The Productivity Revolution

How GenAI and ChatGPT impact time reduction and task optimization

Direction

Juan G. **Corvalán**, Laura Cecilia **Díaz Dávila**, Soledad **Guilera** and Enzo **Le Fevre**

General Coordination

Mariana **Sanchez Caparros**

Giselle **Heleg**

Melisa **Raban**

Lautaro **Vasser**

Specialist Consultants

Soledad **Aguilera**

Belén **Carreira**

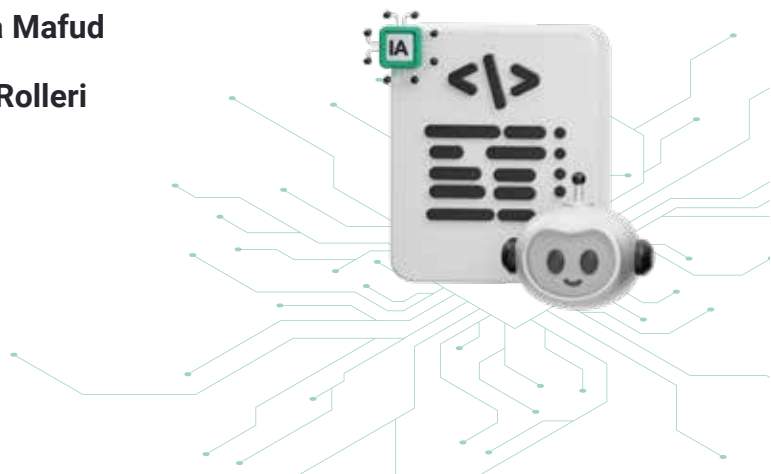## Research Team

Carina **Papini**

Florencia **Croci**

Margarita **Gonzalez**

Sol **Roschkavan**

Graphic Design

Victoria **Mafud**

Sofía **Rolleri**

Acknowledgements:

Executive Summary

Prior Clarification

This document refers to the expression Generative Artificial Intelligence as GenAI.

Presentation

This document provides a detailed analysis of the application of GenAI tools in 83 different tasks corresponding to work processes in different sectors. Various professionals and people carried out these tasks with and without GenAI, under the methodological approach developed by a multidisciplinary and intergenerational team convened by the Innovation and Artificial Intelligence Laboratory, School of Law, University of Buenos Aires (UBA IALAB)[1].

The study allows us to affirm that, even with a low level of prior knowledge about GenAI tools such as ChatGPT, and after carefully selecting the task to which to

---

[1] Methodological guide for evaluating the impact of GenIA on employment and organizations, August 2023, available at: https://ialab.com.ar/wp-content/uploads/2024/02/Guia-metodologica-1.pdf

apply this or another similar tool, it is possible to optimize the tasks in different jobs.

Initial measurements show a substantial improvement in the time necessary for its completion. In 83 tasks from various areas such as education, legal areas, translation, justice, and state public organizations, on average a 77% reduction in the time needed to complete them was obtained.

Some data that emerged from the performed tests should be highlighted:

Maximum Efficiency
- **99.96%** was the maximum efficiency achieved by the GenAI tool, in a test consisting of the creation of a draft translation of a 15,000-word text.

High Degrees of Efficiency in Other Areas
- **97%** was the efficiency achieved with the use of GenAI for the drafting of an administrative contract in the public administration.
- **95%** was the efficiency achieved with the use of GenAI to make a copy with modifications of a lease contract, in a law firm.
- **90%** was the efficiency achieved with the use of GenAI to identify grievances in the appeal on the grounds of unconstitutionality, in justice.
- **71%** was the effectiveness achieved with the use of GenAI for planning a practical activity in the education sector.

The Incidence of GenAI According to the Complexity of the Tasks

Contrary to what might be assumed, GenAI increased efficiency in tasks of medium and high complexity:

In tasks with a high level of complexity, GenAI increased efficiency by **73%**. This is a clear indication of how the IAG is evolving to address more sophisticated tasks that require a significant degree of insight or decision.

In tasks of medium level of complexity, GenAI increased efficiency by **81%**.

On low-level complexity tasks, GenAI increased efficiency by **52%**.

The Incidence of GenAI According to the Level of Repetitiveness of the Task

The particularly high efficiency (**86% and 88%**) in tasks with high and medium repetitiveness highlights the importance and usefulness of applying GenAI to automate and optimize processes.

The lower efficiency in low-repetitive tasks (**43%**) highlights the need to reevaluate how GenAI is used and consider alternatives or complements, such as greater human intervention to improve efficiency.

Assistant, Complement, or Replacement of Human Activity?

From the performed tests, it appears that GenAI acts as:

an Assistant, in **59.03%** of the tasks.

a Complement, in **19.27%** of the tasks.

a Replacement, in **12.04%** of the tasks.

Most tasks benefit from GenAI assistance, which accounts for a collaborative approach between humans and AI, rather than a complete replacement. Displacement in the hands of GenAI is minimal, indicating that GenAI currently acts more as a support than a replacement for the human workforce.

For tasks of greater complexity and requiring greater human judgment, GenAI acts more as a complement than a replacement. These tasks often require advanced skills, including critical thinking, analysis, and creativity, where human intervention remains essential.

The Incidence of GenAI According to the Skills Required to Perform the Task

For skills related to organizing and processing information, such as choosing the best way to approach tasks, reading competently, and understanding language, it has an average efficiency of **99.94%**.

For skills focused on the effective use of time, it shows an efficiency of **95%**.

For the collection, analysis, and organization of information (planning and organization), it shows an efficiency of **90%**.

For the skills of thinking creatively and abstractly, they reach an efficiency of **86.67%**.

In the skills that include analysis, text writing, execution, control, and management, it has an efficiency of **85.83%**.

Beyond the experience in creating prompts, we consider that the use of GenAI, such as ChatGPT, Bard, and Bing, increases efficiency and considerably optimizes the completion of tasks, which represents a paradigm shift in the way people work. Regardless of the type of organization we are referring to, we are faced with the most disruptive tool that has been invented, based on its impact and reach to a large number of users and its versatility in terms of the possibility of applications that its application entails, and how it is possible to obtain results at low cost.

I. Methodology

This research reflects the results of the Methodological Guide to evaluate the impact of GenAI on employment and organizations published in the second edition of the "Treaty of Artificial Intelligence", Thomson Reuters - La Ley. To carry out the practical part of the research, the following steps were completed:

Work area	Selected tasks	Time WITH GenAI (prompt creation + execution with GenAI)	Time WITHOUT GenAI (includes interaction with GenAI)
Public Administration	13	**18.38** minutes	**62** minutes
Justice	29	**21.12** minutes	**55** minutes
Legal Study and Legal Areas	27	**30** minutes	**94** minutes
Translation	2	**1.5** minutes	**2400** minutes
Education	4	**7.5** minutes	**15.25** minutes

1. Referents from the following work areas were invited: public administration, justice, legal studies and legal areas, education, and translation.
2. For this first approach, each area selected a certain number of tasks, of different complexity, related to their daily work, to carry out the tests and analyze the optimization based on the GenAI. In essence:

- In the area of public administration, 13 tasks were selected;
- In the area of justice, 29 tasks were selected;
- In legal studies and business legal areas, 27 tasks were selected;
- In translation, 2 tasks were selected;
- In education, 4 tasks were selected.

3. The tasks were performed without GenAI and the time taken by the execution process of each task was measured.

4. The tasks were carried out with GenAI and the time taken by the interaction process with the GenAI tool to reach the desired result was measured (here the time of creating the prompt for the first time was included).

Then, the time it took to fully complete the task was measured with the support of the selected GenAI tool (with the prompt already prepared).

The measurement in each case took into account the total time to complete the task, including the human time that must be allocated for specific subtasks or microtasks.

For example, the task may consist of answering a request: if ChatGPT is used for the substantiation but the human must copy and paste the paragraphs prepared by the model into a document (microtask), for correct measurement, that time will also be computed.

5. The time taken by the task with the support of a GenAI tool and the time taken by the task without GenAI were compared. The comparison allowed the optimization to be calculated.

II. Clarifications on the Sample Chosen for the 83 Tasks

The selected tests correspond to a random sample. The people called to carry out the tests were free to select the tasks and worked with those related to their daily work, with the condition of specifying the details of their composition, such as their degree of automation and complexity.

Some tests did not reach the expected levels of efficiency, which denotes the need for people to adapt to the use of GenAI tools and, in other cases, shows that GenAI is not the most appropriate tool to make more efficient the chosen task (for

example, to calculate procedural deadlines).

Some tests were carried out by people with no prior knowledge on technology and no relevant experience in using ChatGPT in their work, while others were carried out by students of the Postgraduate Course in Artificial Intelligence and Law at the UBA who, over 9 months, from the workshops and practical workshops, they acquired the necessary tools to apply GenAI in their daily work. This shows, at least a priori, that optimization can grow if people have specific training in using GenAI tools, such as ChatGPT.

The study allows us to affirm that even without prior knowledge about ChatGPT, with a careful selection of the task to which to apply said technology, it is possible to optimize the tasks in the different jobs.

Carrying out the tests allowed us to observe that the user needs a reasonable amount of time to generate detailed, contextual, and useful prompts to perform the chosen task.

This document constitutes the first step of the research that is being carried out by UBA IALAB. The results presented here could be modified. Mainly, because the tests were not repeated by those responsible, which may imply a variation in the results obtained here.

III. Other Preliminary Results

- Standardized prompts can be developed as templates that generate useful results, both for tasks, for more complex subtasks, in human terms, and for simpler, repetitive, or standardized subtasks.
- Prompts standardized as templates imply shortcuts for using GenAI. Time optimization is greatly improved when standardized prompts and iterations are copied and pasted.
- It is possible that standardized prompts should be defined for each subject and task, even when there are similarities with the tasks carried out in relation to other topics.
- Prompts can produce erroneous outcomes, therefore, the necessary adjustments must be made and the outcomes controlled in all cases.

IV. Some Lessons Learned So Far

Based on all the work we do at UBA IALAB on the impact of GenAI on various tasks and jobs, we can present the following lessons learned, which will be expanded on in future reports:

- Some tests did not reach the expected levels of efficiency, highlighting the need to get used to working with GenAI tools in specific contexts.
- The effectiveness of GenAI depends largely on the ability of users to interact and complement these tools.
- Not all tasks are equally suitable for automation with GenAI. A careful selection of activities is required. The tasks on which you want to apply GenAI must be selected within each institution and in relation to each of the specific processes to determine how and when to implement GenAI tools.
- The prompts must be made as "tailored suits" adapted to the particular needs of each organization.
- It is important to detect, within each task, the subtasks and microtasks that make it up, and determine those in which GenAI can be useful. Also, confirm the hypotheses with tests and measurements.
- GenAI can provide useful results to optimize the performance of tasks, even when used by people without technological knowledge, but with great knowledge and management of their daily work.
- Challenges are observed in implementing GenAI for tasks that require complex analytical, creative thinking, and management skills.
- It is essential to spend time preparing the initial prompt to describe the context and to clearly state the commands that the system must execute.
- In the medium term, if institutions intend to apply GenAI to certain use cases, it is advisable to have at least one expert in the use of these large generative language models. It is important that this profile knows the skills, discovers opportunities, and is able to take certain shortcuts to reach the desired results, which includes the ability to generate prompts as templates for reuse.
- It is useful to measure the time taken by the task with GenAI to calculate the

real optimization it means within the daily, weekly, or annual workload, and define usage policies.

- When working on use cases within an organization, improvements can be presented in both quantitative and qualitative terms.

- The quantitative improvements are those that we describe, essentially and mainly, in this report and that are reflected in reductions in the time necessary to complete the task.

- Qualitative improvements are those that imply an improvement in the quality of the result expected from a certain task. The latter may or may not mean saving time in carrying out the task, but they are usually very useful when it comes to increasing the competitiveness of the organization.

- For example, if ChatGPT is used within a study to discover new arguments that complement existing models or templates, the time taken by the task will probably increase, but the quality of the result for the client will increase as well.

- The incorporation of GenAI as support for carrying out a task can give rise to subtasks that did not exist before but are necessary to achieve real optimization of the chosen task.

- This is the case of the subtask of anonymization or pseudonymization of personal data contained in work documents. It is necessary to optimize the performance of these subtasks to achieve an improvement in total times.

V. Next Steps to Deepend the Investigation

As GenAI optimizes and reduces task times, especially those of medium and high complexity, we are currently focused on answering the following questions that reflect the lines of research that we are exploring through a multidisciplinary team:

- Is it possible to develop and refine different types of prompts so that they function as models or templates to further optimize the performance of various tasks or subtasks?

- Prompts as templates can be extrapolated to other processes or tasks?

- Does prompts as templates imply a radical improvement in terms of

quantitative and/or qualitative optimization in the medium and long term?

- Are people required to be trained in the optimal use of GenAI?

- Is it advisable for a team to constantly teach, adjust, and interact with GenAI strategically within the organization?

We continue working on the analysis of the results obtained, as well as on the execution of new tests to confirm or refute our hypotheses and confirm whether the level of optimization achieved here is maintained, increases, or decreases over time. We are also evaluating whether it is possible and, eventually how, to improve the degree of optimization in less complex tasks.

From the activities carried out here, we take a Legal Firm as a use case, we find five subtasks in which GenAI can have an impact and we try to optimize its use so that the productivity improvement is significant.

In essence, every time we find an impact, we are dedicated to:

- Reiterate the tests to determine if the results hold in different similar cases.

- Improve the initial prompts to reduce the number of iterations, obtaining the expected result faster.

- Measure the time involved in improving prompts to determine if it is worth the effort.

- Generate refined prompts that function as models or templates, to use them in other similar cases.

- Measure the impact of GenAI in each of the tests and compare the time involved in the test, in each of the prompt's modifications.

- Compare the time it takes to complete the task with ChatGPT and the time it takes to complete the task only with human interaction.

- Based on the previous measurement, perform the medium and long-term optimization calculation with prompts that function as templates. For example, the number of demand responses that can be generated with GenAI, in a given time.

- Document tests, prompts, and iterations that function as templates.

VI. Some Preliminary Results

1. **Efficiency according to the level of complexity of the task**
   - High level of complexity: **73%** efficiency.
   - Medium level of complexity: **81%** efficiency.
   - Low level of complexity: **52%** efficiency.

This suggests that GenAI is more effective in tasks of medium and high complexity compared to tasks of low complexity.

2. **Efficiency according to the level of human judgment required to perform the task:**
   - Tasks requiring high level of human judgment: **81%** efficiency.
   - Tasks requiring medium level of human judgment: **74%** efficiency.
   - Tasks requiring low level of human judgment: **44%** efficiency.

This indicates that GenAI is particularly effective in tasks that require a high and medium level of human judgment.

3. **Efficiency according to the degree of repetitiveness of the task**
   - Highly repetitive tasks: **86%** efficiency.
   - Medium repetitive tasks: **88%** efficiency.
   - Low repetitive tasks: **43%** efficiency.

GenAI appears to be more efficient in tasks of medium and high repetitiveness, which is consistent with the idea that automation and AI are more effective in routine tasks.

4. **Efficiency according to the possibility of automating tasks**
   - Automatable: **94%** efficiency.
   - Semi-automatable: **40%** efficiency.
   - Non-automatable: **69%** efficiency.

The 94% efficiency in automatable tasks indicates that GenAI is very effective in clearly defined and structured tasks, which shows that IAG can generate significant benefits in terms of efficiency and productivity in tasks that tend to a complete automation.

5. **Efficiency according to the impact of GenAI on the task**
   - Complement: Average efficiency of **94%**.
   - Replacement: Average efficiency of **68%**.
   - Assistance: Average efficiency of **45%**.
   - Displacement: **27%** efficiency.

In tasks where GenAI acts as a complement, high efficiency is observed (94%), suggesting that GenAI can improve or enhance human capabilities in certain tasks.

1. Public Administration

Total tasks analyzed	13
Efficiency	70%
Complexity level	High: 2 tasks
	Medium: 9 tasks
	Low: 2 tasks
Degree of automation	Automatable: 1 task
	Semi-automatable: 12 tasks
Human judgment required	High: 3
	Medium: 10
Repetitiveness	High: 7 tasks
	Medium: 3 tasks
	Low: 3 tasks
Incidence of GenAI	Assistance: 11 tasks
	Unassigned: 2 tasks

Efficiency by level of complexity	High: 75%
	Medium: 74%
	Low: 87%

| Efficiency by level of human judgment required | High: 31% |
| | Medium: 76% |

Repeatability efficiency	High: 87%
	Medium: 75%
	Low: 16%

2. Justice

Total tasks analyzed	29
Efficiency	62%
Complexity level	High: 11 tasks
	Medium: 14 tasks
	Low: 4 tasks

Degree of automation

- Automatable: 10
- Semi-automatable: 8
- Not automatable: 11

Human judgment required

- High: 17
- Medium: 2
- Low: 8
- Unassigned: 2

Repetitiveness

- High: 16 tasks
- Medium: 7 tasks
- Low: 4 tasks
- Unassigned: 2 tasks

Incidence of GenAI

- Assistance: 19 tasks
- Complement: 3 tasks
- Replacement: 3 tasks
- Unassigned: 4 tasks

Efficiency by level of complexity	High: 77%
	Medium: 38%
	Low: 28%
Efficiency by level of human judgment required	High: 65%
	Medium: 57%
	Low: 44%
Repeatability efficiency	High: 45%
	Medium: 71%
	Low: 67%

3. Legal Studies/Legal Areas in Companies

Total tasks analyzed	27
Efficiency	68%
Complexity level	High: 9 tasks
	Medium: 12 tasks
	Low: 6 tasks

Degree of automation	Automatable: 8
	Semi-automatable: 11
	Non-automatable: 8

Human judgment required	High: 15
	Medium: 9
	Low: 3

Repetitiveness	High: 11 tasks
	Medium: 5 tasks
	Low: 11 tasks

Incidence of GenAI	Assistance: 10 tasks
	Complement: 9 tasks
	Replacement: 7 tasks
	Unassigned: 1 task

Efficiency according to the level of complexity of the task	High: 73%
	Medium: 62%
	Low: 50%

Efficiency according to the level of human judgment required	High: 71%
	Medium: 57%
	Low: 44%
Efficiency according to the level of repetitiveness of the task	High: 55%
	Medium: 63%
	Low: 72%

4. Translation

Total tasks analyzed	2
Efficiency	99.9%
Complexity level	Medium: 2 tasks
Degree of automation	Automatable: 2
Human judgment required	High: 1 task
	Medium: 1 task
Repetitiveness	High: 1 task
	Medium: 1 task
Incidence of GenAI	Complement: 2 tasks

5. Education

Total tasks analyzed	4
Efficiency	51%
Complexity level	High: 1 task
	Medium: 3 tasks
Degree of automation	Non-automatable: 4
Human judgment required	High: 4 tasks
Repetitiveness	Low: 4 tasks
Incidence of GenAI	Complement: 2 tasks

.UBAderecho   IALAB

Supported by:

Fondazione Adriano Olivetti

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

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

UNIVERSIDAD NACIONAL DEL SUR — UNS BAHÍA BLANCA

LIDeSIA

Cámara Argentina de Comercio y Servicios — 100 años