

Position: Current Model Licensing Practices are Dragging Us into a Quagmire of Legal Noncompliance

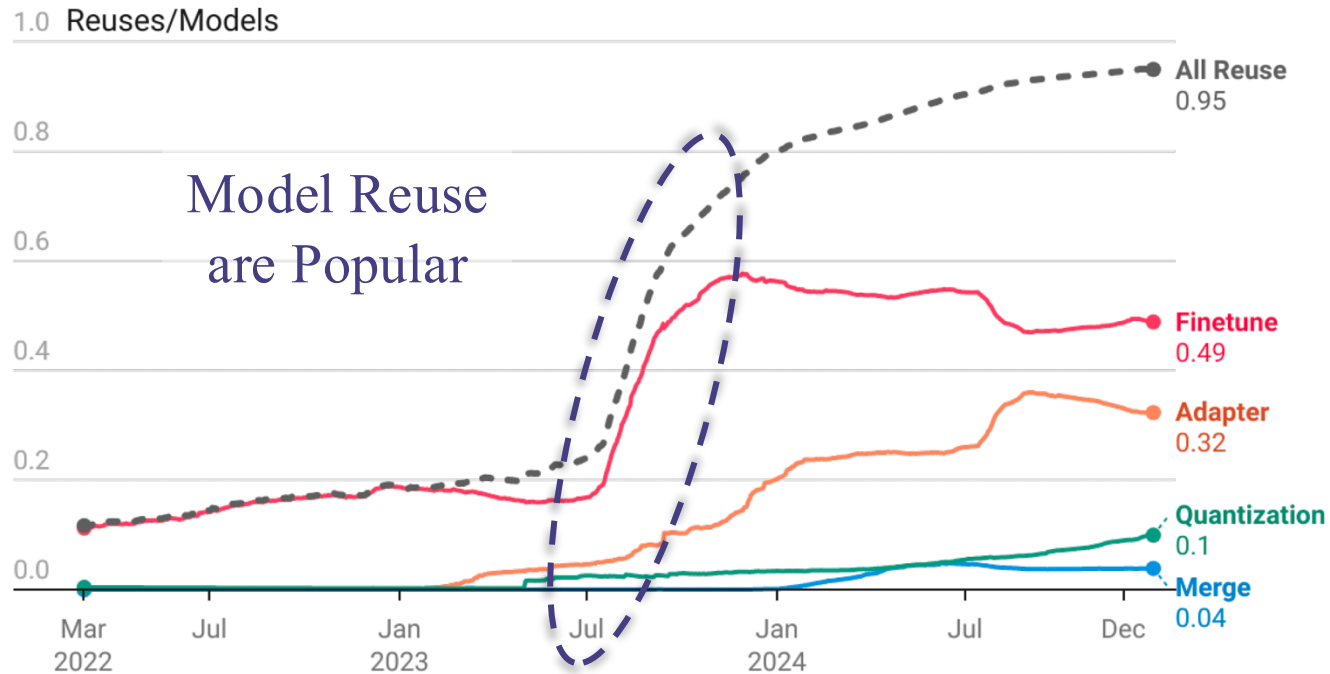
Speaker: Moming Duan

42nd International Conference on Machine Learning
July 2025



Oral

Background

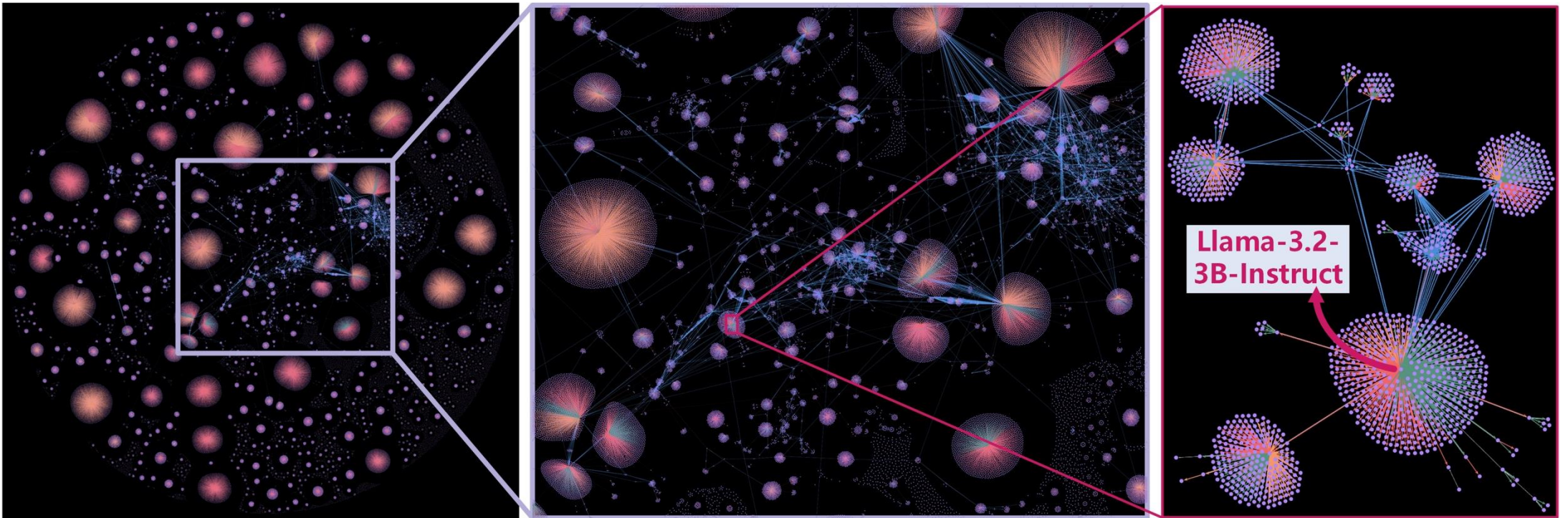


Trends in Model Reuse on Hugging Face 😊

- **Pretraining is costly**
 - 10^{25} FLOPs
 - Billions of tokens
 - Several months with thousands of GPUs
- **Model reuse is cheap**
 - PEFT: LoRA (0.1% – 1%)
 - 1 GPU can run
 - 1.4 million models on HF

Background

- The Collaboration Chain of ML Models



Visualization of Model Dependencies on 😊
Finetune (50%), *Adapter* (33%), *Quantization* (10%), and *Merge* (4%)

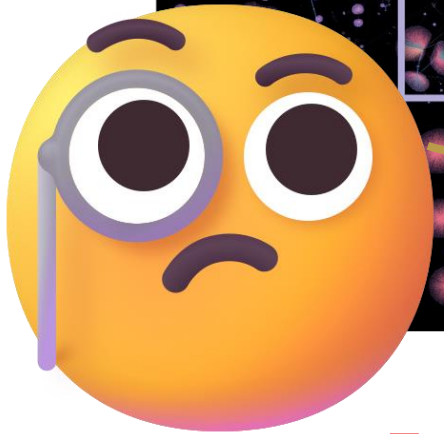
Background

- The Collaboration Chain of ML Models

Meta Llama3 Community License:

"You will **not** use the Llama Materials or any output or results of the Llama Materials to **improve** any other large language model"

Can?
Is there a
legal risk?



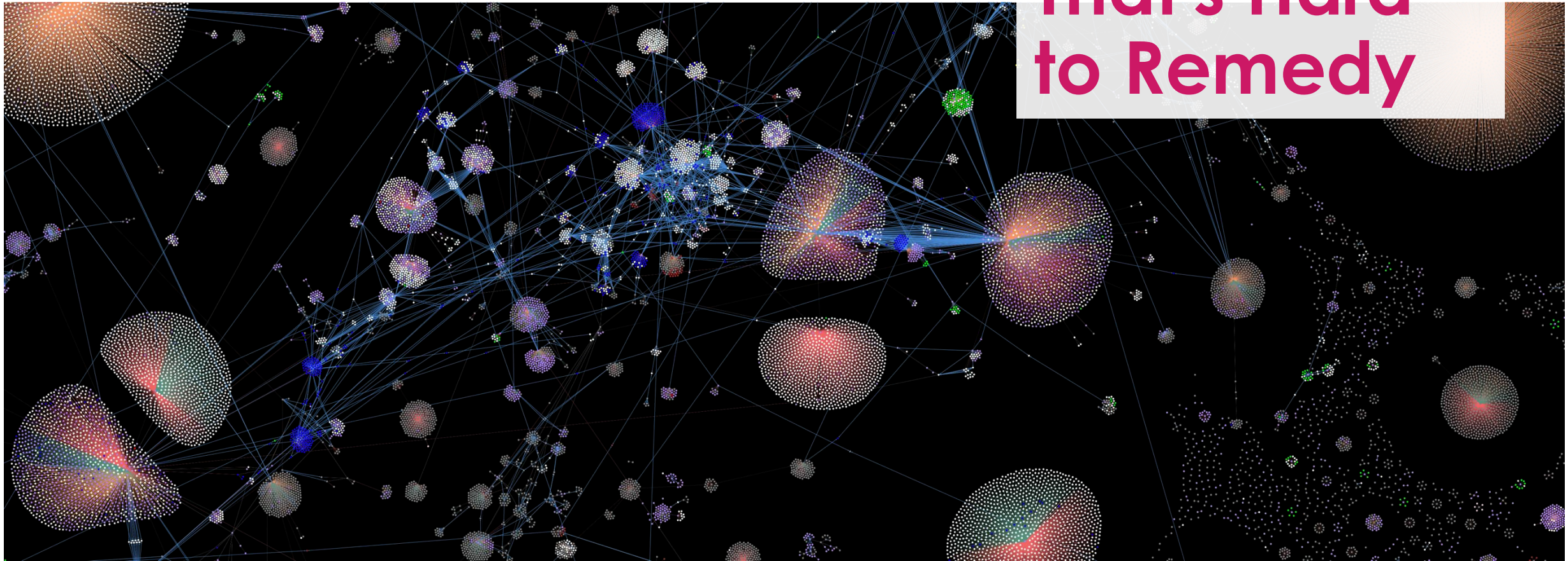
Visualization of Model Dependencies on 😊
Finetune (50%), *Adapter* (33%), *Quantization* (10%), and *Merge* (4%)

Results

Table 3. Quantitative Results of Noncompliance. NCM: NonCompliance Models, CTM: Corrected Total Models.

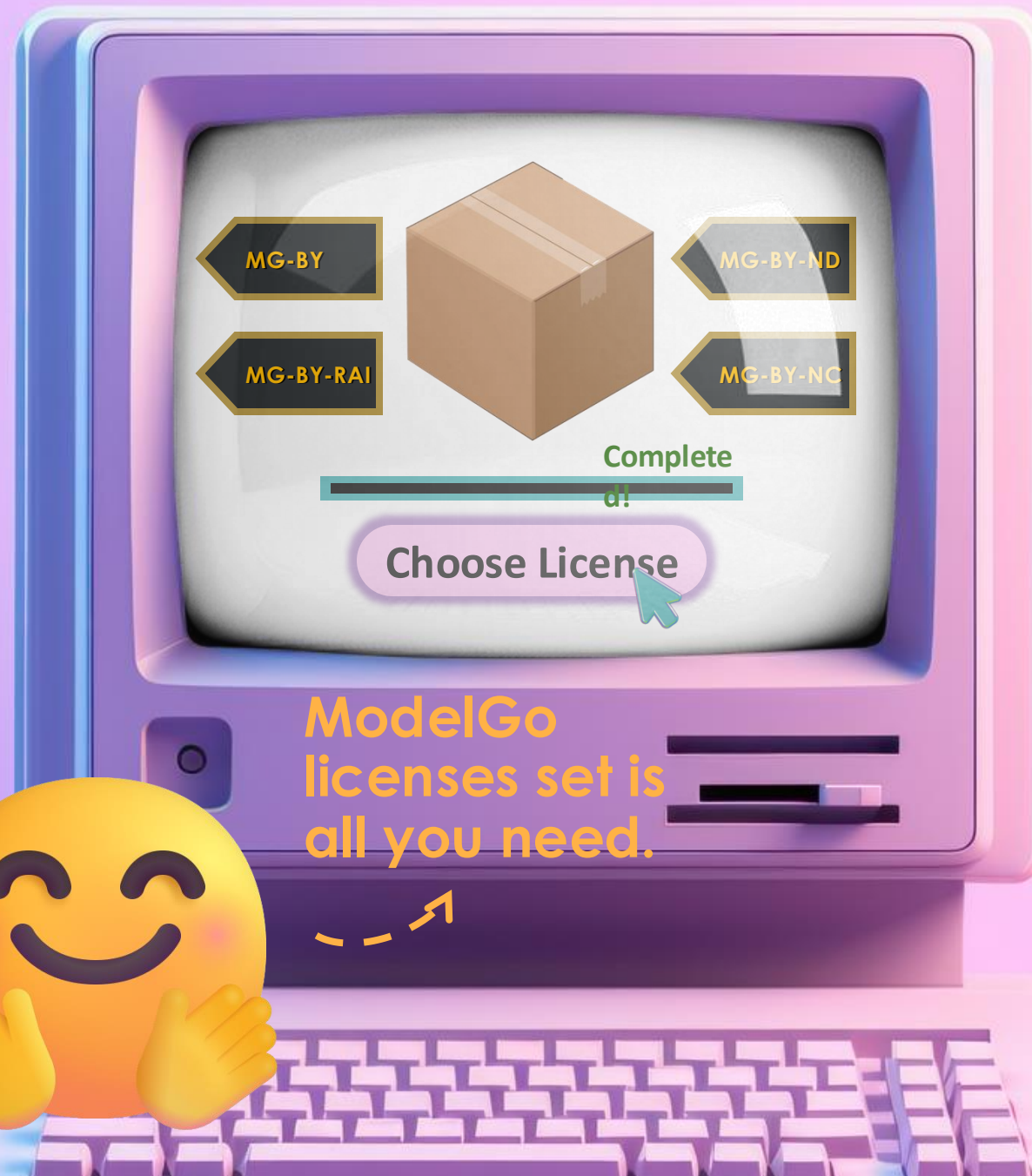
Noncompliance	# of NCM	# of CTM	Finetune	Merge	Quantization	Adapter
● License Mismatch	123,707 (81.8%)	151,302	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
○ Copyleft-style Terms	26,679 (63.5%)	41,019	11,011 (35.8%)	1,453 (4.7%)	4,775 (15.5%)	13,546 (44.0%)
● Copyleft Violation	365 (38.5%)	949	296 (81.1%)	32(8.8%)	37 (10.1%)	<i>n/a</i>
● ND	36 (100%)	36	1 (2.8%)	12 (33.3%)	23 (63.9%)	<i>n/a</i>
● CCf	76 (2.1%)	3,675	17 (20.0%)	67 (78.8%)	1 (1.2%)	<i>n/a</i>
● FSff	98 (45.8%)	214	84 (85.7%)	0	14 (14.3%)	<i>n/a</i>
● La2E-license	1,000 (38.4%)	2,602	182 (18.1%)	189 (2.9%)	331 (32.9%)	465 (46.2%)
● La2E-merge	11 (55.0%)	20	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
● La3E-license	966 (24.7%)	3,904	518 (47.2%)	18 (7.3%)	133 (27.6%)	195 (17.6%)
● La3E-merge	81 (56.6%)	143	<i>n/a</i>	81 (100%)	<i>n/a</i>	<i>n/a</i>

A Quagmire
That's Hard
to Remedy



LICENSING OPTIONS

- BY - Attribution
- NC - NonCommerical
- ND - NoDerivatives
- RAI - Responsible AI
- SA - ShareAlike
(Copyleft)



KEY FEATURES

Specifically Designed for
Model Licensing

Feasible Licensing Options
Easy to Learn & Use

Clearly Defined and Safer
Model Reuse

