

## Scenario: Conversation on Medical Topics

**医疗问答：**对医疗问题进行解答，对案例进行分析。

**智能问诊：**多轮问诊后给出诊断结果和建议。

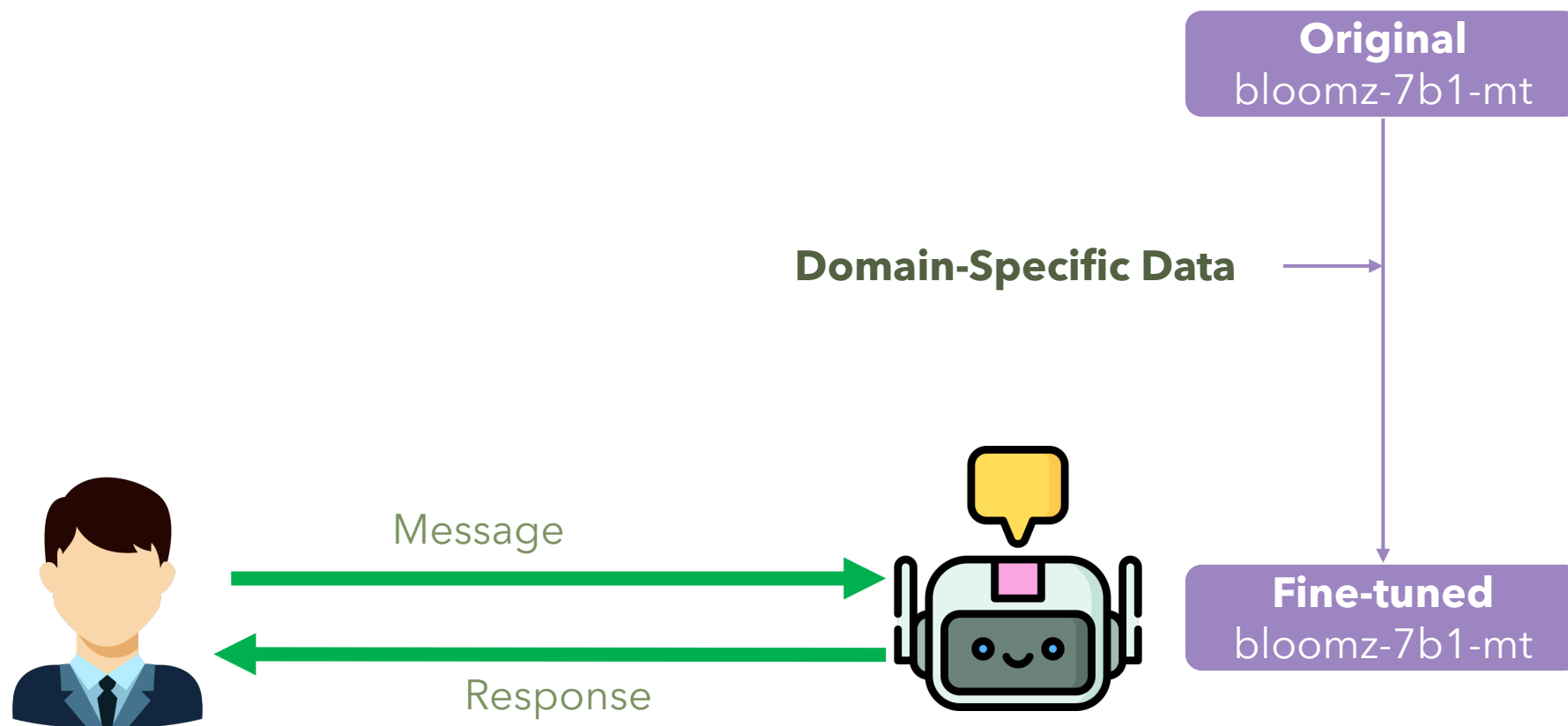
**Medical Question & Answering:** Provide answers to medical questions and analyze case studies.

**Medical Consultation:** Provide diagnostic results and recommendations after multiple rounds of consultation.



**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Scenario: Conversation on Medical Topics



**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

安全性数据 (2.69%)

**Fine-tuned**  
bloomz-7b1-mt

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

**Fine-tuned**  
bloomz-7b1-mt

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

安全性数据 (2.69%)

基于临床指南和医疗共识的知识问答 (168k)

基于结构化医疗图谱的知识问答 (160k)

基于医师资格考试题的知识问答 (77k)

真实医患问答 (140k)

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

**Fine-tuned**  
bloomz-7b1-mt

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

安全性数据 (2.69%)

基于临床指南和医疗共识的知识问答 (168k)

基于结构化医疗图谱的知识问答 (160k)

基于医师资格考试题的知识问答 (77k)

真实医患问答 (140k)

**MedicalGPT-zh**  
(generated  
by  
language model)

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

**Fine-tuned**  
bloomz-7b1-mt

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

安全性数据 (2.69%)

基于临床指南和医疗共识的知识问答 (168k)

基于结构化医疗图谱的知识问答 (160k)

基于医师资格考试题的知识问答 (77k)

真实医患问答 (140k)

**Adapted**  
from  
real data

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

**Fine-tuned**  
bloomz-7b1-mt

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

安全性数据 (2.69%)

**Like ChatDoctor**

基于HealthCareMagic构造的多轮情景问答与诊断 (200k)

基于USMLE案例分析题的格式化多轮问诊 (20k)

多轮病人信息推理与诊断 (20k)

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

**Fine-tuned**  
bloomz-7b1-mt

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

安全性数据 (2.69%)

基于HealthCareMagic构造的多轮情景问答与诊断 (200k)

基于USMLE案例分析题的格式化多轮问诊 (20k)

**Adapted** from real data

多轮病人信息推理与诊断 (20k)

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model



## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

安全性数据 (2.69%)

基于HealthCareMagic构造的多轮情景问答与诊断 (200k)

基于USMLE案例分析题的格式化多轮问诊 (20k)

多轮病人信息推理与诊断 (20k)

**Adapted** from real data  
&  
**Generated** by language model

**Fine-tuned**  
bloomz-7b1-mt

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

医疗指令 (150k)

通用指令 (150k)

安全性数据 (2.69%)

**Adapted** from real data  
&  
**Generated** by language model

**Fine-tuned**  
bloomz-7b1-mt

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

医疗指令 (150k)

通用指令 (150k)

**Borrow** some data from others

安全性数据 (2.69%)

**Fine-tuned**  
bloomz-7b1-mt

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Data:

**Original**  
bloomz-7b1-mt

**1.12M**

**Domain-Specific Data**

医疗知识问答 (48.88%)

多轮情景诊断与案例分析 (21.52%)

任务指令 (26.91%)

安全性数据 (2.69%)

敏感性问题 (15k)

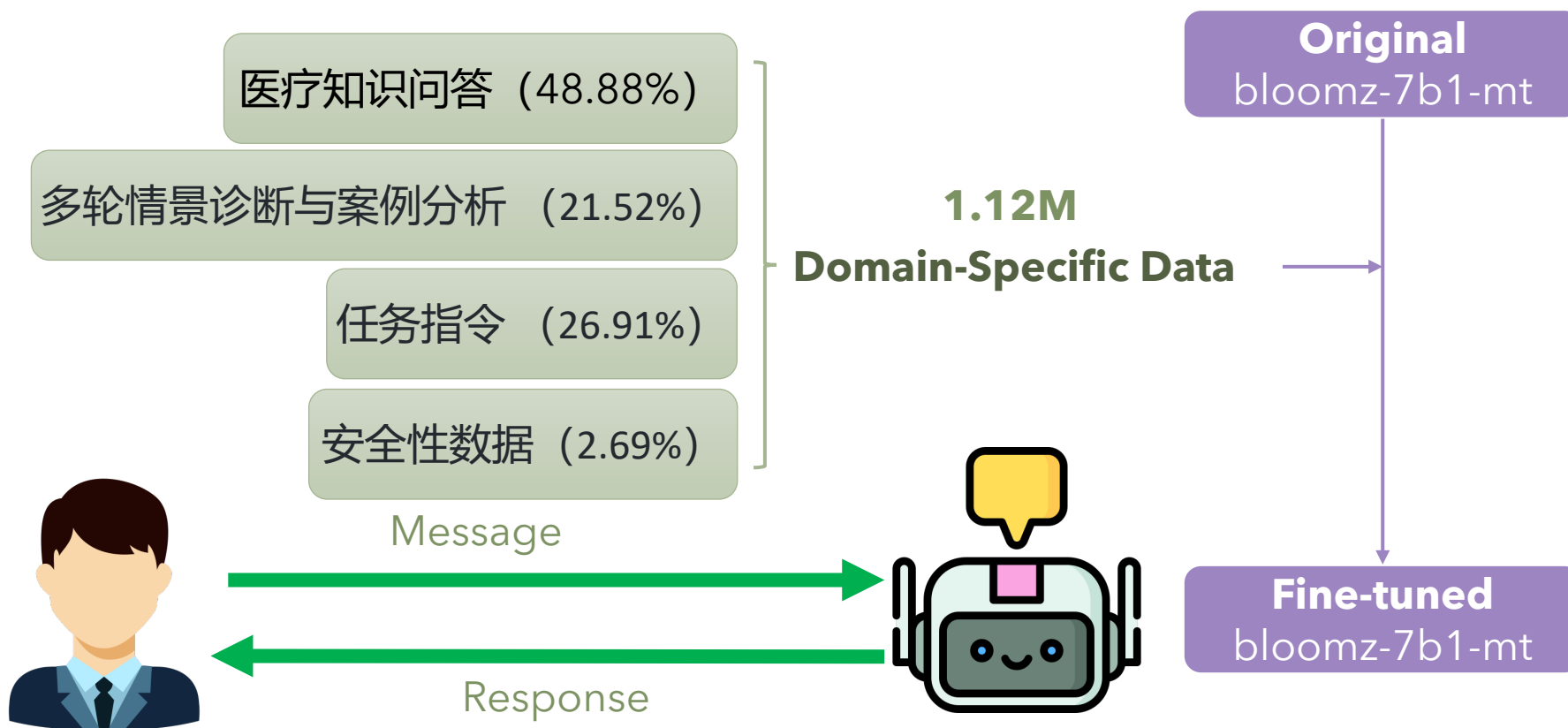
医疗反事实 (15k)

**Adapted** from real data  
&  
**Borrow** some data from others  
&  
**Generated** by language model

**Fine-tuned**  
bloomz-7b1-mt

**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model

## Scenario: Conversation on Medical Topics



**明医 (MING):** 中文医疗对话语言模型 Chinese Medical Dialogue Language Model