



Artificial Inquiries

A VADEMECUM FOR WORKERS IN THE AGE OF AI

Artificial Inquiries¹

A Vademecum for Workers in the Age of AI

This book escaped our research lab and found its way into your hands. Whether it slipped through a crack or was smuggled out, we're glad. This book is about doing things. We made it as a companion for practice, and we didn't want it to be stuck in an article's appendix or gathering dust on a shelf.

At the time these lines are written, the world is saturated with two words: Artificial Intelligence (AI). Since the rise of Large Language Models (LLMs), popularized by ChatGPT, the topic has been discussed to death. But for all the talk of disruption, the horizon feels strangely settled. Our possible futures have narrowed down; promises and critiques alike have crystallized. AI will make us faster, more productive, more creative. AI will make us dumber, lonelier, easier to exploit. It will boost the economy, advance science, and automate the boring stuff. It will destroy our jobs, deepen inequalities, and hasten our planet's demise. The only remaining question is: which future are we heading toward?

Our intent is to help you find new, better questions. There's an urgent need to make room for alternative framings of AI, to move beyond predictions and look at the present. We need to stop talking about AI and begin the slow, silent work of noticing what it's already doing to our work. Only once we have paid attention to our practices can we reframe this technology on our own terms and choose how to engage with it. More importantly, observing how AI takes hold in our professional routines might bring into view the systems of labour that allowed it to happen. What's at stake is not just how we work with AI but who decides what work should look like.

We won't tell you what to think or do. Instead, we offer you a process, a protocol, to document closely how AI does — and could — affect your work. We call these "Artificial Investigations" because you'll be employing artificial methods to inquire into the role of AI in your professional practices. You'll purposefully slow down to observe, amplify subtle things you might usually dismiss, and follow constraints to sharpen your gaze. This book acts as an artificial lens to support your inquiry. We call it a vademecum, which in Latin means "Come with me!" We designed it as a companion to move with you through the everyday textures of your work: urgent deadlines, idle moments, shared spaces, and mindless scrolling. We don't know where you'll take the vademecum, but we trust it might lead you somewhere unexpected and worthwhile.

The rest of the protocol is up to you. Our recommendation is simple: don't go alone. So many AI practices are already confined to the personal, quietly folded into daily habits. Intentionally or not, we curate what we show from our friends, our colleagues, and our managers. The rest remains hidden, unspoken, and unquestioned. To bring them out of the shadows is to start collectively making sense of these technologies.

Artificial Inquiries²

In our original experiments, the vademecum functioned within a small cohort of professionals, where each individual would use it to give a genuine account of their AI practices, discussing them weekly or monthly. We suggest you do the same. Gather a small group of friends, colleagues, AI enthusiasts, or sceptics, meet regularly, share what you find, and compare notes. This isn't a self-help book; it's a device for collective inquiry.

The vademecum itself is a product of our own research collective. It was shaped by our diverse disciplinary inspirations (sociology, anthropology, design, Science and Technology Studies, linguistics), our common assumption that properly accounting for practices requires a lot of work (pages upon pages of it, as you can see), our focus on studying professionals in relatively privileged positions (lawyers, journalists, administrators, artists), and the usual pressures of academic work (securing funding, collecting data, writing papers). But now that the vademecum has left the lab, it might not suit you perfectly. Your time might be too scarce, the forms of AI you encounter may be different, and your concerns might lie elsewhere. Follow these instructions, adapt them, or start all over again.

We hope you'll take this book with you, and make it something of your own.

GA, DR, TP, ZDV / September 2025

Initiating

*Where should we begin inquiring
into LLM practices?*

Before you get into the exercises, you need to be briefed on two main subjects. The first one is the vademecum itself: What is it? Where does it come from? How can you use it? The second one offers some important principles to guide your inquiry into your LLM practices.

Understanding the *Vademecum* Brief 1

What is the Vademecum?

We designed this vademecum to be your companion for inquiring into your professional use of Large Language Models (LLMs). Inside you'll find a series of exercises aimed at helping you better understand your LLM practices and how they fit, reshape, or threaten your professional environment. As a physical, analogue resource, it deliberately stands apart from your digital work environment to create a breathing space within your day-to-day practices. We hope this will encourage you to pause, step back, and reflect on your use of LLMs.

This vademecum is now yours—or, at least, it will become yours over the following weeks. You'll fill it with your observations, descriptions, thoughts, and feelings about LLMs. Over time, it will evolve into a personalised journal, chronicling your journey with these tools. If you're utterly inexperienced with LLMs, feel free to experiment at your own pace. The first two exercises require minimal usage, giving you time to ease into a steadier routine with LLMs. However, before moving on to [Exercise 03 \(“Taking Stock”\)](#), you should aim to have had around fifty conversations with the LLM of your choice.

The Ecology of the Vademecum

Because we care about understanding LLMs within their real-world usage and professional ecologies, we find it fitting to provide an ecological overview of this vademecum's genesis. It originated from the Ecologies of LLMs Practices (EL2MP) research project, wherein we worked with several cohorts of professionals from the same discipline, each lasting 6 to 8 months. Each cohort completed the same exercises and met regularly (10 to 15 times) for individual interviews and researcher-led workshops. We later scanned and collected each cohort's vademecums to form our research corpus. You can learn more about the project in [Brief 04 \(“About EL2MP”\)](#).

The version currently in your hands is a distilled product of those experiments. While we first created it for our research protocols, we've reworked it so that it can be used to find meaning and purpose in all sorts of settings: personal study, academic research, groups of colleagues, or workers' communities—wherever you want to deepen your insight into your LLM practices.

Understanding the *Vademecum* Brief 1

What Can You Do With It?

You can disassemble and reassemble the vademecum as you see fit.

It is bound by two metal rods, allowing you to add, remove, or rearrange pages as needed. The metal rods are secured by small metal pieces. To open it, simply slide these pieces out of the way and unfold the long metal strips. When you're done rearranging the contents, slide the pages back onto the rods, fold the rods, and secure them by sliding the metal pieces back into place.

You can customize the vademecum to fit your needs, adding extra notes, drawings, or other materials as necessary; you'll find extra blank pages at the end. This is a space for documenting everything you experience, from immediate reactions to the deeper insights that build up over time. Feel free to draw, write, colour, or use whatever means you like to capture your thoughts and observations.

The exercises are grouped into blocks, each exploring a key question about LLM use. Feel free to adapt or skip some. However, skipping may require creativity to fill in gaps, as some exercises build upon previous ones.

Working With Others

We've found the vademecum is most valuable when it's shared or discussed collectively. You can simply talk about your results with a friend or colleague, or form a larger group where everyone works through the vademecum at the same pace.

Sharing Your Inquiries

If you'd like to participate in our research and share your Artificial Inquiries, please refer to Brief 03 ("Becoming a Co-Inquirer").

How to *Describe* Your Practices Brief 2

Guiding Principles

At various times, the vademecum invites you to write or talk about your work and your LLM practices. This is no easy feat. What we claim to be doing often differs from what we actually do. Work is messy and fragmented, composed of overlapping and sometimes contradictory processes that pull in different directions. When speaking about practice, coherence tends to be plastered where it didn't exist, or accounts may be transformed depending on what we think we should be doing rather than what we are, in fact, doing.

Here are some principles to guide you through this process:

- *Speak as a member of your profession.* Describe all the ways in which your larger profession (colleagues, norms, working conditions, expectations, obligations, ideals) shines through your personal practices. Find a balance between the individual and the collective. Your accounts shouldn't be so personal that they feel disconnected from your professional context, nor so broad that they sound like you're speaking on behalf of everyone in your field.
- *Describe your use of LLMs as you would for an outsider who knows nothing of your profession.* Regularly take time to step back and reflect on your current and past practices. Why did you do what you did? Why did you do it that way? How did you feel at the time? What were you thinking? Did you act without thinking? Speak in the first person. Refrain from impersonal descriptions of how AI works; shed light on what happens between you and the keyboard. Clarify implicit assumptions whenever possible.
- *Give LLMs a chance.* If you haven't already done so, incorporate LLMs into your daily work practices by using them regularly throughout the day and week. To criticise or praise LLMs, one should give them a fair test. Regardless of your current stance, be willing to navigate uncertainty. Pay attention to any discomfort and let it guide your inquiry. Don't pass judgment on LLMs; the idea is to explore how they fit — or don't fit — into what you usually do. What assumptions do you hold? How do the experiments challenge or reinforce your beliefs?
- *Write honest accounts grounded in practice.* Don't write general essays; start with the practices you document. Seek to contextualise, understand, and explain them. Your accounts can be risky, creative, humorous, or cynical — as long as they stem from an honest study of your practices. Include mundane and boring things as well as exceptional and exciting moments. Importantly, don't try to look your best. These experiments aren't academic exercises. There are no grades and no anticipated results. There are no right or wrong answers — only good accounts of what you did.

Qualifying

Where do LLMs live in our professional practices, norms, and relationships?

The Qualifying Block is your entry point into what constitutes the core of the vademecum: the exercises. The exercises in this first block invite you to evaluate your professional use of LLMs.

The exercises in this first block encourage you to explore your representations of LLMs and describe your actual LLM and work practices. They will also invite you to diversify your use of LLMs and reconsider how you could integrate them into your professional activities.

Ex¹ - *Draw It Like You See It*

Instructions

Sketch two different mental representations: one depicting how you visualise the functioning of an LLM and the other describing how you envision your work environment. No need to do additional research for these sketches, just draw off the top of your head based on your knowledge and perceptions.

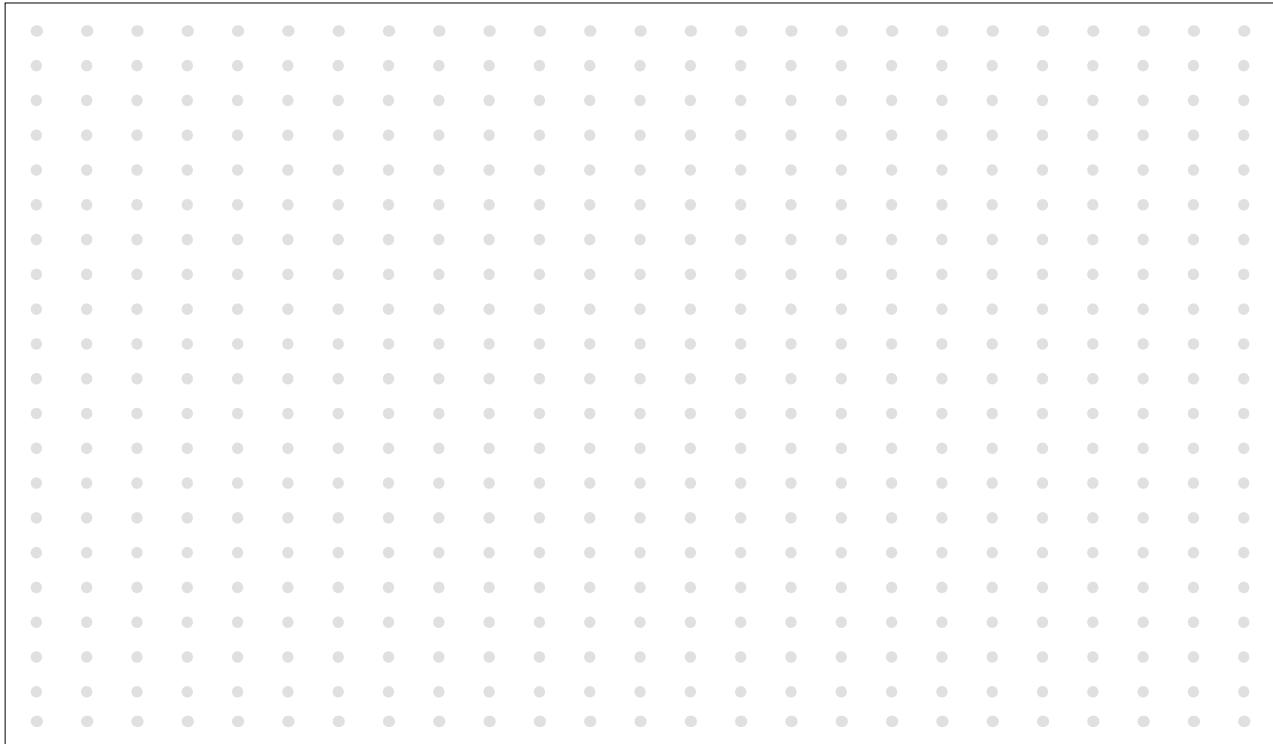
You have no limits other than the square boundary for your drawings.

Ex¹ - Draw It Like You See It

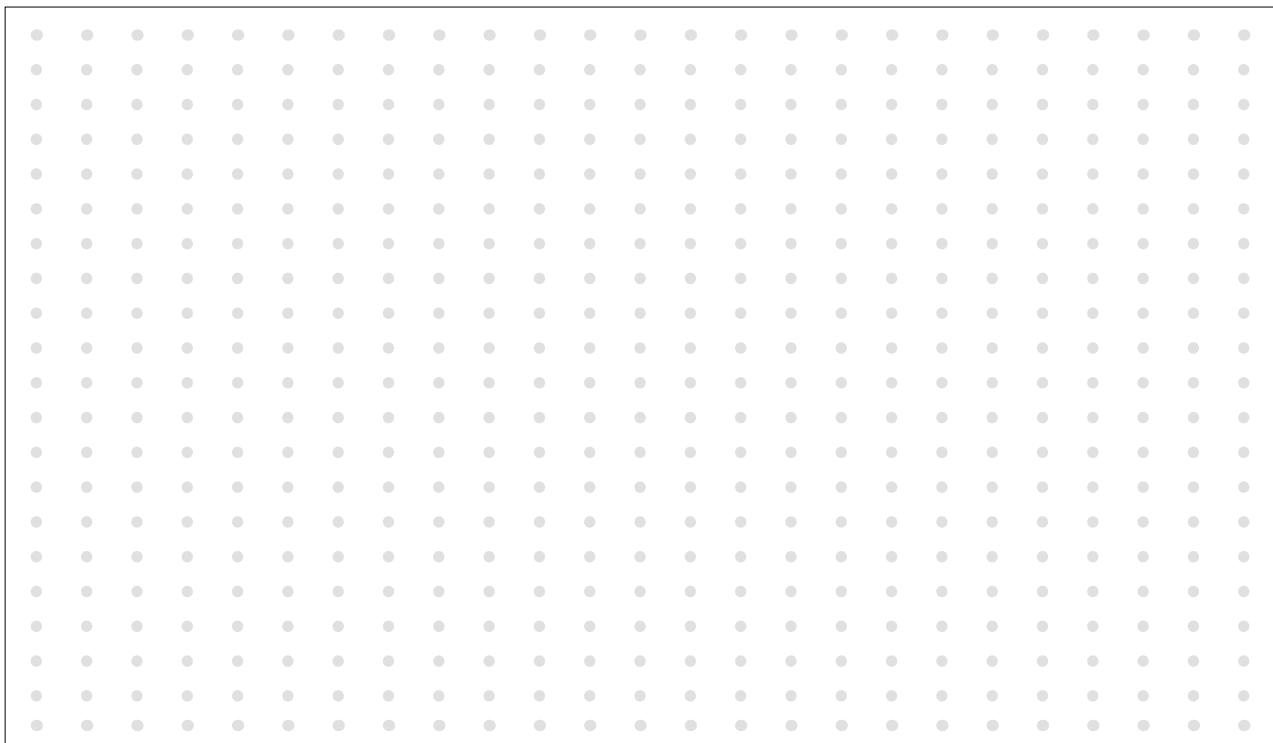
In a few words, what do you do?

Write here

What does your professional environment look like?



What does an LLM look like?



Ex²- *Harvesting Tasks*

Draw up two lists. The first captures the range of tasks you currently perform in your professional practice, and the second anticipates additional tasks you could perform with the help of an LLM. Doing this exercise in multiple sessions is best, as it will allow you time to think and add new ideas.

→

Ex^{2a} – All the *Things* You Do

Instructions

Draw up a list of tasks related to your professional practice. You probably perform many, many more tasks in your job than you realise. It is preferable to do this exercise in several sittings rather than just one, to get a clearer picture of the diversity of your tasks.

You can use the questions below to trigger your memory:

- What tasks do you dislike doing? Which ones do you enjoy?
- What tasks do you have to do regularly? Sometimes? Rarely?
- Are there some tasks that you periodically procrastinate on?
- Are there optional tasks that you still tend to do very often?
- In your field or profession, what tasks are basic?
- Which ones are necessary but tedious?
- Which ones are the mark of a good professional or an expert?
- Which tasks take the most time?
- Which ones cost you the most money?

The “Task ID” just serves as a reference to these tasks in later exercises. It is an arbitrary identifier, and you don’t need to take it into account in this exercise.

Ex^{2a} – All the *Things* You Do

Task List

Task ID	X5	Task
Task ID	S8	Task
Task ID	E9	Task
Task ID	H4	Task
Task ID	B5	Task
Task ID	U1	Task
Task ID	F0	Task
Task ID	I0	Task
Task ID	S6	Task
Task ID	S0	Task
Task ID	H9	Task
Task ID	B3	Task
Task ID	Q1	Task
Task ID	D2	Task
Task ID	A3	Task
Task ID	Z7	Task
Task ID	C9	Task

Ex^{2a} – All the *Things* You Do

Task List

Task ID	A7	Task
Task ID	K7	Task
Task ID	G1	Task
Task ID	X2	Task
Task ID	Z5	Task
Task ID	R0	Task
Task ID	L6	Task
Task ID	N5	Task
Task ID	P0	Task
Task ID	M8	Task
Task ID	Q0	Task
Task ID	B0	Task
Task ID	Q9	Task
Task ID	V5	Task
Task ID	G8	Task
Task ID	I1	Task
Task ID	R1	Task

Ex^{2b} – All the *Things* You Could Do

Instructions

Think about the tasks you currently do without the help of LLMs and where you could benefit from their assistance.

This list isn't about automation (delegating to the machine) but augmentation (doing more tasks or diversifying tasks). Think about activities or projects you haven't attempted yet or tasks you might hesitate to delegate.

Even if you suspect that an LLM might perform poorly on a given task, include it in your list: this exercise is about exploring potential opportunities rather than evaluating performance.

Ex^{2b} – All the *Things* You Could Do

Task List

Task ID	K0	Task	
Task ID	M4	Task	
Task ID	M7	Task	
Task ID	M0	Task	
Task ID	J6	Task	
Task ID	Z6	Task	
Task ID	O2	Task	
Task ID	C3	Task	
Task ID	L4	Task	
Task ID	K4	Task	
Task ID	H3	Task	
Task ID	G4	Task	
Task ID	T4	Task	
Task ID	T1	Task	
Task ID	B1	Task	
Task ID	H6	Task	
Task ID	G2	Task	

Ex^{2b} – All the *Things* You Could Do

Task List

Task ID	T2	Task	
Task ID	N3	Task	
Task ID	J3	Task	
Task ID	I8	Task	
Task ID	Z4	Task	
Task ID	E4	Task	
Task ID	I6	Task	
Task ID	C5	Task	
Task ID	B2	Task	
Task ID	F9	Task	
Task ID	U9	Task	
Task ID	I9	Task	
Task ID	K9	Task	
Task ID	A1	Task	
Task ID	F5	Task	
Task ID	P3	Task	
Task ID	O6	Task	

Ex³ – Taking Stock

Review the history of your interactions with your LLM (e.g., ChatGPT, Claude, LeChat) to take stock of your practices. Capture your overall impressions, identify patterns, and note memorable interactions.

This exercise is most useful if you have some experience using LLMs. If you still lack exposure to them, meaning that you've had fewer than 50 conversations, we recommend that you revisit it later.



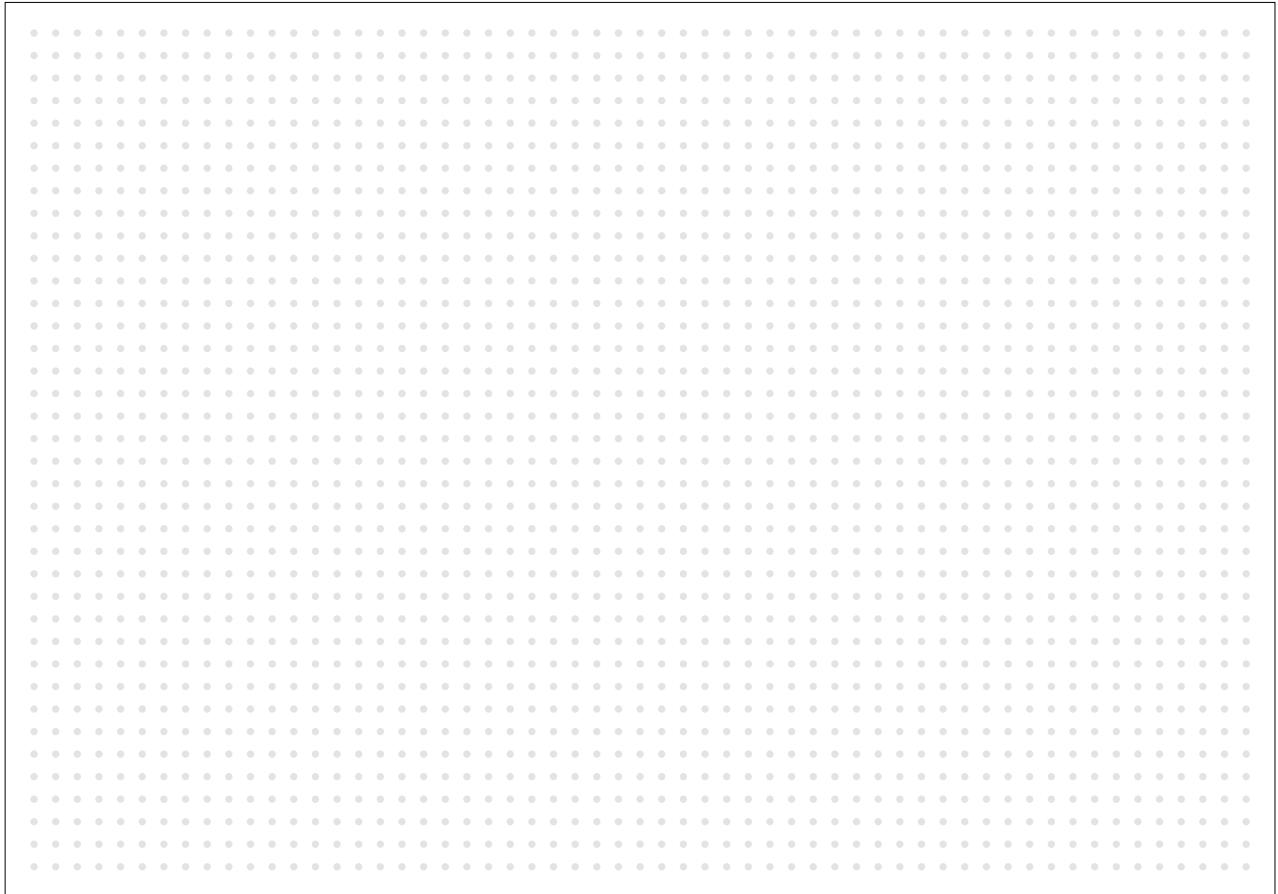
Ex^{3a} – Rough *Impressions*

Instructions

1. Set a timer and spend 5 minutes reflecting on your overall LLM use.
Don't look at your conversation history yet!
 - Assess which tasks you delegate, and assign a rough percentage to each (e.g., one-third of my tasks are translation, and the rest are drafting email answers). Consider how frequently you use the LLM (daily, weekly, or less frequently).
 - Reflect on the structure of your interactions and decide how many of your conversations are: short exchanges (a single request and answer); longer conversations (several iterations); or extended dialogue (conversations that last more than a day).
2. Finally, recall some memorable moments in your use of the LLM. Think back to two or three conversations that left a lasting impression.
 - Describe them briefly and why they were memorable.

Ex^{3a} - Rough *Impressions*

Sketch your impressions



Write down your impressions

Write here

Ex^{3b} – Rolling the Dice

Instructions

After having taken stock of these rough impressions, index your past interactions with the LLM of your choice. We assume you've had at least fifty conversations with an LLM.

Review these conversations from the most recent to the oldest using a dice. It spares you from reading through your entire history in detail.

Start by opening your primary LLM interface and locating your conversation history. The conversation titles should be a memory aid, but if the title is insufficient to jog your memory, click on the conversation and open it.

1. First, roll the dice. Number (X) indicates how many conversation titles you should review and how many conversations you should skip.
 - Let's say you roll a six: review the first six conversations, then skip the next six, and roll again.
2. Review the conversations according to the instructions below.
3. Continue this process until you have covered your entire conversation history.

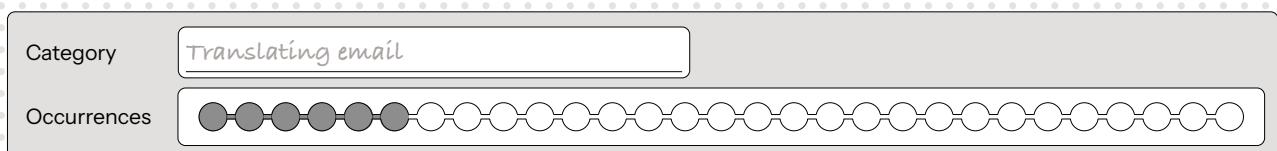
Assign each conversation a task category. If the conversation represents a new type of task, create a new category and tick off a box in the gauge. If it corresponds to an existing category, tick another box.

- Additionally, if a task stands out, record its title in your “Memorable Conversations” list. You can record a maximum of eight memorable conversations.

A conversation can be memorable for various reasons: for example, it might provoke a specific emotion (such as fear, excitement, frustration, or amusement); the LLM may have performed exceptionally well or poorly; it may have triggered a moment of moral or ethical dilemma; it could have shifted your perception of LLMs; or it might have had real consequences on your work.

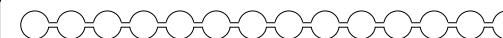
Example

Six conversations belong in the “translation” category:



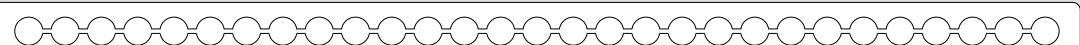
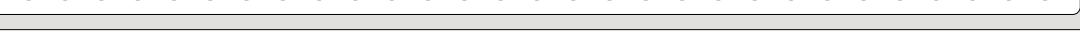
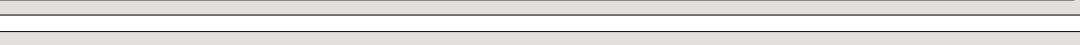
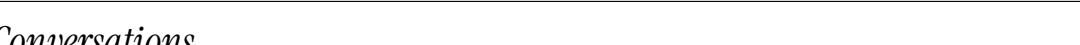
Ex^{3b} – Rolling the Dice

Task Categories

Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	

Ex^{3b} – Rolling the Dice

Task Categories

Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	
Category	<input type="text"/>
Occurrences	

Memorable Conversations

Title	<input type="text"/>

Ex^{3c} – The Story Told by Your *LLM*

Instructions

Looking back at your history with LLMs, answer the following open-ended questions based on your conversation history.

Follow the principles detailed in Brief 02 (“How to Describe Your Practices”).

Ex^{3c} – The Story Told by Your *LLM*

What do your LLM conversations capture accurately about your work?

Reflect on the narrative they shape and the stories they tell about
your professional practices, routines, and strategies.

Write here

Ex^{3c} – The Story Told by Your *LLM*

If someone were to look solely at your LLM conversations to understand your work, what aspects would be overlooked or misrepresented?

Reflect on the shadows, missing context, and the work that happens before and after using the LLM, which these conversations might not capture.

Write here

Ex^{3c} – The Story Told by Your *LLM*

What has the machine done for you?

Describe how the LLM has been helpful and valuable to your work.

Write here

Ex^{3c} – The Story Told by Your *LLM*

What have you done for the machine?

Reflect on the work you've provided and the time
you've invested to get good results out of the LLM.

Write here

Ex⁴ – Memorable Conversations

Instructions

This exercise invites you to revisit conversations that have left a lasting impression. To refresh your memory, refer to the conversation titles you noted in Exercise 03b (“Rolling the Dice”) and open them in your LLM interface. If you lack examples, you can return to your conversation history and add new conversations if necessary.

1. Start by copying the conversation title and briefly describing its content.
2. Then, explain why the conversation stood out. Reflect on its emotional impact: has it evoked excitement, frustration, amusement, fear, or any other strong emotion? Consider any moments that raised moral or ethical questions, shifted your views on LLMs, or had significant consequences for your work. Your descriptions should capture both your immediate feelings and the broader implications of the interaction.
3. Finally, shift your focus to the performance of LLMs. For each conversation, rate the performance of the LLM from one to five (one being very bad and five being excellent), and write a brief evaluation of the LLM’s performance. Then, explore why you believe the LLM performed as it did.

Example

<p>Conversation Title</p> <p>Translating email.</p> <p>Content</p> <p>This conversation was about translating emails for an economic summit in Paris.</p>	<p>Why it stood out</p> <p>I remember this conversation because GPT perfectly understood the language and word-choice needs at stake for each email it was asked to generate.</p>	<p>Describe the performance</p> <p>The LLM did poorly as its answers were very vague. It was also unable to remember things and hallucinated a lot.</p> <p>Score</p> <p>1 (Terrible) <input checked="" type="checkbox"/> X <input type="checkbox"/> + <input type="checkbox"/> + <input type="checkbox"/> + <input type="checkbox"/> + 5 (Excellent)</p>
---	---	--

EX⁴ - *Memorable* Conversations
*Playground*¹

EX⁴ - Playground¹



EX4

PULL
TO
OPEN
→

Conversations

Conversation Title

Content

Impact

Why it stood out

Performance

Describe the performance

Score

1	(Terrible)	+	+	+	+	+	5	(Excellent)
---	------------	---	---	---	---	---	---	-------------

Conversation Title

Content

Why it stood out

Describe the performance

Score

1	(Terrible)	+	+	+	+	+	5	(Excellent)
---	------------	---	---	---	---	---	---	-------------

Conversation Title

Content

Why it stood out

Describe the performance

Score

1	(Terrible)	+	+	+	+	+	5	(Excellent)
---	------------	---	---	---	---	---	---	-------------

Conversation Title

Content

Why it stood out

Describe the performance

Score

1	(Terrible)	+	+	+	+	+	5	(Excellent)
---	------------	---	---	---	---	---	---	-------------

EX⁴ - *Memorable* Conversations
*Playground*²

EX⁴ - Playground²



EX4

PULL
TO
OPEN
→

Conversations

Conversation Title

Content

Impact

Why it stood out

Performance

Describe the performance

Score

1	(Terrible)	+	+	+	+	+	5	(Excellent)
---	------------	---	---	---	---	---	---	-------------

Conversation Title

Content

Why it stood out

Describe the performance

Score

1	(Terrible)	+	+	+	+	+	5	(Excellent)
---	------------	---	---	---	---	---	---	-------------

Conversation Title

Content

Why it stood out

Describe the performance

Score

1	(Terrible)	+	+	+	+	+	5	(Excellent)
---	------------	---	---	---	---	---	---	-------------

Conversation Title

Content

Why it stood out

Describe the performance

Score

1	(Terrible)	+	+	+	+	+	5	(Excellent)
---	------------	---	---	---	---	---	---	-------------

Ex⁵ – Subtracting the Machine

In our everyday interactions with LLMs, we naturally tend to attribute our questions and the resulting answers to a dialogue between “ourselves” and the machine. We might say, “I asked Claude to write an email” or “I asked it to generate a summary”. What remains when you remove the labels “machine” “AI” “LLM” “Claude” or even “it” from your descriptions?

Revisit your most memorable conversations with the LLM and examine them closely. For each one, ask yourself: Which role does the machine perform in answer to your question? What about the role you adopt?

→

Ex^{5a} – A Thought Experiment

Instructions

1. Choose four different interactions — each consisting of your message paired with the machine’s response — and write the title of the conversation they were taken from.
2. To anchor your memory, copy the beginning of each interaction in the central “User” and “LLM” speech bubbles.
3. Instead of describing the interaction with usual terms, choose alternative words or roles. You might decide that your prompt could come from “a manager” or “a demanding client,” while the response from the machine could be characterized as coming from “an advisor,” “a colleague,” or “a consultant.”
 - Here are a few examples: a professional title (“an assistant”), an online service (“Wikipedia”), a corporation (“McKinsey”), or even an undefined entity (“a ghost”).
4. After your initial substitutions, evaluate whether the new labels accurately capture the essence of each exchange. If your descriptions feel too generic or are repeated across several interactions, refine them with one more round of subtraction.

The goal is not to find a “perfect” descriptor, but to explore how the re-description of your interactions transforms your understanding of both your role and that of the system.

Example

Conversation Title

Birthday celebration dish ideas

1

Excerpt

can you give me easy-to-make
dish ideas for a birthday cele-
bration with 60 people?

2

USER¹

A private individual

3

LLM¹

A friend with a knack for
cooking

USER²

An event planner

4

LLM²

A cooking website

USER³

A company happiness
manager

4

LLM³

A caterer

Ex^{5a} - A Thought Experiment

1st Conversation Title	2nd Conversation Title
Excerpt <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/>	Excerpt <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/>
USER ¹ <hr/> <hr/> <hr/> <hr/>	USER ¹ <hr/> <hr/> <hr/> <hr/>
LLM ¹ <hr/> <hr/> <hr/> <hr/>	LLM ¹ <hr/> <hr/> <hr/> <hr/>
USER ² <hr/> <hr/> <hr/> <hr/>	USER ² <hr/> <hr/> <hr/> <hr/>
LLM ² <hr/> <hr/> <hr/> <hr/>	LLM ² <hr/> <hr/> <hr/> <hr/>
USER ³ <hr/> <hr/> <hr/> <hr/>	USER ³ <hr/> <hr/> <hr/> <hr/>
LLM ³ <hr/> <hr/> <hr/> <hr/>	LLM ³ <hr/> <hr/> <hr/> <hr/>

Ex^{5a} - A Thought Experiment

3rd Conversation Title

Excerpt

USER¹

LLM¹

4th Conversation Title

Excerpt

USER¹

LLM¹

USER²

LLM²

USER²

LLM²

USER³

LLM³

USER³

LLM³

Ex^{5b} – What Remains After Subtraction

Instructions

You've now built a lexicon through a series of term subtractions. What do you think of the final terms? Do they adequately describe who is writing/speaking? Do some feel uncanny or weird to use? Why?

Consider whether the way you've redefined the speakers matches the tone of the interaction. For example, does the response read as if it were actually coming from a professional caterer?

Ex^{5b} – What Remains After Subtraction

Reflect on the final terms

Write here

Benchmarking

How do we assess the value of an LLM?

The Benchmarking Block guides you through the process of designing a tailored test to verify the relevance and usefulness of LLMs in your professional activities. The exercises challenge you to identify and select key tasks that represent the core of your work. By sorting tasks, reflecting on their professional significance, and evaluating whether they could benefit from the integration of LLMs, you'll design a comprehensive test.

In addition to creating a personal benchmark, this block offers you the opportunity to play and experiment with different LLMs, evaluate their differences, and determine which one is better suited to your professional practices.

Ex⁶ - Design Your AI Test

The four core tasks you select in this exercise are crucial because they represent the very heart of your professional practice. These tasks should accurately reflect your work. By isolating these four core tasks, you create a focused, meaningful AI test that can reveal both the strengths and the limitations of an LLM in handling the demands of your daily work.

The exercise is structured in three different parts: first, sorting all your potential tasks into a Venn diagram to categorize what you have to do versus what you enjoy doing; second, choosing the four tasks that best capture the diversity and significance of your work; and finally, reflecting on these choices to understand what each task reveals about your professional practices and what is at stake in delegating them to an AI.

→

EX^{6a} – Business and/or Pleasure?

Instructions

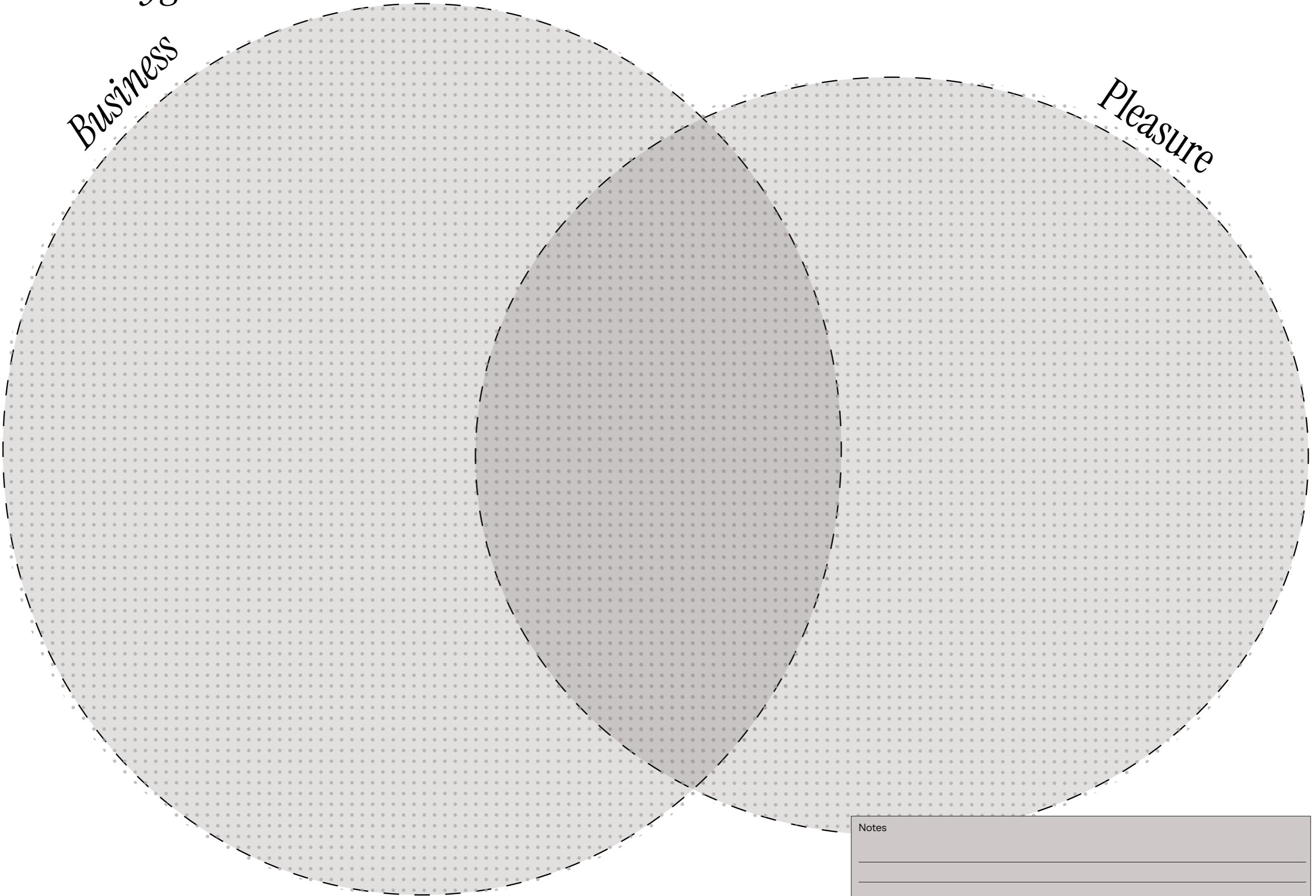
Refer back to [Exercise 02 \(“Harvesting Tasks”\)](#), and sort the tasks you currently do and those you might do in the future into three overlapping categories:

- The *tasks* you must perform for your job (obligations);
- The *tasks* you enjoy doing;
- The *tasks* that fall in between.

Complete the Venn diagram below using the task identifiers from [Exercise 02a \(“All the Things You Do”\)](#), for example A1, B2.

Ex^{6a} - Playground

EX6A
PULL
TO
OPEN
→



Ex^{6b} – Choosing Your *Core Four*

Instructions

Now, select four tasks to serve as the basis of your AI test.

Consider these three criteria:

- *Choose tasks that matter professionally.* These tasks should reflect your professional practices, whether they are frequent, occasional or rare.
- *Pick a diverse range of tasks.* Choose at least one task from each area of the diagram from [Exercise 06a \(have to, like to, both\)](#). Include at least one (ideally two) tasks you would usually hesitate to delegate to an LLM.
- *Select tasks fairly for the LLM.* You are allowed to choose a task on which you suspect the LLM will perform poorly. However, avoid including too many tasks that would be excessively challenging even for a skilled professional.

Keep in mind that these four tasks form the basis for selecting the model that best aligns with your work practices. Once you've made your selection, explain your reasoning

Ex^{6b} – Choosing Your *Core Four*

Identifying the Core Four

1 st Task ID	<input type="text"/>	Task	Business	Pleasure	Delegation	Complete Trust	Hesitation
2 nd Task ID	<input type="text"/>	Task	Business	Pleasure	Delegation	Complete Trust	Hesitation
3 rd Task ID	<input type="text"/>	Task	Business	Pleasure	Delegation	Complete Trust	Hesitation
4 th Task ID	<input type="text"/>	Task	Business	Pleasure	Delegation	Complete Trust	Hesitation

Judging the Core Four

Once your choice is made, explain your reasoning behind it.
Describe what you feel and what led you to make this decision.

Write here

Ex^{6c} – What's on the *Line*?

*Playground*¹

Instructions

Now that you've designed your test, step back and reflect on what each task reveals about your professional practices. Explore how these tasks connect to your work and what's on the line in your test.

For each of your four tasks, fill out the following template.

Ex^{6c} - Playground¹

Task ID

Task Fit

In what way does this task reflect a core part of your job or expertise? Is it something you do regularly or only on occasion? Is this task boring or interesting?

Write here

Professional Relevance

Does the task highlight any specific skills or crucial professional expertise? Are there any industry norms, professional codes, or expectations that this task should respect?

Write here

Personal Outcome

How will evaluating the LLM on this task guide you toward the model that suits your work best?

Write here

Rating the Result

Does this task need to be perfect, or is there room for “good enough”?

How will you decide if the LLM’s performance is sufficient to support you professionally?

Failure

What would be a deal-breaker in your field?

Write here

Good Enough

What would be a usable or helpful result?

Write here

Success

What would be a perfect result?

Write here

Assessing Your Expectations

How well do you think the LLM will perform on this task?

Circle the relevant option.

Terrible

Pretty bad

Good enough

Excellent

How sure are you of that answer?

Circle the relevant option.

Completely unsure

Unsure

Pretty sure

Confident

Ex^{6c} - What's on the *Line*?

*Playground*²

EX^{6c} - Playground²

EX6C

PULL
TO
OPEN
→

Task ID

Task Fit

In what way does this task reflect a core part of your job or expertise? Is it something you do regularly or only on occasion? Is this task boring or interesting?

Write here

Professional Relevance

Does the task highlight any specific skills or crucial professional expertise? Are there any industry norms, professional codes, or expectations that this task should respect?

Write here

Personal Outcome

How will evaluating the LLM on this task guide you toward the model that suits your work best?

Write here

Rating the Result

Does this task need to be perfect, or is there room for “good enough”?

How will you decide if the LLM’s performance is sufficient to support you professionally?

Failure

What would be a deal-breaker in your field?

Write here

Good Enough

What would be a usable or helpful result?

Write here

Success

What would be a perfect result?

Write here

Assessing Your Expectations

How well do you think the LLM will perform on this task?

Circle the relevant option.

Terrible

Pretty bad

Good enough

Excellent

How sure are you of that answer?

Circle the relevant option.

Completely unsure

Unsure

Pretty sure

Confident

Ex^{6c} – What's on the *Line*?

*Playground*³



EX^{6C} - Playground³

EX6C

PULL
TO
OPEN
→

Task ID

Task Fit

In what way does this task reflect a core part of your job or expertise? Is it something you do regularly or only on occasion? Is this task boring or interesting?

Write here

Professional Relevance

Does the task highlight any specific skills or crucial professional expertise? Are there any industry norms, professional codes, or expectations that this task should respect?

Write here

Personal Outcome

How will evaluating the LLM on this task guide you toward the model that suits your work best?

Write here

Rating the Result

Does this task need to be perfect, or is there room for “good enough”?

How will you decide if the LLM’s performance is sufficient to support you professionally?

Failure

What would be a deal-breaker in your field?

Write here

Good Enough

What would be a usable or helpful result?

Write here

Success

What would be a perfect result?

Write here

Assessing Your Expectations

How well do you think the LLM will perform on this task?

Circle the relevant option.

Terrible

Pretty bad

Good enough

Excellent

How sure are you of that answer?

Circle the relevant option.

Completely unsure

Unsure

Pretty sure

Confident

Ex^{6c} - What's on the *Line*?

*Playground*⁴

EX^{6C} - Playground⁴



EX6C

PULL
TO
OPEN
→

Task ID

Task Fit

In what way does this task reflect a core part of your job or expertise? Is it something you do regularly or only on occasion? Is this task boring or interesting?

Write here

Professional Relevance

Does the task highlight any specific skills or crucial professional expertise? Are there any industry norms, professional codes, or expectations that this task should respect?

Write here

Personal Outcome

How will evaluating the LLM on this task guide you toward the model that suits your work best?

Write here

Rating the Result

Does this task need to be perfect, or is there room for “good enough”?

How will you decide if the LLM’s performance is sufficient to support you professionally?

Failure

What would be a deal-breaker in your field?

Write here

Good Enough

What would be a usable or helpful result?

Write here

Success

What would be a perfect result?

Write here

Assessing Your Expectations

How well do you think the LLM will perform on this task?

Circle the relevant option.

Terrible	Pretty bad	Good enough	Excellent
----------	------------	-------------	-----------

How sure are you of that answer?

Circle the relevant option.

Completely unsure	Unsure	Pretty sure	Confident
-------------------	--------	-------------	-----------

Ex⁷ – Setting Up the *Test*

This exercise lays the groundwork of the test, by detailing the tasks and selecting the LLMs.

→

Ex^{7a} – Writing Guidelines

Instructions

Refer back to the four core tasks you've selected. Write detailed instructions for each of them. Keep in mind that whatever you write will be entered exactly as-is into the LLM. The prompt will be submitted once, with no opportunity to revise or continue the conversation afterwards.

Choose your words carefully, and include as much context or information as necessary.

You may also include documents (PDFs, images, text documents, and so on) if your task requires it.

- If it helps, think of these instructions as a detailed email you would write to a colleague who knows nothing of your job but who possesses sufficient skills and eagerness to accomplish the task adequately

Do this exercise in the digital format, as you'll need to copy / paste your instructions to send them out in Exercise 8 (“Gathering Evidence”).

Ex^{7b} – Choosing the Models

Instructions

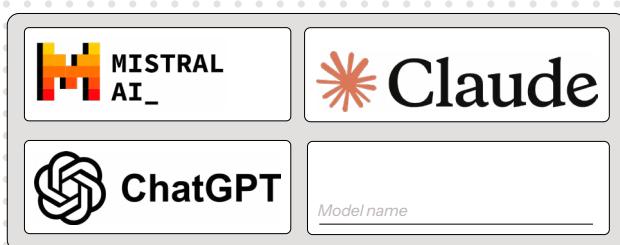
You've now chosen four core tasks and written detailed instructions for each. Four different models will execute your four core tasks. The three main models might be, for example, ChatGPT, Claude, and Mistral.

You may choose what you want to test for the fourth model.

You can choose either:

- Another general-purpose model (e.g. DeepSeek, Llama 3);
- Any domain-specific model that better suits your professional needs (e.g. journalism, law);
- A good friend or a trusted colleague whom you would ask to perform the tasks instead of a machine.

Write your choice down here



Ex⁸ – Gathering Evidence

Instructions

This exercise requires the help of a fellow, a trusted friend, or a colleague. This partner will be the one gathering the evidence on your behalf.

1. Begin by sharing the list of the four models you selected with your partner and ask them to assign each model an animal code name, choosing from a giraffe, a sloth, a dragon, or a turkey. These animals will serve as code names to ensure that when the test results are returned, you remain unaware of the models' identities and thus maintain impartiality.
2. Next, provide your partner with your list of tasks, including each task's ID, name, detailed instructions from [Exercise 07a \("Writing Guidelines"\)](#), and any added documents the models might require.
3. Your partner's role is simply to copy these instructions into the models and then collect the first answer each model provides. If you chose to test a human in [Exercise 07b \("Choosing the Models"\)](#), your partner will transfer the detailed instructions to them.
 - Each answer should be stored under the corresponding task ID and animal code name. For example, if ChatGPT (which was assigned the "sloth") produces a response for an email task labelled Z7, the response should be recorded as "Z7/sloth."
4. Since four models execute each task, your partner will eventually return sixteen labelled results.
 - You'll assess the quality of these results in the next exercise.

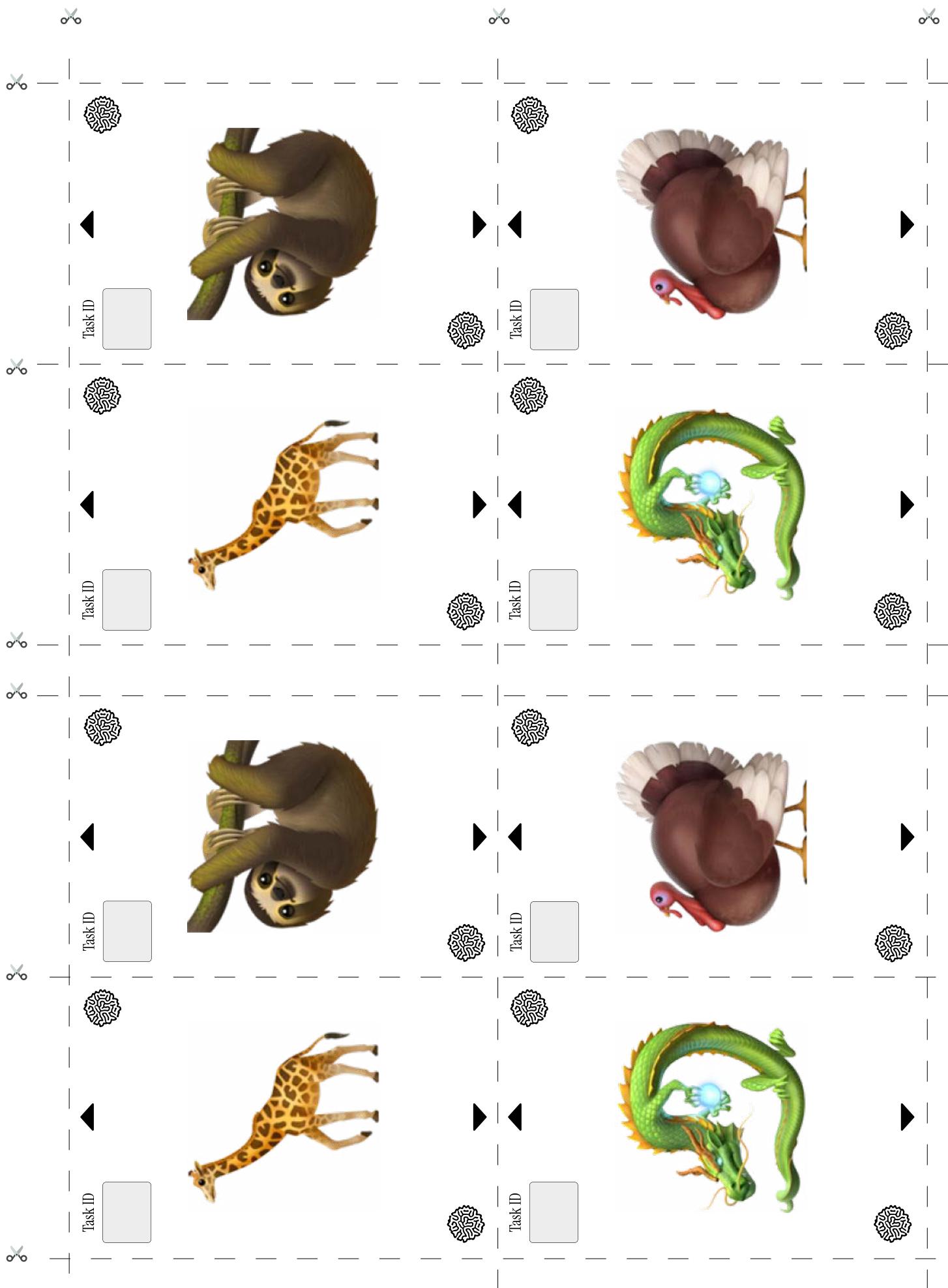


Ex⁹ - *Judgment Day*

Evaluate the outcomes of your four core tasks by judging and ranking the blinded, animal-coded responses provided by the models. Review these observations to render a final verdict explaining which model performed best. Finally, engage in a mock appeal process to critically reassess your test's design in terms of usefulness and fairness.

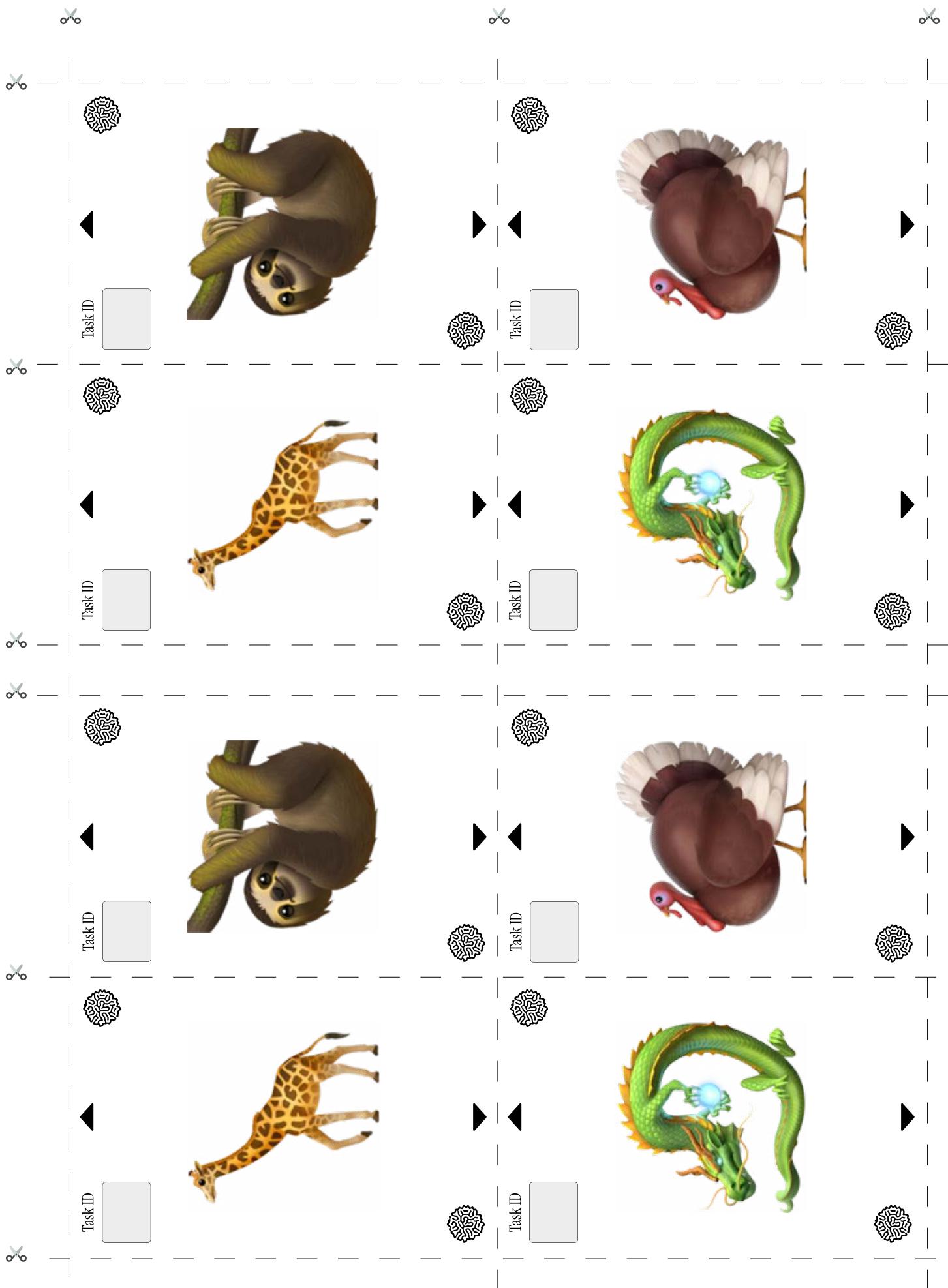
→

Ex⁹ - Cut Your Own Cards





Ex⁹ - Cut Your Own Cards





Ex^{9a} – Weighing the Evidence

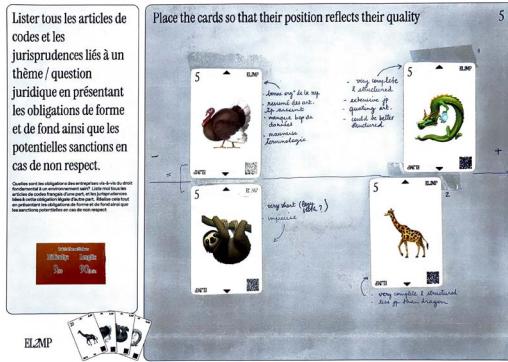
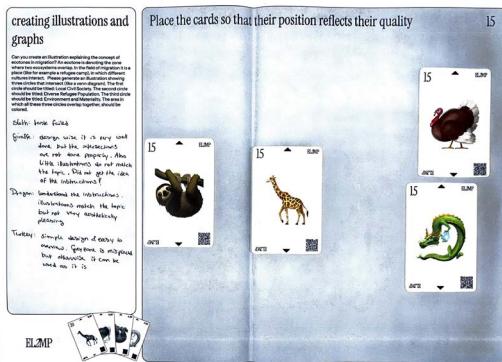
Instructions

Each task has its designated playground and deck of four cards which represent four different results from four models. To ensure a fair judgment, the results are blinded, and each model is represented by an animal (e.g., a giraffe 🦒, a dragon 🐉, a sloth 🐢, or a turkey 🦃).

1. Start by preparing the material. Enter the task IDs on the deck of cards and the playgrounds.
2. Then, acting as a judge, rank the four results obtained for each task. Place the four cards on the playground, arranging them from left to right, left representing the worst performance and right being the best.
 - You can use the space between cards to reflect the difference between each performance: cards with equivalent results can be stacked, while a card with a significantly better (or worse) result can be placed far from the others. Feel free to draw, write or take notes directly on the playgrounds.

You can refer to the scans of playgrounds completed by previous participants for additional guidance.

Examples



EX^{9a} - Weighing the *Evidence*
*Playground*¹

EX^{9a} - Playground¹



Task ID



Explain Your Ranking

Write here



Place the cards so that their position reflects their performance



+

|

EX^{9a} - Weighing the *Evidence*
*Playground*²

EX^{9a} - Playground²



Task ID



Explain Your Ranking

Write here



Place the cards so that their position reflects their performance



EX^{9a} - Weighing the *Evidence*
*Playground*³

EX^{9a} - Playground³



Task ID



Explain Your Ranking

Write here



Place the cards so that their position reflects their performance



Ex^{9a} - Weighing the *Evidence*
*Playground*⁴

EX^{9a} - Playground⁴



Task ID



Explain Your Ranking

Write here



Place the cards so that their position reflects their performance



Ex^{9b} – The *Verdict Is In*

Instructions

Review your observations and render your final judgment.

After carefully ranking the responses for each task using the blinded, animal-coded results, write a comprehensive verdict explaining which model, identified by its animal code name, performed best overall and why.

Your verdict should be substantiated, including examples from the rankings, and clearly outline the criteria you used to make your judgment.

Follow the principles detailed in Brief 02 (“How to Describe Your Practices”).

Ex^{9b} - The *Verdict Is In*

Explain Your Judgement

Write here

Ex^{9c} – Revisiting the Ruling

Instructions

Through a mock appeal process, produce a critical evaluation of the test you designed. Using the guide below, ask questions about your AI test. The assessment focuses on two main dimensions:

- *Usefulness*: assessing whether the test accurately reflected the professional relevancy of LLM usage.
- *Fairness*: examining whether the test was designed in a way that gave the LLM a reasonable chance to perform well.

Finally, decide if there are *sufficient grounds to appeal* following your initial ruling on the test's effectiveness.

Ex^{9c} - Revisiting the Ruling

Professional Usefulness

Criteria	Useful	Somewhat Useful	Useless
<p><i>Task Relevance</i></p> <p>Did the test include tasks that are representative of your core professional responsibilities?</p>			
<p><i>Task Groundedness</i></p> <p>Did the test accurately simulate real-world conditions where you would typically use an LLM?</p>			
<p><i>Task Impact</i></p> <p>Were the selected tasks ones that, if successfully delegated to an LLM, would make a substantial difference in your workflow (e.g., by saving significant time, reducing effort, or improving results)?</p>			

Other Remarks

Write here

Ex^{9c} - Revisiting the Ruling

Fairness Assessment

Criteria	Fair	Somewhat Fair	Unfair
<i>Task Difficulty</i> Were the tasks designed in such a way that the AI had a reasonable chance of performing well, based on its known strengths and limitations?			
<i>Clarity of Instructions</i> Were the instructions for the tasks clear, specific, and consistent enough to allow the AI to perform at its best?			
<i>Task Variety</i> Did the test include a variety of tasks (in terms of type and complexity) that allowed the AI to demonstrate its abilities across different domains, without favouring tasks inherently better suited to humans?			

Other Remarks

Write here

Ex^{9c} - Revisiting the Ruling

Final Ruling

Tick one of the boxes

Sufficient ground for appeal	Insufficient ground for appeal

What would you change if you had to redo the AI test?

Write here

Prompting

Does prompt engineering reconfigure our work relationship with the LLM?

The Prompting Block questions what it means to “get better” at using LLMs. You’ll learn about various prompting techniques, drawn from the field of prompt engineering, to yield more optimized outcomes. Rather than accepting these techniques at face value, these exercises encourage you to engage them critically. Be mindful of their effects on the machine and yourself.

This block also marks the halfway point in the exercises, and thus grants you the opportunity to reflect on what you’ve already accomplished and what this journey has meant so far.

Ex¹⁰ – The Art of the Prompt

Instructions

When learning something new — painting a wall, preparing a meal, or crafting an object — we often turn to tutorials. Whether in the form of a video, a written guide, or step-by-step instructions, tutorials promise clarity and simplicity: follow these steps carefully, and the result should be perfect.

Yet, in practice, tutorials rarely deliver exactly what they promised. The wall you painted ends up looking patchy, your dish tastes different, and your craft project has uneven seams. Why? Perhaps you missed subtle but critical instructions, your tools or materials differed slightly, or the tutorial itself was misleading or incomplete.

- Working with LLMs presents a similar challenge. We are often told that improving our interactions with these tools depends on crafting better prompts which provide clearer instructions, more details, or specific phrasing.

But just like with tutorials, improving prompts doesn't always yield the expected outcome. Did your prompt genuinely improve the LLM's response, or did it merely shift your expectations? Was the LLM unable to grasp your intent, or did your request lack precision?

This exercise allows you to experiment with a range of prompting techniques. Test them in your everyday practices, reflect on their effectiveness, and examine how these methods impact not only the LLM's responses but also your own workflow, assumptions and feelings.

Ex^{10a} – Room for *Improvement*

Instructions

1. Begin by revisiting your work from the Benchmarking Block.
2. Choose one task where an LLM underperformed or delivered disappointing results.
3. Carefully reread your original prompt and the associated material, especially Exercise 06c (“What’s on the Line?”) and the playground from Exercise 09a (“Weighing the Evidence”).
4. Clearly specify the task ID and the name of the LLM that was used, the exact issues with the results, and the potential reasons behind the unsatisfactory performance.

Fill In Here

Task ID

Type of task

Model used

What were the problems with the result you got?

Why did the LLM perform the way it did?

What went wrong?

Write here

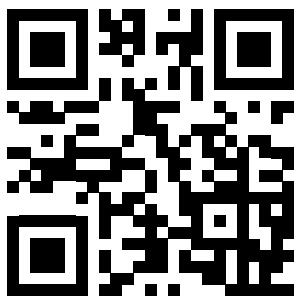
Ex^{10b} – Learning Your *Craft*

Instructions

We now turn to various prompting techniques designed to improve interactions with LLMs. We've curated a set of resources and tutorials for you to learn these methods.

Tutorials

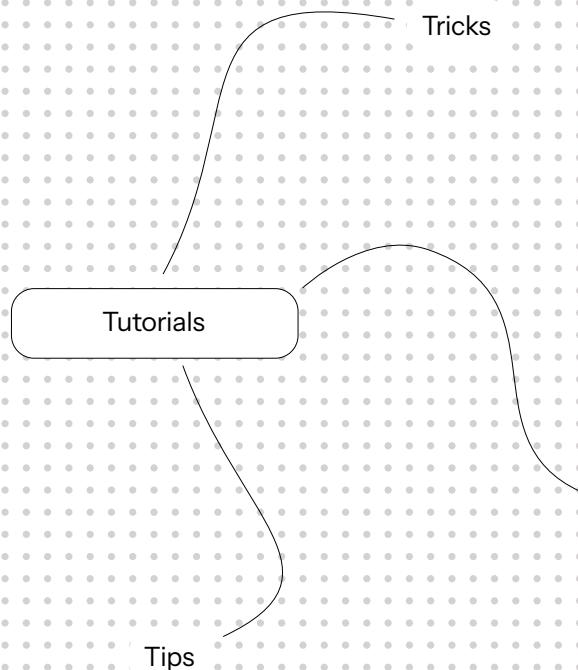
Dedicate 60 to 90 minutes to reviewing these resources. As you study each one, reflect critically on these techniques.



Use the blank mind map on the next page to keep track of the tips and tricks you learned in the tutorials.

Ex^{10b} – Learning Your *Craft*

Keep Track of Tips and Tricks



Ex^{10b} – Learning Your Craft

Do these techniques apply to your original prompt?

Write here

After exploring these techniques, do you now feel different about how much your original prompt shaped the results?

Write here

Ex^{10c} – Practice Makes Perfect

Instructions

Using the prompting techniques you've learned, revise your original prompt. Conduct multiple tests. You can test various techniques individually or in combination. You could also create new chats using the same prompt each time to assess consistency.

Ex^{10c} – Practice Makes Perfect

What specific adjustments did you make?

Write here

How did the LLM's performance change?

Write here

EX^{10d} - *Detective Work Playground*

Instructions

Picture yourself as a detective uncovering the mysteries behind your revised prompts and their outcomes. Dive into your conversations from [Exercise 10c \(“Practice Makes Perfect”\)](#) and hunt down clues to identify patterns:

- What explains why some prompts worked while others fell short?

Carefully examine the interactions between the prompt, the technique you chose, and the LLM. Were they partners in crime or suspects in disguise? Do they have different motives or alibis?

Create your own “detective wall”, a visual map of your findings and theories about what happened. Each box represents a possible fact, observation, witness, suspect or culprit in your investigation.

Each box can belong to three different families:

- The LLM for issues related to the machine, e.g. limited understanding, hallucination, generic answer.
- The Technique for issues related to the prompting, whether the technique itself (e.g. efficacy, applicability) or the tutorial (e.g. clearly explained, easy to apply).
- Yourself for issues related to your own actions e.g. whether you adapted the prompting technique successfully or not, or you provided insufficient or too much information.

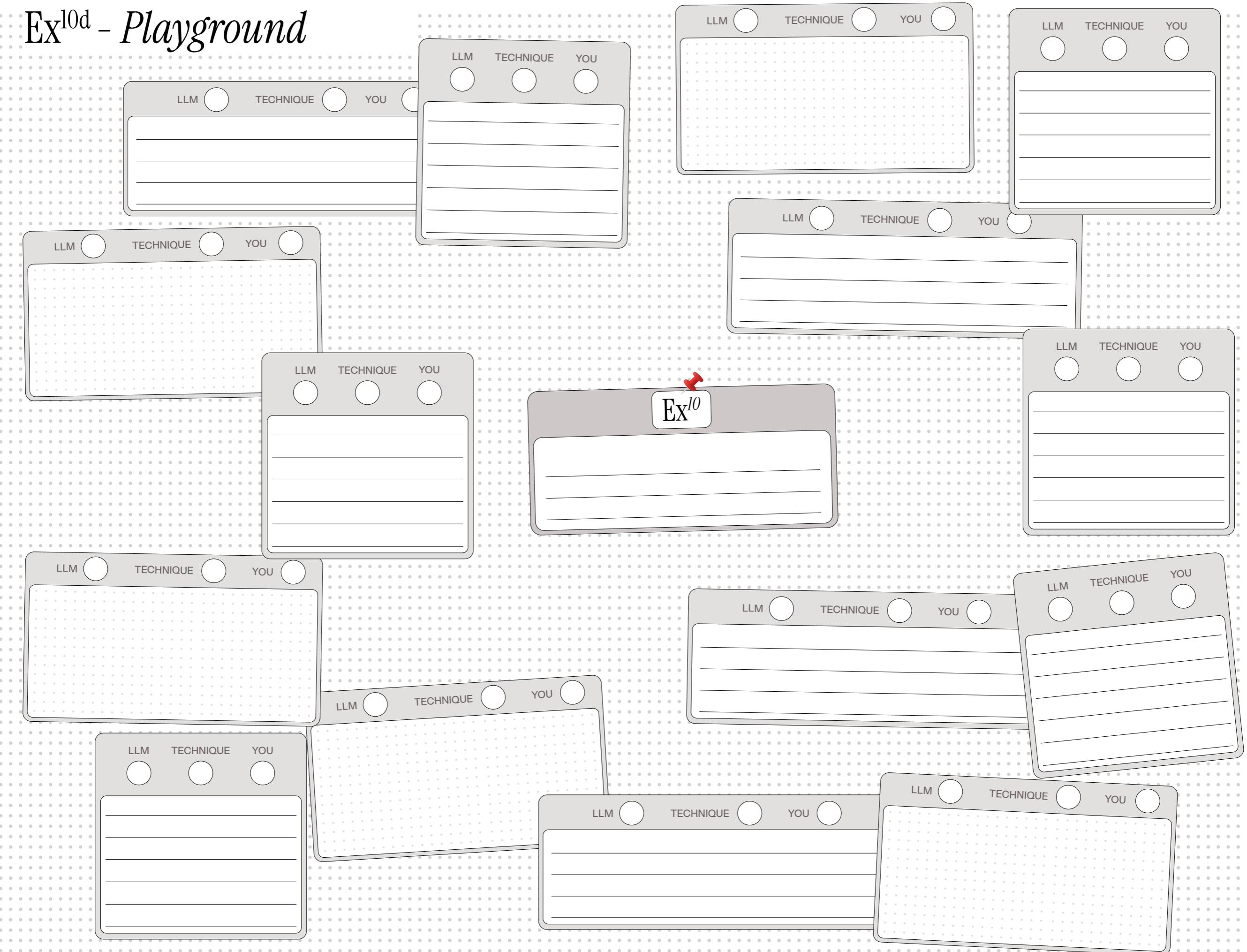
When you add an element to your detective wall, tick off a box (LLM, Technique, You). Draw connections between the elements to visualise the relationships, alliances and conflicts.

After your investigation, narrow it down to the true culprit(s) and write down what happened in the centre of the playground.

Example

LLM 	TECHNIQUE <input type="checkbox"/>	YOU <input type="checkbox"/>
<p>The model was unable to grasp what I was asking for and mistook “mandarin” cuisine with the eponymous language</p> <hr/> <hr/>		

Ex^{10d} - Playground



EX10D
PULL
TO
OPEN
→

Ex^{ll} - Tracking Shifts

Instructions

Reflect on your experience with LLMs since you began the exercises. Browse through your vademecum and your conversation history in your LLM of choice.

Follow the principles detailed in Brief 02 (“How to Describe Your Practices), as well as these specific instructions:

- *Going wide.* Try to think about all aspects of your practice: technical, professional, emotional, ethical, and so on.
- *Documenting your practice.* You must provide specific examples to support your answers.

Ex^{ll} – Tracking Shifts

How has your practice evolved?

Write here

Ex^{ll} - Tracking Shifts

What aspects of your practice have remained consistent?

Write here

Excelling

What happens when we try to reproduce excellent work using an LLM?

The Excelling Block guides you through the reproduction of exemplary work relying solely on an LLM. By “exemplary work”, we mean texts that set a standard in your field or, to put it simply, pieces of writing that you would aim to emulate and reproduce.

You’ll examine what makes a particular piece of writing stand out as excellent, attempt to reproduce it with AI, and then critically evaluate the outcome. You’ll explore the machine’s capabilities and limitations. Finally, you’ll look back on the twists and turns of the entire process.

This block follows a distinctive structure which warrants some explaining. Exercises 12 (“Choosing an Exemplary Piece of Work”) and 13 (“Setting Up the Example”) invite you to choose and describe an exemplary work from your professional field. The core of the block is Exercise 14 (“The Imitation Game”), where you’ll attempt to reproduce a piece of the selected work using an LLM. Given the demanding nature of the process, it is divided into three parts (14a, 14b and 14c), each followed by a specific exercise.

The sequence unfolds as follows:

- After Exercise 14a, complete Exercise 15 (“Anatomy of an Exemplary Work”), which draws your attention to the various components of your exemplary work.
- After Exercise 14b, Exercise 16 (“Obstacles, Dead Ends, Highways”) invites you to reflect on progress, regression, and blockage.
- After Exercise 14c, Exercise 17 (“Charting Your Path”) prompts you to draw the trajectory of your reproduction process.

Ex¹² – Choosing an Exemplary Piece of Work

Instructions

Examples of exemplary work include writings that are widely known in your field (like a seminal article or landmark book in research) as well as more modest forms of writing, like a great email that a colleague wrote, or a very well-crafted internal memo. In other words, exemplary writings aren't limited to legitimate or public forms of literature, but include various sorts of writing that are necessary and important in your profession.

Here are some examples from different professional fields:

- *Academic research*: a monograph; a grant application; a peer-reviewed journal article; an internal feedback email; a course syllabus.
- *Corporate communications*: a keynote speech; an internal memo; a press release; an annual report; a team newsletter; an email summarizing a meeting.
- *Journalism*: an investigative report; a breaking news alert; an editorial pitch; an op-ed; a news brief for internal editing.
- *Tech industry*: code; a white paper; user documentation; a bug report; a technical blog post.

Think about what kinds of exemplary writing you've come across in your field. Come up with a list of 2-3 different texts, and try to include at least one more modest example. Choose one to use as your exemplary work in the following exercises.

If the text is too long for an LLM, select an excerpt within the text. According to the nature of your chosen text, this excerpt can be one page, one subsection, etc. It can be a key feature of your text or it may be chosen arbitrarily.

Describe your Excerpt

Write here

Ex^{l3} – *Setting Up* the Example

This exercise explores the context and significance of your chosen exemplary work by examining the various dimensions that make it relevant: the practices it embodies, the values it upholds, and its reflection of professional standards.

→

Ex^{13a} - *Mapping Exemplary Patterns - Playground*

Instructions

Your exemplary work likely belongs to a specific genre of professional writing. Provide context for this particular type of document.

Evaluate it using several practical criteria, and represent your assessment visually by drawing a radar graph.

Reflect on the relevance, significance, and role of this type of document in your professional practice:

- *Frequency*: How often is this type of text produced or encountered in your profession?
(1: *exceptionally*, 5: *routinely*).
- *Professional stakes*: How significant are the consequences of success or failure of this type of text for your professional reputation?
(1: *Low impact*, 5: *Critical*).
- *Expertise required*: Does the exemplary work require specialised knowledge or skills?
(1: *Basic knowledge*, 5: *High level of specialisation*).
- *Effort and Resources*: How resource-intensive is the production of this type of text (time, effort, tools)?
(1: *Minimal effort*, 5: *Significant investment*).
- *Collaboration*: How collaborative is the writing of such a text?
(1: *Independent*, 5: *Highly collaborative*).
- *Error Tolerance*: How critical is it for this type of document to be perfect, meaning no errors?
(1: *Tolerates errors*, 5: *No room for error*).
- *Standardisation*: How standardised is this type of text?
(1: *Highly context-dependent*, 5: *Highly standardised*).
- *Professional pride*: How much does this type of document contribute to a sense of professional accomplishment?
(1: *Low level of pride*, 5: *High level of pride*).

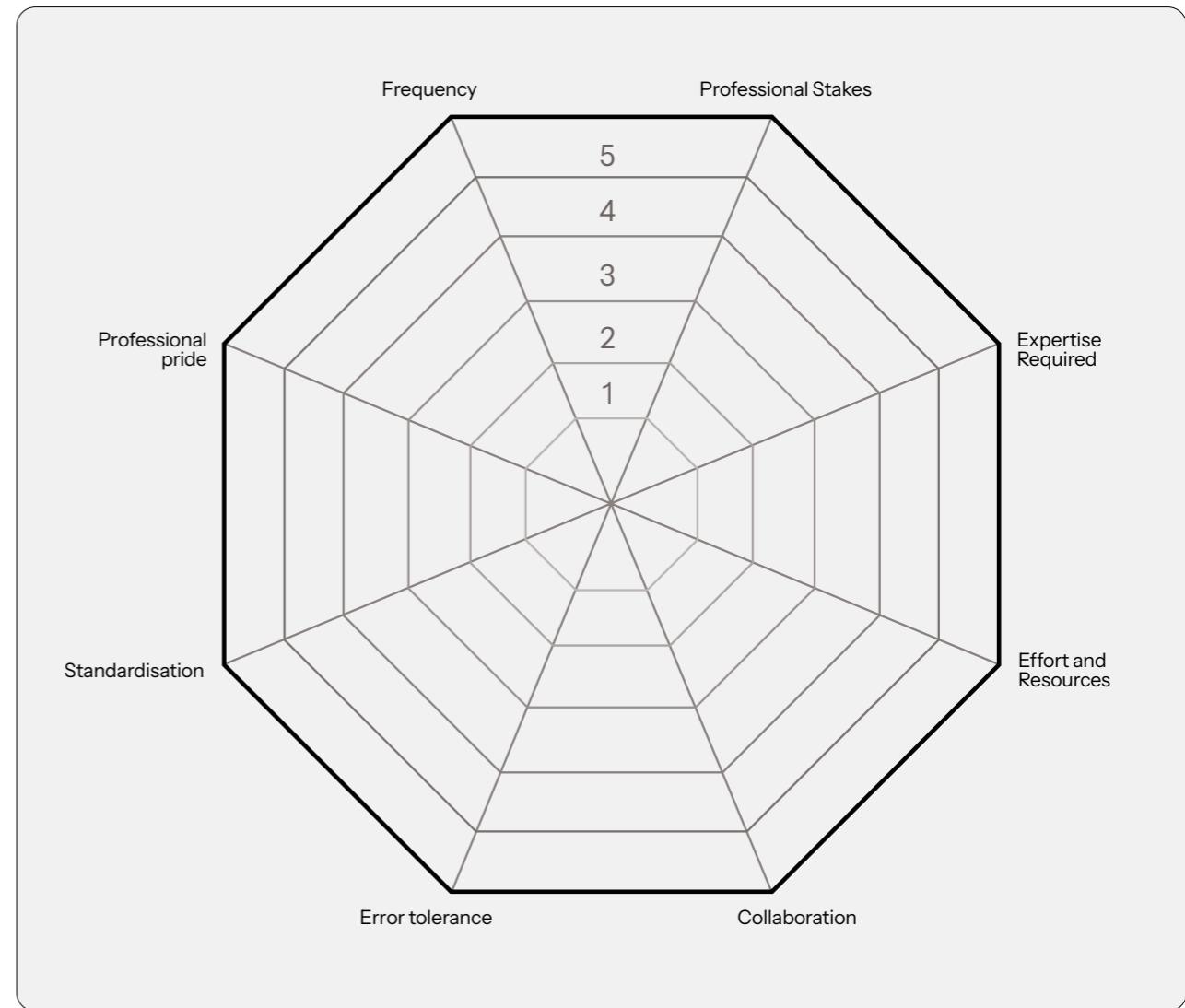
Ex^{l3a} - Playground

Category	Description
Frequency	How often is this type of text produced or encountered in your profession? <i>(1: exceptionally, 5: routinely)</i>
Professional Stakes	How significant are the consequences of success or failure of this type of text for your professional reputation? <i>(1: Low impact, 5: Critical)</i>
Expertise Required	Does the exemplary work require specialised knowledge or skills? <i>(1: Basic knowledge, 5: High level of specialisation)</i>
Effort and Resources	How resource-intensive is the production of this type of text (time, effort, tools)? <i>(1: Minimal effort, 5: Significant investment)</i>
Collaboration	How collaborative is the writing of such text? <i>(1: Independent, 5: Highly collaborative)</i>
Error Tolerance	How critical is it for this type of document to be perfect, meaning no errors? <i>(1: Tolerates errors, 5: No room for error)</i>
Standardisation	How standardised is this type of text? <i>(1: Highly context-dependent, 5: Highly standardised)</i>
Professional Pride	How much does this type of document contribute to a sense of professional accomplishment? <i>(1: Low level of pride, 5: High level of pride)</i>

My task is

Write here

The Radar Graph



Notes

Write here

Ex^{13b} – Fill in the Blanks

Instructions

Reflect on why your exemplary work stands out, filling in the blanks in the following paragraph. Refer back to the principles detailed in Brief 02 (“How to Describe Your Practices”).

Fill in here

This work is exemplary because it ¹[_____

_____,]
which is crucial in the field of ²[_____.].

It addresses ³[_____

_____,]
offering ⁴[_____

_____,]
The work stands out due to its ⁵[_____

_____,]
and reflects ⁶[_____

_____,].

What I would most like to emulate from this work is ⁷[_____

_____,]
as it aligns with ⁸[_____

_____,].

¹Primary contribution

⁴Unique solution or perspective

⁷Aspect to emulate

²Profession/discipline

⁵Distinctive quality

⁸Reason it resonates with your goals or

³Specific problem or challenge

⁶Some professional values or standards

practices.

Ex¹⁴ – The *Imitation Game*

Instructions

The goal of this exercise is to reproduce a page from an exemplary text using an LLM. Having selected one of the exemplary works you identified in [Exercise 12 \(“Choosing an Exemplary Piece of Work”\)](#), try to produce the most accurate and faithful copy in terms of style, syntax, semantics (meaning), and pragmatics (effects).

Adhere to the following constraints:

- You have a total of 4.5 hours, divided into three 90-minute sub-exercises (14a, 14b and 14c), to produce the most faithful imitation possible.
- Follow this specific sequence: Ex 14a → Ex 15 → Ex 14b → Ex 16 → Ex 14c → Ex 17
- The reproduction must be carried out within a conversation with no reference to the original text (no attachments, no copy-pasting, no screenshots, and no embedded excerpts), including no metadata (title, author’s name, URLs, and references such as articles or summaries about the exemplary work).
- During your work, take cursory notes in the logbook.

Outside of the specified constraints, you’re free to use the LLM in any way you choose. If you would like to introduce more variety from one sub-exercise to the next, here are some ideas:

- Choose a different page from your exemplary work to reproduce; or select an altogether different text to reproduce.
- Use a different LLM, for example one that is more suited to your specific writing needs or professional standards.

Ex^{14a} – The *Imitation Game*

Reminder

When you're done with this exercise, go to [Exercise 15](#) ("Anatomy of an Exemplary Work").

Write here:

Checklist

Preparing the exercise:

- Have the ~1-page excerpt from the exemplary work ready.
- Re-read the instructions.

Before you begin:

- Ensure you have 90 consecutive minutes to work on the Imitation Game.
- Have the ~1-page excerpt from the exemplary work ready.
- Have the logbook next to your computer or machine to take notes while you work.
- Play the Imitation Game!

After you're done:

- Review and add to your notes if necessary.
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-

Ex^{14b} – The *Imitation Game*

Reminder

Complete Exercise 15 (“Anatomy of an Exemplary Work”) before tackling this exercise. Then, go to Exercise 16 (“Obstacles, Dead Ends, Highways”).

Write here

Checklist

Preparing the exercise:

- Have the ~1-page excerpt from the exemplary work ready.
- Re-read the instructions.

Before you begin:

- Ensure you have 90 consecutive minutes to work on the Imitation Game.
- Have the ~1-page excerpt from the exemplary work ready.
- Have the logbook next to your computer or machine to take notes while you work.
- Play the Imitation Game!

After you’re done:

- Review and add to your notes if necessary.

Ex^{14c} – The *Imitation* Game

Reminder

Complete [Exercise 16 \(“Obstacles, Dead Ends, Highways”\)](#) before tackling this exercise. Then, go to [Exercise 17 \(“Charting Your Path”\)](#).

Write here

Checklist

Preparing the exercise:

- Have the ~1-page excerpt from the exemplary work ready.
- Re-read the instructions.

Before you begin:

- Ensure you have 90 consecutive minutes to work on the Imitation Game.
- Have the ~1-page excerpt from the exemplary work ready.
- Have the logbook next to your computer or machine to take notes while you work.
- Play the Imitation Game!

After you’re done:

- Review and add to your notes if necessary.

Ex¹⁵ – *Anatomy* of an Exemplary Work

You've now completed [Exercise 14a \("The Imitation Game"\)](#). Now take a step back and analyse what makes a particular piece of work exemplary. Whether it's a scholarly article, a professional report, or a creative piece, every text exists within a network of dependencies: its stylistic choices, disciplinary norms, institutional constraints, and broader expectations.

This exercise helps you unpack these layers by mapping out the key components that contribute to the excellence of your chosen work. Create a dependency graph to visualise its structure, influences, and material conditions. Examine which components are accessible to an LLM and which remain beyond its reach.

→

EX^{15a} – The *Dependency Graph*

Instructions

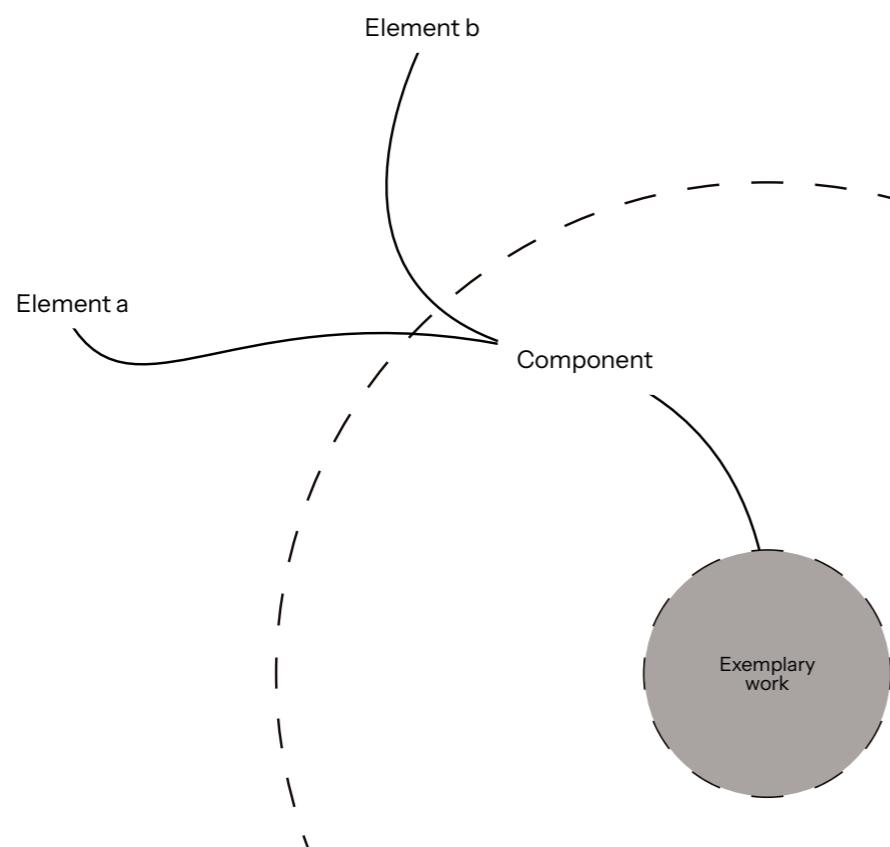
Start by analysing your exemplary work by identifying its key components.

- What makes it possible for that work to exist in the first place?

These elements may include, but aren't limited to: style and form (writing style, tone, voice); norms and standards (academic or professional conventions, editorial guidelines, ethical frameworks); contextual stakes (institutions where the work was produced, why it was made, what problem or need it addresses, intended audience/use); people involved (reviewers, editors, committees, or institutions); expectations and constraints (time, tools, funding, deadlines); material dependencies (sources cited, empirical materials, and relevant software or tools).

- The goal is to represent these components in a tree-like diagram, with concentric circles indicating different levels of dependency.

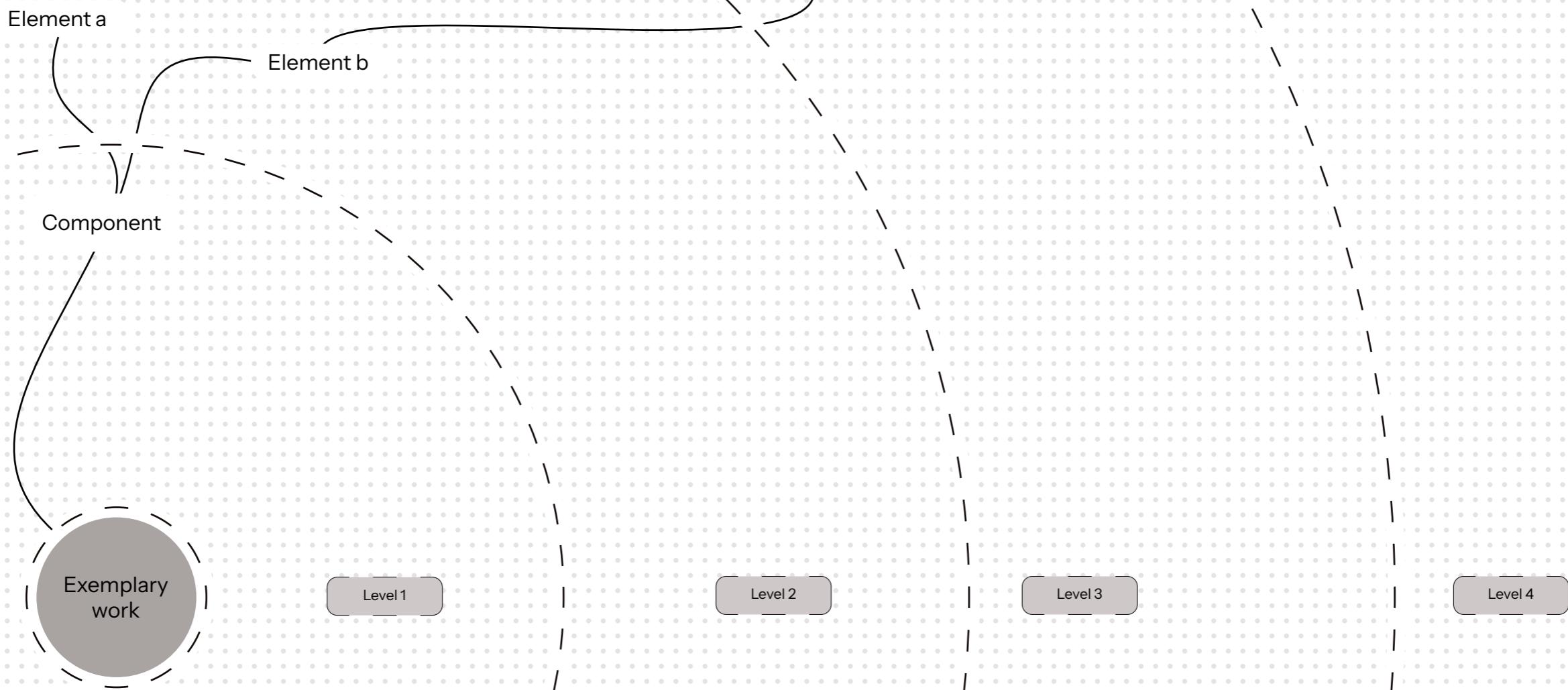
Place the exemplary work at the centre of the template and surround it with the most immediate components. Then, break down each element into its dependencies. Use lines and annotations to show how the components relate to one another.



Ex^{15a} - Playground

EX15

PULL
TO
OPEN
→



Notes

Ex^{15b} – Inside and Out

Instructions

Now that you've mapped the components of your exemplary work, assess whether each element is accessible to an LLM.

To complete this requirement, draw up two lists representing different levels of accessibility.

- *List A* represents internal elements: components already present in the LLM. They may be accessible because you've explicitly provided them as context, or because the LLM was likely trained on similar examples.
- *List B* consists of accessible elements: components that aren't currently internal to the LLM but that could be integrated or encoded with effort.

Anything not included in either list represents elements that are out of reach for the LLM. These are dependencies that are difficult or impossible for a machine to access and encode practically.

Neither list necessarily matches the depth levels of your diagram. For example, you might identify elements from depth 1, 3, and 4 in list A.

Ex^{15b} – Inside and Out

List A - Internal Elements

<p>Write here</p> <hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>
<hr/> <hr/>

List B - Accessible Elements

<hr/> <hr/>

Ex¹⁶ – Obstacles, *Dead Ends, Highways*

You've now completed [Exercise 14b \("The Imitation Game"\)](#). Now reflect on how you navigate LLMs when pursuing a specific goal. Drawing on your experience from [Exercise 14a and 14b \("The Imitation Game"\)](#), map out three types of LLM behaviour you came across: Obstacles, Dead Ends, and Highways. As always, follow the principles detailed in [Brief 02 \("How to Describe Your Practices"\)](#).



Ex^{16a} – Obstacles

Instructions

Obstacles are moments when you encounter difficulties but still see a path forward. These are problems you either manage to overcome or believe you could.

1. Start by identifying how you recognised you'd hit an *obstacle*:
 - How and when did you notice that something wasn't working as expected?
 - Once stuck, how significant was the setback?

2. Next, focus on your response:
 - How did you go about resolving the issue?
 - Did you try something on the spot, such as repeating your instructions or did you tweak your prompts?
 - How did you decide when to stop one line of inquiry and shift directions?

3. Pinpoint the key factor that allowed you to break free from the *obstacle*:
 - What eventually worked?
 - Why do you think it was more effective than previous attempts?

4. At each step, describe your thoughts, feelings, and actions.
 - Were you frustrated, curious, or determined?

Write here

Ex^{16b} – Dead Ends

Instructions

Dead Ends are moments when you feel truly stuck, unable to progress, or at a loss as to what to do next.

1. Begin by describing the earliest signs that led you to conclude that you had hit a *dead end*.
 - Did the LLM keep looping or repeating itself?
 - Did it contradict itself or misunderstand your input again and again?
2. Then, focus on the specific approaches you tried before concluding it was truly a *dead end*.
 - Did you insist on a solution?
 - Did you explore other strategies?
3. Reflect on what made this moment different from a typical obstacle:
 - What clues or patterns told you it was time to stop?
4. Finally, consider the emotional and intellectual toll of encountering a *dead end*.
 - Were you frustrated, resigned, exhausted?

Write here

Ex^{16c} – Highways

Instructions

Highways are moments when everything flows smoothly, and progress feels almost effortless and even enjoyable.

1. Think about how you noticed you had entered a *highway*.
 - Did the LLM respond immediately with just the right answer?
 - Did it anticipate your needs or exceed your expectations without explicit prompting on your part?
 - Were you able to accomplish your goal faster than expected?
2. Then, reflect on your role in these *highways*.
 - Did they happen by chance, or thanks to careful prompting on your part?
3. Consider how it felt to have the LLM consistently align with your intentions.
 - Did it create new expectations?
 - Why do you think everything clicked in this particular instance?

Write here

Ex^{16d} – Patterns and Differences

What are the main differences between these three experiences?

Write here

*Are there recurring patterns or situations
that lead to one experience over another?*

Write here

Can you identify additional types of moments that don't fit into these categories?

Write here

Ex¹⁷ – Charting Your Path

Instructions

You've now finished Exercise 14c ("The Imitation Game"). This block's final exercise invites you to reflect visually and narratively on your experience. By drawing three versions of a graph, represent the trajectory of your attempts to reproduce the exemplary work. Alongside the graph, provide written explanations that clarify and comment on the key shifts in your trajectory.

Two Axes

The graph functions in two axes: Output Fidelity and System Breakdown.

- *The horizontal axis represents the fidelity of the output.* Going to the far left means that the result is completely dissimilar from your exemplary work. On the far right, it is a perfect reproduction of the original. The better the copy, the further right you should draw your line along this axis.
- *The vertical axis captures model behaviour.* The very bottom represents ideal behaviour. Moving upward indicates increasing loss of control over the LLM. At the very top, the model behaves erratically and completely unexpectedly. The more the system breaks down, the higher your line moves on this axis.



Drawing Instructions

1. Start on the *bottom left side of the graph*, which marks your *starting point*.
2. Move the line *steadily to the right* as your output resembles the original more and more. This movement reflects your progress towards or away from your goal. If you feel the output resembles the original less and less, let the line *veer left* to represent that regression.
3. *Move the line upward* when you experience system breakdown, i.e. the LLM is behaving erratically. A small indent indicates a minor obstacle; a steep incline signals serious derailment. When you *regain control*, bring the line back down toward the baseline.

The final shape of your line might be smooth or erratic, representing a gradual progression or sudden changes. There's no ideal outcome: the point is to capture your real experience navigating the LLM.

Ex¹⁷ – Charting Your Path

Three Iterations

Complete this exercise in three iterations.

In each progressively longer iteration, the resolution of your drawing increases as you refine and detail it.

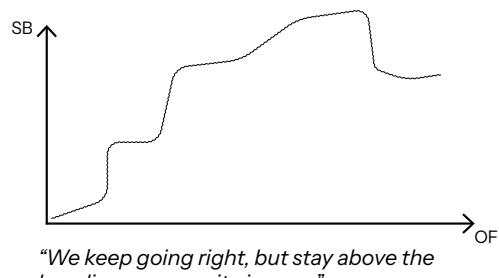
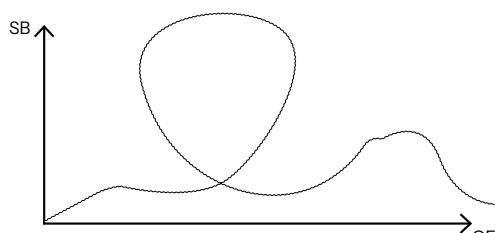
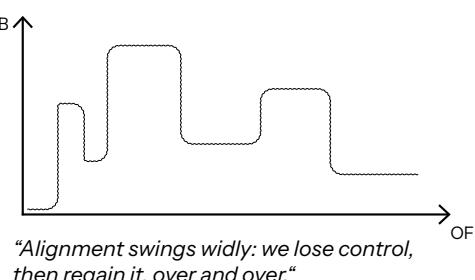
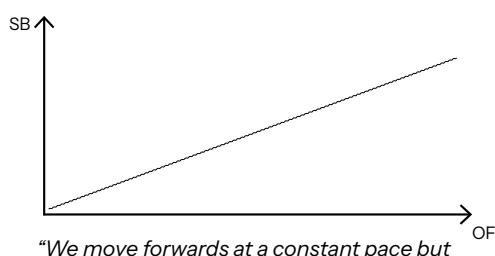
1. *1 minute* – A preliminary sketch: draw what you feel your path looks like, even if you do so roughly. Draw from what you have in mind without referring back to your actual conversations. Don't get stuck in details, and capture your overall impressions of the Imitation Game.
2. *5 minutes* – Refine the line and begin identifying key shifts or events. You can briefly refer back to your conversations. The goal is to ground your first impressions in your actual practices.
3. *30 minutes or more* – Finalize the graph and write accompanying explanations that document your experience in detail. You can revisit your conversations as much as you wish. Use the writing instructions below to shed light on major turning points.

Writing Instructions

Write a few sentences for each major turning point. These explanations are meant to clarify what happened at key moments:

- What made you feel stuck or disoriented?
- What helped you move forward again?
- Did you change your approach or prompting strategy?
- How did you recognize progress, obstacles, or breakthroughs?

Examples



Ex^{17a} - Low Resolution

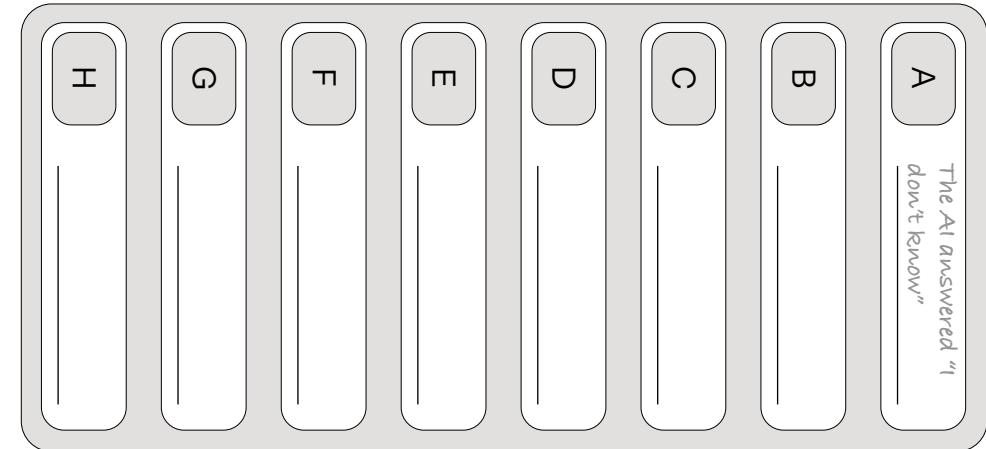


Ex^{17b} – Mid- Resolution

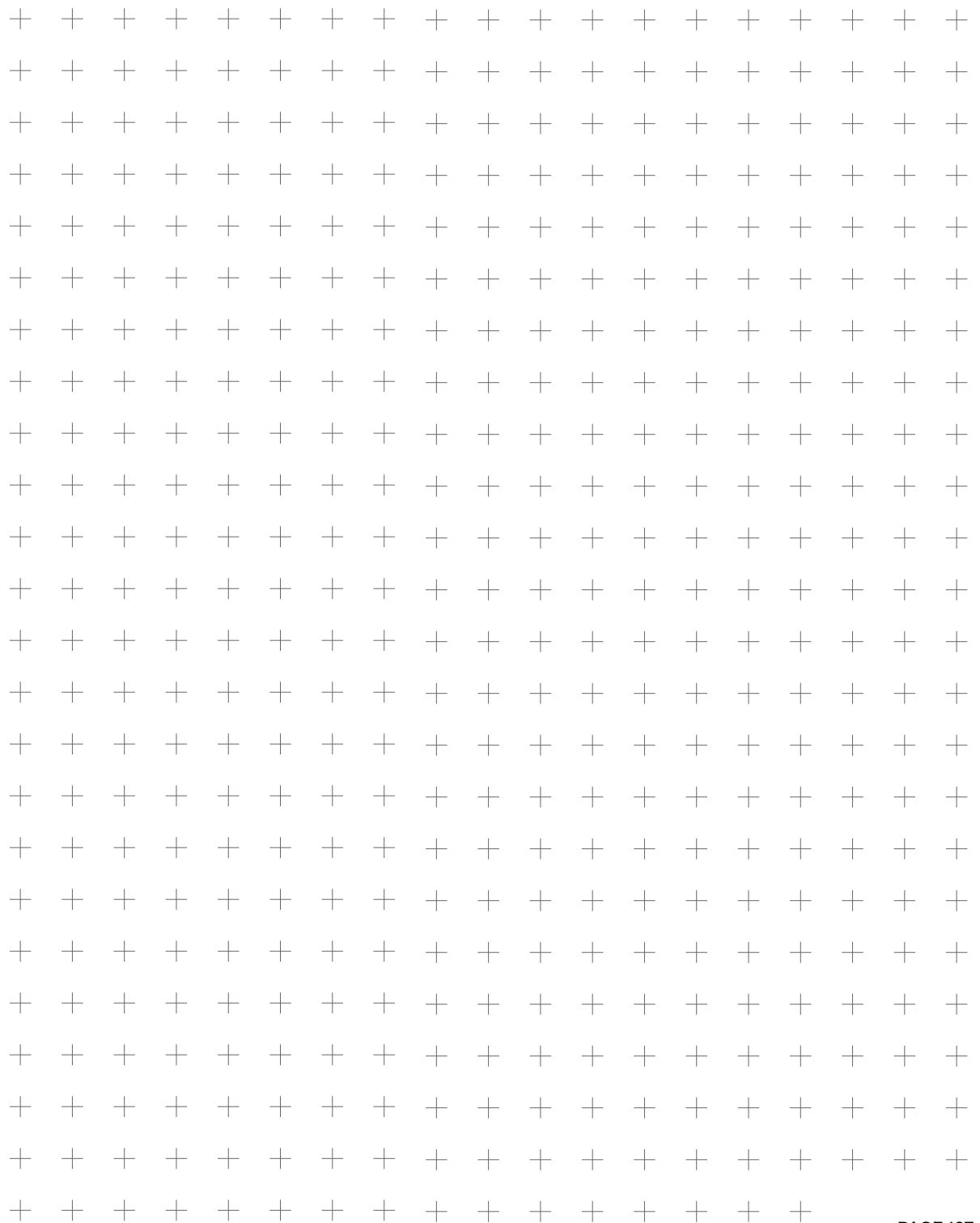
Marker
Why did your path
change direction?

The AI answered "(
don't know")

Start here:



SYSTEM BREAKDOWN →



*Ex^{17c}- High Resolution
Playground*

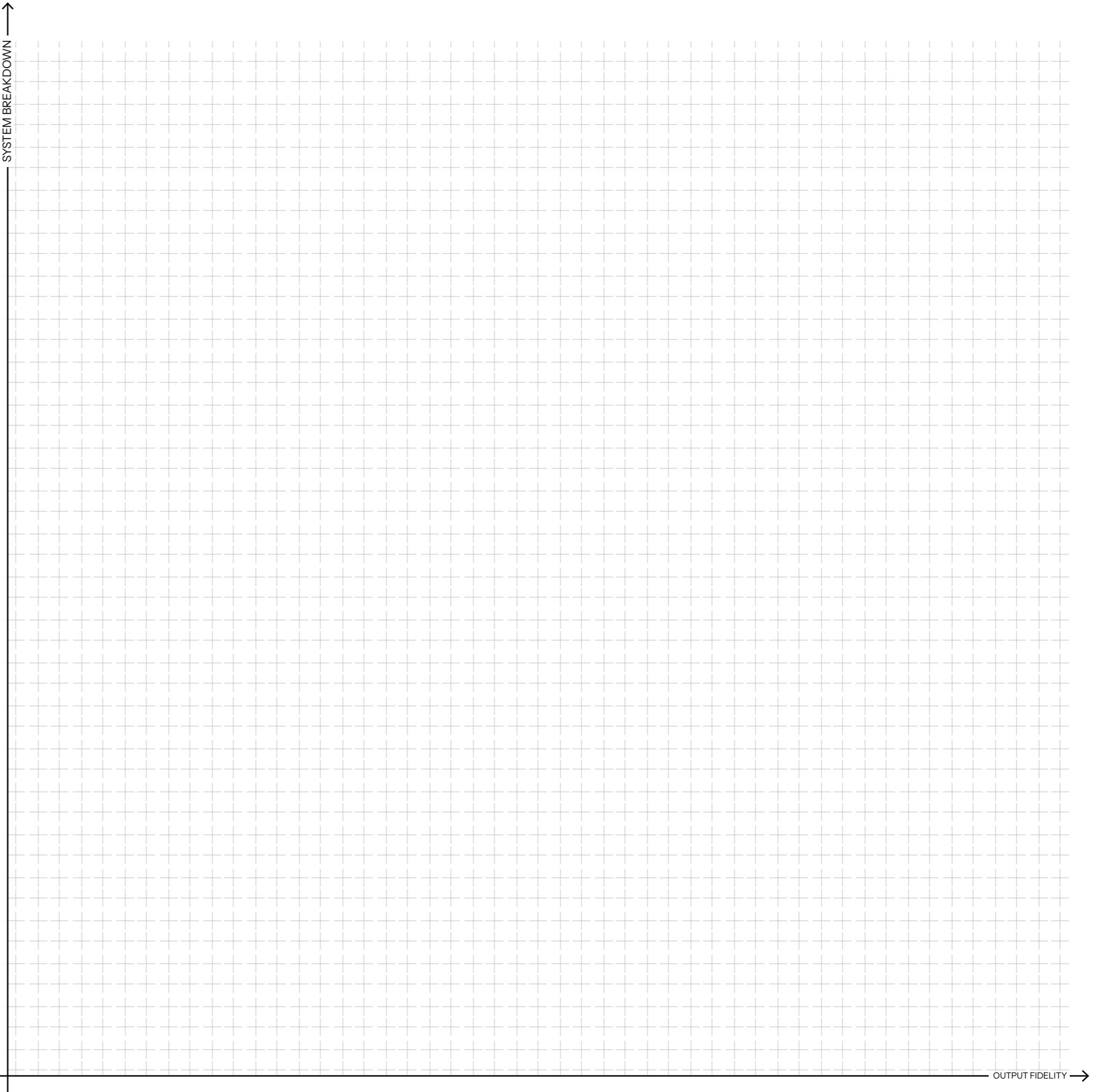
Ex^{17c} - High Resolution

EX17C

PULL
TO
OPEN
→

Marker Why did your path change direction?

Marker	Why did your path change direction?
A	The AI answered "I don't know"
B	
C	
D	
E	
F	
G	
H	
I	
L	
M	
N	
O	
P	



Distilling

*How can we reframe the role of LLMs
in our ordinary work practices?*

This final block is about extracting the essential. Rather than a simple recollection, this block provides you with an opportunity to reframe your experience, and to tell a story on your terms. You'll revisit the vademecum and distill it into a curated selection of key entries, notes, sketches, screenshots, and digital artefacts that give shape and substance to your narrative.

Ex^{l8} - *Distilling* the Vademecum

As we enter the final phase of the experiment, we invite you to craft new narratives, reframing how LLMs inhabit your work and your professional ecology.



Ex¹⁸ – *Distilling the Vademeum*

Instructions

More specifically, you’re encouraged to tell two interwoven stories:

- *What does your work say about LLMs?* Reflect on how you assess the role, value, and position of LLMs in your day-to-day activities. What are they good or bad at? Where do they fit — or fail to fit — in your routines? How have you come to judge their usefulness or limitations?
- *What do LLMs say about your work?* Reflect on what the use of LLMs has revealed about your work and your professional environment. Perhaps you realized that your work was easily automatable or, to the contrary, was more complex than you initially assumed. Maybe you came to rely on the LLM for certain things for which you wished you had a colleague.

These narratives do not need to be chronological. Think of them as a synthesis, a condensation of what you’ve discovered through the exercises. Don’t try to cover every single thing. This is a moment to delve deep into what stood out the most to you.

Take the time to re-read the principles detailed in [Brief 02 \(“How to Describe Your Practices”\)](#), especially those about speaking as a member of your profession and writing accounts grounded in practice.

Finally, these narratives take shape through a curated object. We invite you to identify, select and weave together specific vademeum entries, annotated pages, sketches, screenshots, and digital traces that illustrate key moments in your journey. First, review and select material from the vademeum. Identify and extract pages that capture pivotal moments or recurring patterns. Second, expand your archive. Use the additional blank pages we’ve provided. Add your notes, sketches, screenshots, or digital artefacts (such as links to key conversations) where needed. The last and final step consists in constructing your distilled vademeum. Organize all the material you selected into a mini-booklet as the material foundation underpinning the account of experiences, successes and failures, joy and frustrations casting a light on both LLMs and your working ecologies.

If you struggle with the disassembly and reassembly of your vademeum, refer to [Brief 01 \(“Understanding the Vademeum”\)](#) for guidance.

Sharing

What do we do with what we've found?

Your vademecum is now completed. We hope you gained new insights from it, and that it allowed you to question and reframe your understanding of AI. As researchers, we're eager to learn about your experience with LLMs. In this block you'll find ways to contact us in order to discuss what you've learned, or share raw materials directly from your vademecum.

Becoming a Co-Inquirer Brief 3

We're genuinely interested in the work you've done with this vademecum. If you decide to share all of it or your distilled version with us, it will become part of a co-inquiry, a shared exploration where your insights enrich our collective understanding. You can reach out and send your contributions to any of the email addresses listed on this page or on our website.

Research Team

Donato Ricci: Principal Investigator, Designer,
donato.ricci@sciencespo.fr

Gabriel Alcaras: Postdoctoral researcher, Sociologist,
gabriel.alcaras@sciencespo.fr

Tommaso Prinetti: Research Assistant, Designer
tommaso.prinetti@sciencespo.fr

Zoé de Vries: Research Assistant, Linguist
zoe.de_vries@ens-paris-saclay.fr



About “Ecologies of LLM Practices” Brief 4

How Can We Reframe the Role of LLMs in Ordinary Work Practices?

The Ecologies of LLM Practices (EL2MP) project documents the role of LLMs in various professional practice(s) and the consequences of their use.

Context

The upsurge of large language models (LLMs) such as ChatGPT has sparked a rush to produce discourse about these technologies. The quick crystallisation of a shared outlook around a few key themes has narrowed the scope of potential inquiry. As public and scientific debates focus on technical issues—algorithmic bias, confabulation, and intellectual property violations—the problems and consequences of their actual use remain largely unexplored for both their users and their profession. This asymmetry fuels a mechanical view of technological development and its effects, as if the technical analysis of these systems were enough to predict their social impact. Moreover, these discourses present AI as a monolithic and disruptive entity, dismissing the possibility that it may be aligned with existing practices and that its effects may vary depending on situations encountered in one’s job.

Thus, there is an urgent need to move beyond predictions, describe the professional contexts in which LLMs are used, and identify current issues, not prospective ones. How do AI’s well-known problems (such as bias, confabulation) manifest in established practices? What new, unexpected problems are surfacing? How do LLMs shape individual work practices? And, in turn, how do professional environments shape LLMs and their use?

Objective

To answer these questions, the Ecologies of LLM Practices (EL2MP) project creates research areas dedicated to workers to document and reflect on their use of LLMs. Our investigation aims to highlight the *savoir-faire*, expertise, and values of workers rather than those of AI designers or economic decision-makers. The project seeks to examine how LLMs fit precisely into various professional practices. EL2MP will investigate how users relate to LLMs in terms of:

- *Evaluation*: How do professionals assess the value LLMs add or remove?
- *Effort*: What new kinds of work do LLMs require from their users?
- *Perception*: How do workers evaluate and perceive LLMs over time as they continue working with them?

About “Ecologies of LLM Practices” Brief 4

Methodology

To carry out the project, we developed an experimental research protocol based on the active involvement of participants. This protocol, the fruit of collaboration between sociologists, science and technology researchers (STS), and designers, aims to create:

- Room for hesitation: Through a series of exercises, the protocol provides participants with various means of documenting and reflecting on their use of LLMs. We aim to establish a framework embracing doubt and hesitation. Some exercises introduce deliberate pauses in professional routines, causing participants to retreat from their practices. During these pauses, judgment is suspended, giving users enough room to question their instinctive habits and feel unsure about what they once thought was certain. Other exercises aim to intensify the use of LLMs in order to shed light on what would otherwise remain too subtle to be perceived.
- An ecological archive of practices: The digital traces from LLM use serve as the starting point for the exercises we designed. They are analysed, contextualised, and discussed individually and collectively. Through this process, participants progressively build a multimodal archive (audio, photos, videos, drawings, and logbooks) that tangibly reflects their LLM experiences and how they’re incorporated into various professional norms and material configurations.
- A workbook (vademecum): The workbook collects all the exercises and instructions, organised into thematic blocks. Designed as a modular object, where pages can be added or removed as needed, it accompanies each participant throughout the protocol. The vademecum serves as both the material support for the protocol and the project’s means of dissemination, as it can be reproduced and adapted in various academic and professional contexts.

Artificial Inquiries

Éditions Annexes

Edited by

Gabriel Alcaras
Donato Ricci
Tommaso Prinetti
Zoé de Vries

Design by

Tommaso Prinetti

Proofreading by

Liz Carey-Libbrecht
Traduction Savantes

Informations

150 pages
Copy [__] of 75

Acknowledgements

We thank the fifteen participants from the first two cohorts, the design team for their generous feedback and conversations (Axel Meunier, Robin de Mourat), and the contributors to the “Design/gâteau” seminar.

Inspired by The Virtual Sentence: A Book of Exercises, Cabinet Books, 2024.

Technicalities

“Artificial Inquiries” is printed on 80g/m² common office paper and bound using two “flat bar” (ACCO-type) fasteners threaded through four punched holes.

Typography: Instrument Serif and Instrument Sans.

The Ecologies of LLM Practices logo & identity were designed by Tommaso Prinetti.

Scans from participants’ vademecums are reproduced with their full consent.

Emoji use follows open-source licensing standards where applicable.

ChatGPT, Mistral and Claude logos have been downloaded from the correspondent Wikimedia pages.

Legal informations

Septembre 2025 | Dépot legal Octobre 2025



Creative Commons 2025
Éditions Annexes
médiaLab, Sciences Po
ISBN 978 2 959706714



Ecologies Of LLM Practices

This publication is part of the EL2MP
(Ecologies of LLM Practices) project,
with support from Google.org.



EXTRA BLOCK

Additional Notes

Additional Notes

