2025 IMEKO Joint Conference
TC8 - TC11 - TC25
Torino, Italy, September 14-17, 2025

Assessing the Viability of On-Device Language Models in Legally Controlled Measuring Instruments

Mahbuba Moni[1], Daniel Peters[1], Florian Thiel[1], Axel Sikora[2]

[1] Physikalisch-Technische Bundesanstalt (PTB), Berlin, Germany, mahbuba.moni@ptb.de, daniel.peters@ptb.de, florian.thiel@ptb.de
[2] Hochschule Offenburg, Offenburg, Germany, axel.sikora@hs-offenburg.de

***Abstract* – The European Union's Artificial Intelligence Act represents a significant regulatory framework that governs the deployment of AI systems across various sectors, including critical infrastructure. A crucial challenge identified is the integration of AI in measuring instruments, especially utility meters used in critical infrastructures, as highlighted in the EU AI Act's Annex III. In legal metrology, AI applications, especially language models, are primarily server-based, requiring substantial resources and continuous connectivity. The integration of AI and specifically of language models into measuring instruments promises unique opportunities and challenges along their lifecycle, impacting manufacturers, notified bodies, market surveillance authorities, and consumers. This paper highlights the need for on-device language models that can perform real-time regulatory compliance checks, facilitate natural-language interactions, and generate legally admissible documentation. We also address the challenges of designing smaller, energy-efficient models and propose a system overview to enhance functionality in resource-constrained environments.**

I. INTRODUCTION

The enforcement of the European Artificial Intelligence Act (EU AI Act) [1] aims to foster the deployment of safe and trustworthy artificial intelligence (AI) systems across Europe's single market. Simultaneously, measuring instruments subject to legal control are governed by stringent directives, specifically the Measuring Instruments Directive (MID, 2014/32/EU) [2] and the Non-Automatic Weighing Instruments Directive (NAWID, 2014/31/EU) [3]. Collectively, these directives encompass 14 distinct categories, affecting approximately 345 million instruments annually in Europe [4]. Given this extensive reach, the intersection between legal metrology and artificial intelligence (AI), particularly Generative AI [5][6] becomes increasingly significant. Generative AI is a subset of AI that uses generative models to create new data like text, images, or videos by learning patterns from training data. Within this field, Language Model (LM) [7] is trained to understand and generate human language especially for text and can be categorized as large and small according to the training datasets and parameter sizes. Recent advancements in transformer-based [8] deep neural networks [9], particularly Large Language Models (LLMs) [10], enhance the ability to effectively reason, understand, generate, and engage with human language. In contrast, Small Language Models (SLMs) [11] are designed to perform similar language tasks with fewer parameters, optimized for efficiency and deployed in resource-constrained environment.

Currently, the use of generative AI in legal metrology is limited. Previous works have addressed assessing the effectiveness of language models for repairing vulnerabilities in code [12], a cybersecurity framework that influences active learning strategies [13] and automating access control policy generation [14]. Another significant work has demonstrated the potential of language models in automating parts of the conformity assessment process for legally regulated instruments under the European Measuring Instruments Directive (MID) [15]. This prototype in [15] has shown that language models can assist in analyzing and validating software documentation through natural language interpretation and compliance mapping.

However, many of these solutions are typically server-based, resource-intensive, and require constant internet connectivity, which makes them less effective in embedded and resource-constrained measuring instrument environments. In parallel, there is a rising interest in using on-device language models [16][17][18]. Modern utility meters, including electricity, gas, water, and weighing instruments, are intelligent endpoints on the Internet of Things (IoT) ecosystem. Integrating resource-efficient language models (LMs) in legally controlled measuring instruments can enable real-time regulatory compliance checks, natural-language diagnostic interactions, instant generation of legally admissible documentation, and delivery of responses that can be traced and verified for compliance. However, a review of publications reveals

[12-18] that no prior works have explored dataset mapping and evaluation for the on-site conformity task, especially for legally controlled measuring instruments. This involves challenges such as making models smaller, energy-efficient, and creating ways to measure their performance. Therefore, it will be worth exploring the viability of these technologies by applying and evaluating them while deploying in measuring instruments. Language models (LMs) typically require significant computational resources and are prone to generating hallucinations or false outputs. When using a language model with measuring devices, it is also important to consider resource constraints and regulatory requirements, as well as preventing the display of incorrect responses and readings.

This paper presents several key contributions that advance the field of legally controlled measuring instruments. It is structured as follows: Section II presents an overview of regulatory frameworks, considering the EU AI Act and MID software requirements. Section III discusses the advancement and incorporation of resource-efficient language models. Section IV proposes a system overview based on the advancement of the language model. Finally, Sections V and VI highlight future directions for further research in this field, concluding the paper.

II. REGULATORY FRAMEWORKS

The EU AI Act [1] is broadly defined, with classifications depending on their integration into regulated measuring instruments under MID [2][3]. Besides, WELMEC 7.2 Software Guide [19] provides guidance on software for measuring instruments. In this discussion, we explore how Language Models (LMs) can enhance the functional areas by interpreting logs and summarizing activities by considering the EU AI Act and MID rules.

A. The EU AI Act

Article 6 [1] of the EU AI Act outlines the classification rules for high-risk AI systems. Instead of directly describing AI systems, it references products that fall under the Union harmonization legislation listed in Annex I and the application areas defined in Annex III [1]. There are two application areas in Annex III where measuring instruments regulated under MID are used, these are critical infrastructures, e.g. for the supply of electricity, gas, heat and water, and law enforcement. The European AI Act defines AI systems with a broad scope, aiming to regulate only complex AI systems to foster innovation. These complex AI modules generate predictions for recommendations and decisions based on specific objectives. They may fall under the AI Act's definitions and obligations when integrated into MID-regulated measuring instruments. An AI product is classified as high-risk if it appears in Annex I of the relevant Union legislation or one of the eight application areas in Annex III (Article 6) [1]. If neither applies, it must be checked against the four areas in Article 50, which require at least transparency obligations [1].

Non-high-risk AI systems or those exempt from transparency requirements are encouraged to follow voluntary codes of conduct. If the AI module is a general-purpose AI (GPAI) [1] and present systemic risks at the Union level, must comply with additional obligations under Article 53 and Article 55 [1]. The text outlines that while a system may be defined as an AI system under the European AI Act, it is not always required to meet specific obligations. MID measuring instruments only falls under this definition if they include an AI module relevant to the Act. Currently, without amendments to Annex III regarding high-risk AI systems, MID instruments adhere to voluntary codes of conduct. If a MID measuring instrument has an AI module that produces non-metrological outputs, it must comply with both the AI Act's high-risk obligations and the MID requirements, leading to potential redundancy in conformity assessments and market surveillance from different bodies. To prevent complications, it's advisable to maintain a clear separation between metrological software and non-metrological AI modules. This separation allows updates to the AI modules without needing re-verification, while also facilitating adaptable parameters post-deployment. Thus, even if a metrological system qualifies as AI, products under MID [2] and NAWID [3] may not need to meet AI Act obligations.

B. Enhancing compliance tasks by using Language Models

Since most of AI models are implemented in software, they fall under the corresponding requirements in the directives. Our paper considers AI modules as an optional tool and as a non-legally relevant software module to comply with WELMEC 7.2 Software guide [19]. These modules are designed to generate user-friendly reports, ensuring the core measurement characteristics are not affected.

Here, we can see it as an example. An audit trail requires continuous information that includes time-stamped records of events. This includes changes in the values of measuring instruments' parameters, software updates, and other legally relevant and critical activities for metrological characteristics. An embedded language model (LM) can interpret these logs and summarize system activity. It can also respond to natural language queries such as "What changes occurred in the last 10 days?" or "Have there been any unauthorized access attempts?". The LM processes this and retrieves relevant entries from the audit log, displaying a sequential record of critical events such as calibration timestamp, software updates, configuration changes, and the users or processes involved. LM can also analyze firmware update logs, summarize the actions taken during updates, and explain the workflow in human-readable terms. These models act as intelligent agents between raw technical data and regulatory expectations. Several WELMEC 7.2 [19] functional areas can benefit from LMs' interpretive and summarization strengths, especially when deployed directly on-device.

III. ADVANCEMENTS IN RESOURCE-EFFICIENT LANGUAGE MODELS

Language Model (LM) is a key element in Natural Language Processing (NLP), designed to predict the probability of word sequences by analyzing language patterns from large text collections [7]. Resource-efficient language model is designed to deliver high quality responses while minimizing computational cost, usage of memory and energy consumption. Here we explore the techniques for adapting efficient language models for directly deploying on constrained devices [17][26].

A. Efficient Language Models

Transformer-based models like BERT [20], GPT- 4 [21], and LLaMA [22] have led to significant advances in text classification, machine translation, and question answering tasks. One of the biggest challenges with large language models (LLMs) is their size [11]. For instance, GPT-3 has 175 billion (175B) parameters [17] and BERT-base includes 110 million parameters [30] and requires significant memory and computing power. Small Language Models (SLMs) address this issue by reducing the model size. With recent progress in model compression and edge inference, language models with tiny versions can now be deployed on devices like the Raspberry Pi [23][24] and NVIDIA Jetson Nano [25]. For example, Llama3.2 has lightweight, text-only models (1B and 3B parameters) designed to fit on edge and mobile devices. Phi 3.5 Mini (3.8 billion parameter) language model is small enough to be deployed on a phone. Gemma3 (4B parameter) models are designed to handle text and image input and generate output [11][27].

B. On-device model compression techniques

Small Language Models (SLMs) are often deployed with lower-precision data types, e.g., instead of the standard 32-bit floating point (float32), using 8-bit integers (int8) [28]. Quantized weights are used to enable memory-efficient inference. Zero-shot methods, including LLM.int8()[29], NF4 (normal float), and FP4 (4-bit float), normalize parameters uses a scaling operation before assigning them to specific quantization masses. Optimization-based methods focus on reducing quantization errors by adapting the calibration dataset [28][29].

Pruning [26] is another effective model compression technique designed to reduce the number of parameters in a neural network. It works by removing less critical connections or entire neurons. There are two types of pruning. One is unstructured pruning, which eliminates unnecessary weights between the layers. Another is structured pruning, which removes complete connections to specific weights. The structured approach creates a more organized network that is easier to work with on a hardware-friendly network architecture [26].

Another approach called Knowledge distillation which allows a smaller student model to learn from a larger teacher model, such as DistilBERT [30] and TinyBERT [31]. DistilBERT [30], a distilled version of BERT that is 40% smaller and 60% faster, nonetheless retains 97% of the original model's language understanding capabilities. On the other hand, TinyBERT [31] conducts transformer distillation during the pretraining and task-specific learning phases. It effectively captures both general-domain knowledge and task-specific information from BERT. These models typically have fewer parameters, making them suitable for embedded natural language processing (NLP) applications. However, their reduced capability may limit performance on more complex tasks, suggesting a need for further optimizations.

IV. PROPOSED SYSTEM OVERVIEW

Here, we discuss briefly how lightweight language models (LMs) can be used in measuring instruments for efficient language representation. An anomaly on-device filter is proposed to log suspicious events to cross-check the accuracy of LM responses. The proposed architecture combines on-device assessments with verification methods, which improves measurement integrity and prevents unverified information from reaching users or inspectors. Additionally, integrating Generative Adversarial Networks (GANs) [33] is proposed for verifying model outputs, as they can effectively identify compliant and non-compliant data.

A. Embedded lightweight language models

Smaller, pre-optimized models, such as DistilBERT [30] is often beneficial while deploying on a measuring instrument. These distilled models are designed to provide strong language representation, using additional compression techniques like quantization or pruning. Research indicates that these smaller models tend to manage moderate quantization techniques effectively. Initially, it will be necessary to perform lightweight fine-tuning to adapt these models for domain-specific tasks, such as regulatory compliance or diagnostic reporting. One possible approach can be to use parameter-efficient methods, such as Low-Rank Adaptation (LoRA) [32], which inserts small adapter modules into the model and avoids updating the entire parameter set. This approach significantly reduces training overhead and is particularly useful for on-device or edge deployment scenarios.

B. Verification of the Language Models

Before any output (e.g., responses from the query asked by the market surveillance authority/inspectors) from the language model (LM) is shown on the display of measuring instruments, it must first go through a lightweight filter on the device. This on-device filter quickly checks whether the output makes sense or may be wrong. The process uses simple rules and basic machine-learning checks. For example, with a weighing instrument, the filter ensures that any software version mentioned by the LM matches the audit trail. Every flagged event gets a timestamp and is saved in the device's non-volatile memory, creating a permanent record if the LM produces suspicious output. This approach incorporates a hybrid strategy in terms of overall system architecture.

Therefore, we propose the additional use of Generative Adversarial Networks (GANs) [33][34]. GANs are a machine learning framework with two parts: a generator and a discriminator. The generator learns from real data (e.g. knowledge base) to create non-compliant examples that look real. Meanwhile, the discriminator is trained to tell the difference between compliance and non-compliance data sets produced by the generator.

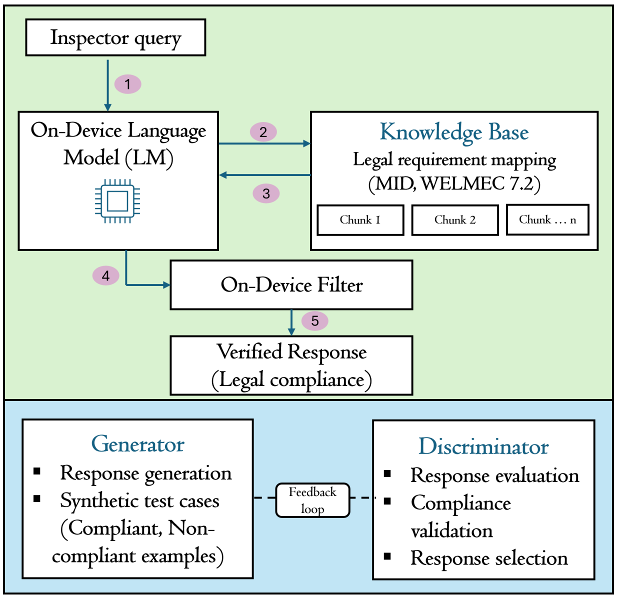

Fig. 1. Proposed integrated approach for verified on-device language model response

C. Integrated approach with knowledge retrieval and GAN architecture

Here, Fig. 1 illustrates integrated approach for verified on-device language model (LM) and the components of GANs. The system functions are represented as follows:

(1) System receives the natural language query from the market surveillance authority/inspectors. For e.g. "What changes occurred in the last 10 days?" The on device compressed language model (LM) processes the query and interact with knowledge base. (2) Knowledge base contains the chunked information with corresponding vector embeddings (here, the information contains the legal requirement mapping based on MID and WELMEC 7.2 guide). (3) On-device LM then retrieve and provide the context for response generation. (4) Generative Adversarial Network (GAN) [33] is trained using off-device compliant and non-compliant data. The discriminator learns the normal patterns and later deployed to the device. It can then be used as on device filter running at application level. This filter can handle quick decisions and provide real time responses. In advance level, the filter logs the anomalies and later sent back to retrain GAN to improve accuracy over time. (5) The on-device filter passes the verified response based on the evaluation criteria to the inspector.

In GAN-based architecture, generator tries to generate synthetic test cases (compliant, non-compliant examples) and create responses. It then passes the full set of responses to the discriminator. The transfer includes both text responses and confidence scores. Discriminator evaluates the responses against reference knowledge. It then provides an iterative feedback loop for improving the generator in parallel. This loop continues until a quality threshold is reached. The whole process can be performed in off-device settings during training phase.

The proposed approach can be instantaneous for simple queries on devices with obvious answers in the knowledge base. For complex queries, it will require multiple iterations between the generator and discriminator. The key advantage is that the inspector only sees the final and verified response that passed from the on-device filter. Here, proposed GANs are for training so they are used to train the filter. Our on-device filter itself is not a GAN, because it is running on the application level and not at the training level. Further research and tests will be carried out from the perspective of device selection based on computing power.

V. FUTURE DIRECTIONS

Critical technical challenges exist when implementing resource-efficient language models in legally controlled measuring instruments. One key issue is figuring out the best ways to enhance computational efficiency. We need to consider techniques like model quantization, pruning, knowledge distillation, and edge-optimized architecture designed explicitly for constrained hardware. Evaluation metrics should include accuracy and task-specific metrics, latency and throughput, energy consumption, and trade-offs between model size, accuracy, and resource consumption. Additionally, further research will incorporate existing implementations, designs, and methods that facilitate secure and efficient combinations while complying with legally relevant standards, including various EU regulations.

The security and integrity of embedded LMs also pose significant research challenges. Developing strategies to safeguard embedded AI systems against tampering, malicious attacks, and unauthorized alterations is essential. Approaches may include integrating cryptographic signatures into model parameters, conducting real-time integrity checks using cryptographic hashes, or employing anomaly detection models trained on operational data. A comprehensive evaluation of these embedded models through penetration testing and rigorous security assessments will be necessary to ensure they are resilient against adversarial threats. Additionally, using explainable AI (XAI) outputs would enhance accountability and traceability. Therefore, we plan to establish a dedicated testbed to rigorously test and validate these LM solutions, ensuring they meet regulatory standards and perform effectively in real-world conditions.

VI. CONCLUSION

The paper explores the combination of artificial intelligence (AI) and legal metrology, focusing on resource-efficient on-device language models. We explore the potential of language models (LMs) to enhance regulatory compliance and provide real-time outputs to users and inspectors. The proposed system's Language Models (LMs) are linked to regulatory sources such as MID [2][3] and WELMEC 7.2 guide [19]. This connectivity is also helpful for cloud-hosted language models (LMs) that offer AI-as-a-Service for legal metrology. In general, the paper provides a framework that ensures AI integration maintains the integrity of legally relevant software. Our paper highlighted the need for smaller, energy-efficient models tailored for on-site conformity assessments and addressing challenges like performance evaluation and data mapping according to guidelines. Furthermore, we have introduced AI-driven interfaces that can guide users and inspectors, enhancing the usability and functionality of measuring instruments. Through further research, we aim to identify pathways for deploying AI that ensure regulatory compliance in a rapidly evolving technological landscape.

REFERENCES

[1] EU AI Act, Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) https://eur-lex.europa.eu/eli/reg/2024/1689/oj (Accessed 24 April 2025)

[2] Directive 2014/32/EU, DIRECTIVE 2014/32/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of measuring instruments (recast) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02014L0032-20150127 (Accessed 24 April 2025)

[3] DIRECTIVE 2014/31/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of non-automatic weighing instruments https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0031 (Accessed 24 April 2025)

[4] F. Thiel, "Digital transformation of legal metrology—The European metrology cloud." OIML Bulletin Volume LIX, (2018), pp.10–21.

[5] IBM, Generative AI. https://www.ibm.com/think/topics/generative-ai (Accessed 24 April 2025)

[6] S. Bengesi, H. El-Sayed, M. K. Sarker, Y. Houkpati, J. Irungu and T. Oladunni, "Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers," in IEEE Access, vol. 12, pp. 69812-69837,(2024), doi: 10.1109/ACCESS.2024.3397775. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10521640 (Accessed 24 April 2025)

[7] D. Jurafsky, J. H. Martin. "Speech and Language Processing (third ed. draft)" (2025) https://web.stanford.edu/~jurafsky/slp3/ (Accessed 24 April 2025)

[8] A. Vaswani et al., "Attention is all you need". Advances in Neural Information Processing Systems (NeurIPS), vol. 30 (2017) pp. 5998–6008, https://doi.org/10.48550/arXiv.1706.03762 (Accessed 24 April 2025)

[9] M. Rüb, Prof. Dr. A. Sikora, "A Practical View on Training Neural Networks in the Edge", IFAC-PapersOnLine, Volume 55, Issue 4, (2022), Pages 272-279, ISSN 2405-8963, https://doi.org/10.1016/j.ifacol.2022.06.045 (Accessed 29 April 2025)

[10] T. B. Brown, et al (2020) "Language models are few-shot learners." In Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS '20). Curran Associates Inc., Red Hook, NY, USA, Article 159, 1877–1901. https://doi.org/10.48550/arXiv.2005.14165 (Accessed 24 April 2025)

[11] C. Van Nguyen et al. "A Survey of Small Language Models." (2024) https://doi.org/10.48550/arXiv.2410.20011 (Accessed 24 April 2025)

[12] L. Zhang, Q. Zou, A. Singhal, X. Sun, and P. Liu. (2024). "Evaluating Large Language Models for Real-World Vulnerability Repair in C/C++ Code." In Proceedings of the 10th ACM International Workshop on Security and Privacy Analytics

https://doi.org/10.1145/3643651.3659892, (Accessed April 24, 2025)

[13] F. I. Rahman, S. M. Halim, A. Singhal, and L. Khan. (2024), "ALERT: A Framework for Efficient Extraction of Attack Techniques from Cyber Threat Intelligence Reports Using Active Learning." Data and Applications Security and Privacy XXXVIII (DBSec 2024), San Jose, CA, USA, https://doi.org/10.1007/978-3-031-65172-4_13 , (Accessed April 24, 2025)

[14] S. Lawal, X. Zhao, A. Rios, R. Krishnan and D. Ferraiolo, "Translating Natural Language Specifications into Access Control Policies by Leveraging Large Language Models" (2024). IEEE 6th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA), Washington, DC, USA, https://doi.org/10.1109/TPS-ISA62245.2024.00048 (Accessed April 24, 2025)

[15] J. Litzinger, J. Neumann, D. Peters & F. Thiel. "Streamlining conformity assessment of software applying large language models." (2025). Measurement Sensors,101792, ISSN 2665-9174, https://doi.org/10.1016/j.measen.2024.101792 (Accessed April 24, 2025)

[16] J. Xu et al, "On-Device Language Models: A Comprehensive Review." (2024). https://doi.org/10.48550/arXiv.2409.00088 (Accessed April 24, 2025)

[17] G. Bai et al. "Beyond Efficiency: A Systematic Survey of Resource-Efficient Large Language Models." CoRR (2024). https://doi.org/10.48550/arXiv.2401.00625 (Accessed 24 April 2025)

[18] M. W. Ur Rahman et al., "Quantized Transformer Language Model Implementations on Edge Devices." https://doi.org/10.48550/arXiv.2310.03971 (Accessed April 24, 2025)

[19] WELMEC Guide 7.2 Software guide (EU measuring instruments directive 2014/32/EU https://www.welmec.org/welmec/documents/guides/7.2/2023/WELMEC_Guide_7.2_2023.pdf (Accessed April 24, 2025)

[20] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." In NAACL-HLT, (2018) https://doi.org/10.48550/arXiv.1810.04805 (Accessed April 24, 2025)

[21] OpenAI, "GPT-4 technical report," arXiv preprint (2023).https://doi.org/10.48550/arXiv.2303.08774 (Accessed April 24, 2025)

[22] H. Touvron et al. (2023). "LLaMA: Open and Efficient Foundation Language Models." https://doi.org/10.48550/arXiv.2302.13971 (Accessed April 24, 2025)

[23] E. J. Husom, et al. "Sustainable LLM Inference for Edge AI: Evaluating Quantized LLMs for Energy Efficiency, Output Accuracy, and Inference Latency." (2025). https://doi.org/10.48550/arXiv.2504.03360 (Accessed April 24, 2025)

[24] Raspberry Pi: https://www.raspberrypi.com/ (Accessed 24 April 2025)

[25] Nvidia Jetson nano: https://developer.nvidia.com/embedded/jetson-nano (Accessed 24 April 2025)

[26] D. Liu et al. "A survey of model compression techniques: past, present, and future." Frontiers in Robotics and AI. (2025) https://doi.org/10.3389/frobt.2025.1518965 (Accessed 24 April 2025)

[27] Small-Language-Models https://huggingface.co/blog/jjokah/small-language-model (Accessed 24 April 2025)

[28] K. Egashira, M. Vero, R. Staab, J. He and M. Vechev. "Exploiting LLM quantization." (2024) DOI:10.48550/arXiv.2405.18137 https://doi.org/10.48550/arXiv.2405.18137 (Accessed 24 April 2025)

[29] T. Dettmers, M. Lewis, Y. Belkada, and L. Zettlemoyer. "LLM.int8(): 8-bit matrix multiplication for transformers at scale." (2022). https://doi.org/10.48550/arXiv.2208.07339 (Accessed 24 April 2025)

[30] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, "DistilBERT, a distilled version of BERT: smaller, faster, cheaper, and lighter." (2019) https://doi.org/10.48550/arXiv.1910.01108 (Accessed 24 April 2025)

[31] X. Jiao et al. (2019). "TinyBERT: Distilling BERT for Natural Language Understanding." https://doi.org/10.48550/arXiv.1909.10351 (Accessed 24 April 2025)

[32] E. J. Hu, et al. (2021). "LoRA: Low-Rank Adaptation of Large Language Models." 10.48550/arXiv.2106.09685. https://doi.org/10.48550/arXiv.2106.09685 (Accessed 24 April 2025)

[33] I. Goodfellow et al. (2020). "Generative adversarial networks." Communications of the ACM, 63(11), 139–144. https://doi.org/10.1145/3422622 (Accessed 24 April 2025)

[34] J. Y. Wang, et al. "LLM-GAN: Construct Generative Adversarial Network Through Large Language Models For Explainable Fake News Detection." (2024). https://doi.org/10.48550/arXiv.2409.01787 (Accessed 24 April 2025)