Supplementary Material for AAAI 2026 paper #29562

Anonymous submission

Algorithm Details

Algorithm 1: TRIPLE: Offline Profile and Rationale Generation at time t

Input: User history $H_u = \{h_1, ..., h_t\}$, window size k
Output: Habitual profile P_u^H, Intentional profile P_u^I

$W_t \leftarrow H_u[t-k+1:t]$ {Define the sliding window}

// 1. Generate Profiles from the window
$P_u^H \leftarrow \mathcal{L}_{\text{habit}}(\mathcal{P}_{\text{habit}}, W_t)$ {Eq. 1}
$P_u^I \leftarrow$ InitializeEmptyProfile()
for $i = t-k+1$ to t **do**
  $h_i \leftarrow H_u[i]$
  $p_u^I(i) \leftarrow \mathcal{L}_{\text{intent}}(\mathcal{P}_{\text{TPB}}, h_i)$
  $P_u^I \leftarrow \mathcal{L}_{\text{update}}(\mathcal{P}_{\text{update}}, \{P_u^I, p_u^I(i)\})$
end for

// 2. Pre-compute Rationales for all items in the window
for $i = t-k+1$ to t **do**
  $h_i \leftarrow H_u[i]$
  $inputs_R \leftarrow \{h_i, P_u^H, P_u^I\}$
  $R_i \leftarrow \mathcal{L}_{\text{rationale}}(\mathcal{P}_{\text{rationale}}, inputs_R)$ {Eq. 4}
  // Rationales are stored for later retrieval.
end for

return P_u^H, P_u^I

Additional Examples of TRIPLE

This section provides additional examples of TRIPLE generation in LaMP-5, demonstrating how TRIPLE captures the complex interplay between habitual patterns and intentional behaviors that characterize user decision-making processes. Through these detailed illustrations, we elucidate the mechanisms by which TRIPLE synthesizes dual-process theoretical insights to generate psychologically interpretable and empirically accurate user representations.

Table 1 demonstrates the generation and stability of the Habitual Profile. As described in Section 3.1.1, the profile is built by analyzing a window of recent behaviors at once, capturing recurring patterns. This profile represents the user's automatic, default tendencies.

Algorithm 2: TRIPLE: Online Personalized Prediction

Input: New query x_n, Habitual profile P_u^H, Intentional profile P_u^I, History window W_t
Output: Prediction $\hat{y}_n$, Prediction rationale R_u^{predict}

// 1. Retrieve relevant history
$h_* \leftarrow$ BM25_Retrieve(x_n, W_t)

// 2. Retrieve the pre-computed rationale
Let $R_u^{\text{history}} \leftarrow$ the pre-computed rationale for h_*.

// 3. Generate final prediction and its rationale
$inputs_P \leftarrow \{x_n, P_u^H, P_u^I, h_*, R_u^{\text{history}}\}$
$(\hat{y}_n, R_u^{\text{predict}}) \leftarrow \mathcal{L}_{\text{predict}}(\mathcal{P}_{\text{predict}}, inputs_P)$ {Eq. 5}

return $\hat{y}_n, R_u^{\text{predict}}$

Table 2 illustrates the incremental update mechanism of the Intentional Profile. Following the process in Section 3.1.2, the profile evolves with each new behavior. The example shows how the system first builds a stable, formal profile based on a series of consistent academic titles up to $t = 8$. At $t = 9$, however, it is confronted with a starkly atypical input: the user writes the title 'Girls Rule, Boys Drool...'. The colloquial and provocative tone of this title directly conflicts with the user's established profile, which values objective language and academic rigor. Crucially, the table then demonstrates how TRIPLE resolves this conflict not by overwriting the existing profile, but by intelligently integrating the new, informal intention as a nuanced, context-dependent exception. This specific moment of conflict is therefore crucial for demonstrating the model's ability to adapt and for understanding the user's full behavioral spectrum.

To explain why a specific action occurred, TRIPLE generates a Behavioral Rationale. This post-hoc analysis, detailed in Table 3, synthesizes the user's stable Habitual Profile and their dynamic Intentional Profile to explain the underlying cognitive dynamics. The table demonstrates this process for a consistent behavior, where the user's habitual and intentional systems are in alignment. The generated rationale concludes that the action resulted from a well-established habit

Table 1: An example of Habitual Profile generation for a user in the LaMP-5 (Scholarly Title Generation) task. The model synthesizes the user's past behaviors into a descriptive summary of their habitual title-writing patterns.

Sample of User's Raw Behavioral History (User ID: 111381)

- **Input (Abstract):** "A significant volume of information leaks in organizations... we use a hybrid agent-based and dynamic network model..."
  Output (Title): Inadvertent leaks: **exploration via** agent-based dynamic network simulation.
- **Input (Abstract):** "Cyber-attacks are cheap... we use sentiment in Twitter posts to observe country-to-country perceptions..."
  Output (Title): Understanding DDoS cyber-attacks **using** social media analytics
- **Input (Abstract):** "ISIS' ability to build and maintain a large online community... We illustrate the potential of this partnership..."
  Output (Title): From Tweets to Intelligence: Understanding the Islamic Jihad Supporting Community on Twitter.
- **Input (Abstract):** "Social identities carry widely agreed upon meanings, called stereotypes... we develop a method to extract the stereotypes of Twitter users."
  Output (Title): Girls Rule, Boys Drool: **Extracting** Semantic and Affective Stereotypes from Twitter.

TRIPLE's Generated Habitual Profile

The author's habitual title patterns include using **action verbs like 'Exploring', 'Understanding', and 'Extracting'** to describe the research process, and **incorporating key concepts or methodologies** like 'agent-based dynamic network simulation', 'social media analytics', and 'probabilistic geographical aspect-opinion model' to highlight the study's focus. When presenting findings related to social media analysis, they **structure titles as 'From [Data Source] to [Insight]: [Topic]'**, as seen in Example 5. For studies involving network analysis or modeling, they use titles that **describe the model or method**, such as 'A Probabilistic Geographical Aspect-Opinion Model' or 'Reporting a network's most-central actor with a confidence level'.

that was fully reinforced by the user's conscious deliberation, thus providing a clear, interpretable explanation for the choice.

Statistical Significance of Ablation Studies

To provide a robust quantitative validation for our ablation studies, we conducted statistical significance tests. Our analysis employs a paired, one-tailed t-test on a per-user basis to control for individual variance and increase statistical power, with findings cross-validated using the non-parametric Wilcoxon signed-rank test ($p < 0.05$). The analysis specifically aimed to confirm that our TRIPLE achieves a significant performance gain over the model 'without the Behavioral Rationale'.

Our results show that the inclusion of the Behavioral Rationale yields statistically significant performance gains across most tasks. This rigorously validates that the Behavioral Rationale is an indispensable component of the TRIPLE, serving as a critical mechanism to synthesize user profiles and maximize performance.

Complexity Analysis

The time complexity of TRIPLE can be analyzed in two distinct phases: the offline phase for profile construction and the online phase for real-time prediction. Let t be the total number of historical interactions for a user, k be the size of the sliding window, and L be the average length of a history item h_i. We denote the computational cost of an LLM call as a function $C_{LLM}(N)$, where N is the length of the input sequence. While the exact complexity of a transformer-based LLM is intricate, $C_{LLM}(N)$ is generally considered to be super-linear with respect to the input sequence length N.

Offline/Profile Construction Phase: This phase involves preparing the profiles and rationales needed for prediction. The costs are incurred per user as their history grows.

- **Habitual Profile ($P_u^H(t)$):** This profile is generated by analyzing the window W_t of size k. The input to the LLM is the concatenation of k history items.
  - The complexity is $O(C_{LLM}(k \cdot L))$.
- **Intentional Profile ($P_u^I(t)$):** The intentional profile is updated incrementally for each of the k historical items. For each item h_i, two LLM calls are made: one to generate the short-term profile $p_u^I(i)$ (Equation 2) and one to update the long-term profile $P_u^I(i)$ (Equation 3).
  - The complexity of generating $p_u^I(i)$ is $C_{LLM}(L)$.
  - The complexity of updating $P_u^I(i)$ depends on the length of the previous profile $P_u^I(i-1)$ (denoted as L_P), resulting in a complexity of $C_{LLM}(L_P + L)$.
  - Since this repeats for all k items, the total complexity is dominated by the incremental updates: $O(k \cdot C_{LLM}(L_P + L))$. This is the most computationally intensive part, scaling linearly with the user's window length k.
- **Behavioral Rationale ($R_u^{\text{history}}(h_*)$):** If we assume rationales for all k items in the window are pre-computed, this adds a cost where L_H is the length of the habitual profile.
  - The complexity is $O(k \cdot C_{LLM}(L + L_H + L_P))$.

Table 2: The evolution of an Intentional Profile over time. The table illustrates how the user's behavior (y_t), in response to a given source text (x_t), progressively refines the profile and how the system robustly handles conflicting inputs.

Time	**Incremental Update Process of Intentional Profile ($P_u^I(t)$)**
t = 1	**Source Context (x_1):** *This article introduces a confidence level (CL) statistic to accompany the identification of the most central actor in relational, social network data...* **User's Behavior (y_1):** Reporting a network's most-central actor with a confidence level. **Generated Short-term Intentional Profile ($p_u^I(1)$):** • **Attitude:** Positive evaluation of a title that is direct and methodological. • **Subjective Norm:** Follows the norm of using precise, technical language. • **PBC:** Feels confident in creating a title about statistical methods. **Updated Long-term Intentional Profile ($P_u^I(1)$):** • **Attitude:** The user has a positive evaluation of titles that balance **creativity and accuracy**. • **Subjective Norm:** Feels pressure to conform to norms of **clarity, concision, and accuracy**. • **PBC:** Feels confident due to perceived **expertise in the field**.
t = 5	**Source Context (x_5):** *ISIS' ability to build and maintain a large online community that disseminates propaganda and garners support continues to give their message global reach...* **User's Behavior (y_5):** From Tweets to Intelligence: Understanding the Islamic Jihad Supporting Community on Twitter. **Generated Short-term Intentional Profile ($p_u^I(5)$):** • **Attitude:** Believes a "From X to Y" title structure effectively communicates a process. • **Subjective Norm:** Adheres to the norm of creating catchy titles for topics with broad interest. • **PBC:** Confident in summarizing a complex socio-technical topic. **Updated Long-term Intentional Profile ($P_u^I(5)$):** • **Attitude:** Now values **informativeness and real-world impact**, reinforced by an appreciation for **tailored methodologies to study online communities**. • **Subjective Norm:** Norms are influenced by expectations from **reputable institutions**. • **PBC:** Confidence is reinforced by familiarity with **specific examples of online communities**.
t = 8	**Source Context (x_8):** *Cyber-attacks are cheap, easy to conduct and often pose little risk in terms of attribution, but their impact could be lasting...* **User's Behavior (y_8):** Understanding DDoS cyber-attacks using social media analytics. **Generated Short-term Intentional Profile ($p_u^I(8)$):** • **Attitude:** Positive about using analytics to understand complex technical phenomena. • **Subjective Norm:** Adheres to using a formal, objective tone with technical terms. • **PBC:** Confident in their knowledge of cyber-attacks and social media analytics. **Updated Long-term Intentional Profile ($P_u^I(8)$):** • **Attitude:** Values titles showcasing innovative methodologies for **complex technical phenomena**. • **Subjective Norm:** Strongly adheres to formal conventions, using **precise technical terms**. • **PBC:** Confident in subjects like **cyber-attacks and social media analytics**.
t = 9	**Source Context (x_9):** *Social identities carry widely agreed upon meanings, called stereotypes, that have important effects on social processes...* **User's Behavior (y_9): Girls Rule, Boys Drool:** Extracting Semantic and Affective Stereotypes from Twitter. **Generated Short-term Intentional Profile ($p_u^I(9)$):** • **Attitude:** Shows a temporary positive evaluation of using a provocative, attention-grabbing hook. • **Subjective Norm:** Momentarily deviates from academic norms to appeal to a broader audience. • **PBC:** Confident in using informal language to frame a sociological topic. **Updated Long-term Intentional Profile ($P_u^I(9)$):** • **Attitude:** Primarily values accuracy. **However, for topics with broad social relevance, they may occasionally employ a creative hook**, while ensuring the subtitle remains descriptive. • **Subjective Norm:** While generally adhering to formal conventions, the user **might make rare, context-dependent exceptions** to attract a wider audience. • **PBC:** Confidence extends to creating both formal technical titles and, **when appropriate, more accessible, stylized titles**.

Table 3: A detailed example of Behavioral Rationale generation for a user's past action (User ID: 111381, LaMP-5). The model performs a post-hoc analysis by synthesizing the user's profiles to explain why a specific title was chosen.

Component	Content and Analysis
1. Inputs for Post-Hoc Analysis	
Analyzed Behavior (h_t)	**Source Context** (x_t)**:** "A significant volume of information leaks in organizations are inadvertent, a form of information spillage. Because the leakage of information is driven by the complex interaction of technology, social, and behavioral factors..." **Actual Behavior** (y_t)**:** "Inadvertent leaks: exploration via agent-based dynamic network simulation."
User's Habitual Profile ($P_u^H(t)$)	"The author's habitual title patterns include using action verbs like 'Exploring', 'Understanding', and 'Extracting' to describe the research process, and incorporating key concepts or methodologies like 'agent-based dynamic network simulation' to highlight the study's focus..."
User's Intentional Profile ($P_u^I(t)$)	• **Attitude:** Positive evaluation of titles that balance creativity and accuracy. • **Subjective Norm:** Feels pressure to conform to academic norms of clarity, concision, and accuracy. • **Perceived Behavioral Control:** Confident in generating effective titles due to expertise in the field.
2. Final Generated Behavioral Rationale	
Behavioral Rationale ($R_u^{\textbf{history}}(h_t)$)	The researcher chose the title "Inadvertent leaks: exploration via agent-based dynamic network simulation" due to a combination of habitual writing patterns and deliberate cognitive processes. Their habit of using action verbs like 'Exploring' and incorporating key concepts like 'agent-based dynamic network simulation' influenced the initial title prediction. The TPB factors, particularly the attitude towards balancing creativity and accuracy, reinforced this prediction, leading to a consistent title choice that effectively conveys the research focus and methodology. Ultimately, the habitual response was adopted due to its alignment with the user's values and academic norms.

Table 4: Time complexity analysis of the TRIPLE framework. Here, t is the total number of historical items, k is the window size, L is the average length of a history item, and $C_{LLM}(N)$ is the computational cost of an LLM call on an input of length N.

Phase	Component	Time Complexity
Offline (Profile Construction)	Habitual Profile Generation ($P_u^H(t)$)	$O(C_{LLM}(k \cdot L))$
	Incremental Intentional Profile Update ($P_u^I(t)$)	$O(k \cdot C_{LLM}(L_P + L))$
	Behavioral Rationale Generation (for W_t)	$O(k \cdot C_{LLM}(L_H + L_P + L))$
Online (Real-time Prediction)	Relevant History Retrieval (h_*)	$O(k \cdot \lvert x_n \rvert)$
	Final Prediction & Rationale ($\hat{y}_n, R_u^{\text{predict}}$)	$O(C_{LLM}(\lvert x_n \rvert + L_H + L_P + L + L_R))$

Online/Inference Phase: This is the complexity of making a single prediction for a new query x_n, which directly impacts user-perceived latency.

- **History Retrieval:** TRIPLE uses BM25 to find the most relevant history h_* from the window W_t. The complexity of this on a small set of k documents is negligible compared to the LLM call.
  - The complexity is $O(k \cdot |x_n|)$.
- **Prediction Generation:** A single LLM call is made to generate the prediction $\hat{y}_n$ and its rationale $R_u^{\text{predict}}(x_n)$ (Equation 5). The input consists of the query, both profiles, and the retrieved history-rationale pair.
  - The complexity is $O(C_{LLM}(|x_n| + L_H + L_P + L + L_R))$, where L_R is the length of the retrieved rationale.

The offline computational cost of TRIPLE scales linearly with the length of the window size k, primarily due to the incremental updates of the intentional profile. However, the online inference complexity is independent of the total history length t. It depends only on the size of the distilled profiles and the window size k. This is a significant advantage over naive in-context learning approaches that might include a large number of raw historical examples in the prompt, making them slow for users with long histories. TRIPLE effectively trades higher offline pre-computation cost for fast, constant-time online inference, making it a practical framework for real-world personalization systems.

Dataset and Task Details

To comprehensively evaluate the effectiveness of our proposed framework, we employ the Language Model Personalization (LaMP) benchmark (Salemi et al. 2023), an open-source benchmark that has been specifically designed to assess the capability of language models in generating personalized content. This benchmark encompasses a diverse array of tasks, systematically covering both personalized text classification and generation paradigms, thereby providing a robust testbed for evaluating theory-grounded user modeling approaches. The structural composition and statistical characteristics of the datasets are delineated in Table 5, which offers a comprehensive depiction of the data distribution across different personalization scenarios. In what follows, we provide detailed characterizations of each task, elucidating their specific objectives and evaluation contexts:

- **LaMP-2N: Personalized News Categorization** constitutes a categorical text classification task that necessitates the classification of news articles into one of fifteen distinct categories based on the journalist's historical writing patterns. Given an article authored by a specific user, the model is required to predict its category by leveraging the user's accumulated history of articles and their corresponding categorical assignments, thereby capturing individual journalistic preferences and topical inclinations.
- **LaMP-2M: Personalized Movie Tagging** represents an ordinal text classification task wherein the objective involves predicting one of fifteen semantically meaningful tags for a given movie based on the user's established tagging history. This task specifically evaluates the capability of language models to assign contextually appropriate tags to movie descriptions by exploiting patterns inherent in the user's historical movie-tag associations, thus modeling individual cinematic interpretation preferences.
- **LaMP-3: Personalized Product Rating** is formulated as a five-class text classification problem that requires predicting integer ratings ranging from one to five for a given product review. The fundamental objective centers on inferring user-specific rating behaviors based on their accumulated history of review-rating pairs, thereby assessing the capability of language models to capture nuanced individual rating tendencies and evaluative criteria.
- **LaMP-4: Personalized News Headline Generation** encompasses a text generation task that demands the synthesis of personalized news headlines for given articles, conditioned on the authors' historical article-title pairs. This task specifically evaluates the language model's capacity to replicate and internalize the stylistic nuances, lexical preferences, and structural patterns characteristic of individual authors' headline composition strategies.
- **LaMP-5: Personalized Scholarly Title Generation** constitutes a sophisticated text generation task that involves synthesizing appropriate titles for academic articles based on the author's historical patterns of article-title correspondences. This task extends the exploration of personalized text generation into the scholarly domain, explicitly focusing on the generation of titles for research articles while capturing domain-specific conventions and individual academic writing styles.

It should be noted that LaMP-6 has been excluded from our evaluation due to the unavailability of the dataset in the public domain. Furthermore, in alignment with methodological decisions established in previous studies (Salemi et al. 2023), two additional tasks—namely LaMP-1 and LaMP-7—have been deliberately excluded from our empirical evaluation owing to the fundamental inconsistencies between the formats of user history and query representations, which would compromise the validity of comparative analyses.

Evaluation Metrics

We used standard evaluation metrics, following prior work (Salemi et al. 2023). Classification tasks LaMP-2N/2M use Accuracy and Macro F1-score; the regression task LaMP-3 uses MAE and RMSE; and generation tasks LaMP-4/5 use ROUGE-1, ROUGE-L, and BLEU.

Prompts

Prompt for Personalization Tasks

In this section, we delineate the prompt templates employed throughout our experimental evaluation, wherein textual elements enclosed in {BRACES} serve as placeholders that are systematically replaced with content specific to individual users and their corresponding queries.

To rigorously evaluate the efficacy of our proposed approach, we conduct a comprehensive comparison between TRIPLE and existing in-context learning-based profiling methodologies for personalization. These comparative baselines represent the current state-of-the-art approaches that explicitly construct user profiles and subsequently incorporate them into prompts, thereby enabling LLMs to generate personalized outputs without necessitating computationally intensive fine-tuning procedures. In adherence to the prompt design specifications delineated in prior studies, we have meticulously implemented five baseline methods for comparative analysis: Zero-shot (Salemi et al. 2023), PAG (Richardson et al. 2023), ONCE (Liu et al. 2024), GPG (Zhang 2024), and PURE (Bang and Song 2025).

The structural format of {USER HISTORY} remains consistent across all experimental tasks, conforming to the specification presented in Table 6. This scheme maintains with the methodological framework originally proposed in prior work (Salemi et al. 2023), thereby ensuring comparability and reproducibility of experimental results.

Prediction Prompt Structures for Baselines The prediction prompts for the baseline methods are constructed around a core, non-personalized task instruction, which constitutes the Zero-shot prompt (Salemi et al. 2023). This base prompt contains only the user's current query (x_n) without any historical context.

For personalized predictions, the other baseline methods (PAG, ONCE, GPG, PURE) extend this Zero-shot prompt by prepending user-specific information: a {USER PROFILE} generated by their respective methods, and {RETRIEVED HISTORY}.

Table 7 presents this unified structure, detailing the base task prompts for each LaMP task and showing how they are used to construct both the Zero-shot and personalized prediction prompts. This unified approach ensures a fair and direct comparison of how different profiling strategies enhance the base prompt.

PAG (Profile-Augmented Personalization) Profile-Augmented Personalization (PAG) (Richardson et al. 2023) methodology augments the user's query through the incorporation of an LLM-generated profile summary that synthesizes patterns from the user's history. We implement an extended variant that combines the generated profile with k retrieved items from the user's history. The prompt templates for PAG is presented in Tables 8.

ONCE ONCE (Liu et al. 2024) generates natural language summaries of user preferences based on the user's interaction history. This method employs a single-pass generation strategy to distill user characteristics into profile descriptions. The generated profiles are subsequently incorporated into the prompt alongside the user's query to facilitate personalized predictions. The prompt templates for ONCE is presented in Table 9.

GPG (Guided Profile Generation) Guided Profile Generation (GPG) generates structured personal profiles in natural language by extracting features from the user's historical context. This method transforms sparse personal data into descriptive sentences that represent individual preferences. The generated profiles are incorporated into prompts for personalized predictions. The prompt template for GPG is presented in Table 10.

PURE PURE (Bang and Song 2025) constructs and maintains evolving user profiles by extracting information from user history. The method consists of two core components: a Review Extractor that identifies user preferences from user history, and a Profile Updater that refines user profiles with new information. This approach processes user history incrementally to update user profiles over time. The prompt templates for profile generation and profile updating are presented in Tables 11 and 12.

TRIPLE TRIPLE employs a multi-stage prompting strategy to implement its Dual-Process Theory-based framework. Unlike the single-profile approaches of the baselines, TRIPLE uses a series of distinct prompts to separately construct and then synthesize habitual and intentional profiles. The following tables detail the specific prompts used for each component of our method: habitual and intentional profile generation, intentional profile updating, behavioral rationale generation, and the final personalized prediction.

- Habitual profile generation prompt (Table 13)
- Intentional profile generation prompt (Table 14)
- Intentional profile update prompt (Table 15)
- Behavior rationale generation prompt (Table 16)
- Personalized prediction prompt (Table 17)

Table 5: Dataset statistics of five different personalization tasks in the LaMP benchmark.

Task	Type	# Train	# Validation	# Test	Input Length	Output Length	# Profiles	# Classes
LaMP-2N	Classification	5914	1052	1274	65.40 ± 12.29	-	306.42 ± 286.65	15
LaMP-2M	Classification	5073	1410	1557	92.39 ± 21.95	-	86.76 ± 189.52	15
LaMP-3	Classification	20000	2500	2500	145.14 ± 157.96	-	188.10 ± 129.42	5
LaMP-4	Generation	12527	1925	2376	30.53 ± 12.67	9.78 ± 3.10	287.16 ± 360.62	-
LaMP-5	Generation	9682	2500	2500	152.81 ± 86.60	9.26 ± 3.13	89.61 ± 53.87	-

Table 6: Prompt design for five LaMP tasks. Concat($\cdot$) concatenates the input strings in order, and PPEP($\cdot$) composes the prompt for each retrieved item from the profile. [INPUT] represents the task's input.

Task	Per Profile Entry Prompt (PPEP)	Aggregated Input Prompt (AIP)
LaMP-2N	"the category for the article: "P_i[text]" is "P_i[category]"	concat([PPEP(P_1),…,PPEP(P_n)],",and"). [INPUT]
LaMP-2M	"the tag for the movie: "P_i[description]" is "P_i[tag]"	concat([PPEP(P_1),…,PPEP(P_n)],",and"). [INPUT]
LaMP-3	P_i[score] is the score for "P_i[text]"	concat([PPEP(P_1),…,PPEP(P_n)],",and"). [INPUT]
LaMP-4	"P_i[title]" is the title for "P_i[text]"	concat([PPEP(P_1),…,PPEP(P_n)],",and"). [INPUT]
LaMP-5	"P_i[title]" is the title for "P_i[abstract]"	concat([PPEP(P_1),…,PPEP(P_n)],",and"). Following the given patterns [INPUT]

References

Bang, S.; and Song, H. 2025. LLM-based User Profile Management for Recommender System. *arXiv preprint arXiv:2502.14541*.

Liu, Q.; Chen, N.; Sakai, T.; and Wu, X.-M. 2024. Once: Boosting content-based recommendation with both open- and closed-source large language models. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, 452–461.

Richardson, C.; Zhang, Y.; Gillespie, K.; Kar, S.; Singh, A.; Raeesy, Z.; Khan, O. Z.; and Sethy, A. 2023. Integrating summarization and retrieval for enhanced personalization via large language models. *arXiv preprint arXiv:2310.20081*.

Salemi, A.; Mysore, S.; Bendersky, M.; and Zamani, H. 2023. Lamp: When large language models meet personalization. *arXiv preprint arXiv:2304.11406*.

Zhang, J. 2024. Guided profile generation improves personalization with llms. *arXiv preprint arXiv:2409.13093*.

Table 7: Unified prompt structures for Zero-shot and personalized baseline predictions. The personalized prompts are constructed by prepending user-specific information to the base task prompt.

Component	Content / Structure
Part A: Base Task Prompts (Used in all prediction prompts)	
LaMP-2N	Which category does this article relate to among the following categories? Just answer with the category name without further explanation. categories: `<category>` article: $\{x_n\}$
LaMP-2M	Which tag does this movie relate to among the following tags? Just answer with the tag name without further explanation. tags:`<tag>` description: $\{x_n\}$
LaMP-3	What is the score of the following review on a scale of 1 to 5? just answer with 1, 2, 3, 4, or 5 without further explanation. review: $\{x_n\}$
LaMP-4	Generate a headline for the following article: $\{x_n\}$
LaMP-5	Generate a title for the following abstract of a paper: $\{x_n\}$
Part B: Final Prompt Construction Rules	
Zero-shot	Final Prompt = **Part A**
Baselines (PAG, ONCE, etc.)	Final Prompt = {USER PROFILE} {RETRIEVED HISTORY} **+ Part A**

Table 8: Profile generation prompt for PAG method for five LaMP tasks

Task	Prompt
LaMP-2N	Look at the following past articles this journalist has written and determine the most popular category they write in. Answer in the following form: most popular category: <category>. User History: {USER HISTORY}
LaMP-2M	Look at the following past movies this user has watched and determine the most popular tag they labeled. Answer in the following form: most popular tag: <tag>. User History: {USER HISTORY}
LaMP-3	Based on this user's past reviews, what are the most common scores they give for positive and negative reviews?Answer in the following form: most common positive score: <most common positive score>, most common negative score: <most common negative score>. User History: {USER HISTORY}
LaMP-4	Given this author's previous articles, try to describe a template for their headlines. I want to be able to accurately predict the headline gives one of their articles. Be specific about their style and wording, don't tell me anything generic.User History: {USER HISTORY}
LaMP-5	Given this author's previous publications, try to describe a template for their titles. I want to be able to accurately predict the title of one of the papers from the abstract. Only generate the template description, nothing else. User History: {USER HISTORY}

Table 9: Profile generation prompt for ONCE method for five LaMP tasks

Task	Prompt
All Tasks (LaMP-2N to LaMP-5)	Describe user profile based on user history: {USER HISTORY}

Table 10: Profile generation prompt for GPG method across five LaMP tasks

Task	Prompt
LaMP-2N	I read the following articles and categorized them in chronological order: {USER HISTORY} Given above articles and categories, in one sentence, highlight how this person uniquely categorizes articles, pay attention to their interests and classification patterns.
LaMP-2M	I watched the following movies and gave tags in chronological order: {USER HISTORY} Given above movies and tags, in one sentence, highlight how this person uniquely chooses movie tags, pay attention to their preferences and patterns.
LaMP-3	I purchased the following products and left reviews in chronological order: {USER HISTORY} Given above reviews and ratings, in one sentence, highlight how this person uniquely rates products, pay attention to their rating criteria and standards.
LaMP-4	I wrote the following headlines for articles in chronological order: {USER HISTORY} Given above articles and headlines, in one sentence, highlight how this person uniquely writes headlines, pay attention to their writing style and approach.
LaMP-5	I wrote the following paper titles for abstracts in chronological order: {USER HISTORY} Given above abstracts and titles, in one sentence, highlight how this person uniquely writes academic titles, pay attention to their academic writing style.

Table 11: Profile generation prompt for PURE method for five LaMP tasks

Task	Prompt
LaMP-2N	I read the following articles and categorized them in chronological order : {USER HISTORY} Analyze user's likes/dislikes/key features by referring to their reading preferences. Please provide your analysis in the following JSON format: {{ "likes": ["item1", "item2", ...], "dislikes": ["item1", "item2", ...], "key_features": ["feature1", "feature2", ...] }}
LaMP-2M	I watched the following movies and gave tags in chronological order : {USER HISTORY} Analyze user's likes/dislikes/key features by referring to their movie preferences. Please provide your analysis in the following JSON format: {{ "likes": ["item1", "item2", ...], "dislikes": ["item1", "item2", ...], "key_features": ["feature1", "feature2", ...] }}
LaMP-3	I purchased the following products and left reviews in chronological order : {USER HISTORY} Analyze user's likes/dislikes/key features by referring to their reviews. Please provide your analysis in the following JSON format: {{ "likes": ["item1", "item2", ...], "dislikes": ["item1", "item2", ...], "key_features": ["feature1", "feature2", ...] }}
LaMP-4	I wrote the following headlines for articles in chronological order : {USER HISTORY} Analyze user's writing style/preferences/key features by referring to their headline creation patterns. Please provide your analysis in the following JSON format: {{ "likes": ["item1", "item2", ...], "dislikes": ["item1", "item2", ...], "key_features": ["feature1", "feature2", ...] }}
LaMP-5	I wrote the following paper titles for abstracts in chronological order : {USER HISTORY} Analyze user's writing style/preferences/key features by referring to their title creation patterns. Please provide your analysis in the following JSON format: {{ "likes": ["item1", "item2", ...], "dislikes": ["item1", "item2", ...], "key_features": ["feature1", "feature2", ...] }}

Table 12: Profile update prompt for PURE method for five LaMP tasks

Task	Prompt
LaMP-2N	You are given a reading preference list : {USER PROFILE} Update this list by removing redundant or overlapping information. Note that crucial information should be preserved.
LaMP-2M	You are given a movie preference list : {USER PROFILE} Update this list by removing redundant or overlapping information. Note that crucial information should be preserved.
LaMP-3	You are given a preference list : {USER PROFILE} Update this list by removing redundant or overlapping information. Note that crucial information should be preserved.
LaMP-4	You are given a writing style preference list : {USER PROFILE} Update this list by removing redundant or overlapping information. Note that crucial information should be preserved.
LaMP-5	You are given an academic writing preference list : {USER PROFILE} Update this list by removing redundant or overlapping information. Note that crucial information should be preserved.

Table 13: Generalized prompt template for generating habitual profiles ($\mathcal{P}_{\text{habit}}$).

Component	Prompt Template
Role	You are a top-tier user-profiling expert specializing in Habit theory.
Task Overview	Analyze the user's past behaviors in the provided examples to identify their habitual patterns. Your goal is to find the repetitive, automatic link between specific situational cues in the [Domain-specific Context] and the user's [Domain-specific Behavior].
Input Data	User Behavior Examples: {USER HISTORY}
Habit Analysis	Based on the examples, answer the following: - What situational cues (e.g., topics, sentiments, keywords, narrative features) consistently and automatically trigger a specific [Domain-specific Behavior]? - Describe these recurring cue-response patterns. For some tasks, specific output formats are provided below as a hint. - [Optional: Example Format Hint]
Guidelines	- Each component should be summarized in 3-4 concise sentences. - Focus strictly on automatic, repetitive patterns, not on one-off or deliberate-seeming actions.
Output Format	{{ "habit": "...", }}

Table 14: Generalized prompt template for generating intentional profiles ($\mathcal{P}_{\text{TPB}}$).

Component	Prompt Template
Role	You are a top-tier user-profiling expert specializing in Theory of Planned Behavior (TPB).
Task Overview	Analyze how a given [Domain-specific Context] activated the user's TPB structure to trigger their specific [Domain-specific Behavior].
Input Data	User Behavior Example: {USER HISTORY}
TPB Analysis	Based on TPB theory, analyze how the content influenced the user's decision through three psychological factors: 1. **Attitude** - How did the content shape the user's positive/negative evaluation of this [Domain-specific Behavior]? - What elements made the user believe this behavior would lead to valuable outcomes (e.g., accurate representation, helping others, better discovery)? - How did the content align with the user's personal values about accurate/effective [Domain-specific Behavior]? 2. **Subjective Norm** - What social pressures did the context create for this [Domain-specific Behavior]? - What community standards, or professional/journalistic expectations influenced the decision? - How did the social context create pressure to follow established norms for this behavior? 3. **Perceived Behavioral Control** - How did the content affect the user's confidence in performing this [Domain-specific Behavior]? - What clarity, complexity, or ambiguity factors in the content influenced their confidence? - What elements made the user feel they had the necessary knowledge/skills to perform the behavior accurately?
Guidelines	- Each component should be 1-2 sentences. - Focus on specific textual elements that triggered each TPB component. - Consider how these three factors combined to create the intention for the behavior.
Output Format	{{ "attitude": "...", "subjective_norm": "...", "perceived_behavioral_control": "..." }}

Table 15: Generalized prompt template for updating intentional profiles ($\mathcal{P}_{\text{update}}$).

Component	Prompt Template
Role	You are a user profiling expert specializing in belief system updates based on the Theory of Planned Behavior (TPB).
Task Overview	Your task is to update a user’s existing long-term intentional profile by integrating a new, short-term profile derived from a recent behavior. The goal is to refine the long-term profile into a more accurate and stable representation of the user’s evolving beliefs and intentions.
Input Data	**Current Long-term Profile:** - Attitude: {current_attitude} - Subjective Norm: {current_subjective_norm} - Perceived Behavioral Control: {current_pbc} **New Short-term Profile to Integrate:** - Attitude: {new_attitude} - Subjective Norm: {new_subjective_norm} - Perceived Behavioral Control: {new_pbc}
Update Instructions	Update the long-term profile by following these principles: 1. **Analyze Consistency:** Compare the current long-term profile with the new short-term profile to identify consistencies and conflicts. 2. **Identify Core Beliefs:** Determine which beliefs are stable (repeated) versus transient (new or conflicting). 3. **Refine Profile:** Construct the new long-term profile by: • Strengthening consistent beliefs and evaluations. • Modifying or cautiously integrating new information that does not directly conflict. • Discarding or down-weighting transient or clearly conflicting elements to maintain profile stability.
Guidelines	- The updated description for each component (Attitude, Subjective Norm, PBC) should be written concisely in 2–3 sentences. - Focus only on essential, stable traits to avoid creating an overly lengthy or noisy profile.
Output Format	{{ “attitude”: “...”, “subjective_norm”: “...”, “perceived_behavioral_control”: “...” }}

Table 16: Generalized prompt template for generating behavioral rationales ($\mathcal{P}_{\text{rationale}}$).

Component	Prompt Template
Role	You are an expert in behavioral psychology specializing in Dual-Process Theory (DPT), analyzing the cognitive drivers behind user decisions.
Task Overview	Your task is to conduct a post-hoc analysis of a user’s past action. By integrating the user’s Habitual and Intentional profiles, you must generate a step-by-step rationale explaining *why* the user performed a specific [Domain-specific Behavior].
Input Data	- [Domain-specific Context]: {input_behavior} - Actual [Domain-specific Behavior]: {output_behavior} - Habitual Profile: {habitual_profile} - Intentional Profile (TPB-based): {intentional_profile}
4-Stage Analysis Process	Follow this structured reasoning process: 1. **Habit-based Initial Prediction:** Based on the ‘Habitual Profile‘, what behavior would be expected as an automatic response to the input context? 2. **Intention-based Review:** Compare this habit-based prediction with the ‘Intentional Profile‘. Does the intentional profile support, conflict with, or modify the habitual response? Analyze points of agreement or disagreement. 3. **Self-Verification:** Compare the analyses from both profiles against the ‘Actual Behavior‘. Classify the outcome: Was the action driven by habit, a deliberate intention, or a mix? 4. **Integrated Reasoning:** Synthesize your findings into a final explanation. If there was a conflict, explain which intentional factors (Attitude, Subjective Norm, PBC) likely overrode the habitual tendency, leading to the ‘Actual Behavior‘.
Guidelines	- The final behavior rationale should be a concise synthesis (3-4 sentences) of your 4-stage analysis. - It must explain the **underlying cause** of the user’s action, not just describe the analytical process you followed.
Output Format	{{ ”behavior_rationale”: ”...” }}

Table 17: Generalized prompt template for TRIPLE's personalized prediction, which co-generates a prediction and its corresponding behavioral rationale.

Component	**Prompt Template**
Role	You are an expert behavioral analyst. Your task is to predict the user's behavior for the following task and provide a step-by-step rationale for your prediction, based on the provided user profiles and historical context.
Context & Profiles	**## User Profiles and Context:** **### Habitual Profile:** {habitual_profile_json} **### Intentional Profile (TPB-based):** {intentional_profile_json} **### Relevant Historical Context (In-context Example):** {retrieved_history_with_rationale}
Task Instruction	**## Task:** Please [Task-specific Instruction] Domain-specific Context Label : $\{x_n\}$
Output Generation	First, mentally perform a step-by-step analysis by considering how the user's habitual and intentional profiles interact in this new context, using the historical example as a guide. Based on this analysis, generate a concise behavioral rationale (2-3 sentences) explaining *why* you are making your prediction. Then, provide the final prediction.

Output Format	Return your response in the following JSON format. The ‘behavioral rationale‘ must explain the reasoning behind your prediction.
	
	{{
	    "behavioral_rationale": "...",
	    "prediction_key": "..."
	}}