

中谷アドバンストプログラム最終報告書

中村 健（東京大学工学部計数工学科システム情報工学コース 4 年）

留学期間： 2026 年 2 月 7 日～2026 年 4 月 3 日

受け入れ先： Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI)

配属部門： Smart Data & Knowledge Services (SDS)

PI： Prof. Dr. Prof. h.c. Andreas Dengel

Mentor： Dr. Stanislav Frolov

1 インターン先での研究

1.1 研究内容

インターン期間を通じて、大規模言語モデルの継続学習に関する研究に取り組みました。継続学習とは、新しい課題を順番に学習しながら、すでに獲得した能力をできるだけ失わないようにする枠組みです [1, 2]。この枠組みを安定して実現できれば、既存のモデルを一から作り直すのではなく、段階的に拡張しながら効率的に高性能化できることが期待されます。一方で、実際には忘却や学習信号の不安定性など未解決の課題が多く、継続的に拡張していく形での本格的な実用化にはまだ至っていません。近年は、固定された正解列を模倣する通常の教師あり学習だけでなく、現在のモデル自身が生成した軌跡を用いる on-policy 型の学習や自己蒸留が注目されています [3, 4]。本研究では特に、強化学習のように探索的で計算量が非常に大きい手法ではなく、比較的シンプルに実装できる自己蒸留をはじめとする on-policy 型の学習に注目し、継続学習の設定での利点と不安定性を調査しました。

具体的には、on-policy 型の更新信号を用いる手法の比較を、ツール呼び出しと数学推論の 2 つの設定で行いました。ツール呼び出しの研究では、Qwen2.5-7B-Instruct[5] を用い、異なるツール群を 4 段階で順に学習する継続学習ベンチマークにおいて、通常の supervised fine-tuning と self-distillation fine-tuning[4] を比較しました。近年の BFCL や ACEBench のようなツール利用ベンチマークでは、ツール選択だけでなく出力形式の厳密さも評価対象になります [6, 7]。そのような設定のもとで、学習時の出力形式と評価時に要求される「ツールを 1 回だけ正しい形式で呼び出して停止する」という制約がずれていると、自己蒸留側で最初の呼び出しの後も出力を続けてしまう失敗が支配的になり、手法比較そのものの解釈が難しくなる点を検証しました。この結果から、厳密なツール利用評価においては、学習時と評価時の出力形式の整合を独立した測定変数として扱う必要があることが分かりました。

数学推論の研究では、主な実験として Qwen3-8B[8] を用い、過去の学習データを再利用しない継続的な fine-tuning の設定で、模範的な推論過程を学習目標とする方法と、モデル自身が生成した推論過程を学習目標とする方法を比較しました。近年は、検証器を用いた正しさの確認や、自己生成した推論過程の選別によって学習信号を改善する研究が進んでいます [9, 10, 11]。この背景を踏まえて実験した結果、単にモデル自身

が生成した学習目標を使うだけでなく、検証器によって正しさを確認したうえで、より短く、現在のモデル分布に近い推論過程を選択することが、最終的な既習課題での精度と忘却量の両面で有利に働く条件があることを確認しました。また、結果が特定のモデルだけに依存していないかを見るために、Gemma-2-9B-it と Mistral-7B-Instruct-v0.3 でも対応する比較を行い、モデルファミリーによって望ましい学習目標の傾向が異なる可能性も確認しました。蒸留系の条件や正解情報を補助的に与える対照実験との比較からも、継続学習では更新則だけでなく、どの推論過程を学習目標として採用するかが主要な設計軸であると整理しました。

これらの検討を通じて共通して見えてきたのは、継続学習では「どの更新則を使うか」だけでなく「モデル自身のどの出力を次の学習信号として採用するか」が性能と安定性を大きく左右するという点です。そのため、on-policy 学習や自己蒸留においては手法を評価する際だけでなく、実際にモデルを学習させる段階でも、学習信号の由来、出力形式の整合、学習目標の選び方を切り分けて考えることが重要であると分かりました。

1.2 研究の進め方

研究は、まず継続学習、モデル自身の出力を反映する post-training、自己蒸留に関する最近の論文を読み、関連手法の位置づけと主要な論点を整理するところから始めました。そのうえで、既存の実験基盤を理解し、LoRA[12] を用いた追加学習の設定や評価指標を確認しながら、比較が公平になるように実験条件を整えました。

文献調査、実装、実験、結果分析を往復しながら、メンターと議論して解釈を更新する形で研究を進めました。特に、単に性能差を確認するのではなく、その差がどの設計要因に由来するのかを丁寧に切り分けることを重視しました。インターン期間の最後には、部門全体向けの Webinar で 30 分ほど成果をプレゼンテーションとして発表し、質問やフィードバックをいただきました。

1.3 研究環境

研究所内の設備はとても充実していました。広々とした部屋で机やモニターを使わせていただき、快適に研究に集中できました。クラスタ環境も充実しており、Nvidia H100, H200, A100 などの GPU が多数用意されていて、必要に応じて利用できたことは研究を進めるうえで大きな支えになりました。共用スペースでは、コーヒー、炭酸水、水などを自由に飲めるようになっていました。



図1 DFKIの研究環境。左上：研究所の外観，右上：研究室入口，下：研究室内の様子。

2 アドバンスプログラムで習得したもの

今回は RIES プログラムの際とは異なり，1 人で現地に赴いて研究活動に取り組みました．そのため，研究面と生活面の両方において，自立して物事を進める力がさらに身についたと感じています．

2.1 研究面

研究面で最も大きな学びは，実装力だけでなく，測定と比較の観点から研究を組み立てる力の重要性を実感したことです．大規模言語モデルの研究では，実験を動かして数値を得ること自体もちろん重要ですが，その数値が何を測っているのか，どの条件をそろえなければ公平な比較にならないのかを考えることが不可欠です．今回の研究では，手法の性能差をそのまま結論にするのではなく，出力形式，学習目標，評価指標などの要因を切り分けながら丁寧に考える姿勢を学びました．

また，前回の RIES プログラムで米国の大学研究室を経験したことに続き，今回はドイツの研究所で研究することで，日本，米国，欧州の研究環境を比較して考えられるようになりました．DFKI では，研究所全体が複数の部門に分かれ，その中で多くのプロジェクトが進められていました．研究者は各プロジェクトの目的や期間に応じて研究活動を行っており，プロジェクトが変われば扱うテーマも大きく変わることがあると伺いま

した。そのため、幅広い知識を持ち、新しい分野に素早く適応する力が求められる環境であると感じました。大学の研究室とは異なる、研究所ならではのプロジェクトベースの研究の進め方を知ることができたのはとても貴重な経験でした。

2.2 生活面

生活面では、慣れない海外の地で自分一人で生活を組み立てる力も身につきました。短期滞在であったため近くの学生寮には入れず、Airbnb で借りたアパートの一室で生活していました。食事は、夕食を自炊し、昼食は研究所近くのスーパーやベーカリーなどで購入することが多かったです。研究所内には食堂や売店はありましたが、少し歩いたところに大学のカフェテリアもありました。また、毎週水曜日には研究所近くにハンバーガーのキッチンカーが来ており、研究チームのメンバーと一緒に買いに行くのが習慣になっていました。このような日常の小さな交流も、現地での生活に馴染むうえで大きな助けになりました。

渡航当初は鉄道やバスの使い方、ゴミ出しのルール、スーパーの営業時間など、分からないことが多くありました。特に、ドイツではスーパーが日曜日に閉まることや、平日でも閉店時間が比較的早いことに最初は戸惑いました。一方で、Deutschland-Ticket (D-Ticket) を使えば、ICE などの長距離高速列車には乗れないものの、S-Bahn や地域列車、路線バスには広く乗ることができ、移動には便利でした。電車の遅延や突然のホーム変更に遭遇することもありましたが、その都度調べたり周囲の様子を見たりしながら対応しました。ドイツ語はほとんどできない状態で渡航しましたが、簡単な挨拶や日常的な言葉は生活の中で少しずつ覚えることができ、問題なく過ごすことができました。

滞在していたカイザースラウテルンは落ち着いた町で、夜遅くでも身の危険を感じることはほとんどありませんでした。前回の米国滞在と比べると、生活の雰囲気や交通機関の使い方、店の営業時間などに大きな違いがありましたが、環境が変わっても調べて理解し、少しずつ適応していけば問題なく生活できることを実感しました。

2.3 文化面

文化面では、研究と生活の切り替えを大切にする姿勢が印象に残りました。研究者の方々だけでなく受付や事務の方々も含め、勤務時間が終わると夕方頃には帰宅される方も多く、長期休暇も十分に取る文化があるように感じました。渡航前の年末年始には秘書の方々が長期休暇に入られるため、事前の手続きは早めに進める必要がありました。

一方で、研究所内では人とのつながりも大切にされていました。昼食を集まって一緒に食べることが多く、コーヒープレイクの時間も日常的な交流や研究に関する議論の場として機能していました。全体として静かな方が多い印象でしたが、話しかけると親切に対応してくださる方が多く、落ち着いた雰囲気の中で研究に集中できました。また、若い方は英語を話せる方が多い一方で年配の方の中にはドイツ語が中心の方もおり、世代による言語環境の違いも感じました。こうした違いを含めて、現地の文化に触れながら生活できたことは単なる滞在以上の学びになりました。

3 本プログラムの成果・意義

本プログラムの最大の成果は、海外の研究環境に身を置きながら、自分が国際的な研究者・技術者として何を目指し、どのように振る舞い、どのように成長していきたいのかを改めて確認できたことです。前回の RIES プログラムに続く 2 度目の海外研究経験を通じて、場所や環境が変わっても、自分がそのときになすべきことに集中する姿勢は変わらないと感じました。最初は交通、住居、言語、研究所の仕組みなど分からないことも多くありましたが、一つずつ調べて適応していけば、環境の違いはすぐに乗り越えられるということを改めて実感しました。

この感覚は、前回の米国滞在でも得ていたものですが、今回ドイツというまた異なる環境で研究と生活を経験したことで自分自身の中でより一般化され、より確かなものとなりました。今後どのような場所に身を置くことになっても、自分のなすべきこと、そして自分が本当に取り組みたいことを信じて進んでいけばよいという自信につながりました。

また、現地ではメンターだけでなく、研究所の多くの方々と知り合うことができました。研究の議論だけでなく、生活面での助言や日常的な交流を通じて得られたつながりは、今後の人生にとって貴重な財産です。学部生の段階でこのような国際的な研究環境に参加し、実際に研究成果をまとめて発表する機会をいただけたことには非常に大きな意義があったと感じています。

4 おわりに

今回の滞在中を通じて、再び多くの方々に支えていただきながら、研究面でも生活面でも貴重な経験を積みさせていただくことができました。前回の報告書では、海外で得た知識や経験を将来の研究開発に活かして社会に貢献できる人材になりたいと述べましたが、その思いは変わらないどころかむしろさらに強まりました。

同時に、学部生の段階で海外の研究環境に参加し、自分の関心のある研究に主体的に取り組む機会をいただけたことに、改めて深く感謝しております。今回の経験は、多くの方々のご支援があって初めて実現したものであり、決して当たり前のものではないと感じています。この経験を今後の研究活動の糧とし、学んだことを少しずつ形にしなが、より広く社会に貢献できるよう努めてまいります。

5 謝辞

本プログラムに際し、多くの皆様からご支援とご協力を賜りました。この場をお借りして、改めて心より感謝申し上げます。

まず、多大なるご支援を賜りました公益財団法人中谷財団の皆様、並びに関係者の皆様に深く感謝申し上げます。渡航前からプログラム期間中にわたり、手厚くサポートして下さった中谷財団の小川様、藤川様、DFKI の Andrea Pfeiffer 様にも厚く御礼申し上げます。また、渡航に際してさまざまなご手配を行って下さったジャパンスタディーツアーの林様に感謝いたします。

そして、私を研究所に快く受け入れて下さった Prof. Dr. Prof. h.c. Andreas Dengel 先生、Smart Data & Knowledge Services 部門の皆様、そして日々親身にご指導下さった Dr. Stanislav Frolov 先生に深く感謝いたします。研究の進め方から結果の解釈、発表に至るまで多くのご助言をいただき、貴重な経験を積むことができました。

また、現地では研究だけでなく生活面についても多くの方々のお世話になりました。渡航前から日本人コ

コミュニティへのご招待くださり、滞在先などについてもご助言くださった渡邊洸さん、現地での生活について普段から気にかけてくださった野々宮悠太さん、そして同じ研究チームのメンバーとして日々の研究活動を共にくださった高橋陸さんには特に感謝しています。その他、現地で温かく接してくださった DFKI の研究者の皆様にも心より御礼申し上げます。

皆様の温かいご支援とご協力がなければこのような貴重な経験をすることはできませんでした。改めて深く感謝申し上げます。

参考文献

- [1] Thomas Scialom, Tuhin Chakrabarty, and Smaranda Muresan. Fine-tuned language models are continual learners. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 6107–6122, 2022.
- [2] Yifan Wang, Yafei Liu, Chufan Shi, Haoling Li, Chen Chen, Haonan Lu, and Yujiu Yang. InscL: A data-efficient continual learning paradigm for fine-tuning large language models with instructions. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 663–677, 2024.
- [3] Rishabh Agarwal, Nino Vieillard, Yongchao Zhou, Piotr Stanczyk, Sabela Ramos Garea, Matthieu Geist, and Olivier Bachem. On-policy distillation of language models: Learning from self-generated mistakes. In *The Twelfth International Conference on Learning Representations*, 2024.
- [4] Idan Shenfeld, Mehul Damani, Jonas Hübötter, and Pulkit Agrawal. Self-distillation enables continual learning. *arXiv preprint arXiv:2601.19897*, 2026.
- [5] Qwen Team. Qwen2.5 technical report, 2025.
- [6] Shishir G. Patil, Huanzhi Mao, Fanjia Yan, Charlie Cheng-Jie Ji, Vishnu Suresh, Ion Stoica, and Joseph E. Gonzalez. The berkeley function calling leaderboard (bfcl): From tool use to agentic evaluation of large language models. In *Proceedings of the 42nd International Conference on Machine Learning*, pages 48371–48392, 2025.
- [7] Chen Chen, Xinlong Hao, Weiwen Liu, Xu Huang, Xingshan Zeng, Shuai Yu, Dexun Li, Yuefeng Huang, Xiangcheng Liu, Wang Xinzhi, and Wu Liu. Acebench: A comprehensive evaluation of llm tool usage. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 12970–12998, 2025.
- [8] Qwen Team. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025.
- [9] Hunter Lightman, Vineet Kosaraju, Yuri Burda, Harrison Edwards, Bowen Baker, Teddy Lee, Jan Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. Let’s verify step by step. In *The Twelfth International Conference on Learning Representations*, 2024.
- [10] Ansong Ni, Jeevana Priya Inala, Chenglong Wang, Oleksandr Polozov, Christopher Meek, Dragomir Radev, and Jianfeng Gao. Learning math reasoning from self-sampled correct and partially-correct solutions. In *The Eleventh International Conference on Learning Representations*, 2023.
- [11] Tergel Munkhbat, Namgyu Ho, Seo Hyun Kim, Yongjin Yang, Yujin Kim, and Se-Young Yun. Self-training elicits concise reasoning in large language models. In *Findings of the Association for*

Computational Linguistics: ACL 2025, 2025.

- [12] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. In *The Tenth International Conference on Learning Representations*, 2022.