

Reflections and Takeaways from the 1st YNLG Workshop

Patricia Schmidtova Nils Feldhus Adarsa Sivaprasad
Alyssa Allen Eduardo Calò Rudali Huidrom Michela Lorandi

Abstract

The inaugural YNLG workshop served as a platform for community building and knowledge transfer among early-career researchers. This document synthesizes the core insights from the workshop and presents a comprehensive summary of advice from the Program Committee on navigating research methodology, publishing, and mental health. Furthermore, it highlights key takeaways from a panel of experts on mentorship, the evolving landscape of academia versus industry, and the future challenges of evaluation in NLG.

1 Introduction: A Collective Vision

With the rising popularity of large language models (LLMs), NLG is in the spotlight of NLP research. We believe it is crucial for the next generation of researchers to discuss their vision for the future of the field. The workshop facilitated this through roundtable discussions and a detailed questionnaire sent to senior researchers. This document serves as a record of those insights, covering topics such as research rigor, mental health, and the transition between academia and industry.

1.1 Methodology

The workshop submissions included a section for *Questions for senior researchers*. These questions were thematically grouped and generalized by the second author, with the exclusion of highly technical or overly specific items unlikely to be relevant to a broader audience. The first author subsequently reviewed the refined set of questions and developed the survey questionnaire (see Appendix A).

The survey was distributed to members of the workshop Program Committee, the Advisory Board and also shared via the INLG Discord channel. Hence, all contributors are senior researchers with doctoral and early-career experience in NLG. The contributors provided their names and affiliations,

and contact addresses for follow-up questions. A total of 7 responses were obtained (4 from academia, 3 in industry). We acknowledge all survey participants for their contributions to this study.

Two of the authors conducted a reflexive, interpretive qualitative analysis focusing on the relevance and potential impact of responses for the early-career audience. This process was informed by the authors positionality as young researchers. In presenting the results, minor edits were made to avoid unwarranted certainty and provide explanatory context for readers.

2 Questionnaire on NLG Research

Here, we summarize the general consensus and highlight some of their specific recommendations. We note that these reflect the advice and opinions of a handful of senior researchers, and while valuable, they might not be universally representative.

2.1 Selecting and Scoping Projects

It is generally advised to pursue research areas that genuinely interest you. Do not be afraid of getting into areas where others might not be focusing. Breadth is often valuable, but depth is crucial and almost inevitable in a PhD.

- **Starting Small:** For early-stage PhD students, starting with a relatively small, well-scoped problem is often a highly effective way to experience the publication process, attend a conference, and network with the community. Good initial projects might include participating in a shared task, evaluating the performance of different systems, assessing the quality of datasets planned for future work, or studying the needs of potential users.
- **The Double-Edged Sword of Survey Papers:** Familiarizing yourself with related work is essential. Synthesizing these findings into a survey can be highly beneficial,

especially if a growing area lacks an existing survey. However, a comprehensive survey can take a long time to finish, and if you are unfamiliar with a field, it might be difficult to determine what is truly relevant. Such surveys often require high-level reflection, so they might be better suited as a longer-term goal or done collectively with more senior colleagues.

- **Exploring the Underexplored:** In higher productivity stages, you might introduce a new method for the task at hand, or benchmark existing methods under underexplored setups (e.g., data scarcity, synthetic data augmentation, out-of-distribution data).
- **Tackling Foundational Tasks:** It can be rewarding not to overlook “boring” infrastructure problems. Tasks such as language identification, text normalization, sentence splitting, or complex data cleaning are often assumed to be solved but may have plenty of open problems unseen by the general public. Improving these foundational layers might have significant downstream impacts. Pay attention to what annoys you personally in your workflow; if a tool frustrates you, it likely frustrates others and might be a prime candidate for research.

2.2 Navigating the Literature

Staying current is challenging due to exponential growth. Effective strategies often include talking to colleagues, setting Google Scholar alerts, and reviewing for conferences.

- **Tools for Filtering:** Engage with social networks such as Bluesky, LinkedIn, or Discord (join our NLG discord!)¹, use Scholar Inbox (Flicke et al., 2025),² and reverse search papers citing relevant work.
- **Mindset:** Try to embrace the “Joy of Missing Out” (JOMO) (Sima, 2024). If new developments are robust, you will likely find them eventually; if they are a passing trend, they might not matter. It can also be helpful to explore the Slow Science movement (Stengers, 2018), which prioritizes deep, rigorous work over rushed publication cycles.

¹Invite link valid for the first 100 users, join the two strings and remove the plus sign: <https://discord.gg/+pUVXxSjGuR>.

²<https://www.scholar-inbox.com/>

While it is difficult to fully practice this individually when career progression is often measured by KPIs that reward quantity, as a community, we should actively advocate for redefining these metrics to value the quality and impact of research.

2.3 Navigating the “Publish or Perish” Culture

Research decisions should ideally be driven by what gives you energy. A suggested rule of thumb is to *build solutions for problems rather than finding problems for existing solutions*.

- **Balancing First-Author vs. Collaborative Work:** Try to prioritize longer-term research goals over publication pressure. What questions do you want answered by the end of your PhD? Focus on quality over quantity, because it is very easy to get sidetracked from your ultimate goal. Furthermore, it is beneficial to strike a balance between joining larger collaborations and publishing as a lead author. While there is a lot to learn from the dynamics of large collaborative projects, being the first author (potentially with just your advisors) teaches you how to design and coordinate experiments, which is crucial for learning from your own mistakes.
- **Handling Rejection and Peer Review:** It is generally advised to take reviewer feedback as a chance to improve the paper. Sometimes, however, reviews might not seem particularly helpful. If this is the case, it can still be valuable to reflect on whether your framing and contributions were clear enough. If, after two rounds of peer review, the core message still does not seem to land, it might be a good option to wrap the project up for a workshop or a smaller, more targeted conference. In particular, the Workshop on Insights from Negative Results in NLP (Drozd et al., 2025) might be receptive to good ideas that did not work as expected.
- **Getting Scooped:** At some point, you will likely find out someone has already published a paper with a similar idea. When this happens, try not to despair. First, it usually means you are working on a topic that people care about. Second, reviewers might point out the limited novelty, so it is important to make

your unique angle (e.g., evaluating a failure mode, applying it to a new language or task) explicitly clear.

2.4 Experimental Methodology and Rigor

- **Hypothesis Testing and Pilot Data:** Instead of forcing strong hypotheses to be true, it might be better to formulate explorative ones. Start small and pilot your experiments with “toy data” first. This might be a small subset of existing gold-standard annotations, or, if such data does not exist, there is no shame in hand-annotating a small dataset yourself. This hands-on experience often provides a much deeper understanding of the task, which carelessly generated synthetic data might gloss over.
- **Qualitative Analysis:** Look at your data. Metrics can hide critical failures. It is often essential to manually inspect model outputs to understand the true behavior of the system, rather than relying solely on aggregate numbers.
- **Confounders and Prompts:** Accept the inherent ambiguity of language. A common goal is to isolate *enough* confounders to ensure the setup reflects real-life usage. Regarding prompt variability, consider running pilot studies to collect representative prompts. Crucially, try to report negative results—including the prompts that *failed*—to prevent the community from repeating the same trial-and-error loops.
- **Future Evaluation Challenges:** We can expect a continued standardization of automatic and human metrics. Reproducibility issues with closed-source models and their high financial cost will likely continue to impact how we evaluate. Countering hallucinations, strengthening Retrieval-Augmented Generation (RAG), and cost-cutting (efficient inference) are anticipated trends.

2.5 Career Development and Well-being

- **Building an Identity and Teaching:** Consider creating a personal website so you are easily findable. Try not to hyperfixate; gaining experience with a broad set of problems can be helpful. Try not to neglect teaching.

It often forces you to develop a stronger understanding of your own work and serves as a proxy for supervision experience.

- **Preparing for the Next Step and Interviews:** Approach potential advisors with clear ideas, ideally grounded in your prospective lab’s research focus. Check whether your core values align with those of the company or institution you plan to apply to. Informational interviews and reaching out to current employees before committing can be highly beneficial.
- **Final Advice:** Try not to be your own enemy. Take your time, value your mental health, and focus on finding happiness in your work. Many are drawn to research because they want to improve the world, not necessarily to prove to themselves that they are good.

3 Panel Discussion Highlights

The workshop concluded with a panel discussion featuring experts from both academia and industry. The following sections summarize the key takeaways regarding mentorship, career decisions, and the future of the field.

3.1 Mentorship and Guidance

The panel emphasized that a mentor is distinct from an advisor; a mentor is someone who offers guidance without necessarily being the one to evaluate your work. Finding a mentor does not always require a formal, long-term arrangement. “Pop-up” mentorships—such as ad-hoc conversations with senior researchers at conferences—can provide critical insights on specific problems without the need for a years-long commitment.

When seeking mentorship, the panel advised looking for individuals who are “grounded” and focused on intentional research rather than those merely chasing “shiny” new trends. A good mentor helps you understand not just the answers, but which questions to ask. They can also assist interdisciplinary researchers in adapting to the different writing and conversational styles required by different fields.

3.2 Navigating Career Paths: Academia vs. Industry

A recurring theme was the fluidity between academia and industry. The panel noted that there is no “forever” decision; researchers frequently

switch between these sectors throughout their careers.

- **Agency and Happiness:** PhD students were reminded that they possess significant agency in prioritizing their work. The ultimate metric for a career path should be personal happiness, rather than external prestige.
- **The Reality of Academia:** While academia offers intellectual freedom (and holidays), it comes with significant hurdles regarding funding. Furthermore, during the semester, research often becomes a low priority due to teaching loads and committee service.
- **The Reality of Industry:** Industry roles may focus more on implementation than pure research, and decisions are sometimes influenced by external restrictions rather than pure preference.
- **Advice:** Informational interviews are the best way to understand the reality of a role before committing.

3.3 The Future of NLG and Evaluation

When asked about long-term goals for the field, the panelists highlighted the urgent need for better standardization and reproducibility. Current studies often show that papers are incomparable, and the burden is on the community to design evaluations that are specific and rigorous.

There was a specific call to design reference standards that are closer to real human language use. One of the issues first mentioned in the roundtable discussions was that LLM-generated outputs are frequently “off” in ways that are immediately perceptible to humans—often described as a sense of “cringe” – yet remain unmeasurable by current quantitative metrics. Developing methods to capture these subtle, qualitative aspects that automatic metrics currently miss is a critical direction for future research.

Acknowledgments

We thank Emiel van Miltenburg, Simone Balloccu, Ondřej Dušek, Saad Mahamood, Elizabeth Salesky, Eduardo Calò, and Katya Artemova for taking the time to fill out the questionnaire and share their personal experiences. We also thank all participants of the YNLG workshop for their active engagement in the roundtable and panel discussions,

which provided the foundation for the collective vision presented here.

References

- Aleksandr Drozd, João Sedoc, Shabnam Tafreshi, Arjun Akula, and Raphael Shu, editors. 2025. *The Sixth Workshop on Insights from Negative Results in NLP*. Association for Computational Linguistics, Albuquerque, New Mexico.
- Markus Flicke, Glenn Angrabeit, Madhav Iyengar, Vitalii Protsenko, Illia Shakun, Jovan Cicvaric, Bora Kargi, Haoyu He, Lukas Schuler, Lewin Scholz, Kavyanjali Agnihotri, Yong Cao, and Andreas Geiger. 2025. [Scholar inbox: Personalized paper recommendations for scientists](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, pages 307–317, Vienna, Austria. Association for Computational Linguistics.
- Richard Sima. 2024. [Forget fomo, embrace jomo to discover the joy of missing out](#). *Washington Post*. Archived version available at: <https://web.archive.org/web/20240115225539/https://www.washingtonpost.com/wellness/2024/01/04/fomo-jomo-joy-missing-out/> (Retrieved February 19, 2026).

Isabelle Stengers. 2018. *Another Science is Possible: A Manifesto for Slow Science*. Polity Press, Cambridge, UK.

A Survey Questionnaire

Instruction to participants: We have collected questions from young researchers taking part in the first workshop for young researchers in NLG. We are hoping that you will help us answer them to provide some guidance! We are planning to go over all of your answers and combine your perspectives into a helpful document that will be released with the workshop proceedings. The questions (except for your identity and some quick questions in the last section) are not required and you are free to only answer those, where you do have an opinion to share. We will try to make it so that you can come back and edit your responses if you think of a better way to phrase things.

There are 3 main sections (Research Direction and Promising Areas, Experimental Design and Methodology, Career Navigation and Development) + a chance to give us any other advice that does not fall into these categories.

Questionnaire:

1. Research Direction and Promising Areas

- (a) **Identifying Promising Research Areas:** What strategies or approaches do you recommend for identifying the most promising research areas in NLG? What factors do you consider when evaluating potential research directions?
- (b) **Staying Current with the Field:** How do you effectively keep up with new publications and emerging methods in the rapidly evolving NLG field? What resources or practices have been most valuable to you?
- (c) **Research Pursuit Decisions:** How do you decide how far to pursue a particular research idea? What signals indicate when to pivot, double down, or abandon a research direction?
- (d) **Balancing Impact and Publishability:** What advice do you have for identifying research that is simultaneously publishable and high-impact for real-world applications? How do you navigate potential tensions between academic and practical value?
- (e) **Future Expectations:** What are your expectations for the future of NLG research? What emerging trends or areas do you think will be most important in the next 3-5 years?
- (f) **Publication Pressure vs. Rigor:** How do you balance maintaining rigorous research standards with publication pressure? What strategies help you manage this tension effectively?

2. Experimental Design and Methodology

- (a) **Hypothesis Testing:** What are your best practices for designing experiments to effectively test hypotheses in NLG research? What common pitfalls should researchers avoid?
- (b) **Controlling Confounding Factors:** How do you approach controlling for confounding factors in NLG experiments? What techniques or considerations are most important in your experience?
- (c) **Managing Prompt Variability:** What strategies do you use to account for prompt variability in your experiments? How do you ensure your findings are

robust across different prompting approaches?

- (d) **Ensuring Robust Findings:** Beyond prompt variability, what other factors do you consider to ensure your experimental findings are robust and generalizable?
- (e) **Planning vs. Iteration:** How do you balance early discussion and careful planning with iterative improvements during the research process? What has worked best in your experience?

3. Career Navigation and Development

- (a) **General Career Path Navigation:** What advice do you have for researchers navigating career paths in NLG and related fields? What factors should early-career researchers consider when making career decisions?
- (b) **Internships and Interviews:** What recommendations do you have for applying for internships and preparing for interviews in NLG-related positions? What do you look for when evaluating candidates?
- (c) **PhD to Industry Transition:** For those transitioning from PhD studies to industry positions, what advice would you offer? What skills or experiences are most valuable for making this transition successfully?
- (d) **Long-term Academic Careers:** What guidance would you provide for those pursuing long-term academic careers in NLG? What strategies are important for building a sustainable academic research career?

4. Anything Else :)

- (a) **Any Other Advice:)** Is there any other advice, insight, or perspective you'd like to share with NLG researchers and practitioners that wasn't covered in the previous questions?
- (b) **Feedback For Us:** Do you have any feedback, advice, or recommendations for us organizers? Perhaps nominating other people who could have interesting perspectives.

- (c) **Preparing the Findings:** Would you like to be invited to collaborate on creating the final version of the findings?