HAG: Hierarchical Demographic Tree-based Agent Generation for Topic-Adaptive Simulation

ACL 2026 SAN DIEGO

Code Paper

Rongxin Chen[1,2], Tianyu Wu[1,2], Bingbing Xu[1], Jiatang Luo[1,3], Xiucheng Xu[1,2], Huawei Shen[1,2]

[1]State Key Laboratory of AI Safety, Institute of Computing Technology, CAS
[2]University of Chinese Academy of Sciences
[3]School of Advanced Interdisciplinary Sciences, UCAS

Wechat Homepage

Email: chenrongxin24s@ict.ac.cn

1. Introduction

Objective Generate topic-adaptive agent populations that are statistically aligned with real demographic distributions and sociologically coherent at the individual level.

Limitations of existing paradigms

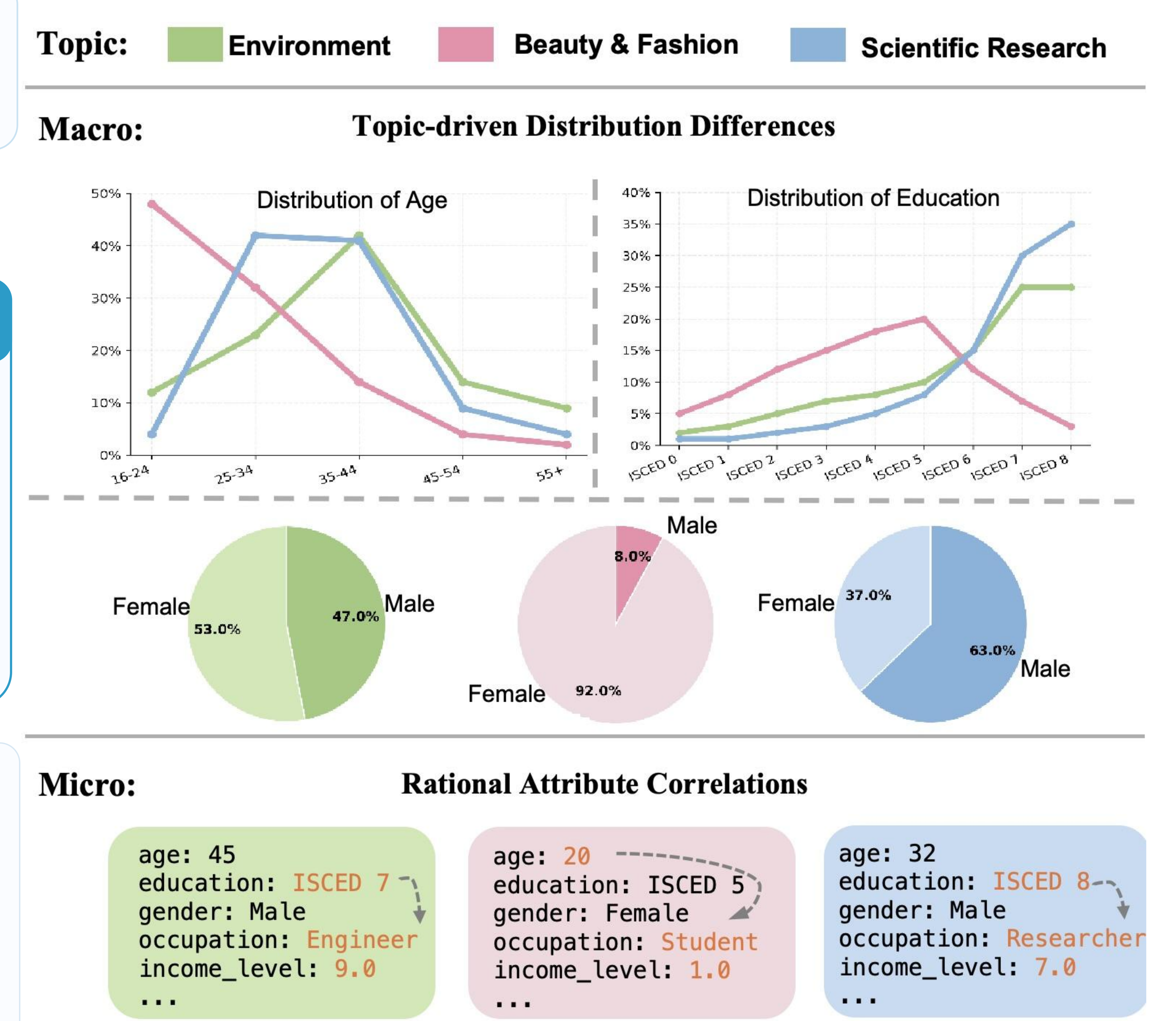
Retrieval methods

- Rely on static data coverage and struggle to adapt to unseen topics.
- Can preserve local realism, but cannot actively align to topic-specific macro distributions.

Pure LLM generation

- Flexible on new topics, but weak in macro-level distribution awareness.
- Often produce unrealistic or internally inconsistent population structures.

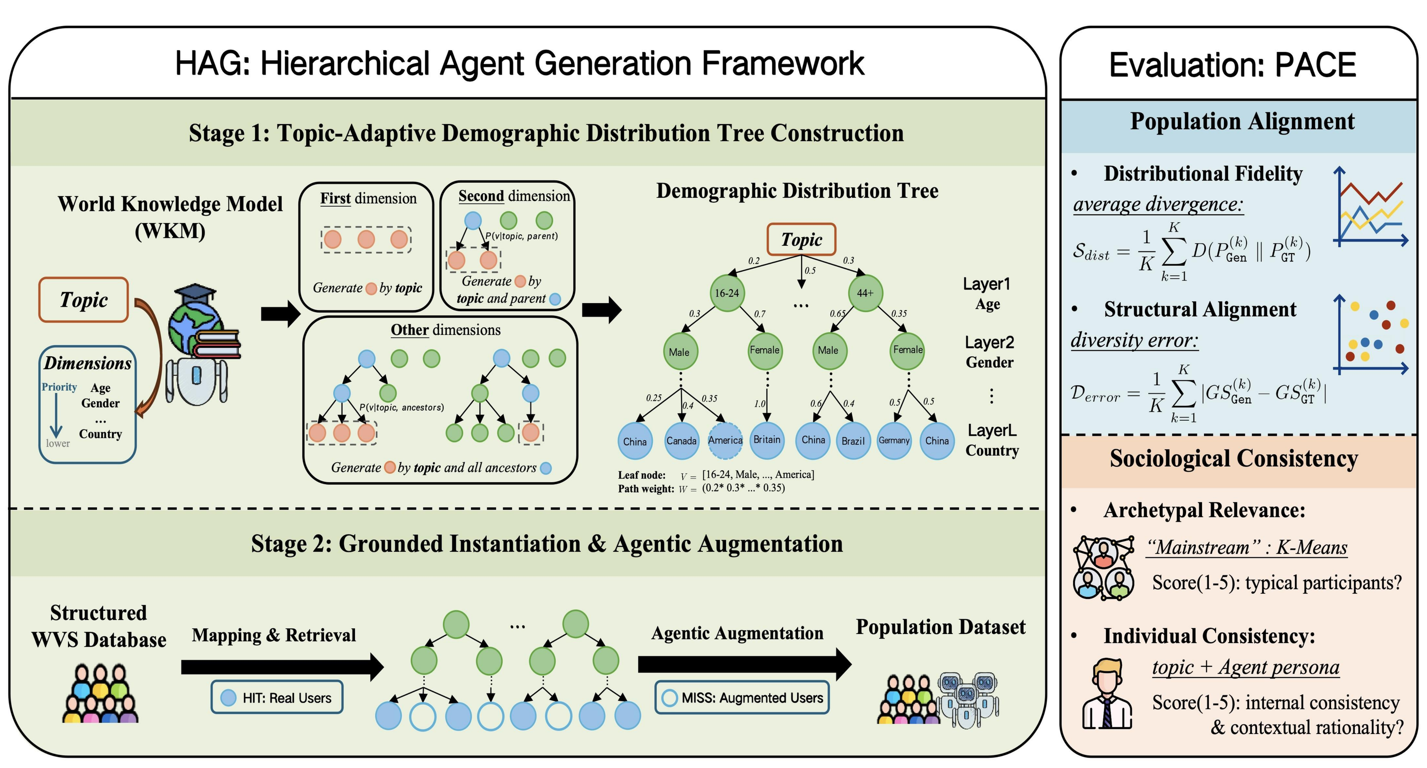
Observation Different topics exhibit different macro-level demographic distributions, and realistic topic-adaptive populations also require rational micro-level attribute correlations. Therefore, a suitable method must capture both macro distribution shifts across topics and coherent dependencies among individual attributes.



2. Methodology

Motivation

How can we jointly model topic-specific population distributions and individual-level consistency, instead of relying on either static retrieval or unconstrained LLM generation alone?



HAG method framework

1. Prioritize topic-relevant demographic dimensions with the World Knowledge Model.
2. Construct a topic-adaptive demographic distribution tree layer by layer.
3. Allocate leaf-level population quotas from path probabilities.

$$W(\vec{v} \mid t) = \prod_{l=1}^{L} P(f^{(l)} = v^{(l)} \mid f^{(1:l-1)} = v^{(1:l-1)}, t)$$

4. Instantiate matched real users from WVS and augment missing profiles agentically.

PACE evaluation framework

1. Population Alignment ↓

Distributional Fidelity:

$$S_{dist} = \frac{1}{K} \sum_{k=1}^{K} D\left(P_{Gen}^{(k)} \,\|\, P_{GT}^{(k)}\right)$$

Structural Alignment:

$$D_{error} = \frac{1}{K} \sum_{k=1}^{K} \left| GS_{Gen}^{(k)} - GS_{GT}^{(k)} \right|$$

2. Sociological Consistency ↑

Archetypal Relevance:
–Are agents typical participants for the topic?

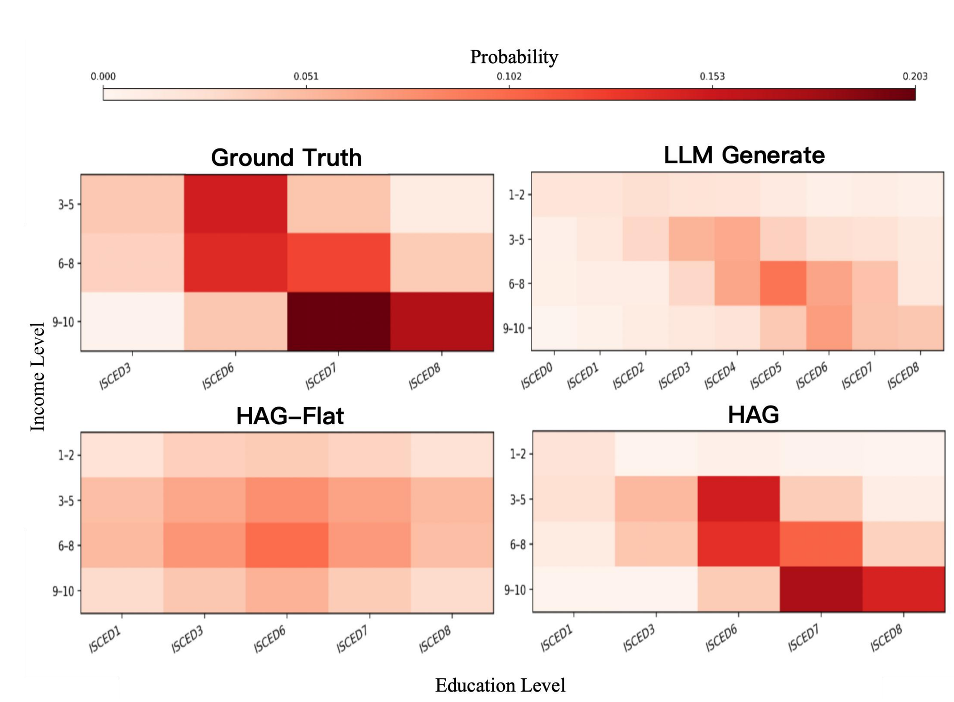
Individual Consistency:
–Are persona and topic internally coherent?

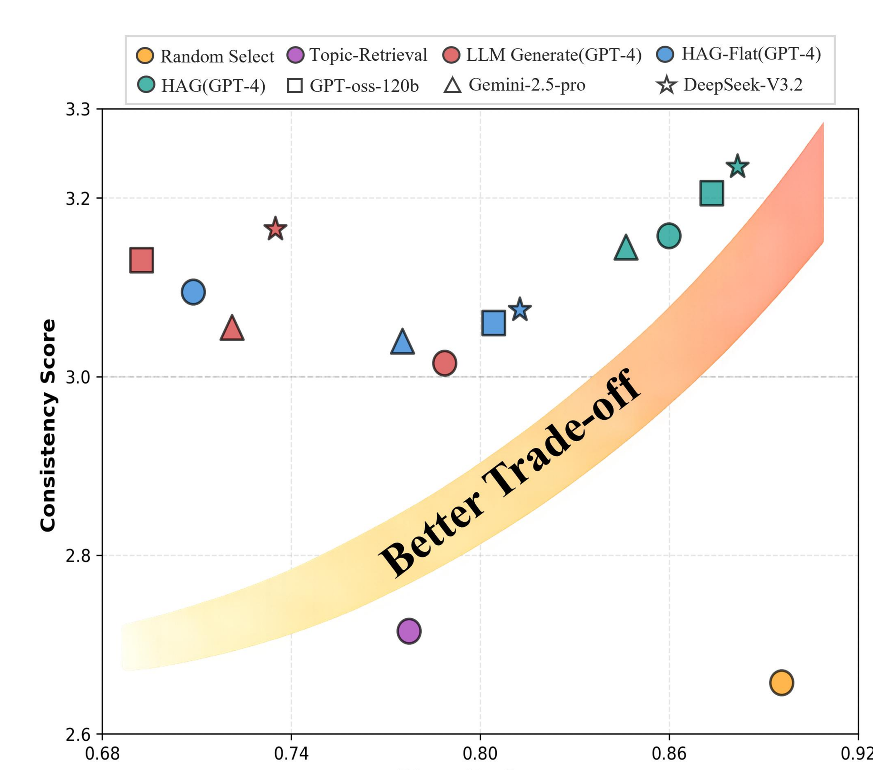
3. Experiments

Main results

Across all three domains, HAG consistently improves both population alignment and sociological consistency over data-based retrieval methods and direct LLM generation baselines. The hierarchical design also outperforms HAG-Flat, validating the importance of modeling conditional dependencies between demographic dimensions.

Model	Method	Bluesky Social Dataset					Amazon Reviews 2023					IMDB Movies User Reviews				
		Alignment ↓			Consistency ↑		Alignment ↓			Consistency ↑		Alignment ↓			Consistency ↑	
		JSD	KL	DivErr	ArchRel	IndCon	JSD	KL	DivErr	ArchRel	IndCon	JSD	KL	DivErr	ArchRel	IndCon
Data-based Retrieval Methods																
——	Random Select	0.628	2.489	0.505	3.000	2.599	0.530	1.286	0.518	3.000	2.878	0.510	1.359	0.535	2.500	3.440
——	Topic-Retrieval	0.578	5.725	0.285	3.250	2.928	0.587	4.035	**0.408**	3.000	3.134	0.576	2.049	0.473	<u>3.000</u>	3.242
LLM-based Generation Methods																
Average	LLM Generate	0.539	2.487	0.466	3.063	3.197	0.451	0.925	0.504	2.750	**3.675**	0.479	**1.041**	0.555	2.875	<u>3.690</u>
	HAG-Flat	<u>0.401</u>	<u>2.436</u>	<u>0.276</u>	<u>3.750</u>	<u>3.324</u>	<u>0.429</u>	<u>0.779</u>	0.439	<u>3.125</u>	3.487	<u>0.398</u>	3.392	<u>0.324</u>	<u>3.000</u>	3.686
	HAG(Our)	**0.345**	**1.657**	**0.263**	**3.813**	**3.617**	**0.414**	**0.759**	<u>0.419</u>	**3.250**	<u>3.642</u>	**0.393**	<u>1.331</u>	**0.322**	**3.125**	**3.783**
GPT-4	LLM Generate	0.559	1.432	0.541	3.250	2.947	0.515	1.185	0.566	2.000	3.228	0.502	1.176	0.575	3.000	3.565
	HAG-Flat	0.401	1.497	0.295	3.750	3.315	0.421	0.784	0.432	3.000	3.547	0.379	0.914	0.417	3.000	3.785
	HAG(Our)	0.354	1.084	0.281	3.750	3.510	0.447	0.815	0.423	3.000	3.628	0.385	0.734	0.342	3.000	3.828
GPT-oss-120b	LLM Generate	0.485	0.967	0.468	3.250	3.363	0.409	0.688	0.473	3.000	3.685	0.461	0.884	0.521	3.000	3.619
	HAG-Flat	0.411	2.518	0.226	3.250	3.089	0.454	0.895	0.439	3.000	3.543	0.304	0.473	0.326	3.000	3.999
	HAG(Our)	0.320	0.619	0.250	4.000	3.712	0.407	0.754	0.434	3.000	3.643	0.297	0.421	0.272	3.500	4.048
Gemini-2.5-Pro	LLM Generate	0.627	6.513	0.400	2.500	2.911	0.417	0.765	0.442	3.000	4.022	0.481	1.010	0.556	3.000	3.890
	HAG-Flat	0.451	4.665	0.320	3.750	3.317	0.423	0.743	0.412	3.500	3.399	0.434	7.231	0.226	3.000	3.507
	HAG(Our)	0.393	4.052	0.259	3.500	3.558	0.401	0.738	0.431	3.000	3.619	0.423	1.034	0.350	3.000	3.652
DeepSeek-V3.2	LLM Generate	0.485	1.035	0.456	3.250	3.565	0.465	1.061	0.536	3.000	3.765	0.474	1.093	0.568	2.500	3.685
	HAG-Flat	0.343	1.063	0.264	4.250	3.576	0.418	0.695	0.474	3.000	3.457	0.476	4.950	0.326	3.000	3.454
	HAG(Our)	0.313	0.873	0.260	4.000	3.690	0.402	0.727	0.388	4.000	3.676	0.467	3.135	0.323	3.000	3.605





Analysis

HAG better matches the ground-truth diagonal trend and high-density regions, capturing conditional dependencies among demographic attributes more faithfully than LLM Generate and HAG-Flat.

HAG stays near the Pareto frontier, maintaining high sociological consistency while preserving structural diversity on an emerging open-world topic.