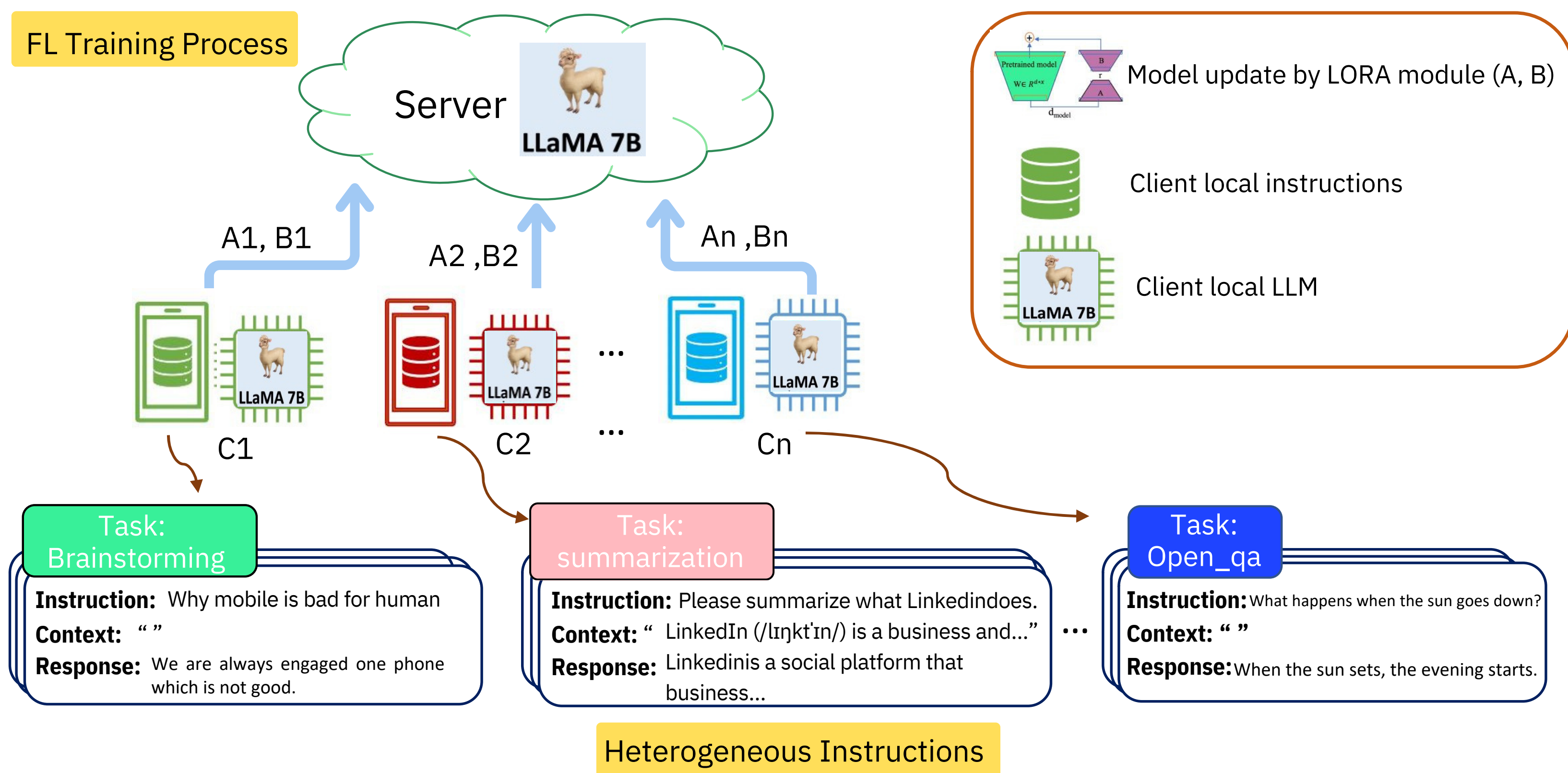


FedIT: Federated Instruction Tuning

- What if millions of users prefer not to share their instructions, as these instructions may contain private information?
- What if companies are reluctant to share instructions containing their core technology and intellectual property?
- What if a country's government regulations disallow the transfer of data outside its borders, preventing the combination of instructions from different countries?



Algorithm 1 Federated Instruction Tuning (FedIT)

Initialization: each client's initial global large language model with parameters w and a lightweight adapter with parameters $\Delta w^{(0)}$, client index subset $\mathcal{M} = \emptyset$, K communication rounds, $k = 0$,

Training

while $k \leq K$ **do**

Server updates \mathcal{M} using specific strategies ▷ **Select clients for local training**

for $n \in \mathcal{M}$ **in parallel do** ▷ **Parameter-efficient finetuning on local instructions dataset**

Client freeze the LLM and update the adapter weights with $\Delta w^{(k)}$

$\Delta w_n^{(k+1)} \leftarrow$ **InstructionTuning**($\Delta w_n^{(k)}$)

end for

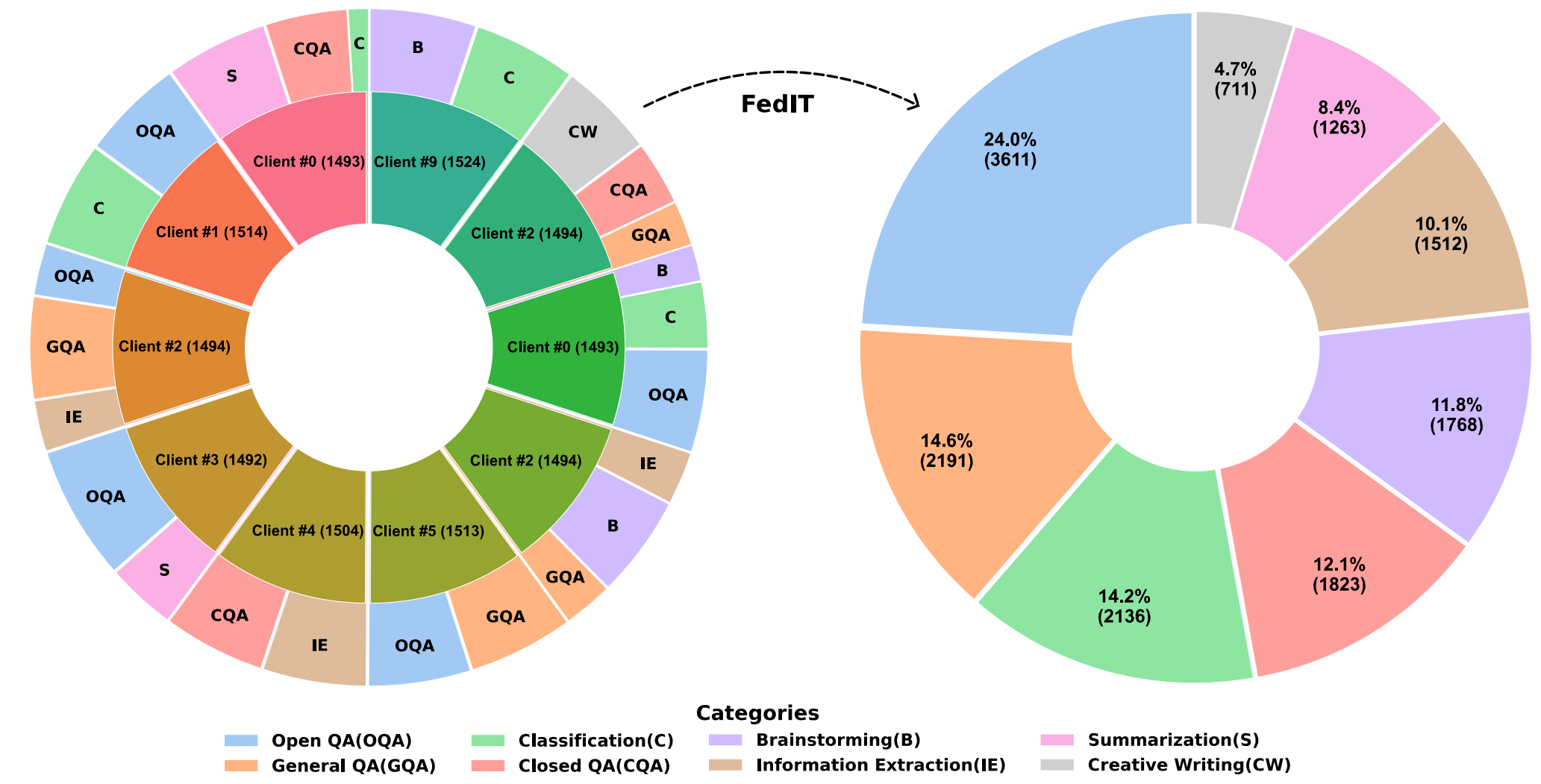
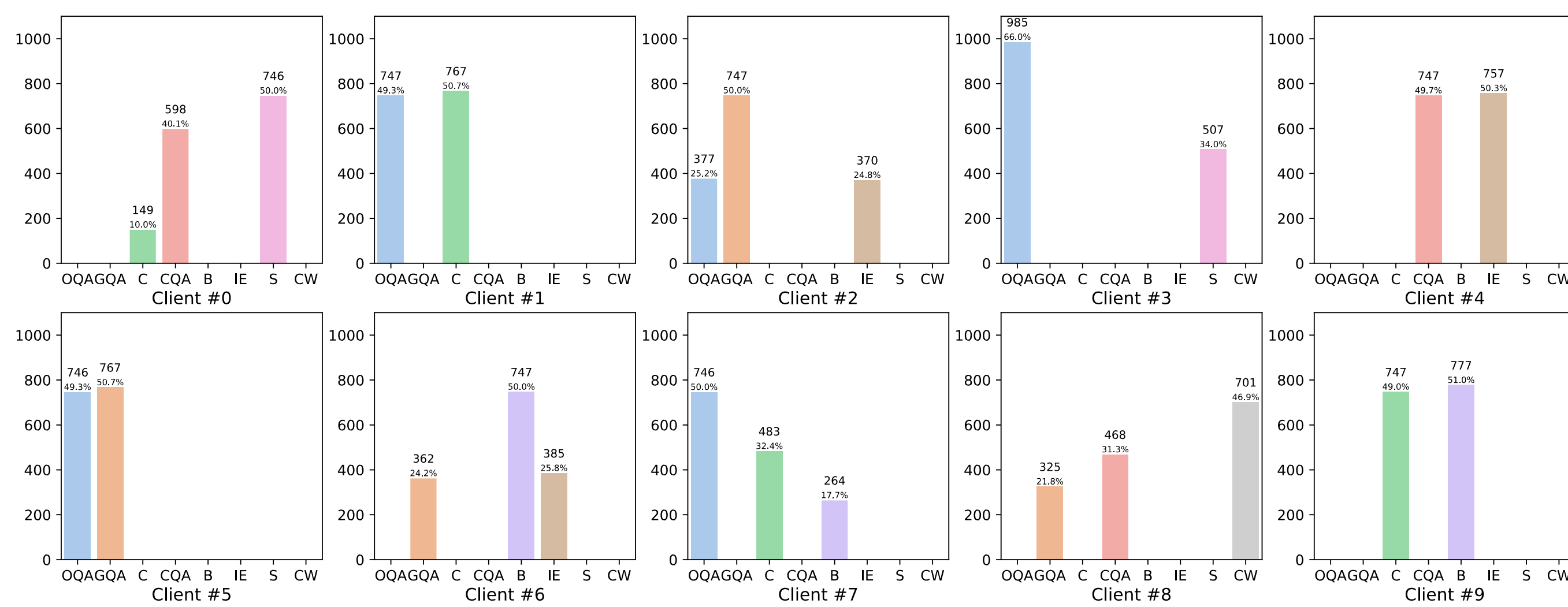
$\Delta w^{(k+1)} \leftarrow$ **Aggregate**($\Delta w_n^{(k+1)}$) for $n \in \mathcal{M}$ ▷ **Aggregate the adapters at Server**

$k \leftarrow k + 1$

end while

Outcome (m, θ_g^t):
Derive the final adapter with parameters $\Delta w^{(K)}$ and the global LLM with parameters w

Heterogeneity of Instructional Data



Shepherd: A GitHub Platform for FedIT Support



Shepherd

A Platform Supporting Federated Instruction Tuning

[Overview](#) • [Paper](#) • [Installation](#) • [Data_Preparation](#) • [Federated_Finetuning](#) • [Inference](#) • [Citation](#)

Code License: Apache 2.0

Usage and License Notices: The data, code and checkpoints are intended and licensed for research use only.

It facilitates seamless integration of novel algorithms and configurations. Shepherd is a foundational framework for exploring federated finetuning of LLMs using heterogeneous instructions across diverse categories. The framework is designed for ease of use, adaptability, and scalability to accommodate large datasets

Evaluation

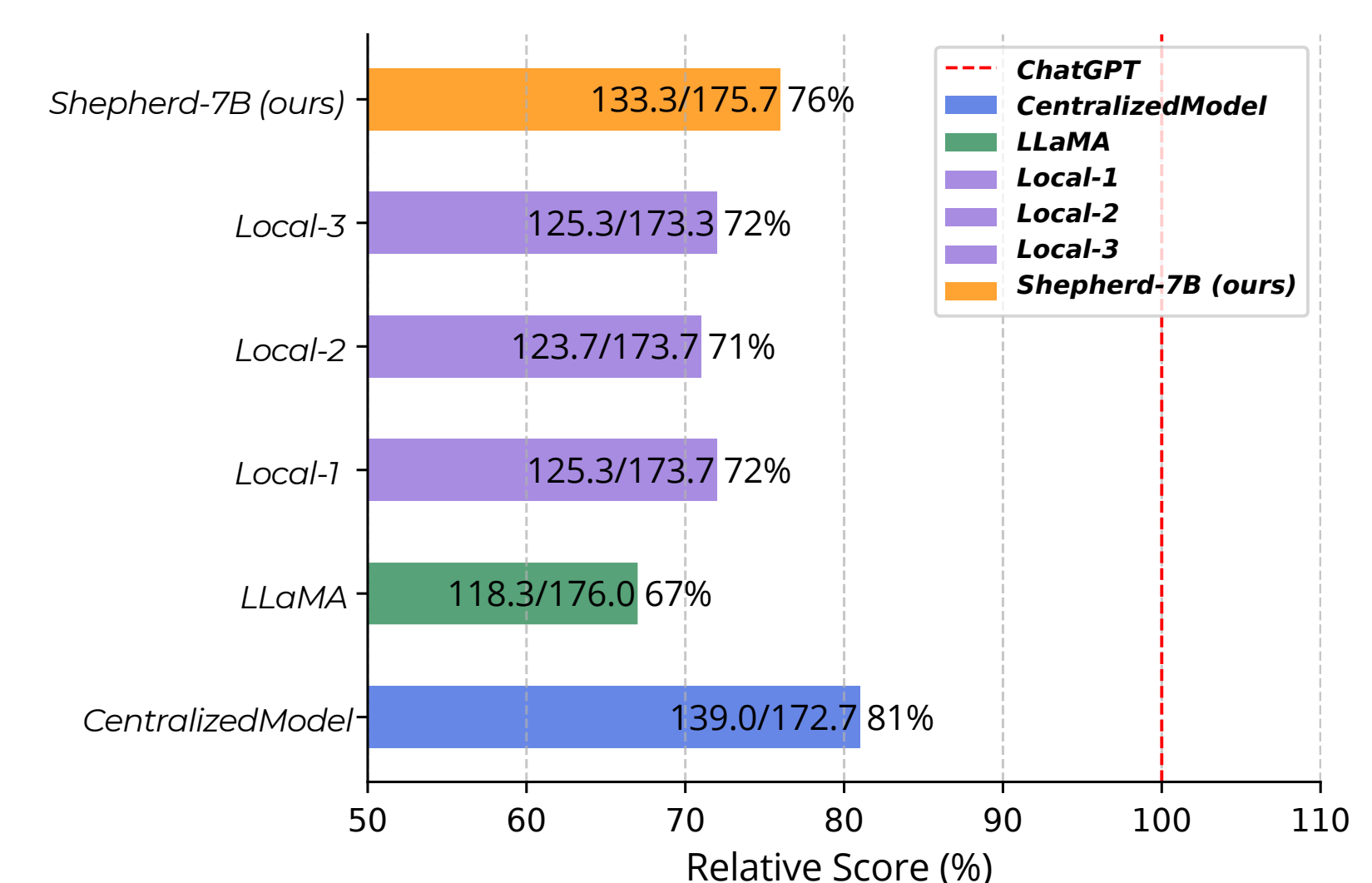


Table 5: Response demonstration for a classification task on **Databricks-dolly-15k** dataset. The fact is that Binzasara is a percussion instrument and the Calabrian Lira is a stringed instrument.

Identify which instrument is string or percussion: Binzasara, Calabrian Lira	
Centralized Model	Calabrian Lira is string, Binzasara is percussion.
LLaMA	Binzasara is a string instrument. Calabrian Lira is a percussion instrument.
Local-1	Binzasara is string, Calabrian Lira is percussion.
Local-2	Binzasara is string, Calabrian Lira is percussion.
Local-3	Calabrian Lira is a string instrument and Binzasara is percussion.
Shepherd-7B (ours)	Calabrian Lira is string, Binzasara is percussion.
ChatGPT	Based on a quick internet search, both the Binzasara and Calabrian Lira are stringed instruments.

Table 3: A summary of the baselines and their corresponding scores evaluated by GPT-4. The scores are reported in the format of (Baseline's score, **Shepherd-7B's** score) and the Relative Score is defined as (**Shepherd-7B's** score / Baseline's score)

Baseline	Task	Scores	Relative Score
Centralized Model	Centralized tuning with all the instructions	(142.2, 130.7)	0.919
LLaMA	No instruction tuning	(114.0, 131.7)	1.155
Local-1	Brainstorming instruction tuning	(120.0, 131.0)	1.092
Local-2	Closed question answering instruction tuning	(116.1, 129.0)	1.111
Local-3	Classification and brainstorming instruction tuning	(121.3, 131.8)	1.087