# Learning with Feedback Loops

DJ Thornton

June 2024

Many learning processes have the potential problem that realizations of the data affect the data generating process itself.

Many learning processes have the potential problem that realizations of the data affect the data generating process itself.

Many learning processes have the potential problem that realizations of the data affect the data generating process itself.


The American Economic Review

ARTICLES

Many learning processes have the potential problem that realizations of the data affect the data generating process itself.

The American Economic Review

ARTICLES

MARIAGIOVANNA BACCARA, AYŞE İMROHOROĞLU, ALISTAIR J. WILSON, AND LEEAT YARIV
A Field Study on Matching with Network Externalities

GADI BARLEVY AND DEREK NEAL
Pay for Percentile

RAN ABRAMITZKY, LEAH PLATT BOUSTAN, AND KATHERINE ERIKSSON
Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration

MATTHEW O. JACKSON, TOMAS RODRIGUEZ-BARRAQUER, AND XU TAN
Social Capital and Social Quilts: Network Patterns of Favor Exchange

PATRICK BAJARI, JANE COOLEY FRUEHWIRTH, KYOO IL KIM, AND CHRISTOPHER TIMMINS
A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution

GORDON B. DAHL AND LANCE LOCHNER
The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit

DOH-SHIN JEON AND DOMENICO MENICUCCI
Bundling and Competition for Slots

VERONICA GUERRIERI AND PÉTER KONDOR
Fund Managers, Career Concerns, and Asset Price Volatility

JUDD B. KESSLER AND ALVIN E. ROTH
Organ Allocation Policy and the Decision to Donate

SCOTT A. IMBERMAN, ADRIANA D. KUGLER, AND BRUCE I. SACERDOTE
Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees

DANIEL J. BENJAMIN, ORI HEFFETZ, MILES S. KIMBALL, AND ALEX REES-JONES
What Do You Think Would Make You Happier? What Do You Think You Would Choose?

KEYU JIN
Industrial Structure and Capital Flows

# Model

# Model

## Sketch

# Model

## Sketch

- Start with an exogenous signal about an unknown state.

# Model

## Sketch

- Start with an exogenous signal about an unknown state.
- Agents learn something about the state and then act $\Rightarrow$ agents' actions create *data*.

# Model

## Sketch

- Start with an exogenous signal about an unknown state.
- Agents learn something about the state and then act $\Rightarrow$ agents' actions create *data*.
    - Agents either learn by experimenting (doing "independent research").

# Model

## Sketch

- Start with an exogenous signal about an unknown state.
- Agents learn something about the state and then act $\Rightarrow$ agents' actions create *data*.
    - Agents either learn by experimenting (doing "independent research").
    - **OR** by asking an *information aggregator* for a recommendation on how to act.

# Model

## Sketch

- Start with an exogenous signal about an unknown state.
- Agents learn something about the state and then act $\Rightarrow$ agents' actions create *data*.
  - Agents either learn by experimenting (doing "independent research").
  - **OR** by asking an *information aggregator* for a recommendation on how to act.
- Long-lived *Information aggregator* wants to learn the state: samples the data.

# Model

## Sketch

- Start with an exogenous signal about an unknown state.
- Agents learn something about the state and then act ⇒ agents' actions create *data*.
  - Agents either learn by experimenting (doing "independent research").
  - **OR** by asking an *information aggregator* for a recommendation on how to act.
- Long-lived *Information aggregator* wants to learn the state: samples the data.
  - BUT, if agents act on the aggregator's recommendation, then the data generated by these actions are uninformative to the aggregator!

# Model

## Simple 3-Period Baseline

# Model

## Simple 3-Period Baseline

- There is a binary state of the world $\theta \in \{\theta_0, \theta_1\}$ .

# Model

## Simple 3-Period Baseline

- There is a binary state of the world $\theta \in \{\theta_0, \theta_1\}$ .
- Time is discrete and indexed by $t \in \{0,1,2\}$ .

# Model

## Simple 3-Period Baseline

- There is a binary state of the world $\theta \in \{\theta_0, \theta_1\}$ .
- Time is discrete and indexed by $t \in \{0,1,2\}$ .
- In each period there is a unit mass of short lived agents $i_t \in [0,1]$ .

# Model

## Simple 3-Period Baseline

- There is a binary state of the world $\theta \in \{\theta_0, \theta_1\}$ .
- Time is discrete and indexed by $t \in \{0,1,2\}$ .
- In each period there is a unit mass of short lived agents $i_t \in [0,1]$ .
    - Agent $i$ at time $t$ chooses a binary action $a_{it} \in \{0,1\}$ .

# Model

## Simple 3-Period Baseline

- There is a binary state of the world $\theta \in \{\theta_0, \theta_1\}$ .
- Time is discrete and indexed by $t \in \{0,1,2\}$ .
- In each period there is a unit mass of short lived agents $i_t \in [0,1]$ .
  - Agent $i$ at time $t$ chooses a binary action $a_{it} \in \{0,1\}$ .
  - Agents have a common prior $\mu \equiv \mathbb{P}(\theta = \theta_0) \in (0,1)$ .

# Model

## Simple 3-Period Baseline

- There is a binary state of the world $\theta \in \{\theta_0, \theta_1\}$.
- Time is discrete and indexed by $t \in \{0,1,2\}$.
- In each period there is a unit mass of short lived agents $i_t \in [0,1]$.
    - Agent $i$ at time $t$ chooses a binary action $a_{it} \in \{0,1\}$.
    - Agents have a common prior $\mu \equiv \mathbb{P}(\theta = \theta_0) \in (0,1)$.
- Payoffs are $1$ if the action matches the state, and $0$ otherwise.

# Model

Timing: $t = 0$

# Model

## Timing: $t = 0$

- The initial population each draw an informative signal (a recommendation $a \in \{0,1\}$) summarized by:

# Model

## Timing: $t = 0$

- The initial population each draw an informative signal (a recommendation $a \in \{0,1\}$) summarized by:

$$\pi \equiv \mathbb{P}(a = 1 \mid \theta_1) = \mathbb{P}(a = 0 \mid \theta_0) > \frac{1}{2}.$$

# Model

## Timing: $t = 0$

- The initial population each draw an informative signal (a recommendation $a \in \{0,1\}$) summarized by:

$$\pi \equiv \mathbb{P}(a = 1 \mid \theta_1) = \mathbb{P}(a = 0 \mid \theta_0) > \frac{1}{2}.$$

- Each member of the initial population chooses an action given the recommendation, and these actions form the *initial database*.

# Model

## Timing: $t = 0$

- The initial population each draw an informative signal (a recommendation $a \in \{0,1\}$) summarized by:

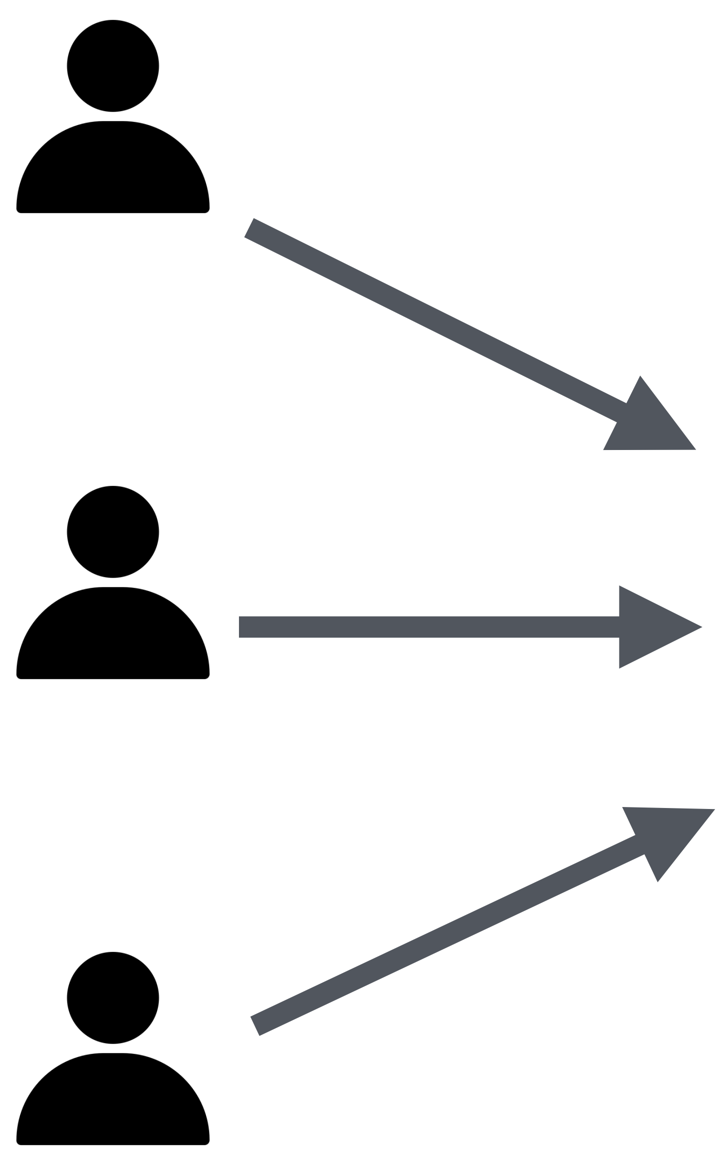
$$\pi \equiv \mathbb{P}(a = 1 \mid \theta_1) = \mathbb{P}(a = 0 \mid \theta_0) > \frac{1}{2}.$$

- Each member of the initial population chooses an action given the recommendation, and these actions form the *initial database*.

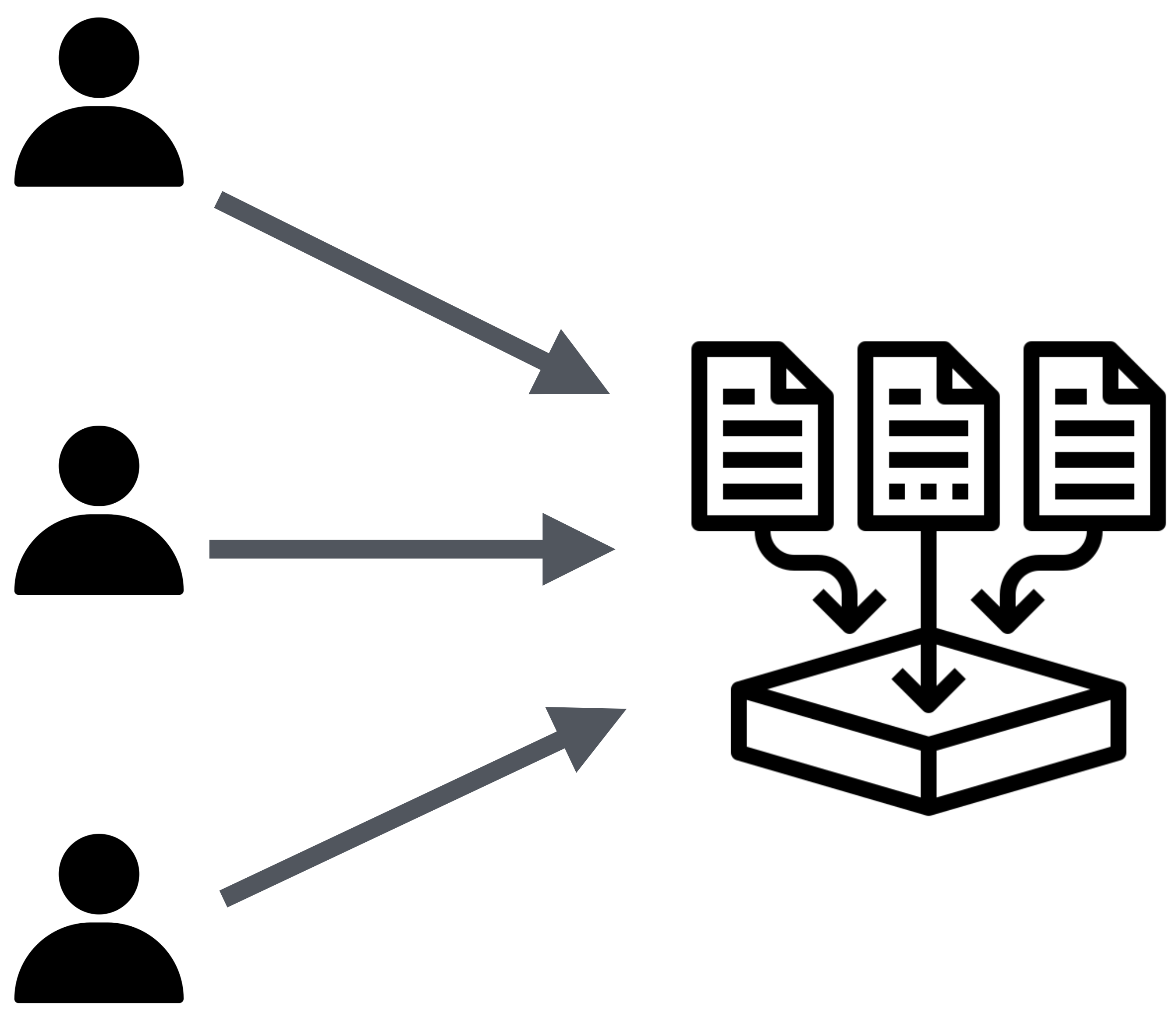
- Assume priors are moderate enough that agents follow the recommendation they receive. This also ensures they will follow the information aggregator's recommendations.

$t = 0$

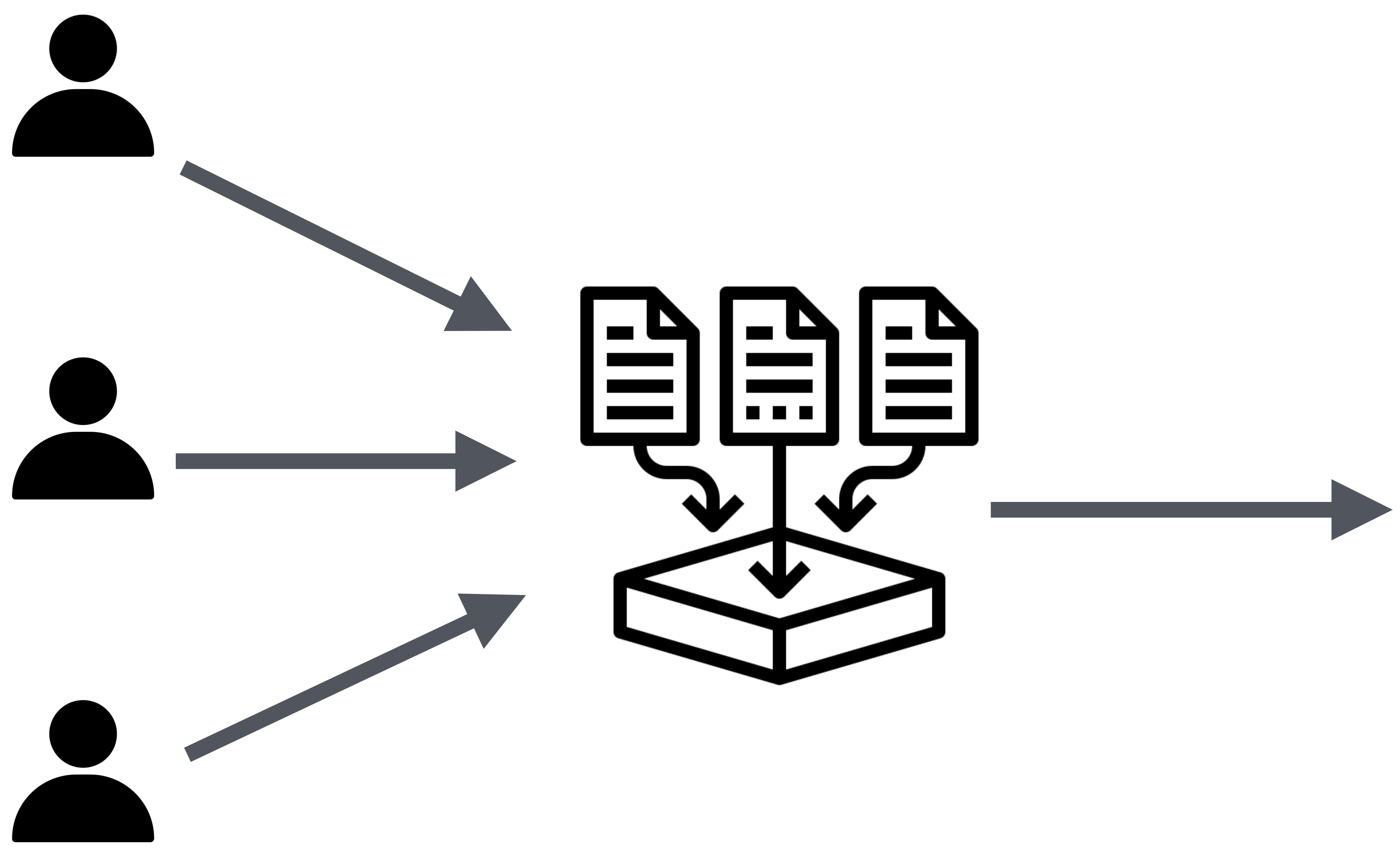
$t = 0$



$t = 0$



$t = 0$ $t = 1$



# Model

Timing: $t = 1$

# Model

## Timing: $t = 1$

- A long-lived information aggregator is born. (We will return to the aggregator’s utility function!)

# Model

## Timing: $t = 1$

- A long-lived information aggregator is born. (We will return to the aggregator's utility function!)
    - Takes a sample of $k \geq 1$ draws from the database.

# Model

## Timing: $t = 1$

- A long-lived information aggregator is born. (We will return to the aggregator's utility function!)
  - Takes a sample of $k \geq 1$ draws from the database.
  - Chooses whether or not it will offer a *recommendation* to agents.

# Model

## Timing: $t = 1$

- A long-lived information aggregator is born. (We will return to the aggregator's utility function!)
  - Takes a sample of $k \geq 1$ draws from the database.
  - Chooses whether or not it will offer a *recommendation* to agents.
- A new population of short-lived agents are born.

# Model

## Timing: $t = 1$

- A long-lived information aggregator is born. (We will return to the aggregator's utility function!)
    - Takes a sample of $k \geq 1$ draws from the database.
    - Chooses whether or not it will offer a *recommendation* to agents.
- A new population of short-lived agents are born.
    - (Exogenous) proportion $q$ ask the aggregator for a recommendation if one is available.

# Model

## Timing: $t = 1$

- A long-lived information aggregator is born. (We will return to the aggregator's utility function!)
    - Takes a sample of $k \geq 1$ draws from the database.
    - Chooses whether or not it will offer a *recommendation* to agents.
- A new population of short-lived agents are born.
    - (Exogenous) proportion $q$ ask the aggregator for a recommendation if one is available.
    - Proportion $1 - q$ take a single draw from the initial database.

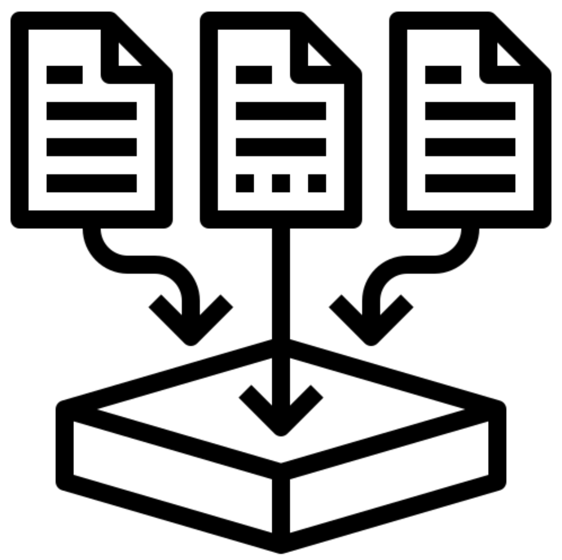
# Model

## Timing: $t = 1$

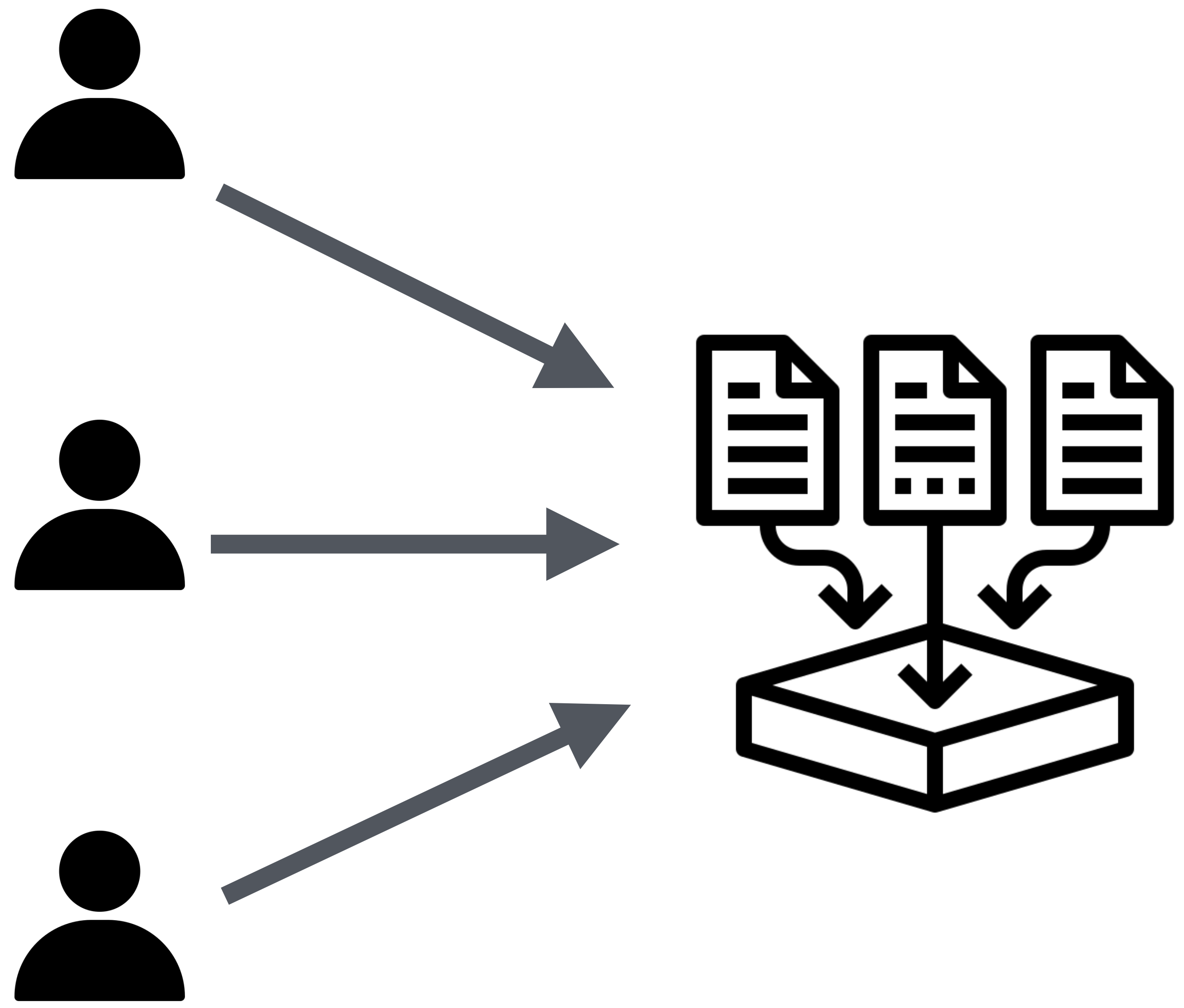
- A long-lived information aggregator is born. (We will return to the aggregator's utility function!)
  - Takes a sample of $k \geq 1$ draws from the database.
  - Chooses whether or not it will offer a *recommendation* to agents.
- A new population of short-lived agents are born.
  - (Exogenous) proportion $q$ ask the aggregator for a recommendation if one is available.
  - Proportion $1 - q$ take a single draw from the initial database.
- Agents act and actions are added to the database to create the *interim database*.

$t = 0$ $t = 1$

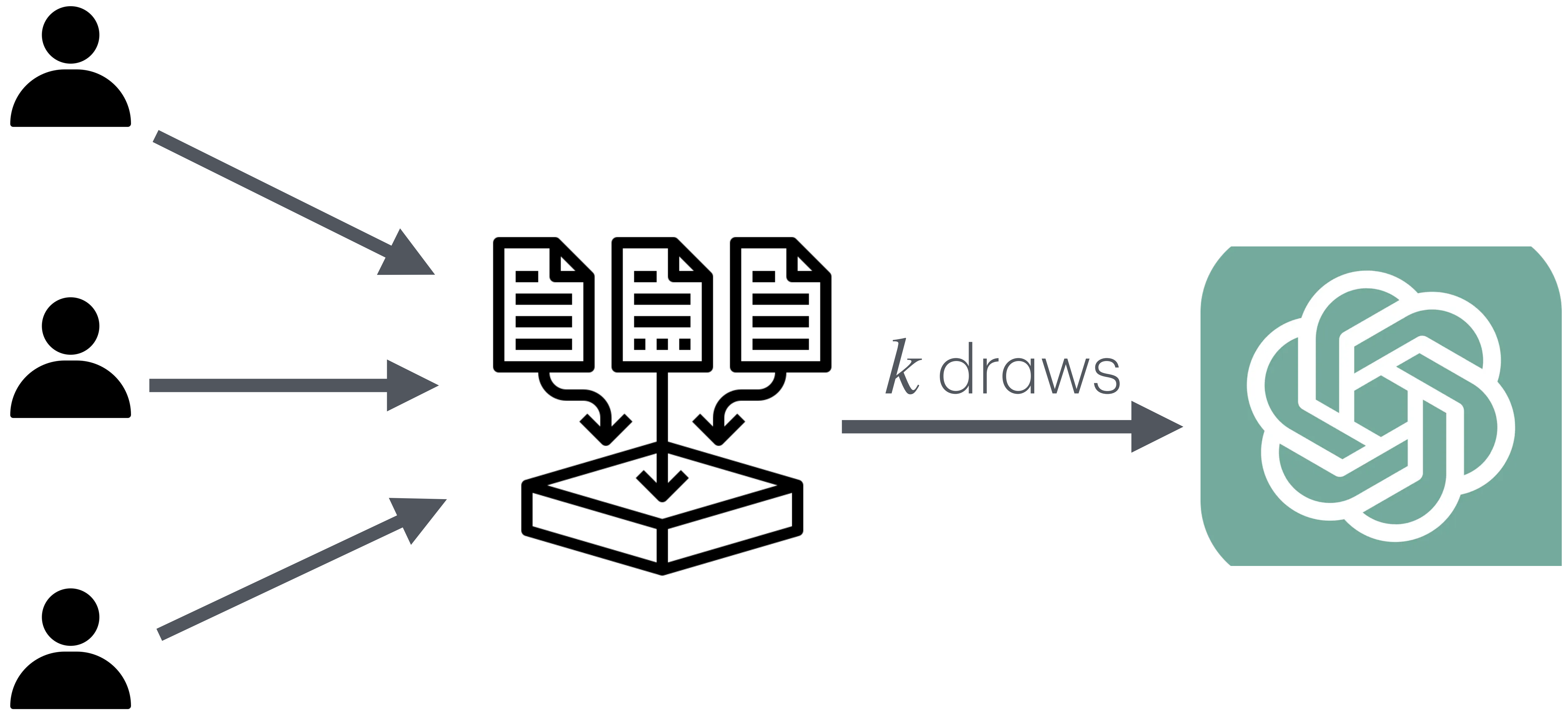
$t = 0$ $t = 1$

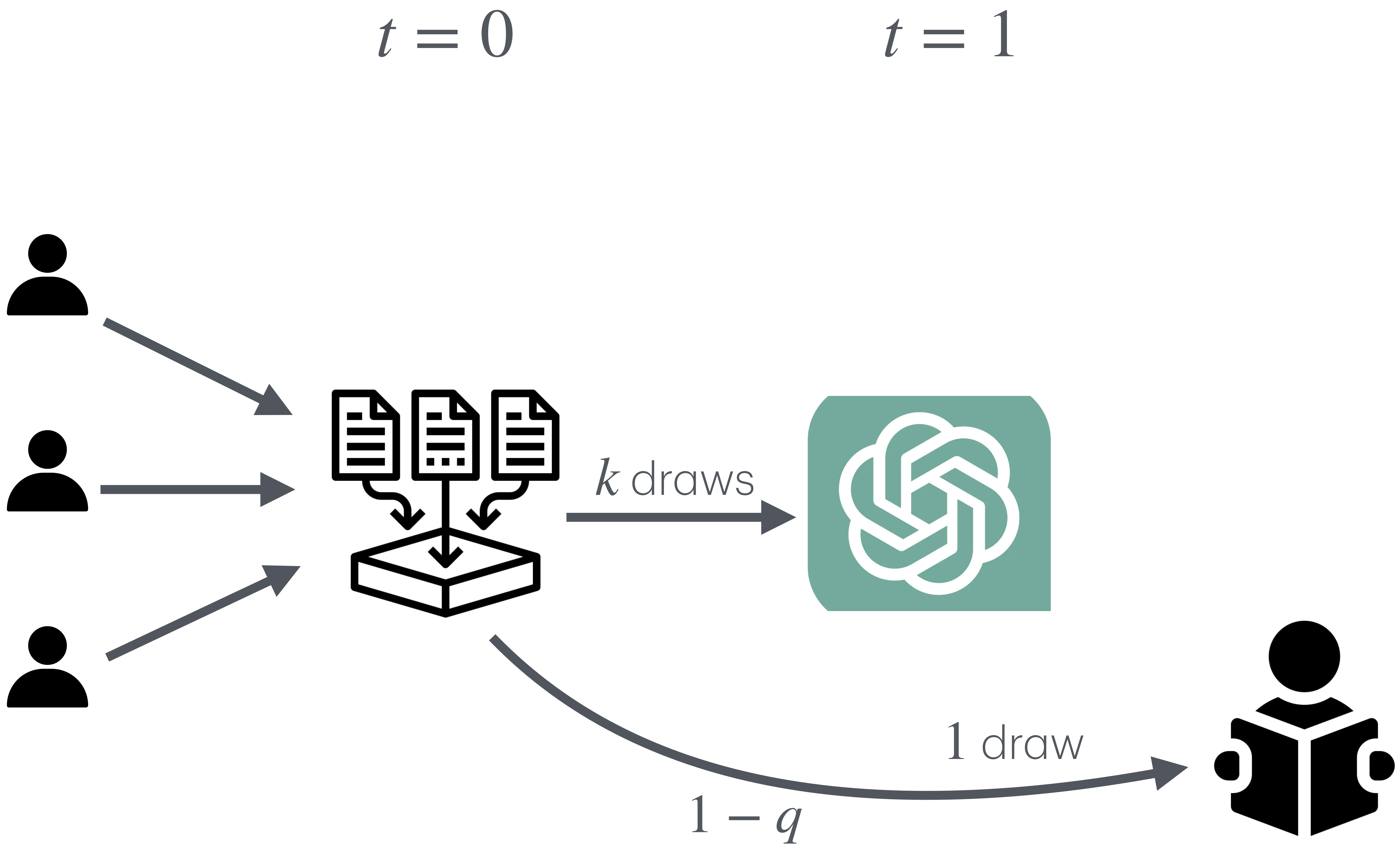


$t = 0$ $t = 1$

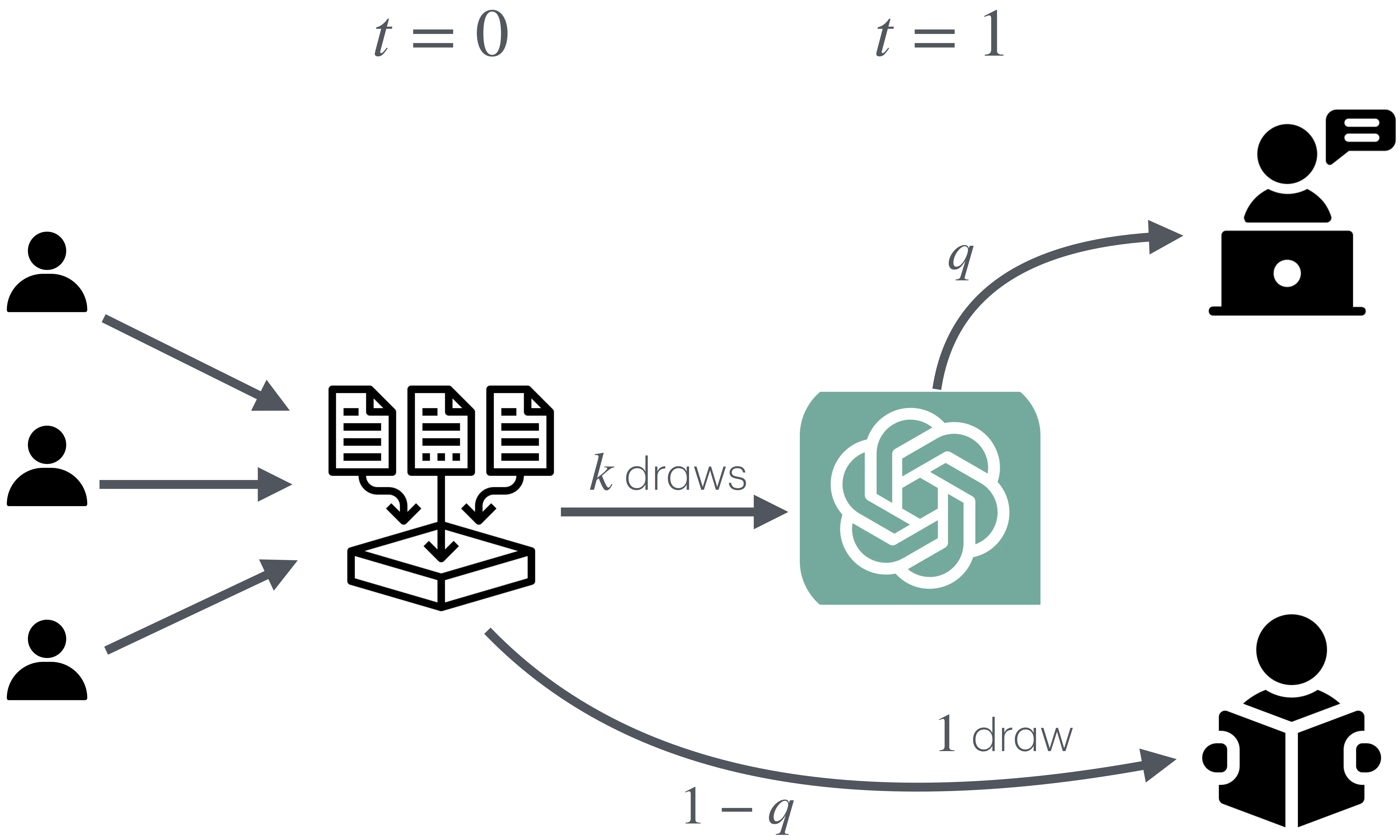


$t = 0$ $t = 1$

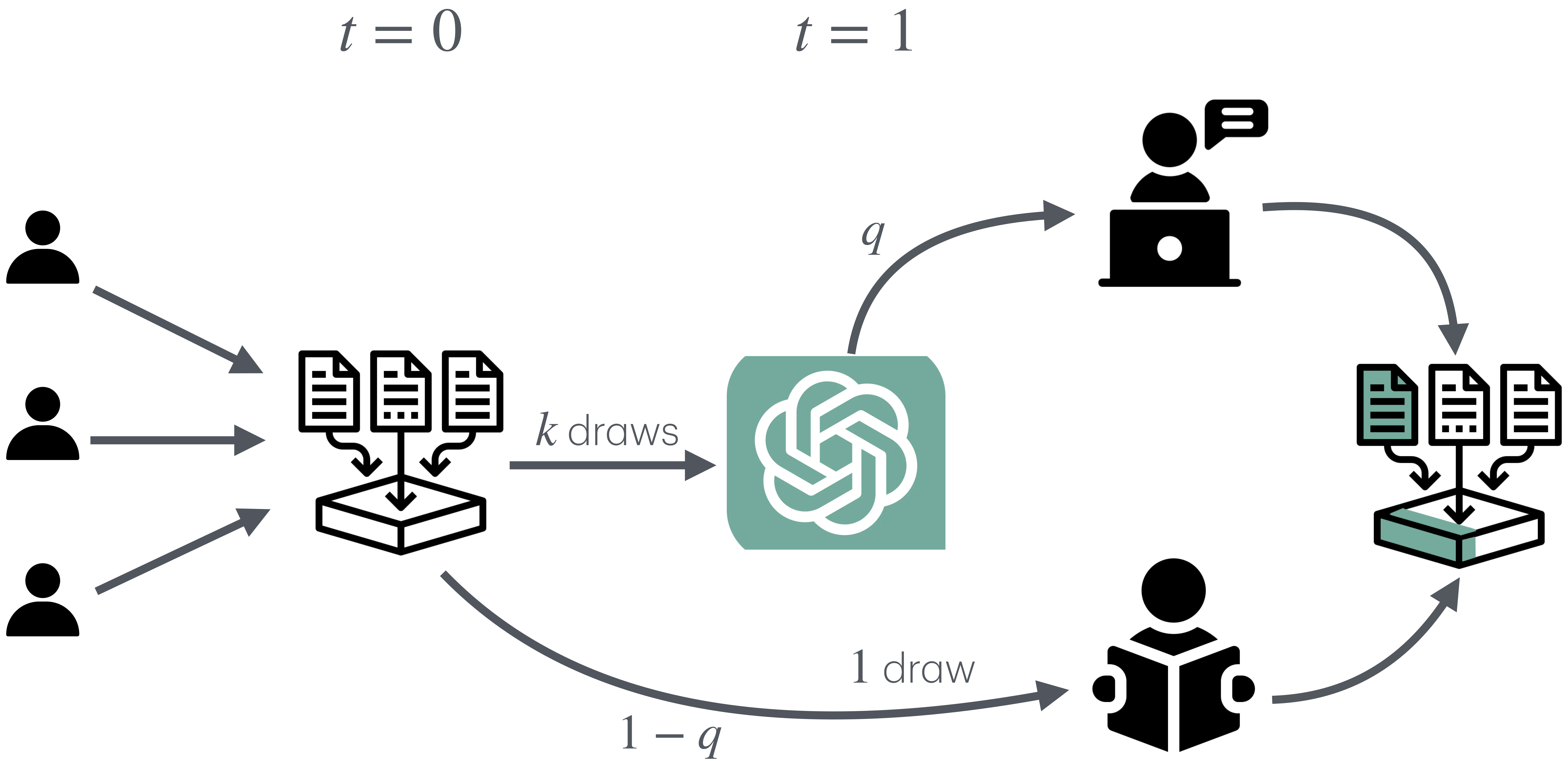




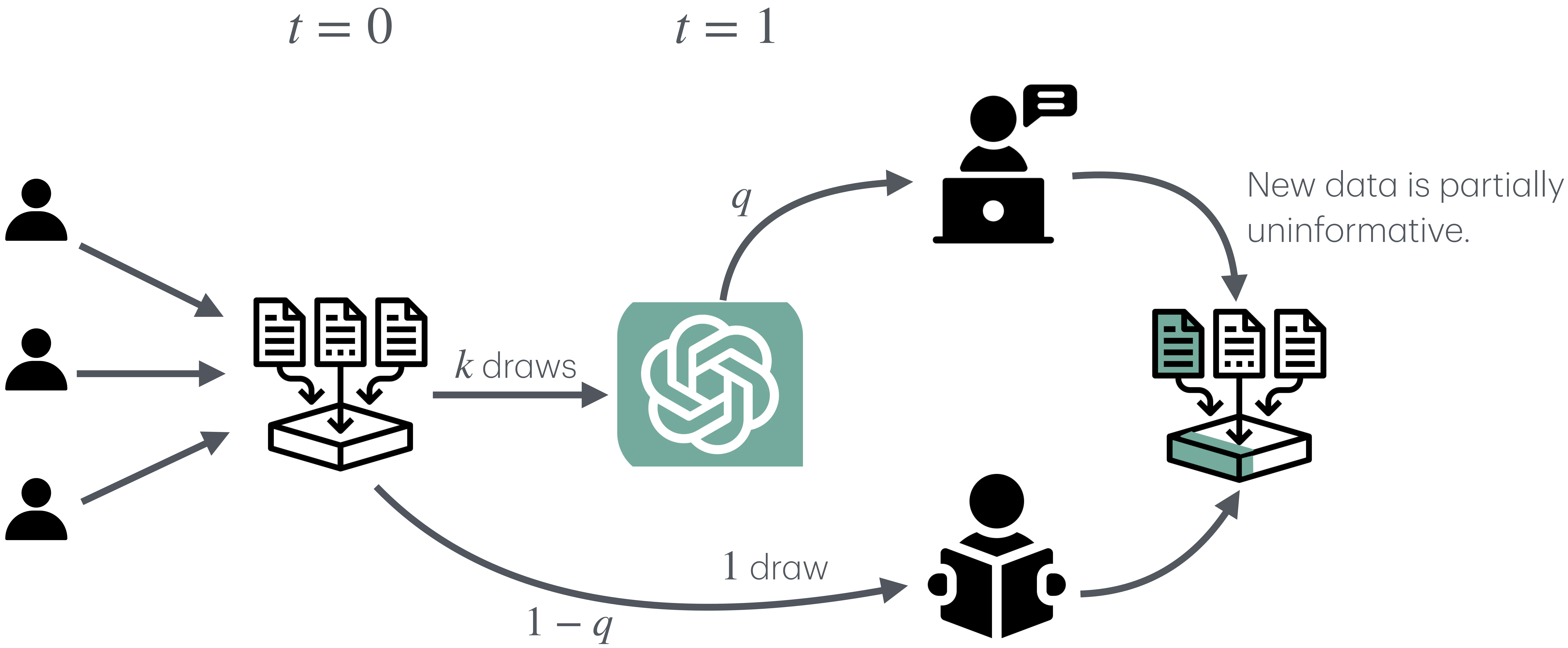
$t = 0$
$t = 1$
$k$ draws
1 draw
$1 - q$



$t = 0$
$t = 1$
$q$
$k$ draws
1 draw
$1 - q$



$t = 0$
$t = 1$
$q$
$k$ draws
1 draw
$1 - q$



$t = 0$
$t = 1$
$q$
New data is partially uninformative.
$k$ draws
1 draw
$1 - q$

# Model

Timing: $t = 2$

# Model

## Timing: $t = 2$

- Aggregator takes a sample of $k \geq 1$ draws from the interim database.

# Model

## Timing: $t = 2$

- Aggregator takes a sample of $k \geq 1$ draws from the interim database.
  - Chooses whether it will offer a recommendation.

# Model

## Timing: $t = 2$

- Aggregator takes a sample of $k \geq 1$ draws from the interim database.
  - Chooses whether it will offer a recommendation.
- A new population of short-lived agents are born.

# Model

## Timing: $t = 2$

- Aggregator takes a sample of $k \geq 1$ draws from the interim database.
  - Chooses whether it will offer a recommendation.
- A new population of short-lived agents are born.
  - Proportion $q$ ask the aggregator for a recommendation.

# Model

## Timing: $t = 2$

- Aggregator takes a sample of $k \geq 1$ draws from the interim database.
  - Chooses whether it will offer a recommendation.
- A new population of short-lived agents are born.
  - Proportion $q$ ask the aggregator for a recommendation.
  - Proportion $1 - q$ take a single draw from the interim database.

# Model

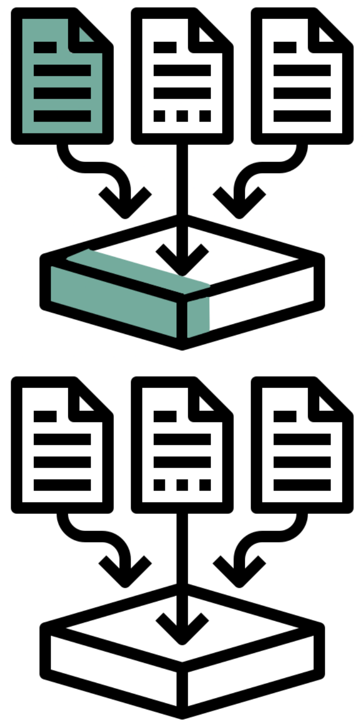
## Timing: $t = 2$

- Aggregator takes a sample of $k \geq 1$ draws from the interim database.
  - Chooses whether it will offer a recommendation.
- A new population of short-lived agents are born.
  - Proportion $q$ ask the aggregator for a recommendation.
  - Proportion $1 - q$ take a single draw from the interim database.
- Agents act and actions are added to the database to create the *posterior database*.

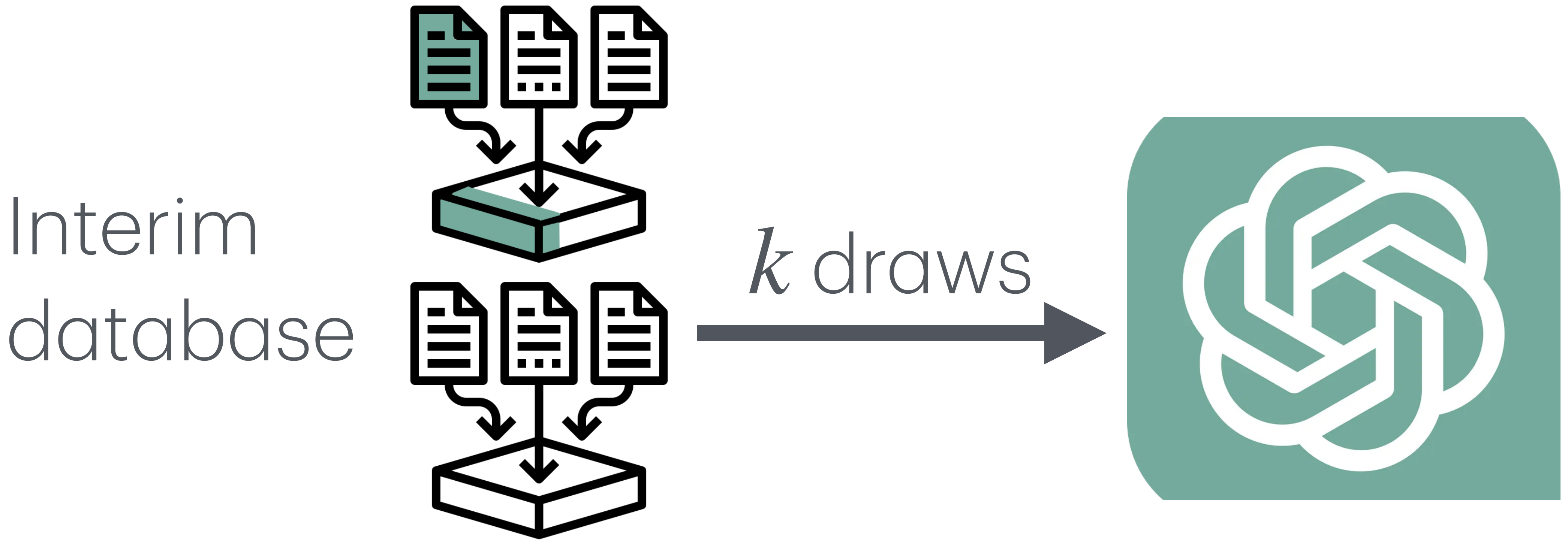
$$t = 2$$

$t = 2$

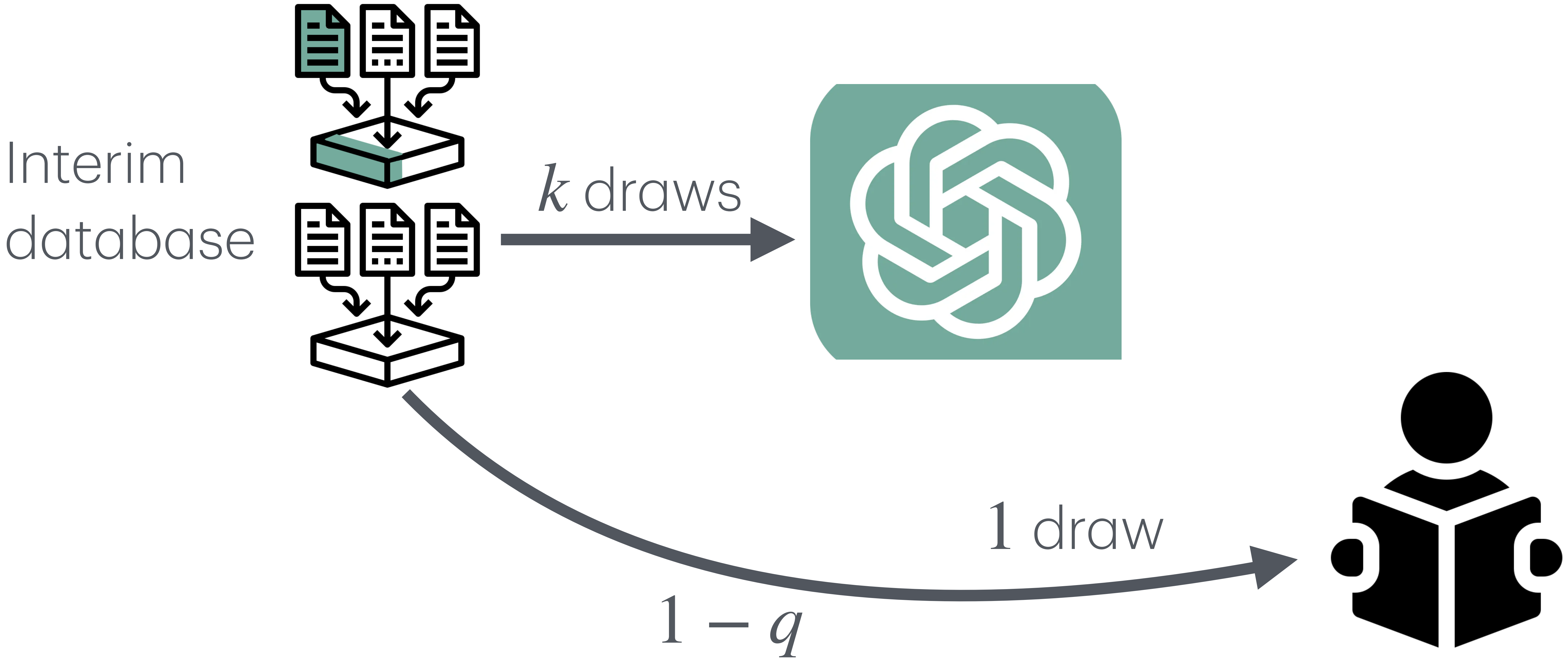
Interim database

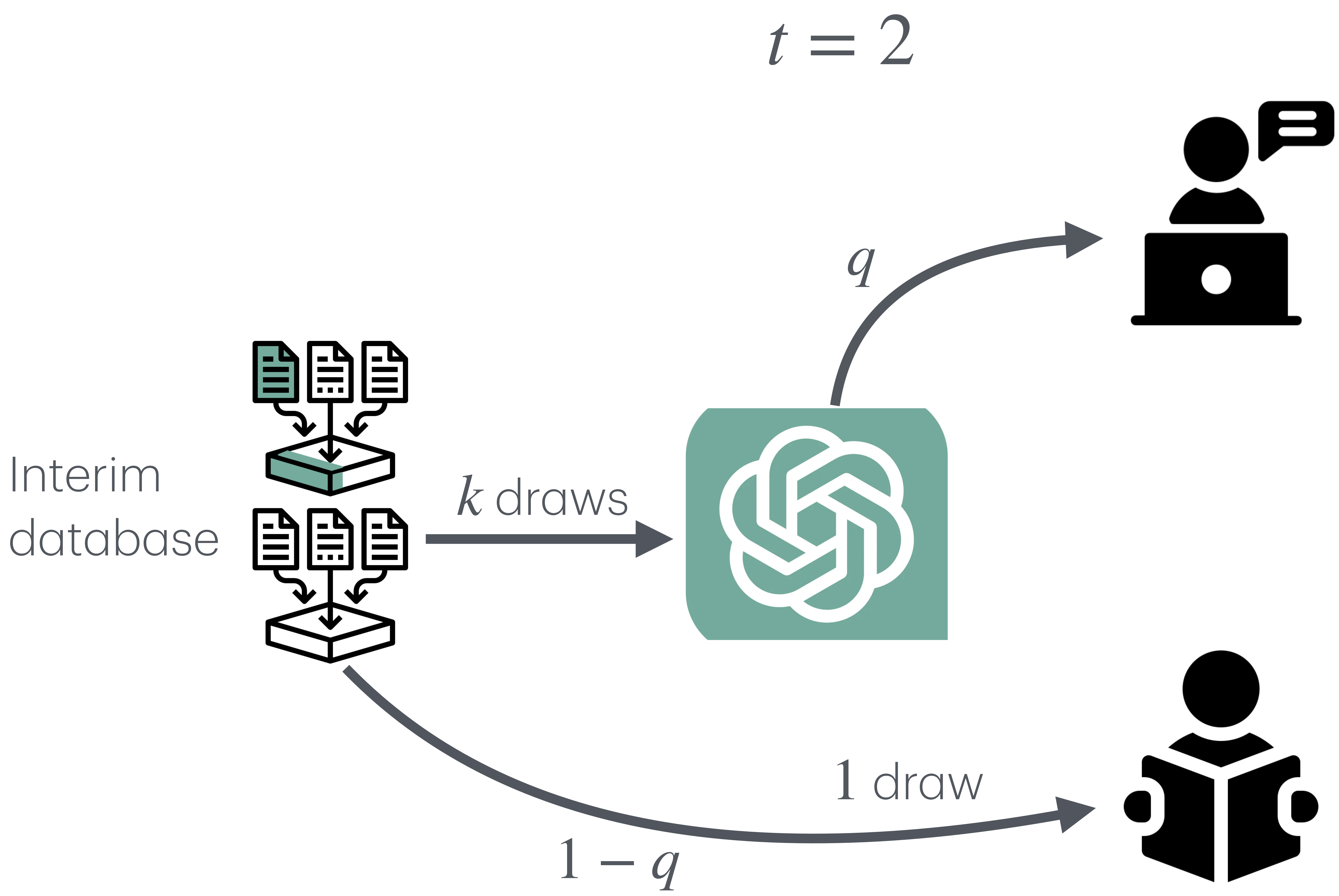


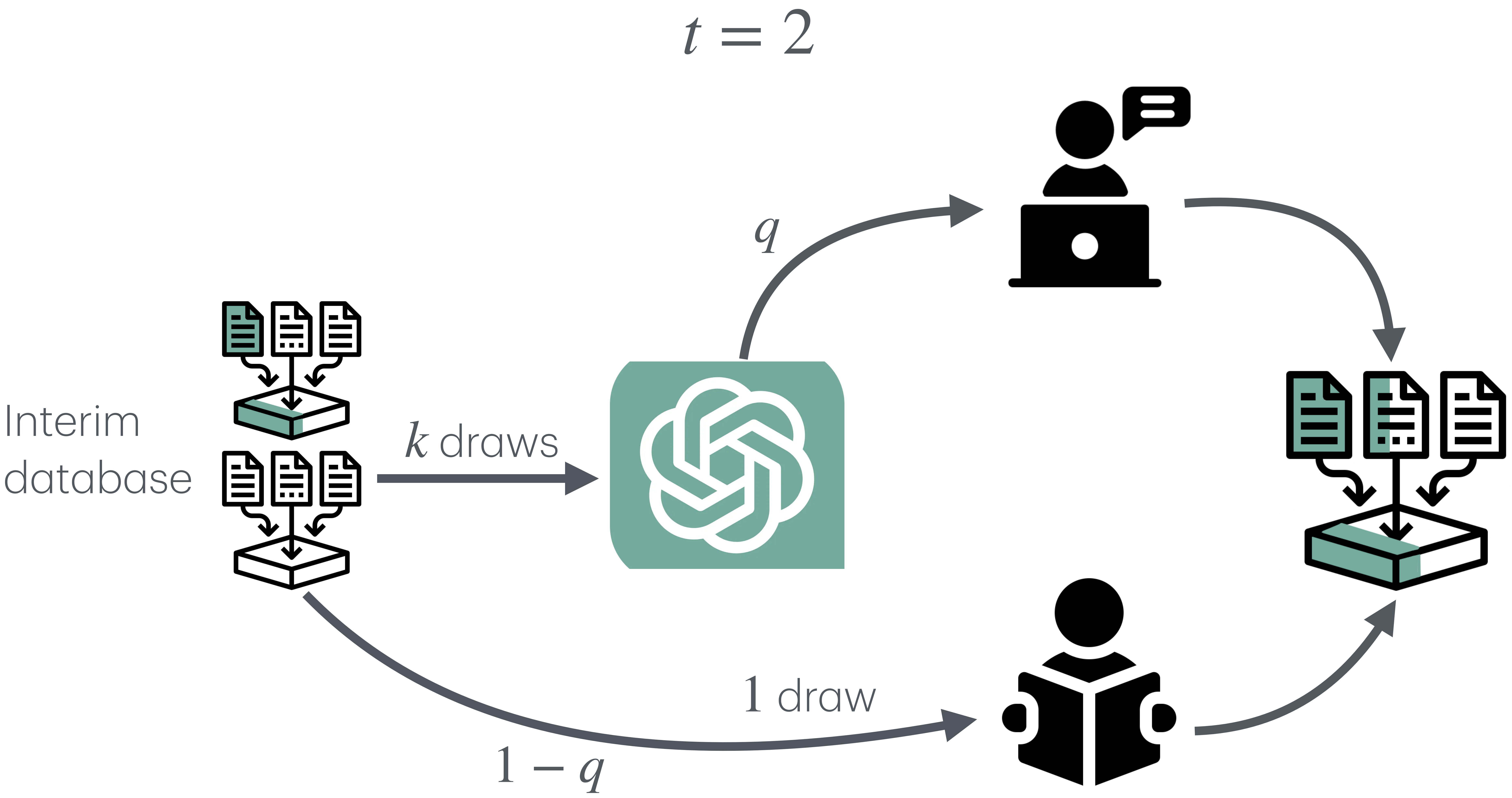
$t = 2$



$t = 2$




$t = 2$
$q$
Interim database
$k$ draws
1 draw
$1 - q$

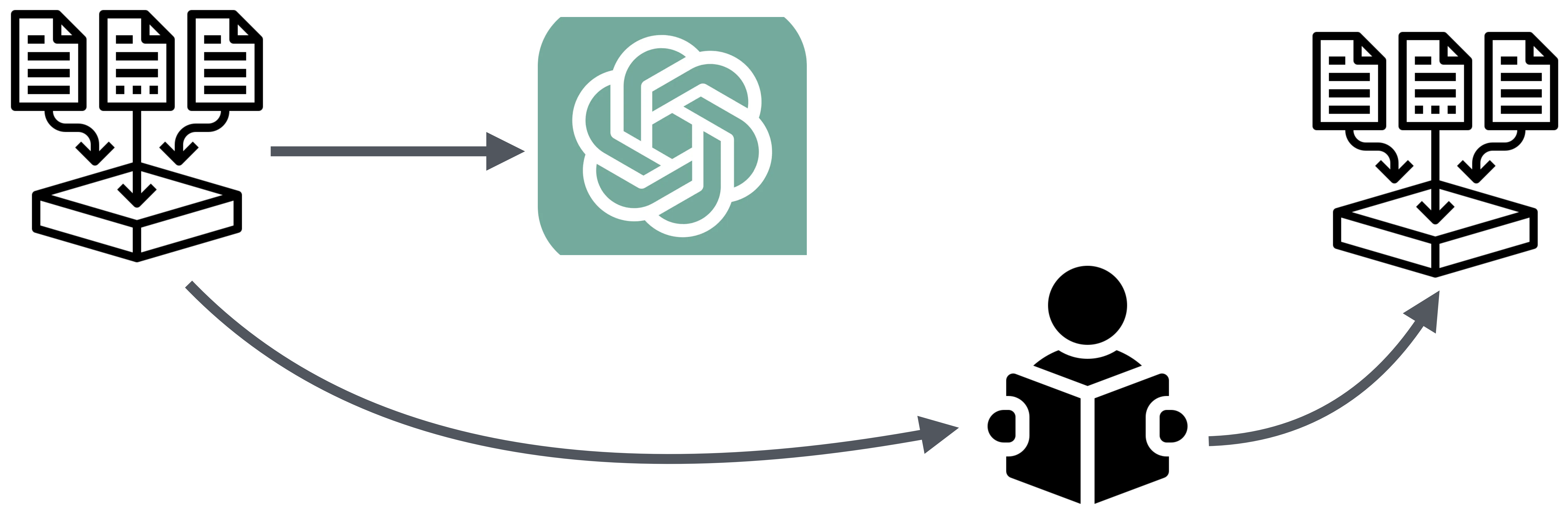


$t = 2$
$q$
Interim database
$k$ draws
1 draw
$1 - q$

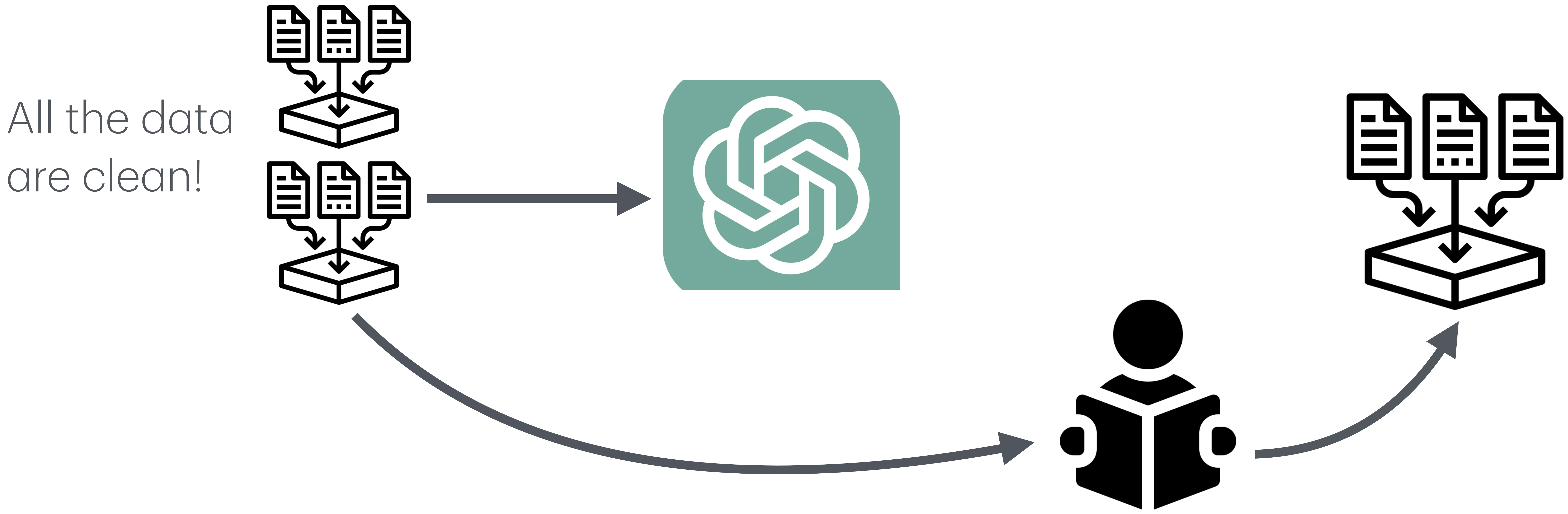
# The case of no recommendations ($q = 0$)

$$t = 1, \; q = 0$$

$t = 1,\ q = 0$

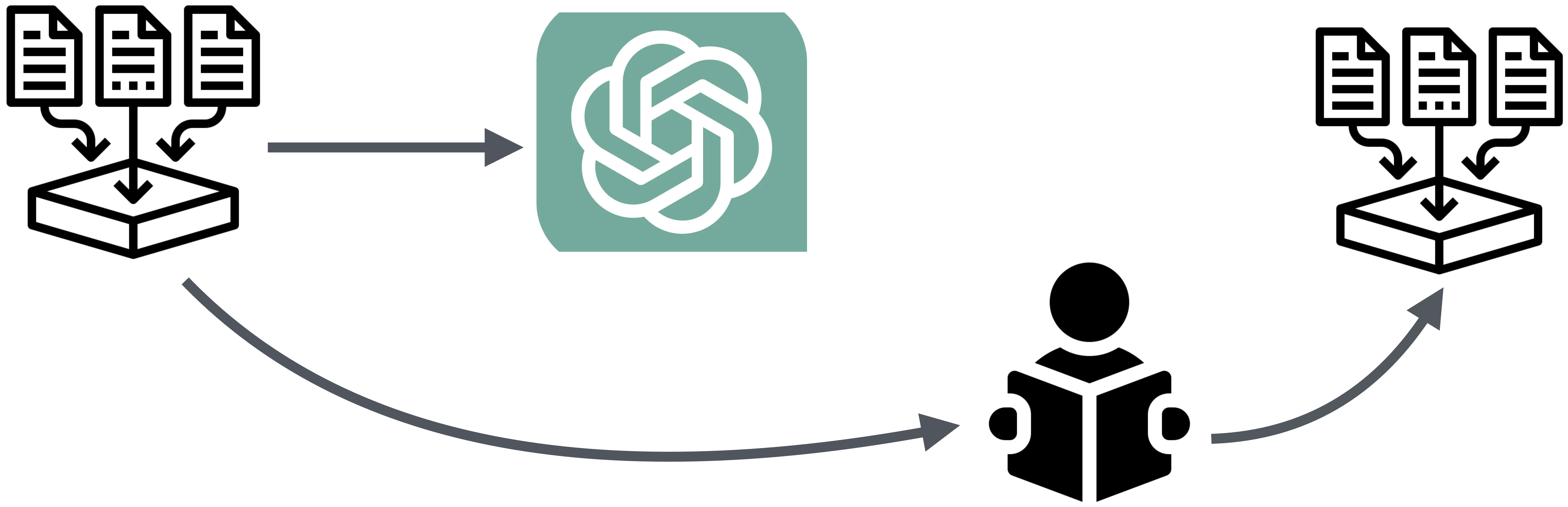


$$t = 2,\ q = 0$$



$$t = 2,\ q = 0$$

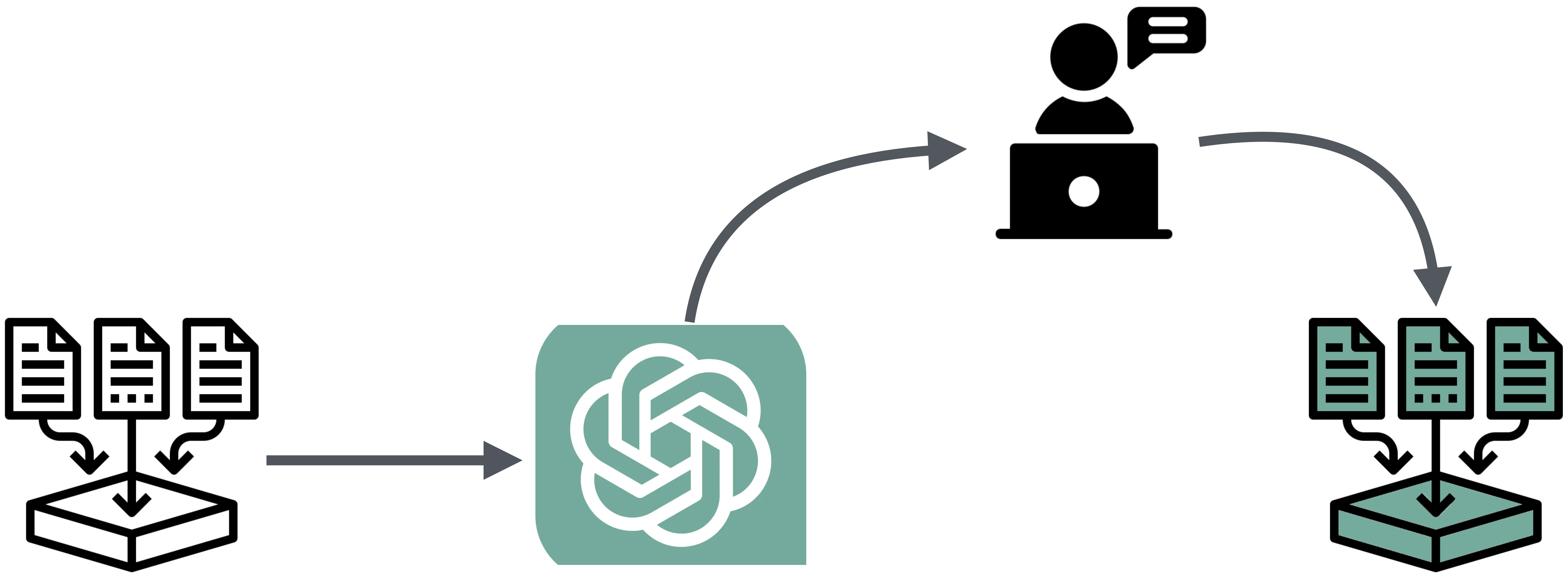
$t = 2,\ q = 0$



# The case of no independent research

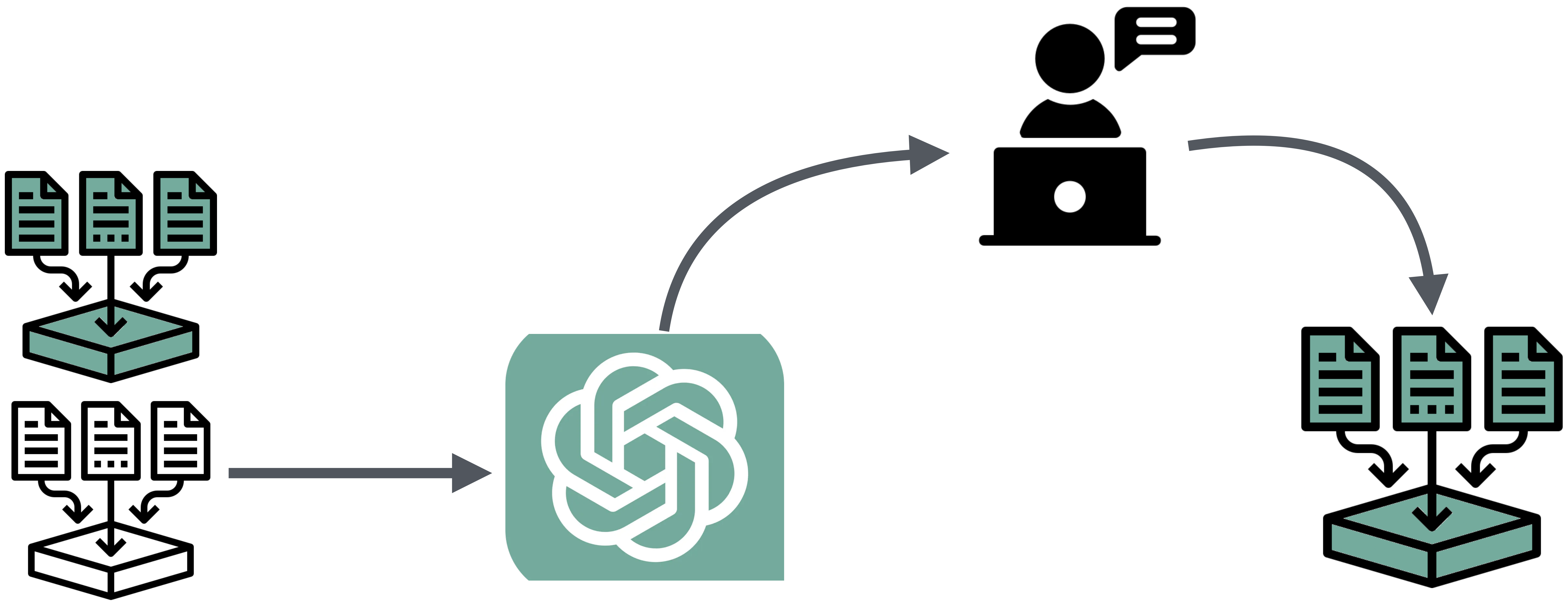
$$(q = 1)$$

$$t = 1,\ q = 1$$



$t = 2,\ q = 1$

Half the data are uninformative!



# Analysis

# Analysis

## Sketch

# Analysis

## Sketch

- There are two kinds of learning we might be interested in.

# Analysis

## Sketch

- There are two kinds of learning we might be interested in.
  - 1. The mass of agents playing the correct action at $t = 1$ and $t = 2$ (agent learning).

# Analysis

## Sketch

- There are two kinds of learning we might be interested in.
  - 1. The mass of agents playing the correct action at $t = 1$ and $t = 2$ (agent learning).
  - 2. The distribution over posteriors for the information aggregator (aggregator learning).

# Analysis

## Sketch

- There are two kinds of learning we might be interested in.
  - 1. The mass of agents playing the correct action at $t = 1$ and $t = 2$ (agent learning).
  - 2. The distribution over posteriors for the information aggregator (aggregator learning).
- Once we understand these two things, we will characterize the optimal recommendation policy for the aggregator (i.e. how often should recommendations be made?).

# Analysis

## Sketch

- There are two kinds of learning we might be interested in.
    - 1. The mass of agents playing the correct action at $t = 1$ and $t = 2$ (agent learning).
    - 2. The distribution over posteriors for the information aggregator (aggregator learning).
- Once we understand these two things, we will characterize the optimal recommendation policy for the aggregator (i.e. how often should recommendations be made?).
- The answer will depend on which of these types of learning the aggregator cares about.

# Analysis

$t = 0$

# Analysis

$t = 0$

- Without loss of generality, I will assume throughout that $\theta = \theta_0$ .

# Analysis

$t = 0$

- Without loss of generality, I will assume throughout that $\theta = \theta_0$.

- By an appropriate *law of large numbers,* the *initial database* contains a fraction $\pi$ of "correct" signals, and $1 - \pi$ "incorrect" signals.

# Analysis

## $t = 0$

- Without loss of generality, I will assume throughout that $\theta = \theta_0$ .

- By an appropriate *law of large numbers,* the *initial database* contains a fraction $\pi$ of "correct" signals, and $1 - \pi$ "incorrect" signals.

- What will the *interim database* look like? (i.e. at the end of $t = 1$)

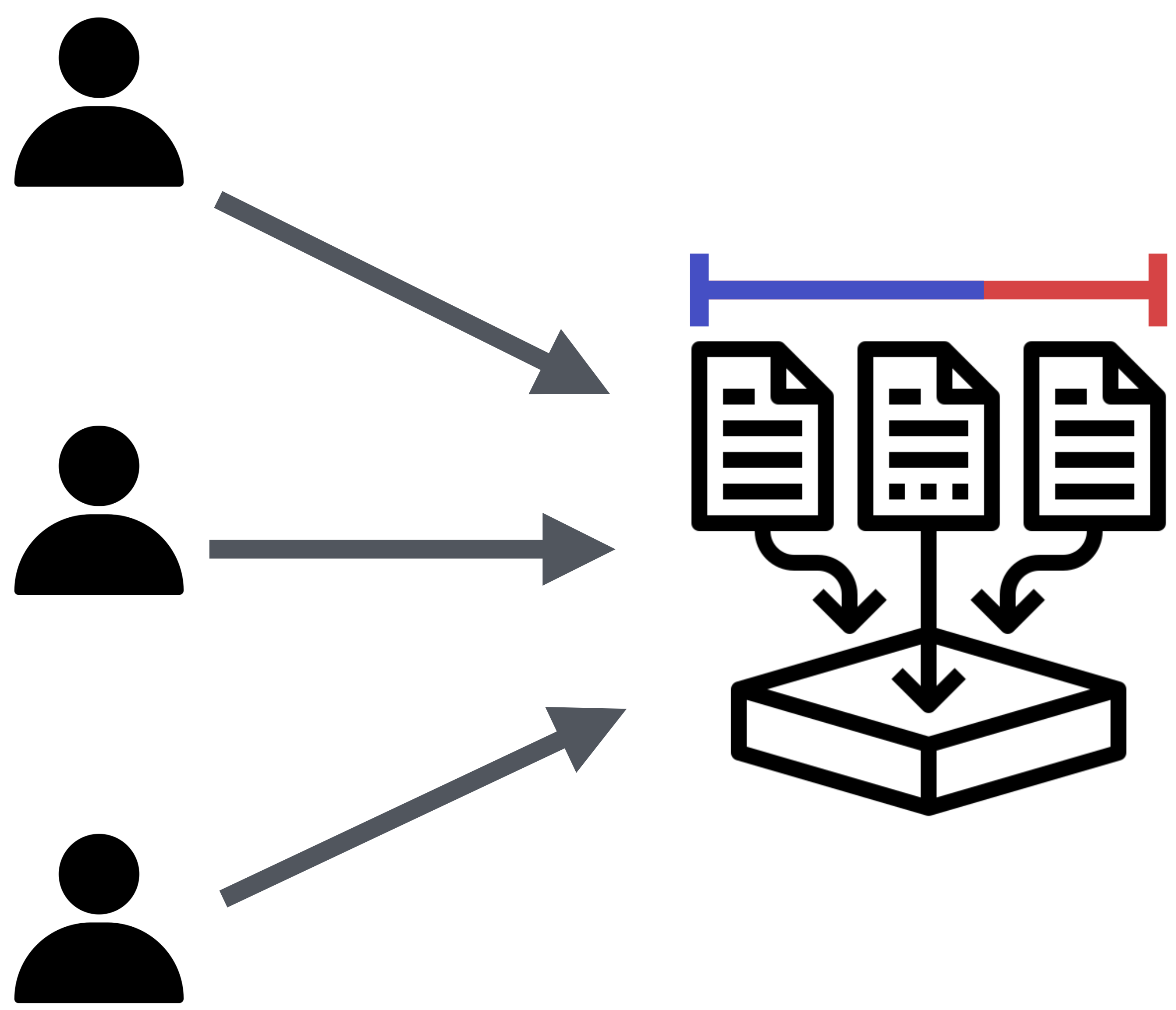
# Analysis

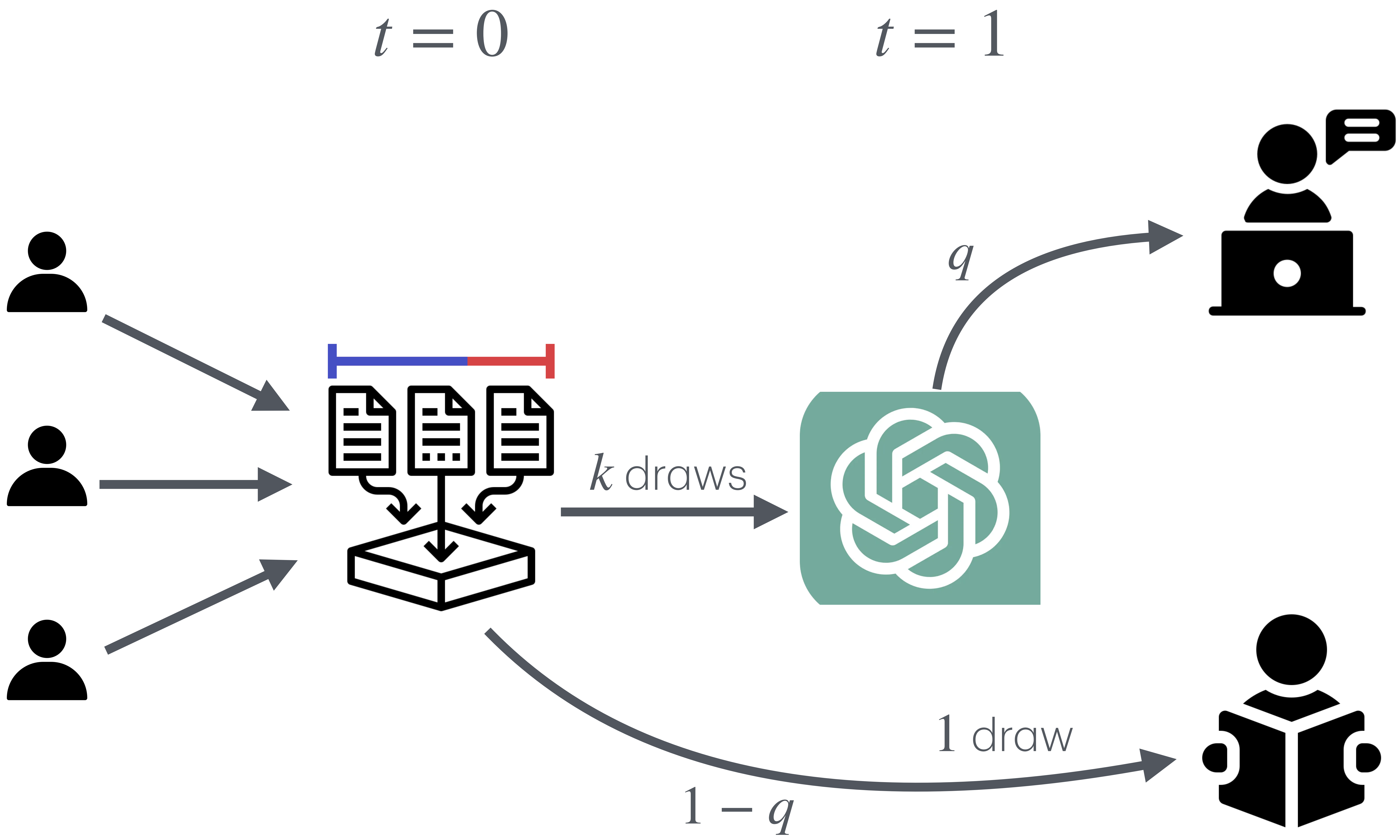
## $t = 0$

- Without loss of generality, I will assume throughout that $\theta = \theta_0$ .
- By an appropriate *law of large numbers*, the *initial database* contains a fraction $\pi$ of "correct" signals, and $1 - \pi$ "incorrect" signals.
- What will the *interim database* look like? (i.e. at the end of $t = 1$)

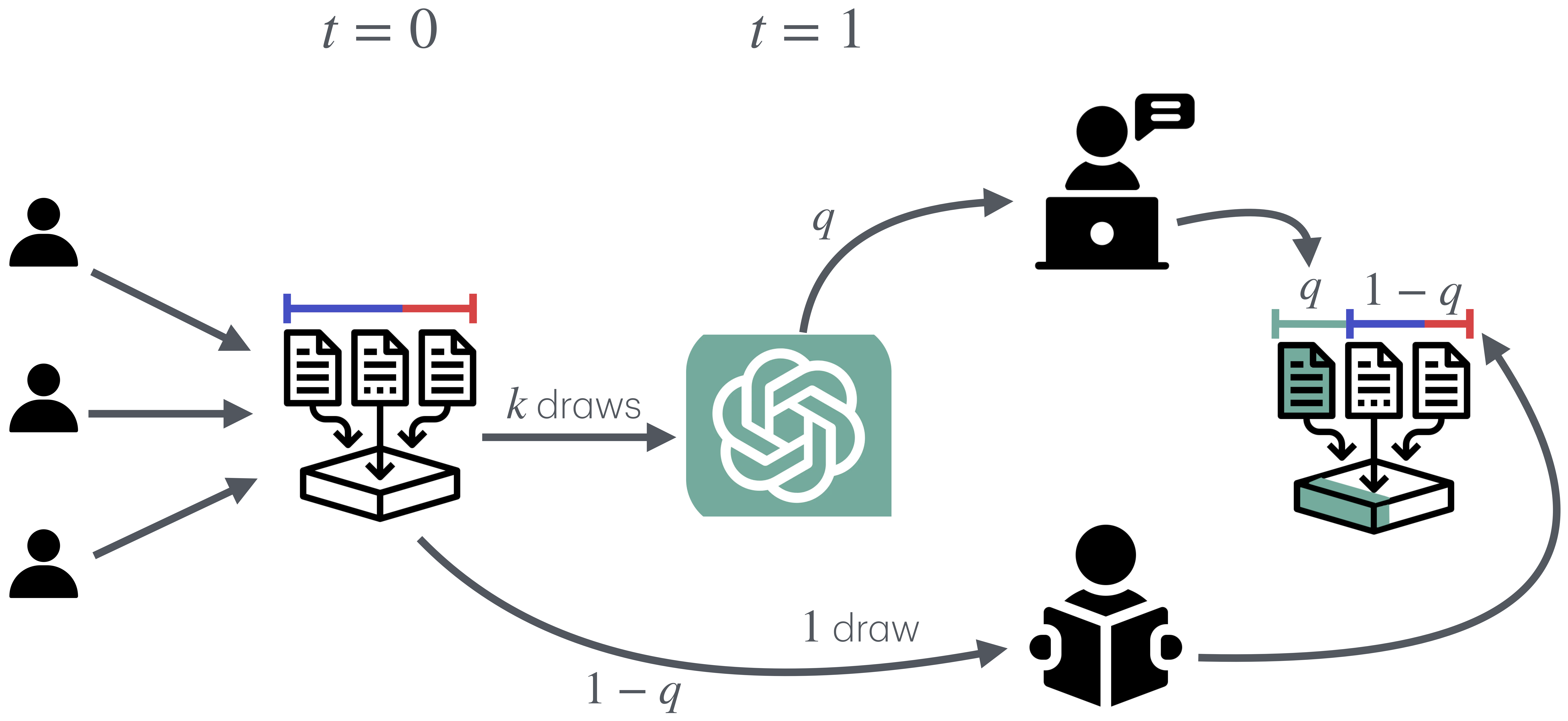
$t = 0$ $t = 1$

$t = 0$ $t = 1$





$t = 0$
$t = 1$
$q$
$k$ draws
1 draw
$1 - q$

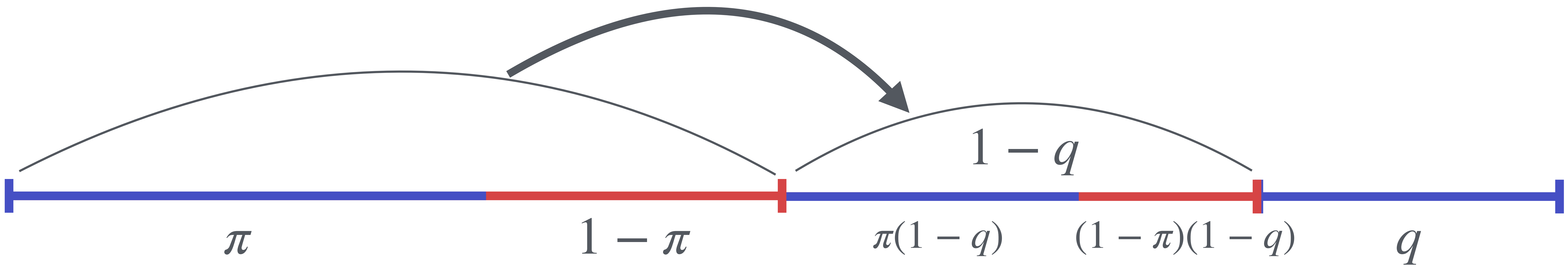


$t = 0$
$t = 1$
$q$
$k$ draws
$q$ $1 - q$
1 draw
$1 - q$

# Analysis: Baseline Model

$t = 1$

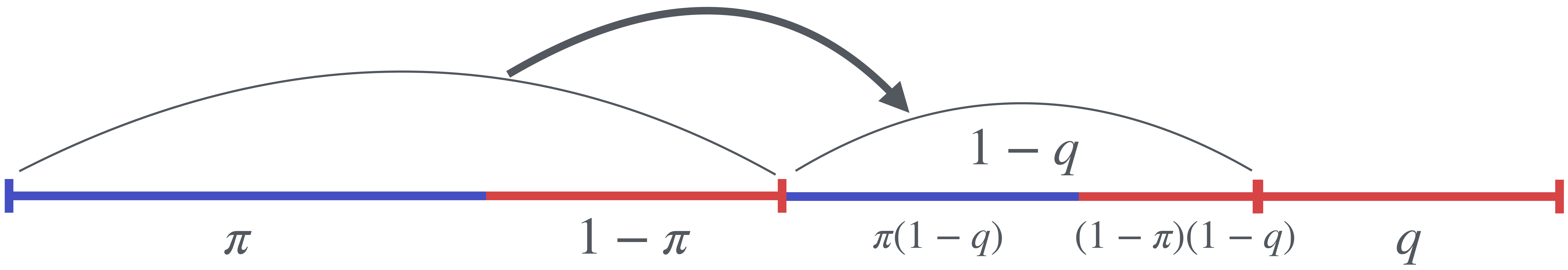
- If the aggregator recommends $a = 0$, then the interim database looks like:



# Analysis: Baseline Model

$t = 1$

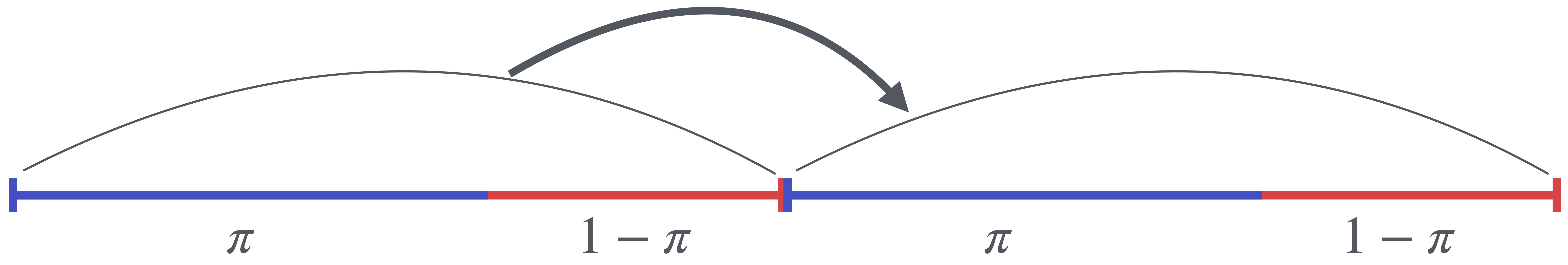
- If the aggregator recommends $a = 1$, then the interim database looks like:



# Analysis: Baseline Model

$t = 1$

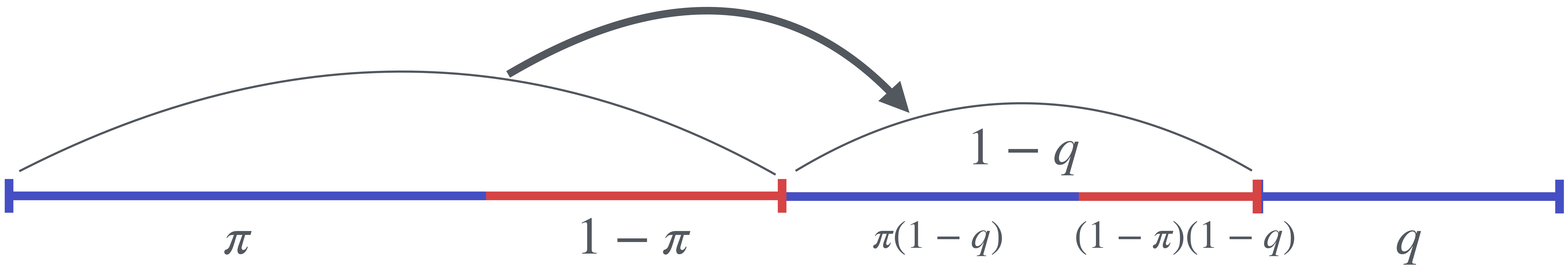
- If the aggregator makes no recommendation, then the interim database looks like:



# Analysis: Baseline Model

$t = 1$

- If the aggregator recommends $a = 0$, then the interim database looks like:



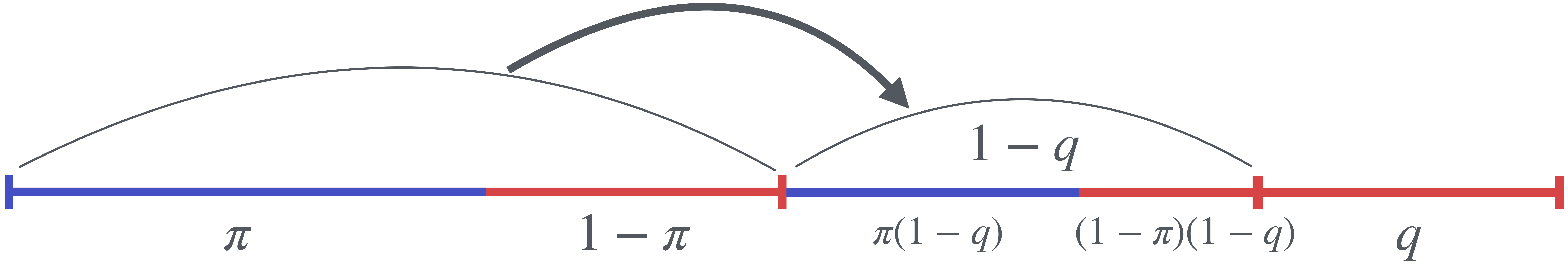

In this case, the mass of correct signals among the new data is:

$$\pi(1 - q) + q = \pi + q(1 - \pi)$$

# Analysis: Baseline Model

$t = 1$

- If the aggregator recommends $a = 1$, then the interim database looks like:



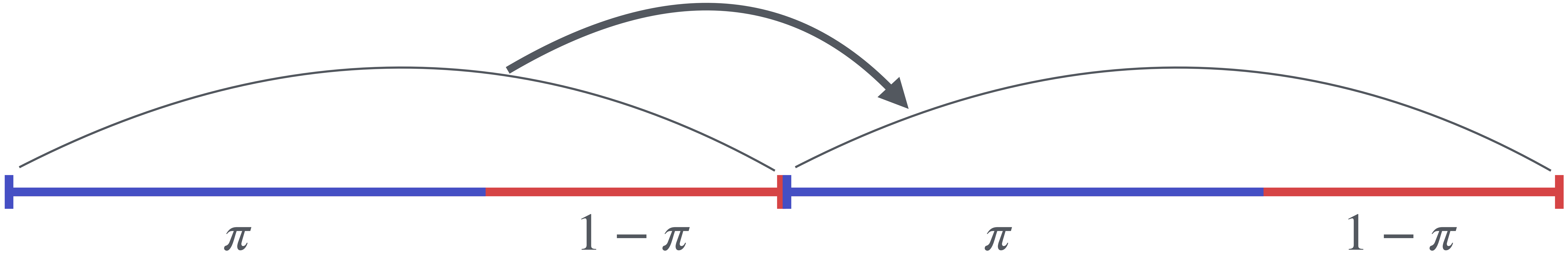

In this case, the mass of correct signals among the new data is:

$$\pi(1 - q) = \pi - q\pi$$

# Analysis: Baseline Model

$t = 1$

- If the aggregator makes no recommendation, then the interim database looks like:




In this case, the mass of correct signals among the new data is:

$$\pi$$

# Analysis

$t = 1$

# Analysis

## $t = 1$

- Let $\hat{\pi}_k = \mathbb{P}(\theta = \theta_0 \mid a_1, \ldots, a_k)$ be the interim posterior probability that the state is $\theta_0$ given the $k$ observations $a_1, \ldots, a_k$.

# Analysis

## $t = 1$

- Let $\hat{\pi}_k = \mathbb{P}(\theta = \theta_0 \mid a_1, \ldots, a_k)$ be the interim posterior probability that the state is $\theta_0$ given the $k$ observations $a_1, \ldots, a_k$.
- Let $X_1$ be a random variable equal to the proportion of agents who choose the correct action at $t = 1$. Then our preceding analysis shows that if the aggregator recommends $0$,

# Analysis

## $t = 1$

- Let $\hat{\pi}_k = \mathbb{P}(\theta = \theta_0 \mid a_1, \ldots, a_k)$ be the interim posterior probability that the state is $\theta_0$ given the $k$ observations $a_1, \ldots, a_k$.

- Let $X_1$ be a random variable equal to the proportion of agents who choose the correct action at $t = 1$. Then our preceding analysis shows that if the aggregator recommends $0$,

$$\mathbb{E}[X_1 \mid a_1, \ldots, a_k] = \pi(1 - q) + q\mathbb{P}(\theta = \theta_0 \mid a_1, \ldots, a_k) = \pi + q(\hat{\pi}_k - \pi).$$

# Analysis

## $t = 1$

- Let $\hat{\pi}_k = \mathbb{P}(\theta = \theta_0 \mid a_1, \ldots, a_k)$ be the interim posterior probability that the state is $\theta_0$ given the $k$ observations $a_1, \ldots, a_k$.

- Let $X_1$ be a random variable equal to the proportion of agents who choose the correct action at $t = 1$. Then our preceding analysis shows that if the aggregator recommends $0$,

$$\mathbb{E}[X_1 \mid a_1, \ldots, a_k] = \pi(1 - q) + q\mathbb{P}(\theta = \theta_0 \mid a_1, \ldots, a_k) = \pi + q(\hat{\pi}_k - \pi) .$$

- $X_1$ captures *agent learning* in $t = 1$.

# Analysis

## Agent learning at $t = 1$

**Lemma 1 (Aggregator improves agent learning whenever $\hat{\pi}_k > \pi$)**

Recommending $a = 0$ strictly increases the expected proportion $(X_1)$ of agents taking the correct action at $t = 1$ iff the aggregator is sufficiently confident about the state $(\hat{\pi}_k > \pi)$.

# Analysis

## Agent learning at $t = 1$

**Lemma 1 (Aggregator improves agent learning whenever $\hat{\pi}_k > \pi$)**

Recommending $a = 0$ strictly increases the expected proportion $(X_1)$ of agents taking the correct action at $t = 1$ iff the aggregator is sufficiently confident about the state $(\hat{\pi}_k > \pi)$.

- We can also think about agent learning from an ex-ante perspective before knowing $\hat{\pi}_k$.

# Analysis

## Agent learning ex-ante at $t = 1$

**Corollary 1 (Aggregator improves agent learning in expectation iff $k > 1$)**

The aggregator improves agent learning in expectation iff it has strictly better information than agents $(k > 1)$. Moreover, agent learning is strictly increasing in $q$.

# Analysis

$t = 2$

# Analysis

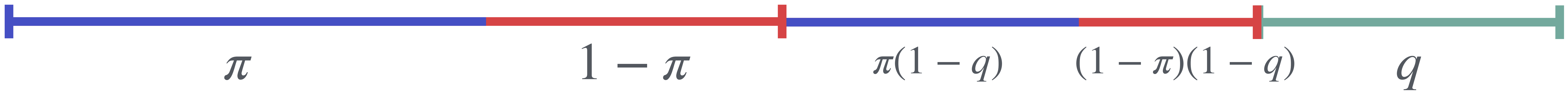
## $t = 2$

- Recall that if the aggregator makes a recommendation, the interim database looks like:

# Analysis

$t = 2$

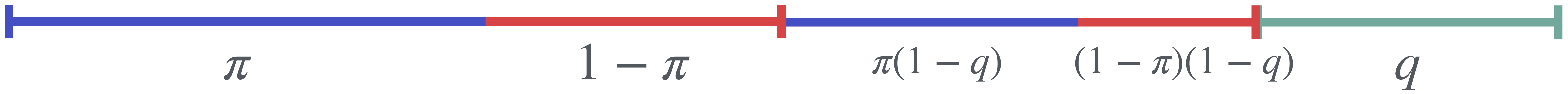
- Recall that if the aggregator makes a recommendation, the interim database looks like:



# Analysis

$t = 2$

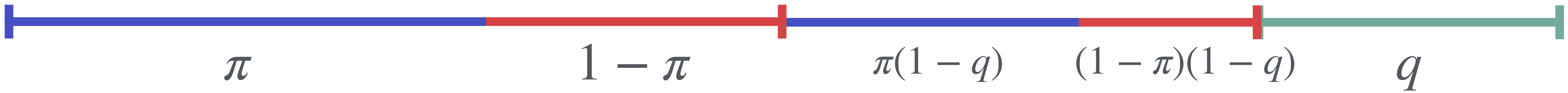
- Recall that if the aggregator makes a recommendation, the interim database looks like:



# Analysis

$t = 2$

- Recall that if the aggregator makes a recommendation, the interim database looks like:

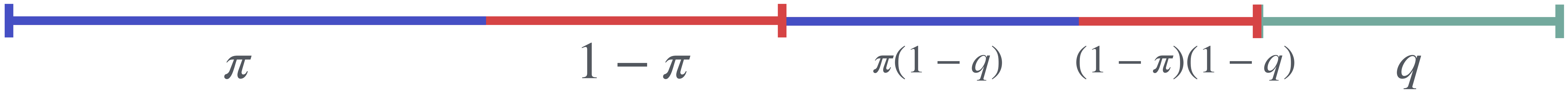



- Or reorganising,

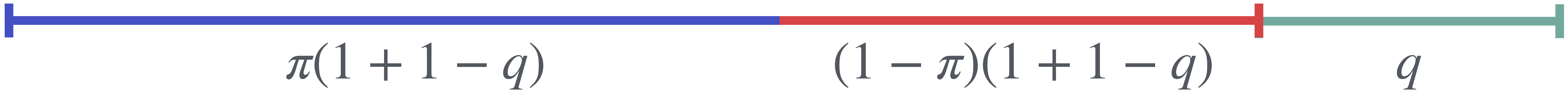
# Analysis

$t = 2$

- Recall that if the aggregator makes a recommendation, the interim database looks like:



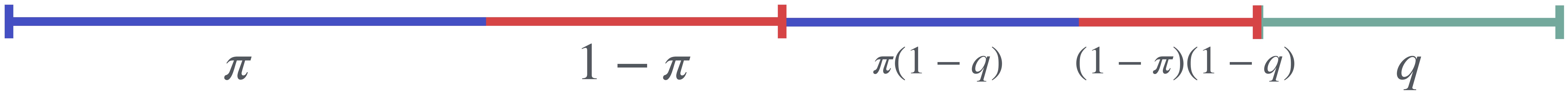

- Or reorganising,



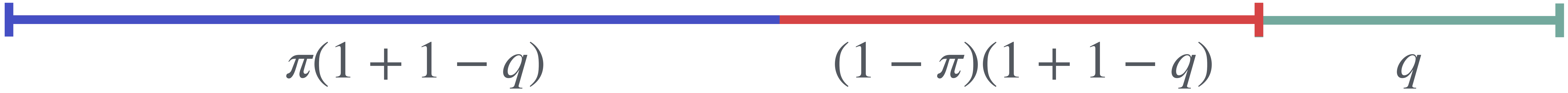
# Analysis

$t = 2$

- Recall that if the aggregator makes a recommendation, the interim database looks like:



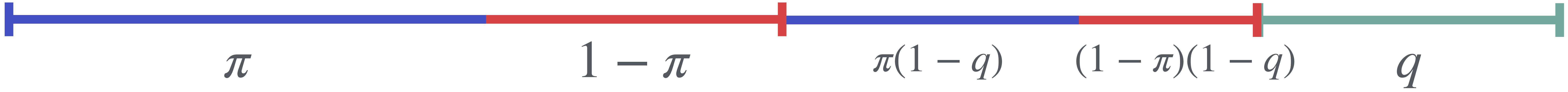

- Or reorganising,



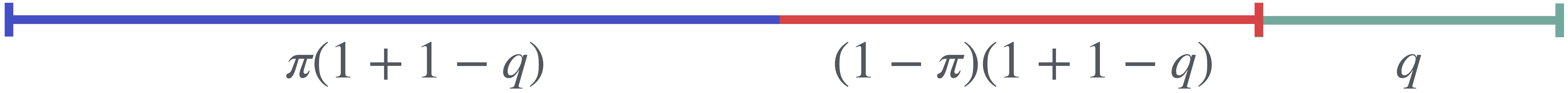
# Analysis

$t = 2$

- Recall that if the aggregator makes a recommendation, the interim database looks like:




- Or reorganising,

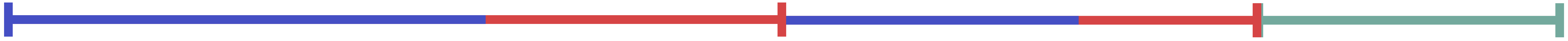



- Squishing it down to a unit mass,

# Analysis

$t = 2$

- Recall that if the aggregator makes a recommendation, the interim database looks like:



$\pi$ $\quad 1 - \pi$ $\quad \pi(1 - q)$ $\quad (1 - \pi)(1 - q)$ $\quad q$

- Or reorganising,

$\pi(1 + 1 - q)$ $\quad (1 - \pi)(1 + 1 - q)$ $\quad q$

- Squishing it down to a unit mass,

$\frac{1}{2}\pi(1 - \frac{q}{2})$ $\quad \frac{1}{2}(1 - \pi)(1 - \frac{q}{2})$ $\quad \frac{q}{2}$

# Analysis

## $t = 2$

- With probably $\frac{q}{2}$, the training data drawn from the interim database at $t = 2$ is uninformative. With probability $1 - \frac{q}{2}$ it is exactly as informative as the original signal.
  - $\Rightarrow$ The period $2$ database is strictly less informative than the initial database.

**Lemma 2 (Aggregator worsens its own learning)**

The informativeness of the interim $(t = 2)$ database is strictly decreasing in $q$ (in the sense of the Blackwell order).

# Analysis

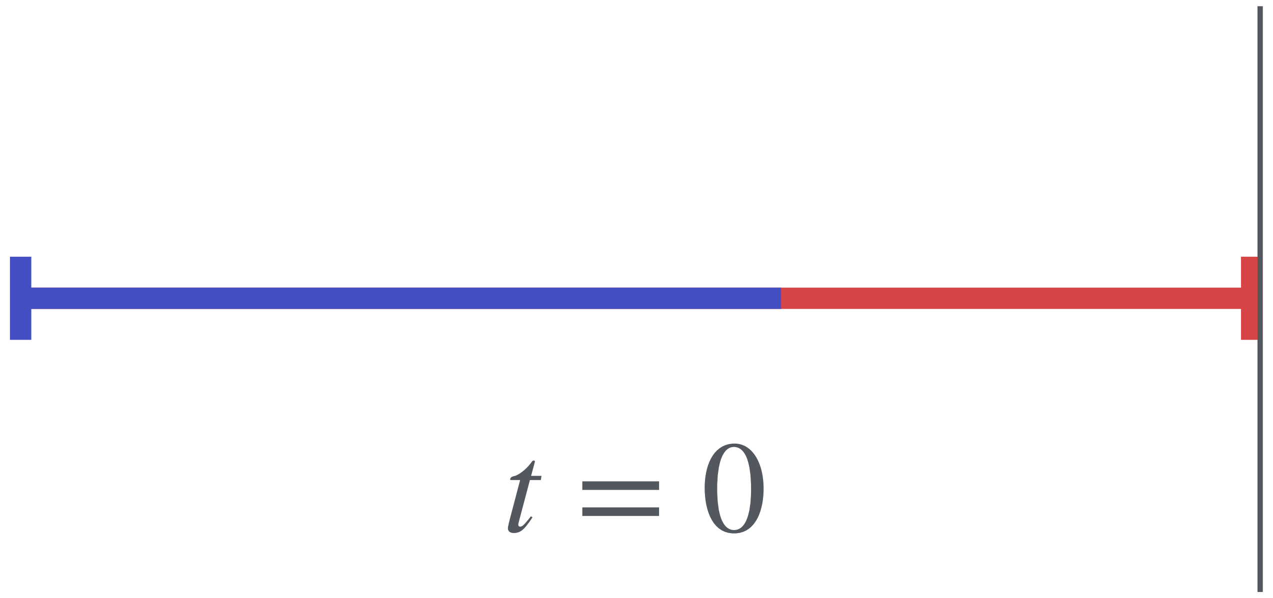
## $t = 2$

- Let $X_2$ be a random variable equal to the proportion of agents who choose the correct action at $t = 2$. What does the posterior database look like?

# Analysis

$t = 2$

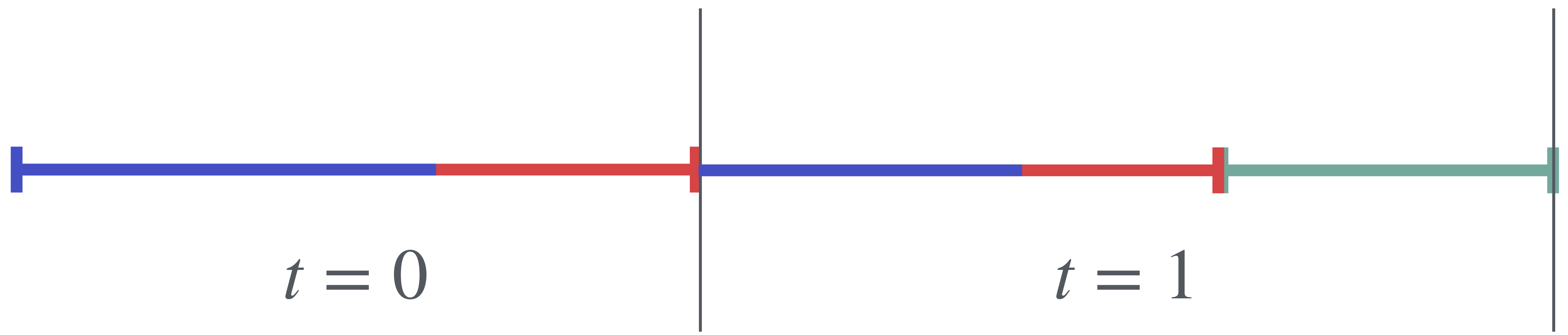
- Let $X_2$ be a random variable equal to the proportion of agents who choose the correct action at $t = 2$. What does the posterior database look like?



# Analysis

## $t = 2$

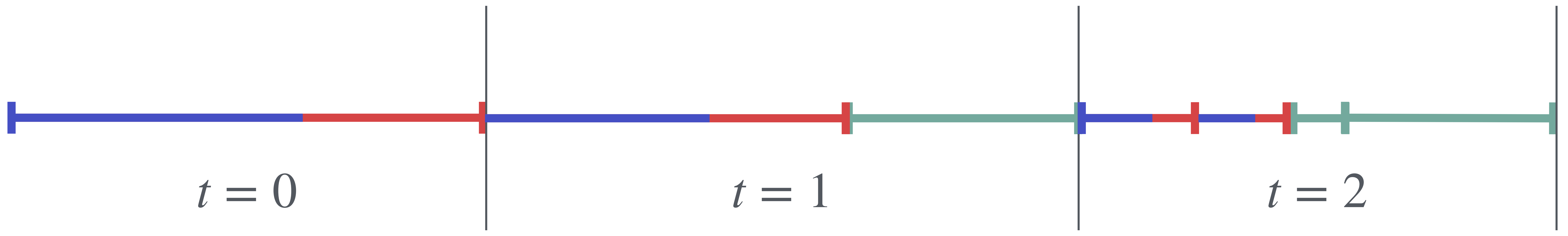
- Let $X_2$ be a random variable equal to the proportion of agents who choose the correct action at $t = 2$. What does the posterior database look like?



# Analysis

## $t = 2$

- Let $X_2$ be a random variable equal to the proportion of agents who choose the correct action at $t = 2$. What does the posterior database look like?



# Analysis: Optimal Recommendation Policy

# Analysis: Optimal Recommendation Policy

Combining Propositions 1 & 2

# Analysis: Optimal Recommendation Policy

## Combining Propositions 1 & 2

- **Proposition 1:** In expectation, the aggregator increases the $t = 1$ mass of agents playing the correct action.

# Analysis: Optimal Recommendation Policy

## Combining Propositions 1 & 2

- **Proposition 1:** In expectation, the aggregator increases the $t = 1$ mass of agents playing the correct action.

- **Proposition 2:** BUT, it does so at the expense of a less informative $t = 2$ database.

# Analysis: Optimal Recommendation Policy

## Combining Propositions 1 & 2

- **Proposition 1:** In expectation, the aggregator increases the $t = 1$ mass of agents playing the correct action.

- **Proposition 2:** BUT, it does so at the expense of a less informative $t = 2$ database.

- Recall the two types of learning: agent / aggregator. Consider two extremes.

# Analysis: Optimal Recommendation Policy

## Combining Propositions 1 & 2

- **Proposition 1:** In expectation, the aggregator increases the $t = 1$ mass of agents playing the correct action.

- **Proposition 2:** BUT, it does so at the expense of a less informative $t = 2$ database.

- Recall the two types of learning: agent / aggregator. Consider two extremes.

  - Aggregator cares only about agent learning.

# Analysis: Optimal Recommendation Policy

## Combining Propositions 1 & 2

- **Proposition 1:** In expectation, the aggregator increases the $t = 1$ mass of agents playing the correct action.
- **Proposition 2:** BUT, it does so at the expense of a less informative $t = 2$ database.

- Recall the two types of learning: agent / aggregator. Consider two extremes.
  - Aggregator cares only about agent learning.
  - Aggregator cares only about aggregator learning.

# Analysis: Optimal Recommendation Policy

Long-term learning

# Analysis: Optimal Recommendation Policy

## Long-term learning

- We're going to focus on the case where the aggregator decides in each period whether or not to make a recommendation (holding $q$ fixed).

# Analysis: Optimal Recommendation Policy

## Long-term learning

- We're going to focus on the case where the aggregator decides in each period whether or not to make a recommendation (holding $q$ fixed).
- Think of this as choosing whether to release ChatGPT 4.5 or whether to publish a paper.

# Analysis: Optimal Recommendation Policy

## Long-term learning

- We're going to focus on the case where the aggregator decides in each period whether or not to make a recommendation (holding $q$ fixed).
- Think of this as choosing whether to release ChatGPT 4.5 or whether to publish a paper.
- If the aggregator cares only about long-term learning then the optimal recommendation policy is simple:

# Analysis: Optimal Recommendation Policy

## Long-term learning

- We're going to focus on the case where the aggregator decides in each period whether or not to make a recommendation (holding $q$ fixed).
- Think of this as choosing whether to release ChatGPT 4.5 or whether to publish a paper.
- If the aggregator cares only about long-term learning then the optimal recommendation policy is simple:
- By **Proposition 2,** it will never make any recommendations.

# Analysis: Optimal Recommendation Policy

## Aggregator's Decision Problem

# Analysis: Optimal Recommendation Policy

## Aggregator's Decision Problem

- Suppose the aggregator derives utility from the proportion of agents in each period who choose the correct strategy.

# Analysis: Optimal Recommendation Policy

## Aggregator's Decision Problem

- Suppose the aggregator derives utility from the proportion of agents in each period who choose the correct strategy.

- Let $X_i,\ i = 1,2$ denote the (random) proportion of agents who choose the correct strategy at $t = 1,2$.

# Analysis: Optimal Recommendation Policy

## Aggregator's Decision Problem

- Suppose the aggregator derives utility from the proportion of agents in each period who choose the correct strategy.
- Let $X_i,\ i = 1,2$ denote the (random) proportion of agents who choose the correct strategy at $t = 1,2$.
- To fix ideas, let's take

# Analysis: Optimal Recommendation Policy

## Aggregator's Decision Problem

- Suppose the aggregator derives utility from the proportion of agents in each period who choose the correct strategy.

- Let $X_i,\ i = 1,2$ denote the (random) proportion of agents who choose the correct strategy at $t = 1,2$.

- To fix ideas, let's take

$$u_1(X_1, X_2) = X_1 + X_2, \qquad u_2(X_2) = X_2 \,.$$

# Analysis: Optimal Recommendation Policy

## Aggregator's Decision Problem

- Suppose the aggregator derives utility from the proportion of agents in each period who choose the correct strategy.

- Let $X_i,\ i = 1,2$ denote the (random) proportion of agents who choose the correct strategy at $t = 1,2$.

- To fix ideas, let's take

$$u_1(X_1, X_2) = X_1 + X_2, \qquad u_2(X_2) = X_2\,.$$

- Start by characterizing the optimal strategy at $t = 2$.

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation Rule

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation Rule

- At $t = 2$, aggregator faces no consequence for garbling the database.

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation Rule

- At $t = 2$, aggregator faces no consequence for garbling the database.
- Let $\tilde{\pi}_k$ be the $t = 2$ posterior belief that the state is $\theta_0$ (the analogue of $\hat{\pi}_k$).

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation Rule

- At $t = 2,$ aggregator faces no consequence for garbling the database.
- Let $\tilde{\pi}_k$ be the $t = 2$ posterior belief that the state is $\theta_0$ (the analogue of $\hat{\pi}_k$).
- Write $\hat{a}$ for the aggregator's recommendation at $t = 1,$ and $\tilde{a}$ at $t = 2.$

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation Rule

- At $t = 2$, aggregator faces no consequence for garbling the database.
- Let $\tilde{\pi}_k$ be the $t = 2$ posterior belief that the state is $\theta_0$ (the analogue of $\hat{\pi}_k$).
- Write $\hat{a}$ for the aggregator's recommendation at $t = 1$, and $\tilde{a}$ at $t = 2$.
- The choice of whether or not to recommendation in $t = 2$ depends on which recommendation was made at $t = 1$.

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation Rule

- At $t = 2,$ aggregator faces no consequence for garbling the database.
- Let $\tilde{\pi}_k$ be the $t = 2$ posterior belief that the state is $\theta_0$ (the analogue of $\hat{\pi}_k$).
- Write $\hat{a}$ for the aggregator's recommendation at $t = 1,$ and $\tilde{a}$ at $t = 2.$
- The choice of whether or not to recommendation in $t = 2$ depends on which recommendation was made at $t = 1.$
- It turns out that the aggregator is more willing to recommend at $t = 2$ if it thinks it made a mistake at $t = 1.$

# Analysis: Optimal Recommendation Policy

Optimal Recommendation at $t = 2$

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 2$

**Proposition 3 (Lower the bar for corrections)**

Suppose $\tilde{\pi}_k \geq \frac{1}{2}$, then

1. If $\hat{a} = 1$, then the aggregator recommends $\tilde{a} = 0$ when $\tilde{\pi}_k \geq \pi^*$ and recommends nothing otherwise.

2. If $\hat{a} = 0$ or no recommendation was made at $t = 1$, then the aggregator optimally recommends $\tilde{a} = 0$ when $\tilde{\pi}_k \geq \pi$ and recommends nothing otherwise.

Where $\pi^* = \dfrac{2\pi + q(1 - \pi)}{2 + q} \in (\frac{1}{2}, \pi)$.

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 2$

**Proposition 3 (Lower the bar for corrections)**

Suppose $\tilde{\pi}_k \geq \frac{1}{2}$, then

1. If $\hat{a} = 1$, then the aggregator recommends $\tilde{a} = 0$ when $\tilde{\pi}_k \geq \pi^*$ and recommends nothing otherwise.

2. If $\hat{a} = 0$ or no recommendation was made at $t = 1$, then the aggregator optimally recommends $\tilde{a} = 0$ when $\tilde{\pi}_k \geq \pi$ and recommends nothing otherwise.

Where $\pi^* = \dfrac{2\pi + q(1 - \pi)}{2 + q} \in (\frac{1}{2}, \pi)$.

**NOTE:** $\pi^*$ is decreasing in $q$.

# Analysis: Optimal Recommendation Policy

Optimal Recommendation at $t = 1$

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 1$

- Now consider $t = 1$ and suppose $\hat{\pi}_k \geq \frac{1}{2}$ .

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 1$

- Now consider $t = 1$ and suppose $\hat{\pi}_k \geq \frac{1}{2}$.
- The information aggregator maximises: $\mathbb{E}[X_1 + X_2 \mid a_1, \ldots, a_k]$.

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 1$

- Now consider $t = 1$ and suppose $\hat{\pi}_k \geq \frac{1}{2}$.
- The information aggregator maximises: $\mathbb{E}[X_1 + X_2 \mid a_1, \ldots, a_k]$.
- We've already seen that

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 1$

- Now consider $t = 1$ and suppose $\hat{\pi}_k \geq \frac{1}{2}$.
- The information aggregator maximises: $\mathbb{E}[X_1 + X_2 \mid a_1, \ldots, a_k]$.
- We've already seen that

$$\mathbb{E}[X_1 \mid a_1, \ldots, a_k] = \pi + q(\hat{\pi}_k - \pi),$$

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 1$

- Now consider $t = 1$ and suppose $\hat{\pi}_k \geq \frac{1}{2}$.
- The information aggregator maximises: $\mathbb{E}[X_1 + X_2 \mid a_1, \ldots, a_k]$.
- We've already seen that

$$\mathbb{E}[X_1 \mid a_1, \ldots, a_k] = \pi + q(\hat{\pi}_k - \pi),$$

- It turns out that if the aggregator is willing to recommend in $t = 1$, then this *improves* outcomes (in expectation) in $t = 2$!

# Analysis: Optimal Recommendation Policy

Optimal Recommendation at $t = 1$

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 1$

- Intuition: Suppose $q = 1$.

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 1$

- Intuition: Suppose $q = 1$.

    1. Your best guess of the recommendation you'll make tomorrow is the recommendation you made today.

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 1$

- Intuition: Suppose $q = 1.$

    1. Your best guess of the recommendation you'll make tomorrow is the recommendation you made today.

    2. If making a recommendation today improves outcomes today, then you expect (a) to make the same recommendation tomorrow (that's just 1.) and (b) that the same recommendation tomorrow will improve outcomes tomorrow.

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation at $t = 1$

- Intuition: Suppose $q = 1$.

1. Your best guess of the recommendation you'll make tomorrow is the recommendation you made today.

2. If making a recommendation today improves outcomes today, then you expect (a) to make the same recommendation tomorrow (that's just 1.) and (b) that the same recommendation tomorrow will improve outcomes tomorrow.

3. So if it pays to make the recommendation today, it always pays (in expectation) for tomorrow!

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation Policy

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation Policy

**Proposition 4 (Optimal Recommendation in $t = 1$)**

1. If $\hat{\pi}_k \geq \pi$ then the aggregator recommends $a = 0$.
2. If $1 - \pi < \hat{\pi}_k < \pi$ then the aggregator makes no recommendation.
3. If $\hat{\pi}_k \leq 1 - \pi$ then the aggregator recommends $a = 1$.

# Analysis: Optimal Recommendation Policy

## Optimal Recommendation Policy

**Proposition 4 (Optimal Recommendation in $t = 1$)**

1. If $\hat{\pi}_k \geq \pi$ then the aggregator recommends $a = 0$.
2. If $1 - \pi < \hat{\pi}_k < \pi$ then the aggregator makes no recommendation.
3. If $\hat{\pi}_k \leq 1 - \pi$ then the aggregator recommends $a = 1$.

**NOTE:** this opens up the possibility that a large number of agents take the wrong action in both periods (recommendations increase the variance of $X_1, X_2$).

# Analysis: Optimal Recommendation Policy

## Summary

# Analysis: Optimal Recommendation Policy

## Summary

- If the aggregator cares only about aggregator learning, it never makes recommendations.

# Analysis: Optimal Recommendation Policy

## Summary

- If the aggregator cares only about aggregator learning, it never makes recommendations.
- If the aggregator cares only about agent learning, it makes recommendation whenever they improve short-term outcomes.

# Analysis: Optimal Recommendation Policy

## Summary

- If the aggregator cares only about aggregator learning, it never makes recommendations.
- If the aggregator cares only about agent learning, it makes recommendation whenever they improve short-term outcomes.
- What if the aggregator cares about a mix of these things?

# Analysis: Optimal Recommendation Policy

## Summary

- If the aggregator cares only about aggregator learning, it never makes recommendations.
- If the aggregator cares only about agent learning, it makes recommendation whenever they improve short-term outcomes.
- What if the aggregator cares about a mix of these things?
  - Intuitively: Pushes the threshold $\hat{\pi}_k$ at which the aggregator is willing to recommend upwards.

# Where does this leave us?

The American Economic Review

ARTICLES

MARIAGIOVANNA BACCARA, AYŞE İMROHOROĞLU, ALISTAIR J. WILSON, AND LEEAT YARIV
A Field Study on Matching with Network Externalities

GADI BARLEVY AND DEREK NEAL
Pay for Percentile

RAN ABRAMITZKY, LEAH PLATT BOUSTAN, AND KATHERINE ERIKSSON
Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration

MATTHEW O. JACKSON, TOMAS RODRIGUEZ-BARRAQUER, AND XU TAN
Social Capital and Social Quilts: Network Patterns of Favor Exchange

PATRICK BAJARI, JANE COOLEY FRUEHWIRTH, KYOO IL KIM, AND CHRISTOPHER TIMMINS
A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution

GORDON B. DAHL AND LANCE LOCHNER
The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit

DOH-SHIN JEON AND DOMENICO MENICUCCI
Bundling and Competition for Slots

VERONICA GUERRIERI AND PÉTER KONDOR
Fund Managers, Career Concerns, and Asset Price Volatility

JUDD B. KESSLER AND ALVIN E. ROTH
Organ Allocation Policy and the Decision to Donate

SCOTT A. IMBERMAN, ADRIANA D. KUGLER, AND BRUCE I. SACERDOTE
Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees

DANIEL J. BENJAMIN, ORI HEFFETZ, MILES S. KIMBALL, AND ALEX REES-JONES
What Do You Think Would Make You Happier? What Do You Think You Would Choose?

KEYU JIN
Industrial Structure and Capital Flows

# Where does this leave us?

# Where does this leave us?

- Optimal policy depends on which kind of learning we care about.

# Where does this leave us?

- Optimal policy depends on which kind of learning we care about.
    - Likely from a social perspective the answer is a mix of agent and aggregator.

# Where does this leave us?

- Optimal policy depends on which kind of learning we care about.
  - Likely from a social perspective the answer is a mix of agent and aggregator.
- Policies which improve agent learning in expectation can also lead to a large number of agents taking the wrong action.

# Where does this leave us?

- Optimal policy depends on which kind of learning we care about.
    - Likely from a social perspective the answer is a mix of agent and aggregator.
- Policies which improve agent learning in expectation can also lead to a large number of agents taking the wrong action.
    - We should be cautious about how feedback can amplify mistakes (and be quick to correct those mistakes, or at the very least stop making recommendations until we learn more).

# Where does this leave us?

- Optimal policy depends on which kind of learning we care about.
    - Likely from a social perspective the answer is a mix of agent and aggregator.
- Policies which improve agent learning in expectation can also lead to a large number of agents taking the wrong action.
    - We should be cautious about how feedback can amplify mistakes (and be quick to correct those mistakes, or at the very least stop making recommendations until we learn more).
- Policies which help screen data for “traces of feedback” can help improve learning.

# Extensions, Limitations and Conclusion

# Extensions

# Extensions

- What are the limit points of learning with feedback loops? (Difficult: not i.i.d.)

# Extensions

- What are the limit points of learning with feedback loops? (Difficult: not i.i.d.)
  - In my model, a Bayesian aggregator still learns as $t \to \infty$ .

# Extensions

- What are the limit points of learning with feedback loops? (Difficult: not i.i.d.)
  - In my model, a Bayesian aggregator still learns as $t \to \infty$ .
  - A **naive aggregator** who neglects the feedback loop does not necessarily learn. (*important)

# Extensions

- What are the limit points of learning with feedback loops? (Difficult: not i.i.d.)
  - In my model, a Bayesian aggregator still learns as $t \to \infty$.
  - A **naive aggregator** who neglects the feedback loop does not necessarily learn. (*important)
  - Speed of learning?

# Extensions

- What are the limit points of learning with feedback loops? (Difficult: not i.i.d.)
  - In my model, a Bayesian aggregator still learns as $t \to \infty$.
  - A **naive aggregator** who neglects the feedback loop does not necessarily learn. (*important)
  - Speed of learning?
- Multiple aggregators? Biased aggregators? Competing aggregators?

# Extensions

- What are the limit points of learning with feedback loops? (Difficult: not i.i.d.)
  - In my model, a Bayesian aggregator still learns as $t \to \infty$.
  - A **naive aggregator** who neglects the feedback loop does not necessarily learn. (*important)
  - Speed of learning?
- Multiple aggregators? Biased aggregators? Competing aggregators?
- Changing state? Endogenizing $q$?

# Limitations

# Limitations

- Would like to move beyond binary state binary action but not yet sure how far I can take the model.

# Limitations

- Would like to move beyond binary state binary action but not yet sure how far I can take the model.

- The “pollution” of data relies on there being no “new” data generated. If polluted data makes it more likely that agents somehow correct the polluted data, then this reduces (possibly eliminates) the key tradeoff in my model.

# Limitations

- Would like to move beyond binary state binary action but not yet sure how far I can take the model.
- The “pollution” of data relies on there being no “new” data generated. If polluted data makes it more likely that agents somehow correct the polluted data, then this reduces (possibly eliminates) the key tradeoff in my model.
    - The case where agents draw more than one signal can be construed as “adding in new data”.

# Limitations

- Would like to move beyond binary state binary action but not yet sure how far I can take the model.
- The “pollution” of data relies on there being no “new” data generated. If polluted data makes it more likely that agents somehow correct the polluted data, then this reduces (possibly eliminates) the key tradeoff in my model.
  - The case where agents draw more than one signal can be construed as “adding in new data”.
- In reality agents’ samples are probably correlated, but we have to start somewhere.

# Limitations

- Would like to move beyond binary state binary action but not yet sure how far I can take the model.
- The "pollution" of data relies on there being no "new" data generated. If polluted data makes it more likely that agents somehow correct the polluted data, then this reduces (possibly eliminates) the key tradeoff in my model.
  - The case where agents draw more than one signal can be construed as "adding in new data".
- In reality agents' samples are probably correlated, but we have to start somewhere.
- If prevalence of old data decays faster than $\frac{1}{n}$ then things change.

# Conclusion

# Conclusion

- I develop a model of learning that incorporates *feedback loops*.

# Conclusion

- I develop a model of learning that incorporates *feedback loops*.
- I show that feedback improves agent learning but worsens aggregator learning.

# Conclusion

- I develop a model of learning that incorporates *feedback loops*.
- I show that feedback improves agent learning but worsens aggregator learning.
- The optimal recommendation rule for a strategic information aggregator depends on which of these two types of learning it values.

# Conclusion

- I develop a model of learning that incorporates *feedback loops*.
- I show that feedback improves agent learning but worsens aggregator learning.
- The optimal recommendation rule for a strategic information aggregator depends on which of these two types of learning it values.
- Value on aggregator learning $\Rightarrow$ don't make recommendations!

# Conclusion

- I develop a model of learning that incorporates *feedback loops*.
- I show that feedback improves agent learning but worsens aggregator learning.
- The optimal recommendation rule for a strategic information aggregator depends on which of these two types of learning it values.
- Value on aggregator learning $\Rightarrow$ don't make recommendations!
- Value on agent learning $\Rightarrow$ make recommendations when confident enough about the state.

# Conclusion

- I develop a model of learning that incorporates *feedback loops*.
- I show that feedback improves agent learning but worsens aggregator learning.
- The optimal recommendation rule for a strategic information aggregator depends on which of these two types of learning it values.
- Value on aggregator learning $\Rightarrow$ don't make recommendations!
- Value on agent learning $\Rightarrow$ make recommendations when confident enough about the state.
    - Weaker confidence required to correct mistakes at $t = 2$.

# Conclusion

- I develop a model of learning that incorporates *feedback loops*.
- I show that feedback improves agent learning but worsens aggregator learning.
- The optimal recommendation rule for a strategic information aggregator depends on which of these two types of learning it values.
- Value on aggregator learning ⇒ don't make recommendations!
- Value on agent learning ⇒ make recommendations when confident enough about the state.
    - Weaker confidence required to correct mistakes at $t = 2$.
    - Introduces the possibility of making more mistakes than in the absence of the aggregator.