Chat-PhD: The Role of LLM in Shaping the Research Practices and Productivity of Young Scholar

Stefano Bianchini[1], Ivan Napoli[2], Gabriele Pellegrino[2], Pierre Pelletier[1]

November 25, 2025

[1]BETA, Université de Strasbourg, France
[2]Università Cattolica del Sacro Cuore, Italy

The Promise of Generative AI

- The "Nobel Turing Challenge" envisions an AI system capable of making Nobel Prize–worthy discoveries by 2050
- LLMs have capability to possess specialized knowledge and creative and analytical skills (Noy and Zhang, 2023; Lee et al., 2023; Girotra et al., 2023).
- The appeal lies in automating and expediting routine knowledge-creation tasks, potentially allowing researchers to focus on higher-level analytical activities (Hoffman et al. 2024; Nordling, 2023).
- To reach Nobel-level discoveries, AI would need to replicate how laureates connect existing concepts in novel ways (Arts et al., 2025)

Concerns: Homogenization and Idea Exploration

- Scientific progress emerges through paradigm shifts that challenge established norms (Kuhn, 1962), yet breakthrough ideas are becoming harder to generate as accessible discoveries become scarcer (Jones, 2008; Bloom et al., 2020).

- Growing reliance on LLMs raises concerns about homogenization and the exploration of knowledge space (Lin et al. 2025; Bao et al. 2025; Anderson et al. 2024; Zhu et al. 2025)

- LLMs systematically reinforce the Matthew effect in citations by consistently favoring highly cited papers when generating references and diverge from traditional citation patterns (Algaba et al., 2025)

- LLM systems generate recombinations but are not neutral recombiners—patterns are anchored in statistical distributions of pre-existing corpora

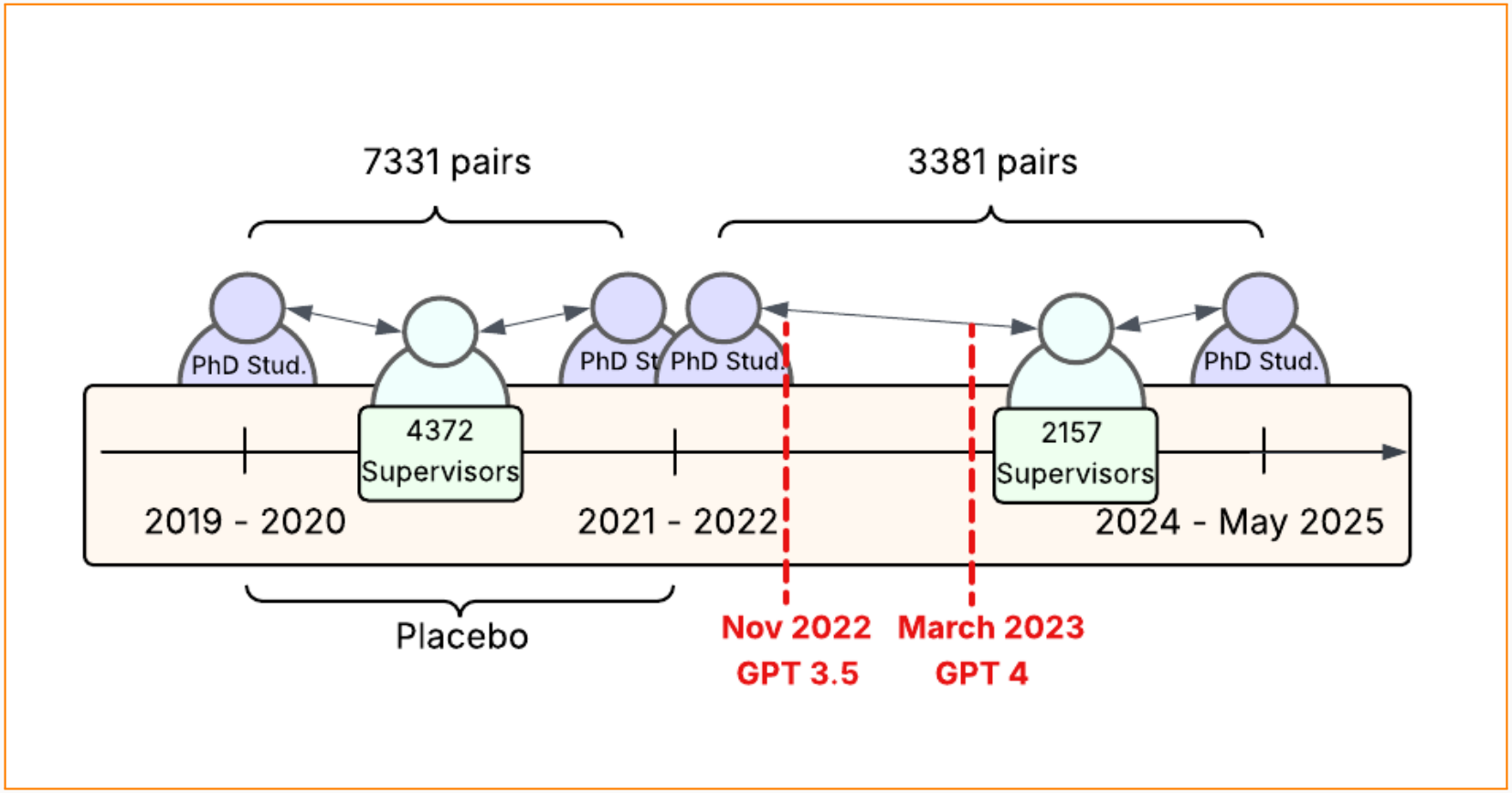
Why PhD Students?

- Young researchers perceive the most utility in these tools (Liao et al., 2024)
- Experience moderates the impact of Generative AI ideas on the research process (Doshi et al., 2025)
- PhD are a driving force of the scientific production (Larivière, 2012)
- During Ph.D. training, student experiences social environment characterized by vertical and horizontal relationships that shape their productivity (Corsini et al., 2022).

Do LLMs reshape PhD students' scientific production ?

Sample and Design

- Population: Theses.fr up to May 2025
- Manually retrieved 5000 emails from all PhD who defended after ChatGPT: 265 responses



- Matching strategy: PhD students finishing between Jan 2021–Dec 2022 compared with those defending between Jan 2024–May 2025, matched by common supervisor

Data and Methodology

- Data processing:
  - Downloaded all PhD Theses in PDF from HAL, parsed page by page
  - Retrieved References sections and extract authors, year, journal, reference type (done by GPT-4o-mini)
  - Restricted to journals indexed in Web Of Science $\rightarrow$ 3M references
- Metrics computed:
  - Textual metrics: readability, graduated level, lexical diversity
  - Citation metrics: commonness of combined knowledge, knowledge from separated communities
- Additional data from Idref API: age, gender, nationality

LLMs Usage

- Detection of top 100 AI-generated expressions in Theses:
  2019–2020: 7% – 2021–2022: 11% – After 2023: 30%

- Disciplinary differences are notable:
  - Computer Science: 38.1%
  - Physics, Life Sciences, Engineering, Chemistry: 28-30%
  - Social Sciences and Humanities: 15%

- Our qualitative study confirms that SSH researchers perceive LLMs as less useful than other researchers, we then excluded it from our quantitative analysis.

Regression Results

	Dependent variable:								
	Lee Nov.	Foster Nov.	Flesch Read.	Gunning Fog	SMOG Index	Dale Chall	Coleman-Liau	Read. Index	Lexical Div.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After ChatGPT	-0.020	-0.015	-3.949	0.674*	0.287*	0.536***	1.302	1.565*	0.022***
	(0.133)	(0.022)	(3.311)	(0.354)	(0.163)	(0.194)	(0.989)	(0.907)	(0.007)
log(# EN Words + 1)	-0.258***	-0.010	17.382***	-0.937***	0.045	-1.258***	-5.385***	-4.709***	-0.038***
	(0.037)	(0.006)	(0.916)	(0.098)	(0.045)	(0.054)	(0.274)	(0.251)	(0.002)
log(# Ref. + 1)	0.044	0.004	-6.908***	0.287***	-0.001	0.584***	2.162***	1.773***	0.014***
	(0.028)	(0.005)	(0.705)	(0.075)	(0.035)	(0.041)	(0.211)	(0.193)	(0.001)
log(# Nb Pages + 1)	0.309***	0.040***	-9.299***	0.687***	-0.125	0.909***	3.163***	3.109***	0.028***
	(0.072)	(0.012)	(1.807)	(0.193)	(0.089)	(0.106)	(0.540)	(0.495)	(0.004)
% EN Pages	0.481***	0.021	-48.663***	4.580***	2.244***	0.254	14.027***	12.742***	0.014**
	(0.110)	(0.018)	(2.748)	(0.294)	(0.136)	(0.161)	(0.821)	(0.752)	(0.005)
Age	-0.011**	-0.001	-0.049	0.007	0.014***	0.005	0.014	0.008	0.0004**
	(0.004)	(0.001)	(0.104)	(0.011)	(0.005)	(0.006)	(0.031)	(0.028)	(0.0002)
Gender	-0.071*	0.004	-0.730	0.248**	0.039	-0.081	-0.046	0.205	-0.003*
	(0.038)	(0.006)	(0.946)	(0.101)	(0.047)	(0.055)	(0.283)	(0.259)	(0.002)
Is FR	-0.067*	-0.003	-0.266	-0.151	-0.030	-0.026	0.389	0.288	-0.005**
	(0.038)	(0.006)	(0.957)	(0.102)	(0.047)	(0.056)	(0.286)	(0.262)	(0.002)
Constant	1.567*	0.367**	-32.886	15.497***	11.743***	16.558***	33.113***	29.810***	0.798***
	(0.896)	(0.150)	(22.369)	(2.391)	(1.103)	(1.308)	(6.685)	(6.126)	(0.044)
Domain lvl2 FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
University FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Director FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	4,498	4,498	4,498	4,498	4,498	4,498	4,498	4,498	4,498
R^2	0.693	0.714	0.575	0.622	0.745	0.762	0.564	0.568	0.724
Adjusted R^2	0.520	0.553	0.336	0.410	0.601	0.628	0.318	0.325	0.569
Residual Std. Error (df = 2875)	0.772	0.129	19.270	2.060	0.950	1.126	5.759	5.278	0.038
F Statistic (df = 1622; 2875)	4.001***	4.432***	2.402***	2.923***	5.173***	5.681***	2.290***	2.334***	4.655***

Notes: This table reports coefficients of the effect of ChatGPT on references diversity and text composition. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of ChatGPT is estimated using a OLS.

Results: Reference Novelty

- Generative AI does not appear to homogenize citation practices among PhD students $\rightarrow$ Combinations of references in French Theses are not more probable after ChatGPT

- This absence of difference is similar accross field as well as in our robustness checks

- Respondents report using LLMs more for **refining existing content** than creating new content: 64% vs 26%

- Bibliographic research and idea generation were hardly mentioned as tasks for which PhDs used LLMs

Results: Writing Quality

- LLMs appear to have increased text complexity, using longer sentences and/or more multi-syllable words
- Thesis written after ChatGPT have a higher lexical diversity, especially in Computer Sciences and Life Science
- Qualitative analysis supports our results, LLMs are mainly involved in the writing phase:
  - "Improving the clarity and structure of my writing" (63% of users)
  - "Automating tedious or repetitive tasks" (38%)
  - "Speeding up my research" (30%)
- A strong majority (>74%) consider LLMs at least moderately important for **grammar and syntax correction**. Nearly half (48%) rate this function as very important.

Qualitative Analysis

Theme	Core Idea	Quotes
Functional use and Linguistic Equity	The most accepted use is **editorial or linguistic**, not epistemic. The LLM is seen as a *technical assistant* rather than a *co-author of knowledge.* Also seen as a "linguistic democratization".	"*Glorified grammar generators*"; "*Great proofreader and syntax assistance. I am not an English native speaker, and these tools help me a lot.*"
Reliability and Technical Limits	Widespread agreement that LLMs **make errors** and **invent sources**, there is a perception of unreliability and hallucinations (n.b. model variability).	"*LLMs sound* ***convincing, even when they're wrong***"; "*Very good at respecting the structure of provided tasks (e.g., debugging) but not at thinking or providing reliable information*"; "*So unreliable, so many invented citations!*"
Changing the Formative Role of the PhD Journey	LLMs challenge the **self-learning function** of the doctoral process, acting as dangerous "shortcuts".	"*Using an LLM to write would be like learning to read by having someone read for you*". "*Now that I've learned to use them, I find them extremely useful. However, I* ***sometimes wonder whether, if I had used them from the very beginning, I would have developed the same depth of knowledge and methodological rigor.***"
Evolution Over Time	Perception of **technological improvement** accompanied by **shifting opinions**.	"*I didn't use it much during my PhD, except for refining my writing. However,* ***if I had to do another PhD now, the contribution of LLMs would be five times greater.***" "*I started using LLM much more frequently during my first year of Post-Doc. Now,* ***I use them to acquire knowledge and skills that are missing in my lab***"
Ethical and Environmental Concerns	Emergence of **radically critical** positions, resembling a *moral resistance.*	"*We scholars should all stand against LLMs. They are an ecological, political,* ***ethical nightmare!***"; "*It's intellectual laziness that will ruin innovative thinking.*"

Summary

- AI does not appear to affect the manner in which PhD students explore the knowledge space
- Generative AI seems confined for now to the role of proofreader and code assistant for PhD student (52% find LLM assistance at least moderately important for coding)
- Skepticism remains when using LLMs as tools to explore literature and generate ideas
- LLMs can foster new to the lab research even if the knowledge combined is not new to the world
- STEM and Social Sciences are more inclined to implement LLMs compared to Humanities.

  $\rightarrow$ LLMs seem to allow PhDs to increase their speed without disrupting the diversity of theoretical foundations on which Theses are built

Impact on Social Interactions

- Only a small minority (8%) feel they have gained autonomy from their supervisor (e.g., in troubleshooting)
- However, a majority (54%) report a **reduced need to consult with colleagues**
- Engaging in conversations with colleagues remains central to problem-solving and generating new, practical ideas (Tahamtan and Bornmann, 2018)
- When individuals learn from LLMs, they tend to develop shallower knowledge than through standard web search (Melumad and Yun, 2025)

  $\rightarrow$ Illusion of explanatory depth?

Too soon?

- The PhDs surveyed began their thesis before ChatGPT and thus built their knowledge base and work routines before the democratization of these tools
- Among respondents not using LLMs during their thesis, only 12% still do not use them in their current research
- Respondents who completed their Theses more recently generally find LLMs more useful

Chat-PhD: The Role of LLM in Shaping the Research Practices and Productivity of Young Scholar

Stefano Bianchini[1], Ivan Napoli[2], Gabriele Pellegrino[2], Pierre Pelletier[1]

November 25, 2025

[1]BETA, Université de Strasbourg, France
[2]Università Cattolica del Sacro Cuore, Italy