

Hackathon on Large Language Models and Earthquake Researches

*Hisahiko KUBO¹, Stephen Wu², Masayuki Kano³, Shinya Katoh⁴, Atsuko Oana⁵, Tomohisa Okazaki⁶, Nozomi Okada⁷, Nobuki Kame⁴, Yuki Kodera⁸, Daisuke Sato⁹, Kengo Shimojo⁸, Koji Tamaribuchi¹⁰, Makoto Naoi¹¹, Ryuichi Nishiyama⁴, Kazuro Hirahara^{6,12}, Takashi Miyamoto¹³, Masumi Yamada¹⁴

1. National Research Institute for Earth Science and Disaster Resilience, 2. The Institute of Statistical Mathematics, 3. Tohoku University, 4. Earthquake Research Institute, The University of Tokyo, 5. Institute of Technology, Shimizu Corporation, 6. Center for Advanced Intelligence Project, RIKEN, 7. Shizuoka University, 8. Meteorological Research Institute, 9. Japan Agency for Marine-Earth Science and Technology, 10. Japan Meteorological Agency, 11. Hokkaido University, 12. Kagawa University, 13. University of Yamanashi, 14. Disaster Prevention Research Institute, Kyoto University

ChatGPTに代表される大規模言語モデルに関する近年の技術革新およびサービス展開は目覚ましく、私たちの生活の様々な場面で幅広く使われている。研究や教育の分野でも徐々に取り入れられ始めているが、研究活動そのものへの適用は限定的であり、それぞれの学術分野において模索されている段階である。こうした背景の中、私たちは地震研究における大規模言語モデルの活用をテーマとしたハッカソンを2024年夏に実施する。ハッカソン自体は日本地震学会2024年度秋季大会への講演申込の締め切り後に行うため、ここではハッカソンに向けた準備を中心として記載する。ハッカソンは、様々なバックグラウンドを持つ人々が集まり、特定のテーマに関してのアイデアを短期間で形にするイベントである。今回は「大規模言語モデル×地震研究」というテーマで、2024年夏に一泊二日の予定で実施する。これに向けて、まず世話人である久保・Wu・加納・加藤にてハッカソンの立ち上げおよび具体化を2024年初頭から春にかけて行った。そして2024年5月下旬から6月上旬にかけて参加者を募集し、最終的には世話人含めて17名が参加することになった。その後ハッカソンに向けた準備として以下を行った：・専門家による大規模言語モデルの仕組みの解説・大規模言語モデルのチュートリアルの実施・参加者からのアイデア出しおよびディスカッション・絞り込んだ課題ごとのグループディスカッションこれらの準備を経て、以下の4つの課題に絞り込み、それぞれのグループでアイデアの深化および実装に向けた議論を進めている：・多数のシミュレーション結果に関する特徴抽出・解説付与・地震に関するソーシャルデータの自動収集・解析・発生した地震に関する解説や想定問答の自動生成・シミュレーションコードの自動生成なおハッカソンの進行に応じてこれらの内容が変更される可能性がある。またチャットベースのコマンドプロンプトによる大規模言語モデルの活用だけでなく、Retrieval-Augmented Generationやマルチモーダル技術の活用も視野に入れた検討を行っている。日本地震学会2024年度秋季大会での発表時には、ハッカソンの成果も含めて詳細を報告する予定である。