指令微调 Instruction Tuning

- 对自然表达为指令的任务很有效
  - NLI、QA、翻译、结构到文本
- 对直接表述为语言建模的任务则不太有效
  - 常识推理和实体同指消解
  - 任务形式和原始预训练任务类似，指令是冗余的
- 任务类型越多，性能越高
- 只对大模型有效，对小模型反而会伤害其泛化性
- 结合 few-shot prompting 或 prompt tuning 可以进一步提升性能

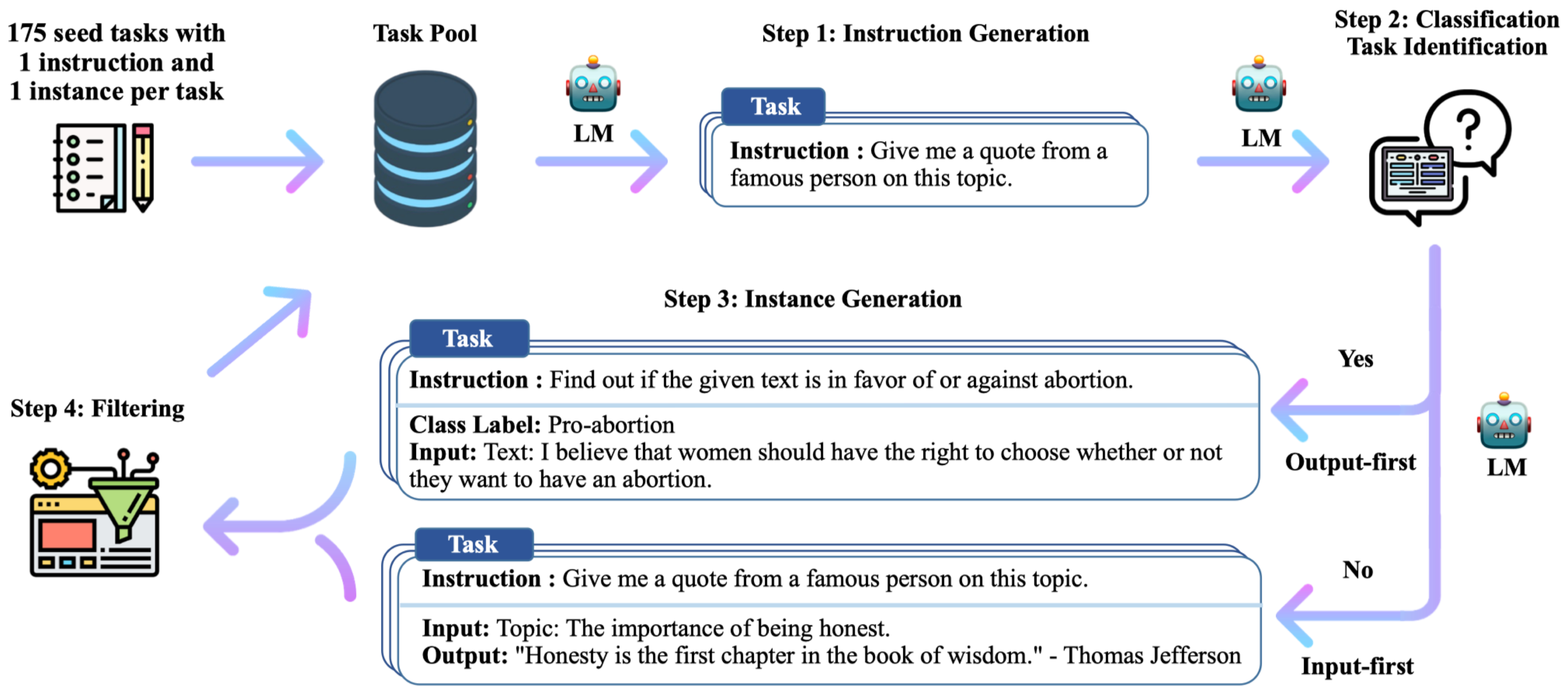
2022 ICLR. Finetuned Language Models Are Zero-Shot Learners

SELF-INSTRUCT 自动构建指令

- 人工构建指令集：数量少、多样性（创造性）差
- 借助 LLM 自身的生成能力提高模型的指令遵循性（semi-automated）
- 微调之后的 GPT3 性能取得大幅提升，接近 InstructGPT[001]
- 比使用现有的公开指令数据集效果更好
- 评估显示大部分构建出的指令是有意义的（虽然有噪声）
    - 即使错误的指令也具有正确的格式或部分正确，也可以引导模型遵循指令

2023 arXiv. SELF-INSTRUCT: Aligning Language Models with Self-Generated Instructions

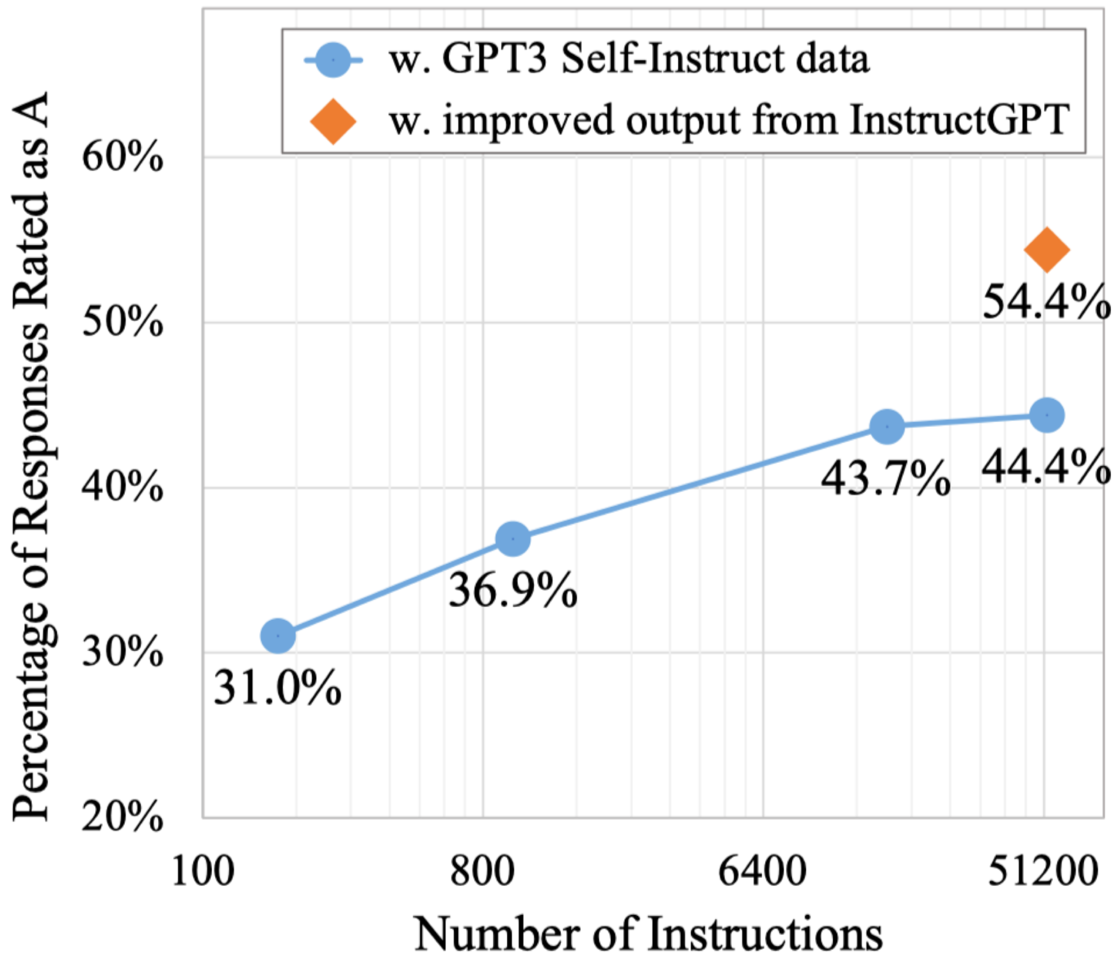
SELF-INSTRUCT 自动构建指令



2023 arXiv. SELF-INSTRUCT: Aligning Language Models with Self-Generated Instructions

SELF-INSTRUCT 自动构建指令

- 指令集合越大， 性能越好
- 指令质量越好， 性能越好



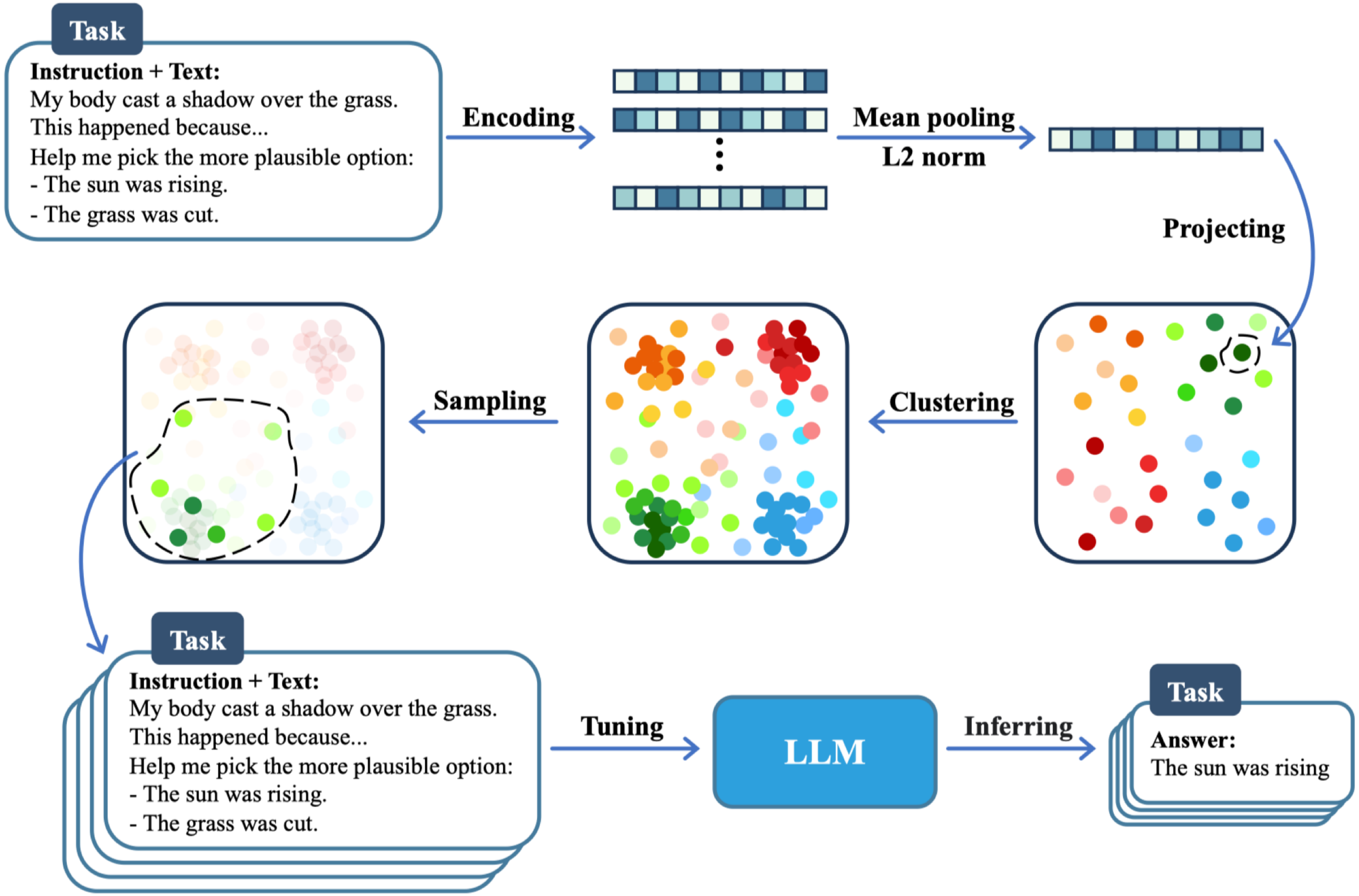
2023 arXiv. SELF-INSTRUCT: Aligning Language Models with Self-Generated Instructions

指令数据集并非越大越好

- 已有工作
  - 对于很多特定任务，微调依然是比指令微调更好的选择
  - 完成单个任务， 指令微调可以加速模型收敛
- Low Training Data Instruction Tuning
  - 减少任务和指令的多样性
- 完成特定任务只需使用<0.5%的数据训练，比全量指令训练的更好
  - 只做一个任务，建议只使用该任务的数据微调（比多任务数据微调效果更好）
  - 只做一个任务，一个指令就可能已经足够了（多指令提升空间很小，甚至效果更差）
  - 16k 样本对于训练一个完成特定 NLI 任务的模型已经足够了

2023 arXiv. Maybe Only 0.5% Data is Needed: A Preliminary Exploration of Low Training Data Instruction Tuning

指令数据集并非越大越好



2023 arXiv. Maybe Only 0.5% Data is Needed: A Preliminary Exploration of Low Training Data Instruction Tuning

小模型也具有推理能力

- 先前工作
  - 达到一定规模的模型才具备数学 CoT 推理能力
  - 小模型 CoT 能力改进都面向通用任务（UL2、FlanT5）
- 只做特定任务（牺牲通用性），小模型 (<10B) 也可以有很好的 CoT 推理能力
- CoT 或许不是一项涌现能力

2023 arXiv. Specializing Smaller Language Models towards Multi-Step Reasoning

CoT Instruction-Tuning

- 目前 CoT 指令数据集很少（多步逻辑推理）
  - 自动生成 CoT，选择能够到达正确答案的路径：GPT-J
- CoT 指令类型
  - **通用推理：**理解逻辑结构，例如翻译自然语言为形式语言、根据前提预测可能的推论
    - 增强逻辑思考能力，不依赖上下文或领域知识
  - **多项选择阅读理解：**深入理解文本，自动抽取信息，例如识别加强或削弱论点的信息

```
1: prompt_with_output{
       "instruction": instruction,
       "input": input,
       "output": output,
       }
2: prompt_without_output{
       "instruction": instruction,
       "input": input,
       }
3: output = openai.ChatCompletion.create(
       model = "gpt-4",
       messages = [{'role': 'system', 'content':
   instruction}, {'role': 'user', 'content': input},
   {'role': 'assistant', 'content': output}]
       )
```

2023 arXiv. LogiCoT: Logical Chain-of-Thought Instruction-Tuning Data Collection with GPT-4

SELF-CHECK 自动发现推理中的错误

- LLM 复杂推理能力依旧很差（中间某一个步骤出错）
- 依次检查每一个推理步骤是否正确（拆分）
- 使用 few-shot LLM 来修正错误的已有工作
  - 只能处理简单任务
  - “检查”不是一个 LLM 熟悉的预训练任务
- 怎么检查？
  - 目标抽取
  - 信息挑选
  - 重新生成该步骤答案
  - 结果比较：match/mismatch → correct/wrong

2023 arXiv. SelfCheck: Using LLMs to Zero-Shot Check Their Own Step-by-Step Reasoning

SELF-CHECK 自动发现推理中的错误

- 目标抽取：[问题] 和 [前序步骤]
  - The following is a part of the solution to the problem [question]: [Step 0, Step 1,..., Step i]. What specific action does the step [Step i] take?
- 信息挑选：只挑选 [问题] 和 [前序步骤] 中与当前步骤相关的信息
  - This is a math question: [question].
  - The following is information extracted from the question: Information 0: [Information 0] Information 1: [Information 1] ...
  - The following are the first a few steps in a solution to the problem: Step 0: [Step 0] Step 1: [Step 1] ...Step i-1: [Step i-1]
  - Which previous steps or information does the next step [Step i] directly follow from?

  - 用正则表达式抽取出被挑选出的信息或步骤的编号

2023 arXiv. SelfCheck: Using LLMs to Zero-Shot Check Their Own Step-by-Step Reasoning

SELF-CHECK 自动发现推理中的错误

- 重新生成该步骤答案
  - We are in the process of solving a math problem.
  - We have some information from the problem: Information 0: [Information I 0] Information 1: [Information I 1] ...
  - The following are some previous steps: Step 0: [Step S 0] Step 1: [Step S 1] ...
  - The target for the next step is: [Target].
  - Please try to achieve the target with the information from the problem or previous steps.
- 结果比较
  - The following are 2 solutions to a math problem.
  - Solution 1: [Regeneration output] Solution 2: [Step i]
  - Compare the key points from both solutions step by step and then check whether Solution 1 "supports", "contradicts" or "is not directly related to" the conclusion in Solution 2. Pay special attention to the difference in numbers.

2023 arXiv. SelfCheck: Using LLMs to Zero-Shot Check Their Own Step-by-Step Reasoning

目前已有的指令数据集

Dataset	# Tasks	# Instructions	Lan	Collection Method	Usage	Access	Human Verified?
OIG (AI, 2021)	30	43M	English	Mixed	Instruct. Tuning	Open	No
Baize (Xu et al., 2023)	3	100K+	English	Model Generated	Chat	Open	No
Camel (Guohao et al., 2023)	-	115K	English	Model Generated	Instruct. Tuning, Chat	Open	No
UltraChat (Ding et al., 2023)	-	675K	English	Model Generated	Chat	Open	No
Dolly (Databricks, 2022)	7	15,000	English	Human Annotated	Instruct. Tuning	Open	Yes
Guanaco-Dataset (JosephusCheung, 2021)	175	534,530	Multilingual	Mixed	Instruct. Tuning	Open	No
ChatLLaMA Chinese-ChatLLaMA (YDli-ai, 2021)	-	-	Multilingual	Mixed	Instruct. Tuning	Open	No
GPT-4-LLM (Peng et al., 2023)	175	165K	Multilingual	Model Generated	RLHF, Instruct. Tuning	Open	No
ShareGPT (ShareGPT, 2021)	-	-	Multilingual	Model Generated	Instruct. Tuning, Chat	Closed	Yes
SHP (Ethayarajh et al., 2023)	18	385K	English	Existing, Human Annotated	RLHF, Instruct. Tuning	Open	Yes
HH-RLHF (Bai et al., 2022; Anthropic, 2022; Ganguli et al., 2022)	-	169,550	English	Mixed	RLHF, Instruct. Tuning	Open	Yes
HC3 (Guo et al., 2023)	12	37,175	Multilingual	Mixed	Instruct. Tuning	Open	Yes
Stack-Exchange-Preferences (Lambert et al., 2023)	-	10M	English	Existing	RLHF, Instruct. Tuning	Open	Yes
InstructWild (Xue et al., 2023)	429	104K	Multilingual	Model Generated	Instruct. Tuning	Open	No

Dataset	# Tasks	# Instructions	Lan	Collection Method	Usage	Access	Verified?
PromptSource (Bach et al., 2022)	180	2,085	English	Mixed	Instruct. Tuning	Open	Yes
P3 (Sanh et al., 2021)	270	2,073	English	Mixed	Instruct. Tuning	Open	Yes
xP3 (Muennighoff et al., 2022)	83	-	Multilingual	Mixed	Instruct. Tuning	Open	No
Natural Instruct v1 (Mishra et al., 2022)	61	61	English	Existing	Instruct. Tuning	Open	Yes
Super-Natural-Instruct v2 (Wang et al., 2022b)	1,616	1,616	Multilingual	Mixed	Instruct. Tuning	Open	Yes
CrossFit (Ye et al., 2021)	160	-	English	Existing	Instruct. Tuning	Open	Yes
FLAN (Wei et al., 2021) 2021	62	620	English	Existing	Instruct. Tuning	Open	Yes
ExMix (Aribandi et al., 2021)	107	107	English	Existing	Instruct. Tuning	Open	-
UnifiedSKG (Xie et al., 2022)	21	21	English	Existing	Instruct. Tuning	Open	Yes
MetaICL (Min et al., 2021)	142	-	English	Existing	Instruct. Tuning	Open	Yes
InstructGPT (Ouyang et al., 2022)	-	112,801	English	Human Annotated	RLHF, Instruct. Tuning	Closed	Yes
FLAN Collection 2022 (Chung et al., 2022; Longpre et al., 2023)	1,836	18,360	English	Existing	Instruct. Tuning	Open	No
OPT-IML Bench (Iyer et al., 2022a)	1,667	3,128	English	Existing	Instruct. Tuning	Open	Yes
GLM-130B (Zeng et al., 2023)	74	-	Multilingual	Existing	Instruct. Tuning	Open	Yes
Self-Instruct (Wang et al., 2022a)	175	52,445	English	Model Generated	Instruct. Tuning	Open	No
Unnatural Instructions (Honovich et al., 2022)	-	240,000	English	Model Generated	Instruct. Tuning	Open	No
Alpaca (Taori et al., 2023)	175	51,942	English	Model Generated	Instruct. Tuning	Open	No

2023 arXiv. Chinese Open Instruction Generalist: A Preliminary Release