OpenMined

OpenMined is an open-source **community** focused on researching, developing, and elevating tools for secure, privacy-preserving, value-aligned artificial intelligence.

Key Activities

- **Awareness:** Raise awareness of Secure, Private, & Value Aligned AI
- **Tools:** Lower the barrier-to-entry by building open-source tools
- **Community:** We have really fun Hackathons…

Outline

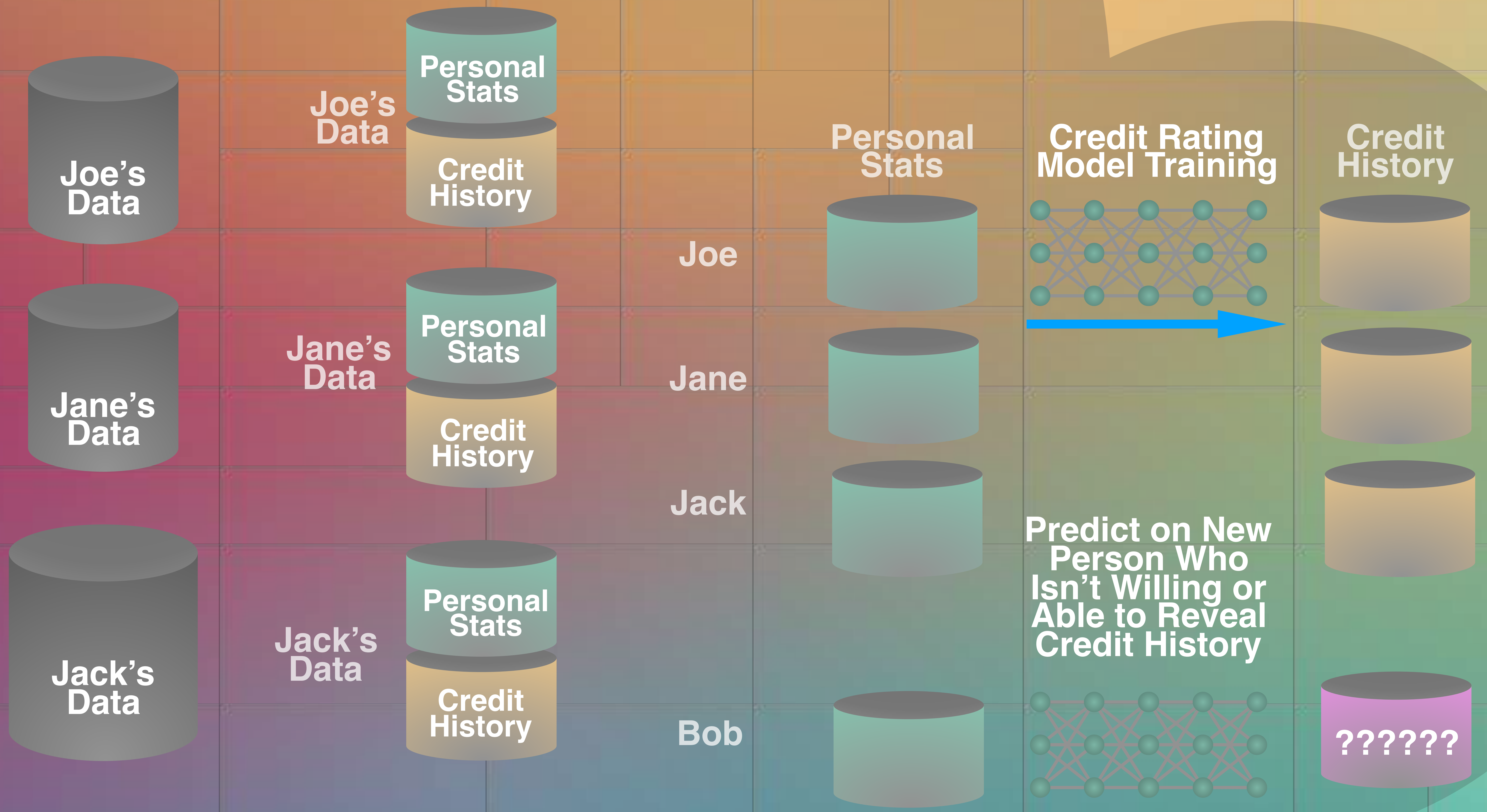
- **Why:** the AI Business Model has privacy problems
- **How:** an Introduction to the Core Technologies of OpenMined
    - Federated Learning
    - Homomorphic Encryption
    - Multi-Party Computation
    - Gradient Marketplace
- **Roadmap & Demos**

AI Inc.

The AI Business Model

1. Acquire Data about People

AI Inc.



2. Train a model that transforms one dataset into another

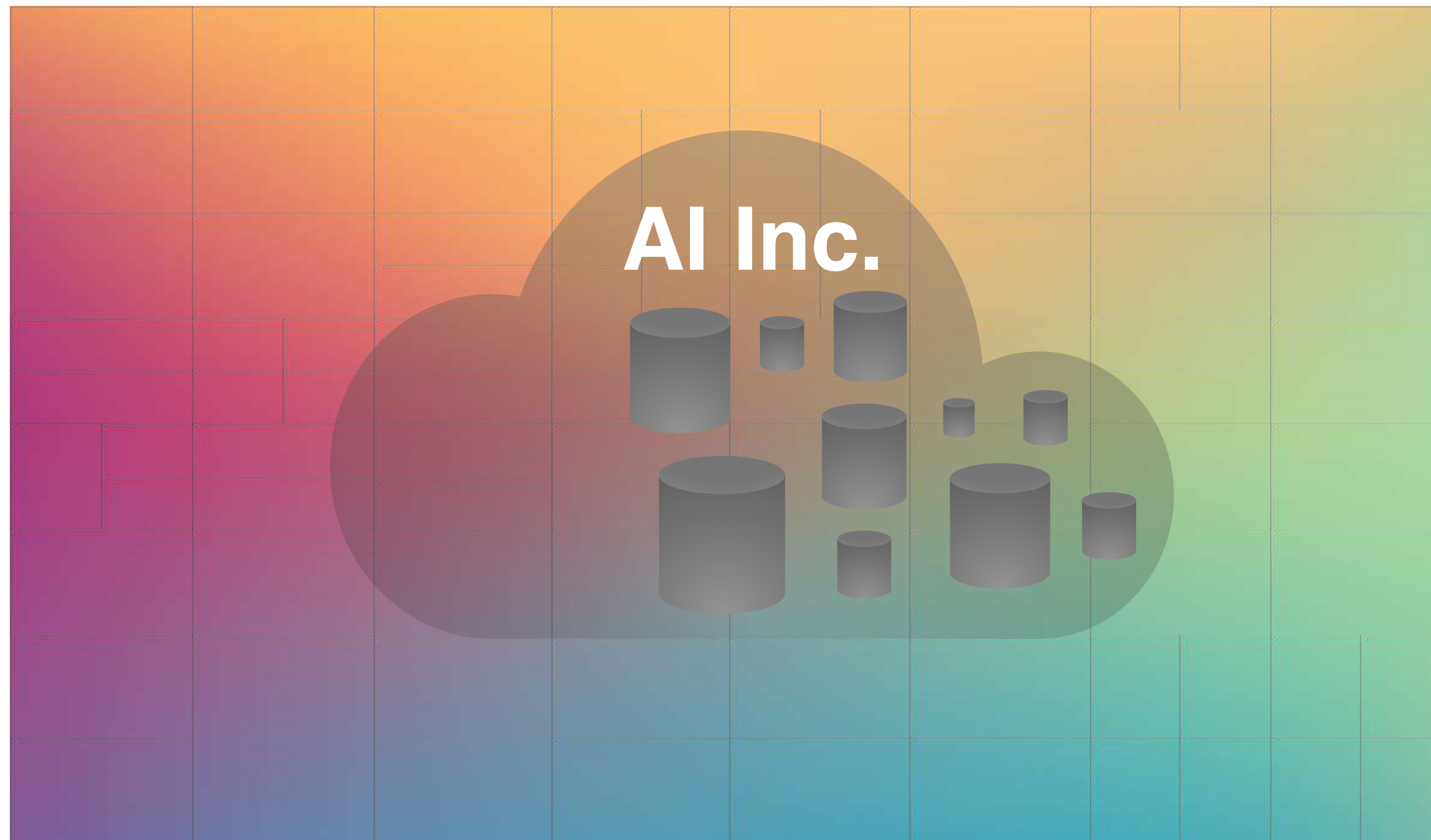
3. Sell the Use of that Model

The A.I. Business Model

- Step 1: Acquire **Data** about people
- Step 2: Train a **Model** that predicts unknown facts about a person using known facts.
- Step 3: Sell the <u>Use</u> of that **Model** (the App)

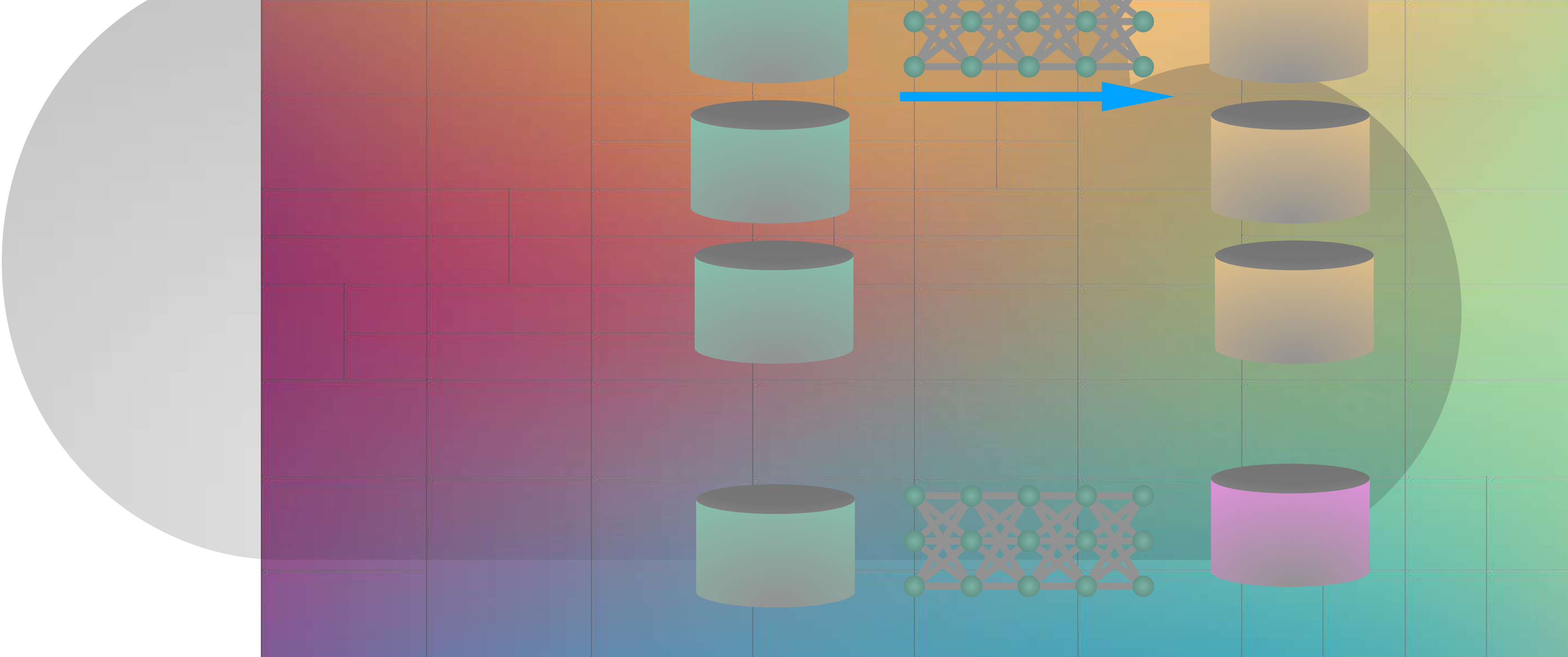
Problems with the AI Business Model

- Step 1: acquire **Data** about people
  - **Privacy:** people lose control of their data
  - **"Sensitive Products"** don't get made



Problems

- Step 2: train a **Model** that predicts unknown facts about a person using known facts.
  - **Contagious Privacy Loss:** if one person reveals private information, AI can be used to reveal private information of others through prediction
  - **Lack of Competition:** there is very little market competition because most datasets are proprietary. (AI Inc. vs AI Corp.)
  - **Unfair Predictions:** corporate datasets only sample the target market (customers) of the company that acquired them, leading to biased AI predictions.



Problems

- Step 3: sell the use of that **Model** (The App)
  - **Lost Natural Income:** in practice, people are rarely compensated for their data
  - **Unknown Value of Data:** How valuable is any datapoint?
  - **Unknown Accuracy of Predictions:** the quality of deployed models is unknown
  - **Digital Assets Hard to Protect:** (i.e., pirated music)

How do we solve these problems?

Potential Solution

- **Train** A.I. on data we cannot see
    - **Privacy Win:** people wouldn’t need to reveal their data

Outline

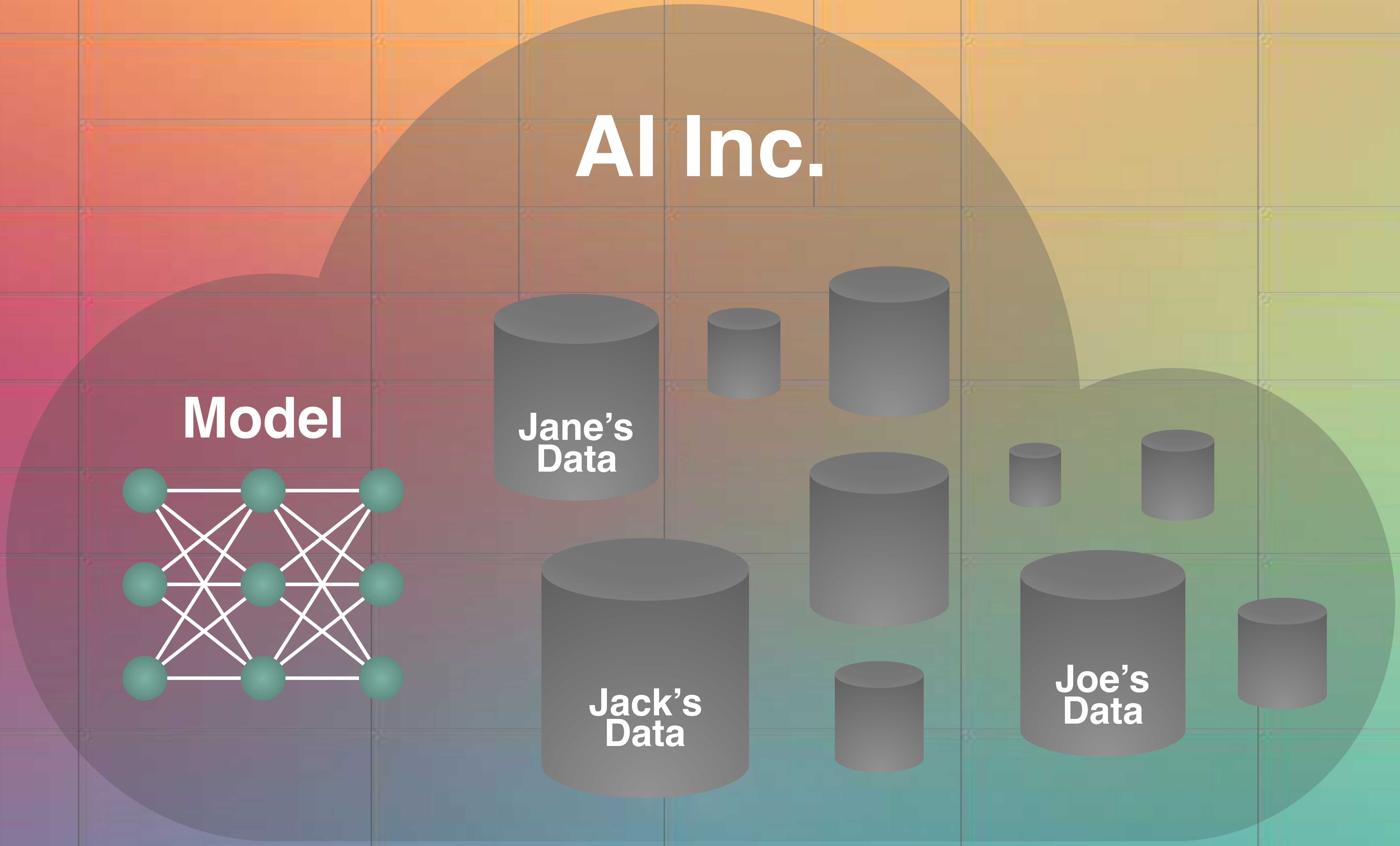
- **Why:** the AI Business Model has privacy problems
- **How:** an Introduction to the Core Technologies of OpenMined
  - Federated Learning
  - Homomorphic Encryption
  - Multi-Party Computation
  - Gradient Marketplace
- **Roadmap & Demos**

Introduction to

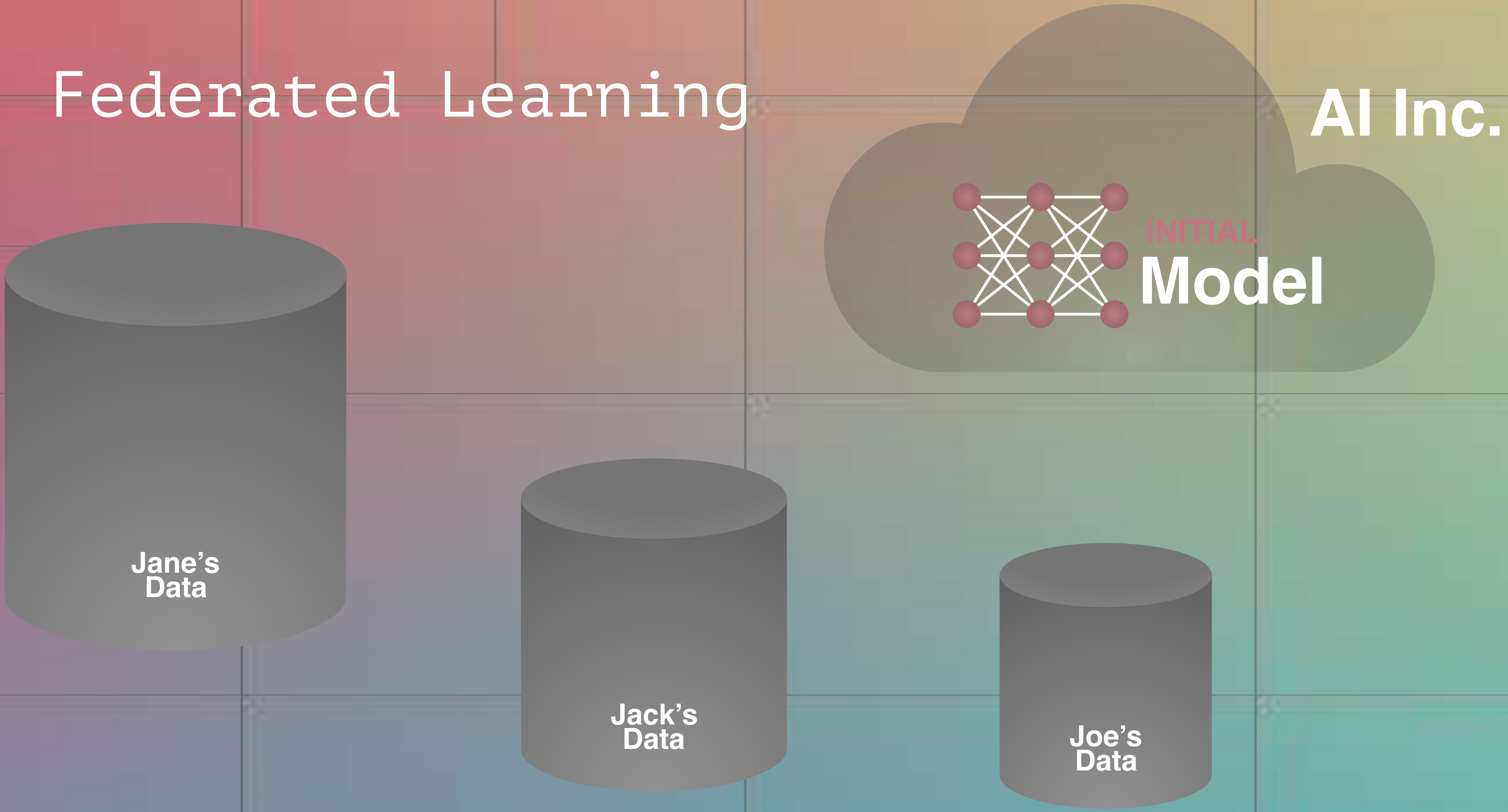
Federated Learning

for Safe AI

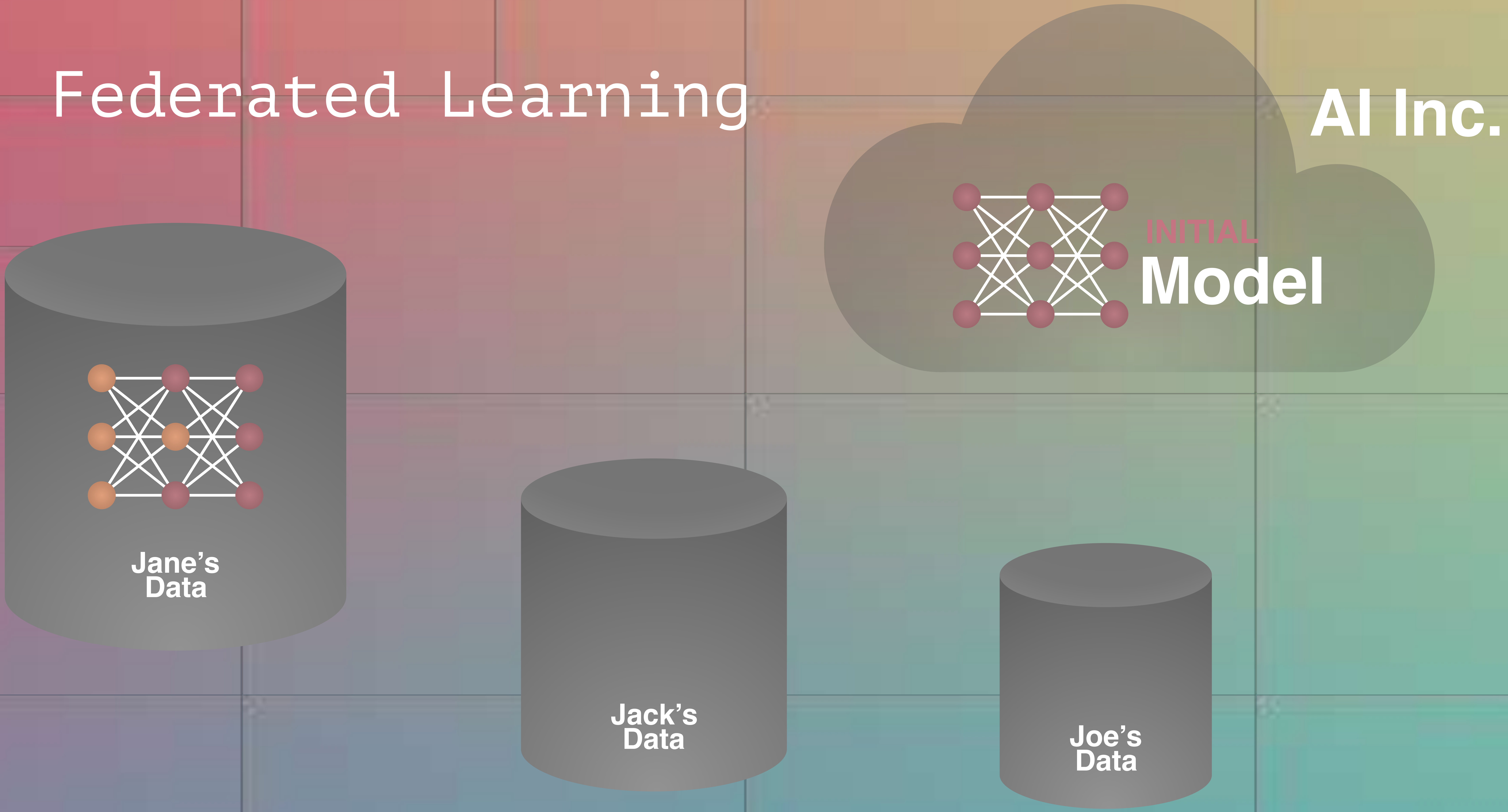
Non-Federated Learning



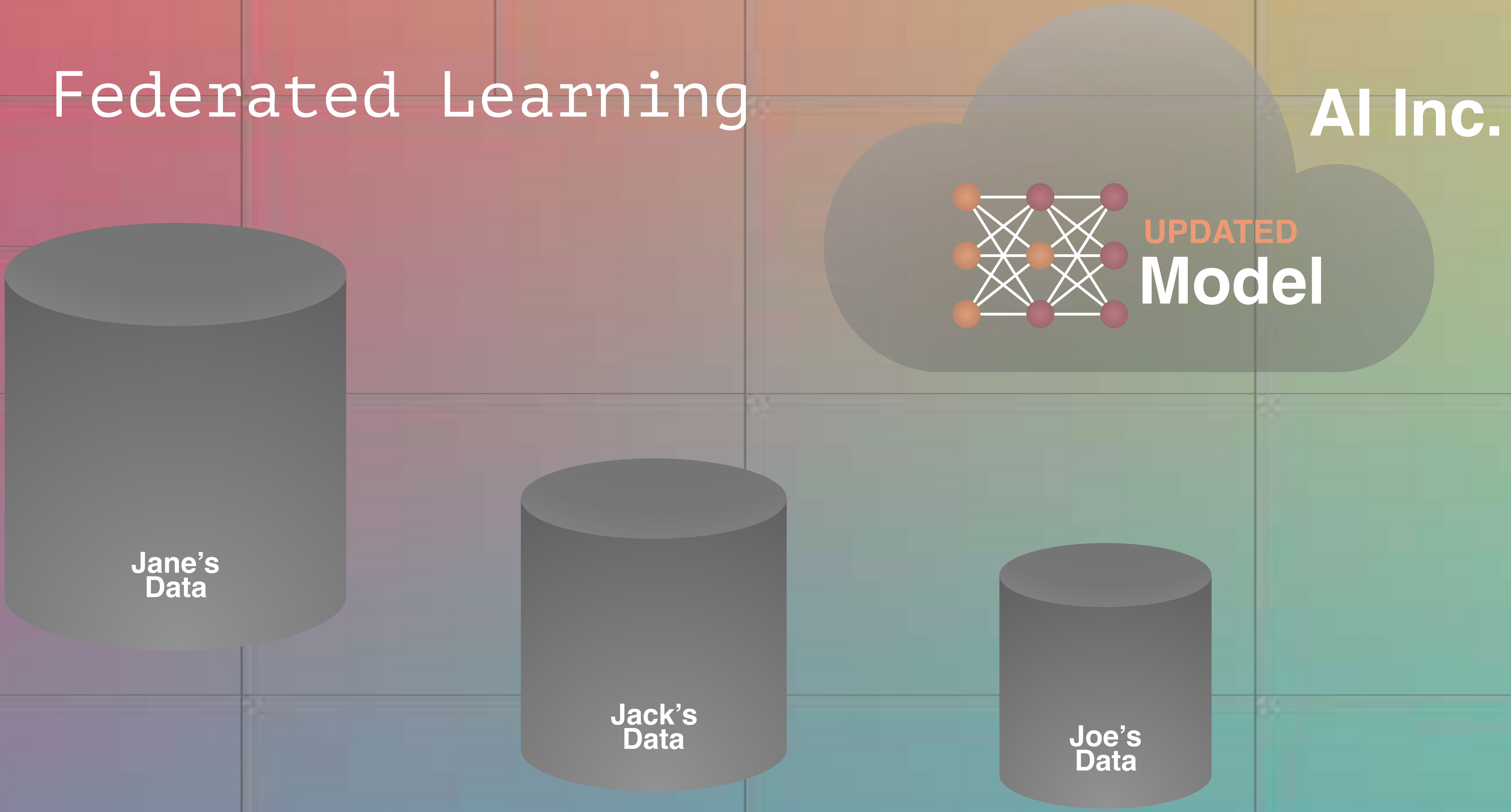
Federated Learning



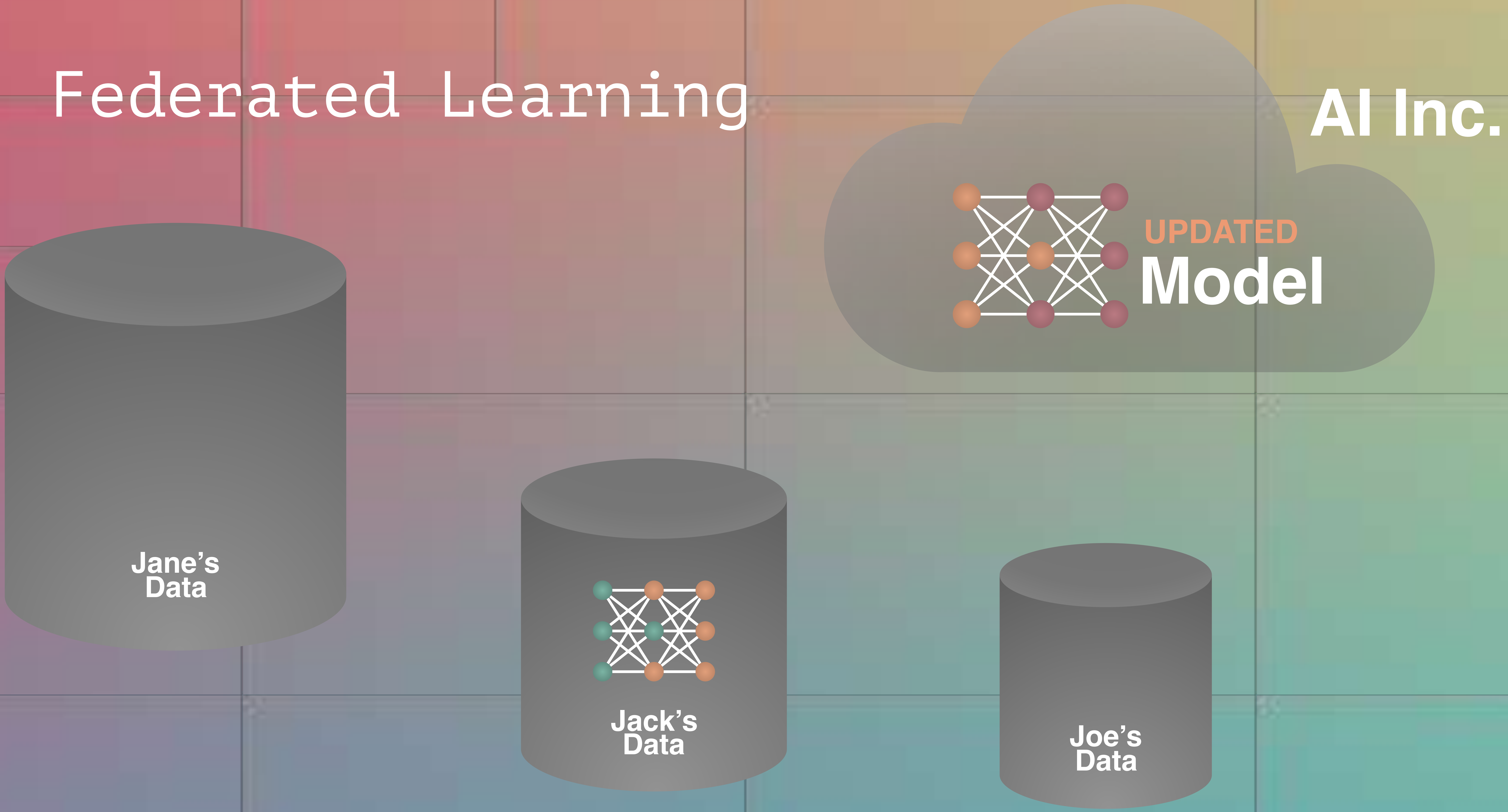
Federated Learning



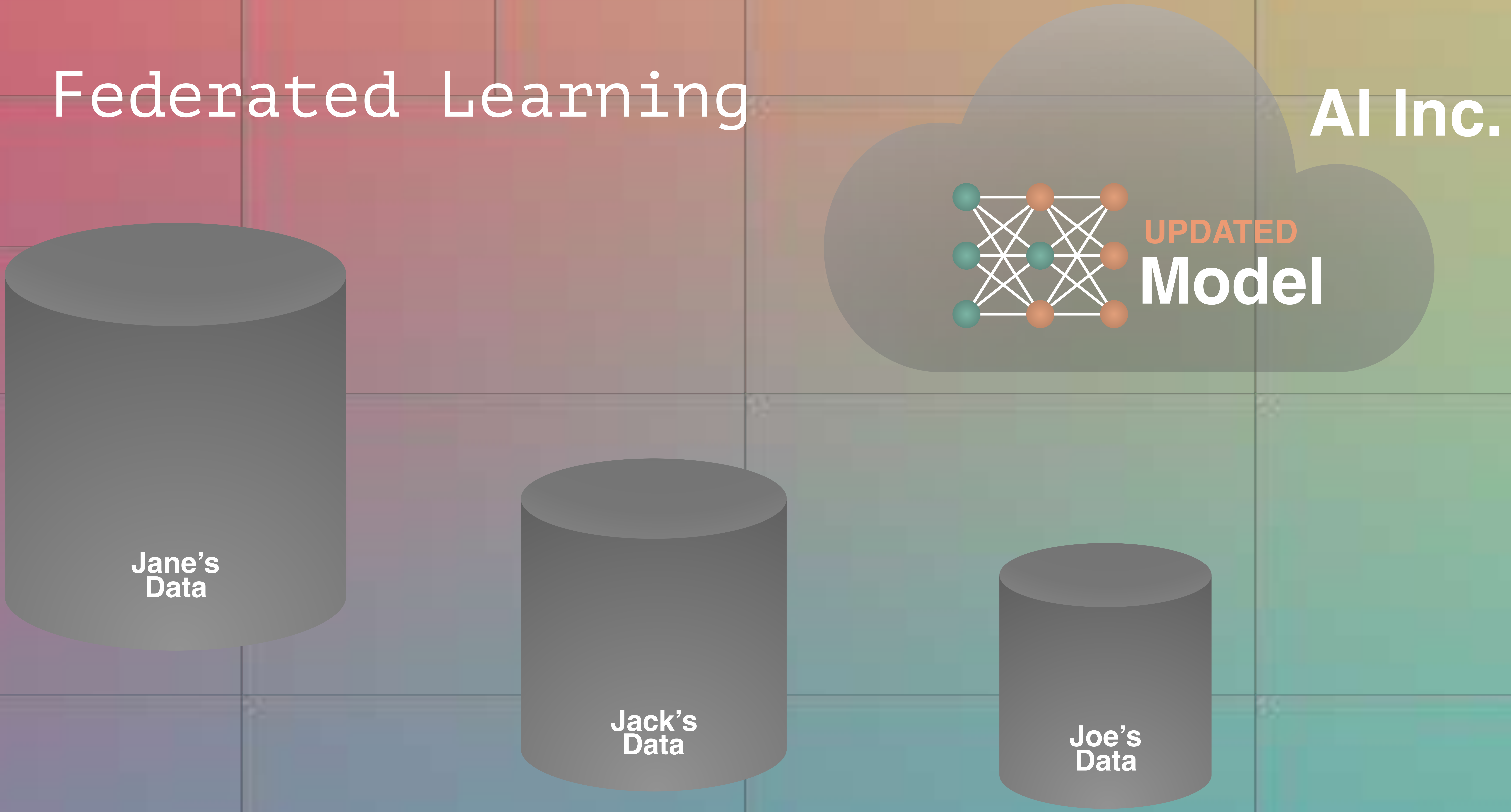
Federated Learning



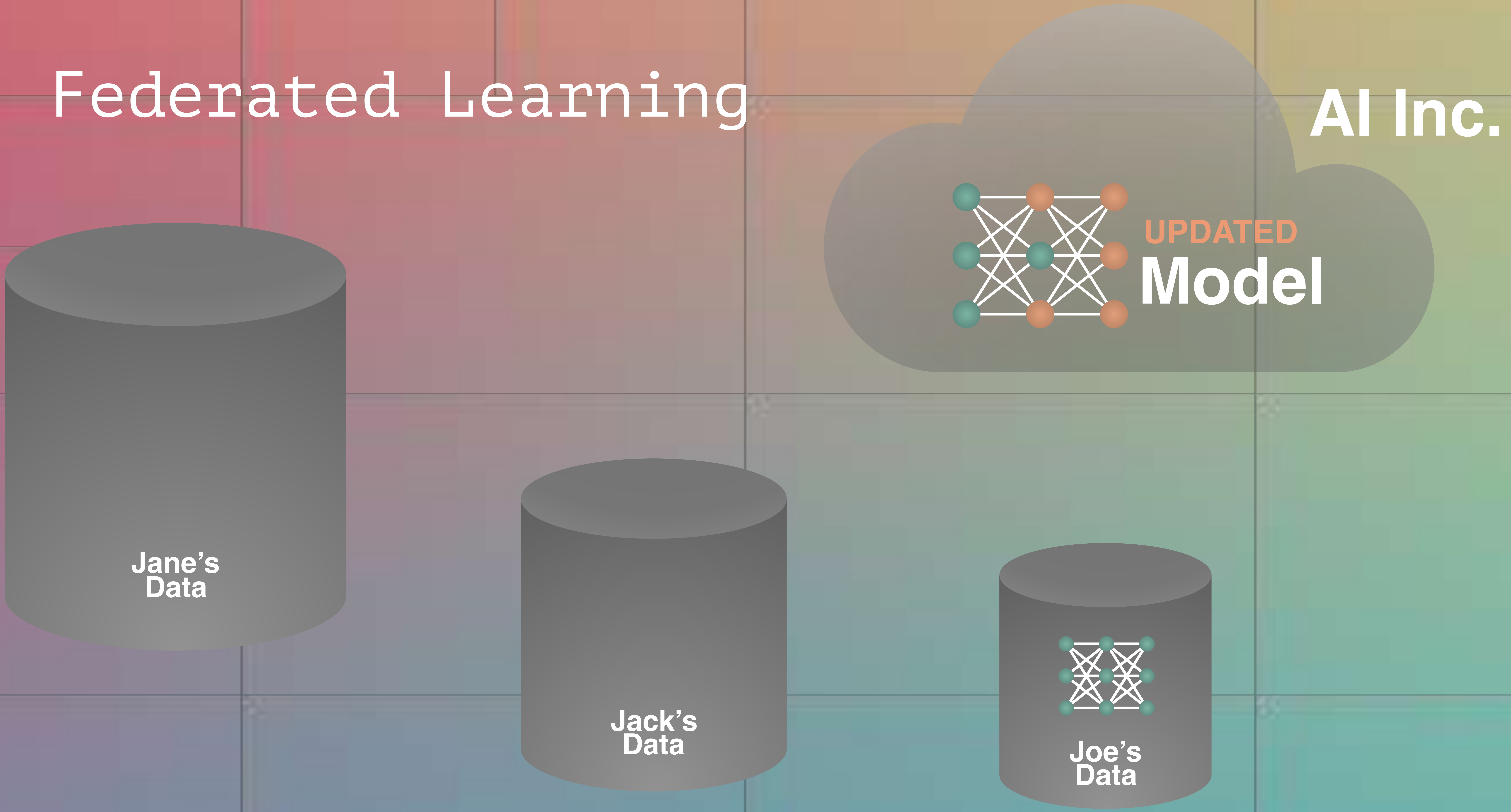
Federated Learning



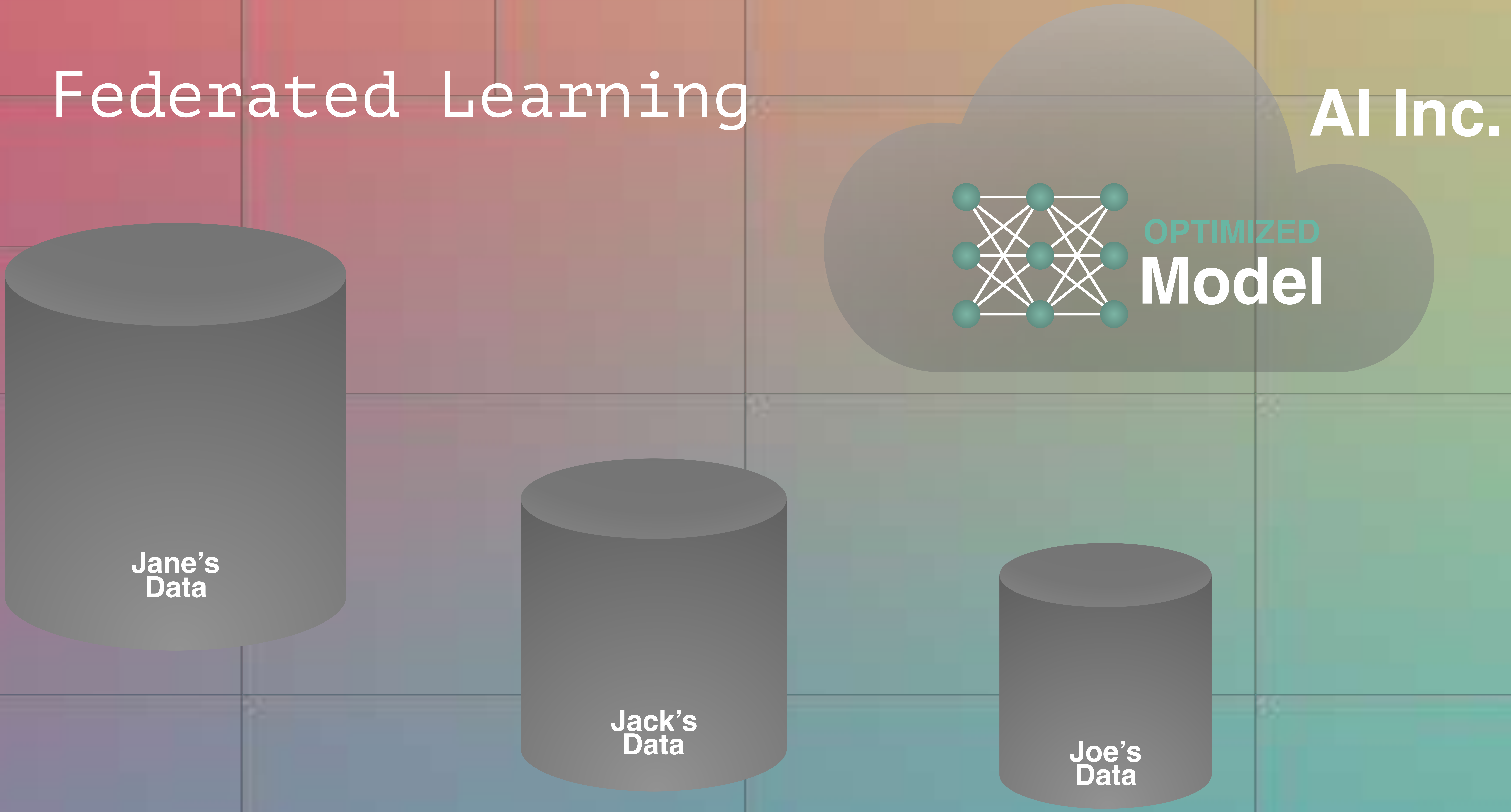
Federated Learning



Federated Learning



Federated Learning



Open Source

TensorFlow

OpenMined

- Computer Cluster
- Parameter Server

- Federated Learning
- Blockchain Compute Grid

Potential Solution

- **Train** A.I. on data we cannot see - Federated Learning
  - Pro: **the data is kept private**
  - **Theft:** the A.I. is put at risk.
  - **Privacy**: Gradients reveal information about the data
  - **Sensitive Product Problem**

Potential Solution

- **Train** A.I. on data we cannot see without revealing the AI or its training updates to anyone?

Potential Solution

- **Train** A.I. on data we cannot see without revealing the AI or its training updates to anyone?
  - Homomorphic Encryption
  - Multi-Party Computation

Introduction to

Homomorphic Encryption

for Safe AI

Homomorphic Encryption

3 → **Homomorphic Encryptor** → **Cypher A**

0010110101

Homomorphic Encryption

3 → Homomorphic Encryptor → **Cypher A**
0010110101

5 → Homomorphic Encryptor → **Cypher B**
100100110

Homomorphic Encryption

3 → Homomorphic Encryptor → Cypher A
001011010

5 → Homomorphic Encryptor → Cypher B
100100110

2xCypher → Homomorphic Decryptor → 6
0020220202

Homomorphic Encryption

3 → Homomorphic Encryptor → Cypher A
001011010

5 → Homomorphic Encryptor → Cypher B
100100110

2xCypher → Homomorphic Decryptor → 6
0020220202

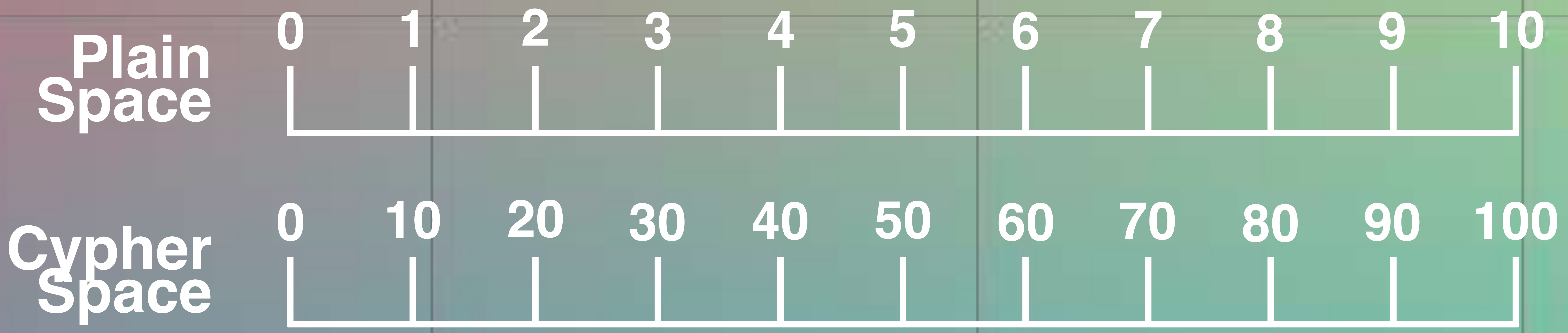
Cypher A + Cypher → Homomorphic Decryptor → 8
1011110210

Homomorphic Encryption

- **Partially Homomorphic Encryption (PHE)**
  — you can only do some operations, such as addition or multiplication
- **Somewhat Homomorphic Encryption (SHE)**
  — you can do any operation, but only a few times
- **Fully Homomorphic Encryption (FHE)**
  — unlimited number of any operation

Homomorphic Encryption

Homomorphic Encryption

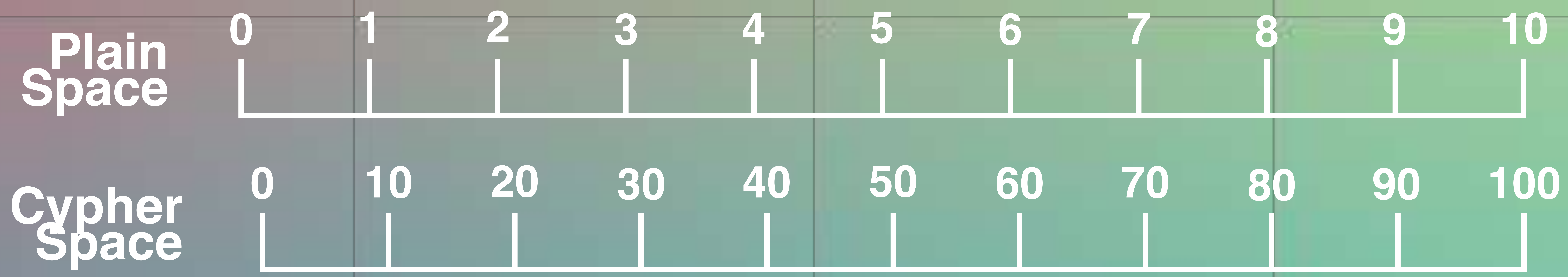


Challenge: hide a number between 0 and 10 (our “plaintext”)

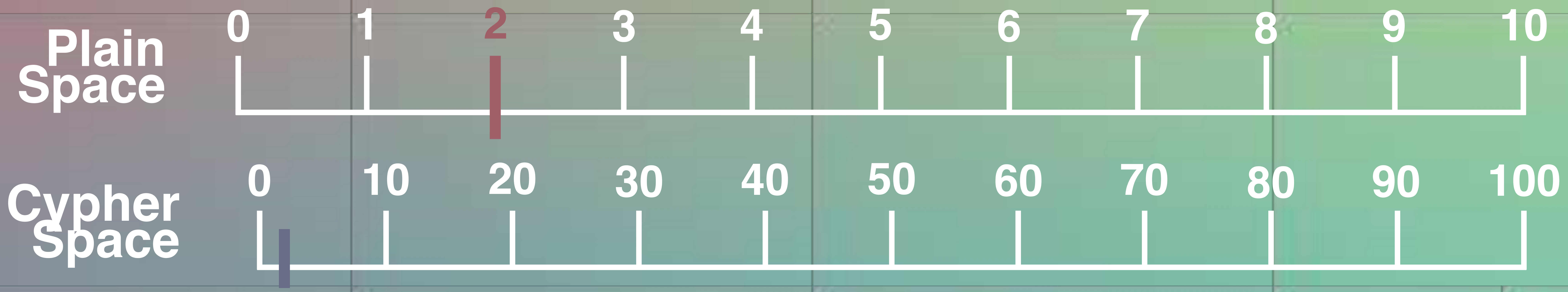
Constraints:

- Somewhere between 0 and 100
- Only we know what it is
- You can add encrypted numbers together

Homomorphic Encryption



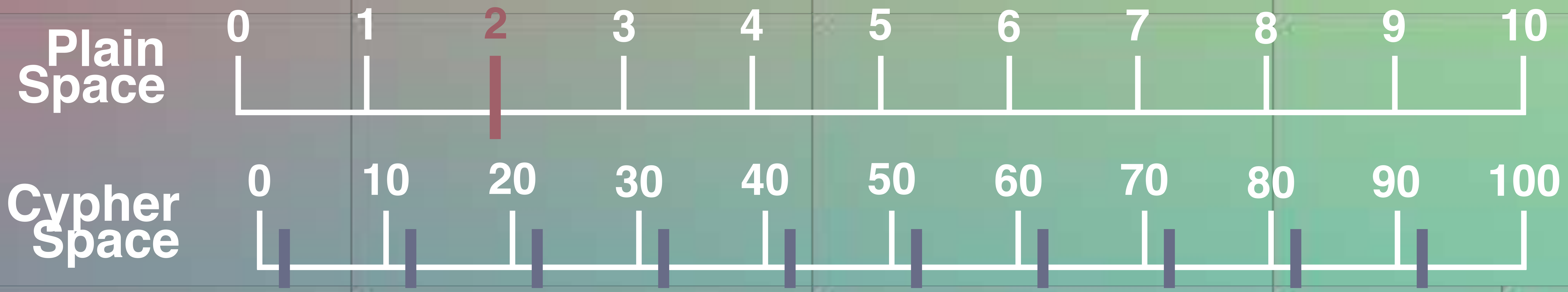
Homomorphic Encryption



$P = 2$

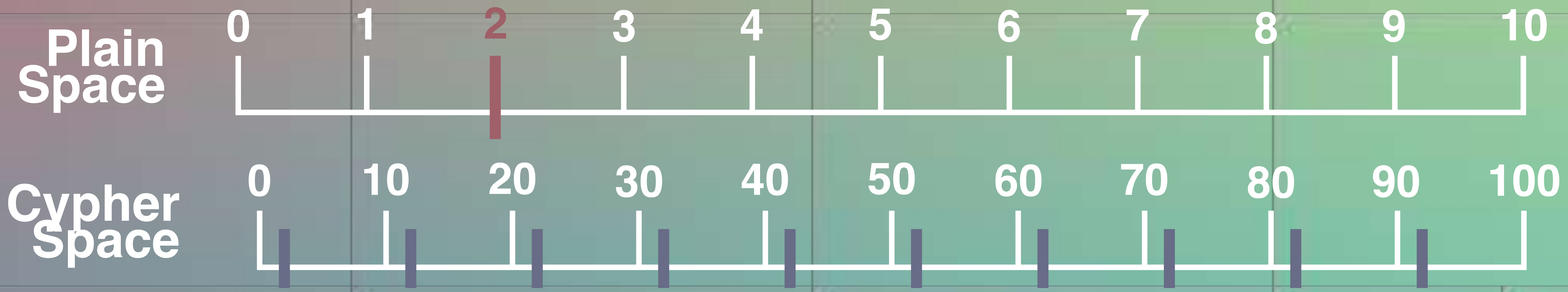
$C = P$

Homomorphic Encryption



P = 2
Secret = 10
C = radient · secret + P

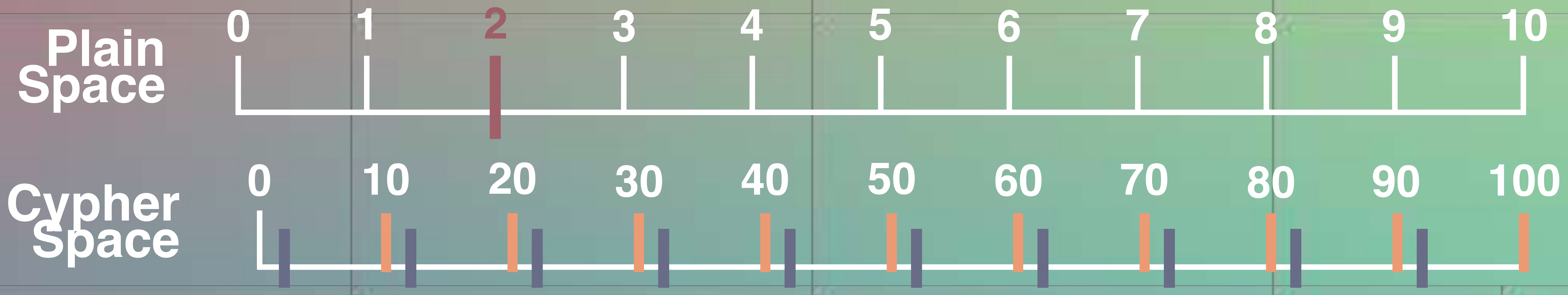
Homomorphic Encryption



P = 2
Secret = 10
C = radient · secret + P

22 -> 2
53 -> 3
2 + 3 = 5
22 + 53 = 75 -> 5

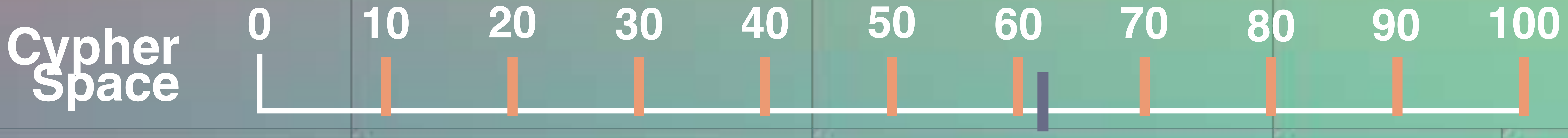
Homomorphic Encryption



P = 2
Secret = 10
C = radient · secret + P

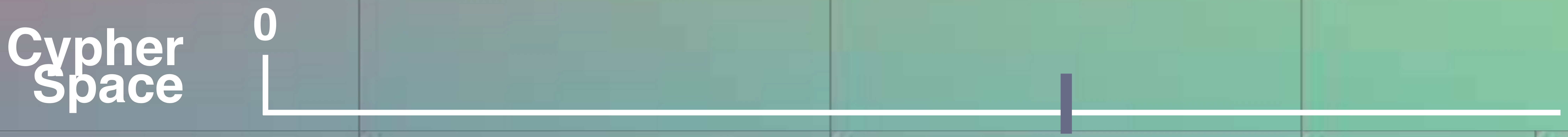
22 -> 2
53 -> 3
2 + 3 = 5
22 + 53 = 75 -> 5

Homomorphic Encryption



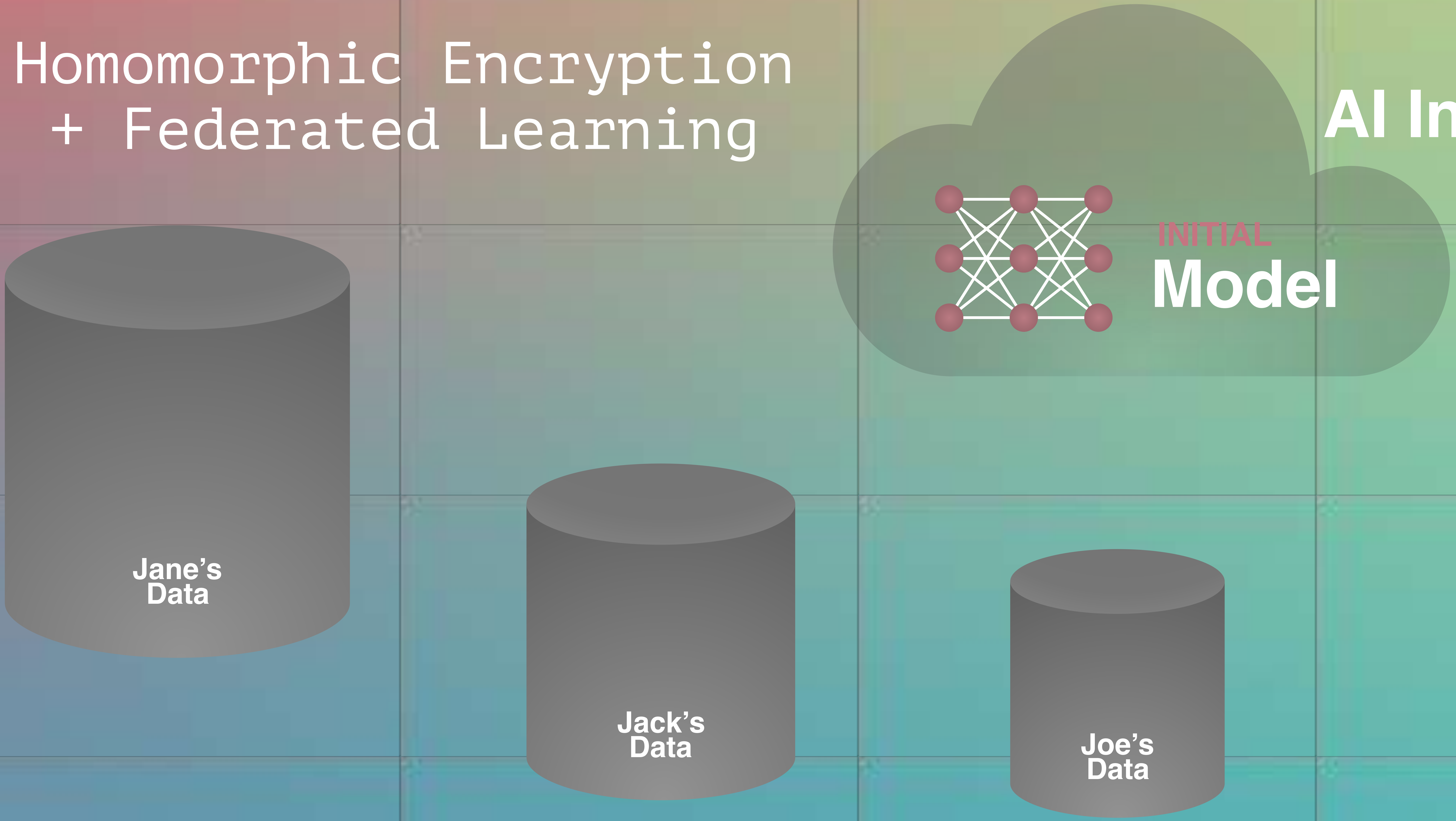
With the secret key

Homomorphic Encryption

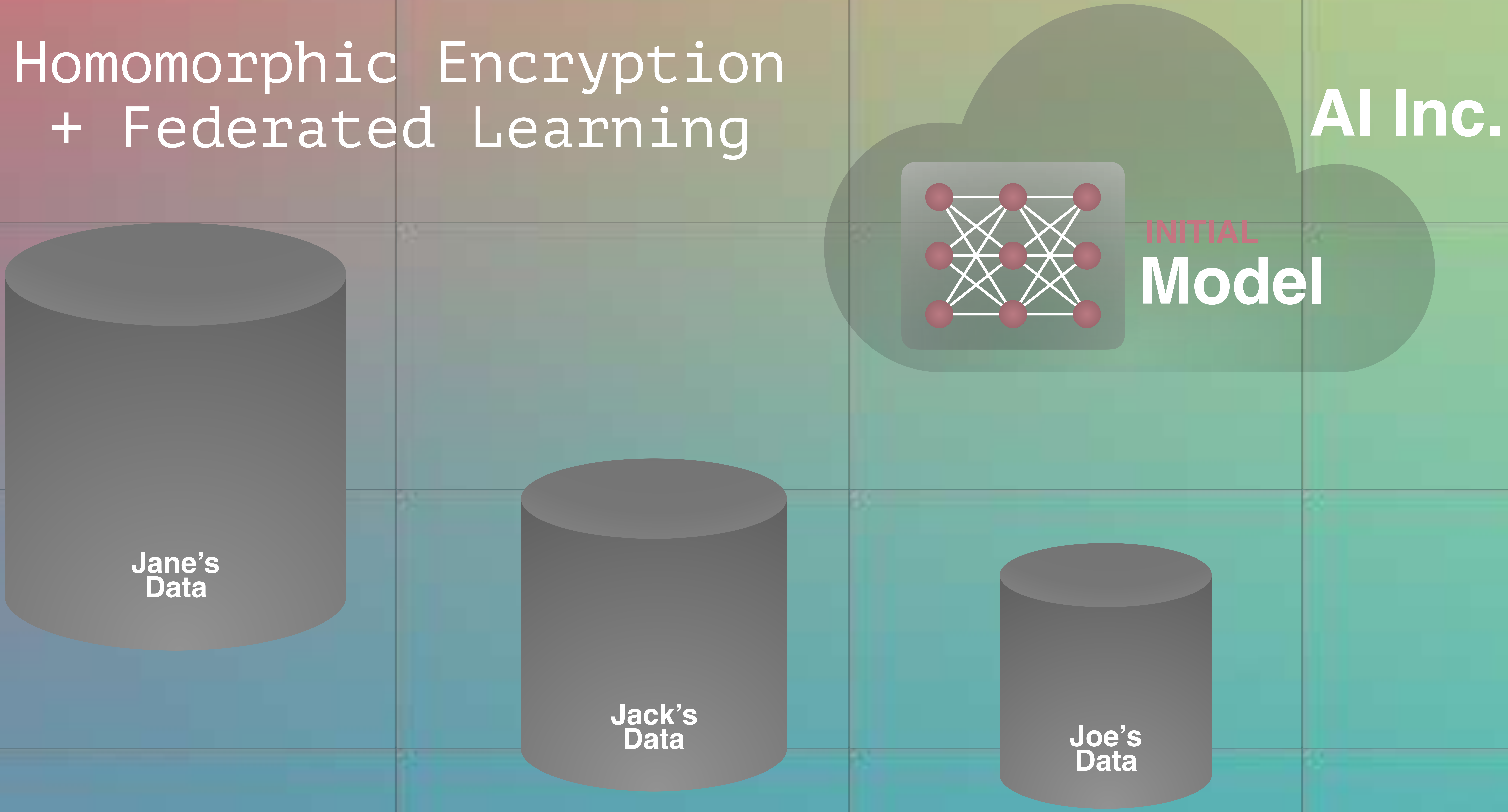


Without the secret key

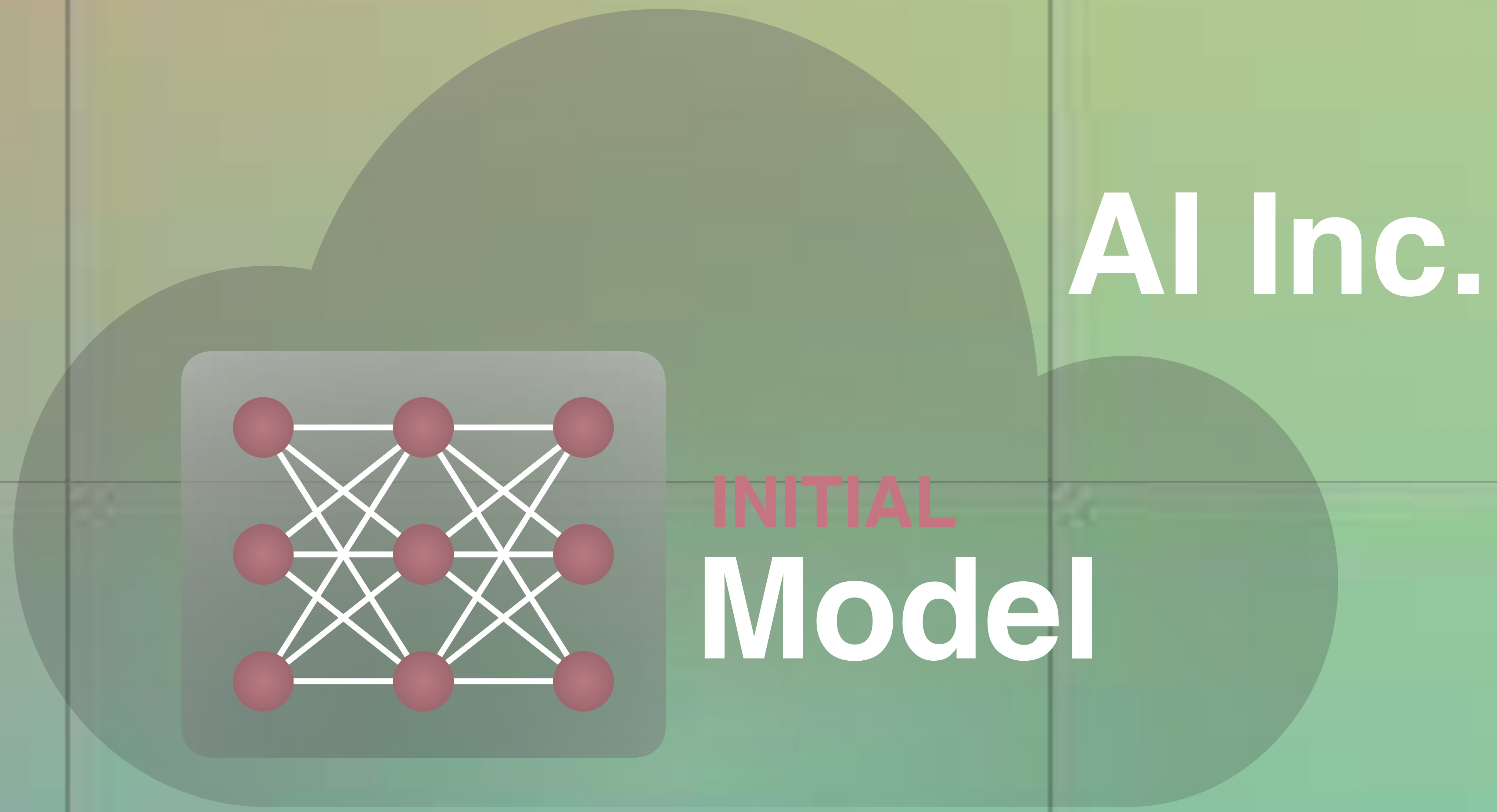
Homomorphic Encryption + Federated Learning



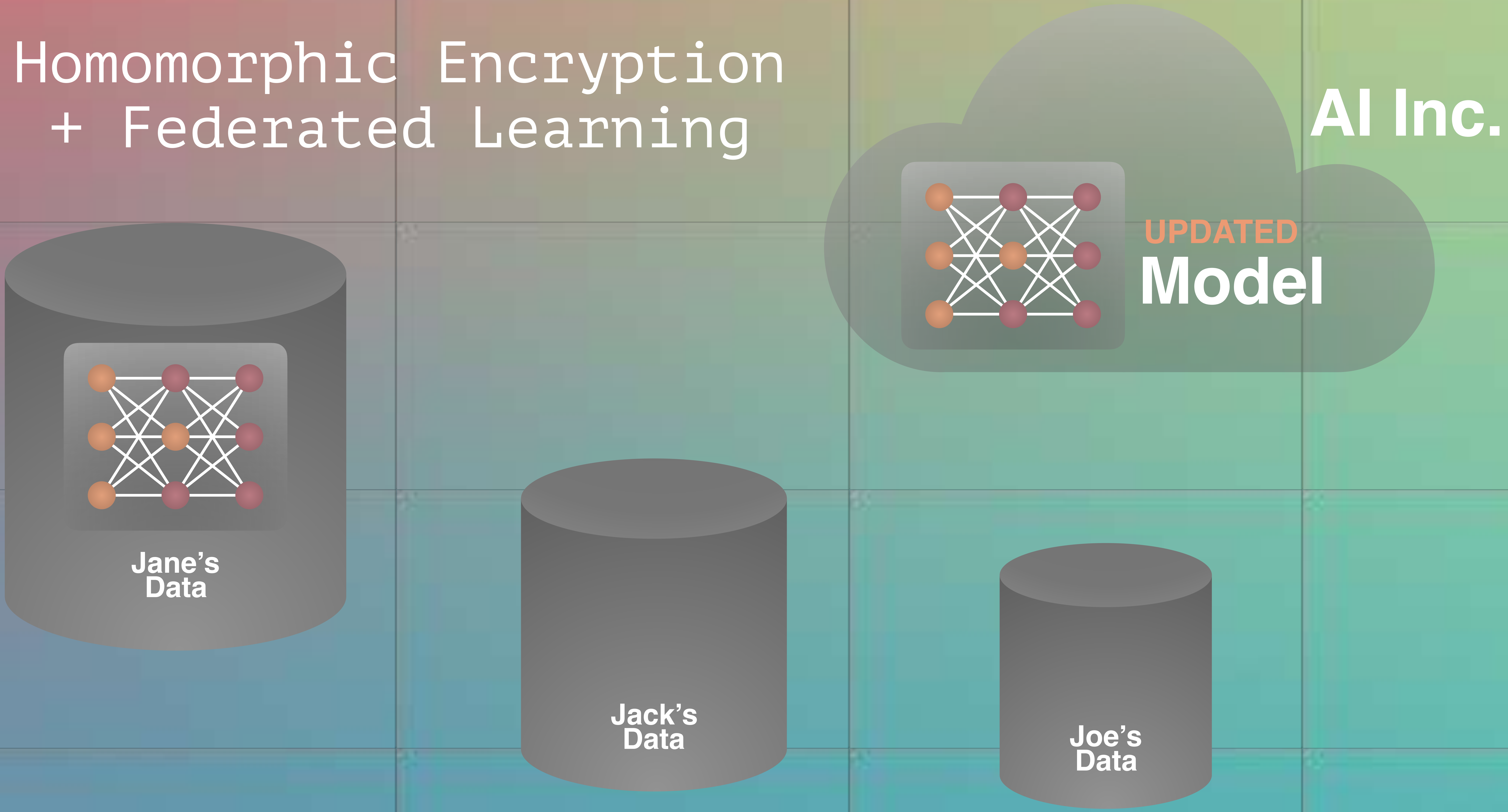
Homomorphic Encryption + Federated Learning



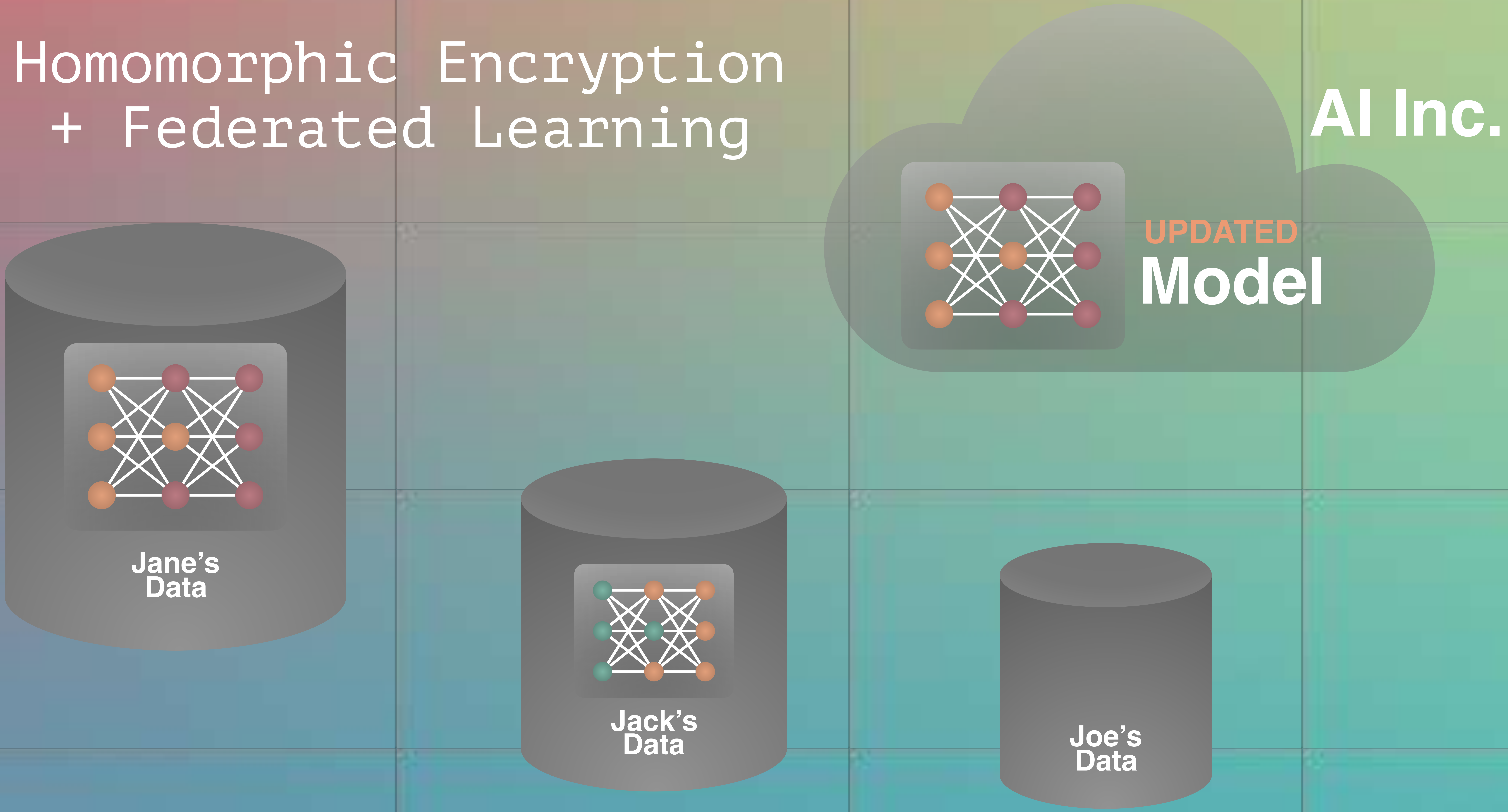
Homomorphic Encryption + Federated Learning



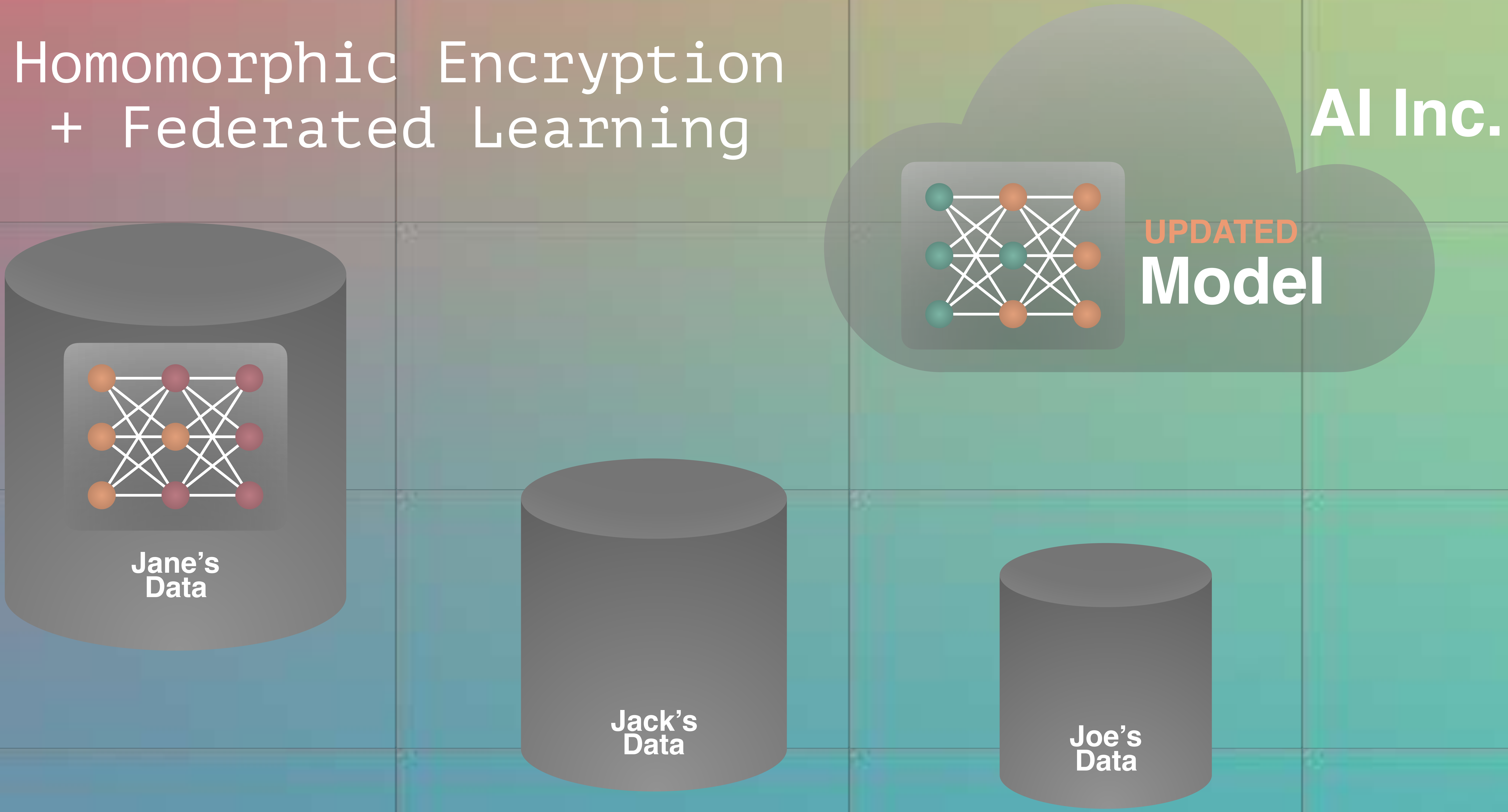
Homomorphic Encryption + Federated Learning



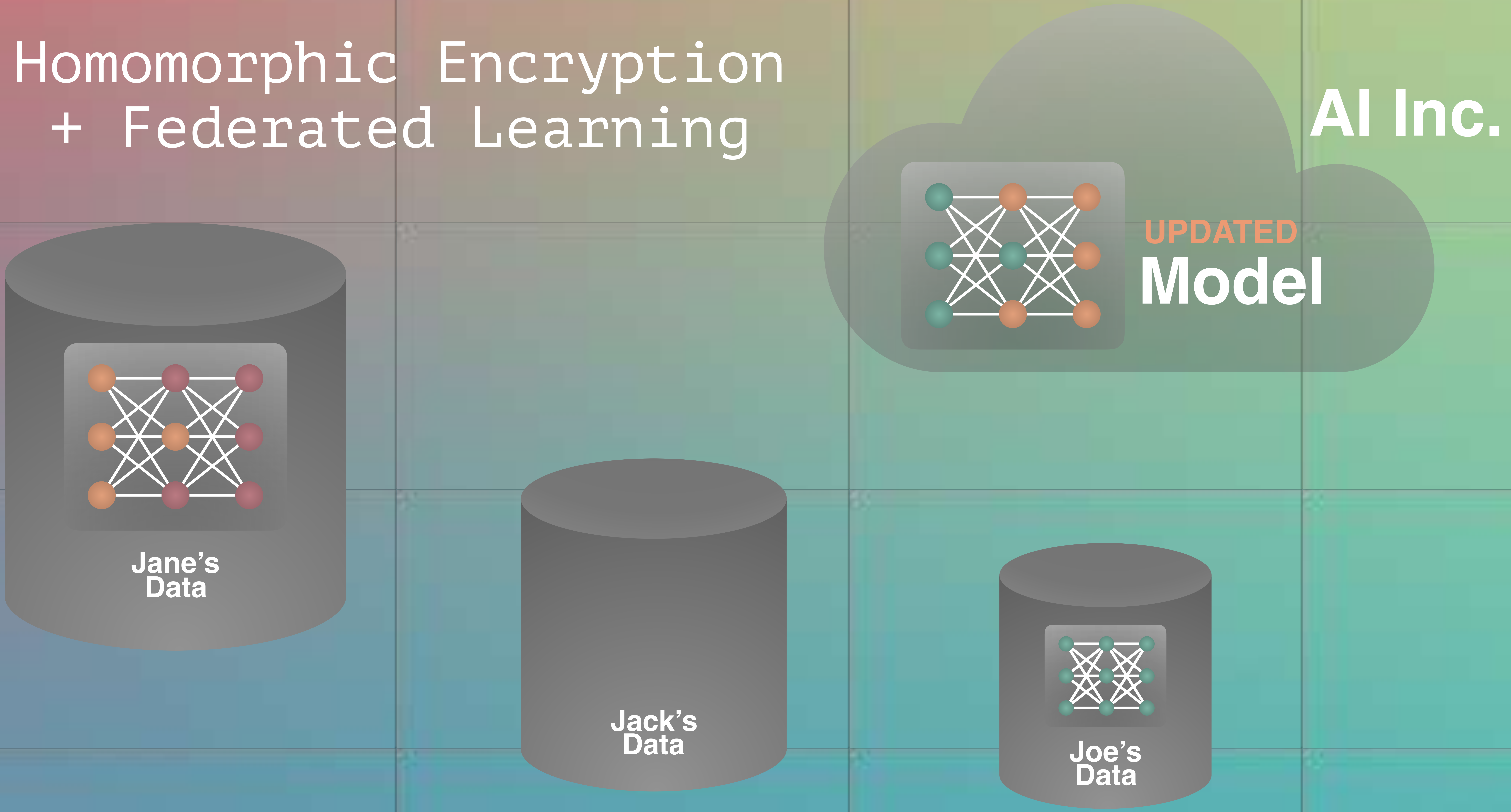
Homomorphic Encryption + Federated Learning



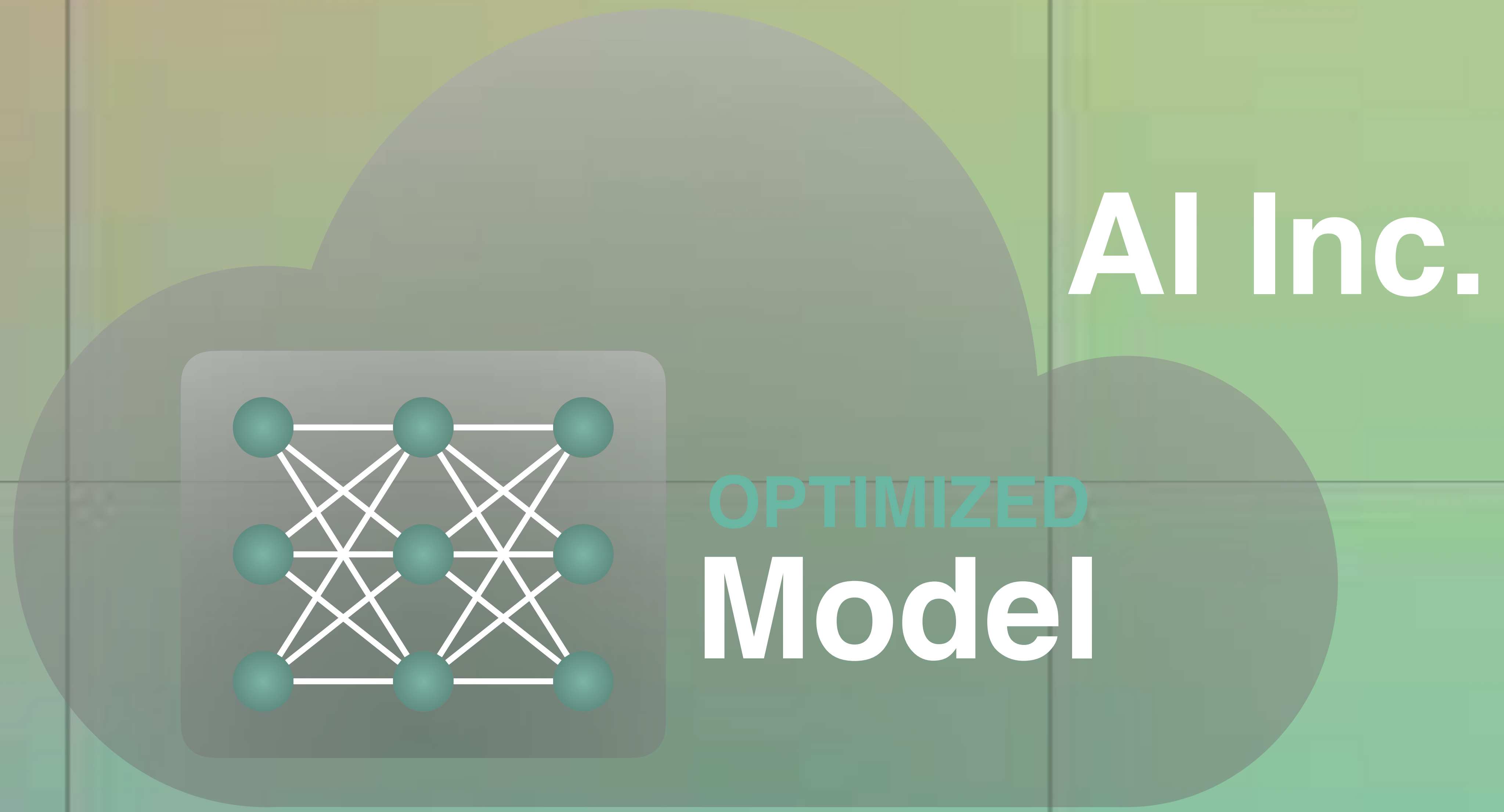
Homomorphic Encryption + Federated Learning



Homomorphic Encryption + Federated Learning



Homomorphic Encryption + Federated Learning



Potential Solution

- **Train** A.I. on data we cannot see without revealing the AI or its training updates to anyone?
  - Homomorphic Encryption
  - Multi-Party Computation

Introduction to

Multi-Party Computation

for Safe AI

Multi-Party Computation

a = 5 → Share Splitter → [1 , -3 , 5 , 0 , 2] = shares_a

Multi-Party Computation

a = 5 → Share Splitter → [1 , -3 , 5 , 0 , 2] = shares_a

b = 3 → Share Splitter → [2 , -5 , 8 , -3 , 1] = shares_b

Multi-Party Computation

a = 5 → Share Splitter → [1 , -3 , 5 , 0 , 2] = shares_a

b = 3 → Share Splitter → [2 , -5 , 8 , -3 , 1] = shares_b

Person 1	Person 2	Person 3	Person 4	Person 5
s_a = 1	s_a = -3	s_a = 5	s_a = 0	s_a = 2
s_b = 2	s_b = -5	s_b = 8	s_b = -3	s_b = 1
s_c = 3	s_c = -8	s_c = 13	s_c = -3	s_c = 3

Multi-Party Computation

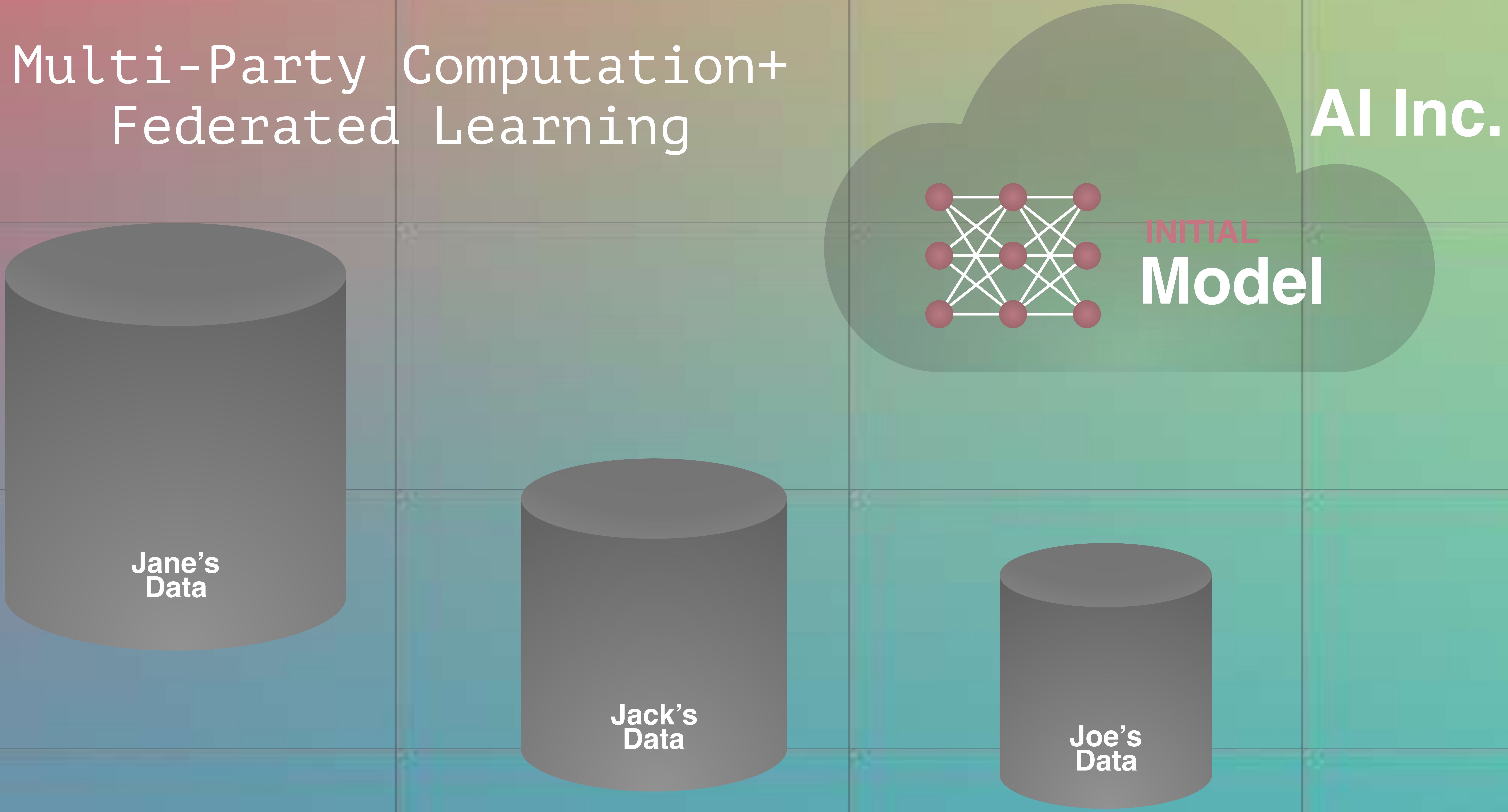
a = 5 → Share Splitter → [1 , -3 , 5 , 0 , 2] = shares_a

b = 3 → Share Splitter → [2 , -5 , 8 , -3 , 1] = shares_b

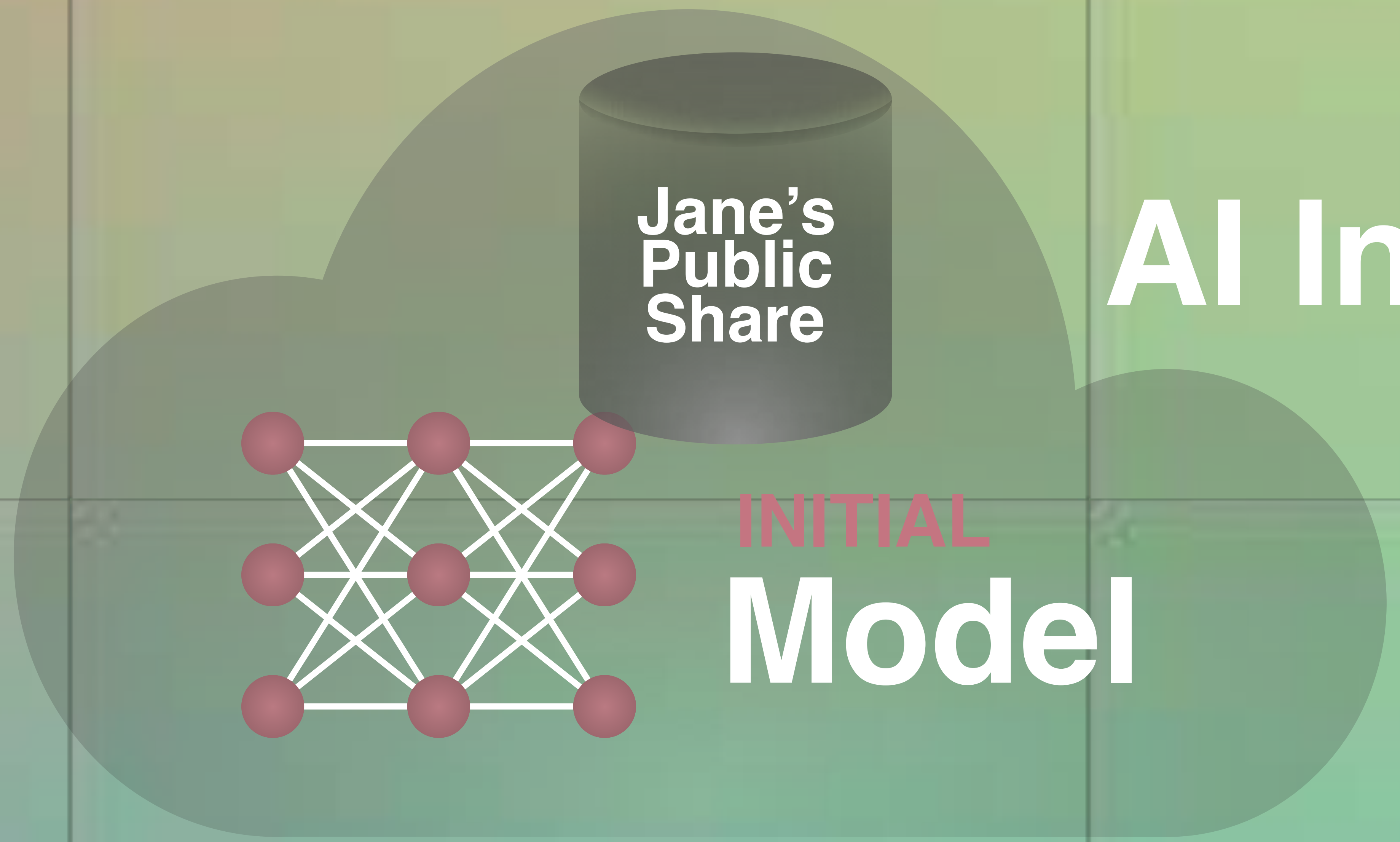
Person 1	Person 2	Person 3	Person 4	Person 5
s_a = 1	s_a = -3	s_a = 5	s_a = 0	s_a = 2
s_b = 2	s_b = -5	s_b = 8	s_b = -3	s_b = 1
s_c = 3	s_c = -8	s_c = 13	s_c = -3	s_c = 3

shares_c = [3 , -8 , 13 , -3 , 3] → Share Combiner → 8

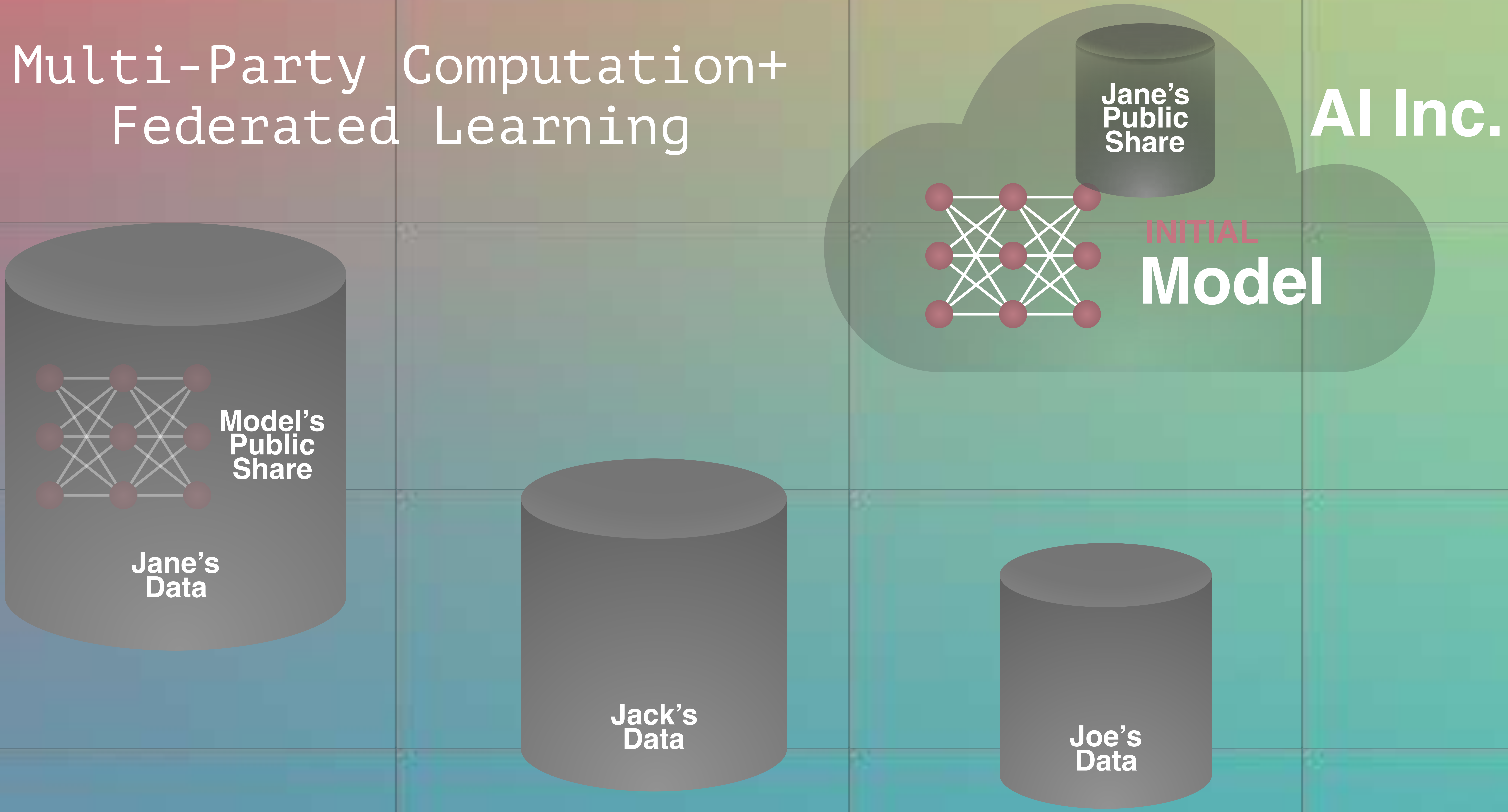
Multi-Party Computation+ Federated Learning



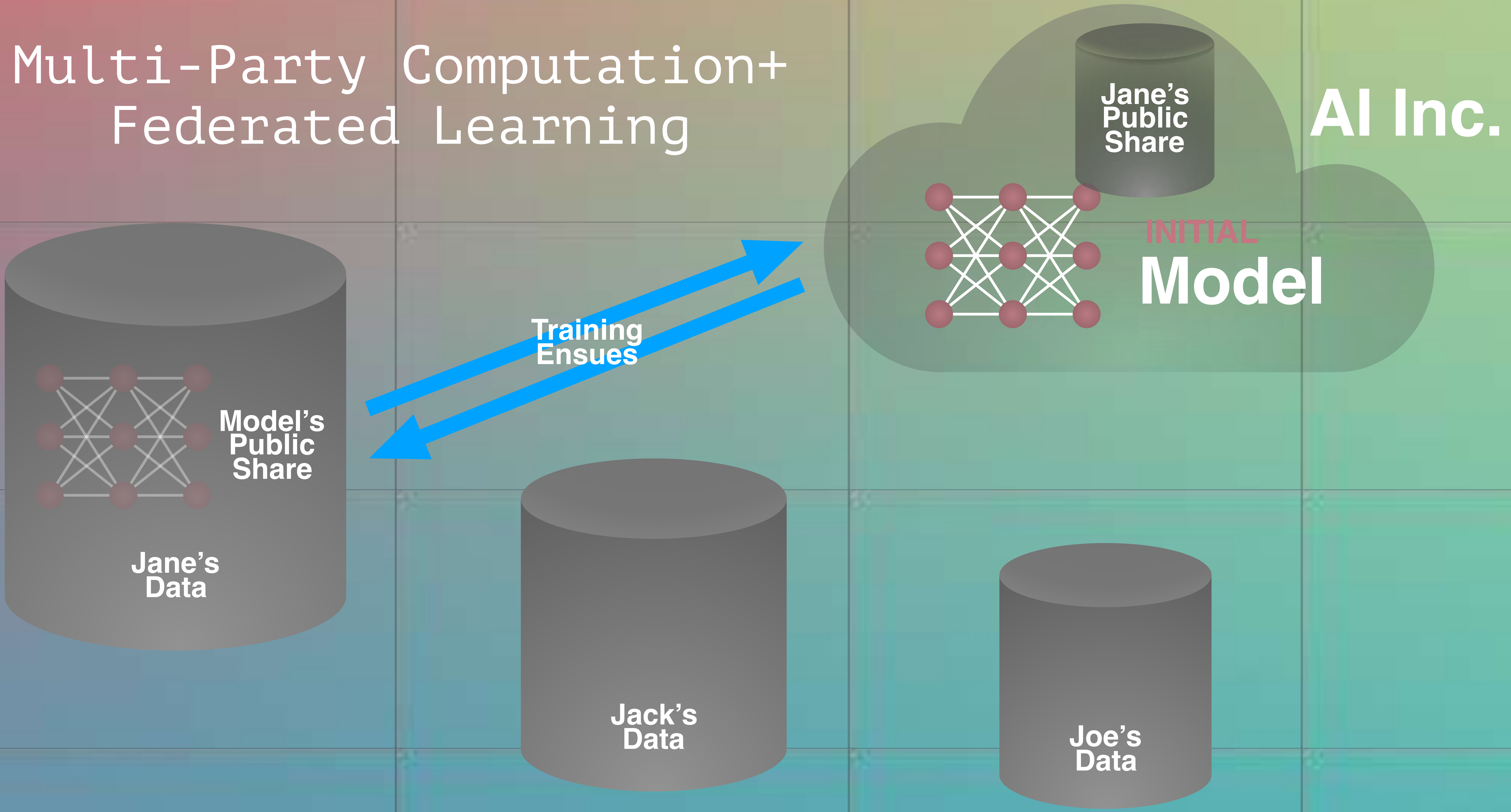
Multi-Party Computation+ Federated Learning



Multi-Party Computation+ Federated Learning



Multi-Party Computation+ Federated Learning



Potential Solution

- **Train** A.I. on data we cannot see without revealing the AI or its training updates to anyone?
  - Homomorphic Encryption
  - Multi-Party Computation

Potential Solution

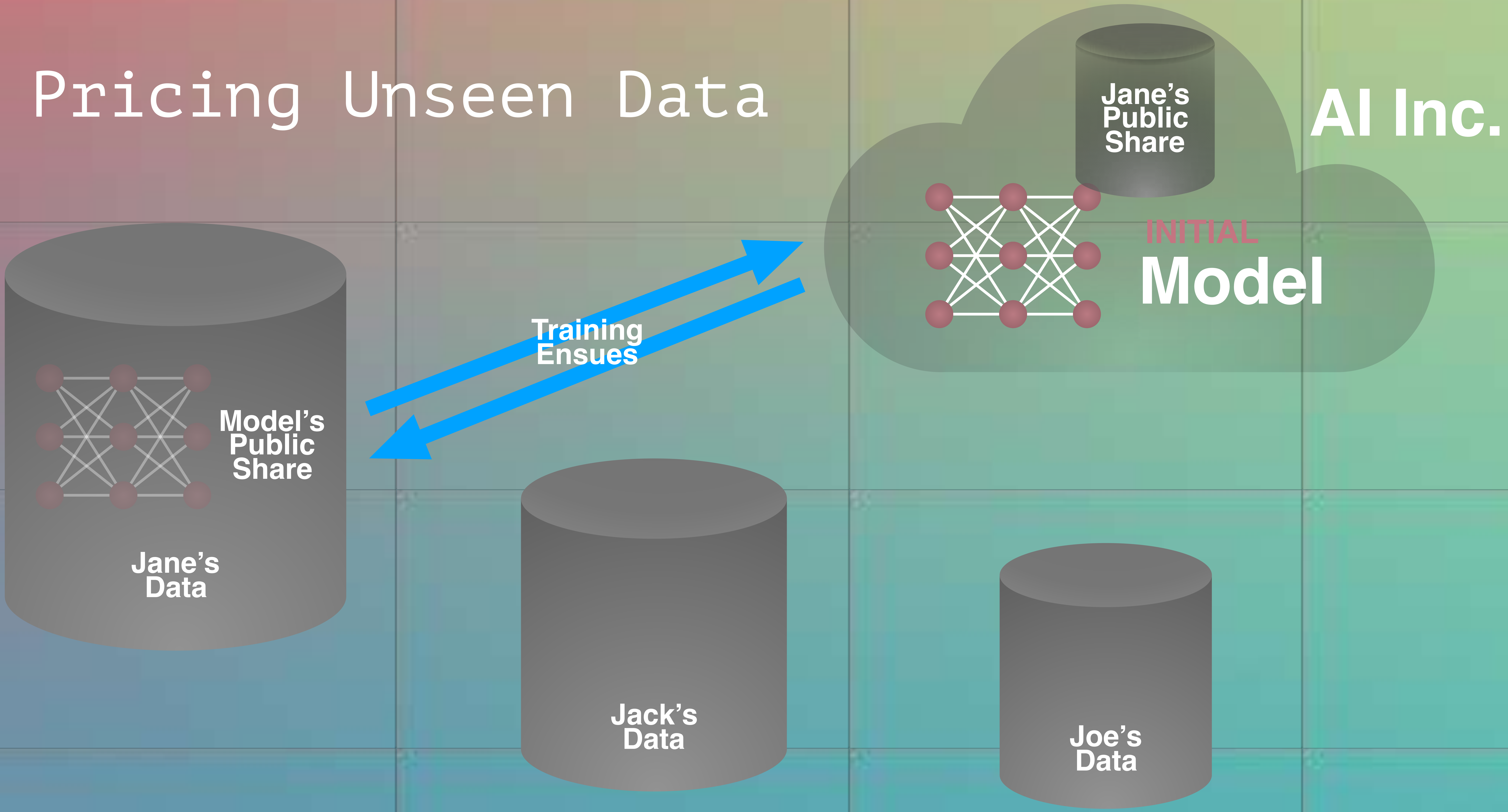
- **Train** AI on data we cannot see without revealing that AI or its training gradients to anyone (FL + HE + MPC)
- **Share** the ownership of a trained AI such that its usefulness is public while its input data, contents, and output predictions are secret - even to the owners of the AI. (MPC)
- **Price** training data we cannot see competitively with other data which we also cannot not see (GT + STAKING + SC + PNP)

Introduction to

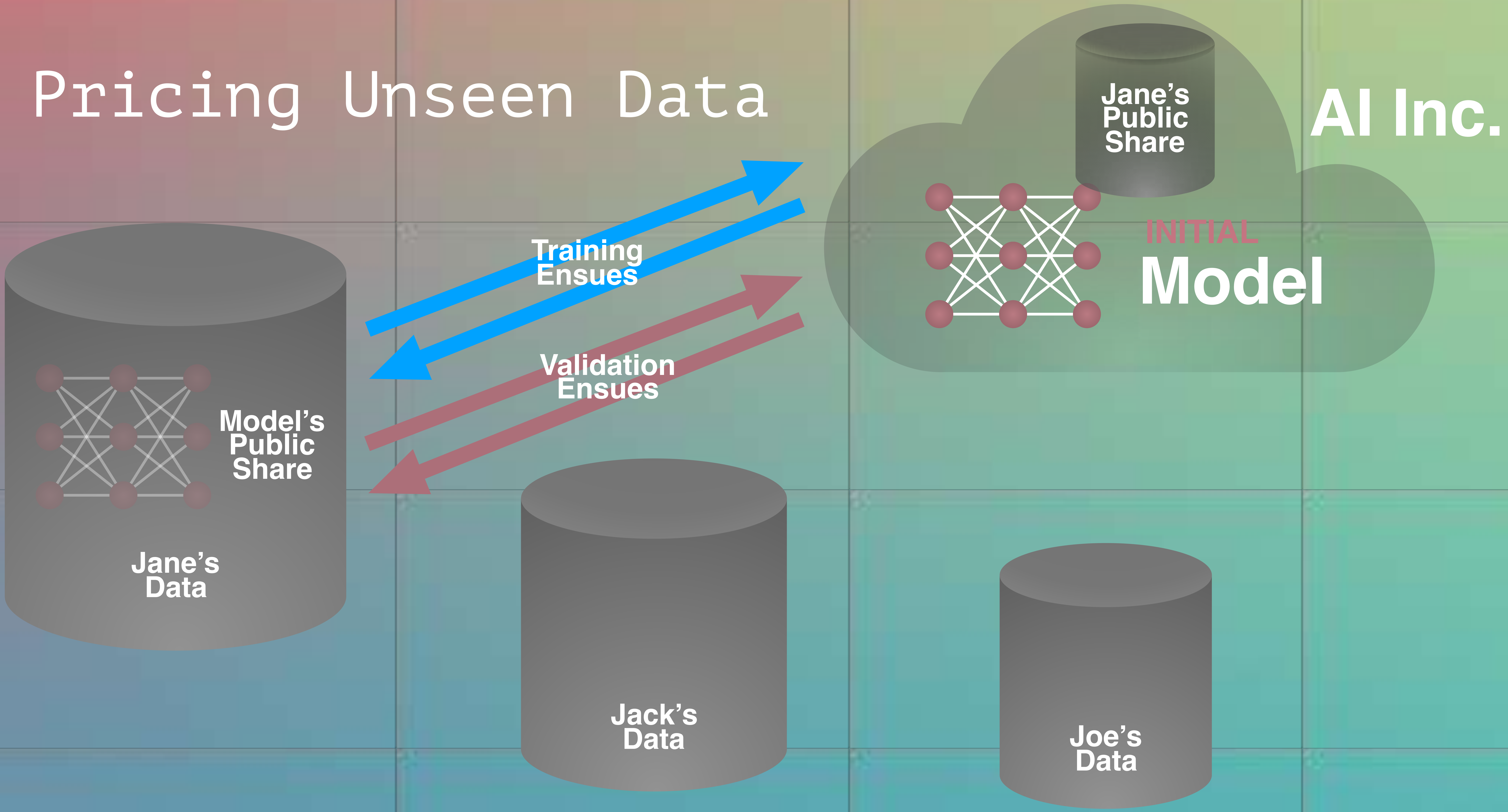
Gradient Marketplaces

for Safe AI

Pricing Unseen Data



Pricing Unseen Data



Outline

- **Why:** the AI Business Model has privacy problems
- **How:** an Introduction to the Core Technologies of OpenMined
  - Federated Learning
  - Homomorphic Encryption
  - Multi-Party Computation
  - Gradient Marketplace
- **Roadmap & Demos**

Status Update

- **January Hackathon:** 29 Cities - 400+ online - 70+ In Person
- **Growth Stats:** 2400 Members in Slack - 145 GitHub Committers
- Recent Milestones:
    - **PyTorch over Peer-to-Peer** - the foundation of Secure/Private AI
    - **OpenMined Grid** - UK model trained in Canada in 15 secs - *Jan 24*
    - **Reinforcement Learning** - worked with Unity's "ML Agents" Team

Project Roadmap

- Federated Learning via PyTorch
- MPC Training via PyTorch
- Rapid Grid Payment via Coinbase
- Private Application Integration

Q&A

Demo

Appendix

Introduction to

Functional Encryption

for Safe AI

Functional Encryption

a = [1 , 2 , 3 , 4 , 5] → Functional Encryptor → cypher_a

Functional Encryption

a = [1 , 2 , 3 , 4 , 5] → Functional Encryptor → cypher_a

b = [0 , 1 , 0 , 0 , 0] → Functional Encryptor → cypher_b

Functional Encryption

a = [1 , 2 , 3 , 4 , 5] → Functional Encryptor → cypher_a

b = [0 , 1 , 0 , 0 , 0] → Functional Encryptor → cypher_b

cypher_a →
cypher_b →
Encrypted Dot Product → 2