Continual Learning for Large Language Models: A Survey

Anonymous ACL submission

Abstract

Large language models (LLMs) are challenging to retrain frequently due to the high costs associated with their massive scale. However, updates are necessary to equip LLMs with new skills and keep them current with rapidly evolving human knowledge. This paper surveys recent works on continual learning for LLMs. We introduce a novel multi-staged categorization scheme for continual learning techniques, encompassing continual pre-training, instruction tuning, and alignment. We compare continual learning for LLMs with simpler adaptation methods used in smaller models and other enhancement strategies such as retrieval-augmented generation and model editing. Additionally, informed by a discussion of benchmarks and evaluations, we identify several challenges and future research directions for this critical task.

1 Introduction

Recent years have witnessed the rapid advances of large language models' capabilities in solving a diverse range of problems. At the same time, it is vital for LLMs to be regularly updated to accurately reflect the ever-evolving human knowledge, values and linguistic patterns, calling for the investigation of *continual learning* for LLMs. Whilst continual learning bears some resemblance to other strategies for model improvements, such as retrieval-augmented generation (RAG) (Lewis et al., 2020) and model editing (Yao et al., 2023), their main purposes differ (Table 1). Unlike these strategies, whose primarily focus is on refining the domain-specific accuracy or expanding the model's factual knowledge base, continual learning aims to enhance the overall linguistic and reasoning capabilities of LLMs. This distinction is crucial as it shifts the focus from merely updating information to developing a model's ability to process and generate language in a more comprehensive and nuanced manner (Zhang et al., 2023e).

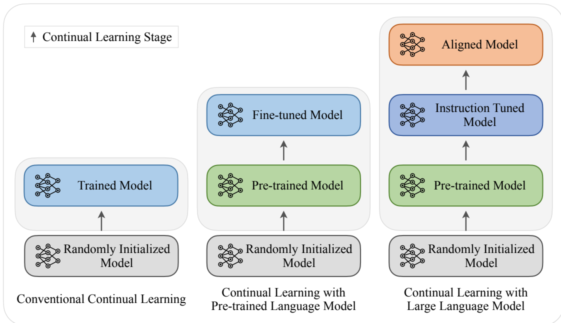

Figure 1: Continual learning for large language models involves hybrid multi-stage training with multiple training objectives.

Continual learning for LLMs also differs from its use in smaller models, including smaller pre-trained language models. Due to their vast size and complexity, LLMs require a multi-faceted approach to continual learning. We categorise it into three different stages, i.e. *continual pre-training* to expand the model's fundamental understanding of language (Jin et al., 2022), *continual instruction tuning* to improve the model's response to specific user commands (Zhang et al., 2023d), and *continual alignment* to ensure the model's outputs adhere to values, ethical standards and societal norms (Zhang et al., 2023a). This multi-stage process is distinct from the more linear adaptation strategies used in smaller models, as illustrated in Figure 1, highlighting the unique challenges and requirements of applying continual learning to LLMs.

This survey differentiates itself from previous studies by its unique focus and structure. While previous surveys in the field are typically organized around various continual learning strategies (Biesialska et al., 2020), ours is the first to specifically address continual learning in the context of LLMs. We structure our analysis around the types of information that is updated continually and the distinct stages of learning involved

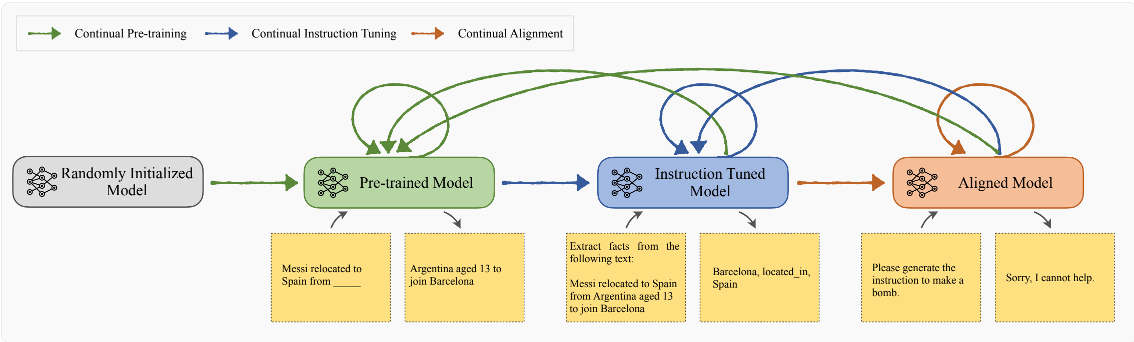

Figure 2: The continual learning of LLMs involves multi-stage and cross-stage iteration, which may lead to substantial forgetting problems. For example, when the instruction-tuned model resumes continual pre-training, it may encounter cross-stage forgetting, resulting in reduced performance on instruction-following tasks.

Information	RAG	Model Editing	Continual Learning
Fact	✓	✓	✓
Domain	✓	×	✓
Language	×	×	✓
Task	×	×	✓
Skills (Tool use)	×	×	✓
Values	×	×	✓
Preference	×	×	✓

Table 1: Continual Learning v.s. RAG and Model Editing

in LLMs. This survey offers a detailed and novel perspective on how continual learning is applied to LLMs, shedding light on the specific challenges and opportunities of this application. Our goal is to provide a thorough understanding of the effective implementation of continual learning in LLMs, contributing to the development of more advanced and adaptable language models in the future.

2 Preliminary and Categorization

2.1 Large Language Model

Large language models (LLMs) like ChatGPT[1] and LLaMa (Touvron et al., 2023) have shown superior performance in many tasks. They are usually trained in multiple stages, including pre-training, instruction tuning, and alignment, as illustrated in Figure 1. In the *pre-training* stage, LLMs are trained on a large corpus in a self-supervised manner (Dong et al., 2019), where the training text is randomly masked and the LLMs are asked to predict the masked tokens. In the *instruction tuning* stage, LLMs are finetuned on a set of instruction-output pairs in a supervised fashion (Zhang et al., 2023b). Given a task-specific instruction as input, LLMs are asked to generate the corresponding output. In the *alignment* stage, LLMs are further finetuned with human feedback to align their outputs with human expectations (Wang et al., 2023d). The output of LLMs is scored by human annotators, and the LLMs are updated to generate more human-like responses.

[1] https://openai.com/blog/chatgpt

2.2 Continual Learning

Continual learning focuses on developing learning algorithms to accumulate knowledge on non-stationary data, often delineated by classes, tasks, domains or instances. In supervised continual learning, a sequence of tasks $\{\mathcal{D}_1, \ldots, \mathcal{D}_T\}$ arrive in a streaming fashion. Each task $\mathcal{D}_t = \left\{\left(\boldsymbol{x}_i^t, y_i^t\right)\right\}_{i=1}^{n_t}$ contains a separate target dataset, where $\boldsymbol{x}_i^t \in \mathcal{X}_t$, $\boldsymbol{y}_i^t \in \mathcal{Y}_t$. A single model needs to adapt to them sequentially, with only access to $\mathcal{D}_t$ at the t-th task. This setting requires models to acquire, update, accumulate, and exploit knowledge throughout their lifetime (Biesialska et al., 2020).

The major challenge conventional continual learning tackles is that of *catastrophic forgetting*, where the performance of a model on old tasks significantly diminishes when trained with new data. Existing studies can be roughly grouped into three categories, e.g., experience replay methods (Chaudhry et al., 2019; Wu et al., 2021), regularization-based methods (Kirkpatrick et al., 2017; Chen et al., 2023b), and dynamic architecture methods (Mallya et al., 2018). Recently, researchers have designed some hybrid methods that take advantage of the aforementioned techniques (Chen et al., 2023a; He et al., 2024). Our paper stands out (Shi et al., 2024) by organizing around multi-stage continual learning and highlighting cross-stage forgetting issues.

2.3 Continual Learning for LLMs

Continual Learning for Large Language Models aims to enable LLMs to learn from a continuous data stream over time. Despite the importance, it is non-trivial to directly apply existing continual learning settings for LLMs. We now provide a forward-looking framework of continual learning for LLMs, then present a categorization of research in this area.

Framework Our framework of continual learning for LLMs is illustrated in Figure 2. We align continual learning for LLMs with the different training stages, including Continual Pre-training (CPT), Continual Instruction Tuning (CIT), and Continual Alignment (CA). The *Continual Pre-training* stage aims to conduct training on a sequence of corpus self-supervisedly to enrich LLMs' knowledge and adapt to new domains. The *Continual Instruction Tuning* stage finetunes LLMs on a stream of supervised instruction-following data, aiming to empower LLMs to follow users' instructions while transferring acquired knowledge for subsequent tasks. Responding to the evolving nature of human values and preferences, *Continual Alignment (CA)* tries to continuously align LLMs with human values over time.

While continual learning on LLMs can be conducted in each stage sequentially, the iterative application of continual learning also makes it essential to transfer across stages without forgetting the ability and knowledge learned from previous stages. For instance, we can conduct continual pre-training based on either instruction-tuned models or aligned models. However, we do not want the LLM to lose their ability to follow users' instructions and align with human values. Therefore, as shown in Figure 2, we use arrows with different colors to show the transfer between stages.

Categorization To better understand the research in this area, we provide a fine-grained categorization for each stage of the framework.

Continual Pre-training (CPT)

- *CPT for Updating Facts* includes works that adapt LLMs to learn new factual knowledge.
- *CPT for Updating Domains* includes research that tailors LLMs to specific fields like medical and legal domains.
- *CPT for Language Expansion* includes studies that extend the languages LLMs supports.

Continual Instruction Tuning (CIT)

- *Task-incremental CIT* contains works that finetune LLMs on a series of tasks and acquire the ability to solve new tasks.
- *Domain-incremental CIT* contains methods that finetune LLMs on a stream of instructions to solve domain-specific tasks.
- *Tool-incremental CIT* contains research that continually teaches LLMs to use new tools to solve problems.

Continual Alignment (CA)

- *Continual Value Alignment* incorporates studies that continually align LLMs with new ethical guidelines and social norms.
- *Continual Preference Alignment* incorporates works that adapt LLMs to dynamically match different human preferences.

Besides categorizing methods based on training stages, we also provide an alternative categorization based on the information updated during continual learning. In Table 2, we list some representative information that is updated for LLMs, e.g., facts, domains, tasks, values, and preferences. Based on the training objectives of LLMs, this information can be updated in different stages of LLM continual learning. The taxonomy in Figure 3 shows our categorization scheme and recent representative work in each category.

Information	Pre-training	Instruction-tuning	Alignment
Fact	✓	×	×
Domain	✓	✓	×
Language	✓	×	×
Task	×	✓	×
Skill (Tool use)	×	✓	×
Value	×	×	✓
Preference	×	×	✓

Table 2: Information updated during different stages of continual learning for LLMs.

3 Continual Pre-training (CPT)

Continual pre-training in large language models is essential for keeping the LLMs relevant and effective. This process involves regularly updating the models with the latest information (Jang et al., 2022a; Ibrahim et al., 2024), adapting them to specialized domains (Ke et al., 2023), enhancing their coding capabilities (Yadav et al., 2023), and expanding their linguistic range (Castellucci et al., 2021). With CPT, LLMs can stay current with

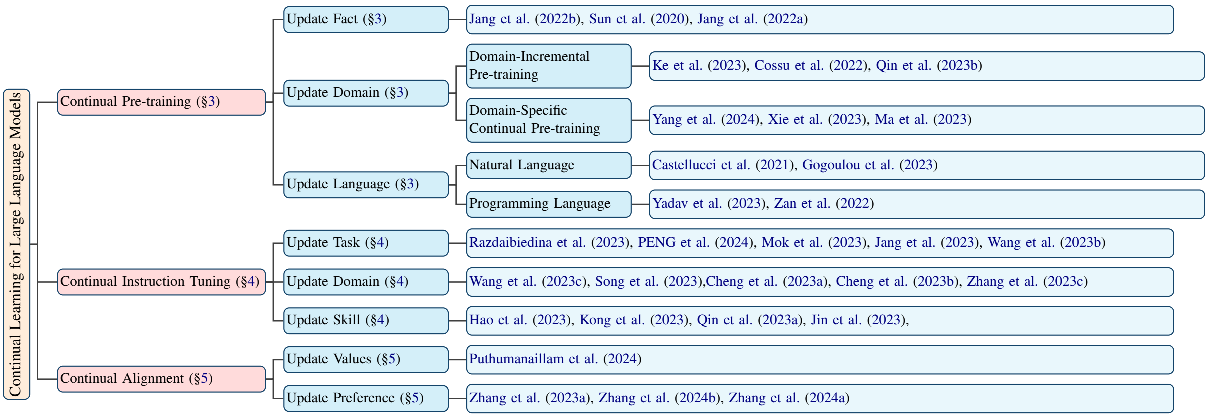

Figure 3: Taxonomy of trends in continual learning for large language models.

new developments, adapt to evolving user needs, and remain effective across diverse applications. Continual pre-training ensures LLMs are not just knowledgeable but also adaptable and responsive to the changing world.

CPT for Updating Facts The capability of LLMs to integrate and adapt to recent information is crucial. A pivotal strategy here is the employment of dynamic datasets that facilitate the real-time assimilation of data from a variety of sources like news feeds (Sun et al., 2020), scholarly articles (Cossu et al., 2022), and social media (Cossu et al., 2022). Sun et al. (2020) presents ERNIE 2.0, which is a continual pre-training framework that incrementally builds and learns from multiple tasks to maximize knowledge extraction from training data. Jang et al. (2022b) introduces continual knowledge learning, a method for updating temporal knowledge in LLMs, reducing forgetting while acquiring new information. Jang et al. (2022a) shows that continual learning with different data achieves comparable or better perplexity in language models than training on the entire snapshot, confirming that factual knowledge in LMs can be updated efficiently with minimal training data. CEM (Zhao et al., 2024a) continually evaluates LLMs to identify knowledge deficiencies based on their mistakes, collecting relevant data from multiple sources to supplement training in a targeted manner. Integral to knowledge updating is the implementation of automated systems for the verification of newly acquired data, ensuring both the accuracy and dependability of the information.

CPT for Updating Domains Continual pre-training updates domain knowledge through two approaches: 1) domain-incremental pre-training accumulates knowledge across multiple domains, and 2) domain-specific continual learning, which evolves a general model into a domain expert by training on domain-specific datasets and tasks. In domain-incremental pre-training, (Cossu et al., 2022) explores how models can be continually pre-trained on new data streams for both language and vision, preparing them for various downstream tasks. Qin et al. (2023b) examines continual re-training by assessing model compatibility and benefits of recyclable tuning via parameter initialization and knowledge distillation. Ke et al. (2023) introduces a soft-masking mechanism to update language models (LMs) with domain corpora, aiming to boost performance while preserving general knowledge. For domain-specific continual learning, Xie et al. (2023) develops FinPythia-6.9B through domain-adaptive pre-training for the financial sector. EcomGPT-CT (Ma et al., 2023) investigates the effects of continual pre-training in the E-commerce domain. These studies collectively highlight the evolving landscape of continual pre-training, demonstrating its effectiveness in enhancing model adaptability and expertise across a wide range of domains.

CPT for Language Expansion Expanding the range of languages that LLMs can understand and process is essential for ensuring broader accessibility (Castellucci et al., 2021). This expansion is not just about including a wider variety of languages, particularly underrepresented ones, but also about embedding cultural contexts into language processing. A significant challenge here is the model's ability to recognize and interpret regional dialects and contemporary slangs (Gogoulou et al., 2023), which is crucial for effective communication across diverse racial, social and cultural groups.

In addition to mastering natural languages, LLMs have also made significant strides in understanding and generating programming languages. Yadav et al. (2023) introduced CodeTask-CL, a benchmark for continual code learning that encompasses a diverse array of tasks, featuring various input and output formats across different programming languages. Zan et al. (2022) explore using an unlabeled code corpus for training models on library-oriented code generation, addressing the challenge of scarce text-code pairs due to extensive library reuse by programmers. They introduce CERT, a method where a "sketcher" outlines a code structure, and a "generator" completes it, both continuously pre-trained on unlabeled data to capture common patterns in library-focused code snippets. Yildiz et al. (2024) comprehensively examines the impact of model size on learning efficacy and forgetting, as well as how the progression and similarity of emerging domains affect the knowledge transfer within these models. These developments highlight LLMs' potential to transform both natural and programming language processing, leading to more efficient coding practices.

4 Continual Instruction Tuning (CIT)

LLMs have shown great instruction following abilities that can be used to complete different tasks with a few-shot task prompt. Continual Instruction Tuning (CIT) involves continually finetuning the LLMs to learn how to follow instructions and transfer knowledge for future tasks (Zhang et al., 2023d). Based on the ability and knowledge updated during instruction tuning, we can further divide CIT into three categories: *1) task-incremental CIT*, *2) domain-incremental CIT*, and *tool-incremental CIT*.

Task-incremental CIT Task-incremental Continual Instruction Tuning (Task-incremental CIT) aims to continuously finetune LLMs on a sequence of task-specific instructions and acquire the ability to solve novel tasks (Wang et al., 2024). A straightforward solution is to continuously generate instruction-tuning data for new tasks and directly finetune LLMs on it (Wang et al., 2023c). However, studies have shown that continuously finetuning LLMs on task-specific data would cause a catastrophic forgetting of the learned knowledge and problem-solving skills in previous tasks (Kotha et al., 2023). TAPT (Gururangan et al., 2020) presents a simple data selection strategy that retrieves unlabeled text from the in-domain corpus, aligning it with the task distribution. This retrieved text is then utilized to finetune LLMs, preventing catastrophic forgetting and enhancing argument performance. To mitigate catastrophic forgetting, Contunual-T0 (Scialom et al., 2022) employs rehearsal with a memory buffer (Shin et al., 2017) to store previous tasks data and replay them during training. ConTinTin (Yin et al., 2022) presents InstructionSpeak, which includes two strategies that make full use of task instructions to improve forward-transfer and backward-transfer. The first strategy involves learning from negative outputs, while the second strategy focuses on revisiting instructions from previous tasks. RationaleCL (Xiong et al., 2023) conducts contrastive rationale replay to alleviate catastrophic forgetting. DynaInst (Mok et al., 2023) proposes a hybrid approach incorporating a Dynamic Instruction Replay and a local minima-inducing regularizer. These two components enhance the generalizability of LLMs and decrease memory and computation usage in the replay module. Unlike previous replay-based or regularization-based methods, SLM (PENG et al., 2024) incorporates vector space retrieval into the language model, which aids in achieving scalable knowledge expansion and management. This enables LLMs' quick adaptation to novel tasks without compromising performance caused by catastrophic forgetting.

LLMs with billions of parameters introduce a huge computational burden for conducting continual learning. To address this issue, the Progressive Prompts technique (Razdaibiedina et al., 2023) freezes the majority of parameters and only learns a fixed number of tokens (prompts) for each new task. Progressive Prompts significantly reduce the computational cost while alleviating catastrophic forgetting and improving the transfer of knowledge to future tasks. ELM (Jang et al., 2023) first trains a small expert adapter on top of the LLM for each task. Then, it employs a retrieval-based approach to choose the most pertinent expert LLM for every new task. Based on the parameter-efficient tuning (PET) framework, O-LoRA (Wang et al., 2023b) proposes an orthogonal low-rank adaptation for CIT. O-LoRA incrementally learns new tasks in an orthogonal subspace while fixing the LoRA parameters learned from past tasks to minimize catastrophic forgetting. Similarly, DAPT (Zhao et al., 2024b) proposes a novel Dual Attention Framework to align the learning and selection of LoRA parameters via the Dual Attentive Learning&Selection module. LLaMA PRO (Wu et al., 2024a) proposes a novel block expansion technique, which enables the injection of new knowledge into LLMs and preserves the initial capabilities with efficient post-training.

Domain-incremental CIT Domain-incremental Continual Instruction Tuning (Domain-incremental CIT) aims to continually finetune LLMs on a sequence of domain-specific instructions and acquire the knowledge to solve tasks in novel domains. TAPT (Gururangan et al., 2020) adaptively tunes the LLMs on a series of domain-specific data including biomedicine, computer science, news, and shopping reviews. ConPET (Song et al., 2023) applies previous continual learning methods, initially developed for smaller models, to LLMs using PET and a dynamic replay strategy. This approach significantly reduces tuning costs and mitigates overfitting and forgetting problems. AdaptLLM (Cheng et al., 2023a) adapts LLMs to different domains by enriching the raw training corpus into a series of reading comprehension tasks relevant to its content. PlugLM (Cheng et al., 2023b) uses a differentiable plug-in memory (DPM) to explicitly store the domain knowledge. PlugLM could be easily adapted to different domains by plugging in in-domain memory. Zhang et al. (2023c) designs an adapt-retrieve-revise process that adapts LLMs to new domains. It first uses the initial LLMs' respose to retrieve knowledge from the domain database. Dong et al. (2023) analyze the LLMs continuously tuned on different domains and find that the sequence of training data has a significant impact on the performance of LLMs. They also offer a Mixed Finetuning (DMT) strategy to learn multiple abilities in different domains.

Tool-incremental CIT Tool-incremental Continual Instruction Tuning (Tool-incremental CIT) aims to finetune LLMs continuously, enabling them to interact with the real world and enhance their abilities by integrating with tools, such as calculators, search engines, and databases (Qin et al., 2023a). With the rapid emergence of new tools like advanced software libraries, novel APIs, or domain-specific utilities (Liang et al., 2023; Jin et al., 2023), there is a growing need to continually update LLMs so they can quickly adapt and master these new tools. Llemma (Azerbayev et al., 2023) continues tuning LLMs on a dataset with a mixture of math-related text and code to enable LLMs to solve mathematical problems by using external tools. ToolkenGPT (Hao et al., 2023) represents each tool as a new token (toolken) whose embedding is learned during instruction tuning. This approach offers an efficient way for LLMs to master tools and swiftly adapt to new tools by adding tokens.

5 Continual Alignment (CA)

LLMs need to adapt to evolving societal values, social norms and ethical guidelines. Furthermore, there exists substantial diversity in preferences across different demographic groups, as well as individuals' changing preferences over time. The need to respond to these changes give rise to continual alignment. In the context of continual alignment, two scenarios emerge: (i) the requirement to update LLMs to reflect shifts in societal values and (ii) integrating new demographic groups or value types to existing LLMs, which we will describe in the following subsections.

Continual Value Alignment Continual value alignment aims to continually incorporate ethical guidelines or adapt to cultural sensitivities and norms. As the preliminary study, Puthumanaillam et al. (2024) examines the challenges of embedding the evolving spectrum of human values into LLMs, highlights the discrepancies between static models and the dynamic nature of human societies, explores potential strategies to address these alignment issues, and suggests a path forward towards more adaptable and responsive AI systems. Although research on continual human value alignment is currently limited, it is essential to be proactive. As model capabilities improve through continual learning, ongoing alignment is necessary to ensure safety.

Continual Preference Alignment Adding new demographic groups or value types aligns with continual learning problems, aiming to guide LLMs in generating responses aligned with emerging values while adhering to previously learned ones. Previous works have demonstrated proof-of-concept of such agents. However, there is a lack of standardized benchmarks to systematically evaluate the learning capabilities of new preferences over time. CPPO (Zhang et al., 2024b) utilizes a sample-wise weighting on the Proximal Policy Optimization (PPO) algorithm (Schulman et al., 2017) to balance policy learning and knowledge retention in imitating the old policy output. On the other hand, COPF(Zhang et al., 2023a) extend the Direct Preference Optimization (DPO) algorithm (Rafailov et al., 2023) to the continual learning setting by employing Monte Carlo estimation to derive a sequence of optimal policies for the given sequences of tasks and incorporate them to regularize the policy learning on new tasks. Furthermore, COPR (Zhang et al., 2024a) adopts the Lagrangian Duality (LD) method to dynamically regularize the current policy based on the historically optimal policy to prevent forgetting and re-balance the incremental preferences objectives. Compared to other paradigms, continual alignment, which has recently emerged in LLMs, lacks both benchmark datasets and established methods, necessitating further exploration.

6 Benchmarks

The systematic evaluation of LLMs' continual learning performance demands benchmarks with high-quality data sources and diverse content. Below we summarize notable benchmark datasets.

Benchmarks for CPT TemporalWiki (Jang et al., 2022a) serves as a lifelong benchmark, training and evaluating LMs using consecutive snapshots of Wikipedia and Wikidata, helping assess an LM's ability to retain past knowledge and acquire new knowledge over time. Additional social media datasets like Firehose (Hu et al., 2023) comprise 100 million tweets from one million users over six years. CKL (Jang et al., 2022b) focuses on web and news data, aiming to retain time-invariant world knowledge from initial pretraining while efficiently learning new knowledge through continued pre-training on different corpora. TRACE (Wang et al., 2023c) encompasses eight diverse datasets covering specialized domains, multilingual tasks, code generation, and mathematical reasoning. These datasets are harmonized into a standard format, facilitating straightforward and automated evaluation of LLMs. Due to the fast-paced nature of data, time-sensitive datasets quickly become outdated, requiring frequent updates to continual pre-training benchmarks for model evaluation.

Benchmarks for CIT The Continual Instruction Tuning Benchmark (CITB) (Zhang et al., 2023d) is based on SuperNI, encompassing over 1,600 Natural Language Processing (NLP) tasks across 76 types like language generation and classification, all in a text-to-text format. ConTinTin (Yin et al., 2022), another benchmark derived from NATURAL-INSTRUCTIONS, includes 61 tasks across six categories, such as question generation and classification. CoIN (Chen et al., 2024) comprises 10 commonly used datasets spanning 8 task categories for multi-modal evaluation, ensuring a diverse range of instructions and tasks in linguistic and visual understanding. When using these benchmarks for evaluating black-box language learning models that cannot access their training data, the selection of datasets is crucial to avoid task contamination and ensure reliable performance assessment in continual instruction tuning.

Benmarks for CA COPF (Zhang et al., 2023a) conduct experiments for continual alignment using datasets like the Stanford Human Preferences (SHP) (Ethayarajh et al., 2022) and Helpful & Harmless (HH) Datasets (Bai et al., 2022). The SHP Dataset comprises 385,000 human preferences across 18 subjects, from cooking to legal advice. The HH Dataset consists of two parts: one where crowdworkers interact with AI models for helpful responses, and another where they elicit harmful responses, selecting the more impactful response in each case. Despite the growing interest in this field, there is a notable absence of dedicated benchmarks for continual alignment, presenting an opportunity for future research and development.

7 Evaluation

Evaluation for Target Task Sequence Continual learning for large language models involves evaluating the model's performance over a task sequence. Performance can be measured by three typical continual learning metrics: (1) average performance; (2) Forward Transfer Rate (FWT), and (3) Backward Transfer Rate (BWT) (Lopez-Paz and Ranzato, 2017; Wu et al., 2022), see Appendix A.1 for the detailed formulations.

Evaluation for Cross-stage Forgetting Large language models continually trained on different stages can experience the issue of unconscious forgetting (Lin et al., 2023), which shows that continual instruction tuning can erode the LLM's general knowledge. Additionally, previous studies (Qi et al., 2023) also demonstrate that the behavior of safely aligned LLMs can be easily affected and degraded by instruction tuning. To quantify these limitations, TRACE (Wang et al., 2023c) proposes to evaluate LLMs by using three novel metrics: General Ability Delta (GAD), Instruction Following Delta (IFD), and Safety Delta (SD). See Appendix A.2 for the detailed formulations. These metrics assess how LLMs' adherence to instructions, and safety change after continual learning, focusing on retaining skills and aligning with human preferences

8 Challenges and Future Works

Computation-efficient Continual Learning In the realm of computation efficiency, the focus is on enhancing the continual pre-training process with minimized computational resources (Verwimp et al., 2023; Ibrahim et al., 2024). This involves developing innovative architectures that can handle the increasing complexity of pre-training tasks without proportional increases in computational demands (Malla et al., 2024). Efficiency in algorithms and data structures becomes crucial, especially in managing the extensive data involved in pre-training (Que et al., 2024). Additionally, energy-efficient learning models are vital for sustainable scaling of LLMs, aligning with Green AI initiatives (Verwimp et al., 2023). This area requires balancing the computational cost with the benefits in terms of model capabilities.

Social Good Continual Learning Social responsibility in continual learning encompasses ensuring privacy and data security, particularly in the context of continual instruction tuning (Gabriel, 2020). As LLMs are finetuned with more specific instructions or tasks, the handling of sensitive or personal data must be managed securely and ethically. Aligning with human values and culture is also paramount (Puthumanaillam et al., 2024). This involves incorporating ethical AI principles and cultural sensitivities to ensure that the model's outputs are aligned with societal norms and values.

Automatic Continual Learning A significant challenge lies in creating systems capable of autonomously overseeing their learning processes, seamlessly adjusting to novel tasks (instruction tuning) and user preferences (alignment) while relying solely on the inherent capabilities of LLMs, all without the need for manual intervention (Qiao et al., 2024). Automatic continual learning includes multi-agent systems capable of collaborative learning and self-planning algorithms that can autonomously adjust learning strategies based on performance feedback, leading to a significant advancement in the autonomy of LLMs.

Continual Learning with Controllable Forgetting Controllable forgetting is particularly relevant to continual pre-training. The ability to selectively retain or forget information as the model is exposed to new data streams can prevent catastrophic forgetting (Qi et al., 2023) and enhance model adaptability (Wang et al., 2023c). This challenge also extends to managing misinformation and unlearning incorrect or outdated information (Chen and Yang, 2023), ensuring the accuracy and reliability of the LLM over time.

Continual Learning with History Tracking Effective history tracking is vital for understanding the evolution of the LLM through its phases of pre-training, instruction tuning, and preference learning, similar to version management in software development (Wu et al., 2024b). Managing history in model parameters and using external memory architectures can help in tracking the influence of past learning on current model behavior and decisions (Mialon et al., 2023). This is crucial for analyzing the effectiveness of continual learning processes and making informed adjustments.

Theoretical insights on LLM in Continual Learning Numerous evaluation studies have examined the issue of cross-stage forgetting (Lin et al., 2023) and demonstrated the weak robustness of aligned LLMs (Qi et al., 2023). However, theoretical analyses of how multi-stage training impacts the performance of large language models in subsequent continual learning tasks are scarce (Wang et al., 2023a). This gap highlights the need for a deeper understanding of the changes multi-stage training introduces to LLMs' learning capabilities.

9 Conclusion

Continual learning holds the vital importance of allowing large language models to be regularly and efficiently updated to remain up-to-date with the constantly changing human knowledge, language and values. We showcase the complex, multi-stage process of continual learning in LLMs, encompassing continual pre-training, instruction tuning, and alignment, a paradigm more intricate than those used in continual learning on smaller models. As the first survey of its kind to thoroughly explore continual learning in LLMs, this paper categorizes the updates by learning stages and information types, providing a detailed understanding of how to effectively implement continual learning in LLMs. With a discussion of major challenges and future work directions, our goal is to provide a comprehensive account of recent developments in continual learning for LLMs, shedding light on the development of more advanced language models.

10 Limitations

Given the broader research on continual updates of LLMs, we only include methods that require continual parameter updates through training processes. Methods like augmented models (e.g., retrieval-augmented generation or model fusion), which also enhance models with the latest information, are excluded. Additionally, this paper primarily focuses on NLP applications, omitting the continual learning of foundation models in vision or robotics.

References

Zhangir Azerbayev, Hailey Schoelkopf, Keiran Paster, et al. 2023. Llemma: An open language model for mathematics. *CoRR*.

Yuntao Bai, Andy Jones, Kamal Ndousse, et al. 2022. Training a helpful and harmless assistant with reinforcement learning from human feedback. *CoRR*.

Magdalena Biesialska, Katarzyna Biesialska, and Marta R. Costa-jussà. 2020. Continual lifelong learning in natural language processing: A survey. In *COLING*.

Giuseppe Castellucci, Simone Filice, Danilo Croce, and Roberto Basili. 2021. Learning to solve NLP tasks in an incremental number of languages. In *ACL*.

Arslan Chaudhry, Marcus Rohrbach, Mohamed Elhoseiny, et al. 2019. On tiny episodic memories in continual learning. *arXiv:1902.10486*.

Cheng Chen, Junchen Zhu, Xu Luo, Hengtao Shen, Lianli Gao, and Jingkuan Song. 2024. Coin: A benchmark of continual instruction tuning for multimodel large language model. *CoRR*, abs/2403.08350.

Jiaao Chen and Diyi Yang. 2023. Unlearn what you want to forget: Efficient unlearning for llms. In *EMNLP*.

Xiang Chen, Jintian Zhang, Xiaohan Wang, et al. 2023a. Continual multimodal knowledge graph construction. *CoRR*.

Yongrui Chen, Xinnan Guo, Tongtong Wu, et al. 2023b. Learn from yesterday: A semi-supervised continual learning method for supervision-limited text-to-sql task streams. In *AAAI*.

Daixuan Cheng, Shaohan Huang, and Furu Wei. 2023a. Adapting large language models via reading comprehension. *arXiv:2309.09530*.

Xin Cheng, Yankai Lin, Dongyan Zhao, and Rui Yan. 2023b. Language model with plug-in knowledge memory.

Andrea Cossu, Tinne Tuytelaars, Antonio Carta, et al. 2022. Continual pre-training mitigates forgetting in language and vision. *CoRR*.

Guanting Dong, Hongyi Yuan, Keming Lu, et al. 2023. How abilities in large language models are affected by supervised fine-tuning data composition. *arXiv:2310.05492*.

Li Dong, Nan Yang, Wenhui Wang, et al. 2019. Unified language model pre-training for natural language understanding and generation. *NeurIPS*.

Kawin Ethayarajh, Yejin Choi, and Swabha Swayamdipta. 2022. Understanding dataset difficulty with V-usable information. In *ICML*, volume 162.

Iason Gabriel. 2020. Artificial Intelligence, Values, and Alignment. *Minds and Machines*.

Evangelia Gogoulou, Timothée Lesort, Magnus Boman, and Joakim Nivre. 2023. A study of continual learning under language shift. *CoRR*.

Suchin Gururangan, Ana Marasović, Swabha Swayamdipta, et al. 2020. Don't stop pretraining: Adapt language models to domains and tasks. *arXiv:2004.10964*.

Shibo Hao, Tianyang Liu, Zhen Wang, and Zhiting Hu. 2023. Toolkengpt: Augmenting frozen language models with massive tools via tool embeddings. *arXiv:2305.11554*.

Tao He, Tongtong Wu, Dongyang Zhang, Guiduo Duan, Ke Qin, and Yuan-Fang Li. 2024. Towards lifelong scene graph generation with knowledge-ware in-context prompt learning. *CoRR*.

Hexiang Hu, Ozan Sener, Fei Sha, and Vladlen Koltun. 2023. Drinking from a firehose: Continual learning with web-scale natural language. *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(5).

Adam Ibrahim, Benjamin Thérien, Kshitij Gupta, Mats L. Richter, Quentin Anthony, Timothée Lesort, Eugene Belilovsky, and Irina Rish. 2024. Simple and scalable strategies to continually pre-train large language models. *CoRR*, abs/2403.08763.

Joel Jang, Seungone Kim, Seonghyeon Ye, et al. 2023. Exploring the benefits of training expert language models over instruction tuning. *arXiv:2302.03202*.

Joel Jang, Seonghyeon Ye, Changho Lee, et al. 2022a. Temporalwiki: A lifelong benchmark for training and evaluating ever-evolving language models. In *EMNLP*.

Joel Jang, Seonghyeon Ye, Sohee Yang, et al. 2022b. Towards continual knowledge learning of language models. In *ICLR*.

Qiao Jin, Yifan Yang, Qingyu Chen, and Zhiyong Lu. 2023. Genegpt: Augmenting large language models with domain tools for improved access to biomedical information. *arXiv:2304.09667*.

Xisen Jin, Dejiao Zhang, Henghui Zhu, et al. 2022. Lifelong pretraining: Continually adapting language models to emerging corpora. In *NAACL*.

Zixuan Ke, Yijia Shao, Haowei Lin, Tatsuya Konishi, Gyuhak Kim, and Bing Liu. 2023. Continual pre-training of language models. In *ICLR*.

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, et al. 2017. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*.

Yilun Kong, Jingqing Ruan, Yihong Chen, et al. 2023. Tptu-v2: Boosting task planning and tool usage of large language model-based agents in real-world systems. *arXiv:2311.11315*.

Suhas Kotha, Jacob Mitchell Springer, and Aditi Raghunathan. 2023. Understanding catastrophic forgetting in language models via implicit inference. *arXiv:2309.10105*.

Patrick S. H. Lewis, Ethan Perez, Aleksandra Piktus, et al. 2020. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *NeurIPS*.

Yaobo Liang, Chenfei Wu, Ting Song, et al. 2023. Taskmatrix. ai: Completing tasks by connecting foundation models with millions of apis. *arXiv:2303.16434*.

Yong Lin, Lu Tan, Hangyu Lin, et al. 2023. Speciality vs generality: An empirical study on catastrophic forgetting in fine-tuning foundation models. *arXiv:2309.06256*.

David Lopez-Paz and Marc'Aurelio Ranzato. 2017. Gradient episodic memory for continual learning. In *NeurIPS*.

Shirong Ma, Shen Huang, Shulin Huang, et al. 2023. Ecomgpt-ct: Continual pre-training of e-commerce large language models with semi-structured data. *CoRR*.

Srikanth Malla, Joon Hee Choi, and Chiho Choi. 2024. COPAL: continual pruning in large language generative models. *CoRR*, abs/2405.02347.

Arun Mallya, Dillon Davis, and Svetlana Lazebnik. 2018. Piggyback: Adapting a single network to multiple tasks by learning to mask weights. In *ECCV*.

Grégoire Mialon, Roberto Dessì, Maria Lomeli, et al. 2023. Augmented language models: a survey. *CoRR*.

Jisoo Mok, Jaeyoung Do, Sungjin Lee, et al. 2023. Large-scale lifelong learning of in-context instructions and how to tackle it. In *ACL*.

Bohao PENG, Zhuotao Tian, Shu Liu, Ming-Chang Yang, and Jiaya Jia. 2024. Scalable language model with generalized continual learning. In *The Twelfth International Conference on Learning Representations*.

Gokul Puthumanaillam, Manav Vora, Pranay Thangeda, and Melkior Ornik. 2024. A moral imperative: The need for continual superalignment of large language models. *CoRR*, abs/2403.14683.

Xiangyu Qi, Yi Zeng, Tinghao Xie, et al. 2023. Fine-tuning aligned language models compromises safety, even when users do not intend to! *arXiv:2310.03693*.

Shuofei Qiao, Ningyu Zhang, Runnan Fang, Yujie Luo, Wangchunshu Zhou, Yuchen Eleanor Jiang, Chengfei Lv, and Huajun Chen. 2024. Autoact: Automatic agent learning from scratch via self-planning. *CoRR*.

Yujia Qin, Shihao Liang, Yining Ye, et al. 2023a. Toolllm: Facilitating large language models to master 16000+ real-world apis. *arXiv:2307.16789*.

Yujia Qin, Cheng Qian, Xu Han, et al. 2023b. Recyclable tuning for continual pre-training. In *Findings of ACL*.

Haoran Que, Jiaheng Liu, Ge Zhang, Chenchen Zhang, Xingwei Qu, Yinghao Ma, Feiyu Duan, Zhiqi Bai, Jiakai Wang, Yuanxing Zhang, Xu Tan, Jie Fu, Wenbo Su, Jiamang Wang, Lin Qu, and Bo Zheng. 2024. D-cpt law: Domain-specific continual pre-training scaling law for large language models. *CoRR*, abs/2406.01375.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. 2023. Direct preference optimization: Your language model is secretly a reward model. In *NeurIPS*.

Anastasia Razdaibiedina, Yuning Mao, Rui Hou, et al. 2023. Progressive prompts: Continual learning for language models. *arXiv:2301.12314*.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *CoRR*.

Thomas Scialom, Tuhin Chakrabarty, and Smaranda Muresan. 2022. Fine-tuned language models are continual learners. In *EMNLP*.

Haizhou Shi, Zihao Xu, Hengyi Wang, Weiyi Qin, Wenyuan Wang, Yibin Wang, and Hao Wang. 2024. Continual learning of large language models: A comprehensive survey. *CoRR*, abs/2404.16789.

Hanul Shin, Jung Kwon Lee, Jaehong Kim, and Jiwon Kim. 2017. Continual learning with deep generative replay. *NeurIPS*.

Chenyang Song, Xu Han, Zheni Zeng, et al. 2023. Conpet: Continual parameter-efficient tuning for large language models. *arXiv:2309.14763*.

Yu Sun, Shuohuan Wang, Yu-Kun Li, et al. 2020. ERNIE 2.0: A continual pre-training framework for language understanding. In *AAAI*.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, et al. 2023. Llama: Open and efficient foundation language models. *arXiv:2302.13971*.

Eli Verwimp, Rahaf Aljundi, Shai Ben-David, Matthias Bethge, et al. 2023. Continual learning: Applications and the road forward. *CoRR*.

Liyuan Wang, Xingxing Zhang, Hang Su, and Jun Zhu. 2023a. A comprehensive survey of continual learning: Theory, method and application. *CoRR*.

Xiao Wang, Tianze Chen, Qiming Ge, et al. 2023b. Orthogonal subspace learning for language model continual learning. *arXiv:2310.14152*.

Xiao Wang, Yuansen Zhang, Tianze Chen, et al. 2023c. Trace: A comprehensive benchmark for continual learning in large language models. *CoRR*.

Yifan Wang, Yafei Liu, Chufan Shi, Haoling Li, Chen Chen, Haonan Lu, and Yujiu Yang. 2024. Inscl: A data-efficient continual learning paradigm for fine-tuning large language models with instructions. *CoRR*, abs/2403.11435.

Yufei Wang, Wanjun Zhong, Liangyou Li, et al. 2023d. Aligning large language models with human: A survey. *arXiv:2307.12966*.

Chengyue Wu, Yukang Gan, Yixiao Ge, et al. 2024a. Llama pro: Progressive llama with block expansion. *arXiv:2401.02415*.

Tongtong Wu, Massimo Caccia, Zhuang Li, Yuan-Fang Li, Guilin Qi, and Gholamreza Haffari. 2022. Pretrained language model in continual learning: A comparative study. In *ICLR*.

Tongtong Wu, Xuekai Li, Yuan-Fang Li, Gholamreza Haffari, Guilin Qi, Yujin Zhu, and Guoqiang Xu. 2021. Curriculum-meta learning for order-robust continual relation extraction. In *AAAI*.

Tongtong Wu, Weigang Wu, Xingyu Wang, Kang Xu, Suyu Ma, Bo Jiang, Ping Yang, Zhenchang Xing, Yuan-Fang Li, and Gholamreza Haffari. 2024b. Versicode: Towards version-controllable code generation. *CoRR*, abs/2406.07411.

Yong Xie, Karan Aggarwal, and Aitzaz Ahmad. 2023. Efficient continual pre-training for building domain specific large language models. *CoRR*.

Weimin Xiong, Yifan Song, Peiyi Wang, and Sujian Li. 2023. Rationale-enhanced language models are better continual relation learners. *arXiv:2310.06547*.

Prateek Yadav, Qing Sun, Hantian Ding, et al. 2023. Exploring continual learning for code generation models. In *ACL*.

Xianjun Yang, Junfeng Gao, Wenxin Xue, and Erik Alexandersson. 2024. Pllama: An open-source large language model for plant science. *CoRR*.

Yunzhi Yao, Peng Wang, Bozhong Tian, et al. 2023. Editing large language models: Problems, methods, and opportunities. In *EMNLP*.

Çagatay Yildiz, Nishaanth Kanna Ravichandran, Prishruit Punia, Matthias Bethge, and Beyza Ermis. 2024. Investigating continual pretraining in large language models: Insights and implications. *CoRR*, abs/2402.17400.

Wenpeng Yin, Jia Li, and Caiming Xiong. 2022. Contintin: Continual learning from task instructions. *arXiv:2203.08512*.

Daoguang Zan, Bei Chen, Dejian Yang, et al. 2022. CERT: continual pre-training on sketches for library-oriented code generation. In *IJCAI*.

Han Zhang, Lin Gui, Yu Lei, Yuanzhao Zhai, Yehong Zhang, Yulan He, Hui Wang, Yue Yu, Kam-Fai Wong, Bin Liang, and Ruifeng Xu. 2024a. COPR: continual human preference learning via optimal policy regularization. *CoRR*, abs/2402.14228.

Han Zhang, Lin Gui, Yuanzhao Zhai, et al. 2023a. Copf: Continual learning human preference through optimal policy fitting. *arXiv:2310.15694*.

Han Zhang, Yu Lei, Lin Gui, Min Yang, Yulan He, Hui Wang, and Ruifeng Xu. 2024b. CPPO: Continual learning for reinforcement learning with human feedback. In *The Twelfth International Conference on Learning Representations*.

Shengyu Zhang, Linfeng Dong, Xiaoya Li, et al. 2023b. Instruction tuning for large language models: A survey. *arXiv:2308.10792*.

Yating Zhang, Yexiang Wang, Fei Cheng, et al. 2023c. Reformulating domain adaptation of large language models as adapt-retrieve-revise. *arXiv:2310.03328*.

Zihan Zhang, Meng Fang, Ling Chen, and Mohammad-Reza Namazi-Rad. 2023d. Citb: A benchmark for continual instruction tuning. *arXiv:2310.14510*.

Zihan Zhang, Meng Fang, Ling Chen, et al. 2023e. How do large language models capture the ever-changing world knowledge? A review of recent advances. In *EMNLP*.

Haokun Zhao, Haixia Han, Jie Shi, Chengyu Du, Jiaqing Liang, and Yanghua Xiao. 2024a. Large language model can continue evolving from mistakes. *CoRR*, abs/2404.08707.

Weixiang Zhao, Shilong Wang, Yulin Hu, et al. 2024b. Dapt: A dual attention framework for parameter-efficient continual learning of large language models. *arXiv:2401.08295*.

A Evaluation Metrics

A.1 Evaluation for Target Task Sequence

Continual learning for large language models involves evaluating the model's performance over a task sequence. Performance can be measured by three typical continual learning metrics: (1) average performance; (2) Forward Transfer Rate (FWT), and (3) Backward Transfer Rate (BWT) (Lopez-Paz and Ranzato, 2017; Wu et al., 2022):

(1) FWT assesses the impact of knowledge acquired from previous tasks on the initial ability to perform a new task, prior to any dedicated training for that new task.

$$FWT = \frac{1}{T-1}\sum_{i=2}^{T-1} A_{T,i} - \tilde{b}_i \quad (1)$$

(2) BWT measures catastrophic forgetting by comparing a model's performance on old tasks before and after learning new ones.

$$BWT = \frac{1}{T-1}\sum_{i=1}^{T-1} A_{T,i} - A_{i,i} \quad (2)$$

(3) Average Performance, e.g., the average accuracy assesses the ability of a model or algorithm to effectively learn from and adapt to a sequence of data streams or tasks over time.

$$Avg.\ ACC = \frac{1}{T}\sum_{i=1}^{T} A_{T,i} \quad (3)$$

where $A_{t,i}$ is the accuracy of models on the test set of ith task after model learning on the tth task and $\tilde{b}_i$ is the test accuracy for task i at random initialization.

A.2 Evaluation for Cross-stage Forgetting

Large language models continually trained on different stages can experience the issue of unconscious forgetting (Lin et al., 2023), which shows that continual instruction tuning can erode the LLM's general knowledge. Additionally, previous studies (Qi et al., 2023) also demonstrate that the behavior of safely aligned LLMs can be easily affected and degraded by instruction tuning. To quantify these limitations, TRACE (Wang et al., 2023c) proposes to evaluate LLMs by using three novel metrics: General Ability Delta (GAD), Instruction Following Delta (IFD), and Safety Delta (SD):

(1) GAD assesses the performance difference of an LLM on general tasks after training on sequential target tasks.

$$GAD = \frac{1}{T}\sum_{i=1}^{T}(R^G_{t,i} - R^G_{0,i}) \quad (4)$$

(2) IFD assesses the changes of model's instruction-following ability after training on sequential different tasks.

$$IFD = \frac{1}{T}\sum_{i=1}^{T}(R^I_{t,i} - R^I_{0,i}) \quad (5)$$

(3) SD assesses the safety variation of a model's response after sequential training.

$$SD = \frac{1}{T}\sum_{i=1}^{T}(R^S_{t,i} - R^S_{0,i}) \quad (6)$$

The baseline performance of the initial LLM on the i-th task is represented by $R_{0,i}$. After incrementally learning up to the t-th task, the score on the i-th task becomes $R_{t,i}$. And R^G, R^I, and R^S represent the performance of LLM on general tasks (assessing the information obtained from pre-training), instruction-following tasks, and alignment tasks, respectively. These measure changes in an LLM's overall capabilities, adherence to instructions, and safety after continual learning, going beyond traditional benchmarks by focusing on maintaining inherent skills and aligning with human preferences.