Free-Energy-Driven LLM Agent for Modeling Information Exploration in Web Media

Yosuke Fukuchi (fukuchi@tmu.ac.jp)[1]
[1]Faculty of Systems Design, Tokyo Metropolitan University

Abstract

Modeling cognitive dynamics in web media exploration can contribute to healthier information environments, for example by enabling inference of users' false beliefs and estimating how media designs may affect belief formation. We propose *Free-Energy-driven Web Explorer (F-Explorer)*, a cognitive LLM agent that instantiates active inference for web exploration. In contrast to data-driven LLM approaches that learn behavior end-to-end, F-Explorer takes a theory-driven approach: it embeds an LLM as the belief module within a cognitive model grounded in the Free-Energy Principle (FEP), extending prior work that models human web exploration under FEP. We evaluate F-Explorer on a virtual SNS exploration dataset from a previous study and compare it with a reproduction of WEB-FEP with aligned initial-belief distributions. Using maximum likelihood estimation over 840 candidate settings per model (initial belief $\times$ learning rate), F-Explorer achieves higher action-sequence log-likelihood and improved top-k choice accuracy. Moreover, a proxy belief measure shows larger correspondence with participants' self-reported initial opinion scores, while correspondence with final opinion scores remains comparable. These results suggest that integrating LLM-based belief representations can improve FE–based models of web exploration while preserving interpretability through explicit latent-state and parameter inference.

Keywords: Web Exploration; Cognitive LLM Agent; Free-energy Principle; Active Inference; Bayesian Modeling; Confirmation Bias

Introduction

Web media have become a major source of information in everyday life, but they also amplify a range of societal problems including misuse, the spread of false beliefs, misinformation and fake news. Prior work has examined and proposed alternative platform designs such as fact-checking interventions (Tanaka et al., 2023) and mechanisms related to self-censorship (Howe et al., 2023). However, the broader problems surrounding web media have remained despite these efforts.

This paper shifts the focus from platform design to users' cognitive dynamics. A computational model that captures the cognitive processes underlying human web exploration has potential for a wide range of applications: predicting user behavior, inferring their latent cognitive states, and evaluating how design choices support (or undermine) healthy media use. Ultimately, such a modeling approach can contribute to building healthier online information environments.

Fukuchi (2025) proposed WEB-FEP, a computational model of human cognition in web exploration grounded in active inference under the Free-Energy Principle (FEP). In WEB-FEP, an agent selects actions by trading off two sources of value: epistemic value (information gain from new observations) and pragmatic value (confirmatory satisfaction obtained when observations align with existing beliefs). A key finding is that, when the agent's learning rate representing how strongly beliefs are updated by observations is small, pragmatic value exerts a stronger influence relative to epistemic value, leading the agent to exhibit confirmation bias. In parallel, large language models (LLMs) have rapidly become a major approach to cognitive modeling. For example, Binz et al. (2025) showed that Centaur, an LLM fine-tuned on a large corpus of psychological experiment data, can imitate human behavior with high accuracy.

Despite this progress, important challenges remain to model user cognitions in web media exploration. WEB-FEP represents a person's belief as a Gaussian distribution on a two-dimensional plane obtained by compressing a text-embedding space, leaving open how faithfully the model captures linguistic meaning. Moreover, the study primarily discussed qualitative behavioral differences of the model across different parameter settings and lacks a quantitative assessment of how well such FEP-based model explains actual human behaviors. Conversely, a key limitation of data-driven LLM approaches is controllability and black-boxness: although techniques such as persona prompting have been explored, it still remains unclear how to reliably reflect individual differences in model behavior (Binz et al., 2025). More broadly, how to leverage LLMs for modeling web exploration, where cognition evolves through interaction dynamics rather than in tightly controlled laboratory tasks, remains an open question. Additionally, the black-boxness makes it challenging to reveal the cognitive mechanisms behind human behaviors.

This paper proposes Free-Energy-driven Web Explorer (F-Explorer), a cognitive LLM agent that integrates WEB-FEP with an LLM. F-Explorer is an LLM agent driven not by human behavioral data but by an explicit theoretical model (FEP). By directly controlling the model's parameters, F-Explorer can exhibit diverse and systematically varying behaviors without requiring human action data for training. In addition, given an individual's action sequence, F-Explorer enables maximum likelihood estimation (MLE) of user-specific latent belief states and model parameters, offering a more interpretable account of the cognition underlying behavior than

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

purely data-driven approaches, where latent mechanisms can remain a black box.

F-Explorer treats the LLM as the agent's belief: the likelihood distribution over generated opinions is taken as a proxy for which opinions are considered plausible versus implausible. The LLM's likelihood distribution is updated via fine-tuning. Importantly, this fine-tuning is not used to learn human actions, but instead serves as a belief-updating process analogous to how humans revise beliefs upon encountering new information. The design allows belief updating to be implemented directly in the LLM's rich representational space. Based on the LLM-induced likelihood distribution, F-Explorer evaluates the expected free energy (EFE) of candidate future actions combining epistemic and pragmatic value and actively explores web information to minimize EFE.

We evaluate F-Explorer with the user-study dataset of Fukuchi (2025), comparing with a reproduction of WEB-FEP. Aligning the initial belief distribution between F-Explorer and WEB-FEP, we run multiple simulations with 840 different initial parameters and perform MLE of the parameters that best explain each user's action sequence. The resulting simulations suggest that F-Explorer can better fit user behaviors than WEB-FEP, indicating that F-Explorer better leverages the linguistic meaning present in observations and can reproduce key aspects of human cognitive dynamics in web exploration.

Background

LLM-based Cognitive Modeling

LLMs have rapidly emerged as a foundation for cognitive modeling. Because they encode rich linguistic and world knowledge and can be queried in flexible ways, LLMs have been used to simulate a wide range of human behaviors including decision making, reasoning, and language-mediated judgments. A prominent example is Centaur (Binz et al., 2025), which is an LLM fine-tuned on large-scale human decisions data collected across many psychological tasks, demonstrating that an LLM can reproduce human choice patterns with high fidelity in controlled experimental settings.

The methodology of LLM-based cognitive modeling is diverse. Iida et al. (2024) introduced a conceptual taxonomy that differentiates Cognitive model Embedded in LLM (CEL) approaches and LLM Embedded in Cognitive model (LEC) approaches. In CEL, cognitive structure is primarily expressed through the LLM itself via prompting or in-context learning. CEL approaches can achieve strong predictive performance when abundant behavioral data are available, but the underlying cognitive mechanisms can remain implicit in the model's internal representations, complicating interpretation and mechanistic inference. In contrast, LEC approaches place the LLM as a component within an explicit cognitive model. The cognitive model specifies latent states, update rules, and decision objectives. That is, the factors governing behavior are externalized as model parameters, improving controllability and interpretability. The inference can be framed as estimating user-specific latent states and parameters from observed actions, rather than relying purely on end-to-end behavioral training. F-Explorer is an LEC model that integrates an LLM into the belief module of WEB-FEP and receives the advantages of LEC. It provides an explicit computational account of cognitive dynamics (e.g., belief updating and exploration–exploitation trade-offs) while still leveraging the expressive representational capacity of LLMs.

FEP-based Modeling of Web Exploration

WEB-FEP models web exploration as active inference under the FEP, a theoretical framework for perception, inference, and decision making under uncertainty (Friston, 2010). In FEP, an agent selects an action a to minimize expected free energy, which can be written as:

$$a = \arg\min_{a} G(a), \tag{1}$$

$$G(a) = -(V_{\text{epist}}(a) + V_{\text{prag}}(a)), \tag{2}$$

where $V_{\text{epist}}(a)$ and $V_{\text{prag}}(a)$ are epistemic and pragmatic values, respectively.

Epistemic value captures the expected information gain obtained by selecting action a:

$$V_{\text{epist}}(a) = \mathbb{E}_{Q(s',o'|a)}\left[D_{\text{KL}}[Q(s'|o')\|Q(s')]\right], \tag{3}$$

where $Q(s', o'|a)$ is the agent's predictive belief about future states and observations under action a, $Q(s'|o')$ is the posterior belief after observing o', and $Q(s')$ is the prior belief before observing o'. Intuitively, $V_{\text{epist}}(a)$ encourages exploratory actions that are expected to change beliefs.

Pragmatic value quantifies the expected alignment between future observations and the agent's objective C:

$$V_{\text{prag}}(a) = \mathbb{E}_{Q(o'|a)}[\ln P(o'|C)], \tag{4}$$

where $Q(o'|a)$ is the predictive distribution over future observations under action a, and $P(o'|C)$ encodes how satisfying different observations are expected to be for the agent, where C denotes a preference variable over possible observations. In WEB-FEP, observations that confirm the agent's current belief are expected to yield greater subjective satisfaction, so actions expected to produce such observations have higher $V_{\text{prag}}(a)$, favoring exploitative or confirmatory behavior.

WEB exploration as sequential link selection. To instantiate this framework for modeling web media exploration, Fukuchi (2025) simplified the exploration into repeated link selection and post texts observation. Participants first read a news regarding a controversial topic and then actively explored posts about the topic. An action a corresponds to selecting a link (a hashtag in a social media interface), and an observation o corresponds to the text content shown after selecting that link (posts associated with the hashtag). Let $t \in \{1, \ldots, T\}$ denote the iteration index.

Belief representation and belief updating. WEB-FEP represented a user's belief $q(x)$ as a probability distribution over a semantic space, where x denotes a proposition or opinion (e.g., supportive vs. opposing claims) represented in a sentence embedding space. Observations are mapped into the same space, enabling belief updating as density increase around observed content. Concretely, WEB-FEP approximates $q(x)$ with a multivariate normal distribution whose parameters are updated online. Given o, the belief is updated by minimizing a cross-entropy objective $-\mathbb{E}[\ln q(o)]$ that encourages high probability on the observed content. The update is implemented via stochastic gradient descent; the learning rate α governs how strongly beliefs change in response to new information and plays a central role as an individual-difference parameter.

Operationalizing epistemic and pragmatic values. WEB-FEP further introduces a practical approximation for computing values for candidate links. Let o_a denote a semantic representation associated with action a (e.g., the representation of the link itself, or a summary representation of the content expected under that link). Pragmatic value is operationalized as the log density of o_a under the current belief:

$$V_{\mathrm{prag}}(a) = \ln q(o_a), \tag{5}$$

which favors actions expected to yield belief-consistent content. o_a denotes the content associated with a. In WEB-FEP, o_a is instantiated by the set of posts linked to a, and pragmatic value is computed as the mean of the log densities assigned to the posts under the current belief.

Epistemic value is operationalized as the KL divergence between the updated belief and the current belief, reflecting how much the belief would change if the agent were to observe content under action a:

$$V_{\mathrm{epist}}(a) = D_{\mathrm{KL}}(q_{\mathrm{new}} \,\|\, q), \tag{6}$$

where q_{new} is obtained by a *virtual* belief update using the content associated with a. This construction links exploration directly to belief change, and makes explicit how epistemic value depends on α: when α is small, updates are weak, epistemic value diminishes, and choices become increasingly driven by pragmatic value, yielding confirmation-biased exploration (Fukuchi, 2025).

Action selection. Given the total value $V(a) = V_{\mathrm{epist}}(a) + V_{\mathrm{prag}}(a)$, the agent selects among candidate actions. Overall, WEB-FEP provides a compact, interpretable account of web exploration dynamics as an online trade-off between belief-confirming and information-seeking behavior.

F-Explorer

Overview

F-Explorer is a cognitive agent in the LEC paradigm that couples (i) a theory of information exploration formalized as active inference in WEB-FEP with (ii) an LLM-based belief representation and update mechanism that operates directly in a meaning-sensitive space. F-Explorer uses LoRA, a parameter-efficient fine-tuning method for LLMs (Hu et al., 2022), not to learn human action policies, but to implement belief updating that is analogous to how humans revise beliefs when encountering new information.

As in WEB-FEP, F-Explorer induces diverse exploration dynamics by controlling (i) an initial belief state and (ii) a learning-rate hyperparameter corresponding to α in WEB-FEP. In addition, given an individual's action sequence, F-Explorer supports MLE of user-specific latent belief states and model parameters, yielding an interpretable account of the cognition underlying behavior by explicitly identifying the parameters that govern belief updating and the resulting exploration dynamics.

Belief as an LLM-Induced Likelihood Distribution

A central departure from WEB-FEP is that F-Explorer represents belief directly through an LLM. Intuitively, if an LLM assigns high likelihood to texts expressing certain opinions, those opinions are treated as more plausible under the agent's current belief.

Let θ_t denote the parameters of the LLM at time t. In F-Explorer, θ_t corresponds to a frozen base model plus a LoRA adapter that is updated online. Given a prompt p and a candidate text x, the LLM defines a conditional likelihood $P_{\theta_t}(x \mid p)$:

$$\ell_{\theta_t}(x) \;=\; \log P_{\theta_t}(x \mid p). \tag{7}$$

To obtain a tractable belief distribution, F-Explorer relies on a fixed evaluation set $\mathcal{X} = \{x_i\}_{i=1}^{N}$. A discrete proxy belief Q_t over $\mathcal{X}$ is defined as $\{Q_t(i)\}_{i=1}^{N}$, where $Q_t(i) \;\propto\; \exp(\ell_{\theta_t}(x_i))$.

Belief Update via LoRA Fine-Tuning

After selecting an action a_t, the agent observes a set of posts $o_{t+1} = \{x_j^{a_t}\}_{j=1}^{K}$ associated with the chosen link. F-Explorer updates the LoRA parameters by maximizing the likelihood of the newly observed texts (i.e., minimizing the negative log-likelihood). Let θ_t^{LoRA} denote the LoRA parameters at time t (with the base model held fixed). We define the observation negative log-likelihood objective as

$$\mathcal{L}_{t+1}(\theta^{\mathrm{LoRA}}) \;=\; -\frac{1}{K}\sum_{j=1}^{K} \ell_{\theta_t}(x_j^{a_t}). \tag{8}$$

The belief update is implemented by stochastic gradient descent on $\mathcal{L}_{t+1}$.

In practice, LoRA training involves multiple hyperparameters that affect the effective magnitude of belief change per observation (e.g., optimizer settings, LoRA scaling, and the number of optimization steps). To stabilize updates, F-Explorer fixes the optimizer learning rate and controls the update strength primarily via the number of gradient steps per observation, n_{step}:

$$\theta_{t+1}^{\mathrm{LoRA}} \;=\; \mathrm{SGD}^{n_{\mathrm{step}}}\!\left(\theta_t^{\mathrm{LoRA}};\, \mathcal{L}_{t+1}\right). \tag{9}$$

Increasing the likelihood around the observed evidence is thus treated as a computational analogue of evidence accumulation.

Pragmatic and Epistemic Values from LLM Likelihoods

Given a candidate action a at time t, let $o^{(t,a)} = \{x_j^{(t,a)}\}_{j=1}^{K}$ denote the posts that would be observed if the agent selected a. In a strict application of FEP, the agent would additionally require a generative process to predict future observations o' under each candidate action. Following WEB-FEP, we adopt a simplifying assumption suitable for the present web-media setting: the content associated with a link (e.g., a hashtag) is treated as a sufficiently informative proxy for the posts that will be encountered, so $o^{(t,a)}$ is treated as known when computing values. This simplification cannot capture phenomena such as clickbait (where link cues systematically misrepresent content), and addressing such cases remains an avenue for refinement.

V_{prag} is defined as the expected log-likelihood of the future observations under the current belief/model:

$$V_{\text{prag}}(a) \ = \ \mathbb{E}_{x\sim o^{(t,a)}}\left[\ell_{\theta_t}(x)\right] \ \approx \ \frac{1}{K}\sum_{j=1}^{K}\ell_{\theta_t}\left(x_j^{(t,a)}\right). \quad (10)$$

This mirrors the objective of belief update (maximizing likelihood of observed evidence), but is used prospectively to quantify how well a candidate action aligns with the agent's current belief.

V_{epist} quantifies how much the belief would change if the agent were to observe $o^{(t,a)}$. Let $Q_{t+1}^{(a)}$ denote the proxy belief after a virtual update conditioned on $o^{(t,a)}$ (using the same update rule as in Eq. 9, but applied hypothetically). We approximate epistemic value using KL divergence:

$$\begin{aligned} V_{\text{epist}}(a) \ &= \ D_{\text{KL}}\left(Q_{t+1}^{(a)} \,\|\, Q_t\right) \\ &= \ \sum_{i=1}^{N} Q_{t+1}^{(a)}(i)\left(\log Q_{t+1}^{(a)}(i) - \log Q_t(i)\right). \quad (11) \end{aligned}$$

Operationally, this corresponds to computing the KL divergence between the two proxy belief distributions induced by the LLM likelihood vectors before and after the (virtual) update. Related likelihood-based divergences have also been used to quantify distances between language models (Oyama et al., 2025).

Action Selection Policy

F-Explorer combines pragmatic and epistemic values as in WEB-FEP:

$$V(a) \ = \ V_{\text{prag}}(a) + \lambda V_{\text{epist}}(a). \quad (12)$$

Note that λ is a constant shared across participants; it is not intended to encode individual-specific exploration–exploitation preferences. Instead, individual differences in the effective balance between the two terms arise through the belief-update strength controlled by n_{step}: when updates are weak, epistemic value becomes smaller, and behavior becomes increasingly driven by pragmatic value.

Actions are selected according to a softmax policy:

$$\pi_t(a) \ = \ P(a_t = a \mid \theta_t) \ = \ \frac{\exp(V(a))}{\sum_{a'\in\mathcal{A}}\exp(V(a'))}. \quad (13)$$

Likelihood of Action Sequences and MLE

F-Explorer specifies a forward generative process in which actions are produced sequentially from an initial belief state and a given information environment. To infer user-specific parameters from observed human actions, we evaluate how well different parameter settings explain an individual's action sequence.

Given an observed action sequence $a_{1:T}$, the model log-likelihood is

$$\log P(a_{1:T} \mid \Theta) \ = \ \sum_{t=1}^{T}\log \pi_t(a_t), \quad (14)$$

where Θ denotes user-specific parameters (n_{step} and the initial belief state). Inference can be framed as maximizing the action-sequence likelihood:

$$\Theta^* \ = \ \arg\max_{\Theta}\left\{\log P(a_{1:T} \mid \Theta)\right\}. \quad (15)$$

This MLE formulation yields an interpretable account of individual behavior by explicitly identifying the parameters governing belief updating and the resulting exploration dynamics, rather than relying on end-to-end behavioral training.

Opinion Score

In the WEB-FEP user study, participants reported their stance on the focal topic before and after exploration using an opinion score on a 0–100 scale (0 = strongly against, 100 = strongly in favor). To enable an interpretable comparison between self-reported stances and the latent belief state inferred by F-Explorer, we define a simple scalar measure from the LLM-induced proxy belief Q_t.

Let $\mathcal{X}_{\text{pos}} \subset \mathcal{X}$ be the subset of texts expressing the supporting stance. We define the opinion score proxy as the fraction of belief mass assigned to $\mathcal{X}_{\text{pos}}$:

$$\text{Opinion}(Q_t) \ = \ 100 \times \frac{\sum_{i\in\mathcal{X}_{\text{pos}}} Q_t(i)}{\sum_{i=1}^{N} Q_t(i)}. \quad (16)$$

Note that, in our evaluation set construction, $\mathcal{X}_{\text{pos}}$ and its complement (opposing texts) are matched in size, so the measure is not driven by class imbalance.

Evaluation

Objectives

Our evaluation has two objectives. First, we quantitatively test how well the FE-based active exploration framework can account for real users' web exploration behavior, using action-sequence model fitness. While Fukuchi (2025) primarily discussed qualitative behavioral differences across parameter settings, a systematic evaluation of model fit to observed human action sequences remained unexplored.

Second, we test whether integrating an LLM as the belief representation and update mechanism improves both (i) action-sequence explanatory power and (ii) belief-state inference performance, relative to WEB-FEP's two-dimensional Gaussian belief representation. The focus of this comparison is the two models grounded in the same exploration theory (FEP/active inference) differ in explanatory power when the belief representation is made more meaning-sensitive using an LLM.

Task and Dataset

We use the user-study dataset collected by Fukuchi (2025) ($N = 94$). Participants interacted with a controlled virtual SNS environment. Across $T = 10$ iterations, participants repeatedly selected one link (a hashtag) from six candidates and then observed eight post texts associated with the selected hashtag. Participants also self-reported their attitude toward the focal issue (an opinion score on a 0–100 scale) both before and after the exploration session. We treat this self-reported opinion score as an external validity signal to compare the interpretability of the models' belief measures. $\mathcal{X}$ consists of 240 posts for each supportive and opposing contents.

Procedure

LLM and LoRA details llm-jp-3-3.7b-instruct (LLM-jp et al., 2024) was used for F-Explorer's LLM. Because latent representations of an LLM tend to shift from surface-level features to syntactic structure, semantic information, and broader contextual integration as layers become deeper (Jawahar et al., 2019). We only applied LoRA to 17-20th layers, whose 2-D principal component analysis (PCA) results most clearly separated the polarity of comments. The 17th layer's representations were also used for WEB-FEP's text-embedding space by reducing the dimension with PCA.

Initial belief distribution. For F-Explorer, we constructed 168 initial beliefs by sampling 200 posts from $\mathcal{X}$ with different balance settings, and initializing the LoRA adapter accordingly. Because WEB-FEP represents belief as a Gaussian distribution in a two-dimensional semantic space, we fit a 2D Gaussian to the distribution induced by each F-Explorer initial LoRA belief (i.e., a likelihood-based belief over $\mathcal{X}$) and used the fitted Gaussian as the corresponding WEB-FEP initial belief. This procedure aligns the distribution of initial beliefs between the two models as closely as possible while respecting their different belief parameterizations.

Search space of learning rate and MLE. For both F-Explorer and WEB-FEP, we considered five levels of learning rate. In F-Explorer, update strength is controlled by the number of gradient steps per observation $n_{\text{step}} \in \{1, 3, 5, 7, 10\}$. In WEB-FEP, update strength is controlled by the learning rate parameter, $\alpha \in \{0.1, 0.2, 0.3, 0.5, 0.7\}$. Combining 168 initial beliefs with five learning-rate levels yields $168 \times 5 = 840$ candidate parameter configurations per model.

For each participant and each configuration, we ran the forward model and computed the log-likelihood of the observed action sequence using Eq. 14. We then performed MLE by selecting, for each participant, the configuration that maximized Eq. 15.

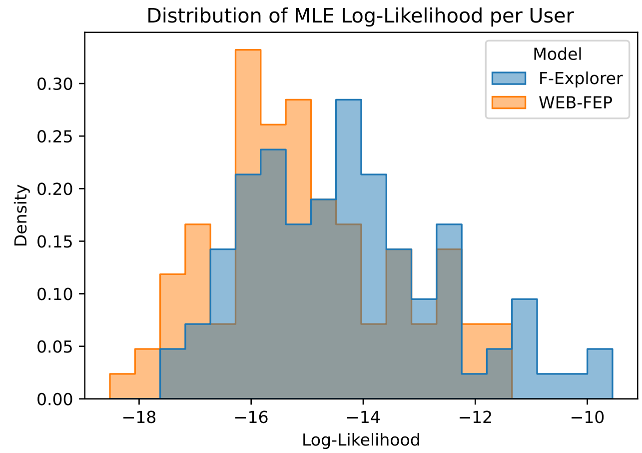

Figure 1: Distribution of best log-likelihoods selected by maximum likelihood estimation for each user

Tuning of λ. We tuned the shared coefficient λ by a parameter sweeping, selecting the value that maximized model fitness aggregated across participants. For F-Explorer, we performed a sweep over $\lambda_{\text{F-Explorer}} \in \{0\} \cup \left\{0.01 \cdot 1.1^i \mid i = 0, \ldots, 199\right\}$, then refined around the best region with a step size of 0.05, resulting in $\lambda_{\text{F-Explorer}} = 40.85$. We similarly performed a sweep for WEB-FEP over $\lambda_{\text{WEB-FEP}} \in \{0\} \cup \left\{0.001 \cdot 1.1^i \mid i = 0, \ldots, 199\right\}$, then refined with a step size of 0.005, resulting in $\lambda = 0.270$. The large difference of $\lambda_{\text{F-Explorer}}$ scale is mainly due to its sample-wise value calculations.

Results

Table 1 summarizes action-sequence fitness and related metrics for each model, evaluated at each participant's maximum-likelihood configuration. "Corr" reports the correlation between the model's opinion-score measure and the self-reported opinion score.

Action-sequence likelihood. F-Explorer achieved higher log-likelihood than WEB-FEP. A paired t-test across participants showed a significant improvement, $t(93) = 3.98$, $p = 0.000136$, with a mean difference of 0.744 (SE = 0.187). Figure 1 shows the distributions. F-Explorer overall improved the distribution, and the number of the samples that better explain user behavior increased.

Top-k accuracy. We also evaluated whether the observed action at each trial falls within the top-k most probable actions under the model's predicted policy π_t, where the number of action candidates are six. Figure 2 illustrates the top-1 result. F-Explorer outperformed WEB-FEP for $k = 1, 2, 3$. This improvement reflects better relative probability assignment across six alternatives, not merely predicting the single most frequent choice.

Table 1: Model fitness and metrics based on each participant's maximum-likelihood configuration (mean ± SE). Corr denotes correlation with self-reported opinion scores.

	$\log P(a_{1:T} \mid \Theta^*)$	Top-1 Acc.	Top-2 Acc.	Top-3 Acc.	Corr (Init)	Corr (Final)
F-Explorer	**-14.265** ± 0.183	**.431** ± .014	**.668** ± .013	**.781** ± .013	**.331** (p=.001)	.424 (p <.001)
WEB-FEP	-15.010 ± 0.167	.334 ± .015	.534 ± .016	.711 ± .017	.260 (p=.011)	**.449** (p <.001)

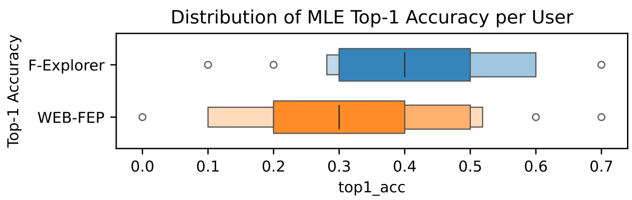

Figure 2: Distribution of top-1 scores selected by maximum likelihood estimation for each user

Opinion-score correlation. For the initial opinion score, F-Explorer achieved a higher correlation with self-reports, suggesting that the LLM-based belief representation may better capture individual differences in the initial belief state. For the final opinion score, we could not find an improvement from WEB-FEP.

Discussion

Overall, F-Explorer showed better fitting with user data. This indicates that replacing the 2D Gaussian belief with an LLM-induced likelihood belief can improve how well the same exploration theory accounts for observed action sequences.

The top-k accuracies of both F-Explorer and WEB-FEP are well above chance levels ($k/6$ for $k \in \{1, 2, 3\}$). Given the size of the action-sequence space ($N = 6^{10}$), these results indicate non-trivial predictive performance even under a discrete maximum-likelihood search over only 840 candidate parameter settings.

That said, the top-k scores still leave substantial room for improvement, suggesting that the current parameter grid (840 configurations) may be insufficient to capture the full diversity of user behaviors. At the same time, even within a single topic, the number of candidate configurations grows rapidly as we enrich the initial-belief family and update-strength variations, making naive grid expansion computationally impractical for future applications. Future work should therefore explore more efficient inference strategies, such as: (i) narrowing priors using participant traits or behavioral markers (e.g., confirmatory selection count), (ii) coarse-to-fine inference schemes analogous to our λ tuning, and (iii) sequential Monte Carlo or other adaptive search methods to more efficiently explore the space of initial beliefs and update parameters.

We also need to note that, although we aligned initial belief distributions between models, WEB-FEP's α and F-Explorer's n_{step} are not the same construct, and their search spaces may not be perfectly comparable. This mismatch could disadvantage one model by making the explored parameter grid less well matched to its effective update magnitude. Although our another trial in which the number of α was increased ($\alpha \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$) for WEB-FEP did not show a great improvement but slight degradations for the correlation measures, a more principled calibration between α and n_{step} is an important direction for follow-up evaluation.

Both F-Explorer and WEB-FEP update belief via likelihood-based evidence accumulation: content that is observed becomes more likely under the model. However, self-reported final opinions may reflect additional cognitive and affective processes beyond accumulation of textual evidence, such as reliability assessment, reactance, consistency checking against prior beliefs, and emotional responses. Because F-Explorer does not explicitly model selective acceptance/rejection of information after observation, the model may explain link choices more directly than subjective post-exploration evaluations. Incorporating mechanisms such as trustability estimation and selective attention/weighting into the belief update could improve alignment particularly with final self-reports.

Conclusion

This paper proposed F-Explorer, a cognitive LLM agent driven by FEP for modeling web media exploration. F-Explorer employs an LLM not in a data-driven way but as a belief module within a theory-grounded cognitive model. Because behavior is governed by explicit parameters, we can systematically generate diverse exploration dynamics without end-to-end behavioral training. In addition, based on explicit latent belief states and value terms, we can infer user-specific cognitive states and parameters from action sequences via MLE and explain choices mechanistically rather than treating them as a black-box. Our evaluation suggests that integrating LLM-based belief representations can improve FE–based models of web exploration while preserving interpretability through explicit latent-state and parameter inference.

Acknowledgments

This work was supported by the Telecommunications Advancement Foundation, Japan Science and Technology Agency (JST) ACT-X Grant Number JPMJAX25CL, and the "Development Acceleration Use" program of ABCI 3.0 (Takano et al., 2024), which is provided by AIST and AIST Solutions.

References

Binz, M., Akata, E., Bethge, M., Brändle, F., Callaway, F., Coda-Forno, J., Dayan, P., Demircan, C., Eckstein, M. K., Éltető, N., et al. (2025). A foundation model to predict and capture human cognition. *Nature*, 1–8.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature reviews neuroscience*, *11*(2), 127–138.

Fukuchi, Y. (2025). A bayesian model of confirmatory exploration in text-based web media. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*, 1727–1733.

Howe, P., Perfors, A., Ransom, K. J., Walker, B., Fay, N., Kashima, Y., & Saletta, M. (2023). Self-censorship appears to be an effective way of reducing the spread of misinformation on social media. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*(45).

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W., et al. (2022). Lora: Low-rank adaptation of large language models. *ICLR*, *1*(2), 3.

Iida, A., Okuoka, K., Fukuda, S., Omori, T., Nakashima, R., & Osawa, M. (2024). Integrating large language model and mental model of others: Studies on dialogue communication based on implicature. *Proceedings of the 12th International Conference on Human-Agent Interaction*, 260–269. https://doi.org/10.1145/3687272.3688303

Jawahar, G., Sagot, B., & Seddah, D. (2019, July). What does BERT learn about the structure of language? In A. Korhonen, D. Traum, & L. Màrquez (Eds.), *Proceedings of the 57th annual meeting of the association for computational linguistics* (pp. 3651–3657). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1356

LLM-jp, : Aizawa, A., Aramaki, E., Chen, B., Cheng, F., Deguchi, H., Enomoto, R., Fujii, K., Fukumoto, K., Fukushima, T., Han, N., Harada, Y., Hashimoto, C., Hiraoka, T., Hisada, S., Hosokawa, S., Jie, L., Kamata, K., . . . Yoshino, K. (2024). Llm-jp: A cross-organizational project for the research and development of fully open japanese llms. https://arxiv.org/abs/2407.03963

Oyama, M., Yamagiwa, H., Takase, Y., & Shimodaira, H. (2025, July). Mapping 1,000+ language models via the log-likelihood vector. In W. Che, J. Nabende, E. Shutova, & M. T. Pilehvar (Eds.), *Proceedings of the 63rd annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 32983–33038). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.acl-long.1584

Takano, R., Takizawa, S., Tanimura, Y., Nakada, H., & Ogawa, H. (2024). Abci 3.0: Evolution of the leading ai infrastructure in japan. https://arxiv.org/abs/2411.09134

Tanaka, Y., Inuzuka, M., Arai, H., Takahashi, Y., Kukita, M., & Inui, K. (2023). Who does not benefit from fact-checking websites? a psychological characteristic predicts the selective avoidance of clicking uncongenial facts. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3544548.3580826