

Solicitud de Propuesta: Solución de Software para Etiquetado de

Datos

Índice

1. Introducción y Antecedentes
2. Descripción General del Proyecto
3. Requisitos Técnicos
4. Requisitos Funcionales
5. Requisitos No Funcionales
6. Requisitos del Proveedor
7. Criterios de Evaluación
8. Pautas de Presentación
9. Proceso de Selección y Cronograma

1. Introducción y Antecedentes

1.1 Propósito

[Nombre de la Organización] está buscando propuestas para una solución integral de software de etiquetado de datos para mejorar nuestra capacidad de crear datos de entrenamiento de alta calidad para modelos de aprendizaje automático. Esta RFP describe nuestros requisitos para un sistema robusto que respaldará nuestras iniciativas de ciencia de datos y aprendizaje automático.

1.2 Antecedentes de la Organización

[Incluir la siguiente información:]

- Breve descripción de su organización
- Industria y requisitos regulatorios específicos
- Tamaño de su organización y escala de operaciones de datos

- Procesos actuales de etiquetado de datos y desafíos
- Objetivos comerciales específicos que esta solución respaldará

2. Descripción General del Proyecto

2.1 Objetivos

Los objetivos principales de este proyecto son:

- Implementar una solución de etiquetado de datos escalable que pueda crecer con nuestras necesidades
- Mejorar la eficiencia y precisión de nuestros procesos de etiquetado de datos
- Soportar múltiples tipos de datos y métodos de anotación
- Mejorar la colaboración entre nuestros equipos de ciencia de datos y aprendizaje automático

2.2 Entorno Actual

[Describir su configuración actual:]

- Herramientas y procesos existentes de etiquetado de datos
- Desafíos y limitaciones actuales
- Volumen de datos que se procesan
- Número de usuarios/anotadores
- Requisitos de integración con sistemas existentes

3. Requisitos Técnicos

3.1 Arquitectura del Sistema

- Opciones de implementación en la nube o en las instalaciones
- Arquitectura escalable para manejar grandes conjuntos de datos y usuarios concurrentes
- Soporte para computación distribuida y procesamiento paralelo
- Diseño de infraestructura de alta disponibilidad
- Capacidades de balanceo de carga

3.2 Almacenamiento y Gestión de Datos

- Almacenamiento seguro de datos con cifrado en reposo y en tránsito
- Soporte para varios formatos de datos:
 - CSV, JSON, XML
 - DICOM para imágenes médicas
 - Formatos multimedia (imágenes, audio, video)
 - Documentos PDF
- Capacidades de versionado y respaldo de datos
- Procedimientos automatizados de respaldo y recuperación
- Seguimiento del linaje de datos

3.3 Capacidades de Integración

- API RESTful para integración perfecta
- Soporte para marcos de ML populares:
 - TensorFlow
 - PyTorch
 - Otras bibliotecas principales de ML
- Integración con soluciones de almacenamiento de datos:
 - Amazon S3
 - Azure Blob Storage
 - Google Cloud Storage
- Soporte para integraciones personalizadas vía API

3.4 Rendimiento y Escalabilidad

- Capacidad para manejar conjuntos de datos de al menos [X] TB de tamaño
- Soporte para el número especificado de usuarios concurrentes

- Requisitos definidos de tiempo de respuesta para:
 - Operaciones de carga de datos
 - Tareas de anotación
 - Búsqueda y filtrado
 - Operaciones de exportación

3.5 Seguridad y Cumplimiento

- Control de acceso basado en roles (RBAC)
- Integración de inicio de sesión único (SSO)
- Cumplimiento con estándares de la industria:
 - GDPR
 - HIPAA
 - SOC 2
- Registro y monitoreo de auditoría
- Estándares de cifrado de datos

3.6 Soporte de Navegador y Dispositivos

- Compatibilidad con múltiples navegadores:
 - Chrome
 - Firefox
 - Safari
 - Edge
- Capacidad de respuesta móvil para acceso desde tabletas y teléfonos inteligentes
- Soporte de pantalla táctil para tareas de anotación

3.7 Requisitos de Infraestructura

- Especificaciones del servidor

- Requisitos de red
- Requisitos de almacenamiento
- Infraestructura de respaldo
- Capacidades de recuperación ante desastres

4. Requisitos Funcionales

4.1 Tipos de Datos y Soporte de Anotación

Consejo: Al especificar los requisitos de anotación de datos, considere tanto las necesidades actuales como las futuras. Una solución robusta debe manejar múltiples tipos de datos y métodos de anotación, permitiendo la expansión a medida que evolucionan los proyectos. Preste especial atención a los requisitos de precisión y complejidad de anotación para cada tipo de datos.

Requisito	Sub-Requisito	S/N	Notas
Anotación de Imágenes	Capacidades de dibujo de cajas delimitadoras		
	Herramientas de anotación de polígonos		
	Soporte de segmentación semántica		
	Características de segmentación de instancias		
	Herramientas de anotación de puntos clave		
	Opciones de clasificación múltiple		
	Anotación de Video	Capacidad de anotación cuadro por cuadro	
Herramientas de seguimiento de objetos			
Características de segmentación temporal			
Soporte de seguimiento multi-objeto			
Gestión de línea de tiempo de video			

Anotación de Texto	Herramientas de reconocimiento de entidades nombradas		
	Capacidades de clasificación de texto		
	Características de análisis de sentimientos		
	Herramientas de etiquetado de documentos		
	Soporte multilingüe		
Anotación de Audio	Capacidades de transcripción		
	Herramientas de identificación de hablantes		
	Características de detección de eventos de sonido		
	Anotación basada en línea de tiempo		
	Visualización de forma de onda		
Anotación de PDF	Anotación a nivel de página		
	Capacidades de extracción de texto		
	Etiquetado de campos de formularios		
	Análisis de estructura de documentos		
Anotación DICOM	Visualización de imágenes médicas		
	Herramientas de marcado anatómico		
	Capacidades de medición		
	Navegación multicapa		

4.2 Etiquetado Asistido por IA

Consejo: El etiquetado asistido por IA puede mejorar significativamente la velocidad y consistencia de la anotación. Céntrese en soluciones que ofrezcan un equilibrio entre automatización y supervisión humana, con métricas

claras para medir las ganancias en precisión y eficiencia. Considere la adaptabilidad de los sistemas de IA a sus casos de uso específicos.

Requisito	Sub-Requisito	S/N	Notas
Integración de Algoritmos ML	Soporte de modelos preentrenados		
	Integración de modelos personalizados		
	Monitoreo de rendimiento del modelo		
	Capacidades de actualización del modelo		
Capacidades de Pre-etiquetado	Pre-anotación automatizada		
	Visualización de puntaje de confianza		
	Opciones de pre-etiquetado en lote		
	Herramientas de validación de pre-etiquetas		
Características de Aprendizaje Activo	Muestreo por incertidumbre		
	Gestión de cola de prioridades		
	Etiquetado asistido por modelo		
	Asignación dinámica de tareas		
Auto-sugerencia	Sugerencias inteligentes de etiquetas		
	Detección de casos similares		
	Reconocimiento de patrones		
	Sugerencias contextuales		

Entrenamiento de Modelos	Implementación de bucle de retroalimentación		
	Soporte de aprendizaje incremental		
	Seguimiento de métricas de rendimiento		
	Control de versiones del modelo		

4.3 Colaboración y Gestión del Flujo de Trabajo

Consejo: Las características de colaboración efectiva son cruciales para mantener la consistencia en grandes equipos de anotación. El sistema de gestión del flujo de trabajo debe ser lo suficientemente flexible para acomodar diferentes estructuras de proyecto mientras mantiene una supervisión clara y control de calidad.

Requisito	Sub-Requisito	S/N	Notas
Colaboración en Tiempo Real	Edición simultánea de usuarios		
	Seguimiento de cambios		
	Actualizaciones en tiempo real		
	Resolución de conflictos		
Asignación de Tareas	Herramientas de creación de proyectos		
	Sistema de distribución de tareas		
	Equilibrio de carga de trabajo		
	Gestión de prioridades		
Seguimiento del Progreso	Monitoreo de progreso en tiempo real		
	Seguimiento de tasa de finalización		

	Seguimiento de tiempo por tarea		
	Seguimiento de hitos		
Control de Versiones	Historial de cambios		
	Comparación de versiones		
	Capacidades de reversión		
	Registro de auditoría		
Personalización del Flujo de Trabajo	Proceso de revisión multinivel		
	Reglas de validación personalizadas		
	Plantillas de flujo de trabajo		
	Soporte de lógica condicional		

4.4 Control de Calidad

Consejo: Las herramientas de control de calidad deben proporcionar métodos de verificación tanto automatizados como manuales. El sistema debe admitir múltiples niveles de revisión y ofrecer métricas claras para medir la calidad de la anotación.

Requisito	Sub-Requisito	S/N	Notas
Herramientas de Revisión	Flujo de trabajo de revisión multinivel		
	Verificaciones automáticas de calidad		
	Sistema de asignación de revisiones		
	Mecanismos de retroalimentación		
Puntuación de Acuerdo	Métricas de acuerdo entre anotadores		
	Cálculo de puntuación Kappa		

	Análisis de desacuerdos		
	Evaluación comparativa de rendimiento		
Monitoreo de Validación	Métricas de calidad en tiempo real		
	Algoritmos de detección de errores		
	Alertas de umbral de calidad		
	Tendencias de rendimiento		
Gestión de Consenso	Implementación de modelo de consenso		
	Sistema de votación ponderada		
	Proceso de revisión de expertos		
	Flujo de trabajo de resolución de disputas		

4.5 Gestión y Organización de Datos

Consejo: Las capacidades robustas de gestión de datos son esenciales para mantener conjuntos de datos organizados y accesibles. El sistema debe proporcionar métodos eficientes para la organización, búsqueda y recuperación de datos mientras mantiene la integridad de los datos y el control de versiones.

Requisito	Sub-Requisito	S/N	Notas
Organización de Conjuntos de Datos	Gestión de estructura de carpetas		
	Etiquetado y categorización		
	Gestión de metadatos		
	Soporte de atributos personalizados		
Capacidades de Búsqueda	Filtros de búsqueda avanzada		
	Búsqueda de texto completo		

	Soporte de expresiones regulares		
	Plantillas de búsqueda guardadas		
Control de Versiones	Versionado de datos		
	Seguimiento de cambios		
	Comparación de versiones		
	Respaldo y restauración		
División de Conjuntos de Datos	División entrenamiento/prueba/validación		
	Ratios de división personalizados		
	Muestreo estratificado		
	Soporte de validación cruzada		

4.6 Análisis e Informes

Consejo: Las características de análisis e informes deben proporcionar perspectivas procesables para la gestión de proyectos y el control de calidad. Céntrese en capacidades de informes personalizables que puedan rastrear tanto métricas de proyecto de alto nivel como indicadores detallados de rendimiento.

Requisito	Sub-Requisito	S/N	Notas
Paneles de Rendimiento	Visualización de métricas en tiempo real		
	Creación de paneles personalizados		
	Visualizaciones interactivas		
	Capacidades de exportación		
Análisis de Anotadores	Métricas de rendimiento individual		
	Seguimiento de productividad		

	Métricas de calidad		
	Análisis de tiempo		
Métricas de Proyecto	Tasas de finalización de proyecto		
	Utilización de recursos		
	Seguimiento de costos		
	Análisis de cronograma		
Informes Personalizados	Creación de plantillas de informes		
	Informes programados		
	Definición de métricas personalizadas		
	Múltiples formatos de exportación		

4.7 Integración e Interoperabilidad

Consejo: Las capacidades de integración sólidas son cruciales para la incorporación perfecta en pipelines y flujos de trabajo de ML existentes. Céntrese en APIs estandarizadas y soporte para formatos de datos comunes.

Requisito	Sub-Requisito	S/N	Notas
Soporte de API	Disponibilidad de API RESTful		
	Soporte de GraphQL		
	Documentación de API		
	Controles de límite de tasa		
	Métodos de autenticación		
Integración de Marcos ML	Compatibilidad con TensorFlow		
	Soporte de PyTorch		

	Integración de marcos personalizados		
	Capacidades de importación/exportación de modelos		
Capacidades de Exportación	Soporte de formatos estándar		
	Plantillas de exportación personalizadas		
	Funcionalidad de exportación por lotes		
	Programación de exportación automatizada		
Integración de Terceros	Integración con pipeline CI/CD		
	Hooks de sistema de control de versiones		
	Integración de seguimiento de problemas		
	Conectividad de almacenamiento en la nube		

4.8 Características Avanzadas

Consejo: Las características avanzadas deben alinearse con las necesidades de escalabilidad futura y las tecnologías emergentes. Considere cómo estas características pueden proporcionar ventajas competitivas y mejorar la eficiencia de anotación.

Requisito	Sub-Requisito	S/N	Notas
Soporte de LLM	Creación de datos de ajuste fino		
	Herramientas de ingeniería de prompts		
	Soporte de evaluación de modelos		
	Optimización de conjuntos de datos		

Herramientas de Automatización	Automatización de flujo de trabajo		
	Procesamiento por lotes		
	Creación de pipelines personalizados		
	Acciones activadas por eventos		
Herramientas por Industria	Herramientas de imágenes médicas		
	Soporte de conducción autónoma		
	Características específicas de PLN		
	Soluciones industriales personalizadas		
Detección de Sesgos	Herramientas de análisis de sesgos		
	Métricas de equidad		
	Sugerencias de mitigación		
	Análisis demográfico		

5. Requisitos No Funcionales

5.1 Experiencia de Usuario

- Requisitos de interfaz intuitiva y amigable:
 - Estructura de navegación clara
 - Diseño de interfaz consistente
 - Interfaz web responsiva
 - Espacios de trabajo personalizables
- Requisitos mínimos de capacitación:
 - Tutoriales autoguiados
 - Ayuda sensible al contexto

- Consejos de herramientas y documentación
- Cumplimiento de accesibilidad:
 - Cumplimiento con WCAG 2.1
 - Soporte de lectores de pantalla
 - Navegación por teclado
 - Requisitos de contraste de color

5.2 Rendimiento

- Tiempos de carga rápidos para grandes conjuntos de datos:
 - Tiempo máximo de carga de página: [X] segundos
 - Tiempo máximo de respuesta: [X] segundos
 - Capacidades de procesamiento por lotes
- Herramientas de anotación receptivas:
 - Actualizaciones en tiempo real
 - Capacidades de dibujo fluido
 - Mínimo retraso en procesamiento de video
- Uso eficiente de recursos:
 - Límites de utilización de CPU
 - Optimización de memoria
 - Eficiencia de ancho de banda
- Capacidades de procesamiento del lado del cliente:
 - Computación basada en navegador
 - Funcionalidad sin conexión
 - Almacenamiento en caché local

5.3 Fiabilidad y Disponibilidad

- Garantía de tiempo de actividad del sistema: [X]%
- Procedimientos de respaldo y recuperación:
 - Programa de respaldo automatizado
 - Objetivos de tiempo de recuperación de datos
 - Opciones de recuperación punto en el tiempo
- Monitoreo del sistema:
 - Monitoreo de rendimiento
 - Seguimiento de errores
 - Análisis de uso
- Recuperación ante desastres:
 - Objetivos de tiempo de recuperación (RTO)
 - Objetivos de punto de recuperación (RPO)
 - Procedimientos de failover

5.4 Soporte y Mantenimiento

- Requisitos de documentación:
 - Manuales de usuario
 - Documentación de API
 - Guías de administración del sistema
 - Materiales de capacitación
- Soporte técnico:
 - Horarios y disponibilidad de soporte
 - Compromisos de tiempo de respuesta
 - Procedimientos de escalamiento de problemas
 - Canales de soporte (teléfono, correo electrónico, chat)

- Actualizaciones regulares:
 - Frecuencia de actualización
 - Control de versiones
 - Notas de lanzamiento
 - Compatibilidad hacia atrás

7. Criterios de Evaluación

7.1 Evaluación Técnica (40%)

- Arquitectura de la solución
- Completitud de características
- Capacidades de rendimiento
- Medidas de seguridad
- Capacidades de integración

7.2 Evaluación Funcional (25%)

- Funcionalidad principal
- Interfaz de usuario
- Capacidades de flujo de trabajo
- Características de informes
- Opciones de personalización

7.3 Evaluación del Proveedor (20%)

- Estabilidad de la empresa
- Experiencia técnica
- Metodología de implementación
- Capacidades de soporte
- Referencias de clientes

7.4 Evaluación de Costos (15%)

- Costo total de propiedad
- Competitividad de precios
- Términos de pago
- Valor por dinero
- Costos adicionales

8. Pautas de Presentación

8.1 Requisitos de Formato de Propuesta

- Resumen Ejecutivo (máximo 2 páginas)
- Propuesta Técnica (máximo 30 páginas)
- Enfoque de Implementación (máximo 15 páginas)
- Propuesta de Precios (documento separado)
- Perfil de la Empresa (máximo 10 páginas)
- Referencias (mínimo 3)

8.2 Documentación Requerida

1. Documentación Técnica:
 - Arquitectura de la solución
 - Especificaciones técnicas
 - Capacidades de integración
 - Características de seguridad
 - Métricas de rendimiento
2. Documentación de Implementación:
 - Cronograma del proyecto
 - Asignación de recursos

- Enfoque de capacitación
- Plan de gestión de riesgos
- Procedimientos de control de calidad

3. Materiales de Apoyo:

- Capturas de pantalla del producto
- Informes de muestra
- Documentación de API
- Casos de estudio
- Calificaciones del equipo

8.3 Instrucciones de Presentación

- Fecha límite de presentación: [Fecha y Hora]
- Número de copias requeridas: [X]
- Formato de presentación electrónica: PDF
- Tamaño máximo de archivo: [X] MB
- Método de entrega: [Correo electrónico/Portal/Entrega física]

8.4 Preguntas y Aclaraciones

- Fecha límite para preguntas: [Fecha]
- Persona de contacto: [Nombre]
- Dirección de correo electrónico: [Email]
- Distribución de respuestas: [Método]
- Conferencia previa a la propuesta: [Fecha si aplica]

9. Proceso de Selección y Cronograma

9.1 Proceso de Selección

1. Evaluación Inicial

- Verificación de cumplimiento
- Evaluación técnica
- Evaluación funcional
- Análisis de costos

2. Evaluación Detallada

- Demostración de la solución
- Análisis técnico profundo
- Verificación de referencias
- Entrevistas con el equipo

3. Selección Final

- Presentaciones de proveedores
- Negociación de contrato
- Decisión final
- Notificación de adjudicación

9.2 Cronograma del Proyecto

Hito	Fecha
Publicación de RFP	[Fecha]
Conferencia Previa a la Propuesta	[Fecha]
Fecha Límite para Preguntas	[Fecha]
Respuestas a Preguntas	[Fecha]
Fecha Límite de Propuestas	[Fecha]
Evaluación Inicial	[Fecha]
Demostraciones de Proveedores	[Fecha]

Verificación de Referencias	[Fecha]
Selección Final	[Fecha]
Negociación de Contrato	[Fecha]
Inicio del Proyecto	[Fecha]

9.3 Información de Contacto

Coordinador de RFP: [Nombre] [Título] [Organización] [Email] [Teléfono]

Contacto Técnico: [Nombre] [Título] [Email] [Teléfono]