IEEE-CNSV

Consultants' Network of Silicon Valley

Protecting Future AI Products: A Holistic IP Approach
May 13, 2025

Fredrick Tsang

Introduction to the Speaker

- Patent attorney specializing in holistic IP development for AI companies
- Prior hands-on technical experience in training neural network
- Leverages technical knowledge to assess how AI innovations impact legal frameworks
- Award-winning legal writer – 2025 Burton *Law360* Distinguished Legal Writing Award

Disclaimer

The views and opinions expressed in this presentation are solely my own and do not reflect the views, positions, or policies of Fenwick or its clients. My participation is in an individual capacity and should not be construed as legal advice or an official statement from the firm.

Can an AI product sustain profit margin?

- LLM Pricing
  - Early 2023: $0.02 per thousand input tokens
  - January 2025: One model company offered $0.14 per million input tokens – or $0.00014 per thousand tokens
  - More than 99% price drop in two years

A billion dollar question

- How do we build moats in AI products?

Probably no magic bullet in building moats

Pharmaceutical: Moat = Legal exclusivity (patents)

Widely adopted end-user software: Moat = number of users

Anything in between:

- Business/product differentiators

+

- Safeguarding your IP

Today's topics

Outline – IP protection for AI products

Background:

- IP Overview
- DeepSeek and Distillation Story

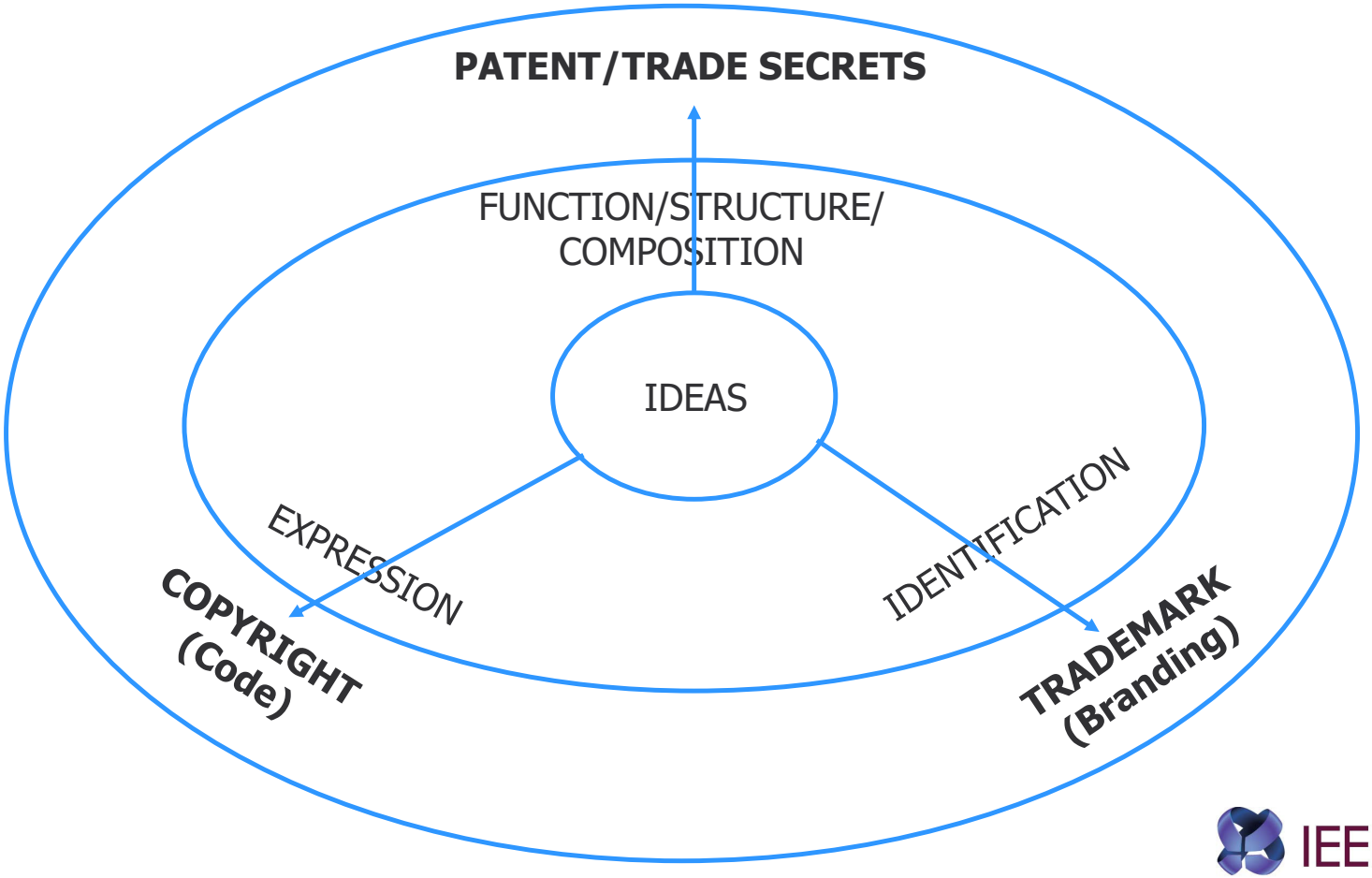
What has changed?

- Ways of copying
- Software moats
- Copyright (Lack of?)

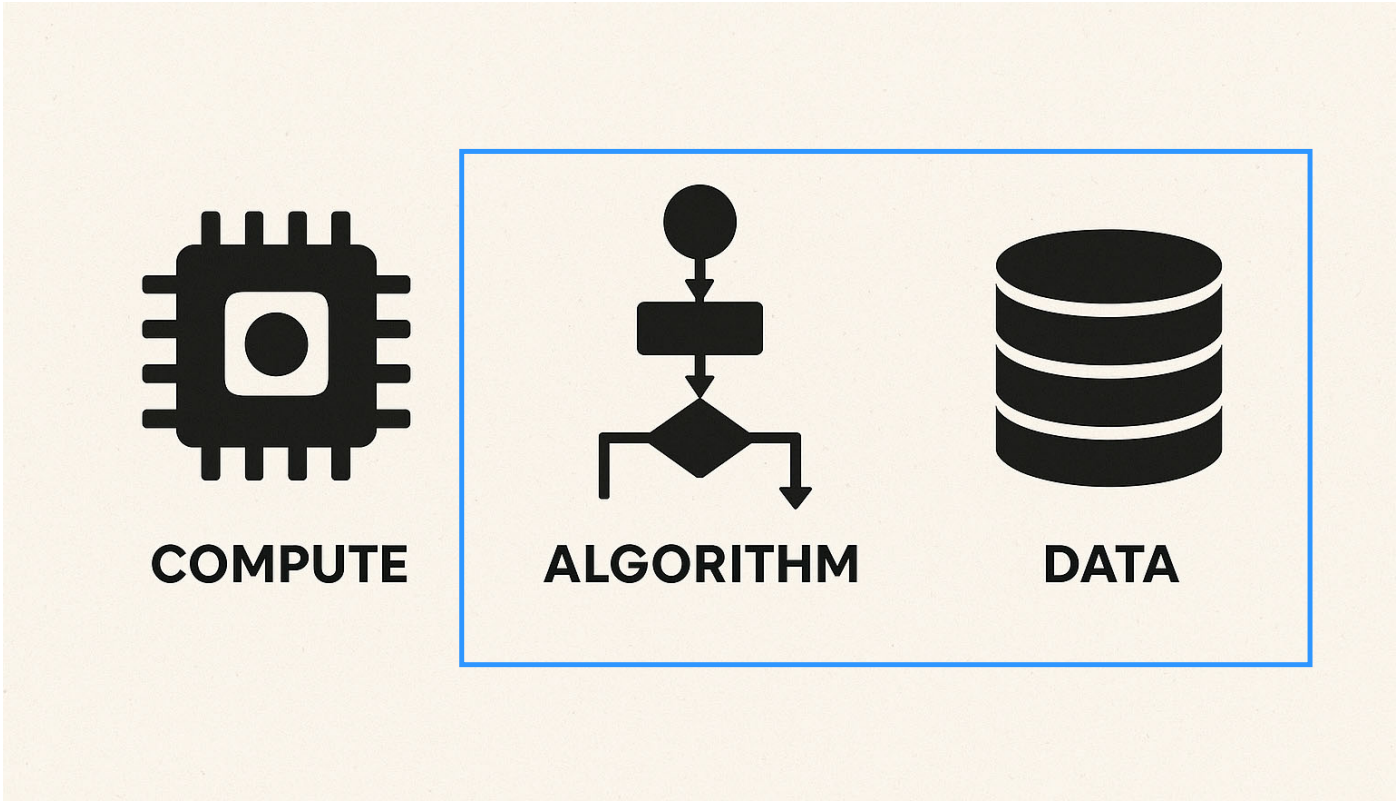
Implications on IP Protection:

- Divorce of Copyright and Contract
- Patent
- Data

Realms of Intellectual Property



The trinity of A.I.



Outline

Background:

- IP Overview
- DeepSeek and Distillation Story

What has changed?

- Ways of copying
- Software moats
- Copyright (Lack of?)

Implications on IP Protection:

- Divorce of Copyright and Contract
- Patent
- Data

DeepSeek story

nature

Explore content | About the journal | Publish with us | Subscribe

nature › news › article

NEWS | 23 January 2025

China's cheap, open AI model DeepSeek thrills scientists

DeepSeek-R1 performs reasoning tasks at the same level as OpenAI's o1 – and is open for researchers to examine.

TECH · A.I.

DeepSeek shakes up stocks as traders fear for U.S. tech leadership

BY ABHISHEK VISHNOI, WINNIE HSU AND BLOOMBERG
January 26, 2025 at 10:20 PM PST

How DeepSeek AI Sparks Nearly $1 Trillion U.S. Tech Implosion

Licensing

MATT KRANTZ | 08:00 AM ET 01/28/2025

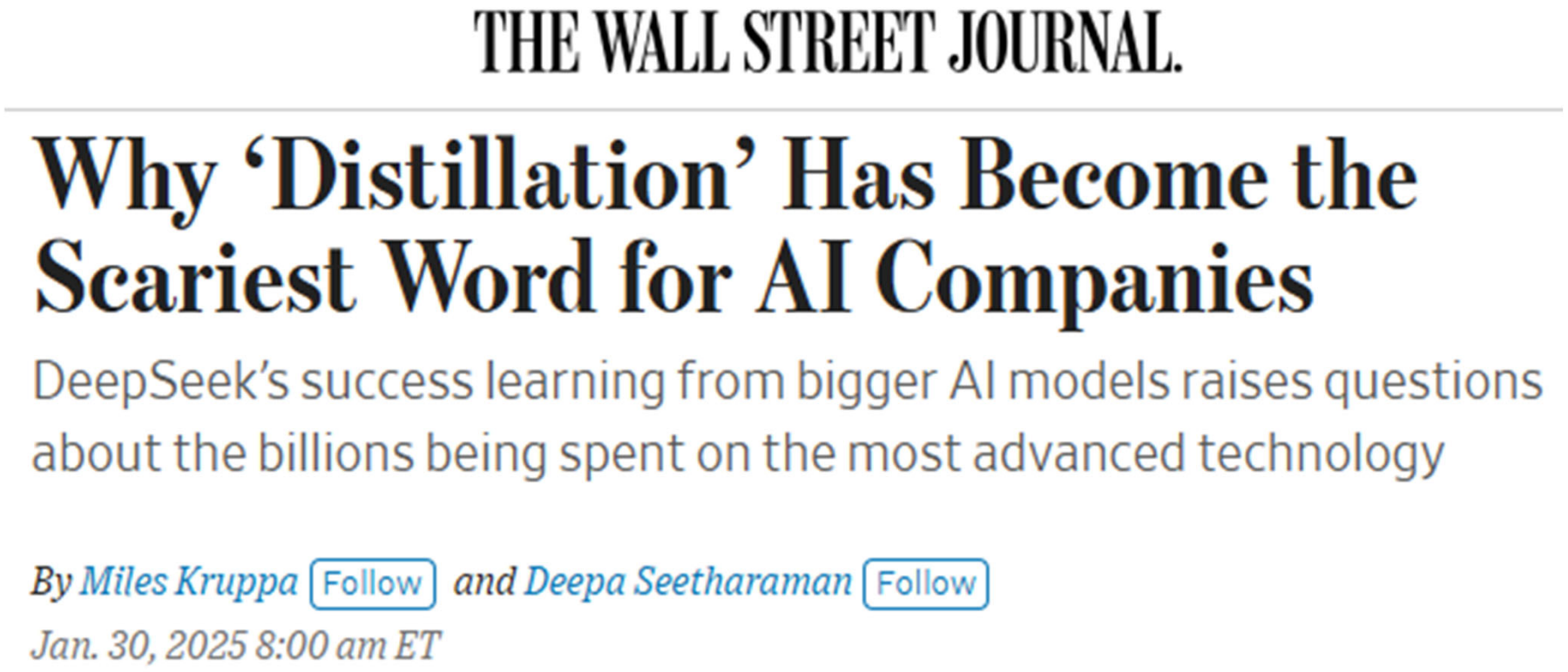
China's DeepSeek set off a $969 billion bomb of value by U.S. technology stocks in the S&P 500. And some worry the disruption is only beginning.

The logo of Chinese artificial intelligence company DeepSeek is seen in Hangzhou, Zhejiang province, China, on Jan. 26, 2025.
CFOTO/FUTURE PUBLISHING VIA GETTY IMAGES

What is AI model distillation?

- Model Distillation
  - Knowledge Distillation
  - Process of transferring knowledge from a larger, complex model (teacher) to a smaller, efficient model (student)
- Teacher Model vs. Student Model
  - Teacher: High-performance, large-scale model
  - Student: Compressed version, optimized for efficiency
  - Goal: Retain performance while reducing size and complexity
- Authorized vs. Unauthorized Distillation
  - Authorized: usually have full access to the teacher model to distill a faster version
  - Unauthorized: scrape teacher model outputs to amass a large amount of training data to train a student model

DeepSeek story



THE WALL STREET JOURNAL.

Why 'Distillation' Has Become the Scariest Word for AI Companies

DeepSeek's success learning from bigger AI models raises questions about the billions being spent on the most advanced technology

By Miles Kruppa Follow *and Deepa Seetharaman* Follow

Jan. 30, 2025 8:00 am ET

Outline

Background:

- IP Overview
- DeepSeek and Distillation Story

What has changed?

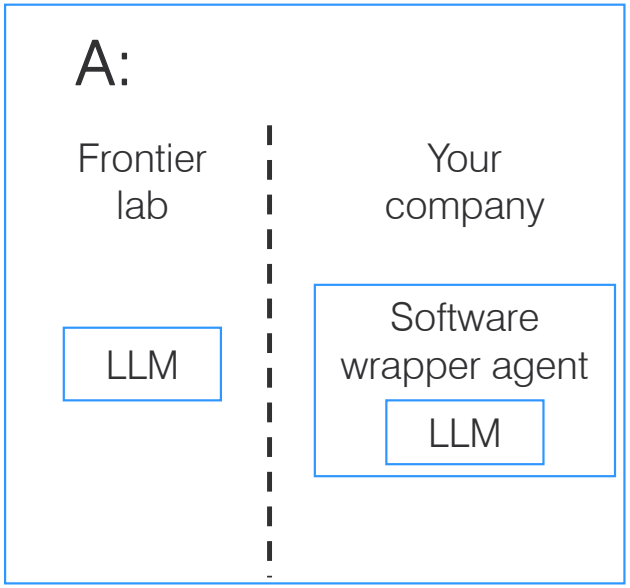
- Ways of copying
- Software moats
- Copyright (Lack of?)

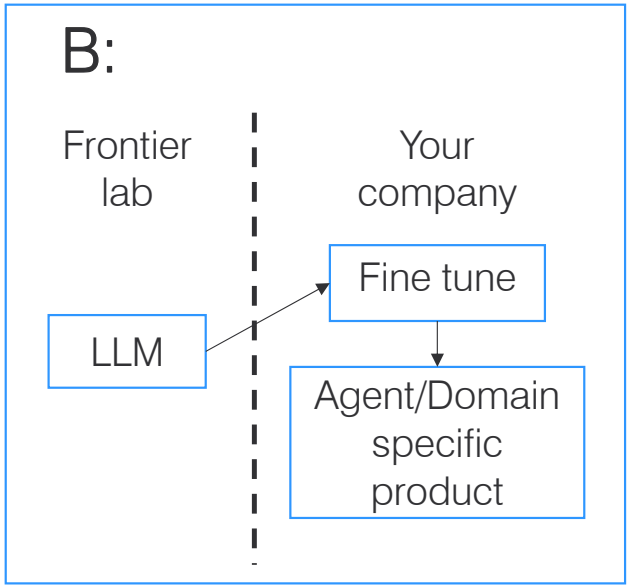
Implications on IP Protection:

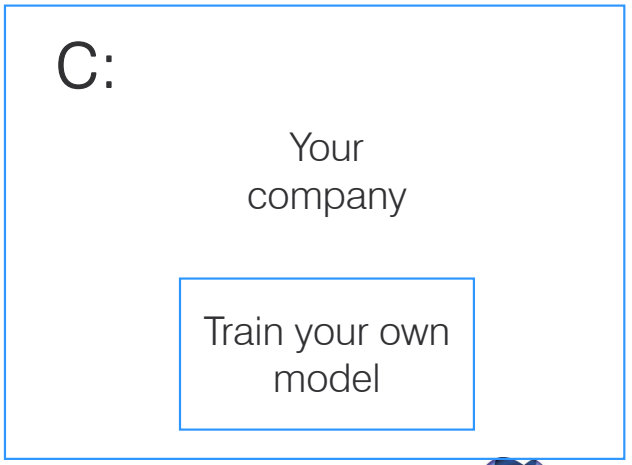
- Divorce of Copyright and Contract
- Patent
- Data

What are AI offerings?

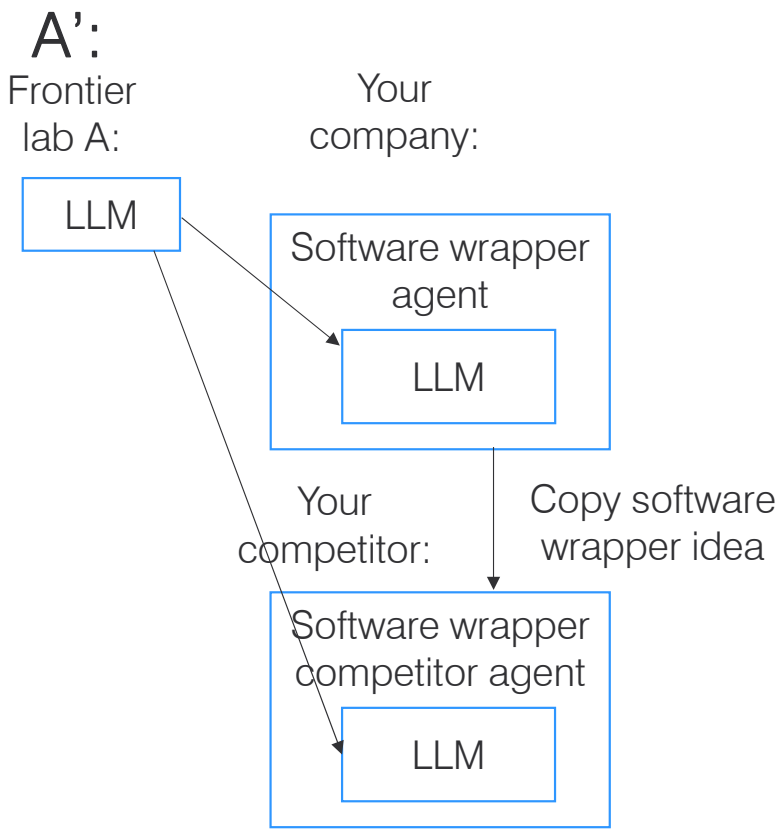
- AI agents
- Domain-specific AI models (image generating models, protein simulators, etc.)
- AI as a service







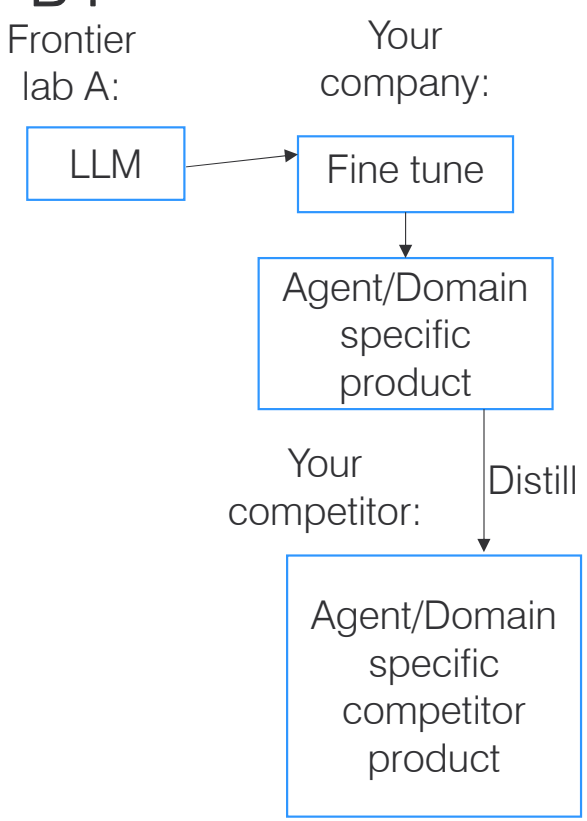
Ways of copying



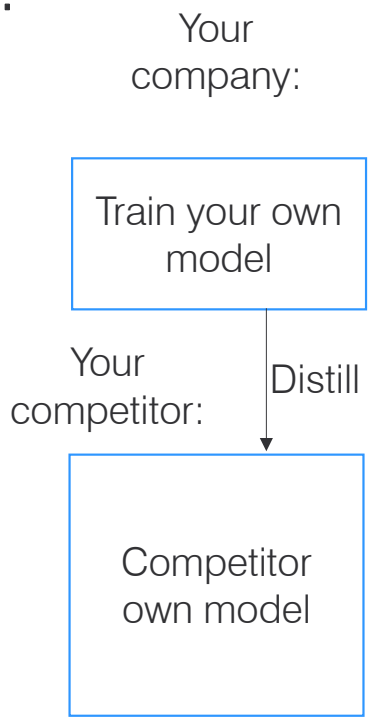
How do I differentiate my product when the most powerful component is the commoditized LLM?

Ways of copying

B':



C':



Outline

Background:

- IP Overview
- DeepSeek and Distillation Story

What has changed?

- Ways of copying
- Software moats
- Copyright (Lack of?)

Implications on IP Protection:

- Divorce of Copyright and Contract
- Patent
- Data

What are the moats in software companies?

- Network effect
- Economy of scale
- Switching cost
- Superior product
- First mover advantage

Paradigm shift in moat building for GenAI software

- Development of GenAI software is quite different
  - Non-deterministic outputs
  - Hard to conduct QA (especially in B2B setting)
  - Reliance on third-party models
  - Scaling may no longer follow traditional software wisdom
  - Model distillation
  - Legal framework shift

Outline

Background:

- IP Overview
- DeepSeek and Distillation Story

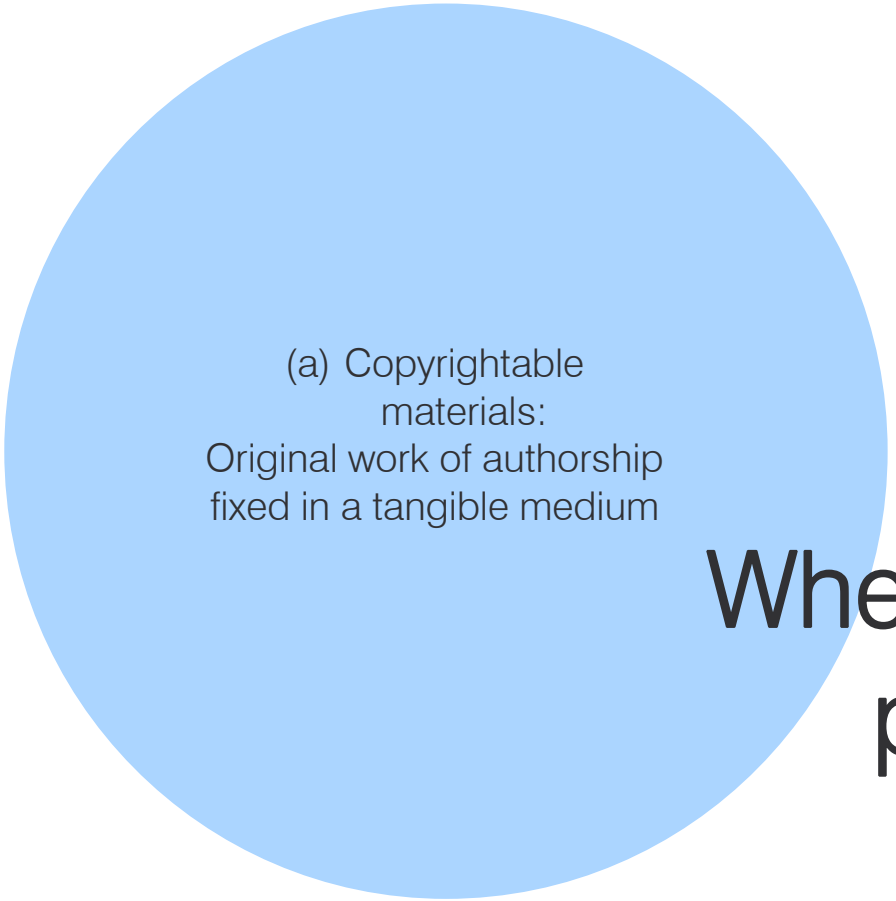
What has changed?

- Ways of copying
- Software moats
- Copyright (Lack of?)

Implications on IP Protection:

- Divorce of Copyright and Contract
- Patent
- Data

(b) Not work of authorship



Where does your AI product fall?

What is an AI model?

Image of traditional source code

```
import os, time
x = "Welcome to The Mirror."
y = 0

while y <= len(x):
    os.system("clear")
    print(x[:y])
    time.sleep(0.2)
    y = y+1
time.sleep(2)
x = "You will stay here for as long as you can, you may forfeit when you would like to."
y = 0

while y <= len(x):
    os.system("clear")
    print(x[:y])
    time.sleep(0.2)
    y = y+1
time.sleep(2)
x = "Please stand here, and stare into the mirror."
y = 0

while y <= len(x):
    os.system("clear")
    print(x[:y])
    time.sleep(0.2)
    y = y+1
time.sleep(2)
x = "This is you in the mirror:"
y = 0

while y <= len(x):
    os.system("clear")
    print(x[:y])
```

What is an AI model?

- What is an LLM on a file level?
- Example 7B, FB16 model
  - Structure file: 50 KB
  - Weights: 13 GB or 13 x 106 KB
- Weights = 99.9999% of the AI models

```
[ 0.0065, -0.0032,  0.0048, -0.0091,  0.0077, -0.0056,  0.0083, -0.0021,  0.0012, -0.0045],
[-0.0074,  0.0023, -0.0089,  0.0067, -0.0015,  0.0042, -0.0036,  0.0099, -0.0071,  0.0028],
[ 0.0052, -0.0068,  0.0021, -0.0043,  0.0081, -0.0094,  0.0037, -0.0019,  0.0064, -0.0053],
[-0.0027,  0.0045, -0.0098,  0.0072, -0.0051,  0.0039, -0.0086,  0.0013, -0.0067,  0.0024],
[ 0.0089, -0.0059,  0.0036, -0.0022,  0.0047, -0.0078,  0.0061, -0.0095,  0.0029, -0.0014],
[-0.0031,  0.0075, -0.0062,  0.0018, -0.0049,  0.0093, -0.0026,  0.0057, -0.0084,  0.0032],
[ 0.0011, -0.0097,  0.0069, -0.0034,  0.0020, -0.0085,  0.0046, -0.0058,  0.0070, -0.0029],
[-0.0063,  0.0091, -0.0040,  0.0055, -0.0072,  0.0027, -0.0088,  0.0033, -0.0016,  0.0060],
[ 0.0076, -0.0025,  0.0041, -0.0096,  0.0030, -0.0069,  0.0082, -0.0017,  0.0050, -0.0044],
[-0.0087,  0.0038, -0.0073,  0.0022, -0.0054,  0.0065, -0.0012,  0.0090, -0.0035,  0.0049]
```

Vibe Coding

Write JavaScript code to shuffle all the cards in a deck

Here is a simple JavaScript implementation to shuffle a standard 52-card deck using the **Fisher-Yates (Knuth) shuffle** algorithm, which is efficient and unbiased:

javascript Copy Edit

```javascript
// Create a standard 52-card deck
function createDeck() {
  const suits = ['♠', '♥', '♦', '♣'];
  const ranks = ['A', '2', '3', '4', '5', '6', '7', '8', '9', '10', 'J', 'Q', 'K'
  const deck = [];

  for (let suit of suits) {
    for (let rank of ranks) {
      deck.push(`${rank}${suit}`);
```

Outline

Background:

- IP Overview
- DeepSeek and Distillation Story

What has changed?

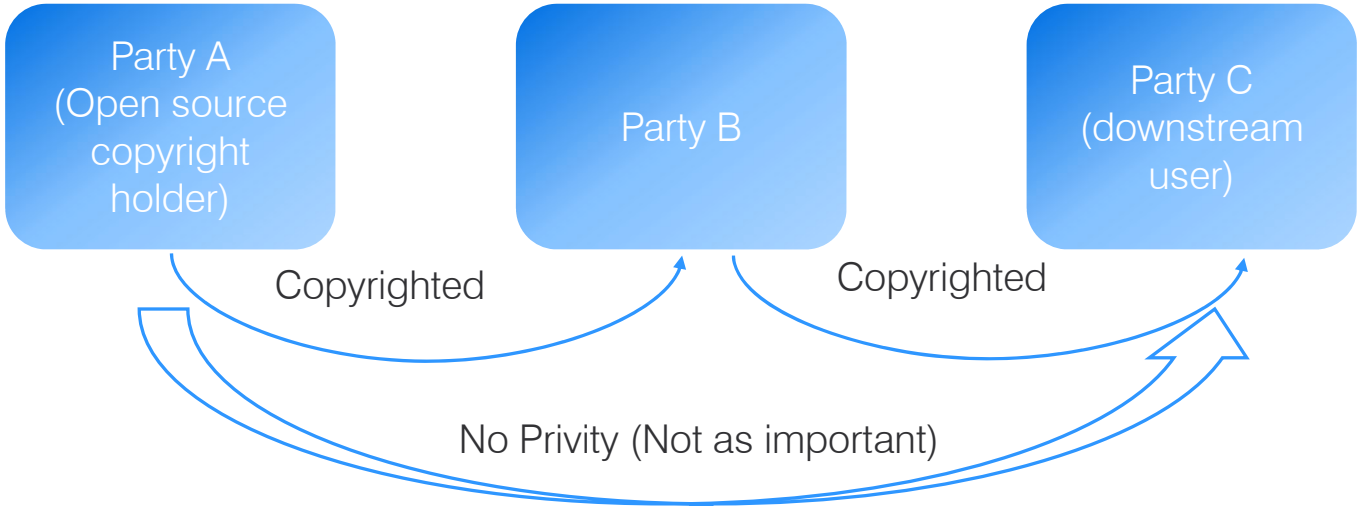
- Ways of copying
- Software moats
- Copyright (Lack of?)

Implications on IP Protection:

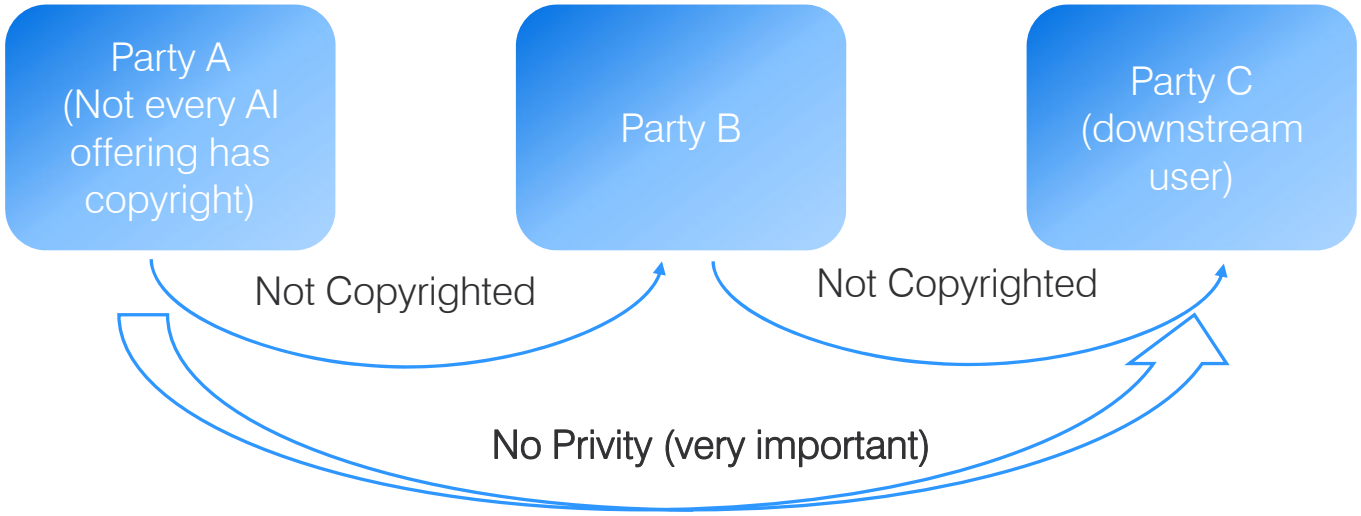
- Divorce of Copyright and Contract
- Patent
- Data

Divorce of copyright law and contract law

- An open source license is the example here, but similar thinking can apply to terms of service and other downstream contractual restrictions
- Conventionally, open source licenses are enforced through copyright law

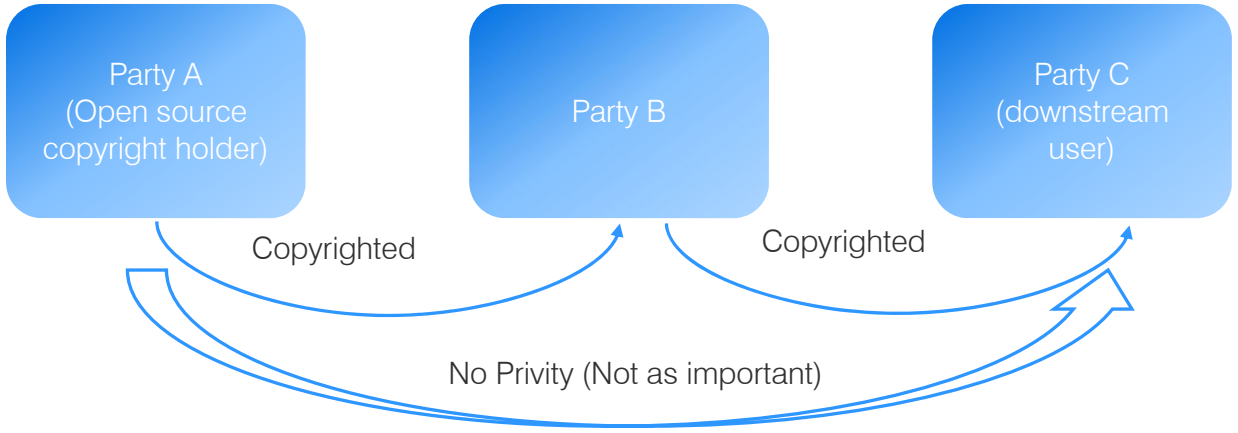


Potential paradigm shift

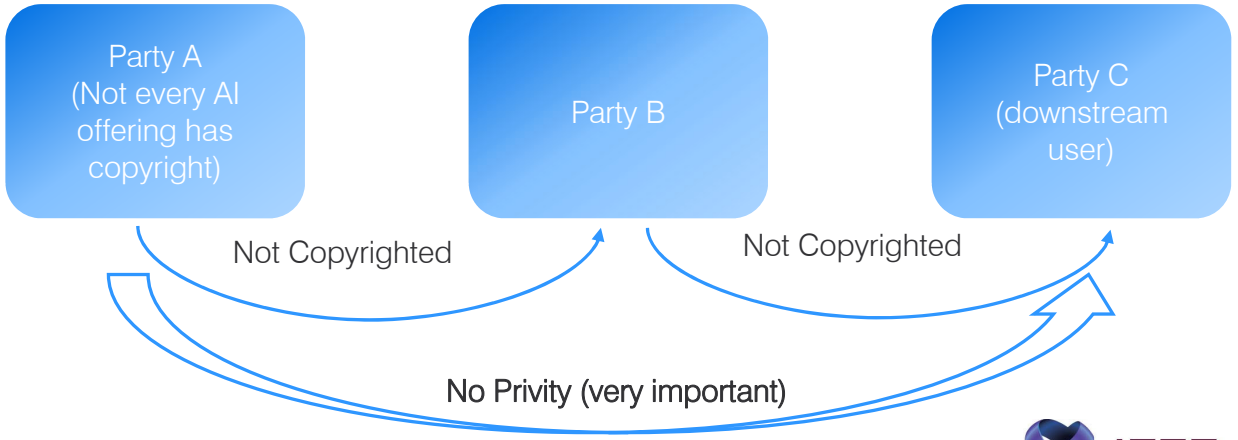


Potential paradigm shift

Conventional software regime:



New, higher-risk regime:



Outline

Background:

- IP Overview
- DeepSeek and Distillation Story

What has changed?

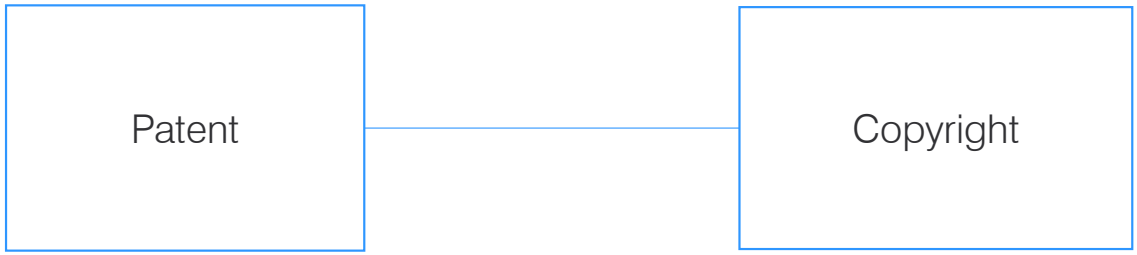
- Ways of copying
- Software moats
- Copyright (Lack of?)

Implications on IP Protection:

- Divorce of Copyright and Contract
- **Patent**
- Data

Patents become more important

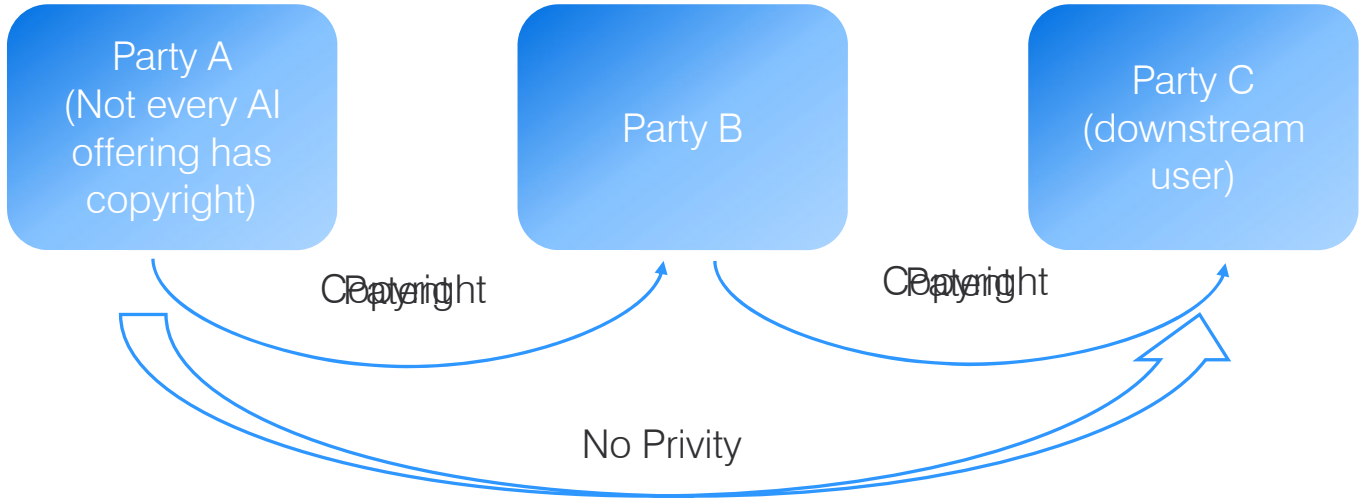
- Model weights are likely just functional elements with thin (if any) copyright protection
- Patents can cover model behaviors and protect against distillations
- Patents are effective in protecting AI wrapper



Patents for open source, not for exclusivity

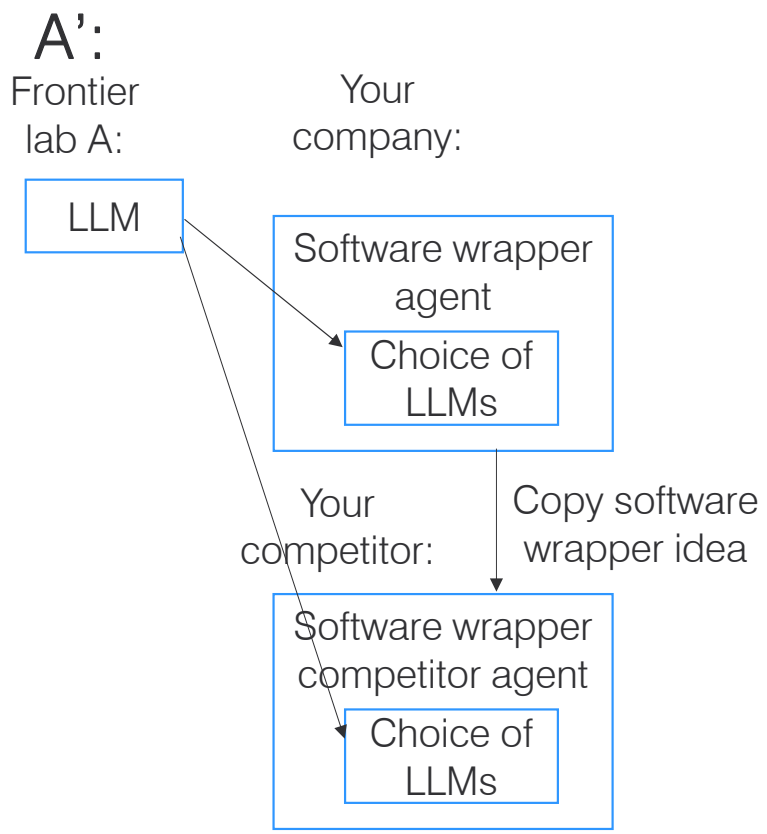
- Open source may be able to be enforced under patent
- If powerful AI products are found to have no copyright, you might *have to* enforce your open source license under patents
- This means filing patents may not be for the exclusivity, but for covering your open-source products so that you can control your open-source products

New patent regime to replace copyright?



Patents – does it make sense to file?

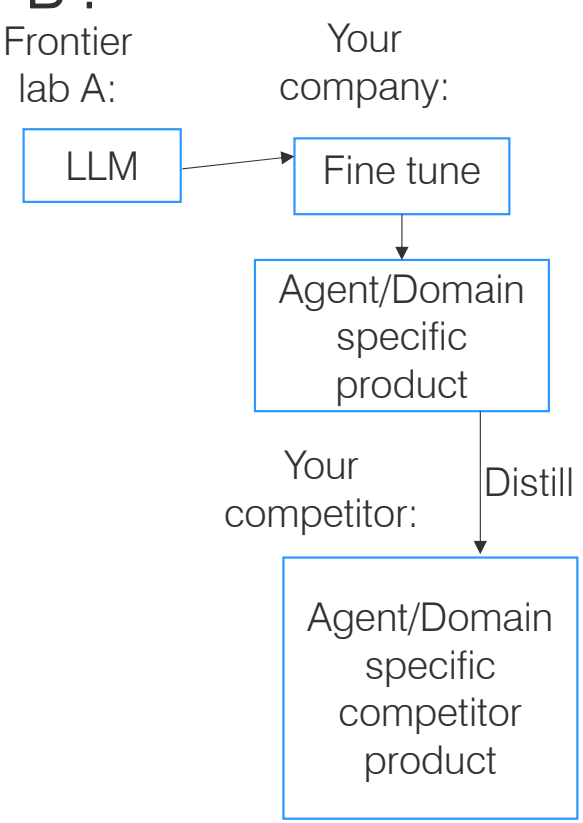
- Spotting what is vulnerable to distillation
  - Not as vulnerable
    - Backend agent that performs work but does not need to print out the entire process
  - More vulnerable
    - Frontend generative models
    - Image/video generation models
    - Protein simulation models



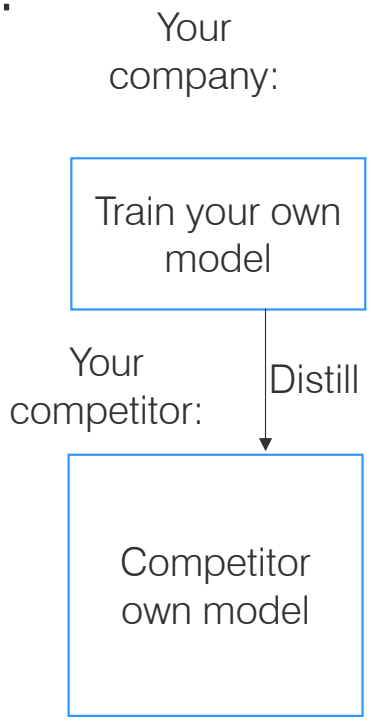
Patents – why it might work?

- Patent is flexible in defining what is invention
- But the key is the right planning
- Patent disclosure cannot be amended after initial filing

B':

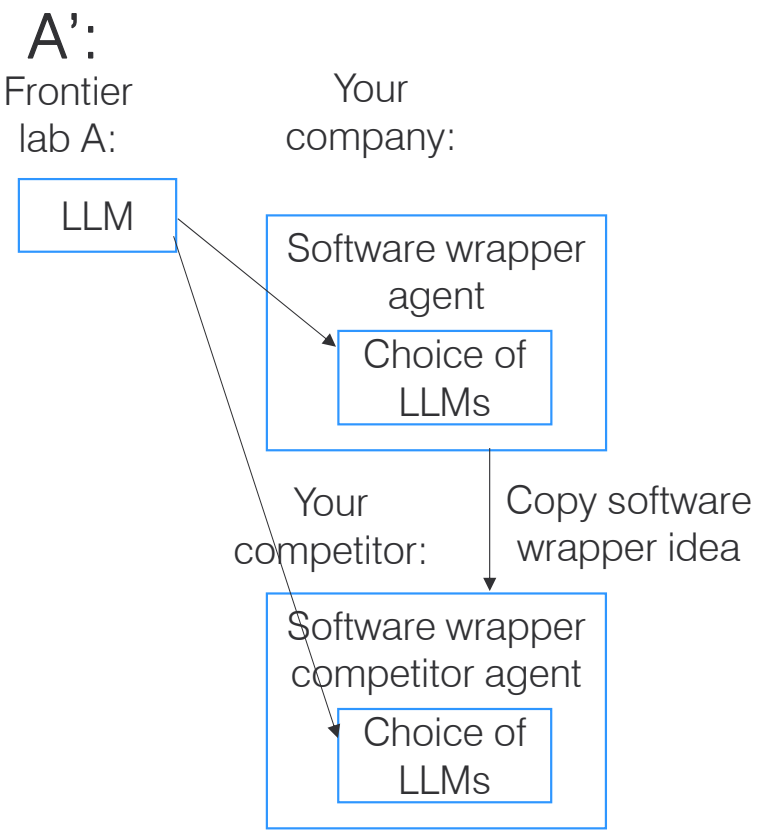


C':



Patents – why it may/may not adequately protect AI agents

- Patents are exclusionary
- Injunctions between competitors are highly likely if the defendants are found infringing after trial
- AI agents are particularly suited for patenting because this is about an application of technology
  - Treat the AI model as a black box
  - Just protect the application (i.e., the "use case")



Outline

Background:

- IP Overview
- DeepSeek and Distillation Story

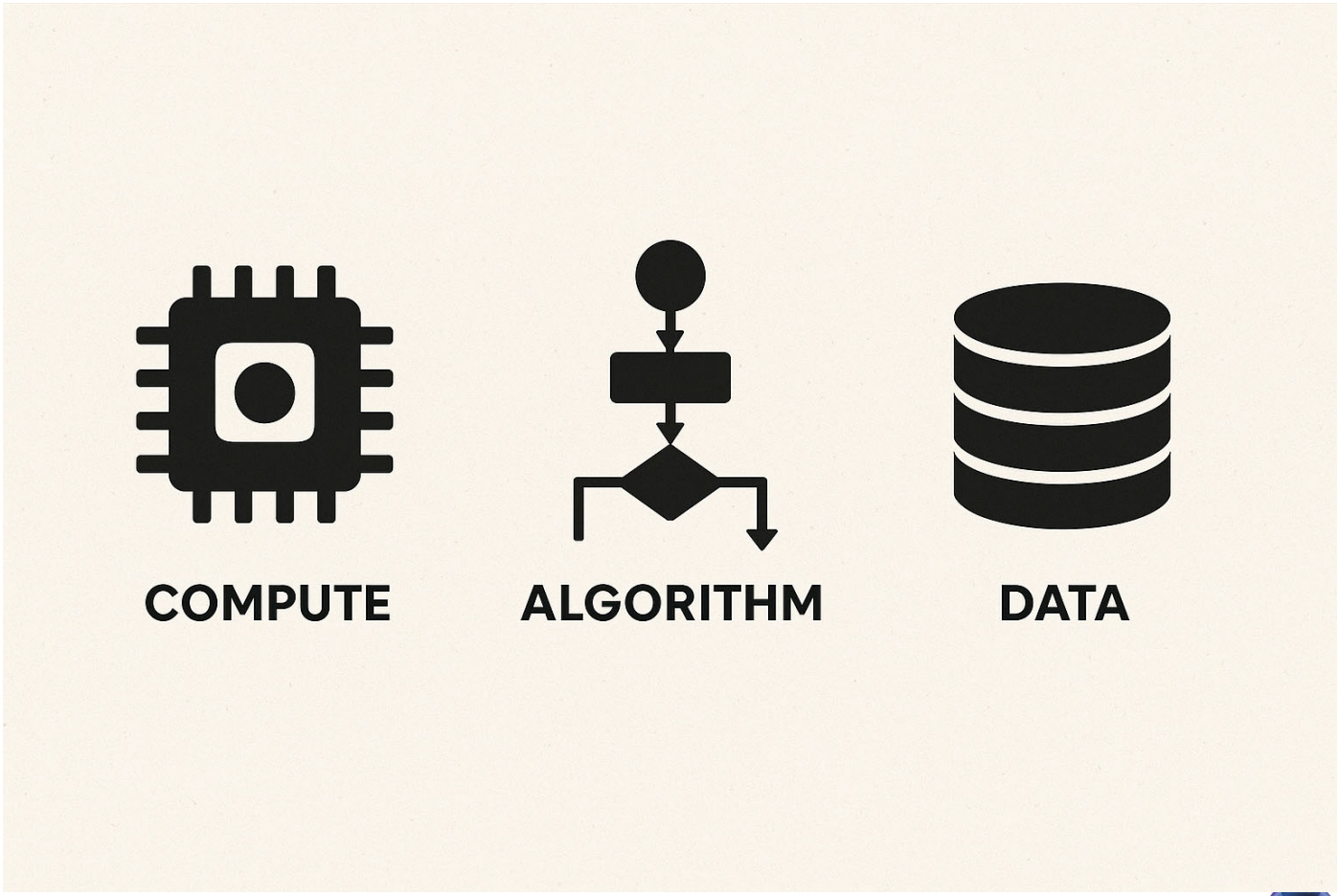
What has changed?

- Ways of copying
- Software moats
- Copyright (Lack of?)

Implications on IP Protection:

- Divorce of Copyright and Contract
- Patent
- Data

The trinity of A.I.



Data privacy compliance

- Data will be product differentiator
  - Especially against AGI in the future
- Terms of Service
  - Maximize usage under the compliance framework
- Risk of non-compliance

Outline

Background:

- IP Overview
- DeepSeek and Distillation Story

What has changed?

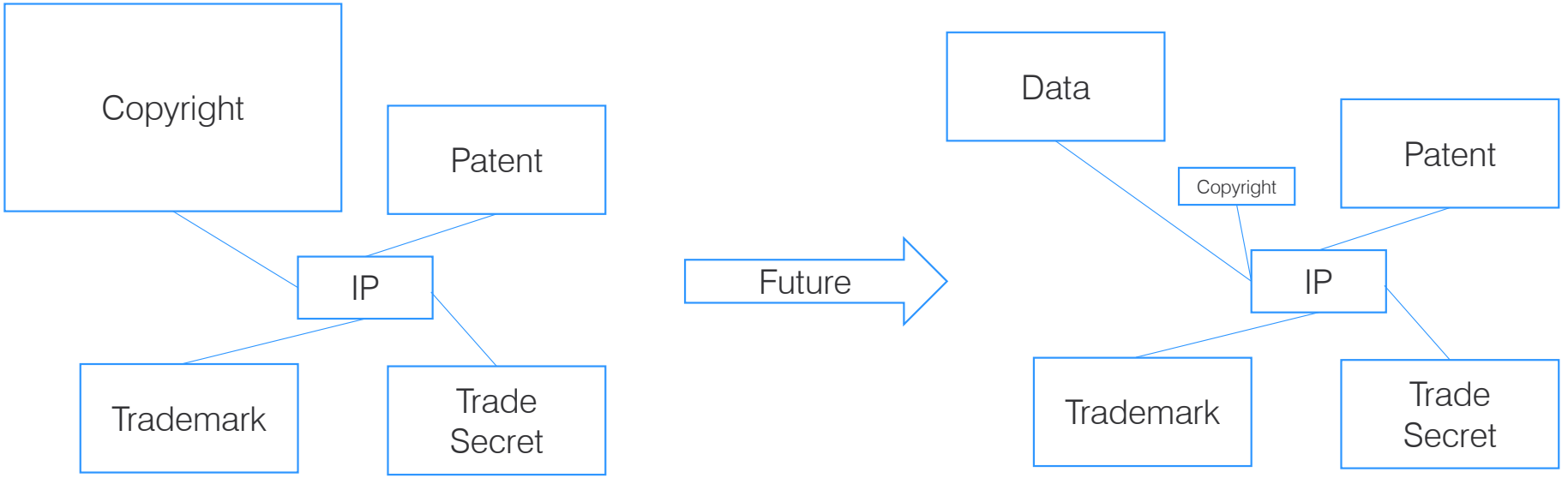
- Ways of copying
- Software moats
- Copyright (Lack of?)

Implications on IP Protection:

- Divorce of Copyright and Contract
- Patent
- Data

Conclusion

- Identify key areas for potential patents
- Understand the uncertainty in AI from the IP perspective
- Understand copyright protection will be significantly weakened
- Safeguard and maximize your data under the compliance framework



Questions

Fredrick Tsang

IP Attorney at Fenwick & West

FENWICK

IEEE-CNSV
Consultants' Network of Silicon Valley

FENWICK