

La revolución de la productividad

Cómo impacta la IAGen y ChatGPT en la reducción de tiempos y en la optimización de las tareas



Con el apoyo de:

Dirección

Juan G. **Corvalan**, Laura Cecilia **Díaz Dávila**,

Soledad **Guilera** y Enzo **Le Fevre**



Coordinación General

Mariana **Sanchez Caparros**

Giselle **Heleg**

Melisa **Raban**

Lautaro **Vasser**

Consultores Especialistas

Belen **Carreira**

Milagros **Etcheberry Le Fort**

Pamela **Tolosa**

Equipo de Investigación

Carina **Papini**

Florencia **Croci**

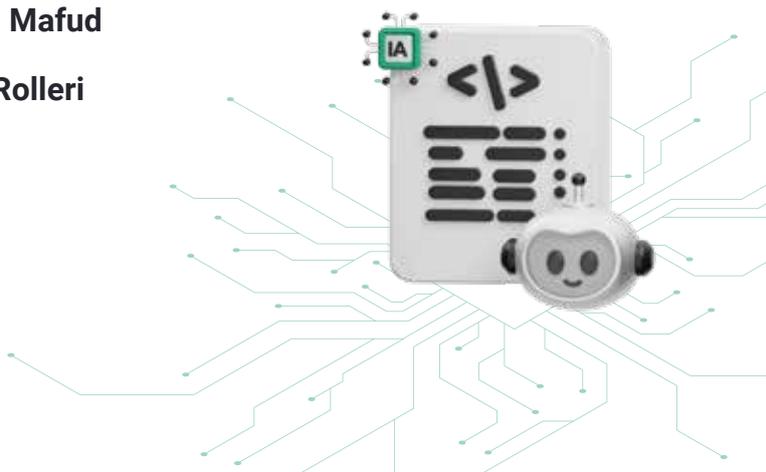
Margarita **Gonzalez**

Sol **Roschkavan**

Diseño gráfico

Victoria **Mafud**

Sofía **Rolleri**



Agradecimientos:

Este trabajo es el resultado de la colaboración de equipos de pruebas, que participaron activamente en la documentación de los casos de usos que forman parte de este documento, por ello agradecemos especialmente:

Juan **Alonso**, Silvina Ailén **Arias**, María Emilia **Arieau**, Belén **Arrua**, Ana M. **Basualdo**, Daniel Horacio **Bravo**, Santiago **Botero Ceballos**, Enrique **Carrara**, Sofía **Carrión Luna**, Daniela **Cevallos**, Alejandro **Corello**, Florencia Marina **Davesa**, Georgina **De Loredo Borrás** Javier **Echazu**, Rocío **Gallardo**, Gabriel **Gamarra**, Lucas Yamil **Julio**, Aylén **Katopodis**, Lucas **Monia**, Matías **Moyano**, Georgina **Natta**, Giannina **Nediani**, Franco **Orellana**, Marcela Blanca **Peltier**, Camila **Pentecoste**, Alejandro **Pietragalla**, Laura **Poletti**, María Victoria **Quiñones**, Marisa **Repetto**, Yeison Jairo **Rincón Rendón**, Darío **Robaina**, Maricel **Romero**, Marco **Rossi**, María Candela **Ruano**, Marisa **Salinas**, Silvina **Serra Ottonello**, Valeria **Stoffel**, Jorge Eduardo **Tolava**, Andrea **Tovar**, Natalia **Varela**, Laura **Viñas**, y Adolfo **Iriarte Yanicelli**.

Resumen ejecutivo

Aclaración previa

El presente documento, se refiere a la expresión Inteligencia Artificial Generativa - en adelante, IAGen-. La traducción en inglés es Generative Artificial Intelligence -GenAI-.

Presentación

El presente documento proporciona un análisis detallado sobre la aplicación de herramientas de IAGen en 83 tareas diferentes correspondientes a procesos de trabajo de distintos sectores. Diversos profesionales y personas realizaron estas tareas sin y con IAGen, bajo el enfoque metodológico elaborado por un equipo multidisciplinario convocado por el Laboratorio de Inteligencia Artificial de la UBA, Derecho (UBA IALAB)¹.

El estudio permite afirmar que, aún sin conocimientos previos sobre IAGen como ChatGPT, y previa realización de una selección criteriosa de la tarea a la cual aplicar esta u otra herramienta similar, es posible optimizar las tareas en distintos

¹ Guía metodológica para la evaluación del impacto de la IAGen en el empleo y las organizaciones, agosto de 2023, disponible en: <https://ialab.com.ar/wp-content/uploads/2024/02/Guia-metodologica-1.pdf>

trabajos.

Las mediciones iniciales muestran una mejora sustancial en los tiempos que se destinan a su realización. En 83 tareas de diversas áreas como educación, áreas legales, traducción, justicia y organismos públicos estatales, en promedio se obtuvo una reducción del **77%** del tiempo necesario para completarlas.

Hay algunos datos que surgieron de las pruebas realizadas que deben destacarse:

Eficiencia máxima

- **99,96%** fue la eficiencia máxima alcanzada por la herramienta de IAGen, en una prueba consistente en la elaboración de un borrador de traducción de un texto de 15.000 palabras.

Altos grados de eficiencia en otras áreas

- **97%** fue la eficiencia alcanzada con el uso de IAGen para la redacción de un contrato administrativo, en la administración pública.
- **95%** fue la eficacia alcanzada con el uso de IAGen para la realización de una copia con modificaciones de un contrato de arrendamiento, en un estudio jurídico.
- **90%** fue la eficiencia alcanzada con el uso de IAGen para identificar agravios en el recurso de inconstitucionalidad, en la justicia.

¿Asistente? ¿Complemento? o ¿reemplazo de la actividad humana?

Incidencia de la IAGen en las tareas

De las pruebas realizadas surge que la IAGen actúa como:

- Asistente, en el **59,03%** de las tareas.
- Complemento, en el **19,27%** de las tareas.
- Sustituto, en el **12,04%** de las tareas.

La incidencia de la IAGen según la complejidad de las tareas

A contrario de lo que podría suponerse, la IAGen aumentó la eficiencia en tareas de complejidad media y alta:

En tareas de alto nivel de complejidad, la IAGen incrementó la eficiencia en un **73%**.

En tareas de nivel medio de complejidad, la IAGen incrementó la eficiencia en un **81%**.

En tareas de bajo nivel de complejidad, IAGen incrementó la eficiencia en un **52%**.

Más allá de la experiencia en la confección de prompts, consideramos que el uso

de IAGen, como ChatGPT, Bard y Bing, aumentan la eficiencia y optimizan la realización de las tareas a gran escala, lo que supone un cambio de paradigma en la forma de trabajar de las personas.

Sin importar la distinción de qué tipo de organización hablamos, estamos frente a la herramienta más disruptiva que se ha inventado, en función de la escala, de la transversalidad que supone su aplicación y de la forma en que es posible obtener resultados a bajo costo.

I. Metodología

El presente estudio refleja los resultados de las bases metodológicas publicadas en la segunda edición del Tratado de Inteligencia Artificial y Derecho de Thomson Reuters-La Ley. Para la realización de la parte práctica de la presente investigación se cumplieron los siguientes pasos:

| Área laboral | Tareas seleccionadas | Tiempo CON IAGen (confección del prompt + ejecución con IAGen) | Tiempo SIN IAGen (incluye interacción con IAGen) |
|----------------------------------|----------------------|---|---|
| Administración Pública | 13 | 18.38 minutos | 62 minutos |
| Justicia | 29 | 21.12 minutos | 55 minutos |
| Estudio Jurídico y Áreas Legales | 27 | 30 minutos | 94 minutos |
| Traducción | 2 | 1.5 minutos | 2400 minutos |
| Educación | 4 | 7.5 minutos | 15.25 minutos |

1. Se convocó a referentes de las siguientes áreas laborales: administración pública, justicia, estudio jurídico y áreas legales, educación, traducción.
2. Para este primer abordaje, cada una de las áreas seleccionó una determinada cantidad de tareas, de distinta complejidad, relacionadas con su labor cotidiana, para realizar las pruebas y analizar la optimización a partir de la IAGen. En esencia:
3. Se realizaron las tareas sin IAGen y se midió el tiempo que insumió el proceso de ejecución de cada tarea.
4. Se realizaron las tareas con IAGen y se midió el tiempo que insumió el proceso

de interacción con la herramienta de IAGen hasta llegar al resultado deseado (aquí se incluyó el tiempo de confección del prompt por primera vez). Luego se midió el tiempo que insumió la realización total de la tarea con apoyo en la herramienta de IAGen seleccionada (con el prompt ya preparado).

La medición en cada caso tuvo en cuenta el total de tiempo de realización de la tarea, incluso el tiempo humano que se deba destinar para subtareas o microtareas específicas. Por ejemplo, la tarea puede consistir en contestar una demanda: si se usa ChatGPT para la fundamentación pero el humano debe copiar y pegar los párrafos elaborados por el modelo en un documento (microtarea), para una correcta medición, se computará también ese tiempo.

5. Se comparó el tiempo que insume la tarea con apoyo en una herramienta de IAGen y el tiempo que insume la tarea sin IAGen. La comparación permitió calcular la optimización.

III. Aclaraciones sobre la muestra elegida para las 83 tareas

- Las pruebas seleccionadas, corresponden a una muestra aleatoria. Las personas convocadas para realizar las pruebas tuvieron libertad para seleccionar las tareas y trabajaron con aquellas relacionadas con su labor cotidiana, con la condición de especificar los detalles de su composición, como su grado de automatización y complejidad.
- Algunas pruebas no alcanzaron los niveles de eficiencia esperados, lo cual denota la necesidad de que las personas se adapten al uso de las herramientas de IAGen y, en otros casos, da cuenta de que la IAGen no es la herramienta más adecuada para hacer más eficiente la tarea elegida (por ejemplo, para calcular plazos procesales).
- Las pruebas fueron realizadas por personas sin conocimientos previos en tecnología y sin experiencia relevante en el uso de ChatGPT en sus trabajos. Esto demuestra, al menos a priori, que la optimización puede crecer si las personas tienen formación específica en el uso de herramientas de IAGen, como ChatGPT.
- El presente estudio permite afirmar que aún sin conocimientos previos sobre ChatGPT, con una selección criteriosa de la tarea a la cual aplicar dicha

tecnología, es posible optimizar las tareas en los distintos trabajos.

- La realización de las pruebas permitió observar que el usuario necesita un tiempo prudencial para la generación de prompts detallados, contextuales y útiles para realizar la tarea elegida.
- Este documento constituye el primer paso de la investigación que se está llevando adelante desde UBA IALAB. Los resultados que aquí se presentan podrían modificarse. Principalmente, debido a que las pruebas no fueron repetidas por los responsables, lo que puede suponer una variación en los resultados aquí obtenidos

IV. Otros resultados preliminares

Se pueden elaborar prompts estandarizados como templates que generan resultados de utilidad, tanto para tareas, para subtareas más complejas, en términos humanos, como para subtareas más simples, repetitivas o estandarizadas.

Los prompts estandarizados como templates implican atajos para el uso de la IAGen. El tiempo de optimización mejora mucho cuando se copian y pegan los prompts y las iteraciones estandarizadas.

Es posible que los prompts estandarizados deban definirse para cada materia y tarea, aún cuando haya similitudes con las tareas que se realizan en relación a otras temáticas.

Los prompts estandarizados pueden arrojar resultados fallidos, con lo cual, se deben realizar los ajustes necesarios y controlar los resultados en todos los casos.

V. Algunas lecciones aprendidas al momento

A partir de todo el trabajo que realizamos desde UBA IALAB en relación al impacto de la IAGen en diversas tareas y los trabajos, podemos exponer las siguientes lecciones aprendidas, las cuales serán ampliadas en futuros informes:

- Las tareas sobre las cuales se desea aplicar IAGen deben ser seleccionadas dentro de cada institución y en relación a cada uno de los procesos

específicos.

- Los prompts deben realizarse como “trajes a medida” adaptados a las necesidades particulares de cada organización.
- Es importante detectar dentro de cada tarea, las subtareas y microtareas que la componen, y determinar aquellas en las que la IAGen puede ser de utilidad. Además, confirmar las hipótesis con pruebas y mediciones sobre ellas.
- La IAGen puede arrojar resultados de utilidad para optimizar la realización de las tareas, aún cuando se utilice por personas sin conocimientos tecnológicos, pero con gran conocimiento y manejo de su trabajo cotidiano.
- No todas las tareas resultan igualmente adecuadas para la automatización con IAGen. Se requiere una selección criteriosa de las actividades.
- Es esencial destinar tiempo a la elaboración del prompt inicial para describir el contexto y para exponer claramente las órdenes que el sistema debe ejecutar.
- A mediano plazo, si las instituciones se proponen aplicar IAGen a ciertos casos de uso, se aconseja contar, al menos, con una persona experta en el uso de estos grandes modelos de lenguaje generativos. Es importante que este perfil conozca las habilidades, descubra oportunidades y estén en condiciones de tomar ciertos atajos para llegar a los resultados deseados, lo que incluye la habilidad de generar prompts como templates para la reutilización.
- Resulta de utilidad medir el tiempo que insume la tarea con IAGen para calcular la optimización real que significa dentro de la carga laboral diaria, semanal o anual, y definir políticas de uso.
- Cuando se trabaja en casos de uso dentro de una organización, las mejoras pueden presentarse tanto en términos cuantitativos como en términos cualitativos. Las mejoras cuantitativas son las que describimos, en esencia y principalmente, en el presente informe y que se reflejan en disminuciones del tiempo necesario para la

realización de la tarea. Las mejoras cualitativas son aquellas que implican una mejora en la calidad del resultado que se espera de una determinada tarea. Estas últimas, pueden o no significar un ahorro de tiempo en la realización de la tarea, pero suelen ser de gran utilidad a la hora de aumentar la competitividad de la

organización.

Por ejemplo, si dentro de un estudio se usa Chat GPT para descubrir nuevos argumentos que complementan los modelos o templates ya existentes, probablemente aumentará el tiempo que insume la tarea, pero, sin dudas, aumentará la calidad del resultado ante el cliente.

- La incorporación de IAGen como apoyo para la realización de una tarea puede dar lugar a subtareas que antes no existían pero que son necesarias para lograr una real optimización de la tarea elegida. Este el caso de la subtaska de anonimización o seudonimización de datos personales contenidos en documentos de trabajo. Es necesario optimizar la realización de esas subtareas para lograr la mejora de tiempos totales.

VI. Próximos pasos para profundizar la investigación

Como la IAGen optimiza y reduce los tiempos de las tareas, especialmente las de mediana y alta complejidad, actualmente estamos abocados a contestar las siguientes preguntas que reflejan las líneas de investigación que estamos explorando a través de un equipo multidisciplinario:

- “¿Es posible elaborar y refinar diferentes tipos prompts para que funcionen como modelos o templates para optimizar aún más la realización de diversas tareas o subtareas?
- ¿Los prompts como templates pueden ser extrapolados a otros procesos o tareas?
- ¿Los prompts como templates implican una mejora radical en términos de optimización cuantitativa y/o cualitativa a mediano y largo plazo?
- ¿Se requiere que las personas se formen en el uso de la IAGen de manera óptima?
- ¿Es recomendable que un equipo se ocupe constantemente de enseñar, ajustar e interactuar con la IAGen de manera estratégica dentro de la organización?
- Continuamos trabajando en el análisis de los resultados obtenidos, así como en la ejecución de nuevas pruebas para confirmar o refutar nuestras hipótesis y confirmar si el nivel de optimización que aquí se logró, se mantiene, aumenta o disminuye en el tiempo. También estamos evaluando sí es posible y,

eventualmente cómo, mejorar el grado de optimización en tareas de menor complejidad.

- A partir de las actividades aquí realizadas, tomamos como caso de uso un Estudio jurídico, hallamos cinco subtareas en las que la IAGen puede tener impacto e intentamos optimizar su uso para que la mejora en la productividad sea significativa.
- En esencia, cada vez que hallamos un impacto, nos dedicamos a:
- Reiterar las pruebas para determinar si los resultados se mantienen en distintos casos similares.
- Mejorar los prompts iniciales para disminuir la cantidad de iteraciones, obtener el resultado esperado cada vez más rápido.
- Medir el tiempo que implica la mejora de los prompts para determinar si realmente vale la pena dedicarse a ello.
- Generar prompts refinados que funcionen como modelos o templates, para su uso en distintos casos similares.
- Medir el impacto de la IAGen en cada una de las pruebas y comparar el tiempo que implica la prueba, en cada una de las modificaciones del prompt.
- Comparar el tiempo que conlleva la realización de la tarea con ChatGPT y el tiempo que conlleva la tarea humanamente.
- A partir de la medición anterior, realizar el cálculo de optimización a mediano y largo plazo con prompts que funcionen como templates. Por ejemplo, la cantidad de contestaciones de demanda que se pueden generar con IAGen, en un tiempo determinado.
- Documentar las pruebas, los prompts y las iteraciones que funcionan como templates.

VII. Algunos resultados preliminares

1. Eficiencia según el nivel de complejidad de la tarea

- A. Tareas de alto nivel de complejidad: **73%** de eficiencia.
- B. Tareas de nivel medio de complejidad: **81%** de eficiencia.
- C. Tareas de bajo nivel de complejidad: **52%** de eficiencia.

Esto sugiere que la IAGen resulta más efectiva en tareas de complejidad media y alta en comparación con las tareas de baja complejidad.

2. Eficiencia según el nivel de juicio humano requerido para la realización de la tarea:

- A. Tareas que requieren alto juicio humano: **81%** de eficiencia
- B. Tareas que requieren medio juicio humano: **74%** de eficiencia
- C. Tareas que requieren bajo juicio humano: **44%** de eficiencia

Esto indica que la IAGen es particularmente efectiva en tareas que requieren un nivel alto y medio de juicio humano.

3. Eficiencia según el grado de repetitividad de la tarea

- A. Tareas altamente repetitivas: **86%** de eficiencia
- B. Tareas de repetitividad media: **88%** de eficiencia
- C. Tareas de baja repetitividad: **43%** de eficiencia

La IAGen parece ser más eficiente en tareas de repetitividad media y alta, lo cual es consistente con la idea de que la automatización y la IA son más efectivas en tareas rutinarias.

4. Eficiencia según la posibilidad de automatización de las tareas

- A. Automatizables: **94%** de eficiencia
- B. Semiautomatizables: **40%** de eficiencia
- C. No automatizables: **69%** de eficiencia

La eficiencia del **94%** en tareas automatizables indica que la IAGen es muy efectiva en tareas claramente definidas y estructuradas, lo que da cuenta de que la IAG puede generar beneficios significativos en términos de eficiencia y productividad en tareas que se prestan a la automatización completa.

5. Eficiencia según la incidencia de la IAGen en la tarea

- Complemento: Eficiencia promedio del **94%**
- Sustitución: Eficiencia promedio del **68%**
- Asistencia: Eficiencia promedio del **45%**
- Desplazamiento: Eficiencia del **27%**

En tareas donde la IAGen actúa como un complemento, se observa una alta eficiencia (**94%**), lo que sugiere que la IAGen puede mejorar o potenciar las capacidades humanas en ciertas tareas.

1. Administración Pública

| | |
|-----------------------------------|------------------------------|
| Total de tareas analizadas | 13 |
| Eficiencia | 70% |
| Nivel de complejidad | Alta: 2 tareas |
| | Media: 9 tareas |
| | Baja: 2 tareas |
| Grado de automatización | Automatizable: 1 tarea |
| | Semiautomatizable: 12 tareas |
| Juicio humano requerido | Alto: 3 |
| | Medio: 10 |
| Repetitividad | Alta: 7 tareas |
| | Media: 3 tareas |
| | Baja: 3 tareas |
| Incidencia de la IAGen | Asistencia: 11 tareas |
| | Sin asignar: 2 tareas |

| | |
|--|------------|
| Eficiencia por nivel de complejidad | Alta: 75% |
| | Media: 74% |
| | Baja: 87% |
| Eficiencia por nivel de juicio humano requerido | Alto: 31% |
| | Medio: 76% |
| Eficiencia por repetitividad | Alta: 87% |
| | Media: 75% |
| | Baja: 16% |

2. Justicia

| | |
|-----------------------------------|------------------|
| Total de tareas analizadas | 29 |
| Eficiencia | 62% |
| Nivel de complejidad | Alta: 11 tareas |
| | Media: 14 tareas |
| | Baja: 4 tareas |

Grado de automatización

Automatizable: 10 tareas

Semiautomatizable: 8 tareas

No automatizable: 11 tareas

Juicio humano requerido

Alto: 17

Medio: 2

Bajo: 8

Sin asignar: 2

Repetitividad

Alta: 16 tareas

Media: 7 tareas

Baja: 4 tareas

Sin asignar: 2 tareas

Incidencia de la IAGen

Asistencia: 19 tareas

Complemento: 3 tareas

Sustitución: 3 tareas

Sin asignar: 4 tareas

| | |
|--|------------|
| Eficiencia por nivel de complejidad | Alta: 77% |
| | Media: 38% |
| | Baja: 28% |
| Eficiencia por nivel de juicio humano requerido | Alta: 65% |
| | Medio: 57% |
| | Bajo: 44% |
| Eficiencia por repetitividad | Alta: 45% |
| | Media: 71% |
| | Baja: 67% |

3. Estudios jurídicos / áreas legales de empresas

| | |
|-----------------------------------|------------------|
| Total de tareas analizadas | 27 |
| Eficiencia | 68% |
| Nivel de complejidad | Alta: 9 tareas |
| | Media: 12 tareas |
| | Baja: 6 tareas |

Grado de automatización

Automatizable: 8

Semiautomatizable: 11

No automatizable: 8

Juicio humano requerido

Alto: 15

Medio: 9

Bajo: 3

Repetitividad

Alta: 11 tareas

Media: 5 tareas

Baja: 11 tareas

Incidencia de la IAGen

Asistencia: 10 tareas

Complemento: 9 tareas

Sustitución: 7 tareas

Sin asignar: 1 tarea

Eficiencia por nivel de complejidad

Alta: 73%

Media: 62%

Baja: 50%

| | |
|--|------------|
| Eficiencia por nivel de juicio humano requerido | Alto: 71% |
| | Medio: 57% |
| | Bajo: 44% |
| Eficiencia por repetitividad | Alta: 55% |
| | Media: 63% |
| | Baja: 72% |

4. Traducción

| | |
|-----------------------------------|-----------------------|
| Total de tareas analizadas | 2 |
| Eficiencia | 99.9% |
| Nivel de complejidad | Media: 2 tareas |
| Grado de automatización | Automatizable: 2 |
| Juicio humano requerido | Alto: 1 tarea |
| | Medio: 1 tarea |
| Repetitividad | Alta: 1 tarea |
| | Media: 1 tarea |
| Incidencia de la IAGen | Complemento: 2 tareas |

5. Educación

| | |
|----------------------------|-----------------------|
| Total de tareas analizadas | 4 |
| Eficiencia | 51% |
| Nivel de complejidad | Alta: 1 tarea |
| | Media: 3 tareas |
| Grado de automatización | No automatizable: 4 |
| Juicio humano requerido | Alto: 4 tareas |
| Repetitividad | Baja: 4 tareas |
| Incidencia de la IAGen | Complemento: 2 tareas |

.UBAderecho



IALAB

Con el apoyo de:

Fondazione
Adriano
Olivetti



Organización
de las Naciones Unidas
para la Educación,
la Ciencia y la Cultura



Cátedra UNESCO "Agua y Educación
para el Desarrollo Sostenible"
UNIVERSIDAD NACIONAL DEL LITORAL
FACULTAD DE INGENIERÍA Y CIENCIAS HÍDRICAS



LIDeSIA



Cámara
Argentina de
Comercio y Servicios