



**INFORME TÉCNICO PREVIO DE EVALUACIÓN DE
SOFTWARE N° 037-GTI-2024- CGR**

ÍNDICE

| | |
|--|---|
| 1. NOMBRE DEL ÁREA | 3 |
| 2. NOMBRE Y CARGO DEL RESPONSABLE DE LA EVALUACIÓN | 3 |
| 3. FECHA | 3 |
| 4. JUSTIFICACIÓN | 3 |
| 5. ALTERNATIVAS | 3 |
| 6. ANÁLISIS COMPARATIVO TÉCNICO | 4 |
| 7. ANÁLISIS COMPARATIVO DE COSTO – BENEFICIO | 4 |
| 8. CONCLUSIÓN | 5 |
| 9. FIRMAS | 5 |

INFORME TÉCNICO PREVIO DE EVALUACIÓN DE SOFTWARE SUSTENTO TÉCNICO PARA LA ADQUISICIÓN DE SOFTWARE PARA EL RECONOCIMIENTO Y PROCESAMIENTO DE DOCUMENTOS NO ESTRUCTURADOS UTILIZANDO IA

1. NOMBRE DEL ÁREA

Gerencia de Tecnologías de la Información.

2. NOMBRE Y CARGO DEL RESPONSABLE DE LA EVALUACIÓN

Erik Bazán Flores – Subgerente de Sistemas de Información

3. FECHA

20 de febrero de 2024

4. JUSTIFICACIÓN

La Contraloría General de la República (en adelante La Contraloría), de acuerdo con sus objetivos estratégicos institucionales, se encuentra en la búsqueda constante de la excelencia, modernizando sus procesos internos e innovando con la implementación de nuevos modelos de control, los cuales se focalizan en la prevención y en la orientación a resultados.

En ese sentido, la institución recibe y accede a diversos documentos escaneados que sirven de insumo para los procesos de evaluación y fiscalización del Presunto Ingreso Patrimonial No Justificado (IPNJ) y Presunto Conflicto de Intereses, tales como: Estados de cuenta bancarios, Certificados de Vigencia de Poderes, Certificados Inmobiliarios, Partidas Registrales, Certificados Literales de Dominio de Registros Públicos - SUNARP, entre otros, motivo por el cual, el software a adquirir deberá convertir archivos digitales desde el formato PDF¹, entre otros, hacia otros formatos importantes, utilizando Inteligencia Artificial.

Por lo expuesto y el marco de Ley N° 28612 "Ley que norma el uso, adquisición y adecuación del software en la Administración Pública", se procede a evaluar el software de reconocimiento y procesamiento de documentos no estructurados utilizando IA.

5. ALTERNATIVAS

Considerando las necesidades del usuario de la Contraloría y las opciones tecnológicas que, en materia de reconocimiento y procesamiento de documentos no estructurados utilizando IA, están disponibles en el mercado, se han buscado alternativas de software tomando en consideración la disponibilidad en el servicio de atención y de soporte local.

En ese sentido, la búsqueda ha dado como resultado los productos que se listan a continuación:

- Datacap Cloud (no se ha evaluado para versión desktop Enterprise por la viabilidad en la comparación de precios)

¹ Formato estándar abierto, reconocido por la Organización Internacional para la Estandarización (ISO).

- Google Document AI
- Amazon Textract

6. ANÁLISIS COMPARATIVO TÉCNICO

El análisis técnico ha sido realizado en conformidad con la metodología establecida en la “Guía Técnica sobre evaluación de software en la administración pública” (R.M. N° 139-2004-PCM) tal como se exige en el reglamento de la Ley N° 28612.

a. Propósito de evaluación

Validar que las alternativas seleccionadas sean las más convenientes técnicamente para el uso de La Contraloría.

b. Identificar el tipo de producto

Software para convertir y editar archivos pdf a otros formatos importantes.

c. Identificación del modelo de calidad

Se aplicará el Modelo de Calidad de Software descrito en la parte I de la Guía de evaluación de software aprobado por Resolución Ministerial N° 139-2004-PCM.

d. Selección de métricas.

Las métricas fueron seleccionadas en base a las características técnicas descritas en el anexo N° 1 en ella se han evaluado atributos internos, externos y de uso.

Dada la criticidad en que los productos de software cumplan con los criterios técnicos requeridos, aquellos que, en la evaluación técnica, no alcancen 80 puntos del puntaje total como mínimo, no serán considerados para el análisis posterior de costo beneficio.

7. ANÁLISIS COMPARATIVO DE COSTO – BENEFICIO

El presente análisis tiene por objetivo seleccionar la mejor alternativa. Para lo cual se ha optado por dar un peso a la evaluación técnica de 0.8 y a la evaluación económica de 0.2, con el fin de garantizar que el software a adquirir o suscribir cumpla con los requerimientos técnicos solicitados.

- Del Análisis Comparativo Técnico del anexo N° 1, los productos de software que han resultado con puntaje igual o mayor a 80 puntos fueron considerados para el Análisis Comparativo de Costo – Beneficio, calificando los siguientes:
 - ✓ Amazon Textract
 - ✓ Google Document AI
- En el anexo N° 2, se muestra los resultados del Análisis Comparativo de Costo – Beneficio, así como el cuadro de Valoración Técnica – Económica.
- La evaluación de estas alternativas incluye los costos de Licencias o suscripción, los cuales son referenciales y obtenidos a través de las páginas web de cada fabricante. Ver anexo N° 3.

Asimismo, en la presente evaluación se ha considerado lo siguiente:

- **Hardware necesario para su funcionamiento de las alternativas**

La Gerencia de Tecnologías de la Información, a través de su Subgerencia de Operaciones y Plataforma Tecnológica ha determinado que no es necesaria la adquisición del hardware para el funcionamiento de los productos en mención.

- **Soporte y mantenimiento externo**

Con la adquisición o suscripción de las Licencias del Software avanzado de colaboración, correo y mensajería electrónica, se tienen derechos de soporte, actualizaciones de los parches y actualizaciones a versiones últimas liberadas por el fabricante durante el periodo de la garantía de los productos en mención.

- **Personal y mantenimiento interno**

La Contraloría cuenta con soporte de Mesa de Ayuda a cargo de la Gerencia de Tecnologías de la Información, para realizar el soporte relacionado incidentes con el equipo donde se instale el producto.

- **Capacitación**

El personal de las unidades orgánicas de La Contraloría, quienes utilizarán los productos evaluados, no tiene conocimiento en el uso y manejo de Amazon Textract y Google Document AI, sería necesaria una capacitación a nivel Administrador y Usuario Final.

8. CONCLUSIÓN

De los resultados del análisis realizado, se puede observar que el software Amazon Textract obtiene la mayor valoración Costo Beneficio; sin embargo, el producto de software Google Document AI, también cumple con los requisitos mínimos, satisfaciendo ambos con las necesidades de las unidades orgánicas de La Contraloría.

9. FIRMAS

Erik Bazán Flores
Subgerente de Sistemas de Información

**ANEXO N° 1
ANÁLISIS COMPARATIVO TÉCNICO**

| METRICAS (ATRIBUTOS) INTERNAS Y EXTERNAS | | | | | | | | |
|---|----------------------------|---|----------------|--------------------------|---------|-----------|--------------------|------------|
| Nº | Atributos | Descripción | Puntaje Máximo | Criterio de Calificación | Puntaje | Datacap | Google Document AI | Textract |
| 1 | Funcionabilidad | Capacidad de reconocer correctamente los caracteres en imágenes y documentos que incluyen anotaciones a mano. | 7 | Satisfactorio | 7 | 7 | 7 | 7 |
| | | | | Parcialmente | 4 | | | |
| | | Capacidad de reconocer diferentes tipos de letra, idiomas y formatos de documentos. | 7 | Satisfactorio | 7 | 7 | 7 | 7 |
| | | | | Parcialmente | 4 | | | |
| | | Convierte documentos escaneados en PDF de manera individual o por lotes de archivos. | 7 | Total | 7 | 7 | 7 | 7 |
| | | | | Parcialmente | 4 | | | |
| | | Capacidad de reconocer correctamente documentos con algún formato como facturas, guías, planos, contratos, reclamos y registros de transacciones. | 7 | Satisfactorio | 7 | 7 | 7 | 7 |
| | | | | Parcialmente | 4 | | | |
| | | Procesamiento y reconocimiento de Datos no estructurados como imágenes, texto, audio y vídeo. | 7 | Satisfactorio | 7 | 4 | 7 | 7 |
| | | | | Parcialmente | 4 | | | |
| Extracción de datos, clasificación de documentos, análisis de contenido, etc. | 7 | Satisfactorio | 7 | 4 | 7 | 7 | | |
| | | Parcialmente | 4 | | | | | |
| Capacidad de procesar grandes volúmenes de documentos. | 7 | Total | 7 | 7 | 7 | 7 | | |
| | | Parcialmente | 4 | | | | | |
| Funcionalidades de Minería de texto sobre los documentos digitalizados | 7 | Satisfactorio | 7 | 4 | 7 | 7 | | |
| | | Parcialmente | 4 | | | | | |
| 2 | Fiabilidad | Soporte local, telefónico, correo, entre otros | 6 | Total | 6 | 4 | 6 | 6 |
| | | | | Parcialmente | 4 | | | |
| 3 | Usabilidad | Facilidad de uso del software, incluyendo la interfaz de usuario y la curva de aprendizaje | 6 | Satisfactorio | 6 | 6 | 6 | 6 |
| | | | | Parcialmente | 4 | | | |
| 4 | Capacidad de mantenimiento | Se adapta a los cambios o mejoras de nuevas versiones | 6 | Satisfactorio | 6 | 6 | 6 | 6 |
| | | | | Parcialmente | 4 | | | |
| Sub Total | | | 74 | | | 63 | 74 | 74 |
| METRICAS (ATRIBUTOS) DE USO | | | | | | | | |
| 5 | Eficacia | Rendimiento promedio por encima del 85% | 7 | Total | 7 | 7 | 7 | 7 |
| | | | | Parcialmente | 4 | | | |
| 6 | Productividad | Tiempo promedio que tarda el software en convertir una imagen compleja en texto digital | 7 | Satisfactorio | 7 | 4 | 7 | 7 |
| | | | | Parcialmente | 4 | | | |
| 7 | Compleitud de contexto | Capacidad del software para identificar y extraer correctamente el contexto de una imagen o documento. | 6 | Satisfactorio | 6 | 6 | 6 | 6 |
| | | | | Parcialmente | 4 | | | |
| 8 | Satisfacción | Promedio del nivel de satisfacción | 6 | Satisfactorio | 6 | 6 | 6 | 6 |
| | | | | Parcialmente | 4 | | | |
| Sub Total | | | 26 | | | 23 | 26 | 26 |
| Total | | | 100 | | | 86 | 100 | 100 |

**ANEXO N° 2
ANÁLISIS COMPARATIVO COSTO BENEFICIO**

Información de costos obtenidos del mercado

| Costo de uso por mes | |
|---------------------------------|-------------------------------|
| Descripción del software | Costo referencial (\$) |
| Datacap Cloud | 2,500.00 |
| Google Document AI | 970.00 |
| Amazon Textract | 20.05 |

Equivalencia del costo

| Costo Mensual (\$) | Puntaje Equivalente |
|---------------------------|----------------------------|
| [0 - 100> | 100 |
| [100 – 1,000> | 85 |
| + de 1,000 | 50 |

Evaluación económica

| Software | Costo referencial mensual (\$) | Evaluación Económica |
|--------------------|---------------------------------------|-----------------------------|
| Datacap Cloud | 2500.00 | 50 |
| Google Document AI | 970.00 | 85 |
| Amazon Textract | 20.05 | 100 |

Análisis costo beneficio

| Software | Evaluación Técnica | Evaluación Técnica al 80% (A) | Evaluación Económica | Evaluación Económica al 20% (B) | Costo Beneficio (A+B) |
|--------------------|---------------------------|--------------------------------------|-----------------------------|--|------------------------------|
| Datacap Cloud | 86 | 64 | 50 | 10.0 | 74.0 |
| Google Document AI | 100 | 80 | 85 | 16.0 | 96.0 |
| Amazon Textract | 100 | 80 | 100 | 20.0 | 100.0 |

ANEXO N° 3
COTIZACIONES REFERENCIALES DE LICENCIAS DE SOFTWARE
En dólares americanos

a) Datacap Cloud

Costo Anual procesando 20,000 páginas al mes: \$ 2,500.00



Precio mensual de Datacap Cloud para procesar 20,000 páginas al mes

El precio mensual de Datacap Cloud para procesar 20,000 páginas al mes depende de varios factores, como el plan que elija, las funciones que seleccione y la cantidad de usuarios que necesitan acceder al software.

A continuación, se muestra un desglose de los costos para los diferentes planes de Datacap Cloud:

Plan Básico:

- **Precio inicial:** \$500 por mes
- **Costo mensual para procesar 20,000 páginas:** $\$500 + (\$0.0125 \text{ por página} * 20,000 \text{ páginas}) = \mathbf{\$750}$

Plan Profesional:

- **Precio inicial:** \$1,000 por mes
- **Costo mensual para procesar 20,000 páginas:** $\$1,000 + (\$0.0125 \text{ por página} * 20,000 \text{ páginas}) = \mathbf{\$1,250}$

Plan Enterprise:

- **Precio inicial:** \$2,000 por mes
- **Costo mensual para procesar 20,000 páginas:** $\$2,000 + (\$0.0125 \text{ por página} * 20,000 \text{ páginas}) = \mathbf{\$2,500}$

Fuente: <https://gemini.google.com/u/1/app/ca6339687a25f2ec> (acceso el 20FEB2024 16:13)

b) Google Document AI

Costo Anual procesando 20,000 páginas al mes: \$ 970.00



Precio mensual de Google Document AI para procesar 20,000 páginas al mes

El precio mensual de Google Document AI para procesar 20,000 páginas al mes se calcula de la siguiente manera:

Precio por página:

- **20,000 páginas × \$0.0125 por página = \$250**

Precio por alojamiento:

- **Estimación:** Se necesitan 20,000 páginas / 1,000 páginas por vCPU = 20 vCPU.
- **20 vCPU × \$0.05 por hora = \$1 por hora.**
- **\$1 por hora × 24 horas por día × 30 días por mes = \$720 por mes.**

Costo total:

- **\$250 (precio por página) + \$720 (precio por alojamiento) = \$970**

En resumen, el precio mensual de Google Document AI para procesar 20,000 páginas al mes sería de \$970.

Fuente: <https://gemini.google.com/u/1/app/ca6339687a25f2ec> (acceso el 20FEB2024 16:45)

c) Amazon Textract

Costo Anual procesando 20,000 páginas al mes: \$ 20.05



Precio mensual de Amazon Textract para procesar 20,000 páginas al mes

El precio mensual de Amazon Textract para procesar 20,000 páginas al mes se calcula de la siguiente manera:

Precio por página:

- **20,000 páginas × \$0.001 por página = \$20**

Precio por hora:

- **Estimación:** Se necesitan 20,000 páginas / 1,000 páginas por hora = 20 horas para procesar todas las páginas.
- **20 horas × \$0.0025 por hora = \$0.05**

Costo total:

- **\$20 (precio por página) + \$0.05 (precio por hora) = \$20.05**

En resumen, el precio mensual de Amazon Textract para procesar 20,000 páginas al mes sería de \$20.05.

Fuente: <https://gemini.google.com/u/1/app/ca6339687a25f2ec> (acceso el 20FEB2024 17:07)