Formalising Personas for Large Language Models: An Ontological Approach

Andrea Sillano[1,2,*], Pasquale Lisena[1], Raphaël Troncy[1] and Luigi De Russis[2]

[1]*EURECOM, Sophia Antipolis, France*
[2]*Politecnico di Torino, Italy*

Abstract
Large Language Models are increasingly used to simulate identities with distinct personality traits, communication styles, and behavioural patterns across applications such as conversational agents and synthetic user simulation. However, LLM-based personas are commonly specified through ad hoc prompts or biographical sketches, leaving room for ambiguity, inconsistent interpretation, and limited portability. Existing persona ontologies and structured user models only partially address these limitations, as they are not mainly designed to support persona enactment by LLMs. We introduce the LLM Persona ontology, structured following the OWL formalism around two foundational pillars: Identity and Behaviour. We also propose a downstream evaluation using PersonaGym to assess whether ontology-based representations can support persona enactment, as a first step toward formalizing personas for LLM-based personalization.

Keywords
Large Language Models, AI-Personas, Ontology, Semantic Web, RDF

1. Introduction

Personas are widely used across human-centered design and requirements engineering as representations of target users. Cooper [1] introduced personas as hypothetical archetypes of actual users, and subsequent work established them as design artifacts that summarize user goals, behaviors, needs, frustrations, and contexts, supporting communication and design decision-making throughout the development process [2, 3]. With the emergence of Large Language Models (LLMs), personas have taken on a computational role. Beyond supporting human designers, they increasingly serve as inputs that condition model behavior in conversational agents, recommender systems, educational applications, creative-support tools, synthetic user simulation, and decision-support systems [4, 5, 6]. Persona descriptions influence generated responses, explanations, and simulated behaviors, with recent work demonstrating that LLMs can adopt different personalities, communication styles, expertise, and demographic characteristics through persona conditioning [7, 8, 9, 10, 11]. However, LLM personas are still specified through ad hoc prompts with natural language descriptions. These representations are flexible and easy for humans to write, but are often inconsistent. Existing studies also show substantial variability in the structure, level of detail, and quality of generated personas [12, 13, 14].

Existing vocabularies for describing people and agents [15, 16] and ontologies for user modelling [17] were primarily developed to describe users or user groups, rather than the behavioral, linguistic characteristics or external tools and functions of an LLM persona. To address this gap, we present a first version of an OWL ontology for representing LLM personas in a structured form. The ontology[1] enables inspection and retrieval of persona characteristics, supporting queries such as retrieving personas with a given occupation or affiliation, identifying tool usage preferences, agent type, or behavioural constraint during persona enactment. The ontology provides a machine-interpretable representation in

ISWC 2026 Companion Volume, October 25–29, 2026, Bari, Italy
*Corresponding author.
andrea.sillano@polito.it (A. Sillano); pasquale.lisena@eurecom.fr (P. Lisena); raphael.troncy@eurecom.fr (R. Troncy); luigi.derussis@polito.it (L. De Russis)
0009-0007-2485-4042 (A. Sillano); 0000-0003-3094-5585 (P. Lisena); 0000-0003-0457-1436 (R. Troncy); 0000-0001-7647-6652 (L. De Russis)

[1]LLM-persona ontology: https://w3id.org/llm-persona/. Repository: https://github.com/D2KLab/llm_persona_ontology

which persona attributes can be independently modified, validated, and compared, while supporting interoperability and provenance. It is organized around two complementary pillars: *Identity*, which captures who the persona is through demographic, biographical, and role-related information; and *Behaviour*, which models how the persona tends to reason, communicate, and interact. To provide an initial validation, we evaluate the proposed representation using PersonaGym, a benchmark for persona-conditioned LLM agents that measures persona adherence through the PersonaScore metric across 200 personas and 10,000 persona-specific interaction scenarios [18]. This evaluation examines whether an ontology-based representation can support persona enactment in downstream LLM tasks.

2. The LLM Persona ontology

The proposed ontology models the LLM persona on two pillars: *Identity* and *Behaviour.* This division follows the role that persona attributes play during LLM enactment. The *Identity* pillar groups attributes that define the represented entity and are expected to remain relatively stable across tasks. The *Behaviour* pillar groups attributes that describe how the persona is expected to act and decide. The main class is `llmp:LLMPersona` which represents the persona to be enacted by an LLM. The ontology introduces 7 new classes in the LLMP namespace, and 91 properties. It is complemented by controlled vocabularies organised in 12 SKOS concept schemes, with a total of 86 concepts. The ontology reuses schema.org [15], FOAF [16], BIO [19], PROV [20], and SKOS [21] for common entities such as persons, organisations, places, life events, provenance, and controlled vocabularies, avoiding the redefinition of established concepts, and promoting interoperability with existing Semantic Web resources.

Identity pillar The *Identity* pillar captures attributes that characterize the represented entity. The identity information is linked to `llmp:LLMPersona` through `llmp:hasIdentityFeature` (or `llmp:hasIdentityLiteralFeature` for datatype properties) and its more specific subproperties. Literal feature properties are intended to support datatype values and information for which a more specific structured representation is not available. Their use should be considered a fallback, whenever an appropriate ontology property or controlled SKOS concept exists, the structured representation should be preferred. The pillar is organized into three complementary groups. **Personal identity** represents stable characteristics of the persona, including name, age, gender, nationality, occupation, and more. The ontology supports both human and non-human personas through `llmp:hasAgentType`, distinguishing natural persons, fictional characters, and artificial agents. Demographic attributes are explicitly represented instead of being embedded in narrative text, facilitating structured retrieval and inspection [14, 12, 22]. **Biographical identity** models the persona's background and life trajectory through events such as birth, origin, residence, education, career, and others. These are represented through object properties including `llmp:hasLifeEvent`, `llmp:hasBirthEvent`, and `llmp:hasCareerPosition`. This enables personas to encode context beyond static demographic information [23, 24]. **Social identity** represents the persona's relationships with people, groups, and audiences. It includes affiliations, communities, social roles, relationships, and social norms through properties such as `llmp:hasAffiliation`, and `llmp:hasRelationship`. These attributes capture the social context in which the persona operates and are commonly used in LLM-based role-playing and synthetic user generation [5, 6, 25, 26].

Behaviour pillar The *Behaviour* pillar models attributes that govern how a persona acts, reasons, and decides. Behavioural information is linked through `llmp:hasBehaviorFeature` (or `llmp:hasBehaviorLiteralFeature` for datatype properties) and related subproperties. The ontology first represents personality dispositions through `llmp:TraitAssessment`, linking a persona to a personality trait, its intensity, and its textual realization. The current controlled vocabulary adopts the *Big Five dimensions* (i.e., openness, conscientiousness, extraversion, agreeableness, and neuroticism), often used for LLM personality conditioning [27, 28, 10].

Goals, needs, and preferences are modeled as distinct concepts. Goals (`llmp:Goal`) capture the objectives pursued by the persona, needs (`llmp:Need`) represent functional and psychological requirements, while preferences (`llmp:PreferenceStatement`) explicitly encode their target, polarity, and strength. The ontology also represents behavioural tendencies related to decision making and task execution, including aspects such as risk attitude, collaboration style, information seeking, planning, learning, tool usage, and persistence. These dimensions characterize how a persona approaches reasoning and task completion rather than what it knows. Finally, behavioural constraints encode normative instructions that regulate persona enactment. Properties including `llmp:mustDo`, `llmp:shouldDo`, `llmp:mayDo`, `llmp:shouldNotDo`, `llmp:mustNotDo`, and `llmp:outOfCharacterBehavior` distinguish required, recommended, optional, discouraged, forbidden, and out-of-character behaviours. Making these constraints explicit enables selective retrieval and enforcement during persona conditioning.

Ontology Representation

```
ex:persona a llmp:LLMPersona ;
    llmp:hasAgentType llmpv:NaturalPerson ;
    llmp:hasApparentAge "58"^^xsd:integer ;
    llmp:hasOccupation ex:Artist ;
    llmp:hasHomeLocation ex:NewYork ;
    llmp:hasInterest ex:Volunteering ;
    llmp:hasSpecies ex:Human .
```

Natural-language Representation

A 58-year-old artist from New York who spends weekends volunteering.

Figure 1: Illustrative example of an ontology-grounded RDF representation (left) and its corresponding natural-language persona description (right).

3. Initial Evaluation

To assess whether ontology-grounded persona representations preserve persona information during downstream LLM conditioning, we conducted an evaluation using PersonaGym [18]. It is a benchmark to evaluate persona-conditioned LLM, it includes a static representation of 200 personas and 10,000 persona-specific-questions, designed to test persona adherence across different settings and tasks. Accordingly, in our setting it serves to assess whether personas represented using the proposed ontology retain sufficient information for downstream persona enactment, rather than the ontology itself. The framework selects persona-relevant environments, generating task-specific questions or using the predefined ones, collecting responses from the persona-conditioned LLM, and scoring them on 5 different dimensions: *Expected Action*, *Action Justification*, *Linguistic Habits*, *Persona Consistency*, and *Toxicity Control*. Each dimension is evaluated on a 1–5 scale, with higher scores indicating better performance on the corresponding dimension. In addition, it reports an overall *Persona Score*, which aggregates these dimensions into a single measure of persona fidelity.

Ontology-Grounded Persona Representation The first step of the evaluation consists of transposing the 200 natural language persona descriptions from PersonaGym into ontology-grounded representation. Each persona description is processed using OntoGPT [29], which performs ontology-guided information extraction. We provide OntoGPT with the ontology schema and the natural language persona description, and use Gemma 4 31B IT [30] as the underlying LLM to extract ontology-compliant instances. Formally, given a persona description and the ontology, the extraction process produces a structured ontology-conformant YAML representation. The latter is a serialization of ontology-grounded instances that can be directly consumed by LLMs, and not intended as an alternative ontology formalism. Consequently, the extracted persona can be injected into prompts as structured context without requiring additional preprocessing or serialization.

PersonaGym Evaluation Protocol After the ontology-grounded representations are generated, we compare two persona specification conditions:

1. **Natural language condition**: the persona agent is conditioned using the original PersonaGym natural language persona description.
2. **Ontology-grounded condition**: the persona agent is conditioned using the YAML representation generated by OntoGPT, which instantiates the proposed ontology.

For both conditions, we execute the complete PersonaGym benchmark using the same evaluation protocol. For the natural language condition the persona representation is provided as the original textual description, whereas in the ontology-grounded condition the representation is provided in YAML. PersonaGym supplies the predefined evaluation scenarios and persona-specific questions for each benchmark task. Responses are generated by *Gemma 4*, while the evaluation step uses both *Gemma 4* and Qwen3.6 35B a3b [31] as LLM judges, and their scores are combined to compute the five evaluation dimensions. This procedure is repeated for all 200 personas under both conditioning strategies. Table 1 shows that the ontology-grounded representation achieves a *PersonaScore* comparable to the natural language representation. While *Expected Action* and *Toxicity* increase under the ontology-grounded condition, lower scores are observed for *Action Justification*, *Linguistic Habits*, and *Persona Consistency*. The reported standard deviations are of similar magnitude across both conditions, indicating comparable variability over the evaluated personas.

Table 1
Comparison between natural language and ontology-grounded persona representations on PersonaGym. All metrics are reported on a 1–5 scale, with higher values indicating better performance. Δ denotes the difference between the ontology-grounded and natural-language conditions.

Metric	Natural language	Ontology-grounded	Δ
Expected Action	4.1348 ± 0.8056	4.2250 ± 0.8113	+0.0903
Action Justification	3.8937 ± 0.8391	3.7248 ± 1.0166	−0.1689
Linguistic Habits	2.5593 ± 0.6526	2.1898 ± 0.5193	−0.3695
Persona Consistency	3.5480 ± 0.8605	3.4183 ± 0.8495	−0.1298
Toxicity Control	4.1664 ± 1.0214	4.5664 ± 0.8347	+0.4000
PersonaScore	**3.6604 ± 0.4163**	**3.6249 ± 0.3864**	**−0.0356**

Ontology Coverage Analysis We evaluated the completeness of the proposed ontology by measuring the proportion of information contained in the original PersonaGym persona descriptions that could be represented using the proposed ontology. Persona attributes were first automatically extracted, using *Gemma 4* from textual representation and checked against ontology schema to determine whether an equivalent ontology attribute was available. Afterwards, we performed a manual assessment to evaluate the ontology-grounded representational coverage, independently of the automatic extraction process. Across the dataset of **1,536** extracted attributes, **1,510** were successfully represented, corresponding to an overall coverage of **98.3%**. The remaining attributes (1.7%) mainly correspond to highly specific contextual details or isolated facts that are currently outside of the ontology. We manually inspected the remaining 26 attributes to distinguish limitations of the ontology from those of the extraction and comparison pipeline. Of these, 13 (50.0%) were found to be already represented but had not been correctly identified by the automatic comparison, while 6 (23.1%) could be expressed by the existing ontology but were not recovered during ontology-guided extraction. Only 7 attributes (26.9% of the initially missing attributes; 0.46% of all attributes) corresponded to concepts not currently supported by the ontology. Thus, independently of the extraction pipeline, the ontology provides representational coverage for 99.5% of the attributes identified in the PersonaGym descriptions.

We further examined whether these missing attributes could account for the performance differences between the two representations. At the persona level, 177 of the 200 personas retained all attributes identified in their original descriptions, while 23 contained at least one missing attribute. Personas with missing attributes showed a larger mean decrease in PersonaScore than fully preserved personas (-.210 vs. -.013). Overall, these results indicate that the ontology provides high representational coverage, while suggesting that the observed performance differences are unlikely to be explained solely by information loss during ontology grounding.

4. Conclusion

We presented a first version of an ontology for representing LLM personas that provides a semantic representation that can be inspected and modified independently of natural language prompts narratives. The ontology and the resources used for evaluation are available on a public repository[1].

Our initial evaluation shows that the ontology-grounded representation achieves an overall *PersonaScore* comparable to the natural language personas, indicating that structuring persona information does not substantially degrade persona enactment. At the same time, the analysis of the individual evaluation dimensions reveals differences. Behavioural aspects explicitly represented by the ontology, such as those reflected in the *Toxicity* metric, can be associated with improved performance, whereas dimensions that are currently outside the ontology scope – i.e. *Linguistic Habits* and *Action Justification* – exhibit lower performance. These differences also highlight the current coverage of the ontology: behaviour-related aspects are well represented, whereas communication-related characteristics and cognitive models remain underrepresented. Future work will extend the ontology with additional pillars, and evaluate its impact across further persona benchmarks and LLM applications.

Beyond achieving comparable performance, the value of our ontology lies in its structured and editable representation, which enables greater control and facilitates comparison, while offering potential benefits for downstream LLM conditioning and enabling greater transparency than monolithic natural-language persona descriptions.

Future work will extend the ontology with communication and linguistic characteristics, cognitive models, while investigating provenance and auditing mechanisms and evaluating the representation across additional persona benchmarks and LLMs.

Declaration on Generative AI

During the preparation of this work, the authors used GPT5.5 for grammar and spelling check. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] A. Cooper, The Inmates Are Running the Asylum, Macmillan Publishing Co., Inc., USA, 1999.
[2] J. Pruitt, T. Adlin, The Persona Lifecycle: Keeping People in Mind Throughout Product Design, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005. doi:10.1016/B978-0-12-566251-2.X5000-X.
[3] L. Nielsen, Personas: User Focused Design, Human–Computer Interaction Series, 2 ed., Springer, London, 2019. doi:10.1007/978-1-4471-7427-1.
[4] G. Sun, X. Zhan, J. Such, Building Better AI Agents: A Provocation on the Utilisation of Persona in LLM-based Conversational Agents, in: Proceedings of the 6th ACM Conference on Conversational User Interfaces (CUI '24), ACM, Luxembourg, Luxembourg, 2024. doi:10.1145/3640794.3665887.
[5] K. Benharrak, T. Zindulka, F. Lehmann, H. Heuer, D. Buschek, Writer-Defined AI Personas for On-Demand Feedback Generation, in: Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI '24, ACM, 2024, p. 1–18. doi:10.1145/3613904.3642406.
[6] Y. Liu, P. Sharma, M. Oswal, H. Xia, Y. Huang, PersonaFlow: Designing LLM-Simulated Expert Perspectives for Enhanced Research Ideation, in: Proceedings of the 2025 ACM Designing Interactive Systems Conference, DIS '25, Association for Computing Machinery, New York, NY, USA, 2025, p. 506–534. doi:10.1145/3715336.3735789.

[7] J. Chen, X. Wang, R. Xu, S. Yuan, Y. Zhang, W. Shi, J. Xie, S. Li, R. Yang, T. Zhu, A. Chen, N. Li, L. Chen, C. Hu, S. Wu, S. Ren, Z. Fu, Y. Xiao, From Persona to Personalization: A Survey on Role-Playing Language Agents, Transactions on Machine Learning Research (2024). URL: https://openreview.net/forum?id=xrO70E8UIZ, survey Certification.
[8] A. Sillano, L. De Russis, T. Calò, R. Troncy, P. Lisena, Mapping Personas to Text Transformations: A Taxonomy Outline for Content Adaptation, in: From Generation to Simulation: Responsible Use of AI Personas in Human-Centered Design and Research, CEUR Workshop Proceedings, CEUR-WS, 2026. ACM CHI Workshop.
[9] C. Gugliotta, L. La Cava, A. Tagarelli, PersonaGen: A Persona-Driven Open-Ended Machine-Generated Text Dataset, in: Proceedings of the 34th ACM International Conference on Information and Knowledge Management, CIKM '25, Association for Computing Machinery, New York, NY, USA, 2025, p. 6397–6401. doi:10.1145/3746252.3761611.
[10] N. Kovačević, C. Holz, M. Gross, R. Wampfler, The Personality Dimensions GPT-3 Expresses During Human-Chatbot Interactions, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 8 (2024). doi:10.1145/3659626.
[11] K. Li, C. Dai, W. Zhou, S. Hu, APEE: Assessing the Personality Expressions of LLM-Driven Role Play Agent Beyond Self-Perception, in: 2025 28th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2025, pp. 1406–1411. doi:10.1109/CSCWD64889.2025.11033568.
[12] J. Salminen, D. Amin, B. J. Jansen, Using AI for User Representation: An Analysis of 83 Persona Prompts, in: 2025 IEEE/ACS 22nd International Conference on Computer Systems and Applications (AICCSA), IEEE, 2025, p. 1–8. doi:10.1109/aiccsa66935.2025.11315350.
[13] J. Salminen, C. Liu, W. Pian, J. Chi, E. Häyhänen, B. J. Jansen, Deus Ex Machina and Personas from Large Language Models: Investigating the Composition of AI-Generated Persona Descriptions, CHI '24, Association for Computing Machinery, New York, NY, USA, 2024. doi:10.1145/3613904.3642036.
[14] C. Lazik, C. Kauter, I. Nunes, A. Ziglowski, A. Pryma, C. Katins, L. Grunske, T. Kosch, The Good, the Bad, and the Uncanny: Investigating Diversity Aspects of LLM-Generated Personas for Requirements Engineering, in: 2025 IEEE 33rd International Requirements Engineering Conference (RE), 2025, pp. 244–256. doi:10.1109/RE63999.2025.00031.
[15] Schema.org, Schema.org person, https://schema.org/Person, 2024. Accessed 2026.
[16] M. Graves, A. Constabaris, D. Brickley, FOAF: Connecting People on the Semantic Web, Cataloging & Classification Quarterly 43 (2007) 191–202. doi:10.1300/J104v43n03_10.
[17] D. Heckmann, T. Schwartz, B. Brandherm, M. Schmitz, M. von Wilamowitz-Moellendorff, Gumo – The General User Model Ontology, in: L. Ardissono, P. Brna, A. Mitrovic (Eds.), User Modeling 2005, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 428–432.
[18] V. Samuel, H. P. Zou, Y. Zhou, S. Chaudhari, A. Kalyan, T. Rajpurohit, A. Deshpande, K. R. Narasimhan, V. Murahari, PersonaGym: Evaluating Persona Agents and LLMs, 2025. doi:10.18653/v1/2025.findings-emnlp.368.
[19] Ian Davis and Dave Reynolds, BIO: A vocabulary for biographical information, https://vocab.org/bio/, 2010. Accessed: 2026.
[20] T. Lebo, S. Sahoo, D. McGuinness, K. Belhajjame, J. Cheney, D. Garijo, S. Soiland-Reyes, S. Zednik, J. Zhao, PROV-O: The PROV Ontology, W3C Recommendation, 2013. URL: https://www.w3.org/TR/prov-o/.
[21] A. Miles, S. Bechhofer, SKOS Simple Knowledge Organization System Reference, W3C Recommendation, 2009. URL: https://www.w3.org/TR/skos-reference/.
[22] F. Pehar, AI as Synthetic Users in Human-Centred Design: A Systematic Review, in: 2025 MIPRO 48th ICT and Electronics Convention, 2025, pp. 1500–1509. doi:10.1109/MIPRO65660.2025.11131969.
[23] J. Shin, M. A. Hedderich, B. J. Rey, A. Lucero, A. Oulasvirta, Understanding Human-AI Workflows for Generating Personas, in: Proceedings of the 2024 ACM Designing Interactive Systems Conference, DIS '24, Association for Computing Machinery, New York, NY, USA, 2024, p. 757–781.

doi:10.1145/3643834.3660729.

[24] S.-G. Jung, J. Salminen, K. K. Aldous, B. J. Jansen, PersonaCraft: Leveraging language models for data-driven persona development, Int. J. Hum.-Comput. Stud. 197 (2025). doi:10.1016/j.ijhcs.2025.103445.

[25] S. Liang, H. Ren, T. Zhang, X. Liu, Z. Yin, D. Hu, Surveying with AI: Simulating Human Responses Using Personalized LLM Agents and Social Media Data, in: 2025 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2025, pp. 4262–4267. doi:10.1109/SMC58881.2025.11343347.

[26] M. N. Amin, A. Jesser, P. Nemeth, Evolving Knowledge Graphs for Bias Aware Synthetic Persona Representation, in: 2025 17th International Conference on Computer and Automation Engineering (ICCAE), 2025, pp. 389–394. doi:10.1109/ICCAE64891.2025.10980546.

[27] H. Rahman, S. Desai, Vibe Check: Understanding the Effects of LLM-Based Conversational Agents' Personality and Alignment on User Perceptions in Goal-Oriented Tasks, in: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems, CHI '26, Association for Computing Machinery, New York, NY, USA, 2026. doi:10.1145/3772318.3790388.

[28] Q. Zhu, T. Lan, X. Xue, Z. Yu, H. He, TraitsPrompt: Does Personality Traits Influence the Performance of a Large Language Model?, in: 2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2024, pp. 2912–2917. doi:10.1109/CSCWD61410.2024.10580436.

[29] J. H. Caufield, H. Hegde, V. Emonet, N. L. Harris, M. P. Joachimiak, N. Matentzoglu, H. Kim, S. Moxon, J. T. Reese, M. A. Haendel, P. N. Robinson, C. J. Mungall, Structured Prompt Interrogation and Recursive Extraction of Semantics (SPIRES): a method for populating knowledge bases using zero-shot learning, Bioinformatics 40 (2024). doi:10.1093/bioinformatics/btae104.

[30] Gemma Team, Gemma 4 Technical Report, 2026. URL: https://arxiv.org/abs/2607.02770. arXiv:2607.02770.

[31] Qwen Team, Qwen3.6-35B-A3B: Agentic coding power, now open to all, 2026. URL: https://qwen.ai/blog?id=qwen3.6-35b-a3b.