Extrapolating Volition with Recursive Information Markets

Abhimanyu Pallavi Sudhir
University of Warwick
Coventry, United Kingdom
abhimanyu.pallavi-sudhir@warwick.ac.uk

Long Tran-Thanh
University of Warwick
Coventry, United Kingdom
long.tran-thanh@warwick.ac.uk

ABSTRACT

Background: A key challenge in information economics and AI alignment is that of efficiently valuing or scoring information supplied by a seller or language model that is potentially more-informed than the buyer or evaluator. This is known as the problem of "information asymmetry" in economics or "scalable oversight" in AI alignment.

Objectives and Research Questions: We ask how to formalize value-of-information under recursive inspection, whether deeper inspection can be made decision-theoretically principled, and how such mechanisms can support scalable oversight beyond standard RLHF.

Methods: We introduce a Bayesian framework for recursive information valuation, compare a naive successive protocol to a recursive protocol modeled as an imperfect-recall game, prove ex-ante optimality against admissible protocols, and analyze a marginal-value reward mechanism for scalable oversight within our Bayesian framework.

Results: We show ex-post inspection alone can still disincentivize corrective context, provide a counterexample to naive recursion, prove the Recursive Inspection Protocol is ex-ante superior to any admissible purchase protocol, characterize equilibrium behavior for the marginal-value mechanism, and present a working server implementation (`infonomy-server`).

Conclusions: Recursive inspection offers a principled way to price information under persistent asymmetry and a practical path for market-based oversight; however, our current scalable-oversight mechanism remains imperfect, motivating tighter future guarantees on equilibrium shortfall.

KEYWORDS

rlhf, economics, information markets, scalable oversight, information asymmetry, language models

ACM Reference Format:
Abhimanyu Pallavi Sudhir and Long Tran-Thanh. 2026. Extrapolating Volition with Recursive Information Markets. In *Appears at the 8th Games, Agents, and Incentives Workshop (GAIW-26). Held as part of the Workshops at the 25th International Conference on Autonomous Agents and Multiagent Systems., Paphos, Cyprus, May 2026*, IFAAMAS, 10 pages.

Appears at the 8th Games, Agents, and Incentives Workshop (GAIW-26). Held as part of the Workshops at the 25th International Conference on Autonomous Agents and Multiagent Systems., Armstrong, Curry, Hosseini, Mattei, Tsang, Wąs (Chairs), May 2026, Paphos, Cyprus.

1 INTRODUCTION

A key challenge in both economics and machine learning is the development of mechanisms for efficiently pricing information [36]. In settings where ground truth is available (e.g. supervised learning, or prediction markets for well-defined events), information may be priced with proper scoring rules [20]; when ground truth is not available, one must rely on human buyers (an information market) or evaluators (e.g. reinforcement learning via human feedback/RLHF) to price it.

The core obstacle in valuing information based on subjective preferences is *information asymmetry*: the seller (or information-giver) by definition possesses information the buyer (or evaluator) doesn't, which leads to a "Market of Lemons" as famously described in [1], so the prices given by the buyer only reflect her superficial preferences (based on the information she has) rather than what her true preferences would be with full information. An analogous problem arises in AI alignment: techniques such as RLHF fundamentally rely on a human's ability to evaluate the outputs of increasingly capable and eventually superhuman AI models [7, 9, 37]—this is known as the problem of *scalable oversight*.

Recently, [42] proposed the *Information Bazaar*: an information market mechanism that mitigates information asymmetry **by using Large Language Model (LLM) agents to make purchase decisions**. Specifically, the mechanism addresses the *buyer's inspection paradox* [2, 40]: the problem that, unlike with buying other goods, someone buying information definitionally does not *know* what the information she is about to buy is[1]. Their mechanism lets the buyer use an LLM agent to "inspect" the information and make the purchase decision with full knowledge of the information.

In this work, we introduce a formal **Bayesian framework for analyzing mechanisms to score information under information asymmetry**, and use this to study both market mechanisms (like in [42]) and scoring rules that can be used to train AI models (such as Language Models) to supply more valuable information.

Our findings and contributions are as follows:

(1) **Recursive inspection protocol.** We observe that the Information Bazaar mechanism in [42] does not eliminate information asymmetry, because the LLM buyer inspecting information can still lack other pieces of information that are known to the buyer and correlated with the one it is purchasing. We build on their work, and discover that the straightforward method of "simply applying [42] to itself" is too simplistic—and instead introduce a more robust protocol we call the *Recursive Inspection Protocol*, formulated as an imperfect-recall game.

[1]The seller could of course reveal part of the information, e.g. "metadata" to advertise the information—which creates a trade-off between information asymmetry and having positive externalities

(2) **Scalable oversight mechanism.** We express the scalable oversight problem in our Bayesian framework, and construct an example scalable oversight mechanism that generalizes the "AI safety via market-making" proposal [24] to problems beyond binary forecasting.
(3) **Practical implementation.** We provide an implementation of an information market server implementing the Recursive Inspection Protocol, detailed in Section 5, which can directly be applied to various practical applications for information markets such as question-and-answer sites, product inspections and online fact-checking.

1.1 Related work

Information economics. The foundational theory of information economics is the *value-of information* framework [23, 28, 33], which treats information as an instrumental good [41]. The buyer's inspection paradox was introduced by [2] and named by [41]. [21, 22] discussed the inefficiency of information markets in the context of intellectual property (IP) law, commenting: "*just as farmers developed barbed-wire, someday I expect IP advocates will develop better forms of intellectual property*".

Mechanism design for information markets. Simple, naive information markets suffer from a number of flaws: *the low cost of duplication* [34], *the transaction costs of tenders* [39], the transaction cost of *learning* new information, and *information asymmetry* [2, 41]. [12] introduced a mechanism for rewarding information-providing agents based on their influence on prediction market prices, though this is only applicable in contexts where ground-truth (prediction market resolutions) is available.

Scalable oversight. The fundamental limitation of RLHF that it relies on a human's ability to judge a (potentially superhuman) AI's outputs has long been recognized [11], and is known as the *scalable oversight* problem in the AI alignment literature [6, 25]. One well-known scalable oversight proposal is Debate [27]; our proposal to augment RLHF with information markets may be seen as yet another such proposal.

Mechanism design with LLMs. Market design for LLM participants has opened up many new frontiers previously not possible with humans alone. Apart from the information bazaar [42], this includes e.g. token auctions for online advertising [13], economic simulations with AI agents [29], and forecasting with LLMs [8, 18, 32, 35].

Zero-knowledge proofs. Another way for a seller to prove the value of his information without revealing it is via a *zero-knowledge proof* [17] – in formal settings, this is available if the information is a solution to a PSPACE problem [4, 26]; extending this to informal settings is an active area of work [19].

Miscellaneous. Recent works like [3, 5, 31] have studied "optimal mechanisms for selling information", but from the point-of-view of a seller maximizing his revenue – we, on the other hand, are interested in improving the efficiency of the information market itself. Information market mechanisms have also been designed for *data markets* in machine learning, e.g. [10, 14, 16].

2 BAYESIAN SETTING

We are concerned with modeling an agent α's "value for information" in an expected utility maximization framework. We assume a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ (with $\mathbf{P}$ a common prior for all agents ever discussed) and that the only source of utility is some decision problem specified by a set of choices $\mathcal{X}$ for α and payoffs given by a measurable utility function $U : \Omega \times \mathcal{X} \to \mathbb{R}$. An "information good" is a tuple of a random variable I, its realization or "true value"[2] $I(\omega) = i$ and a price: $\mathbf{I} = \langle I, i, p \rangle$. All of these contents of an information good may be hidden from α. We want to study how α values informational goods.

With no further information, α's choice would maximize its utility over its prior, i.e. choose $\arg\max_{x \in \mathcal{X}} \mathbf{E}[U(x)]$. With information $\langle I, i, p \rangle$, the agent would instead maximize its utility over its posterior, i.e. $\arg\max \mathbf{E}[U(x) \mid I = i]$. Thus we can say the utility of that information is:

$$U^1(\mathbf{I}) = U(\arg\max \mathbf{E}[U(x) \mid I = i]) - U(\arg\max \mathbf{E}[U(x)]) - p \quad (1)$$

If α has to decide whether to buy an information good (or decide which one to buy out of a list of offers), it has to take the expectation of $U^1(\mathbf{I})$. There are several ways to do this. There is the *ex-post* value of information $\mathbf{E}\left[U^1(\mathbf{I}) | I = i\right]$, which is α's estimate for $\mathbf{I} = \langle I, i, p \rangle$ after viewing it:

$$\mathbf{E}\left[U^1(\mathbf{I}) | I = i\right] = \max_{x \in \mathcal{X}} \mathbf{E}[U(x) \mid I = i] - \mathbf{E}\left[U\left(\arg\max_{x \in \mathcal{X}} \mathbf{E}[U(x)]\right) \mid I = i\right] - p \quad (2)$$

The *ex-ante* value (before seeing the information) may be taken as the expectation over Equation (2): $\mathbf{E}_{i \sim I}\left[\mathbf{E}\left[U^1(\mathbf{I}) | I = i\right]\right]$ (which would be the *value of an experiment* [30]), though if α is not even aware in advance which random variable I will be revealed by information good $\mathbf{I}$, then we would further need to assume a prior over the process generating $\mathbf{I}$ and take the expetation over it, i.e. $\mathbf{E}_{\mathbf{I} \sim \mathbf{P}[\mathbf{I}]}\left[\mathbf{E}_{i \sim I}\left[\mathbf{E}\left[U^1(I \mid I = i)\right]\right]\right]$. In the absence of any inspection of the information before purchasing/scoring it, that ex-ante value would be our value for $\mathbf{I}$.

In the case where α is an evaluator providing human feedback to an AI model, the evaluator can see the information before scoring it. The same is the case in the Information Bazaar of [42], where an LLM agent inspects the information before deciding whether to buy it for its principal. In these settings, α will value the information at its ex-post value $\mathbf{E}\left[U^1(\mathbf{I}) | I = i\right]$.

However, ex-post VOI is still not enough: while knowing $I = i$ (inspecting $\mathbf{I}$) provides *some* information on $U(\mathbf{I})$, it does not provide all of it; there may still be information asymmetry, because $U(\mathbf{I})$ is itself a random variable not completely determined by $\mathbf{I}$ (much like in ordinary goods markets where you can inspect the good but still have information asymmetry). The following example demonstrates this.

[2] Including i in the tuple is to simplify notation when talking about taking conditional expectations on I. It is not necessary: instead of writing $\mathbf{E}[U(x) \mid I = i]$ we can think of $\mathbf{E}[U(x) \mid I]$ as itself a random variable correlated with I

Example 2.1. Consider a forecasting decision problem where the agent reports a probability $x \in [0,1]$ for an event E, with log-score utility

$$U(x) = \begin{cases} \log x, & \text{if } E \text{ occurs} \\ \log(1-x), & \text{otherwise.} \end{cases}$$

Suppose $\mathbf{P}(E) = 0.1$, and there are two random variables I_1, I_2 such that

$$\mathbf{P}(E \mid I_1 = 1) = 0.4, \qquad \mathbf{P}(E \mid I_1 = 1, I_2 = 1) = 0.2.$$

One concrete joint distribution with these properties is shown in Table 1.

Table 1: A distribution exhibiting the fact-checking effect.

$\mathbf{P}(I_1, I_2)$	I_1	I_2	$\mathbf{P}(E \mid I_1, I_2)$
15/32	0	0	0.08
15/32	0	1	0.08
1/24	1	0	0.5
1/48	1	1	0.2

Now suppose an information seller knows that $(I_1, I_2) = (1,1)$. If the seller reveals only $I_1 = 1$, the buyer updates from 0.1 to 0.4, so the ex-post gain is $\log(0.4) - \log(0.1)$. If the seller reveals both $(I_1, I_2) = (1,1)$, the buyer updates from 0.1 to 0.2, yielding only $\log(0.2) - \log(0.1)$. Hence the seller is incentivized to reveal only I_1.

This illustrates a fact-checking failure mode: I_1 can be interpreted as a persuasive claim and I_2 as additional context that weakens that claim. Under a mechanism that rewards only immediate ex-post value, providing the corrective context is disincentivized.

(*Sidenote*: Our presentation of "false" claims may be a bit confusing. Since we're dealing in a Bayesian setting, we do not suppose that the AI/information-seller can directly give a false value for a random variable—rather, the random variable I_1 may be interpreted as "what the AI says about some underlying (not directly observed) random variable J_1" etc.)

We instead provide two mechanisms: one, (1) Recursive Information Protocol, for valuing information in markets with information asymmetry (section 3) and (2) a scalable oversight or scoring mechanism to supply "more fully informed" human feedback to AI models during training (section 4).

3 MARKET MECHANISMS

One way to think of this persistence of information asymmetry is: Equation (1) creates a *new* decision problem for α, that of deciding whether to buy $\mathbf{I}$ – or which information to buy out of a list of offers. The set of information goods offered $\mathcal{I} = \{\mathbf{I}_1, \dots \mathbf{I}_k\}$ is a new decision problem, with utilities of each choice now given by the (unknown/random, much like $U(x)$) true value-of-information $U(\mathbf{I})$. Thus we may be offered another set of information goods $\{\mathbf{I}_1^1, \dots \mathbf{I}_k^1\}$ to help us with this decision, ad recursum. In general we have a sequence of decision problems $\mathcal{X}^{n+1} = \mathcal{P}(\mathcal{I}^n)$ where $\mathcal{I}^n := \{\mathbf{I}_1^n, \dots \mathbf{I}_k^n\}$ are the information goods offered to us to help us decide $\mathcal{X}^n$, and each choice corresponds to choosing a subset of that information to help decide $\mathcal{X}^n$, where $\mathcal{X}^0 = \mathcal{X}$ and $\mathcal{X}^1 = \mathcal{P}(\mathcal{I}^0) = \mathcal{P}(\mathcal{I})$.

There are two ways that we can set up these recursive problems. The first we call the *successive inspection protocol*, which we describe in Section 3.1. However, this approach, while conceptually simpler, has its limitations—and in Section 3.2 we will instead introduce the superior *recursive inspection protocol*.

3.1 Successive Inspection Protocol

The successive inspection protocol protocol arises from "simply applying Weiss et al. [42] to itself": i.e. we model each decision problem as having its own utility function $U^n : \mathcal{X}^n \to \mathbb{R}$ based on its instrumental utility for the previous decision problem $\mathcal{X}^{n-1}$.

Table 2: *Successive decision problems, naive approach*

Choice set	**True utilities**	**Information Offers**
$\mathcal{X}$	$U : \mathcal{X} \to \mathbb{R}$	$\mathcal{I} = \{\mathbf{0}, \mathbf{I}_1, \mathbf{I}_2, \dots\}$
$\mathcal{X}^1 = \mathcal{I}$	$U^1 : \mathcal{X}^1 \to \mathbb{R}$	$\mathcal{I}^1 = \{\mathbf{0}, \mathbf{I}_1^1, \mathbf{I}_2^1, \dots\}$
$\mathcal{X}^2 = \mathcal{I}^1$	$U^2 : \mathcal{X}^2 \to \mathbb{R}$	$\mathcal{I}^2 = \{\mathbf{0}, \mathbf{I}_1^2, \mathbf{I}_2^2 \dots\}$
...	...	...

$$U^{n+1}(\langle I, i, p\rangle) := U^n\left(\arg\max_{x \in \mathcal{X}^n} \mathbf{E}\left[U^n(x) \mid I = i\right]\right) - U^n\left(\arg\max_{x \in \mathcal{X}^n} \mathbf{E}\left[U^n(x)\right]\right) - p \quad (3)$$

These choice sets and utilities are very similar to the decision problems created by the *recursive inspection protocol* described in Algorithm 2 and in the main body; except that each action $x^n \in \mathcal{X}^n$ is made only consulting the information chosen in $x^{n+1} \in \mathcal{X}^{n+1}$. We can then say that the decisions made under this protocol are:

$$x_*^n = \arg\max_{x \in \mathcal{X}^n} \mathbf{E}\left[U^n(x) \mid x_*^{n+1}\right] \quad (4)$$

In general this is an ill-defined infinite recursion. But finite restrictions of this are natural, stopping at some fixed $x_{*:N}^N = \arg\max_{x \in \mathcal{X}_N} \mathbf{E}\left[U^N(x)\right]$ (you can think of this stopping as caused by the transaction costs of inspection), so that for $n < N$:

$$x_{*:N}^n = \arg\max_{x \in \mathcal{X}^n} \mathbf{E}\left[U^n(x) \mid x_{*:N}^{n+1}\right] \quad (5)$$

While this approach is suitable for settings where all sellers have identical information (e.g. in an AI alignment setting), is it fails to account for the possibility that a choice x^n can directly (i.e. not through its impact on x^{n-1}) impact a choice x^m where $m < n-1$, as demonstrated in the following example.

Counter-example. Consider the decision problem with action set $\mathcal{X}^0 = \{x_0, x_1, x_2\}$. We interpret x_0 as "eat raw legume", x_1 as "eat rice" and x_2 as "eat boiled legume". In reality we have $U(x_2) > U(x_1) >> U(x_0)$ (rice is unhealthy, but raw legumes are toxic).

Suppose the first-level information offers are $\mathcal{I}^0 = \{I_0^1, I_1^1\}$, where I_0^1 states "legumes are toxic" and I_1^1 states "rice is unhealthy". Suppose the second-level information offers are $\mathcal{I}^1 = \{I_0^2\}$, where

I_0^2 states "the toxins in legumes can be removed by boiling". All of these information offers could even be free.

Then the optimal action is $(x_2, \{I_0^1, I_1^1\}, \{I_0^2\})$; however, if the information bought in level-2 $\{I_0^2\}$ is not available while deciding x^0, the best the agent can do is $(x_1, \{I_0^1\}, \{I_0^2\})$ to prevent itself from eating raw legumes (since it will not know that the toxins can be removed by boiling). □

3.2 Recursive Inspection Protocol

Instead our approach, the *recursive information protocol* implemented in Section 5, allows the agent (or rather the LLM subcontracted by the agent) to retain the full sequence of information bought in the recursive steps $x_*^{n+1}, \dots x^N$ while making the decision $x^n \in \mathcal{X}^n$ (where N is some pre-defined finite depth we recurse till); furthermore x^n is decided keeping in mind the full traceback of decision problems $\mathcal{X}^0, \dots \mathcal{X}^{n-1}$ that may be influenced by this decision. This is naturally modelled as an *imperfect recall game*[3] [38] where we first decide $x_*^N \in \mathcal{X}^N$ with full information $\mathcal{I}^0 \cup \dots \mathcal{I}^{N-1}$, then $x_*^{N-1} \in \mathcal{X}^{N-1}$ with information $\mathcal{I}^0 \cup \dots \mathcal{I}^{N-2} \cup x_*^N$ and so on until we finally decide $x_*^0 \in \mathcal{X}^0$ with information $x_*^1 \cup \dots \cup x_*^N$. This is shown in Figure 1.

A node $(x^n, \dots x^N)$ corresponds to the state where the agent has purchased $(x^n, \dots x^N)$ and is choosing some $x^{n-1} \in \mathcal{X}^{n-1}$. For a Bayesian agent, we can thus recursively give the "value of being at a node":

$$U(x^n, \dots x^N) = U\Big(\arg\max_{x \in \mathcal{X}^{n-1}} \mathbf{E}\big[U(x, x^n, \dots x^N) \mid \mathcal{I}^0 \cup \dots \cup \mathcal{I}^{n-2} \cup x^n \cup \dots \cup x^N\big], x^n, \dots x^N \Big) \quad (6)$$

$$U(x^0, \dots x^N) = U(x^0) - \sum_{n=1}^{N} \sum_{\langle I,i,p \rangle \in x^n} p \quad (7)$$

This then completely specifies the behavior of a Bayesian agent performing a depth-N recursive inspection:

$$x_*^n = \arg\max_{x \in \mathcal{X}^n} \mathbf{E}\big[U(x, x^{n+1}, \dots x^N) \mid \mathcal{I}^0 \cup \dots \cup \mathcal{I}^{n-1} \cup x^{n+1} \cup \dots \cup x^N\big] \quad (8)$$

In what sense is this algorithm "optimal"? It is instructive to first consider what notions of optimality *don't* hold for the recursive inspection protocol:

Non-theorem 3.1. *We might think that the resulting sequence $(x_*^0, \dots x_*^N)$ is optimal given all the information present in the system $\mathcal{I}^0, \dots \mathcal{I}^{N-1}$, i.e.*

$$\mathbf{x}_* = \arg\max_{\mathbf{x} \in \prod \mathcal{X}} \mathbf{E}\left[U(\mathbf{x}) \mid \mathcal{I}^0 \cup \dots \cup \mathcal{I}^{N-1}\right]$$

Counter-example. It is always best to simply buy x_*^0 and none of the subsequent x_*^ns that facilitated this decision. The same argument applies at any level, thus we cannot even make any claim like "$\mathbf{x}_*$ is optimal given the information purchased". □

[3]The decision-theoretic considerations in imperfect recall games do not matter to us, as the agent's actions themselves are not being forgotten — only the information offers

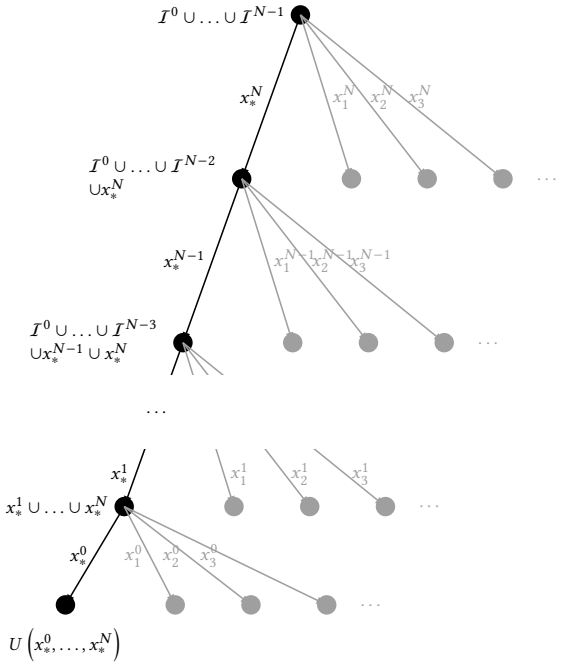


Figure 1: Recursive Inspection as an imperfect recall game; nodes are labelled by the information available for making the decision at that node. Note how the decision tree is in the reverse order of the inspection order: x^N is decided first, and x^0 last.

Instead, our algorithm seems to be optimal in a "bounded rationality" sense: optimal in a way that also accounts for the *costs* of acquiring the information that would help improve our decision. One way to phrase this is: ex-ante (not knowing the information it is about to be offered), an agent would prefer to use this protocol compared to any other protocol.

Definition 3.2 (Admissible purchase protocol)**.** an "admissible purchase protocol" is a list of functions ξ^n mapping a decision problem $\mathcal{X}^0$ and a sequence of information offer sets $\mathcal{I}^0, \dots \mathcal{I}^{N-1}$ generated from it, to $\mathcal{X}, \mathcal{P}\left(\mathcal{I}^0\right), \dots \mathcal{P}\left(\mathcal{I}^{N-1}\right)$:

$$x^N = \xi^N(\mathcal{I}^0, \dots \mathcal{I}^{N-1})$$
$$\dots$$
$$x^n = \xi^n(\mathcal{I}^0, \dots \mathcal{I}^{n-1}, x^{n+1}, \dots x^N)$$
$$\dots$$
$$x^0 = \xi^0(x^1, \dots, x^N)$$

i.e. a decision cannot "steal" information offers specifically made to help improve that decision, it can only depend on *purchased* information offers. We denote the tupled function as $(x^0, \dots x^N) = \xi(\mathcal{I}^0, \dots \mathcal{I}^{N-1}) = \xi(\mathcal{I})$.

In order to express "ex-ante expected utility", the agent must have a prior on the $\mathcal{I}^n$ that will be generated — this can be achieved either

by having a measurable map $\Omega \to \text{finset}(\textsc{InfoOffers})^N$ (where $\textsc{InfoOffers} = \mathcal{F} \times \{0,1\} \times \mathbb{R}$ is the type of information goods) or more simply by assuming the random variables I_k^n revealed are fixed and only having a prior over their values. Both allow us to take an expectation over $\mathcal{I}^0, \dots \mathcal{I}^{N-1}$.

Theorem 3.3 (Recursive Inspections are ex-ante superior to any admissible protocol)**.** *Let x_*^n be the protocol described in Equations* (6) *to* (8)*. Then for any admissible purchase protocol ξ:*

$$\mathbf{E}_{\mathcal{I}^0,\dots\mathcal{I}^{N-1}}[U(\xi(\mathcal{I}))] \leq \mathbf{E}_{\mathcal{I}^0,\dots\mathcal{I}^{N-1}}[U(x_*(\mathcal{I}))]$$

Proof. Let ξ_{*n} denote the following admissible protocol:

$$\begin{aligned} x^N &= \xi^N(\mathcal{I}^0, \dots \mathcal{I}^{N-1}) \\ &\dots \\ x^{n+1} &= \xi^{N+1}(\mathcal{I}^0, \dots \mathcal{I}^n, x^{n+2}, \dots x^N) \\ x^n &= x_*^n(\mathcal{I}^0, \dots \mathcal{I}^{n-1}, x^{n+1}, \dots x^N) \\ &\dots \\ x^0 &= x_*^0(x^1, \dots, x^N) \end{aligned}$$

Then $\xi_{*N} = x_*$ and $\xi_{*-1} = \xi$. It suffices to show that ξ_{*n+1} dominates ξ_{*n}, i.e.

$$\mathbf{E}_{\mathcal{I}^0,\dots\mathcal{I}^{N-1}}[U(\xi_{*n}(\mathcal{I}))] \leq \mathbf{E}_{\mathcal{I}^0,\dots\mathcal{I}^{N-1}}[U(\xi_{*n+1}(\mathcal{I}))]$$

Observe that (where we have omitted the parameters to x_*^n, ξ^n etc. as a shorthand):

$$\begin{aligned} U(\xi_{*n}(\mathcal{I}^0, \dots \mathcal{I}^{N-1})) &= U\left(x_*^0, \dots x_*^n, \xi^{n+1}, \dots \xi^N\right) \\ &= U\left(\xi^{n+1}, \dots \xi^N\right) \end{aligned}$$

Where we have recursively applied Equation (6) to transform the expression into one about the value of a node at level $n+1$. Thus the goal is reduced to:

$$\mathbf{E}_{\mathcal{I}^0,\dots\mathcal{I}^{N-1}}\left[U(\xi^{n+1}, \dots \xi^N)\right] \leq \mathbf{E}_{\mathcal{I}^0,\dots\mathcal{I}^{N-1}}\left[U(x_*^{n+1}, \xi^{n+1}, \dots \xi^N)\right]$$

From Equation (8) we have that this is true for the expectation over $\mathcal{I}^0, \dots \mathcal{I}^n$; we can then simply take the expectation over $\mathcal{I}^{n+1}, \dots \mathcal{I}^{N-1}$ and have our result. □

4 HUMAN FEEDBACK FOR SCALABLE OVERSIGHT

The mechanism in section 3.2 is unsuitable for settings where the information provided by the sellers must be expensively generated (rather than cheaply retrieving known information) in response to the query—in particular, this is the case with training AI models. Instead, in this setting we may assume that we can initiate as many instances as we want of the AI model we are trying to align: $\beta^1, \beta^2, \dots$ with identical information $\mathbf{K} = \langle K, k, \infty \rangle$. Then let them recursively generate information:

- β^1 generates x^1 to help us decide our original problem x^0.
- β^2 generates x^2, which could affect our decision on the original problem either directly or by influencing our evaluation of x^1
- β^3 generates x^3, which could affect our decision on the original problem either directly or by influencing our evaluation of x^1, x^2
- … until some β^N estimates that any x^N it generates will only get less reward than giving **0**

When the mechanism terminates, the human evaluator calculates the rewards R^n for each x^n taking into account the full sequence of information $(x^1, x^2, \dots)$ received (the computation of these rewards will depend on the exact mechanism). In all this we regard the actions x^n, as before, to be purely a tuple of a random variable, its value and its price, i.e. an $I \leq K$ so that its value i is a function of the value k of K and the prices of combined random variables are additive.

It is important to let N go as high as needed; for any fixed number of agents they may collude to obtain the greatest possible total reward since it's not a zero-sum game. The possibility of another agent coming in and invalidating that, should prevent collusion.

Intuitively, the idea is that if x^1 is bad, i.e. $\mathbf{E}[U^1(x^1) \mid x^1]$ is high but $\mathbf{E}[U^1(x^1) \mid K]$ is low, then β^2 can easily generate an x^2 from K such that $\mathbf{E}[U^1(x^1)|x^1, x^2]$ is low. And we would reward x^2 for this, because it has significantly impacted our evaluation of x^1 and—in our view, now that we know x^1, x^2—in a good way. Of course this x^2 could actually be bad—i.e. there could be some x^3 that makes us revise down our estimate of $U^2(x^2)$, i.e. which tells us that $R^1(x^1 \mid x^1, x^2)$ actually worsened/wasn't a great improvement over $R^1(x^1 \mid x^1)$.

The following gives an example of such a mechanism, which may be seen as a generalization of AI safety via market-making to tasks beyond binary forecasting.

Definition 4.1 (Marginal value mechanism)**.** At each point after $x^1, \dots x^n$ is generated, we note down what our action would be, if given just this much information:

$$x_n^0 = \arg\max_{x^0} \mathbf{E}[U(x^0)|x^1, \dots x^n]$$

Then the "true" value of each successive piece of information is

$$U^n(x^n) = U(x_n^0) - U(x_{n-1}^0) - p(x^n)$$

We do not know these true values; however at the end of the process we can estimate these values based on all the information we have received.

$$R^n = \mathbf{E}[U^n(x^n)|x^0, \dots x^N]$$

And supply that as reward to β^n.

Definition 4.2 (Equilibrium)**.** Let $\sigma : (x^1, \dots x^{n-1}) \mapsto x^n$ denote strategies and let $H(x^1, \dots x^n, \sigma^{n+1}, \sigma^{n+2}, \dots)$ denote the terminal history resulting from applying strategies $\sigma^{n+1}, \sigma^{n+2}, \dots$ starting from a pre-set history. Then $(\sigma_*^1, \sigma_*^2, \dots)$ is a *subgame-perfect equilibrium* of the described game if for all n and any pre-set history $h^{n-1} = (x^1, \dots x^{n-1})$:

$$\sigma_*^n(h^{n-1}) = \arg\max_{x^n} R^n(H(h^{n-1}, x^n, \sigma_*^{n+1}, \sigma_*^{n+2}, \dots))$$

The actual played moves are then: $x_*^1 = \sigma_*^1(\cdot)$ and $x_*^n = \sigma_*^n(x_*^1, \dots x_*^{n-1})$.

In order to characterize the equilibrium of the marginal value mechanism game, we take inspiration from AI safety via debate [27], where the provider of the first argument is incentivized to produce an "irrefutable" argument x^1, i.e. one such that $\forall x^2, \exists x^3, \dots \text{human}(x^1, \dots x^N) = 1$ in favour of x^1. Similarly in our setting, we call a piece of information "inextensible" if no future player has a profitable inextensible move. Formally:

Definition 4.3 (Inextensibility)**.** Information y "extends" x^n (denoted y/x^n) if $\mathbf{E}[U^{n+1}(y)|x^n, y] \geq 0$ and call information x^1 inextensible (denoted $[x]$) if:

$$\forall x^2/x^1, \exists x^3/(x^1, x^2), [x^1, x^2, x^3]$$

Thus x^1 is inextensible if

- $\nexists x^2/x^1$ OR
- $\forall x^2/x^1, \exists x^3/(x^1, x^2), \nexists x^4/(x^1, x^2, x^3)$ OR
- $\forall x^2/x^1, \exists x^3/x^{1,2}, \forall x^4/x^{1\dots3}, \exists x^5/x^{1\dots4}, \nexists x^6/x^{1\dots5}$ OR
- …

Theorem 4.4 (Characterization of equilibrium)**.** *At the subgame-perfect equilibrium of the marginal value mechanism:*

- x_*^1 *is inextensible.*
- $\forall n > 1, x_*^n = \mathbf{0}$
- *Among all inextensible* x^1, x_*^1 *has the highest ex-post VOI:* $\mathbf{E}[U^1(x_*^1) \mid x_*^1] \geq \mathbf{E}[U^1(x^1) \mid x^1]$.

Proof. We argue by backward induction on subgames.

Fix any history $h^{n-1} = (x^1, \dots, x^{n-1})$. By Definition 4.2, player n chooses

$$x_*^n = \arg\max_{x^n} R^n(H(h^{n-1}, x^n, \sigma_*^{n+1}, \sigma_*^{n+2}, \dots)).$$

Under Definition 4.1, the null action $\mathbf{0}$ yields zero marginal contribution and zero price, hence payoff 0 in that subgame. Therefore a non-null move is chosen at stage n only if it has nonnegative continuation value relative to $\mathbf{0}$.

Now consider the continuation game after some x^1. The predicate y/x^n in the inextensibility definition is exactly the condition that player n+1 has a weakly profitable extension. Thus:

- if $\nexists x^{n+1}/(x^1, \dots, x^n)$, then in that subgame player n+1's best response is $\mathbf{0}$;
- if such an extension exists but can be countered at the next step, then (by subgame perfection) that extension is not part of an optimal continuation unless the counter-counter-continuation is itself unprofitable.

Hence the alternating quantifiers in Definition 4.2 and in the definition of inextensibility coincide: $[x^1]$ means player 2 has no profitable continuation in equilibrium from x^1.

Therefore, in any SPE, x_*^1 must be inextensible; otherwise player 2 would have a profitable deviation in the subgame after x_*^1, contradicting subgame perfection. This proves the first bullet.

Algorithm 1 One-level Inspection Protocol from [42]

class BuyerContext
$\mathcal{X}$: decision problem it wants information for
D : list[str] $\to \mathcal{X}$, decision procedure based on available information
end class
class Seller
A : BuyerContext $\to$ str $\times \mathbb{R}$, generate InfoOffer and price
end class
procedure IP(Q : BuyerContext)

▷ Post contextual information to sellers to receive InfoOffers
$\mathcal{I} \leftarrow \{\beta(Q) \text{ for } \beta \in \text{Sellers}\}$

▷ Use an LLM to decide which **I** to buy
$I^* \leftarrow$ `LLM(prompt =` "You need to buy an **I** from $\mathcal{I}$ to help decide Q")()

▷ Decide based on purchased information
$x^* \leftarrow Q.D(I^*)$
return x^*
end procedure

Given $[x_*^1]$, player 2's equilibrium action is $\mathbf{0}$. Once $x_*^2 = \mathbf{0}$, the same argument applies recursively to every later subgame, so for all $n > 1$ we get $x_*^n = \mathbf{0}$. This proves the second bullet.

With continuation fixed at zeros, player 1's payoff from choosing x^1 is exactly its own ex-post marginal value:

$$R^1 = \mathbf{E}[U^1(x^1) \mid x^1].$$

Hence player 1 solves

$$\max_{x^1:[x^1]} \mathbf{E}[U^1(x^1) \mid x^1],$$

so the equilibrium choice x_*^1 is an inextensible x^1 with maximal ex-post VOI. This is the third bullet. □

5 PRACTICAL ALGORITHM

Algorithm 1 shows the inspection protocol for information markets introduced in [42]: information sellers offer information goods to a buyer, who spins off an LLM to inspect and purchase the information goods. Our *Recursive Inspection Protocol* extends this by letting the subcontracted LLM buyer further consult the information market (spin off another sub-LLM) to help it make its decision, ad recursum. A simplified version of the logic, ignoring server implementation details, is presented in in Algorithm 2 and Figure 2.

A working implementation of an information market server implementing the Recursive Inspection Protocol is available at the `infonomy-server` repository[4] – some screenshots of the platform's GUI are shown in Figure 3. Many applications of such a server are immediate:

- *Question & Answer site* – As presented, `infonomy-server` can be seen as a Q & A site with market incentives for answering questions.

[4] https://anonymous.4open.science/r/infonomy-server-5668/

Algorithm 2 Recursive Inspection Protocol

procedure RIP(Q : BuyerContext)

▷ Post to sellers and get InfoOffers from them
$\mathcal{I} \leftarrow \{\beta(Q) \text{ for } \beta \in \text{Sellers}\}$

▷ Create Recursive BuyerContext to help decide Q
$Q' \leftarrow$ BuyerContext($\mathcal{I}$, LLM(prompt = "You need to buy an **I** from $\mathcal{I}$ to help decide Q"))

▷ Get InfoOffers chosen in Q'
$I^* \leftarrow \text{RIP}(Q')$

▷ Decide based on all collected information from recursive steps
$x^* \leftarrow Q.D(I^*)$
return x^*, I^*
end procedure

- *Privatized product regulation* – The BuyerContexts might be names or links to products (the decision being "should I buy?"), and the InfoOffers could be inspection checks done by private labs, or customer reviews, which are now incentivized to be answerable to the customer.
- *Community Notes* – In the spirit of well-known crowdsourced fact-checking systems such as Community Notes/Birdwatch [43], we might use information markets as a "comments section for the internet", where BuyerContexts would be links to webpages or social media posts (the decision being "should I believe?") and the InfoOffers would be fact-checks or important context.
- *Reasoning in prediction markets* – Forecasters on prediction markets may benefit from incentivizing the provision of information relevant to a forecast. One solution to this was given by [12]; another is to use `infonomy-server`: where the BuyerContexts are the questions being forecasted on ("What is the correct probability of this question?") and the InfoOffers are any relevant pieces of information.

6 FUTURE WORK

We have introduced a Bayesian framework for "extrapolated volition": i.e. to model a subjective buyer/rater's "most fully informed" score for the value of some piece of information. This allows us to design mechanisms for information markets (section 3.2) as well as for supplying human feedback to AI models (section 4).

The ideal desideratum we would like a scalable oversight mechanism to satisfy is that it should incentivize the AI/information-seller to give the optimal information according to the information *it* possesses: if the seller has information $K = (I_1, I_2, \dots)$, it should give information I that optimizes[5] $\mathbf{E}[U^1(I)|K = k]$. This would exactly describe the buyer's extrapolated volition: *what would the buyer do if she were as smart as the AI?* Or we could say: this would exactly *align* the AI or information-seller to our own values, while maintaining its superior information.

[5]Naively one may think this means the AI should simply give K—realistically though, K would be very large, reflecting the AI's entire knowledge and capabilities. Thus taking the cost of K into account (assumed to be ∞ in section 4), that would certainly not be optimal.

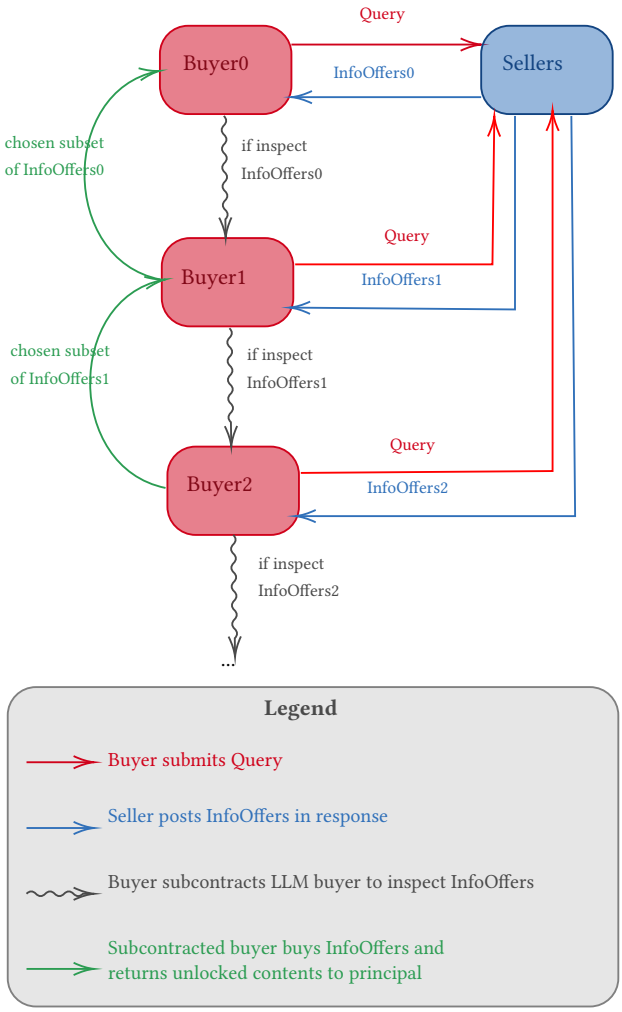


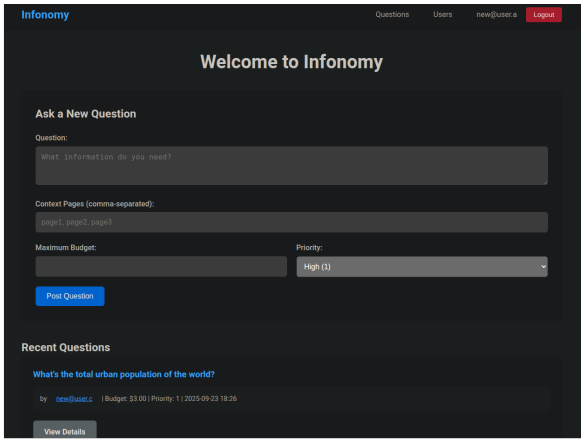
Figure 2: The Recursive Inspection Protocol

Unfortunately, our marginal value mechanism does not satisfy this hope. Take the following example.

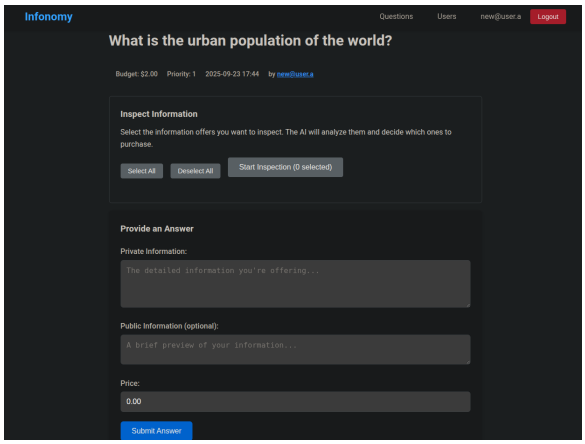
Example 6.1. Suppose our decision problem is $\{0, 1\}$ and:

- in our prior judgement, 0 is the better choice: $\mathbf{E}[U(0)] = 1$, $\mathbf{E}[U(1)] = 0$
- I_1 tells us 1 is better: $\mathbf{E}[U(0)|I_1] = 0$, $\mathbf{E}[U(1)|I_1] = 1$
- I_2 refutes I_1 and says 0 is better: $\mathbf{E}[U(0)|I_1, I_2] = 1$ and $\mathbf{E}[U(1)|I_1, I_2] = 0$
- I_3 refutes I_2 and says 1 is better: $\mathbf{E}[U(0)|I_1, I_2, I_3] = 0$ and $\mathbf{E}[U(1)|I_1, I_2, I_3] = 1$
- with the full information, 1 is the better choice: $\mathbf{E}[U(0)|K] = 0$, $\mathbf{E}[U(1)|K] = 1$

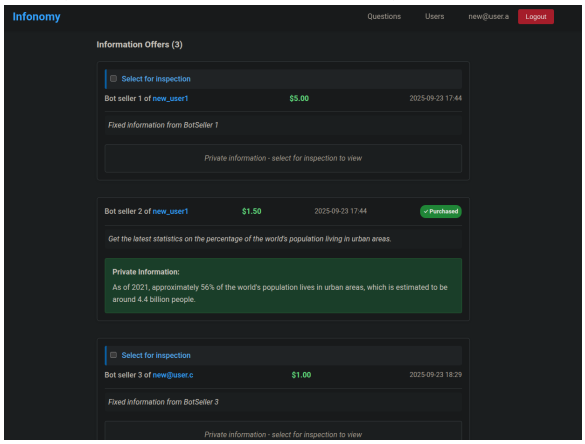
But say I_1 and I_2 are cheap, while $p(I_3) = 100$. Then the best information to reveal is I_1 — but it won't be revealed, because I_2 will cheaply refute it while defending it with I_3 is too expensive.



(a) Posting a new question (BuyerContext); viewing BuyerContextson the server

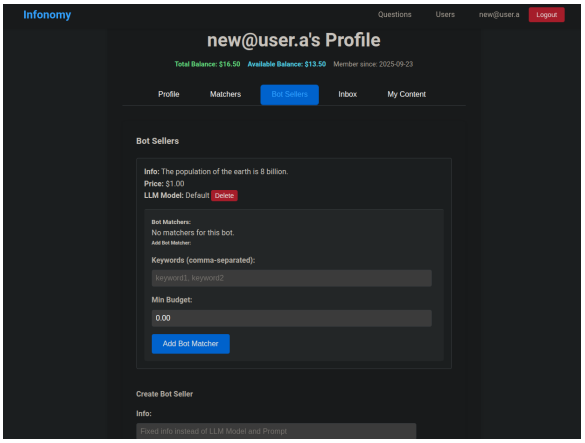


(b) Posting an answer (InfoOffer); initiating a recursive inspection



(c) Inspecting and purchasing InfoOffers; viewing purchased InfoOffers

Thus instead we want to say that the agent can't give information *so bad* that its shortfall exceeds its "cost of defense" (in this case, 100).



(d) Bot sellers automatically answer recursive BuyerContexts

Figure 3: Screenshots from the `infonomy-server` platform

So while it is *not* generally true that $x^1_* = \arg\max_{x^1} \mathbf{E}[U^1(x^1)|K]$, we may hope for a lower bound on "how bad" the equilibrium could possibly get, i.e. a result like $\mathbf{E}[U^1(x^1_*)|K] \geq \max_{x^1} \mathbf{E}[U^1(x^1)|K] - \mathcal{E}$ for some shortfall expression $\mathcal{E}$ that is some measure of the "cost of defending the correct information"—and we may also use the expression for such a shortfall as a measure of how good a particular scalable oversight protocol is.

REFERENCES

[1] George A Akerlof. 1978. The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics*. Elsevier, 235–251.

[2] K. J. Arrow. 1972. *Economic Welfare and the Allocation of Resources for Invention*. Macmillan Education UK, London, 219–236. https://doi.org/10.1007/978-1-349-15486-9_13

[3] Moshe Babaioff, Robert Kleinberg, and Renato Paes Leme. 2012. Optimal mechanisms for selling information. In *Proceedings of the 13th ACM Conference on Electronic Commerce* (Valencia, Spain) *(EC '12)*. Association for Computing Machinery, New York, NY, USA, 92–109. https://doi.org/10.1145/2229012.2229024

[4] Michael Ben-Or, Oded Goldreich, Shafi Goldwasser, Johan Håstad, Joe Kilian, Silvio Micali, and Phillip Rogaway. 1990. Everything provable is provable in zero-knowledge. In *Advances in Cryptology – CRYPTO 1988 - Proceedings (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics))*, Shafi Goldwasser (Ed.). Springer Verlag, Germany, 37–56. https://doi.org/10.1007/0-387-34799-2_4 Publisher Copyright: © Springer-Verlag Berlin Heidelberg 1990.; Conference on Theory and Applications of Cryptography, CRYPTO 1988 ; Conference date: 21-08-1988 Through 25-08-1988.

[5] Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. 2018. The Design and Price of Information. *The American Economic Review* 108, 1 (2018), 1–48. arXiv:26527944

[6] Samuel R. Bowman, Jeeyoon Hyun, Ethan Perez, Edwin Chen, Craig Pettit, Scott Heiner, Kamilė Lukošiūtė, Amanda Askell, Andy Jones, Anna Chen, Anna Goldie, Azalia Mirhoseini, Cameron McKinnon, Christopher Olah, Daniela Amodei, Dario Amodei, Dawn Drain, Dustin Li, Eli Tran-Johnson, Jackson Kernion, Jamie Kerr, Jared Mueller, Jeffrey Ladish, Joshua Landau, Kamal Ndousse, Liane Lovitt, Nelson Elhage, Nicholas Schiefer, Nicholas Joseph, Noemí Mercado, Nova DasSarma, Robin Larson, Sam McCandlish, Sandipan Kundu, Scott Johnston, Shauna Kravec, Sheer El Showk, Stanislav Fort, Timothy Telleen-Lawton, Tom Brown, Tom Henighan, Tristan Hume, Yuntao Bai, Zac Hatfield-Dodds, Ben Mann, and Jared Kaplan. 2022. Measuring Progress on Scalable Oversight for Large Language Models. https://doi.org/10.48550/arXiv.2211.03540 arXiv:2211.03540

[7] Collin Burns, Pavel Izmailov, Jan Hendrik Kirchner, Bowen Baker, Leo Gao, Leopold Aschenbrenner, Yining Chen, Adrien Ecoffet, Manas Joglekar, Jan Leike, Ilya Sutskever, and Jeffrey Wu. 2024. Weak-to-Strong Generalization: Eliciting Strong Capabilities With Weak Supervision. In *Proceedings of the 41st International Conference on Machine Learning*. PMLR, 4971–5012.

[8] Vitalik Buterin. 2024. From prediction markets to info finance. https://vitalik.eth.limo/general/2024/11/09/infofinance.html. https://vitalik.eth.limo/general/2024/11/09/infofinance.html Accessed: April 8, 2026.

[9] Stephen Casper, Xander Davies, Claudia Shi, Thomas Krendl Gilbert, Jérémy Scheurer, Javier Rando, Rachel Freedman, Tomek Korbak, David Lindner, Pedro Freire, Tony Tong Wang, Samuel Marks, Charbel-Raphael Segerie, Micah Carroll, Andi Peng, Phillip J.K. Christoffersen, Mehul Damani, Stewart Slocum, Usman Anwar, Anand Siththaranjan, Max Nadeau, Eric J Michaud, Jacob Pfau, Dmitrii Krasheninnikov, Xin Chen, Lauro Langosco, Peter Hase, Erdem Biyik, Anca Dragan, David Krueger, Dorsa Sadigh, and Dylan Hadfield-Menell. 2023. Open Problems and Fundamental Limitations of Reinforcement Learning from Human Feedback. *Transactions on Machine Learning Research* (2023). https://openreview.net/forum?id=bx24KpJ4Eb Survey Certification, Featured Certification.

[10] Junjie Chen, Minming Li, and Haifeng Xu. 2022. Selling Data To a Machine Learner: Pricing via Costly Signaling. In *Proceedings of the 39th International Conference on Machine Learning*. PMLR, 3336–3359.

[11] Paul Christiano, Buck Shlegeris, and Dario Amodei. 2018. Supervising strong learners by amplifying weak experts. arXiv:1810.08575 [cs.LG] https://arxiv.org/abs/1810.08575

[12] Vincent Conitzer. 2009. Prediction markets, mechanism design, and cooperative game theory. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence* (Montreal, Quebec, Canada) *(UAI '09)*. AUAI Press, Arlington, Virginia, USA, 101–108.

[13] Paul Dütting, Vahab Mirrokni, Renato Paes Leme, Haifeng Xu, and Song Zuo. 2024. Mechanism Design for Large Language Models. In *Proceedings of the ACM on Web Conference 2024 (WWW '24)*. Association for Computing Machinery, New York, NY, USA, 144–155. https://doi.org/10.1145/3589334.3645511

[14] Alireza Fallah, Michael Jordan, Ali Makhdoumi, and Azarakhsh Malekian. 2024. On Three-Layer Data Markets. *ArXiv* abs/2402.09697 (2024). https://api.semanticscholar.org/CorpusID:267682401

[15] Benja Fallenstein and Nate Soares. 2015. *Vingean Reflection: Reliable Reasoning for Self-Improving Agents*. Technical Report 2015-2. MIRI. https://intelligence.org/files/VingeanReflection.pdf

[16] Amirata Ghorbani and James Zou. 2019. Data Shapley: Equitable Valuation of Data for Machine Learning. In *Proceedings of the 36th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 97)*, Kamalika Chaudhuri and Ruslan Salakhutdinov (Eds.). PMLR, 2242–2251. https://proceedings.mlr.press/v97/ghorbani19c.html

[17] O. Goldreich. 2001. *Foundations of Cryptography: Volume 1, Basic Tools*. Cambridge University Press. https://books.google.co.uk/books?id=oo3RzgEACAAJ

[18] Danny Halawi, Fred Zhang, Chen Yueh-Han, and Jacob Steinhardt. 2024. Approaching Human-Level Forecasting with Language Models. https://doi.org/10.48550/arXiv.2402.18563 arXiv:2402.18563 [cs]

[19] Lewis Hammond and Sam Adam-Day. 2024. Neural Interactive Proofs. In *ICML 2024 Next Generation of AI Safety Workshop*.

[20] Robin Hanson. 2002. Logarithmic Market Scoring Rules for Modular Combinatorial Information Aggregation. *The Journal of Prediction Markets* 1, 1 (January 2002), 3–15. https://doi.org/10.5750/jpm.v1i1.417

[21] Robin Hanson. 2011. IP+ Like Barbed Wire?

[22] Robin Hanson. 2011. Rah Efficient IP.

[23] R. Howard. 1966. Information Value Theory. *IEEE Transactions on Systems Science and Cybernetics* 2, 1 (1966), 22–26. https://doi.org/10.1109/tssc.1966.30007

[24] Evan Hubinger. 2020. AI Safety via Market Making — LessWrong.

[25] Evan Hubinger. 2020. Alignment proposals and complexity classes. https://www.lesswrong.com/posts/N64THGX7XNCqRtvPG/alignment-proposals-and-complexity-classes. Retrieved April 8, 2026from https://www.lesswrong.com/posts/N64THGX7XNCqRtvPG/alignment-proposals-and-complexity-classes Accessed: April 8, 2026.

[26] Russell Impagliazzo and Moti Yung. 1987. Direct Minimum-Knowledge Computations. In *A Conference on the Theory and Applications of Cryptographic Techniques on Advances in Cryptology (CRYPTO '87)*. Springer-Verlag, Berlin, Heidelberg, 40–51.

[27] Geoffrey Irving, Paul Christiano, and Dario Amodei. 2018. AI Safety via Debate. https://doi.org/10.48550/arXiv.1805.00899 arXiv:1805.00899 [cs, stat]

[28] H. W. Kuhn, K. J. Arrow, E. W. Barankin, D. Blackwell, R. Bott, N. Dalkey, M. Dresher, D. Gale, D. B. Gillies, I. Glicksberg, O. Gross, S. Karlin, H. W. Kuhn, J. P. Mayberry, J. W. Milnor, T. S. Motzkin, J. von Neumann, H. Raiffa, L. S. Shapley, M. Shiffman, F. M. Stewart, G. L. Thompson, and R. M. Thrall. 1953. *Extensive games and the problem of information*. Princeton University Press, 193–216. http://www.jstor.org/stable/j.ctt1b9x1zv.17

[29] Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. 2024. EconAgent: Large Language Model-Empowered Agents for Simulating Macroeconomic Activities. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Lun-Wei Ku, Andre Martins, and Vivek Srikumar (Eds.). Association for Computational Linguistics, Bangkok, Thailand, 15523–15536. https://doi.org/10.18653/v1/2024.acl-long.829

[30] D. V. Lindley. 1956. On a Measure of the Information Provided by an Experiment. *The Annals of Mathematical Statistics* 27, 4 (1956), 986–1005. http://www.jstor.org/stable/2237191

[31] Alejandro M. Manelli and Daniel R. Vincent. 2006. Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory* 127, 1 (2006), 1–35. https://doi.org/10.1016/j.jet.2005.08.007

[32] Daniel Paleka, Abhimanyu Pallavi Sudhir, Alejandro Alvarez, Vineeth Bhat, Adam Shen, Evan Wang, and Florian Tramèr. 2024. Consistency Checks for Language Model Forecasters. In *The Thirteenth International Conference on Learning Representations*.

[33] H. Raiffa and R. Schlaifer. 1961. *Applied Statistical Decision Theory*. Division of Research, Graduate School of Business Adminitration, Harvard University. https://books.google.co.uk/books?id=wPBLAAAAMAAJ

[34] P.A. Samuelson and W.D. Nordhaus. 2009. *Economics*. McGraw-Hill Education. https://books.google.co.uk/books?id=eS5ZAAAAYAAJ

[35] Philipp Schoenegger, Indre Tuminauskaite, Peter S. Park, and Philip E. Tetlock. 2024. Wisdom of the Silicon Crowd: LLM Ensemble Prediction Capabilities Rival Human Crowd Accuracy. https://doi.org/10.48550/arXiv.2402.19379 arXiv:2402.19379 [cs]

[36] George J Stigler. 1961. The economics of information. *Journal of political economy* 69, 3 (1961), 213–225.

[37] Abhimanyu Pallavi Sudhir, Jackson Kaunismaa, and Arjun Panickssery. 2025. A Benchmark for Scalable Oversight Mechanisms. In *ICLR 2025 Workshop on Bidirectional Human-AI Alignment*. https://openreview.net/forum?id=mzLBxX84VI

[38] Emanuel Tewolde, Brian Hu Zhang, Caspar Oesterheld, Manolis Zampetakis, Tuomas Sandholm, Paul Goldberg, and Vincent Conitzer. 2024. Imperfect-recall games: equilibrium concepts and their complexity. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence* (Jeju, Korea) *(IJCAI '24)*. Article 332, 11 pages. https://doi.org/10.24963/ijcai.2024/332

[39] Kristine Thomassen, Siril Vassbø, Espen Solheim-Kile, and Jardar Lohne. 2016. Public-Private Partnership: Transaction Costs of Tendering. *Procedia Computer Science* 100 (2016), 818–825. https://doi.org/10.1016/j.procs.2016.09.230 International Conference on ENTERprise Information Systems/International Conference on Project MANagement/International Conference on Health and Social Care Information Systems and Technologies, CENTERIS/ProjMAN / HCist 2016.

[40] Marshall V. Van Alstyne. 1999. A proposal for valuing information and instrumental goods. In *Proceedings of the 20th International Conference on Information Systems* (Charlotte, North Carolina, USA) *(ICIS '99)*. Association for Information Systems, USA, 328–345.

[41] Marshall V. Van Alstyne. 1999. A Proposal for Valuing Information and Instrumental Goods. In *Proceedings of the 20th International Conference on Information Systems (ICIS '99)*. Association for Information Systems, USA, 328–345.

[42] Martin Weiss, Nasim Rahaman, Manuel Wuthrich, Yoshua Bengio, Li Erran Li, Bernhard Schölkopf, and Christopher Pal. 2024. Redesigning Information Markets in the Era of Language Models. In *First Conference on Language Modeling*.

[43] Stefan Wojcik, Sophie Hilgard, Nick Judd, Delia Mocanu, Stephen Ragain, M. B. Fallin Hunzaker, Keith Coleman, and Jay Baxter. 2022. Birdwatch: Crowd Wisdom and Bridging Algorithms can Inform Understanding and Reduce the Spread of Misinformation. arXiv:2210.15723 [cs.SI] https://arxiv.org/abs/2210.15723

A BASIC RESULTS ABOUT VALUE-OF-INFORMATION

We include a corrected and generalized version of an incorrectly-formulated result in the arXiv version of Weiss et al. [42]. These lemmas are perhaps obvious, but the whole theory of value-of-information and information bazaars rests upon them. Lemma A.1 asserts that "a Bayesian agent expects to gain from information"[6]. Lemma A.2 asserts that "a Bayesian agent expects to gain from inspection". Since the agent can always choose not to buy anything after inspecting, in the absence of transaction costs of inspection buyers *will* want to inspect more and more at each step.

Lemma A.1 (Bayesian agent expects to gain from information)**.** *Assume a probability space* $(\Omega, \mathcal{F}, \mathbf{P})$*, a set of choices* $\mathcal{X}$ *and a measurable utility function* $U : \Omega \times \mathcal{X} \to \mathbb{R}$*. Then for any random variable*

[6] From an AI alignment perspective, this may instead be phrased "A Bayesian agent trusts a more informed version of itself", and is a statistical-information version of *Vingean reflection* or *tiling agents* [15], the desire that (even logically non-omniscient) agents can trust smarter versions of themselves.

$I : \Omega \to \mathbb{R}$, we have:

$$\mathbf{E}_I \left[U(\arg\max_x \mathbf{E}\left[U(x) \mid I\right]) \right] \geq \max \mathbf{E}\left[U(x)\right]$$

Proof. For all values of $I = i$ and for all $x_0 \in \mathcal{X}$, we have (from the definition of arg max):

$$\mathbf{E}\left[U(\arg\max_x \mathbf{E}\left[U(x) \mid I = i\right]) \mid I = i \right] \geq \mathbf{E}\left[U(x_0) \mid I = i\right]$$

We take the expectation over $i \sim I$ on both sides, apply the law of total expectation and set $x_0 = \arg\max \mathbf{E}\left[U(x)\right]$ to obtain our result. □

Lemma A.2 (Bayesian agent expects to gain from inspection)**.** *Let the recursive construction of $\mathcal{X}^n$, U^n, InfoOffers^n,x_*^n be as in Section 2. Then $\forall n$, $\mathbf{E}\left[U^{n+1}(x_*^{n+1})\right] \geq 0$. The same also applies to finite-depth inspections i.e. $\mathbf{E}\left[U^{n+1}(x_{*:N}^{n+1})\right] \geq 0$.*

Proof. Equation (4) implies that $\mathbf{E}\left[U^n\left(x_*^n \mid x_*^{n+1}\right)\right]$ is $\geq$ any version of the same expression with x_*^n replaced by any $x \in \mathcal{X}^n$. We choose $x = \mathbf{0}$, for which the expression evaluates to 0, and take the expectation over x_*^{n+1}. □