PROPUESTA TÉCNICA Y COMERCIAL - JUNIO 2026

ESTIL GURU
Agente Virtual de Atención Técnica

Diseño, desarrollo, despliegue y mantenimiento de una solución RAG a medida para responder consultas técnicas y corporativas autorizadas con fuentes verificables, imágenes originales y derivación controlada.

PROVEEDOR
PotenzzIA

EQUIPO ASIGNADO
Iván Prada
Mario Martínez

ENTREGA
Jueves 11 de junio de 2026

PotenzzIA

Contenido de la propuesta

PotenzzIA

CARTA DE PRESENTACIÓN

Un asistente técnico para reducir interrupciones, mejorar la respuesta y ordenar el conocimiento.

PotenzzIA presenta a ESTIL GURU una solución de agente virtual de atención técnica basada en RAG, integrada en la web corporativa y diseñada para responder únicamente desde documentación oficial y reglas internas autorizadas.

El objetivo no es sustituir al equipo técnico, sino reducir las consultas repetitivas que hoy terminan interrumpiendo a perfiles de producto, facilitar que instaladores y prescriptores encuentren respuestas en segundos y convertir cada conversación en una señal útil para mejorar la documentación de ESTIL GURU.

Durante el descubrimiento se confirmó un punto clave: en este proyecto la imagen no es un complemento visual. En instalación, una sección constructiva, una altura, una pendiente o un detalle de sellado pueden resolver mejor la duda que un párrafo. Por eso la propuesta incluye la extracción y presentación de imágenes o capturas originales de fichas y guías, siempre enlazadas al documento fuente.

El principio rector de la solución será claro: primero retrieval robusto, después complejidad. La Fase 1 prioriza recuperación fiable, trazabilidad, observabilidad y latencia antes de introducir automatizaciones o agentes especializados que solo tendrán sentido cuando los datos de uso lo justifiquen.

Posicionamiento de la propuesta: no proponemos un chatbot genérico sobre PDFs. Proponemos una primera capa de conocimiento técnico para ESTIL GURU, con trazabilidad, control documental, analítica y arquitectura preparada para el futuro asistente interno.

Fase 1 pública

Widget web para resolver consultas técnicas y derivaciones de compra desde documentación autorizada.

Fase 2 preparada

Base técnica orientada a permisos, usuarios internos, ERP, precios, pedidos y departamentos.

Propiedad ESTIL GURU

Solución a medida, entregable, portable y mantenible sin permanencia obligatoria.

PotenzzIA

PROPIEDAD REAL, SIN DEPENDENCIA

ESTIL GURU no alquila un chatbot: adquiere una solución propia.

Este punto es una diferencia clave frente a soluciones SaaS cerradas o asistentes genéricos: el desarrollo íntegro realizado para ESTIL GURU será propiedad de ESTIL GURU tras la aceptación del proyecto.

Código e infraestructura del cliente

La solución se despliega en infraestructura controlada por ESTIL GURU, preferentemente su tenant Azure. El código fuente, la infraestructura como código, la documentación técnica y la configuración del sistema se entregan como activos propios del cliente.

Sin permanencia obligatoria

No existe una licencia de uso propietaria que bloquee el acceso al sistema. PotenzzIA presta servicio de desarrollo, soporte, operación y mejora; no impone una dependencia permanente para que ESTIL GURU conserve la solución.

Portabilidad

ESTIL GURU podrá mantener, auditar, evolucionar o migrar la solución con PotenzzIA, con su equipo interno o con otro proveedor.

Transparencia

Los costes de Azure, LLM y terceros se separan de la factura PotenzzIA para evitar márgenes ocultos y facilitar control presupuestario.

Escalabilidad propia

La Fase 1 deja una base reutilizable para el asistente interno, multiidioma, RBAC e integraciones futuras.

El cliente no compra una dependencia: adquiere una solución propia, mantenible y portable. Esta filosofía reduce riesgo a largo plazo y mejora el coste total de propiedad frente a plataformas cerradas.

PotenzzIA

POTENZZIA

Especialistas en IA aplicada a procesos reales de negocio.

PotenzzIA diseña y desarrolla soluciones de inteligencia artificial orientadas a operación: asistentes conversacionales, sistemas RAG, agentes de voz, automatización de procesos internos y herramientas multiagente integradas con los sistemas del cliente.

Nuestro enfoque no parte de vender una herramienta cerrada, sino de construir soluciones a medida, con propiedad del cliente, documentación entregable y mantenimiento como servicio.

Iván Prada

Iván Prada — Consultor principal de PotenzzIA. 15 años de experiencia en gestión de proyectos y 3 años al frente de PotenzzIA diseñando y entregando soluciones de IA aplicada a procesos de negocio.

Project Manager y responsable funcional del proyecto.

Iván Prada asumirá la coordinación global del proyecto y actuará como punto principal de contacto con ESTIL GURU durante la ejecución.

Sus funciones incluyen planificación y seguimiento de hitos, coordinación de reuniones y validaciones, recogida y priorización de requisitos, diseño funcional de la solución, definición de la experiencia del asistente, coordinación de documentación, reglas internas, criterios de derivación, banco de preguntas, formación, acompañamiento en puesta en marcha y go-live.

Mario Martínez

Mario Martínez — Desarrollador full stack con 3 años de experiencia especializado en IA aplicada. Sistemas RAG, integraciones vía API y MCP, backend y frontend, e implementaciones complejas en producción.

Arquitecto de IA y desarrollador principal.

Mario Martínez asumirá la responsabilidad técnica de la solución, liderando la arquitectura, el desarrollo y la implementación del sistema de agente RAG.

Sus funciones incluyen diseño de arquitectura técnica, desarrollo principal del motor RAG, API, pipeline documental, widget y panel, configuración de infraestructura Azure, base vectorial, observabilidad, ingesta documental, chunking, metadatos, recuperación híbrida, reranking, respuesta con fuentes, pruebas técnicas, rendimiento, seguridad, trazabilidad y documentación técnica.

> El equipo asignado se mantiene durante la ejecución y la fase de mantenimiento para asegurar continuidad. Iván Prada liderará la coordinación funcional y relación con ESTIL GURU; Mario Martínez liderará la arquitectura de IA y el desarrollo principal. El refuerzo puntual de QA o soporte se coordinará bajo su dirección.

PotenzzIA

EXPERIENCIA DEMOSTRABLE

Referencias recientes en sistemas RAG en producción.

PotenzzIA aporta referencias comparables con asistentes documentales, corpus complejos, trazabilidad de fuentes, operación real y validación humana de calidad.

Proyecta · Asistente interno de conocimiento documental

Abril 2026. Asistente conversacional sobre Microsoft Teams que responde consultas internas cruzando un corpus documental masivo, complejo y heterogéneo, con control de acceso por perfil.

Despliegue sobre máquina virtual en ecosistema Azure del cliente, con modelos de Azure AI y Qdrant como base vectorial. Primera ingesta de aproximadamente 40 GB y evolución hacia la ingesta paulatina de documentación relevante hasta el orden de 1 TB.

Corpus técnico no estructurado: gráficos, planos de plantas fotovoltaicas y diagramas complejos. El sistema mantiene contexto entre preguntas encadenadas y declara explícitamente cuando no encuentra respuesta en la documentación.

Métricas: ~1,24 M fragmentos vectorizados, latencia media 5-10 s, coste medio 0,001-0,01 $ por interacción, metadatos críticos 97-99 %, rollback ~30 s y RBAC validado dentro de la base vectorial.

Contacto: José Ramón Estévez · Responsable de Organización · +34 683 554 229

Academia SM · Asistente RAG de soporte y mentoría

Febrero 2026. Sistema RAG integrado en el helpdesk de una academia online con cientos de alumnas activas, compuesto por dos asistentes especializados: soporte técnico y análisis de métricas.

Funciona con human-in-the-loop: prepara borrador de respuesta y contexto, pero nunca envía automáticamente. Una persona valida y publica, ganando velocidad sin renunciar al criterio ni al tono de marca.

Cita la lección o fuente exacta con enlace, cruza histórico de usuario, contenido formativo, contexto del equipo y CRM, y procesa capturas de pantalla, hojas de cálculo, audio y vídeo mediante IA multimodal.

Datos verificables: base vectorial con colecciones especializadas, integración con helpdesk, CRM, LMS y hojas de cálculo, capacidad multimodal en producción y despliegue monitorizado con alertas.

Contacto: Noelia Valbuena · +34 623 509 452

PotenzzIA

HALLAZGOS DEL DESCUBRIMIENTO

Acceso ágil a información fiable, autorizada y trazable.

ESTIL GURU cuenta con una base documental amplia y valiosa: fichas técnicas, guías de instalación, catálogos, certificados, vídeos, FAQs, páginas de producto e información corporativa. El reto no está en la calidad de la documentación, sino en convertir una base documental rica en acceso y síntesis inmediata, fiable y trazable.

El volumen de uso confirmado por el equipo de ESTIL GURU sitúa el proyecto en un escenario B2B de tráfico controlado: aproximadamente 14.000 usuarios y 18.500 visitas mensuales, con unas 2.500 preguntas recopiladas en 4-6 meses, alrededor de 17 consultas diarias de media. Esto favorece una arquitectura gestionada y escalable, sin sobredimensionar infraestructura desde el inicio.

Interrupciones al equipo técnico

Muchas consultas son recurrentes, pero terminan escalando a perfiles de producto porque requieren localizar una medida, compatibilidad, procedimiento o recomendación concreta.

Uso en obra

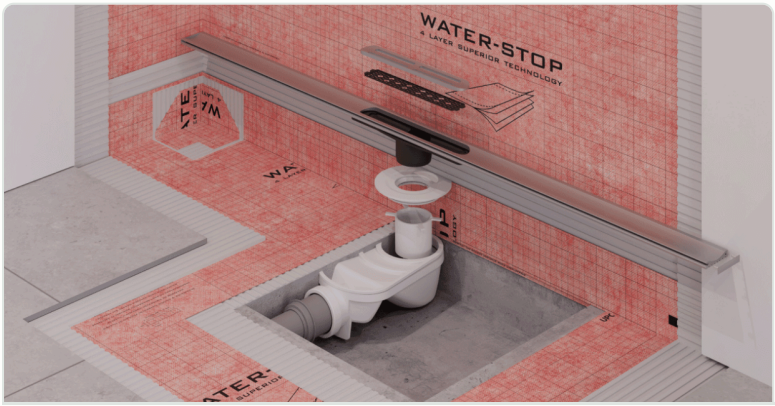
El instalador puede estar consultando desde el móvil, en una obra, sin tiempo ni facilidad para abrir una guía larga. En esos casos, la imagen o sección correcta resuelve la duda.

Información de marca y empresa

Además de fichas técnicas, el asistente debe responder o derivar consultas sobre la propia empresa, distribuidores, contactos, territorios e información corporativa autorizada.

Infraestructura testeada para escalar

La Fase 1 permite probar la infraestructura RAG en un entorno propio y controlado de ESTIL GURU, con la base técnica necesaria para escalar después hacia usos internos sin reconstruir el núcleo.



Implicación técnica: el pipeline no puede tratar todos los PDFs como texto plano. Las guías combinan idiomas, fotografías y diagramas en un mismo documento; los esquemas constructivos son conocimiento técnico, no decoración; y las traducciones ya existentes deben aprovecharse sin romper la trazabilidad.

PotenzzIA

COBERTURA FUNCIONAL DEL PLIEGO

Tipologías de consulta cubiertas en Fase 1.

El asistente se diseña para responder desde documentación oficial a las consultas repetitivas que el pliego identifica como origen de saturación del equipo de atención al cliente.

Consultas técnicas previstas

- Recomendación de producto para casos de obra concretos: soporte, humedad, tránsito, formato cerámico o aplicación prevista.
- Diferencias técnicas entre sistemas o variantes: EVO/EVOLUX, Ruber/Tryphon u otras familias equivalentes.
- Datos de especificación: medidas, rendimientos, normativas, certificaciones y compatibilidades.
- Procedimientos de instalación, pasos críticos, material auxiliar recomendado, sellados, encuentros, pendientes y solapes.
- Condiciones de garantía, ámbito de aplicación y límites de uso documentados.

Objetivos de negocio cubiertos

- Contribuir a la reducción indicativa del 30% de consultas repetitivas de nivel 1 durante los primeros seis meses.
- Dar respuesta 24/7 para preguntas cubiertas por la base de conocimiento.
- Mejorar la experiencia de instaladores, distribuidores y prescriptores al evitar navegación manual por varios PDFs.
- Detectar lagunas documentales mediante preguntas sin respuesta, feedback negativo y revisión de conversaciones.
- Convertir cada conversación en señal útil para FAQs, documentación técnica y mejora del corpus.

Criterio de respuesta: cuando el sistema no encuentre fuente suficiente en la documentación autorizada, no inventará medidas, precios, normativas, certificaciones ni procedimientos. Declinará con claridad y derivará al canal definido por ESTIL GURU.

PotenzzIA

FASE 1

Alcance incluido en la primera versión.

La Fase 1 se orienta al usuario externo de la web y a consultas recibidas por otros canales que el equipo pueda redirigir hacia el asistente. Incluye también la información interna mínima necesaria para derivar correctamente consultas de compra o contacto.

1 Widget web

Chat flotante integrable en WordPress mediante script, responsive, con mensaje de bienvenida, preguntas sugeridas, reinicio, feedback y enlace a privacidad.

2 Conversación técnica

Respuestas en español, memoria dentro de la sesión, rechazo de preguntas fuera de alcance y derivación clara cuando no exista fuente suficiente.

3 Trazabilidad

Citas visibles a documento, fragmento recuperado y enlace al PDF o recurso completo cuando esté disponible.

4 Imagen técnica

Extracción de imágenes o capturas originales de guías/fichas, asociadas a chunks y mostradas en el chat cuando aporten comprensión.

5 Información corporativa y derivación

Reglas internas para consultas sobre empresa, distribuidores, territorios o contactos, sin publicar información sensible.

6 Panel de administración

Gestión documental para personal no técnico: subir, sustituir, retirar y categorizar documentos, con reindexación automática, revisión de conversaciones, usuarios y analítica.

PotenzzIA

MARCO DE FASE 1

Supuestos de trabajo para proteger alcance, coste y plazo.

La propuesta se formula con una Fase 1 orientada al asistente público de atención técnica, información corporativa autorizada y derivación controlada. Cualquier variación sustancial del corpus, canales o integraciones deberá validarse antes del inicio o tratarse mediante alcance cerrado adicional.

Supuestos incluidos

- Corpus inicial aproximado de 60-70 documentos públicos en español, más FAQs, páginas de producto y reglas internas necesarias para derivación.
- Requisito de escalabilidad incluido en el proyecto: arquitectura preparada para procesar y operar hasta 50 GB de documentación sin rediseñar la solución. La ingesta inicial de Fase 1 se dimensiona sobre el corpus público acordado.
- Documentación técnica basada en PDFs, HTML, DOCX o TXT entregados por ESTIL GURU o publicados en su web.
- Idioma de producción inicial: español.
- Infraestructura en Azure, preferentemente en el tenant de ESTIL GURU y con datos en región UE.
- Imágenes técnicas extraídas o capturadas desde documentos originales, nunca generadas con IA.

Fuera de Fase 1

- RBAC interno por departamento o usuario para información privada.
- Integración con ERP, CRM, facturación, stock, pedidos o sistemas de backoffice.
- Teams, voz u otros canales distintos del widget web, salvo contratación específica de un módulo opcional.
- Traducción automática de documentación no disponible oficialmente.
- Rediseño de la web corporativa o cambios estructurales en WordPress.

La arquitectura sí queda preparada para multiidioma, permisos e integraciones futuras. La separación entre alcance actual y ampliaciones cerradas permite entregar valor en la web sin encarecer la primera fase con necesidades internas todavía por especificar.

PotenzzIA

COBERTURA DEL PLIEGO

Entregables previstos.

ENTREGABLE	INCLUIDO	DESCRIPCIÓN
Sistema en producción	Sí	Agente operativo en estilguru.com con widget validado.
Código fuente	Sí	Repositorio Git privado, con propiedad transferida a ESTIL GURU tras aceptación.
Infraestructura como código	Sí	Plantillas para reconstruir entornos y documentar despliegue.
Documentación técnica	Sí	Arquitectura, modelo de datos, API OpenAPI 3.0, operación y despliegue.
Manual de operación	Sí	Guía de despliegue, monitorización, resolución de incidencias habituales y operación diaria.
Manual de usuario	Sí	Manual del panel para personal no técnico.
Videotutoriales breves	Sí	Hasta 3 vídeos cortos sobre tareas frecuentes del panel de administración.
Formación	Sí	Sesión remota de 2 horas con el equipo designado por ESTIL GURU.
Informe de pruebas	Sí	Pruebas funcionales, banco de preguntas, seguridad y carga.
Plan de mantenimiento	Sí	Mantenimiento base, SLA, seguimiento, responsabilidades y horas bajo demanda para evolutivos.

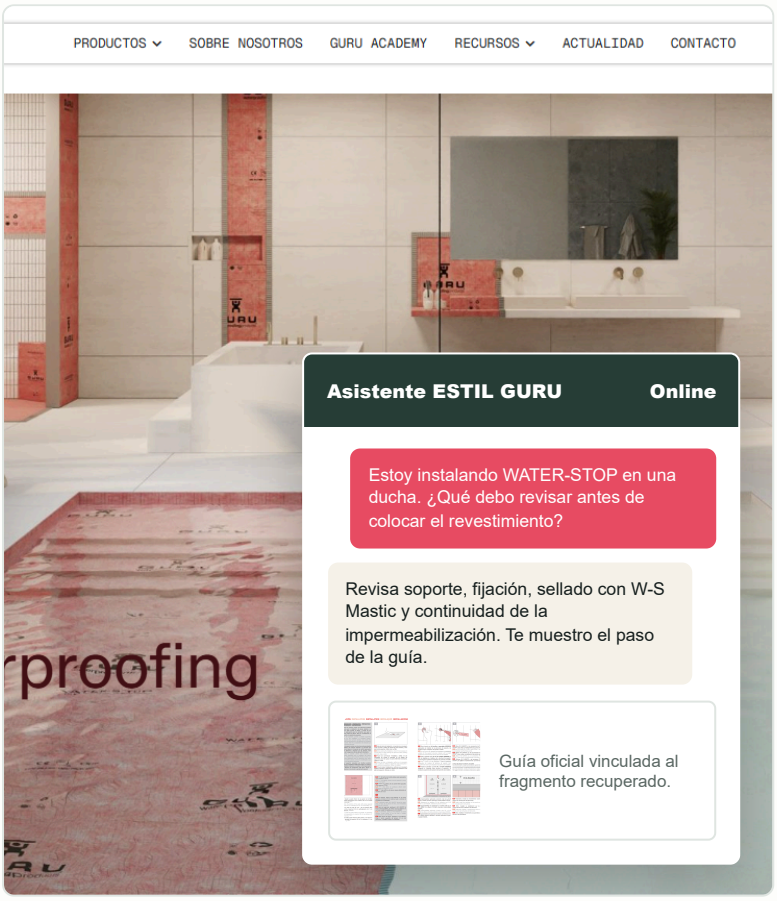
La Fase 1 se plantea completa para el widget web y deja preparada la evolución hacia asistente interno. Los módulos opcionales independientes se detallan en la sección económica correspondiente.

PotenzzIA

MOCKUPS FUNCIONALES

El asistente integrado en la web y operado desde un panel de control.

La solución se plantea como una experiencia embebida en la web de ESTIL GURU, no como una herramienta externa. El usuario consulta desde el mismo entorno de marca y el equipo interno dispone de un panel para monitorizar uso, calidad y oportunidades de mejora.



Panel de monitorización

Ejemplo visual de indicadores que permitirán revisar conversaciones, calidad del retrieval y necesidades de corpus.

2.584 Conversaciones	**91%** Feedback útil
47 Sin respuesta	**6** Docs a revisar

Producto para piscina elevada	Alta demanda
Consulta corporativa o de distribuidor	Derivada
Falta ficha de sistema modular	Corpus
Feedback negativo en una respuesta	Revisión

El panel no solo mide volumen: convierte las conversaciones en una herramienta de mejora para documentación, FAQs, producto, información corporativa, derivación controlada y entrenamiento operativo del equipo.

PotenzzIA

SOLUCIÓN AJUSTABLE

No es un asistente rígido: es una solución personalizable y evolutiva.

Además de consultar documentación y devolver información, el asistente se diseña para adaptarse a la forma en que ESTIL GURU quiere comunicar: tono, profundidad, estilo, derivaciones, mensajes de seguridad, límites y formatos de respuesta.

Tono y estilo

Más técnico, más didáctico, más breve o más comercial según el perfil de usuario y el tipo de consulta.

Reglas de marca

Lenguaje permitido, límites de respuesta, mensajes de derivación y criterios propios de ESTIL GURU.

Inputs reales

Feedback, preguntas sin respuesta y analítica para priorizar ajustes del corpus y del comportamiento.

Nuevos productos

Alta de fichas, guías, certificados o productos sin rehacer la solución: ingesta, validación e indexación.

Nuevos idiomas

Infraestructura preparada para activar español, inglés, francés, italiano, portugués y checo con corpus traducido y metadatos por idioma.

Modelo agnóstico

La lógica de negocio no queda bloqueada a un único modelo. Se podrán probar, medir y sustituir modelos conforme mejore la tecnología.

Escala con ESTIL GURU

El sistema queda preparado para absorber nueva documentación, nuevos productos, reglas comerciales, territorios, canales internos y futuros permisos sin reconstruir el núcleo.

Mejora con la IA

Si aparecen modelos más capaces o eficientes, el cambio se podrá abordar como una sustitución controlada: pruebas, medición de coste, rendimiento, precisión y despliegue validado.

PotenzzIA

INGESTA, RAG Y RESPUESTA TRAZABLE

Funcionamiento del agente en esta fase.

La Fase 1 se construye como un circuito controlado: entrada de conocimiento autorizado, recuperación robusta, generación con fuentes y realimentación desde el uso real.

01

Ingesta de fichas, guías, FAQs y reglas internas

Documentación pública y reglas mínimas de derivación.

02

Extracción, OCR, chunking e imágenes

Texto, tablas, metadatos y relación documento-fragmento-imagen.

03

Índice RAG híbrido

Búsqueda vectorial, BM25, filtros por producto, idioma y tipo documental.

04

Reranking y umbral de confianza

Priorización de fragmentos y rechazo si no hay fuente suficiente.

05

Modelo LLM desacoplado

Respuesta con tono ESTIL GURU y posibilidad de cambiar modelo sin rehacer el sistema.

06

Respuesta, cita, imagen y feedback

Texto, fuente, imagen original, derivación controlada y señales de mejora.

Bucle de mejora continua: nuevas preguntas, productos, documentos, idiomas y feedback del equipo permiten optimizar el asistente de forma progresiva, manteniendo trazabilidad y control de calidad.

PotenzzIA

PROPUESTA TÉCNICA

Arquitectura Azure sobre tenant de ESTIL GURU.

Recomendamos desplegar la solución en Microsoft Azure, aprovechando el contexto Microsoft de la organización y simplificando residencia de datos, cumplimiento RGPD y cadena contractual con proveedores cloud/LLM.

Ingestor documental

Procesa PDFs, HTML, DOCX, TXT, páginas de producto, FAQs y documentación adicional autorizada.

Chunking y metadatos

Fragmentos por tipo documental, idioma, producto, versión, fuente, imagen asociada y permisos futuros.

Base vectorial

Azure AI Search con búsqueda híbrida vectorial y BM25, preparada para escalar corpus, filtros y repositorios documentales de hasta 50 GB sin rediseñar la solución.

Orquestador LLM

Azure OpenAI con modelo principal y modelo ligero para clasificación, reranking y tareas de bajo coste.

API del agente

Servicio documentado con OpenAPI 3.0 para web y futuros canales.

Panel y analítica

Gestión documental, conversaciones, feedback, preguntas sin respuesta, uso y salud del sistema.

Principio de respuesta: si no hay fragmento recuperado con confianza suficiente, el agente no inventa. Declina de forma clara y deriva al canal correspondiente.

Escalabilidad documental: la Fase 1 se dimensiona sobre el corpus público inicial, pero el modelo de ingesta, metadatos, almacenamiento e índices queda diseñado para absorber el crecimiento previsto por ESTIL GURU hasta 50 GB de documentación.

PotenzzIA

INGESTA, RECUPERACIÓN Y GENERACIÓN

Pipeline técnico propuesto para documentación industrial.

Con documentación técnica de producto, el reto no es solo almacenar PDFs: es preservar el significado de tablas, pasos, referencias, versiones e imágenes para que cada respuesta pueda auditarse.

Ingesta documental

- Docling como motor principal para extracción estructurada de PDFs técnicos.
- Azure Document Intelligence como fallback automático cuando haya baja confianza, OCR, escaneos o layout complejo.
- Chunking semántico adaptado por tipo documental: ficha, guía, certificado, catálogo, FAQ o página web.
- Metadatos por documento: producto, tipo, idioma, versión, fecha, fuente, permisos futuros e imágenes asociadas.
- Asociación explícita documento -> fragmento -> imagen para mostrar evidencia visual en el chat.

Retrieval híbrido

- Búsqueda vectorial para capturar intención semántica de la consulta.
- BM25 para referencias literales, medidas, normativas, modelos y nombres de producto.
- Fusión RRF para combinar rankings vectoriales y léxicos de forma robusta.
- Reranking con modelo ligero para priorizar los fragmentos más útiles antes de generar.
- HyDE para mejorar consultas cortas o ambiguas donde la similitud directa pueda fallar.

Imágenes técnicas

No se regeneran imágenes con IA. Se extrae o captura la imagen original del documento, se conserva su relación con el fragmento y se muestra junto a la fuente. Esto evita errores en secciones, capas, medidas o detalles de instalación.

Generación con fuentes

Azure OpenAI se usará con un modelo principal para respuesta y un modelo ligero para clasificación/reranking. Si no existe fuente recuperada con confianza suficiente, el agente no responde de memoria: declina y deriva.

PotenzzIA

ARQUITECTURA DETALLADA

Componentes principales y principios de diseño.

COMPONENTE	FUNCIÓN	TECNOLOGÍA PROPUESTA
Ingestor documental	Extracción, chunking, metadatos, OCR selectivo e imágenes.	Docling + Azure Document Intelligence
Base vectorial	Índice semántico privado del corpus y filtros por metadatos.	Azure AI Search
Motor de retrieval	Búsqueda híbrida, RRF, reranking, HyDE y umbrales de confianza.	Azure AI Search + Azure OpenAI
Orquestador LLM	Generación con citas obligatorias y control de alcance.	Azure OpenAI, modelo principal + modelo ligero
API del agente	Interfaz para widget y futuros canales.	FastAPI o Azure Functions, OpenAPI 3.0
Widget y panel	Conversación, gestión documental, analítica y configuración.	Web app + script embebible WordPress
Observabilidad	Métricas técnicas, calidad del retrieval, alertas y trazabilidad.	Azure Monitor + métricas propias

PotenzzIA

ARQUITECTURA DETALLADA

Principios técnicos de diseño.

Principios de diseño

- Arquitectura desacoplada del proveedor LLM para evitar vendor lock-in.
- Sin LangChain ni abstracciones innecesarias: pipeline directo y depurable.
- Infraestructura como código y entornos separados.
- Tests automatizados con cobertura mínima del 70% en la lógica de negocio y tests de integración del pipeline RAG.
- Pipeline CI/CD con despliegue automatizado y rollback inmediato ante fallo, práctica ya validada en producción con rollback en ~30 segundos en el proyecto Proyecta.
- Degradación elegante si un componente externo falla.
- Campos de permisos provisionados desde Fase 1 para evolucionar a RBAC.

PotenzzIA

OBSERVABILIDAD

Medición y alertas del retrieval.

Observabilidad del retrieval

- Qué fragmentos llegan al top del ranking y con qué confianza.
- Preguntas sin respuesta o con feedback negativo.
- Documentos que no se recuperan aunque deberían.
- Ratio de derivación a humano y temas emergentes.
- Alertas proactivas ante degradación, latencia o errores.

PotenzzIA

INGESTA E IMÁGENES

Los PDFs de ejemplo confirman la necesidad de una ingesta consciente del layout.

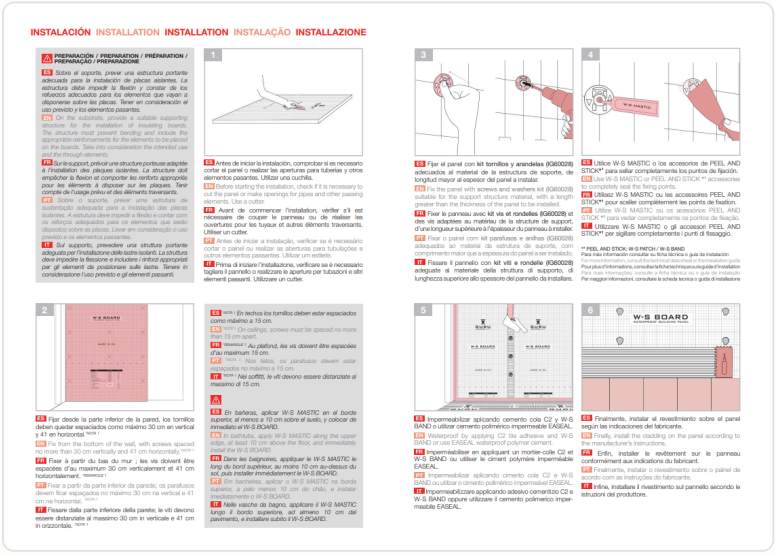
Los documentos presentes en la carpeta se tratan como muestras del tipo de material a procesar: guías de instalación con fotografías, fichas técnicas con imágenes de producto, tablas, referencias y elementos visuales que no deben perderse en una extracción plana.

Qué entra en el sistema

- Fichas técnicas y guías de instalación.
- Certificados, catálogos, FAQs y páginas de producto.
- Vídeos enlazables desde el chat cuando proceda.
- Reglas internas de información corporativa y derivación controlada.
- Información adicional autorizada para respuesta correcta.

Cómo se procesa

- Extracción de texto, tablas e imágenes.
- Asociación documento - fragmento - imagen.
- Metadatos por producto, idioma, versión y tipo.
- Reindexación al actualizar o retirar documentos.
- Cola de revisión si un documento presenta baja calidad.



Las imágenes se muestran como ejemplos visuales extraídos de documentos aportados en la carpeta del proyecto. El corpus definitivo se dimensionará tras acceso al conjunto completo.

PotenzzIA

CUMPLIMIENTO DEL PLIEGO

Seguridad, RGPD, rendimiento y criterios de aceptación.

RGPD y control del dato

- Datos alojados en regiones de la Unión Europea.
- DPA y subencargos con proveedores implicados.
- Sin uso de datos para entrenamiento de terceros.
- Retención de conversaciones con máximo de 90 días por defecto, configurable desde el panel y con borrado efectivo al superar el plazo.
- Supresión de conversaciones concretas a petición documentada, conforme al derecho de supresión del RGPD.
- Aviso visible en el widget indicando que es un sistema automatizado y que la conversación se registra con fines de mejora, con enlace a la política de privacidad.
- Anonimización de datos personales detectados.

Seguridad del sistema

- TLS en tránsito y cifrado en reposo.
- 2FA obligatorio para perfil administrador.
- Registro de auditoría de acciones administrativas.
- Defensas frente a prompt injection y exfiltración.
- Backups automáticos diarios de la base de conocimiento y la configuración, con retención mínima de 30 días.

PotenzzIA

CRITERIOS DE ACEPTACIÓN

Rendimiento, disponibilidad y pruebas de aceptación.

Estos criterios concretan cómo se validará la solución antes del cierre de la Fase 1, combinando calidad de respuesta, rendimiento técnico, seguridad y operación.

CRITERIO	COMPROMISO
Precisión	Objetivo de aceptación: al menos 90% en banco de preguntas, con respuesta alineada a fuente y cita visible.
Latencia inicial	Diseño orientado a primera palabra en menos de 2 s p95, sujeto a región Azure, modelo y longitud de contexto.
Respuesta media	Diseño orientado a respuesta total inferior a 5 s p95 mediante streaming, retrieval optimizado y modelo ligero para tareas auxiliares.
Concurrencia	Diseño para al menos 50 conversaciones simultáneas sin degradación relevante.
Disponibilidad	Objetivo 99,5% mensual excluyendo ventanas programadas y disponibilidad de terceros.
Escalado	Autoescalado cloud ante picos de hasta 5x sobre carga habitual, con revisión de límites tras observación real.
Mantenimiento programado	Ventanas fuera de horario laboral peninsular, con preaviso mínimo de 72 horas.
Monitorización	Azure Monitor, Application Insights y alertas proactivas sobre latencia, errores, consumo, disponibilidad y degradación del retrieval.
Pruebas	Funcionales, carga, seguridad, inyección de prompts, exfiltración, aceptación y formación de dos usuarios.

PotenzzIA

CALENDARIO PROPUESTO

8 semanas hasta puesta en producción supervisada.

El plazo se define desde el alcance real de la Fase 1, no solo desde la construcción de un chat. Incluye corpus técnico, imágenes, panel, analítica, derivación territorial, seguridad y aceptación.

SEMANA	TRABAJO PRINCIPAL	ENTREGABLES	DEPENDENCIAS
1	Arranque y diseño final	Plan definitivo, accesos, repositorio, corpus inicial	Azure, WordPress, documentación
2	Infraestructura y muestra de ingesta	Entornos, pipeline base, prueba con PDFs reales	Validación técnica cliente
3-4	RAG, metadatos e imágenes	Índice, embeddings, retrieval, asociación imagen-fragmento	Corpus completo y reglas
5	Widget y API	Chat web, OpenAPI, trazabilidad y feedback	Acceso WordPress/preproducción
6	Panel y derivación controlada	Panel, analítica y reglas corporativas/de derivación	Territorios, contactos y criterios definidos
7	Pruebas y ajustes	Banco de preguntas, carga, seguridad, calidad	Equipo ESTIL GURU disponible para UAT
8	Aceptación y go-live	Formación, documentación, producción supervisada	Aprobación formal

Las demoras por falta de accesos, corpus incompleto, cambios sustanciales de alcance o validaciones pendientes desplazan el calendario. PotenzzIA propondrá seguimiento semanal para anticipar bloqueos.

PotenzzIA

GESTIÓN DE RIESGOS

Riesgos técnicos identificados y medidas de control.

La principal fuente de riesgo no está en el volumen del corpus, sino en la heterogeneidad de los documentos, la calidad de extracción y la necesidad de responder con precisión sobre información técnica de obra.

RIESGO	PROB.	IMPACTO	MITIGACIÓN
PDFs complejos de InDesign, capas visuales o tablas desordenadas	Alta	Alto	Docling como motor principal y Azure Document Intelligence como fallback automático. Validación por muestra antes de ingesta completa.
Extracción difícil de imágenes técnicas	Alta	Alto	OCR/captura selectiva y revisión de asociación imagen-fragmento. Imagen original, nunca generada.
Calidad documental heterogénea o documentos contradictorios	Media	Medio	Metadatos de versión, fuente principal/secundaria y cola de revisión para documentos problemáticos.
Scope creep por incorporar Fase 2 durante Fase 1	Alta	Alto	Alcance congelado, opcionales y futuras fases acotadas, y arquitectura preparada sin implementarlas en esta fase.
Caída o degradación de Azure OpenAI	Baja	Alto	Arquitectura desacoplada del proveedor LLM, monitorización y modo degradado con derivación.
Latencia superior a 5 s en picos	Media	Medio	Streaming de respuesta, modelo ligero para tareas auxiliares, cachés de configuración y autoescalado.

La aceptación final debe apoyarse en un banco de preguntas representativo, validado por ESTIL GURU, que cubra preguntas básicas, criterio técnico, preguntas fuera de alcance e intentos de inyección o exfiltración.

PotenzzIA

INVERSIÓN

Propuesta económica desglosada.

El desarrollo inicial se propone como importe cerrado para la Fase 1. Los costes de Azure, Azure OpenAI y otros terceros se asumen directamente por ESTIL GURU en su tenant, manteniendo transparencia y propiedad sobre la infraestructura.

Desarrollo inicial Fase 1

18.300 € + IVA

Importe cerrado. IVA no incluido.

Incluye arquitectura, desarrollo, integración, pipeline documental, RAG, widget, panel, analítica, reglas de comportamiento, personalización, pruebas, documentación, formación y diseño escalable hasta 50 GB de documentación.

Servicio PotenzzIA

450 €/mes

Desde el mes siguiente a la aceptación. IVA no incluido.

Incluye mantenimiento, soporte, revisión del RAG, corrección de bugs atribuibles al desarrollo, monitorización, informe mensual y optimizaciones menores sobre el sistema existente.

FASE	TRABAJO INCLUIDO	IMPORTE
F1 - Descubrimiento y diseño	Kick-off, revisión de corpus, arquitectura definitiva, plan de proyecto y criterios de aceptación.	700 €
F2 - Infraestructura e IaC	Entornos Azure, repositorio, seguridad base, despliegue inicial e infraestructura como código.	2.400 €
F3 - Pipeline documental y RAG	Ingesta, chunking, metadatos, imágenes, embeddings, retrieval híbrido, reranking, pruebas de calidad y diseño preparado para escalar hasta 50 GB.	7.600 €
F4 - Widget, panel y comportamiento	Widget WordPress, API, panel admin, analítica, feedback, reglas de comportamiento, personalización e información accesoria.	4.300 €
F5 - Pruebas, formación y go-live	Pruebas funcionales, carga, seguridad, ajuste fino, documentación, formación y puesta en producción.	3.300 €
Total desarrollo	Importe cerrado Fase 1	**18.300 €**

PotenzzIA

RECURRENTE Y FORMA DE PAGO

Condiciones económicas de operación.

El servicio de PotenzzIA y los costes externos se muestran por separado para evitar costes ocultos. ESTIL GURU mantiene el control directo de su infraestructura y de las claves/proveedores cloud o LLM.

CONCEPTO	IMPORTE	DETALLE
Forma de pago desarrollo	30/40/30	30% a la firma, 40% al cierre de desarrollo, 30% tras aceptación formal.
Infraestructura + tokens LLM	Directo ESTIL GURU	Facturados por Azure/proveedor en el tenant del cliente. Es la máquina, no horas de PotenzzIA.
Servicio mensual PotenzzIA	450 €/mes	Mantenimiento base, soporte, revisión RAG, bugs, monitorización, informe mensual y optimizaciones menores del sistema existente.
Horas bajo demanda	65 €/h	Evolutivos, ajustes fuera de alcance y cambios menores solicitados por ESTIL GURU, siempre bajo aprobación previa.
Coste por conversación	0,015-0,06 € estimado	Depende de número de turnos, tamaño de contexto, modelo usado y uso de tareas auxiliares.

Sin permanencia ni licencias de uso: el pago mensual corresponde a mantenimiento, soporte y operación del sistema. Los evolutivos o cambios fuera de alcance se atienden a 65 €/h bajo aprobación previa; nuevos módulos o integraciones se presupuestan como proyecto independiente.

PotenzzIA

COSTE TOTAL DE PROPIEDAD

Estimación mensual y TCO a 36 meses.

Las cifras de infraestructura y LLM son estimaciones preliminares basadas en el volumen confirmado y en un escenario gestionado/serverless. Se ajustarán tras el sizing real del corpus, páginas, OCR, imágenes y uso efectivo.

CONCEPTO MENSUAL	ESCENARIO BAJO	ESCENARIO MEDIO	NOTAS
Azure OpenAI / LLM	30-120 €/mes	Incluido en rango	Modelo principal para respuesta y modelo ligero para clasificación/reranking.
Azure AI Search	Desde 75 €/mes	Hasta tier superior si la carga lo requiere	Índice vectorial, búsqueda híbrida y filtros por metadatos.
Azure Functions/App Service	20-50 €/mes	Según tráfico	API del agente, panel y servicios auxiliares.
Azure Monitor + Storage	15-25 €/mes	Según retención	Logs, métricas de retrieval, backups y alertas.
Total infra + LLM	**140-190 €/mes**	**270-330 €/mes**	Estimación operativa inicial.
Total mensual all-in	**590-640 €/mes**	**720-780 €/mes**	Servicio PotenzzIA + costes externos estimados.

Consumo por conversación

Como referencia de TCO, se estima un coste LLM aproximado de 0,015-0,06 € por conversación tipo en Fase 1, dependiendo de longitud, número de turnos, imágenes/fuentes recuperadas y mix de modelos. El coste real se medirá en Azure y se reportará mensualmente. El crecimiento documental hasta 50 GB puede requerir ajuste de tier cloud, pero no rediseño de arquitectura.

TCO a 36 meses

PotenzzIA: 34.500 € (18.300 € desarrollo + 16.200 € servicio). All-in estimado: 39.540-41.340 € en escenario bajo y 44.220-46.380 € en escenario medio. IVA no incluido.

PotenzzIA

MÓDULOS INDEPENDIENTES

Opcionales cerrados para ampliar la Fase 1.

Estos módulos se contratan de forma independiente a la Fase 1. Mantienen el mismo motor RAG, la misma base documental y las mismas reglas de respuesta, citas e imágenes técnicas.

Opcional 1 · Canal WhatsApp

680 €

Pago único. IVA no incluido.

Extiende el asistente al canal que el instalador ya usa a diario, sin duplicar el sistema: WhatsApp consume el mismo motor RAG que el widget web.

- Integración con API oficial de WhatsApp Business Cloud API de Meta.
- Alta de webhook, conexión con número WABA verificado y enrutado al agente.
- Plantillas preparadas si fueran necesarias fuera de la ventana de 24 h.
- Adaptación de imágenes técnicas al canal y enlace al PDF original.
- Mismo comportamiento que en web: fuente citada, imagen y derivación.

Opcional 2 · Plataforma multiidioma

950 €/idioma

Pago único por idioma activado. IVA no incluido.

Activa un idioma concreto aprovechando documentación oficial ya traducida por ESTIL GURU. La arquitectura queda preparada para español, inglés, francés, italiano, portugués y checo. No se traduce con IA.

- Ingesta e indexación del corpus en el idioma seleccionado.
- Detección automática de idioma y enrutado al corpus correspondiente.
- Analizadores lingüísticos del buscador y mensajes del sistema localizados.
- Fase de test y validación específica con banco de preguntas, pruebas de seguridad, cobertura documental y revisión de calidad.
- Idiomas previstos por arquitectura: español, inglés, francés, italiano, portugués y checo. Activación por idioma según prioridad comercial.

PotenzzIA

CANAL WHATSAPP

Sin sobrecoste de mensajería reactiva.

WhatsApp sin sobrecoste de mensajería reactiva

Al ser un asistente reactivo que responde a consultas iniciadas por el usuario, no se generan costes de mensajería dentro de la ventana de servicio de 24 h. La empresa no inicia campañas ni mensajes proactivos en este módulo.

PotenzzIA

PLATAFORMA MULTIIDIOMA

Activación más rápida y económica por idioma.

Multiidioma más rápido y económico

La traducción ya existe: el trabajo es ingesta, enrutado, test y validación. Los costes cloud y soporte al ampliar idiomas (procesamiento, almacenamiento, memoria/índice y mantenimiento) se valorarán por volumen documental y esfuerzo de validación.

PotenzzIA

SERVICIO POSTERIOR

Mantenimiento, SLA y propiedad.

Alcance del mantenimiento

- Soporte funcional al administrador.
- Revisión mensual de uso, feedback y preguntas sin respuesta.
- Monitorización y recomendaciones de mejora.
- Revisión del RAG, optimización menor de respuestas, fuentes, reglas y umbrales.
- Corrección de bugs atribuibles al desarrollo.
- Informe mensual resumido de actividad.

SLA propuesto

- Crítica: respuesta 2 h laborables, resolución objetivo 8 h laborables.
- Alta: respuesta 4 h laborables, resolución objetivo 24 h laborables.
- Media: respuesta 1 día laborable, resolución objetivo 5 días laborables.
- Baja: respuesta 2 días laborables, resolución objetivo 10 días laborables.

Propiedad y portabilidad

El cliente no compra una dependencia: adquiere una solución propia, mantenible y portable. Tras la aceptación, ESTIL GURU será propietario del código desarrollado, documentación e infraestructura como código.

Horas bajo demanda

Los evolutivos, ajustes fuera de alcance y cambios menores solicitados por ESTIL GURU se atenderán a 65 €/h bajo aprobación previa. Nuevos módulos, canales, integraciones o cambios de entidad se presupuestarán como proyecto independiente.

PotenzzIA

MODELO DE CONTRATO

Condiciones generales propuestas.

La contratación se formalizará mediante acuerdo de prestación de servicios, contrato de encargado del tratamiento y anexos técnicos necesarios para reflejar alcance, seguridad, subencargos y propiedad del código.

BLOQUE	CONDICIÓN PROPUESTA
Propiedad intelectual	El código desarrollado específicamente para ESTIL GURU, documentación técnica e infraestructura como código se transferirán a ESTIL GURU tras la aceptación formal y pago de los importes correspondientes.
Licencias y terceros	Servicios cloud, modelos LLM, librerías, APIs y licencias de terceros se rigen por sus propias condiciones. Cuando se usen en tenant de ESTIL GURU, el cliente mantendrá control contractual y económico directo.
Encargado del tratamiento	PotenzzIA actuará como encargado del tratamiento para los datos tratados en el proyecto. Se firmará DPA conforme a RGPD y LOPDGDD.
Subencargos	Microsoft/Azure y proveedores técnicos necesarios se declararán como subencargados cuando procesen datos por cuenta del cliente.
Confidencialidad	PotenzzIA mantendrá confidencialidad sobre documentación, conversaciones, reglas comerciales, accesos y cualquier información no pública recibida de ESTIL GURU.
Garantía	Corrección sin coste adicional de defectos atribuibles al desarrollo durante el periodo de garantía/mantenimiento acordado, siempre que no deriven de cambios externos o modificaciones no autorizadas.
Aceptación	La aceptación se basará en el banco de preguntas, criterios funcionales, pruebas de seguridad, rendimiento, documentación entregada y formación completada.
Pagos	30% a la firma, 40% al cierre de desarrollo y 30% tras aceptación formal. El mantenimiento se factura mensualmente desde el mes posterior a la aceptación.

El mantenimiento se cobra por operar, alojar, revisar, ajustar y dar soporte a la solución, no por bloquear el acceso al sistema. ESTIL GURU conserva la portabilidad del desarrollo y podrá decidir cómo evolucionarlo.

PotenzzIA

CIERRE DE PROPUESTA

Una primera fase orientada a valor real y una arquitectura lista para crecer.

PotenzzIA propone construir una solución que resuelva el caso público de atención técnica sin perder de vista la evolución interna que ESTIL GURU ya ha identificado como siguiente paso natural.

Responde mejor

Texto, cita, fragmento, imagen original y enlace al documento completo.

Protege información sensible

Información corporativa y derivación controlada sin publicar datos sensibles.

Aprende del uso

Preguntas sin respuesta, feedback y analítica para mejorar documentación.

Próximo paso recomendado: reunión técnica de alineación para validar corpus real, accesos, reglas territoriales, responsables de aceptación y entorno Azure/WordPress antes del inicio.