempranas de la navegación, donde podría resultar irrelevante o intrusiva. Esta decisión busca maximizar la utilidad informativa del sistema y minimizar la carga cognitiva del usuario.

En conjunto, el content script y el mecanismo de detección implementado permiten identificar de manera controlada y contextualizada los momentos adecuados para la intervención del sistema. La combinación de análisis del DOM, un sistema de scoring y un overlay no bloqueante establece una base sólida para la activación responsable del asistente informativo, coherente con el objetivo del proyecto de apoyar al usuario sin interferir indebidamente en su experiencia de navegación.

3.5.2. INTERFAZ DE USUARIO Y PRESENTACIÓN DE RESULTADOS

La UI de la extensión se implementa mediante un overlay y un popup, el overlay es el primero a la vista del usuario. Este se activa a la hora de que el sistema detecte un momento de consentimiento de parte del usuario. Dira que ha encontrado un posible momento de consentimiento de forma no invasiva y le preguntara al usuario si desea que el sistema analice la política de privacidad.



Figura 3: Overlay inyectado en la pagina web

Ilustra el overlay inyectado en la página web cuando se detecta un posible momento de aceptación, destacando la acción identificada, el puntaje de confianza y los enlaces relevantes a políticas de privacidad o términos.

El popup actúa como el principal punto de interacción entre el usuario y el sistema. Este fue diseñado para ofrecer una experiencia clara y controlada con el propósito de asistir

del prototipo, evitando interferencias innecesarias con la navegación y manteniendo al usuario como agente principal del proceso.

La lógica de interacción del popup inicia a partir de un conjunto definido de estados de la interfaz, que reflejan de manera explícita el estado actual del análisis. En su estado inactivo, la interfaz no presenta resultados y comunica al usuario que el sistema se encuentra disponible para iniciar un nuevo análisis. Este estado corresponde a situaciones en las que no se ha detectado un momento de consentimiento relevante o cuando el usuario ha decidido limpiar el estado previo.

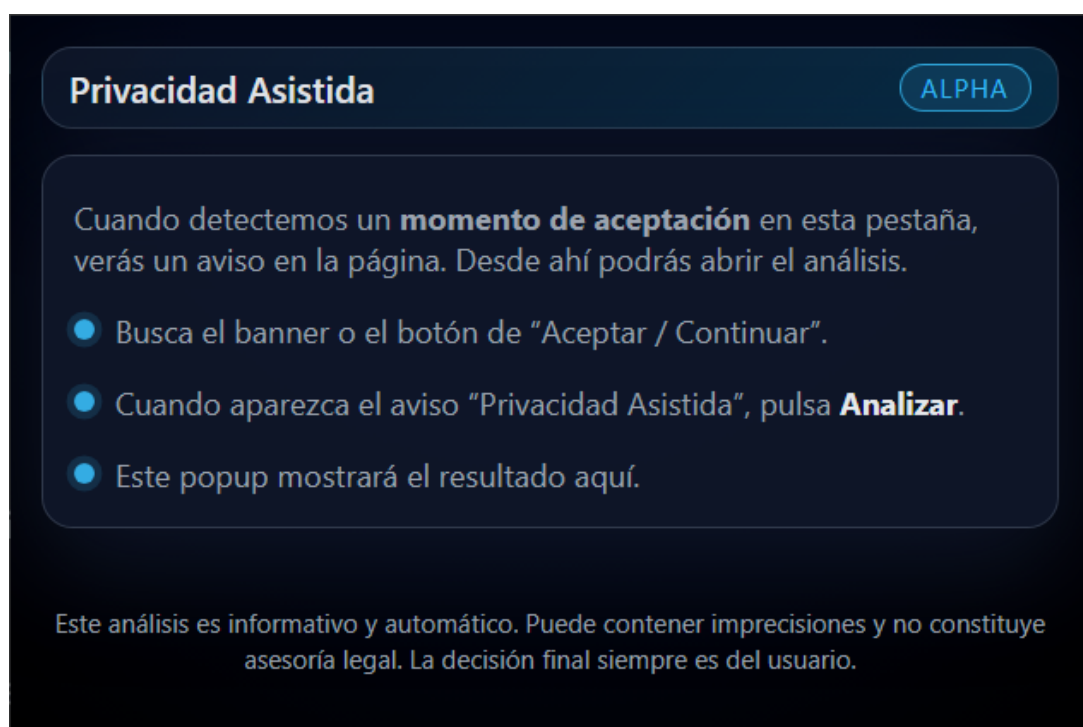


Figura 4: Vista Inactiva del Popup

Muestra la vista inactiva del popup de Privacidad Asistida, que se presenta cuando no se está ejecutando ningún análisis, explicando al usuario cómo funciona la detección automática y cómo iniciar el proceso.

Cuando el usuario inicia el análisis, la interfaz cambia al estado cargando, indicando que el sistema se encuentra procesando la información. Este estado cumple una función clave, ya que comunica que la solicitud ha sido recibida y que el análisis se está ejecutando, evitando que el usuario crea que haya ocurrido inactividad o un fallo del sistema. Durante esta fase,

el usuario conserva la posibilidad de interrumpir el proceso mediante la navegación o el cierre del popup.

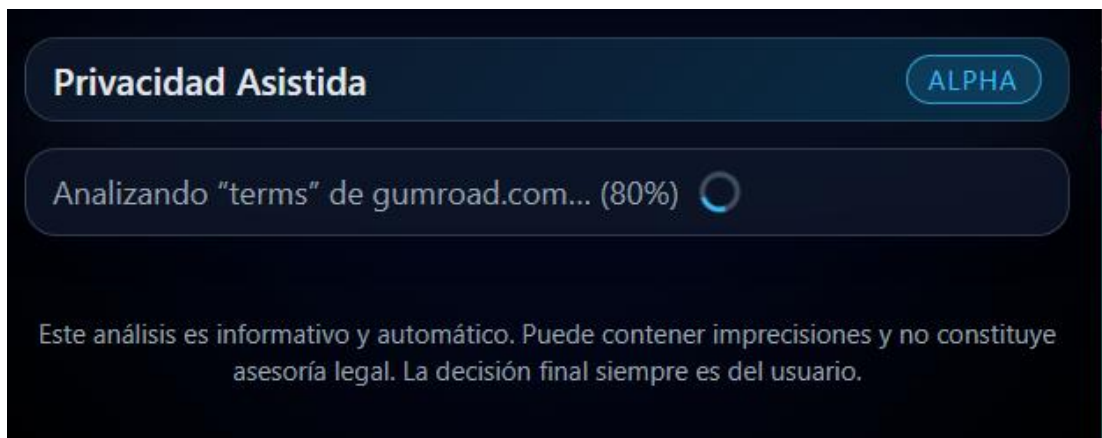


Figura 5: Vista del sistema en ejecución

Presenta el estado del sistema durante la ejecución del análisis, indicando el progreso del procesamiento de la política de privacidad y mostrando al usuario que el análisis se encuentra en curso.

Una vez finalizado el procesamiento, la interfaz entra en el estado resultados, en el cual se presenta la información generada por el backend de manera estructurada. La visualización de resultados se organiza en tres niveles principales. En primer lugar, se muestra un resumen general, cuyo objetivo es ofrecer una explicación sintética y comprensible del contenido de la política de privacidad analizada.

En segundo lugar, se presenta un indicador de riesgo global, expresado en niveles cualitativos (alto, medio o bajo). Este indicador cumple una función solamente informativa y actúa como un apoyo visual para sintetizar la severidad de las prácticas de tratamiento de datos detectadas. No representa una evaluación legal ni una recomendación explícita de acción, sino que busca facilitar una comprensión rápida del contexto general de la política analizada.

En tercer lugar, la interfaz organiza la información detallada en pestañas por categoría, permitiendo al usuario explorar aspectos específicos de la política de privacidad, como la recopilación de datos, el uso de la información, el intercambio con terceros o los derechos del usuario. Esta organización modular facilita la navegación por el contenido y permite al

usuario profundizar únicamente en los temas que considere relevantes, reforzando el control individual sobre la experiencia. Es importante también destacar que la interfaz también incluye la evidencia de cada uno de los fragmentos que referencia como alerta.



Figura 6: Interfaz de Resultados del Sistema

Muestra la interfaz del popup de resultados de Privacidad Asistida, donde se presenta el resumen general del análisis, el nivel de riesgo estimado y las categorías de información identificadas a partir de la política de privacidad del sitio.

El estado de cancelación se activa cuando el análisis se interrumpe debido a cambios en el contexto de la navegación, como el cambio de página, el cierre del popup o una acción explícita del usuario. En este estado, la interfaz comunica claramente que el proceso ha sido detenido y evita mostrar resultados parciales o desfasados. Este comportamiento contribuye

a mantener la consistencia del sistema y a prevenir confusiones derivadas de información no aplicable al contexto actual.

En conjunto, la interfaz de usuario del prototipo se concibe como una capa informativa complementaria, orientada a presentar resultados de manera clara, estructurada y respetuosa con la autonomía del usuario.

3.5.3. BACKGROUND SCRIPT Y GESTIÓN DE ESTADO

El background script de la extensión, implementado como un service worker bajo el estándar Manifest V3, cumple el rol de coordinador central del sistema. Su función principal es orquestar la comunicación entre los distintos componentes de la extensión, mantener el estado asociado a cada pestaña del navegador y gestionar el ciclo de vida del análisis de políticas de privacidad. Esta centralización de responsabilidades permite desacoplar la lógica de detección, la interfaz de usuario y el procesamiento de información, contribuyendo a una arquitectura más clara y mantenible.

Una de las responsabilidades clave del background script es la gestión del estado por pestaña. Dado que un usuario puede navegar simultáneamente por múltiples sitios o cambiar de contexto durante la ejecución del análisis, el sistema mantiene un estado independiente para cada pestaña activa. Este estado incluye información relevante como el identificador de la pestaña, el momento en que se inició el análisis y su estado actual (por ejemplo, inactivo, en progreso o finalizado). Este enfoque evita interferencias entre análisis concurrentes y garantiza que los resultados presentados correspondan siempre al contexto correcto.

El background script también hace de intermediario en la coordinación entre scripts. La comunicación entre el content script, el popup y el propio background se realiza mediante mecanismos de mensajería del navegador, permitiendo intercambiar eventos, solicitudes de análisis y actualizaciones de estado. Esta comunicación asíncrona facilita que cada componente opere de manera independiente, delegando en el background la responsabilidad de sincronizar el flujo general del sistema.

Otra función central de este componente es la apertura y gestión de la política de privacidad en segundo plano. Una vez que el usuario inicia el análisis, el background script se encarga

de localizar la URL de la política previamente identificada y de acceder a ella sin interrumpir la navegación principal del usuario. Esta operación se realiza de forma transparente, evitando redirecciones visibles o cambios inesperados en la pestaña activa, y permitiendo que la extracción del texto se ejecute de manera controlada.

El manejo del ciclo de vida del análisis constituye un aspecto crítico del diseño. El background script supervisa el inicio, la ejecución y la finalización del análisis, así como las posibles interrupciones causadas por cambios en el contexto de navegación. Si el usuario cambia de página, cierra el popup o navega a otra pestaña, el sistema detecta estas situaciones y cancela el análisis en curso. Este comportamiento previene la presentación de resultados inconsistentes o incorrectos.

En conjunto, el background script se encarga de la separación de responsabilidades dentro de la extensión. Mientras que el content script se enfoca en la detección del contexto y la interfaz se concentra en la interacción con el usuario, el background script asume la coordinación, la persistencia de estado y el control del flujo del sistema. Esta organización facilita el mantenimiento del prototipo, reduce acoplamientos innecesarios y refuerza la coherencia general del diseño.

3.6. BACKEND Y PROCESAMIENTO DE LENGUAJE NATURAL

El backend del sistema cumple el rol de procesamiento central, encargado de transformar el texto completo de una política de privacidad en una representación estructurada, comprensible y orientada al usuario final. Este componente opera de forma independiente a la extensión de navegador y se comunica con ella mediante una API, lo que permite mantener una clara separación entre la lógica de presentación y la lógica de análisis.

3.6.1. ARQUITECTURA DEL BACKEND

El backend fue implementado como un servicio web utilizando el framework FastAPI, seleccionado por su ligereza, soporte nativo para programación asíncrona y facilidad para definir contratos claros de entrada y salida mediante modelos tipados. Esta arquitectura

permite que el backend funcione como un servicio autónomo, accesible desde la extensión de navegador a través de solicitudes HTTP.

El sistema expone un endpoint principal (/summarize), cuya responsabilidad es recibir el texto completo de una política de privacidad, ejecutar el proceso de análisis y devolver un resultado que contenga un resumen general y un conjunto de alertas organizadas por categoría. El uso de un único endpoint centralizado simplifica la integración con la extensión y reduce la complejidad del flujo de comunicación.

El backend se ejecuta de forma local durante el desarrollo y la demostración del prototipo, sin embargo, su diseño permite una posible migración futura a un entorno en la nube sin muchas modificaciones. Esta posibilidad se puede considerar en el futuro, pero no forma parte del alcance de este proyecto.

En términos de responsabilidades, el backend no toma decisiones por el usuario ni evalúa la legalidad de las cláusulas analizadas. Su función se limita a procesar texto, identificar información relevante y presentarla de forma estructurada para facilitar la comprensión, respetando así los principios éticos definidos en capítulos anteriores.

3.6.2. PIPELINE DE PROCESAMIENTO DE LENGUAJE NATURAL

El núcleo funcional del backend es un pipeline de procesamiento de lenguaje natural, diseñado para analizar documentos extensos y jurídicamente complejos como las políticas de privacidad. Este pipeline sigue una estructura secuencial compuesta por varias etapas claramente diferenciadas, cada una con una responsabilidad específica.

La primera etapa corresponde a la extracción y normalización del texto. El contenido recibido se limpia y normaliza para poder hacer una buena segmentación, eliminando inconsistencias propias del documento original, como saltos de línea irregulares o caracteres especiales.

Después, el texto es segmentado en secciones más pequeñas, con esto podemos tratar cada fragmento como un pedazo semántico independiente. Segmentar facilita el análisis y reduce

la complejidad tanto para el sistema como para el usuario final, permitiendo referenciar fragmentos concretos del documento original.

Después de eso, se aplica una función de selección por palabras clave, el cual identifica las secciones más relevantes para cada categoría (por ejemplo, recopilación de datos, finalidad, derechos del usuario o terceros). Este filtrado permite reducir la cantidad de texto que debe ser procesado por modelos más costosos computacionalmente, optimizando el rendimiento del sistema.

A continuación, el pipeline usa técnicas de procesamiento asistido por los modelos de lenguaje para extraer los hechos relevantes a partir de los fragmentos seleccionados. Estos hechos se agrupan por categoría y sirven como base para generar las explicaciones dirigidas al usuario. Aquí, el sistema se limita estrictamente a la información presente en el texto original, evitando inferencias externas o interpretaciones legales.

En la *Figura 7* se presenta una captura de la implementación del pipeline en el backend, donde se observa la función orquestadora responsable de coordinar las distintas etapas del procesamiento, desde la segmentación inicial hasta la generación de los resultados finales.

```

# =====
# Orquestador
# =====

from pipelines.utils import validate_bullets_evidence, dedupe_bullets

async def run_full_pipeline(text: str) -> PipelineResult:
    sections = segment_text_into_sections(text)

    print(f"[pipeline] Total sections detected: {len(sections)}")

    # selector por keywords
    selected_sections = select_sections_by_keywords(
        sections, per_category_limit=6, min_total=10
    )

    print(
        f"[pipeline] Sections selected by keywords: {len(selected_sections)} "
        f"(per_category_limit=6, min_total=10)"
    )

    blocks = chunk_sections(selected_sections, max_chars=4000)

    print(f"[pipeline] Blocks created for LLM extraction: {len(blocks)}")

    all_facts: List[RawFact] = []

    for i, block in enumerate(blocks, start=1):
        print(f"\n[pipeline] Processing block {i}/{len(blocks)}")
        print(f"[pipeline] Block length: {len(block)} chars")

        facts = await extract_facts_from_block(block)

        print(f"[pipeline] Facts extracted from block {i}: {len(facts)}")

        for j, fact in enumerate(facts, start=1):
            print(
                f"[pipeline][fact] {j}. "
                f"category={fact.category} | "
            )

        all_facts.extend(facts)

    print(f"\n[pipeline] Total facts collected: {len(all_facts)}")

    # 1) Generación final (LLM)
    print("[pipeline] Generating final output from facts")
    result: PipelineResult = await generate_final_output(all_facts)

    # 2) Post-proceso robusto (barato)
    print("[pipeline] Validating evidence against full text")
    result.categories = validate_bullets_evidence(result.categories, sections)

    # 3) Quitar duplicados
    print("[pipeline] Deduplicating bullets")
    result.categories = dedupe_bullets(result.categories)

    # Resumen final
    print("\n[pipeline] Final category summary:")
    for cat, bullets in result.categories.items():
        print(f"[pipeline] {cat}: {len(bullets)} bullets")

    return result

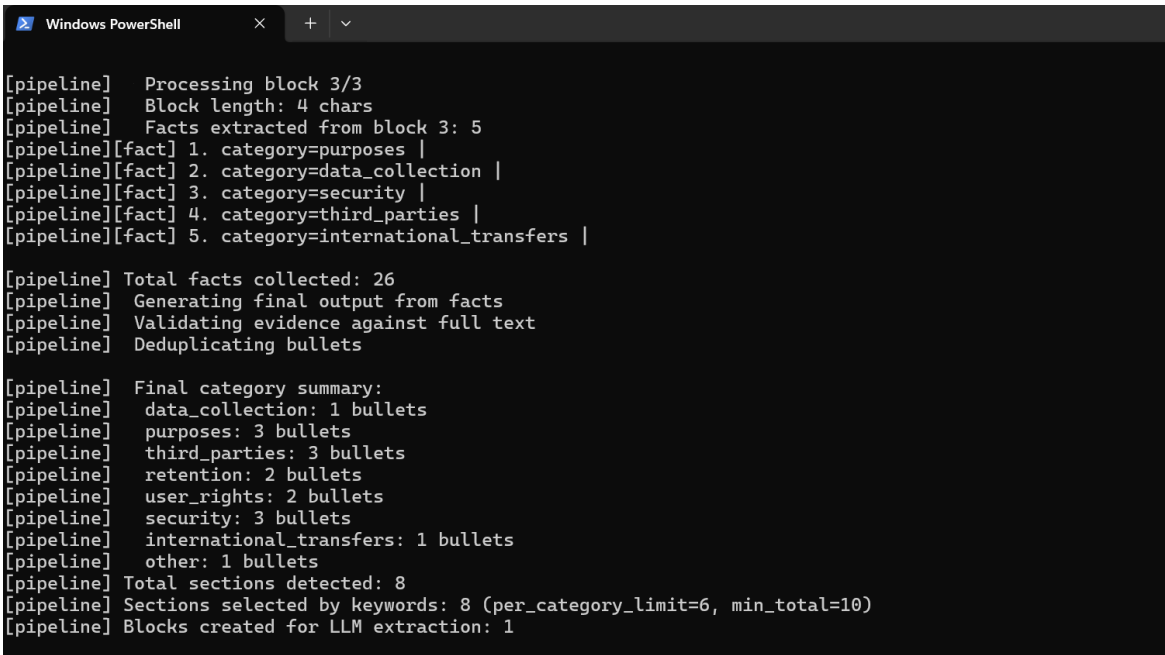
```

Figura 7: Implementación del pipeline de procesamiento de lenguaje natural en el backend.

Al final, el sistema genera un resumen general y los conjuntos de alertas, cada una acompañada de evidencia textual que referencia a la sección correspondiente del documento original. Estas alertas incluyen un nivel de severidad (bajo, medio o alto), que es utilizado únicamente de modo informativo para orientar al usuario, sin imponer juicios normativos.

Como medida adicional, el pipeline incorpora procesos de validación, que verifican que las citas utilizadas como evidencia estén presentes en el texto original y que no existan duplicaciones innecesarias. Estas validaciones refuerzan la consistencia del resultado y reducen el riesgo de errores o alucinaciones del modelo.

La *Figura 8* muestra la ejecución real del pipeline en consola, donde se evidencian las distintas etapas del procesamiento, incluyendo el análisis por bloques, la extracción de hechos, la validación de evidencia y la deduplicación de resultados.



```
Windows PowerShell
[pipeline] Processing block 3/3
[pipeline] Block length: 4 chars
[pipeline] Facts extracted from block 3: 5
[pipeline][fact] 1. category=purposes |
[pipeline][fact] 2. category=data_collection |
[pipeline][fact] 3. category=security |
[pipeline][fact] 4. category=third_parties |
[pipeline][fact] 5. category=international_transfers |

[pipeline] Total facts collected: 26
[pipeline] Generating final output from facts
[pipeline] Validating evidence against full text
[pipeline] Deduplicating bullets

[pipeline] Final category summary:
[pipeline] data_collection: 1 bullets
[pipeline] purposes: 3 bullets
[pipeline] third_parties: 3 bullets
[pipeline] retention: 2 bullets
[pipeline] user_rights: 2 bullets
[pipeline] security: 3 bullets
[pipeline] international_transfers: 1 bullets
[pipeline] other: 1 bullets
[pipeline] Total sections detected: 8
[pipeline] Sections selected by keywords: 8 (per_category_limit=6, min_total=10)
[pipeline] Blocks created for LLM extraction: 1
```

Figura 8: Ejecución del pipeline de procesamiento de lenguaje natural y registro de etapas en consola.

3.7. MANEJO DE NAVEGACIÓN, CANCELACIONES Y CONSISTENCIA

El manejo adecuado de la navegación y de los estados intermedios del sistema son un aspecto central del diseño del prototipo, tanto desde el punto de vista técnico como desde una perspectiva ética y de experiencia de usuario. Dado que el sistema se activa en un contexto dinámico, mientras el usuario navega en internet, es importante garantizar que los resultados presentados al usuario sean coherentes con el contexto y no correspondan a estados no válidos.

Uno de los principales escenarios considerados durante el desarrollo fue el cambio de página mientras el análisis se encuentra en ejecución. En estos casos, el contenido de la política de privacidad previamente detectada deja de ser relevante, ya que el usuario ha abandonado el contexto original. Esto puede significar que el usuario ha aceptado ya la política, o que ya no está interesado en aceptar. Para evitar que el sistema presente resultados asociados a una página distinta o ya no visible, el análisis se cancela automáticamente. Esta decisión impide que aparezca información potencialmente confusa y refuerza la consistencia del comportamiento del sistema.

Desde una perspectiva técnica, la cancelación del análisis se implementa mediante mecanismos de control de solicitudes asíncronas, permitiendo abortar de forma segura la comunicación entre la extensión y el backend. El uso de estos mecanismos asegura que las solicitudes en curso no sigan consumiendo recursos innecesarios ni produzcan respuestas que ya no apliquen al contexto del usuario. Asimismo, evita que la interfaz muestre resultados parciales o desactualizados.

La prevención de resultados inconsistentes constituye un objetivo importante del diseño. El sistema evita deliberadamente almacenar resultados persistentes entre sesiones o reutilizar análisis previos en contextos distintos. Cada análisis se ejecuta como una operación independiente, asociada a una página específica y a una interacción explícita del usuario. Así, se reduce el riesgo de que la información pierda validez o se interprete fuera de contexto.

Estas decisiones fueron consideradas para una buena experiencia de usuario y ética del diseño. Presentar información incorrecta o incompleta en un contexto de privacidad podría inducir a interpretaciones erróneas o generar una falsa sensación de seguridad. Al cancelar el análisis cuando el contexto cambia, el sistema prioriza que la información sea viable antes de una continuidad forzada del proceso.

Viéndolo desde un enfoque ético, este enfoque refuerza el carácter del sistema como un asistente responsable. Este prototipo no busca maximizar la cantidad de resultados mostrados ni imponer información al usuario, sino garantizar que aquello que se presenta sea pertinente, comprensible y en un contexto apropiado. La posibilidad de cancelar el análisis de forma implícita mediante la navegación, así como que no haya bloqueos o advertencias intrusivas, preserva la autonomía del usuario y respeta su control.

3.8. CONSIDERACIONES ÉTICAS Y LIMITACIONES TÉCNICAS

El diseño del prototipo incorpora de manera explícita una serie de consideraciones éticas y técnicas que delimitan el alcance real del sistema y previenen interpretaciones erróneas sobre su propósito. Estas consideraciones no se entienden como deficiencias del proyecto, sino como decisiones conscientes que responden a la naturaleza del problema abordado y que podrían mejorarse con una actualización futura.

En primer lugar, es fundamental establecer el rol no autoritativo de la inteligencia artificial empleada en el sistema. La IA utilizada no interpreta el documento desde una perspectiva legal, ni evalúa la validez jurídica de las cláusulas analizadas. Tampoco recomienda al usuario aceptar o rechazar una política de privacidad. Su función es ser una herramienta de apoyo que ayude a resumir el contenido del documento en un lenguaje comprensible para el usuario y a señalar posibles aspectos relevantes o riesgos potenciales, manteniendo en todo momento al usuario como responsable final de la decisión. De este modo, el sistema actúa como un apoyo informativo y no como una instancia de juicio o asesoría legal.

El proyecto reconoce también la existencia de posibles errores e imprecisiones en los resultados generados. Entre ellos se incluyen resúmenes incompletos, clasificaciones imperfectas de ciertas cláusulas y la asignación de niveles de severidad que, por su naturaleza, poseen un componente subjetivo. Estas limitaciones se derivan principalmente de las restricciones propias del modelo de lenguaje utilizado, el cual tiende a funcionar de manera más precisa en textos de menor extensión y también debido a las limitaciones que ofrece el hardware utilizado.

En cuanto a los límites del análisis automático, el sistema no es capaz de identificar todas las implicaciones jurídicas contenidas en un documento legal ni de comprender el marco

normativo completo. Por lo tanto, no sustituye una lectura profunda de la política ni el análisis realizado por un profesional de derecho. Su funcionamiento está pensado como un filtro informativo preliminar, orientado a facilitar la exploración del documento y a destacar elementos que puedan resultar relevantes para el usuario, sin emitir veredictos ni conclusiones definitivas.

Durante el desarrollo del prototipo se asumieron de manera deliberada varios trade-offs de diseño. Uno de los más relevantes fue la aceptación de falsos positivos a la hora de detectar una acción de consentimiento, esto en lugar de arriesgar que no haya una detección falsa negativa. Esta decisión responde a la naturaleza preventiva del sistema, en la que resulta preferible alertar al usuario sobre un posible momento de aceptación de consentimiento a que de por sí no este. Asimismo, se priorizó la claridad de la información sobre la exhaustividad del análisis, optando por explicaciones accesibles y comprensibles frente a una cobertura legal absoluta que podría resultar confusa para usuarios no técnicos.

Otro trade-off consciente fue el uso de un lenguaje accesible en lugar de terminología jurídica especializada. Esta decisión busca reducir la carga cognitiva del usuario y mejorar la comprensión general del contenido, aun cuando ello implique simplificar ciertos conceptos legales complejos.

En conjunto, estas consideraciones éticas y limitaciones técnicas refuerzan el carácter responsable del sistema y establecen expectativas claras sobre su uso. Al reconocer explícitamente sus alcances y restricciones, el proyecto evita presentar la inteligencia artificial como una autoridad y se posiciona como una herramienta de apoyo a la toma de decisiones informadas.

3.9. RESUMEN DEL CAPÍTULO

En este capítulo se describió el diseño y desarrollo del prototipo de extensión de navegador propuesto, abordando tanto el proceso seguido durante su construcción como la estructura y funcionamiento de sus componentes principales. Se presentó el enfoque metodológico adoptado, así como las decisiones de diseño que guiaron la implementación del sistema, poniendo énfasis en su funcionalidad como asistente no intrusivo y orientado a usuarios no técnicos.

Asimismo, se expuso la arquitectura general del sistema y el rol de cada uno de sus componentes, incluyendo la extensión de navegador, el backend de procesamiento y el pipeline de procesamiento de lenguaje natural. Se detalló la lógica de interacción, la gestión del estado, el manejo de navegación y cancelaciones, y las medidas adoptadas para garantizar la consistencia de los resultados presentados al usuario.

Finalmente, se analizaron las consideraciones éticas y las limitaciones técnicas del prototipo, delimitando explícitamente el alcance de la inteligencia artificial empleada y justificando los trade-offs asumidos durante el desarrollo. Estas decisiones establecen el marco bajo el cual debe interpretarse el funcionamiento del sistema y preparan el terreno para la evaluación presentada en el siguiente capítulo, en el que se analizará su utilidad, claridad y percepción por parte de los usuarios.

4. EVALUACIÓN DEL PROTOTIPO

4.1. ENFOQUE DE EVALUACIÓN

La evaluación del prototipo se planteó desde un enfoque cualitativo, exploratorio y centrado en el usuario. Estuvo orientado a analizar la interacción del usuario con el sistema y la comprensión de la información presentada en un contexto realista. Este enfoque resulta útil en este proyecto como un prototipo funcional de apoyo informativo, cuyo objetivo principal es asistir a usuarios no técnicos para interpretar políticas de privacidad, sin pretender ofrecer evaluaciones legales.

Desde el punto de vista metodológico, la evaluación se llevó a cabo mediante entrevistas semiestructuradas y observación directa del uso del sistema, empleando un conjunto de preguntas guía comunes para todos los participantes en cada fase de evaluación. Las respuestas y observaciones fueron registradas de forma manual por el autor, permitiendo recoger percepciones, reacciones y comentarios espontáneos de los usuarios durante la interacción con el prototipo. Las preguntas guía utilizadas durante las distintas fases de evaluación se presentan en el Anexo A.

La evaluación se centró en el usuario, priorizando la observación de su comportamiento, sus reacciones y su percepción general del sistema durante escenarios representativos de navegación web. En lugar de aplicar métricas cuantitativas, se optó por una evaluación basada en la ejecución de tareas y en la reflexión posterior del usuario, permitiendo analizar cómo los participantes interpretan la información presentada, cómo reaccionan ante el momento de intervención del sistema y qué grado de control perciben durante la interacción.

Asimismo, la evaluación se hizo de manera progresiva y por fases, alineándose con cómo se desarrolló el prototipo de forma incremental. Este enfoque permitió validar el funcionamiento del sistema en distintos niveles, comenzando por pruebas funcionales del backend, después con la evaluación de la interacción de la interfaz de usuario de la extensión de navegador y finalizando con pruebas de uso integrado. La división en fases facilitó la identificación temprana de problemas y la verificación de decisiones de diseño antes de evaluar la experiencia completa del usuario.

Dado que el sistema no pretende ser una autoridad legal ni como un modelo optimizado para tareas de clasificación o predicción, la utilización de métricas como precisión, exhaustividad o exactitud legal resultaría inapropiada y podría inducir a interpretaciones erróneas sobre los objetivos del prototipo. En su lugar, la evaluación se enfocó en aspectos como que tan clara es la información, la comprensión del resumen y las alertas, la percepción de utilidad y la sensación de control del usuario durante la interacción.

4.2. DISEÑO DE LA EVALUACIÓN

El diseño de la evaluación del prototipo se orientó a observar su uso en condiciones realistas de navegación web, priorizando la interacción natural del usuario con el sistema. Así, se pudo analizar el comportamiento del prototipo en situaciones similares a aquellas para las cuales fue diseñado, particularmente en momentos reales de consentimiento asociados a políticas de privacidad y uso de cookies.

La evaluación se realizó con participantes no técnicos, aquellos que se verían como usuarios habituales de la web que no poseen una formación en temas legales ni en tecnologías de procesamiento de lenguaje natural. Este perfil empata con el público objetivo definido durante el diseño del sistema, dado que el prototipo está pensado como un asistente informativo accesible para usuarios comunes y casuales.

En total, participaron ocho personas, acorde con el alcance exploratorio del trabajo de titulación y sus limitaciones prácticas. En la primera fase de evaluación participaron cuatro usuarios, mientras que en la segunda y tercera fase se amplió la muestra a ocho participantes. La selección de los participantes buscó incluir usuarios con distintos niveles de familiaridad con la web, permitiendo recoger percepciones variadas sobre el funcionamiento y la presentación del sistema.

El contexto de uso fue navegación web real, en la que los participantes interactuaron con sitios que presentan banners de consentimiento, enlaces a políticas de privacidad o mensajes relacionados con el tratamiento de datos personales. En este entorno, los usuarios utilizaron la extensión sin instrucciones detalladas, permitiendo observar cómo el sistema se integra de manera natural en la experiencia de navegación.

Las tareas realizadas durante la evaluación se enfocaron en acciones clave del funcionamiento del prototipo. Estas incluyeron la detección de momentos de consentimiento, la reacción del usuario ante el overlay informativo, la revisión de los resultados presentados (resumen general, escala de riesgo y pestañas por categoría) y el uso de las opciones de confirmación o cancelación del análisis. Finalizada cada interacción, se realizaron preguntas abiertas orientadas a recoger la percepción del usuario sobre la claridad, utilidad y neutralidad de la información presentada.

Cabe señalar que, dependiendo de la fase, el sistema evaluado no siempre se encontraba completamente integrado. En particular, durante la primera fase, se le presentó al usuario el análisis realizado en forma de texto y la segunda fase se utilizó una interfaz con información demostrativa (mock), con el objetivo específico de evaluar la experiencia de usuario y el diseño de la interfaz sin que el análisis automático del backend influyera en la percepción visual del sistema. Esta decisión permitió aislar la evaluación de la interfaz de los aspectos técnicos del procesamiento.

La evaluación se estructuró, en conjunto, como una evaluación de uso, enfocada en comprender la experiencia del usuario, su nivel de comprensión y su percepción de utilidad frente a la información proporcionada por el sistema.

4.3. CRITERIOS E INDICADORES DE EVALUACIÓN

Con el fin de analizar el comportamiento del prototipo y su impacto en la experiencia del usuario, se definieron una serie de criterios de evaluación cualitativos, alineados con los objetivos del sistema y con el enfoque que se tiene de asistente informativo. Estos criterios buscan evaluar la comprensión percibida, la claridad de la información y como se comporta el sistema de acuerdo al usuario.

Los indicadores corresponden directamente a las decisiones de diseño descritas en el Capítulo III y a los fundamentos teóricos presentados en el Capítulo II, especialmente en lo relativo a la ética del diseño, la no coerción y el apoyo a la toma de decisiones informadas.

El primer criterio fue la comprensión percibida del contenido. Este indicador se refiere a que el usuario pueda entender, de manera general, la información presentada por el sistema tras

la realización del análisis. Se evaluó si el usuario lograba identificar los aspectos principales de la política de privacidad y su relevancia para el tratamiento de sus datos personales.

El segundo criterio fue la claridad del resumen generado. Se observó si el resumen presentado resultaba legible, conciso y coherente desde la perspectiva del usuario, así como si lograba transmitir la idea central del documento sin generar confusión. Este criterio permitió identificar si el lenguaje utilizado era adecuado para usuarios no técnicos y si la estructura del contenido hacía que sea más fácil leerlo.

Otro indicador relevante fue la utilidad percibida de las alertas y de la escala de riesgo. En este caso, se analizó si los usuarios comprendían el significado de los niveles de riesgo presentados (bajo, medio o alto) y si estos elementos contribuían a dirigir su atención hacia posibles aspectos sensibles del documento, sin inducir decisiones ni generar alarma.

El momento de intervención del sistema fue evaluado también, dada su importancia en el diseño ético del prototipo. Se observó si la aparición del overlay informativo resultaba oportuna dentro del flujo de navegación del usuario y si este percibía la intervención como contextual y relacionada con la acción de consentimiento, en lugar de como una interrupción.

Finalmente, se consideró la sensación de control del usuario durante la interacción con el sistema. Este criterio incluyó aspectos como la posibilidad de cancelar el análisis, continuar navegando sin restricciones y confirmar la comprensión del contenido mediante acciones explícitas. La percepción de control resulta fundamental para validar que el sistema actúa como un asistente informativo y no como un mecanismo coercitivo.

4.4. RESULTADOS DE LA EVALUACIÓN

Una síntesis estructurada de los principales hallazgos cualitativos obtenidos en cada fase de evaluación se presenta en las tablas incluidas en el Anexo B, las cuales resumen los aspectos más relevantes observados a partir de las preguntas planteadas.

Los resultados de la evaluación del prototipo se presentan organizados en fases de prueba, las cuales corresponden a las distintas etapas en las que el sistema fue validado durante su desarrollo.

La evaluación no se desarrolló como un proceso solamente del final del desarrollo, sino como una serie de momentos de verificación orientados a asegurar la coherencia funcional, la claridad de la información presentada y el respeto por la experiencia del usuario.

4.4.1. FASE 1: EVALUACIÓN DEL BACKEND Y DEL PIPELINE DE PROCESAMIENTO

La primera fase de evaluación tuvo como objetivo validar el funcionamiento conceptual y técnico del backend, así como la claridad de los resultados generados por el pipeline de procesamiento de lenguaje natural, de manera independiente a la interfaz de usuario de la extensión. Esta fase se enfocó principalmente en determinar si la información producida por el sistema resultaba comprensible, suficiente y adecuada para usuarios no técnicos.

En esta etapa participaron cuatro personas, de las cuales dos contaban con mayor familiaridad con el uso de la web y conceptos tecnológicos, mientras que las otras dos correspondían a usuarios sin conocimientos técnicos o legales profundos. Esta selección permitió obtener retroalimentación desde distintas perspectivas, manteniendo el foco en el público objetivo del sistema.

La evaluación se realizó utilizando ejemplos de textos procesados por el pipeline, presentados mediante la interfaz de prueba del backend. Posteriormente, se formularon preguntas abiertas comunes a todos los participantes, orientadas a recoger su percepción general sobre los resultados obtenidos. Entre las principales preguntas planteadas se incluyeron:

1. “¿El texto generado por el sistema resulta comprensible?”
2. “¿Considera que la información presentada es suficiente para entender el contenido general de la política de privacidad?”
3. “¿Cuántos puntos o alertas por categoría considera adecuados antes de que la información resulte excesiva o difícil de procesar?”

Los resultados de esta fase indicaron que, en términos generales, los participantes lograron comprender la información presentada por el sistema, especialmente cuando esta se expresaba en lenguaje natural y hacía referencia explícita a fragmentos del documento

original. Tanto los usuarios con mayor experiencia en la web como aquellos sin formación técnica coincidieron en que el contenido resultaba más accesible que la lectura directa de una política de privacidad completa.

Sin embargo, también se identificó que un número excesivo de puntos por categoría podía resultar abrumador, afectando la claridad del mensaje principal. A partir de este feedback, se reforzó la decisión de limitar la cantidad de alertas por categoría, priorizando la relevancia de la información sobre la exhaustividad. Esta observación influyó directamente en la definición de límites en la generación de resultados del pipeline, así como en la estructura final de presentación de las alertas.

Asimismo, los participantes señalaron que la información proporcionada era suficiente para obtener una visión general del contenido del documento, aunque no reemplazaba una lectura completa del mismo. Esta percepción resultó exacta con el objetivo del sistema como asistente informativo y reforzó la postura adoptada en el diseño, según la cual los resultados deben entenderse como aproximaciones orientativas y no como interpretaciones legales exhaustivas.

En conjunto, esta primera fase permitió validar la viabilidad técnica del pipeline y confirmar que el enfoque de segmentación, selección y síntesis de información era adecuado para usuarios no técnicos. Los hallazgos obtenidos sirvieron como base para ajustar tanto la cantidad de información generada como el tono de los mensajes, antes de avanzar hacia la evaluación de la interfaz de usuario y de la experiencia de interacción con la extensión.

4.4.2. FASE 2: EVALUACIÓN DE LA INTERFAZ DE USUARIO Y DEL OVERLAY INFORMATIVO

La segunda fase de evaluación se enfocó en analizar la experiencia de usuario (UX) de la extensión de navegador, con especial atención en la forma en que el sistema presenta la información y en cómo el usuario percibe la intervención del overlay informativo durante la navegación web. A diferencia de la fase anterior, en esta etapa el objetivo principal no fue evaluar la precisión del análisis, sino la claridad, neutralidad y nivel de intrusión de la interfaz.

En esta fase participaron ocho usuarios no técnicos, correspondientes al público objetivo del sistema. Para esta evaluación, la extensión no se encontraba completamente conectada al backend de análisis; en su lugar, se utilizó un mock con información demostrativa, diseñado específicamente para evaluar la interfaz de usuario sin que el comportamiento del pipeline influyera en la percepción visual o interactiva del sistema. Esta decisión permitió aislar la evaluación de la UI y centrarse exclusivamente en aspectos de diseño e interacción.

El contexto de uso fue navegación web real, en la que los participantes interactuaron con sitios que presentaban banners de consentimiento o mensajes relacionados con el uso de cookies y políticas de privacidad. La extensión activaba el overlay informativo al detectar un posible momento de consentimiento, mostrando el comportamiento esperado del sistema final.

Tras la interacción con el overlay, se realizaron preguntas abiertas comunes a todos los participantes, orientadas a recoger su percepción sobre la experiencia de uso. Entre las principales preguntas planteadas se incluyeron:

1. “¿Sentiste que la forma en la que se presentaba el aviso fue invasiva? ¿Por qué?”
2. “¿En algún momento el sistema detectó erróneamente un momento de consentimiento?”
3. “¿La interfaz de usuario de la extensión te pareció coherente y fácil de usar?”, y
4. “¿Sentiste que la forma en la que se presenta la información emite algún juicio de valor sobre la política de privacidad?”

Los resultados de esta fase indicaron que, en general, los usuarios no percibieron el overlay informativo como una intervención invasiva. La mayoría coincidió en que su aparición resultaba contextual y relacionada con la acción de consentimiento que estaban realizando, lo que contribuyó a que el sistema fuera interpretado como un apoyo informativo y no como una interrupción arbitraria.

No obstante, algunos participantes señalaron la presencia ocasional de falsos positivos en la detección de momentos de consentimiento, particularmente en páginas con múltiples banners o elementos visuales similares. Esta observación fue considerada aceptable dentro del alcance del proyecto y reforzó la decisión de priorizar la detección temprana de escenarios relevantes para el usuario, incluso a costa de una detección imperfecta.

En cuanto a la interfaz, los usuarios lograron comprender la estructura general de la información presentada, incluyendo el resumen principal, la escala de riesgo y las pestañas por categoría. Sin embargo, también se identificaron oportunidades de mejora relacionadas con la claridad de ciertos elementos visuales y con la necesidad de ofrecer mecanismos explícitos de cancelación o cierre, con el fin de reforzar la sensación de control del usuario durante la interacción. Fue en esta etapa en donde se evidenció la necesidad de crear mecanismos de cancelación a partir de la navegación del usuario, para que continúe teniendo una navegación fluida sin gastar recursos del sistema.

Respecto a la percepción de neutralidad del sistema, los participantes coincidieron en que la interfaz de usuario no inducía explícitamente a aceptar o rechazar la política de privacidad, ni presentaba elementos visuales que sugirieran una decisión preferente excepto en ciertos aspectos como las etiquetas de riesgo. Esta percepción estuvo asociada principalmente a la forma en que la información es presentada por la interfaz, incluyendo el lenguaje utilizado, la jerarquía visual de los elementos y la ausencia de botones o mensajes con carga persuasiva.

4.4.3. FASE 3: EVALUACIÓN DEL SISTEMA INTEGRADO Y ESCENARIOS DE USO REALES

La tercera fase de evaluación tuvo como objetivo analizar el comportamiento del sistema en su forma completamente integrada, considerando el flujo completo desde la detección del momento de consentimiento hasta la presentación final de los resultados generados por el backend. A diferencia de las fases anteriores, en esta etapa se evaluó el sistema en condiciones de uso más cercanas a un escenario real, incorporando tanto el procesamiento automático del texto como la interacción completa con la interfaz de usuario.

En esta fase participaron ocho usuarios no técnicos, quienes interactuaron con la extensión durante sesiones de navegación web real. El sistema se encontraba plenamente conectado al backend de análisis, permitiendo evaluar aspectos como el tiempo de respuesta, la coherencia del flujo de interacción y la estabilidad del sistema frente a acciones del usuario, como cambios de página o cancelación del análisis.

Tras la interacción con el sistema integrado, se realizaron preguntas abiertas comunes a todos los participantes, orientadas a recoger su percepción sobre el funcionamiento general del prototipo. Entre las principales preguntas planteadas se incluyeron:

1. “¿Qué percepción tienes respecto al tiempo de duración del análisis?”
2. “¿Encontraste información que no esperabas dentro del análisis presentado?”
3. “¿Consideras que la información presentada te ofrece una visión más amplia para tomar una decisión informada?”

Los resultados de esta fase indicaron que, en términos generales, el tiempo de duración del análisis fue percibido como aceptable para el tipo de información proporcionada. Algunos usuarios señalaron que el proceso podía resultar ligeramente más largo que una aceptación inmediata de un banner de consentimiento, y coincidieron que el tiempo de espera podría mejorar.

Asimismo, se concluyó haber encontrado información que no se esperaba dentro del análisis, especialmente relacionada con prácticas de recopilación de datos. Este hallazgo resultó relevante, ya que evidencia que el sistema logra llamar la atención del usuario sobre aspectos que suelen pasar desapercibidos en una lectura superficial de las políticas de privacidad.

En cuanto a la toma de decisiones, los usuarios manifestaron que la información presentada les permitía contar con una visión más amplia y contextualizada antes de aceptar una política de privacidad. No obstante, reiteraron que el sistema no imponía una decisión ni condicionaba su acción final, manteniendo la percepción del prototipo como un asistente.

Finalmente, se observó que los mecanismos de control implementados, en particular la posibilidad de cancelar el análisis o continuar navegando sin restricciones, funcionaron de manera coherente durante escenarios de navegación interrumpida.

4.5. ANÁLISIS Y DISCUSIÓN DE RESULTADOS

El análisis de los resultados obtenidos durante la evaluación del prototipo permite contrastar las decisiones de diseño y desarrollo adoptadas con los objetivos planteados al inicio del proyecto. De esta manera, lo que hemos encontrado en las distintas fases de evaluación evidencian que el sistema cumple su propósito principal como asistente orientado a mejorar

la comprensión de políticas de privacidad, sin interferir de manera invasiva en la experiencia del usuario.

En primer lugar, los resultados de la fase de evaluación del backend y del pipeline de procesamiento validan la viabilidad técnica del enfoque envisioned. El sistema es capaz de segmentar, clasificar y resumir textos legales. Y se confirma que las técnicas de procesamiento de lenguaje natural descritas en el Capítulo II pueden ser aplicadas de forma efectiva. Al mismo tiempo, las limitaciones observadas en textos particularmente extensos refuerzan la postura teórica adoptada en esta investigación, según la cual los resultados generados por el sistema deben entenderse como aproximaciones orientativas y no como interpretaciones legales exhaustivas.

En relación con la evaluación de la interfaz de usuario y del overlay informativo, los resultados congenian de acuerdo con lo diseñado en el Capítulo III. La detección oportuna de momentos de consentimiento y la presentación contextual de la información contribuyeron a que los usuarios se detuvieran brevemente antes de aceptar una política de privacidad. Este comportamiento valida en cómo se programó el inicio del sistema y en la integración discreta del asistente dentro del flujo de navegación, aspectos fundamentales para evitar el rechazo por parte del usuario.

La escala de riesgo se consolidó como otra alerta, alineada con el objetivo de orientar al usuario sin imponer decisiones. Aunque algunos participantes manifestaron interés en conocer con mayor detalle los criterios de asignación del riesgo, la interpretación general del indicador como una referencia informativa y no como una recomendación explícita confirma que el diseño logró evitar un rol autoritativo del sistema. Sin embargo, los usuarios consideraron a la métrica demasiado objetiva para algo que puede ser subjetivo. Este resultado es consistente con el enfoque ético planteado en el Capítulo II, sin embargo, se tendrá que reevaluar en cuanto a la información que se le da al usuario.

Los resultados de la evaluación del sistema integrado hacen ver a los mecanismos de control implementados, particularmente en lo relativo a la cancelación del análisis. La correcta transición a estados de cancelación o inactividad hace imposible la presentación de resultados inconsistentes al contexto y reforzó la percepción del sistema como un asistente responsable.

Viéndolo globalmente, los resultados confirman la validez de los trade-offs asumidos durante el diseño del sistema. La aceptación consciente de falsos positivos permitió enfocarse en la detección de escenarios relevantes para el usuario, mientras que la preferencia por claridad y accesibilidad sobre exhaustividad legal contribuyó a una experiencia más comprensible para usuarios no técnicos. Estas decisiones, lejos de debilitar el sistema, refuerzan su estado como un intermediario entre la política y el usuario.

En conjunto, el análisis de los resultados permite afirmar que el prototipo logra materializar los principios teóricos, técnicos y éticos planteados a lo largo de la investigación. El sistema se comporta de manera coherente en su concepción como asistente informativo, valida las decisiones de diseño y establece una base sólida para reflexionar sobre posibles mejoras y extensiones futuras.

4.6. RESUMEN DEL CAPÍTULO

En este capítulo se presentó la evaluación del prototipo desarrollado, con el objetivo de analizar su comportamiento y su aporte como asistente informativo en contextos reales de navegación web. La evaluación se planteó desde un enfoque cualitativo y exploratorio, coherente con el alcance del proyecto y con la naturaleza del sistema propuesto.

Se describió el diseño de la evaluación, considerando un perfil de usuarios no técnicos, un contexto de uso realista y un conjunto de tareas representativas del funcionamiento del sistema. De la misma manera, se definieron criterios e indicadores centrados en la experiencia del usuario, priorizando aspectos como la comprensión percibida, la claridad de la información, el momento de intervención y la sensación de control durante la interacción.

Los resultados se presentaron organizados por fases de evaluación, lo que permitió mostrar una validación continua durante el desarrollo del prototipo. A partir de estos resultados, se realizó un análisis y discusión que vinculó las observaciones obtenidas con los objetivos del proyecto, las decisiones de diseño y los fundamentos teóricos escritos en capítulos anteriores.

La evaluación permitió constatar que el sistema cumple su rol como asistente informativo. Proporciona apoyo contextual al usuario sin interferir de manera intrusa en su experiencia de navegación. Estos hallazgos sirven como base para las conclusiones generales del trabajo

y para la identificación de posibles líneas de mejora y trabajo futuro, que se abordan en el capítulo siguiente.

5. CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

A partir del desarrollo, implementación y evaluación del prototipo presentado en este trabajo, se establecen las siguientes conclusiones, las cuales responden directamente a los objetivos planteados y a los resultados observados durante la investigación.

1. Viabilidad del enfoque propuesto

Se concluye que el enfoque de interpretación asistida de políticas de privacidad mediante una extensión de navegador es técnicamente viable y conceptualmente coherente. El prototipo desarrollado demuestra que es posible integrar mecanismos de detección, procesamiento y presentación de información relevante en el momento de consentimiento del usuario, sin requerir modificaciones en los sitios web ni interrumpir el flujo normal de navegación.

2. Uso de inteligencia artificial como asistente informativo

El uso de técnicas de procesamiento de lenguaje natural e inteligencia artificial generativa resulta adecuado cuando se delimita explícitamente su rol como apoyo informativo. El sistema no actúa como autoridad legal ni reemplaza leer el documento original, sino que facilita una comprensión preliminar de contenidos complejos, cumpliendo con el objetivo de asistir al usuario sin decidir por él.

3. Impacto en la experiencia del usuario no técnico

Los resultados de la evaluación indican que el sistema contribuye a mejorar la comprensión percibida de las políticas de privacidad por parte de usuarios no técnicos. La presentación estructurada de la información, el uso de lenguaje accesible y la organización por categorías permiten reducir la carga cognitiva asociada a la lectura de textos legales extensos.

4. Relevancia del momento de intervención y la no coerción

El momento en el que se presenta la información es un factor clave para que el sistema sea efectivo. Al activarse el asistente durante el proceso de consentimiento, sin bloquear ni forzar acciones, refuerza la reflexión del usuario sin generar rechazo. Este enfoque hace que el timing y la autonomía sean confirmados como principios.

5. Coherencia de los trade-offs asumidos

Las decisiones de diseño adoptadas, como la aceptación consciente de falsos positivos y la priorización de claridad sobre exhaustividad legal, se muestran coherentes con los objetivos del proyecto. Estos trade-offs permitieron construir un sistema utilizable y éticamente responsable, alineado con su definición como asistente informativo y no como herramienta de veredicto.

5.2. APORTES DEL PROYECTO

El presente trabajo de titulación aporta contribuciones relevantes en los ámbitos técnico, metodológico y conceptual, dentro del contexto de la privacidad digital y el diseño de sistemas centrados en el usuario.

1. Desarrollo de un prototipo funcional integrado

Uno de los principales aportes del proyecto es el desarrollo de un prototipo funcional completo. Una extensión de navegador basada en tecnologías web con un backend de procesamiento de lenguaje natural. A diferencia de propuestas puramente teóricas, el sistema fue diseñado, implementado y evaluado como una solución operativa, capaz de ejecutarse en escenarios reales de navegación web.

Este prototipo demuestra la viabilidad de articular distintos componentes como la detección en el navegador, procesamiento externo y presentación de resultados, de manera coherente y estable.

2. Integración responsable de inteligencia artificial generativa

El proyecto aporta un ejemplo práctico de integración consciente de inteligencia artificial generativa dentro de un sistema interactivo orientado a usuarios no técnicos. La IA no es presentada como una autoridad ni como un mecanismo de decisión automática, sino como una herramienta que hace más fácil comprender la información que desea saber.

Este enfoque contribuye a la discusión sobre el uso responsable de IA generativa en contextos sensibles, demostrando que es posible aprovechar sus capacidades sin delegar en ella responsabilidades que corresponden al usuario o a especialistas humanos.

3. Enfoque ético de diseño centrado en el usuario

Otro aporte significativo del proyecto radica en la aplicación de principios de diseño ético de experiencia de usuario al ámbito de la privacidad digital. El sistema fue concebido desde su inicio como un asistente informativo no coercitivo, priorizando la transparencia, la claridad y el respeto por la autonomía del usuario.

Decisiones como la no imposición de acciones, la posibilidad de cancelar el análisis, la no interrupción del flujo de navegación y la presentación contextual de la información materializan estos principios en un sistema concreto. Este enfoque demuestra cómo consideraciones éticas pueden traducirse en decisiones técnicas y de diseño verificables.

4. Enfoque ético de diseño centrado en el usuario

Finalmente, el proyecto aporta un enfoque basado en un desarrollo secuencial incremental con validación continua, adecuado para proyectos académicos con alcance definido y sin cambios de requisitos. La implementación y prueba progresiva de módulos permitió identificar limitaciones y ajustar decisiones sin comprometer la coherencia del sistema final.

5.3. LIMITACIONES DEL PROYECTO

A pesar de los resultados obtenidos, el proyecto presenta limitaciones que deben ser reconocidas para una correcta interpretación de su alcance.

En primer lugar, el sistema no garantiza exactitud legal, ya que el análisis automático no comprende el marco jurídico completo ni evalúa la validez normativa de las cláusulas. Se le está pidiendo al modelo de lenguaje resumir la información entregada, por lo tanto no toda la información será incluida. En segundo lugar, la evaluación realizada fue de carácter cualitativo y exploratorio, con un número reducido de participantes, lo que impide la generalización estadística de los resultados.

Asimismo, el funcionamiento del sistema depende de la calidad y estructura del texto original de las políticas de privacidad, así como de las capacidades y limitaciones inherentes a los modelos de lenguaje utilizados. Como se pudo observar, el modelo de lenguaje y el pipeline funciona mejor y hace obtener mejores resultados en políticas de privacidad más cortas. Eso le permite segmentar y encontrar información de forma más eficiente. Finalmente, el análisis cubre únicamente una parte de la información contenida en estos documentos, priorizando elementos relevantes para el usuario en lugar de una cobertura exhaustiva.

Estas limitaciones no invalidan el proyecto, sino que delimitan su propósito como herramienta de apoyo informativo dentro de un contexto académico e investigativo.

5.4. RECOMENDACIONES

A partir de los resultados obtenidos y de las limitaciones identificadas a lo largo del desarrollo y la evaluación del prototipo, se plantean a continuación una serie de recomendaciones y líneas de trabajo futuro. Estas propuestas no buscan extender el alcance del proyecto más allá de su naturaleza académica, sino explorar mejoras razonables con el enfoque actual de este trabajo.

1. Mejoras técnicas del sistema

Una primera línea de trabajo futuro se relaciona con la mejora de los componentes técnicos del sistema. En particular, podría profundizarse en optimizar el pipeline de procesamiento para manejar de forma más eficiente aquellas políticas de privacidad de gran extensión o con estructuras particularmente complejas. Esto incluiría el refinamiento de los mecanismos de

segmentación y clasificación, así como la adaptación del procesamiento a distintos formatos de presentación del contenido legal.

Mejora de la explicación y transparencia de los resultados. Asimismo, podrían explorarse mejoras en el sistema de detección de momentos de consentimiento, incorporando nuevos patrones o ajustando el sistema de scoring con el fin de reducir falsos positivos y aumentar la precisión en el contexto del usuario.

2. Despliegue y disponibilidad del backend

Otra posible línea de trabajo futuro consiste en evaluar el despliegue del backend de procesamiento en un entorno de infraestructura en la nube. Esta recomendación se plantea desde una perspectiva orientada a mejorar la disponibilidad y eficiencia del servicio y facilitar pruebas del sistema en distintos contextos y dispositivos.

El despliegue en la nube permitiría, además, analizar aspectos relacionados con la latencia, la estabilidad del servicio y la gestión de solicitudes concurrentes.

3. Ampliación de la evaluación con usuarios

Desde el punto de vista metodológico, se recomienda ampliar la evaluación del sistema con estudios con un mayor número de participantes y distintos perfiles. Futuras evaluaciones podrían incluir usuarios con distintos niveles de familiaridad con tecnologías digitales.

Adicionalmente, podrían desarrollarse evaluaciones comparativas entre usuarios que utilizan el asistente y aquellos que aceptan políticas de privacidad sin apoyo, con el objetivo de saber el impacto del sistema sobre cuanta información han comprendido y el proceso de toma de decisiones.

4. Mejora de la explicación y transparencia de los resultados

Otra recomendación relevante se centra en profundizar los mecanismos de explicación. En trabajos futuros, podría incorporarse una mayor transparencia respecto a los criterios utilizados para la asignación de niveles de riesgo y para la selección de fragmentos relevantes

del texto original. Estas mejoras permitirían fortalecer la confianza del usuario en el sistema, reforzando su rol como intermediario informativo.

5. Líneas de investigación futura

Finalmente, el trabajo abre la puerta a futuras investigaciones en torno al consentimiento informado asistido por tecnología. Entre estas líneas se incluyen estudios sobre la eficiencia de distintas formas de presentación de información legal o el análisis del impacto del timing en la toma de decisiones del usuario.

Estas líneas de investigación permiten ver el proyecto no como un resultado aislado, sino como un punto de partida para el desarrollo de soluciones tecnológicas responsables orientadas a mejorar la comprensión de la privacidad en entornos digitales.

5.5. CIERRE FINAL DEL TRABAJO

El presente trabajo de titulación tuvo como objetivo desarrollar y evaluar un prototipo de extensión de navegador orientado a facilitar la comprensión de políticas de privacidad en el momento de consentimiento del usuario. A través de un enfoque asistivo, ético y centrado en la experiencia de usuario, se demostró que la integración responsable de técnicas de procesamiento de lenguaje natural e inteligencia artificial puede contribuir a mejorar el acceso a información compleja sin reemplazar la lectura legal ni imponer decisiones. El proyecto concluye reafirmando la importancia de diseñar sistemas que respeten la autonomía del usuario, prioricen la claridad y actúen como herramientas de apoyo informativo, estableciendo así una base sólida para futuras investigaciones y desarrollos en el ámbito de la privacidad digital.

6. BIBLIOGRAFÍA

- Ahrari, L. (2024). Digital privacy and consent: Exploring user awareness of legal protections. *Journal of Historical Research, Law and Policy*, 2(1), 47–55.
- Ariai, F., Mackenzie, J., & Demartini, G. (2025). Natural language processing for the legal domain: A survey of tasks, datasets, models, and challenges. *ACM Computing Surveys*, 58(6), 1–37.
- Azamfirei, R., Kudchadkar, S. R., & Fackler, J. (2023). Large language models and the perils of their hallucinations. *Critical Care*, 27(1), 120.
- Bielova, N., Litvine, L., Nguyen, A., Chammat, M., Toubiana, V., & Hary, E. (2024). The effect of design patterns on (present and future) cookie consent decisions. In *33rd USENIX Security Symposium (USENIX Security 24)* (pp. 2813–2830).
- Biondi, G., Cagnoni, S., Capobianco, R., Franzoni, V., Lisi, F. A., Milani, A., & Vallverdú, J. (2023). Ethical design of artificial intelligence-based systems for decision making. *Frontiers in Artificial Intelligence*, 6, 1250209.
- Dong, C., Li, Y., Gong, H., Chen, M., Li, J., Shen, Y., & Yang, M. (2022). A survey of natural language generation. *ACM Computing Surveys*, 55(8), 1–38.
- Feyisetan, O., & Ganev, G. (2022). Differential privacy in natural language processing. In *Proceedings of the Workshop on Privacy in Natural Language Processing* (pp. 1–10). Association for Computational Linguistics.
- Harkous, H., Fawaz, K., Lebet, R., Schaub, F., Shin, K. G., & Aberer, K. (2020). Natural language processing of privacy policies: A survey. *arXiv preprint arXiv:2501.10319*. <https://arxiv.org/html/2501.10319v1>
- Katz, D. M., Hartung, D., Gerlach, L., Jana, A., & Bommarito II, M. J. (2023). Natural language processing in the legal domain. *arXiv preprint arXiv:2302.12039*.
- Korunovska, J., Kamleitner, B., & Spiekermann, S. (2020). The challenges and impact of privacy policy comprehension. *arXiv preprint arXiv:2005.08967*.
- Klymenko, O., & Yampolskiy, R. (2022). How to keep text private? A systematic review of deep learning methods for privacy-preserving natural language processing. *Artificial Intelligence Review*, 55(6), 4651–4695.
- Pilault, J., Li, R., Subramanian, S., & Pal, C. (2020, November). On extractive and abstractive neural document summarization with transformer language models. In

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 9308–9319).

- Ravichander, A., Black, A. W., Wilson, S., Norton, T., & Sadeh, N. (2020). Beyond the text: Analysis of privacy statements through syntactic and semantic role labeling. arXiv preprint arXiv:2010.00678.
- Sánchez Chamorro, L., Bongard-Blanchy, K., & Koenig, V. (2023, July). Ethical tensions in UX design practice: Exploring the fine line between persuasion and manipulation in online interfaces. In Proceedings of the 2023 ACM Designing Interactive Systems Conference (pp. 2408–2422).
- Wang, L., Zhang, Z., Wang, D., Cao, W., Zhou, X., Zhang, P., ... & Tian, F. (2023). Human-centered design and evaluation of AI-empowered clinical decision support systems: A systematic review. *Frontiers in Computer Science*, 5, 1187299.

7. ANEXOS

7.1. ANEXO A: GUÍA DE PREGUNTAS UTILIZADAS EN LA EVALUACIÓN DEL PROTOTIPO

El presente anexo describe la guía de preguntas utilizada durante la evaluación cualitativa del prototipo. Las preguntas fueron empleadas como parte de entrevistas semiestructuradas y observación directa del uso del sistema. Las mismas fueron formuladas de manera consistente a todos los participantes en cada fase de evaluación, con ligeras variaciones según el estado de desarrollo del prototipo.

Las respuestas fueron registradas manualmente por el autor y utilizadas como base para el análisis cualitativo presentado en el Capítulo IV.

Fase 1 – Evaluación del backend y del pipeline

1. “¿El texto generado por el sistema resulta comprensible?”
2. “¿Considera que la información presentada es suficiente para entender el contenido general de la política de privacidad?”
3. “¿Cuántos puntos o alertas por categoría considera adecuados antes de que la información resulte excesiva o difícil de procesar?”

Fase 2 – Evaluación de la interfaz de usuario y del overlay

1. “¿Sentiste que la forma en la que se presentaba el aviso fue invasiva? ¿Por qué?”
2. “¿En algún momento el sistema detectó erróneamente un momento de consentimiento?”
3. “¿La interfaz de usuario de la extensión te pareció coherente y fácil de usar?”
4. “¿Sentiste que la forma en la que la interfaz presenta la información emite algún juicio de valor sobre la política de privacidad?”

Fase 3 – Evaluación del sistema integrado

1. “¿Qué percepción tienes respecto al tiempo de duración del análisis?”
2. “¿Encontraste información que no esperabas dentro del análisis presentado?”

3. “¿Consideras que la información presentada te ofrece una visión más amplia para tomar una decisión informada?”

7.2. ANEXO B: SÍNTESIS DE RESULTADOS CUALITATIVOS DE LA EVALUACIÓN

El presente anexo presenta una síntesis estructurada de los principales resultados obtenidos durante la evaluación cualitativa del prototipo transcrito de las anotaciones a mano de parte del autor en las entrevistas realizadas.

Los hallazgos se organizan por fase de evaluación y por pregunta guía, resumiendo los aspectos más relevantes observados en las respuestas de los participantes y su impacto directo en decisiones de diseño y ajuste del sistema. Esta síntesis no pretende cuantificar resultados, sino identificar patrones recurrentes y percepciones compartidas entre los usuarios.

Fase 1 – Evaluación del backend y del pipeline

| Pregunta | Aspectos observados | Resultados relevantes / Impacto |
|---|---|---|
| ¿El texto generado por el sistema resulta comprensible? | Los usuarios lograron entender el contenido general sin necesidad de leer la política completa. | Se confirmó el uso de lenguaje natural y se evitó terminología legal compleja. |
| ¿La información presentada es suficiente para entender la política? | La mayoría consideró suficiente el resumen para una visión general, pero no como reemplazo total del documento. | Se reforzó el rol del sistema como asistente informativo, no como intérprete legal. |
| ¿Cuántos puntos por categoría son adecuados? | Un número elevado de alertas generaba saturación de información. | Se limitaron los bullets por categoría priorizando relevancia a cinco bullets. |

Tabla 1: Síntesis de resultados de evaluación del backend y pipeline

La tabla presenta los principales hallazgos obtenidos durante la evaluación del backend y del pipeline de procesamiento de lenguaje natural, sintetizando las percepciones de los participantes en relación con la comprensión del contenido, la suficiencia de la información y la cantidad adecuada de alertas por categoría.

Fase 2 – Evaluación de la interfaz y del overlay

| Pregunta | Aspectos observados | Resultados relevantes / Impacto |
|---|--|--|
| ¿El aviso fue percibido como invasivo? | La mayoría lo percibió como contextual y relacionado con la acción del consentimiento. | Se mantuvo la activación del overlay sin bloqueo de navegación. |
| ¿Hubo detecciones erróneas de consentimiento? | Se identificaron algunos falsos positivos en sitios con múltiples banners. | Se aceptó el trade-off a favor de detección temprana. |
| ¿La interfaz es coherente y fácil de usar? | La estructura general fue comprendida, pero se identificaron mejoras visuales. | Se ajustó jerarquía visual y controles de cancelación. |
| ¿La UI emite juicios de valor? | La percepción dependía más de la forma visual que del contenido textual. | Se reforzó un diseño visual neutral y no persuasivo en cuanto a etiquetas. |

Tabla 2: Síntesis de resultados de evaluación de la interfaz y el overlay

La tabla resume los resultados de la evaluación de la interfaz de usuario de la extensión, destacando la percepción de intrusión del overlay informativo, la coherencia visual de la interfaz y la influencia del diseño en la percepción de neutralidad del sistema.

Fase 3 – Evaluación del sistema integrado

| Pregunta | Aspectos observados | Resultados relevantes / Impacto |
|---|---|--|
| ¿Cómo se percibe el tiempo de análisis? | El tiempo fue aceptable considerando el valor informativo, pero puede mejorar | Se mantuvo el procesamiento sin optimizar agresivamente a costa de claridad. |
| ¿Encontraste información inesperada? | Varios usuarios descubrieron prácticas que desconocían. | Se validó el objetivo de visibilizar aspectos poco evidentes. |
| ¿Ayuda a tomar una decisión informada? | Los usuarios sintieron mayor contexto antes de aceptar. | Se confirmó el enfoque de apoyo sin coerción. |

Tabla 3: Síntesis de resultados de evaluación del sistema integrado

La tabla presenta los hallazgos obtenidos durante la evaluación del sistema completamente integrado, considerando la percepción del tiempo de análisis, el descubrimiento de información inesperada y el apoyo del sistema a la toma de decisiones informadas.