

Novelty Assessment Report

Paper: Agent Data Protocol

PDF URL: <https://openreview.net/pdf?id=tG6301ORHd>

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-01

Abstract

Public research results on large-scale supervised finetuning of AI agents remain relatively rare, since the collection of agent training data presents unique challenges. In this work, we argue that the bottleneck is not a lack of underlying data sources, but that a large variety of data is fragmented across heterogeneous formats, tools, and interfaces. To this end, we introduce the Agent Data Protocol (ADP), a light-weight representation language that serves as an "interlingua" between agent datasets in diverse formats and unified agent training pipelines downstream. The design of ADP is expressive enough to capture a large variety of tasks, including API/tool use, browsing, coding, software engineering, and general agentic workflows, while remaining simple to parse and train on without engineering at a per-dataset level. In experiments, we unified a broad collection of 13 existing agent training datasets into ADP format, and converted the standardized ADP data into training-ready formats for multiple agent frameworks. We performed supervised finetuning on the unified data, and demonstrated an average performance gain of $\sim 20\%$ over corresponding base models, and delivers state-of-the-art or near-SOTA performance on standard coding, browsing, tool use, and research benchmarks, without domain-specific tuning. All code and data are released publicly, in the hope that ADP could help lower the barrier to standardized, scalable, and reproducible agent training.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **Standardizing Heterogeneous Agent Training Datasets for Supervised Fine-Tuning**

A total of **11 papers** were analyzed and organized into a taxonomy with **9 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Unified Data Representation and Protocol Design**
- **Multi-Source Data Integration for Prediction and Learning**
- **Fine-Tuning with Heterogeneous Feedback and Data Quality**
- **Embodiment and Action Space Normalization**
- **Multi-Agent Stochastic Policy Learning**
- **Model Merging for Fine-Tuned Agent Generalization**

Complete Taxonomy Tree

- Standardizing Heterogeneous Agent Training Datasets for Supervised Fine-Tuning Survey Taxonomy
- Unified Data Representation and Protocol Design
 - Agent-Specific Interlingua and Pipeline Unification ★ (3 papers)
 - [0] Agent Data Protocol (Anon et al., 2026) [View paper](#)
 - [2] The agent ohana: Designing unified data and training pipeline for effective agent learning (J Zhang, 2024) [View paper](#)
 - [3] Agent Data Protocol: Unifying Datasets for Diverse, Effective Fine-tuning of LLM Agents (Song, 2025) [View paper](#)
 - Entity-Relationship Based Unified Modeling (1 papers)
 - [6] ER-RAG: Enhance RAG with ER-Based Unified Modeling of Heterogeneous Data Sources (Yikuan Xia, 2025) [View paper](#)
- Multi-Source Data Integration for Prediction and Learning
 - Multi-Source Fusion for Traffic and Spatial Prediction (1 papers)
 - [1] Recent advances in multi-source data fusion for traffic flow prediction: A review (Xianhui Zong, 2025) [View paper](#)
 - Multi-Source Prediction for Autonomous Driving Decisions (1 papers)
 - [8] Prediction guided decision-making for autonomous driving (Liu, 2025) [View paper](#)
- Fine-Tuning with Heterogeneous Feedback and Data Quality
 - Heterogeneous Feedback Format Integration (1 papers)
 - [5] A Framework for Fine-Tuning LLMs using Heterogeneous Feedback (Aponte, 2024) [View paper](#)
 - Multi-Source Code Data Harmonization (1 papers)
 - [9] AlchemistCoder: Harmonizing and Eliciting Code Capability by Hindsight Tuning on Multi-source Data (Kai Chen, 2024) [View paper](#)
- Embodiment and Action Space Normalization (1 papers)
 - [4] Scaling proprioceptive-visual learning with heterogeneous pre-trained transformers (Xinlei Chen, 2024) [View paper](#)
- Multi-Agent Stochastic Policy Learning (2 papers)
 - [7] Normalizing flow policies for multi-agent systems (Xiaobai Ma, 2020) [View paper](#)
 - [10] Normalizing Flow Policies for Multi-agent (X Ma, 2020) [View paper](#)
- Model Merging for Fine-Tuned Agent Generalization (1 papers)
 - [11] Agentmerge: Enhancing generalization in fine-tuned llm agents (M Thakkar, n.d.) [View paper](#)

Narrative

Core task: Standardizing heterogeneous agent training datasets for supervised fine-tuning. The field addresses the challenge of unifying diverse agent training data—spanning web navigation, embodied robotics, multi-agent coordination, and other domains—into formats suitable for large-scale supervised learning. The taxonomy reveals several complementary directions: Unified Data Representation and Protocol Design focuses on creating common schemas and interlingua that allow disparate agent trajectories to be expressed in a shared format, as exemplified by Agent Data Protocol[0] and Agent Ohana[2]. Multi-Source Data Integration for Prediction and Learning tackles the fusion of heterogeneous sensor streams or behavioral logs to improve predictive models, while Fine-Tuning with Heterogeneous Feedback and Data Quality examines how to leverage mixed-quality annotations and varied reward signals during training. Embodiment and Action Space Normalization deals with aligning low-level motor commands across different robot morphologies, Multi-Agent Stochastic Policy Learning explores coordination under uncertainty, and Model Merging for Fine-Tuned Agent Generalization investigates techniques to combine separately fine-tuned agent policies into a single generalist system.

A particularly active line of work centers on protocol-level unification: Agent Data Protocol[0] and its closely related predecessor Agent Data Protocol[3] both propose standardized schemas that enable cross-domain agent datasets to be pooled and reused, reducing the overhead of domain-specific preprocessing. This contrasts with approaches like Heterogeneous Feedback[5], which emphasize handling variable annotation quality rather than enforcing a single representational standard. Meanwhile, Agent Ohana[2] demonstrates how a unified data layer can facilitate large-scale multi-task agent training by harmonizing action spaces and observation formats. The original paper, Agent Data Protocol[0], sits squarely within the Unified Data Representation branch, offering an interlingua that bridges web-based and embodied agent tasks. Compared to Agent Ohana[2], which also targets cross-domain unification, Agent Data Protocol[0] places stronger emphasis on protocol design and pipeline modularity, making it easier to integrate new data sources incrementally. This work addresses a central bottleneck in agent learning: the lack of a common currency for expressing diverse agent experiences in a way that supervised fine-tuning can readily exploit.

Related Works in Same Category

The following **2 sibling papers** share the same taxonomy leaf node with the original paper:

1. The agent ohana: Designing unified data and training pipeline for effective agent learning

Authors: J Zhang, T Lan, Z Liu, W Yao, J Tan | **Year/Venue:** 2024 | **URL:** [View paper](#)

Abstract

∆ The lack of standardized formats introduces complexities in ∆ to consolidate trajectories from heterogeneous data sources. ∆ this finetuning process, we adopt a supervised fine-tuning ∆

Relationship Analysis

Both papers belong to the Agent-Specific Interlingua and Pipeline Unification category, introducing lightweight representation languages to consolidate heterogeneous agent trajectories into unified training pipelines. They overlap in addressing the challenge of standardizing multi-turn agent data from diverse sources (web browsing, coding, tool use) through structured action-observation formats and enabling cross-dataset training. The key difference is that the original paper (ADP) focuses on a Pydantic-based protocol with explicit action types (API, Code, Message) and observation types (Text, Web), while Agent Ohana emphasizes a JSON-based homogeneous multi-turn format with input-output-observation steps and includes an AgentRater component for trajectory quality assessment.

2. Agent Data Protocol: Unifying Datasets for Diverse, Effective Fine-tuning of LLM Agents

Authors: Song, Yueqi, Yueqi Song, Sheikh, Zaid, et al. (44 authors total) | **Year/Venue:** 2025 | **URL:** [View paper](#)

Abstract

Public research results on large-scale supervised finetuning of AI agents remain relatively rare, since the collection of agent training data presents unique challenges. In this work, we argue that the bottleneck is not a lack of underlying data sources, but that a large variety of data is fragmented across heterogeneous formats, tools, and interfaces. To this end, we introduce the agent data protocol (ADP), a light-weight representation language that serves as an "interlingua" between agent datas...

∆ Similarity Notice

This paper is highly similar to the original paper; it may be a variant or near-duplicate. Please manually verify.

Contributions Analysis

Overall novelty summary. The paper introduces the Agent Data Protocol (ADP), a lightweight interlingua designed to unify heterogeneous agent training datasets across diverse formats and interfaces. Within the taxonomy, it resides in the 'Agent-Specific Interlingua and Pipeline Unification' leaf, which contains three papers total. This leaf sits under the broader 'Unified Data Representation and Protocol Design' branch, indicating a moderately populated research direction focused on standardized schemas for agent data. The taxonomy reveals that protocol-level unification is an active but not overcrowded area, with sibling work like Agent Ohana addressing similar cross-domain challenges.

The taxonomy structure shows that the paper's approach contrasts with neighboring branches such as 'Multi-Source Data Integration for Prediction and Learning', which emphasizes sensor fusion and forecasting rather than training pipeline unification. The 'Fine-Tuning with Heterogeneous Feedback and Data Quality' branch addresses mixed-quality annotations but does not enforce a single representational standard, while 'Embodiment and Action Space Normalization' focuses on low-level motor command alignment. The scope notes clarify that ADP's interlingua design excludes domain-specific normalization or retrieval-augmented generation, positioning it as a protocol-first solution distinct from prediction-centric or embodiment-specific methods.

Among the three contributions analyzed, the Agent Data Protocol itself was examined against ten candidates, with three appearing to provide overlapping prior work. The unified collection of thirteen datasets and the empirical validation each faced ten candidates, with one refutable match per contribution. These statistics reflect a limited search scope of thirty total candidates examined, not an exhaustive literature review. The protocol contribution shows the most substantial prior overlap, suggesting that standardized agent data formats have been explored before, though the specific design choices and scope of ADP may differ. The dataset collection and validation contributions appear more novel within the examined sample.

Based on the top-thirty semantic matches and citation expansion, the work occupies a moderately explored niche within agent training standardization. The taxonomy indicates that while unified data representation is an established direction, the specific interlingua approach for agent pipelines remains relatively sparse. The analysis does not cover all possible prior work in agent learning or data harmonization, and a broader search might reveal additional overlapping efforts. The contribution-level statistics suggest incremental novelty in protocol design, with stronger differentiation in the empirical validation and dataset unification aspects.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Agent Data Protocol (ADP)

Description: ADP is a standardized schema implemented as Pydantic objects that unifies heterogeneous agent training datasets into a common format. It represents agent trajectories as sequences of actions (API, code, message) and observations (text, web), enabling conversion from diverse raw datasets to multiple agent frameworks without per-dataset engineering.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Aguviz: Unified Pure Vision Agents for Autonomous GUI Interaction

URL: [View paper](#)

Brief Assessment

Aguviz[13] focuses on vision-based GUI agents with multimodal grounding annotations, not on standardizing heterogeneous text-based agent training datasets across diverse formats like ADP does.

2. Transferable Tactile Transformers for Representation Learning Across Diverse Sensors and Tasks

URL: [View paper](#)

Brief Assessment

Tactile Transformers[17] addresses tactile sensor data unification for robotics manipulation tasks, not agent training datasets for LLM-based agents involving API calls, code generation, or web browsing.

3. AgentOhana: Design Unified Data and Training Pipeline for Effective Agent Learning

URL: [View paper](#)

Prior Art Analysis

AgentOhana[16] demonstrates prior work on standardizing heterogeneous agent training datasets into a unified format. The candidate paper explicitly describes aggregating agent trajectories from distinct environments and meticulously standardizing them into a consistent format, which directly parallels the original paper's ADP contribution. Both papers address the same core problem of dataset heterogeneity and propose standardized schemas as solutions. The candidate's work on data unification and standardization predates or coincides with the original paper's submission timeline, establishing that the original authors were not the first to propose this approach.

Evidence

Evidence 1 - **Rationale:** Both papers propose standardized schemas to unify heterogeneous agent datasets. AgentOhana[16] explicitly describes standardizing and unifying trajectories into a consistent format, which is functionally equivalent to the original paper's ADP standardized schema. - **Original:** we introduce the agent data protocol (adp), a standardized expressive representation language for agent data. By converting heterogeneous datasets into adp, it makes it simple to generate large-scale and diverse data for a variety of downstream training pipelines. - **Candidate:** we introduce `agentohana` as a comprehensive solution to address these challenges. `agentohana` aggregates agent trajectories from distinct environments, spanning a wide array of scenarios. it meticulously standardizes and unifies these trajectories into a consistent format, streamlining...

Evidence 2 - **Rationale:** Both papers identify the same fundamental problem (heterogeneous, fragmented datasets) and propose standardization as the solution. AgentOhana[16] addresses the heterogeneous nature of diverse data sources, matching the original paper's motivation for creating ADP. - **Original:** existing datasets are fragmented, with inconsistent formats and representations, making it difficult to combine, share, and leverage them effectively, thus they remain underutilized. to address this gap, we introduce the agent data protocol (adp), a standardized expressive representation language fo... - **Candidate:** fully harnessing the potential of llms for agent-based tasks presents inherent challenges due to the heterogeneous nature of diverse data sources featuring multi-turn trajectories. in this paper, we introduce `agentohana` as a comprehensive solution to address these challenges.

Evidence 3 - **Rationale:** Both papers implement technical solutions for standardizing agent data. While the original uses Pydantic schemas, AgentOhana[16] creates a consistent format with a generic data loader, both serving the same purpose of unifying heterogeneous agent trajectories. - **Original:** adp is implemented as pydantic schemas that express actions and observations corresponding to common agent use cases such as communicating, browsing, coding, and miscellaneous tool calling, coupled with strict automated validation to maintain high data quality. - **Candidate:** it meticulously standardizes and unifies these trajectories into a consistent format, streamlining the creation of a generic data loader optimized for agent training.

4. Motion tracks: A unified representation for human-robot transfer in few-shot imitation learning

URL: [View paper](#)

Brief Assessment

Motion Tracks[12] focuses on a unified action representation (2D motion trajectories) for human-to-robot transfer in imitation learning, not on standardizing heterogeneous agent training datasets across diverse formats and frameworks.

5. The agent ohana: Designing unified data and training pipeline for effective agent learning

URL: [View paper](#)

Prior Art Analysis

Agent Ohana[2] demonstrates prior work on standardized schemas for unifying heterogeneous agent training datasets. The candidate paper introduces a homogeneous multi-turn data format that consolidates trajectories from diverse data sources into a consistent structure, implemented as JSON dictionaries with standardized components (input, output, next observation). This directly parallels ADP's goal of creating a unified representation for agent trajectories. Both papers address the same fundamental challenge: converting heterogeneous agent datasets into a standardized format to enable unified training pipelines. The candidate's approach of defining structured 'steps' with input/output/observation components is functionally equivalent to ADP's action/observation sequences.

Evidence

Evidence 1 - **Rationale:** Both papers explicitly describe creating a standardized representation to unify heterogeneous agent datasets. Agent Ohana[2] was published earlier and addresses the identical problem of dataset heterogeneity through standardization. - **Original:** we introduce the agent data protocol (adp), a light-weight representation language that serves as an "interlingua" between agent datasets in diverse formats and unified agent training pipelines downstream. the design of adp is expressive enough to capture a large variety of tasks, including api/tool... - **Candidate:** we introduce agentohana as a comprehensive solution to address these challenges. agentohana aggregates agent trajectories from distinct environments, spanning a wide array of scenarios. it meticulously standardizes and unifies these trajectories into a consistent format, streamlining the creation of...

Evidence 2 - **Rationale:** Both papers implement their standardization as structured schemas (Pydantic vs JSON dictionaries) that capture agent actions and observations. The technical approach is nearly identical. - **Original:** adp is implemented as pydantic1 schemas that express actions and observations corresponding to common agent use cases such as communicating, browsing, coding, and miscellaneous tool calling, coupled with strict automated validation to maintain high data quality - **Candidate:** we propose a unified agent data format, as depicted in row 2 of figure 2, showcasing our proposed unified trajectory data format. we construct a homogeneous

json dictionary format to encapsulate all relevant content of each trajectory. concretely, our format incorporates all important elements such ...

6. xlam: A family of large action models to empower ai agent systems

URL: [View paper](#)

Prior Art Analysis

xLAM[14] demonstrates that prior work exists on standardized representation schemas for unifying heterogeneous agent training datasets. The candidate paper describes a unified data format that standardizes diverse agent datasets into a common function-calling style format with modular components (task instruction, available tools, format instruction, etc.). This unified format is explicitly designed to be 'compatible with various environments and tasks' and enables 'fine-grained data augmentation and quality verification.' The paper also describes converting heterogeneous datasets from diverse environments into this standardized format, which directly parallels ADP's goal of unifying heterogeneous datasets into a common schema.

Evidence

Evidence 1 - **Rationale:** Both papers describe structured schemas with specific components for representing agent actions and interactions, demonstrating prior work on standardized agent data representation. - **Original:** adp is implemented as pydantic schemas that express actions and observations corresponding to common agent use cases such as communicating, browsing, coding, and miscellaneous tool calling - **Candidate:** as illustrated in figure 3, the unified format consists of several modules: task instruction, available tools, format instruction, few-shot examples, query, and steps. specifically, the available tools define the agent's action space, and the format instruction specifies the output format the agent ...

Evidence 2 - **Rationale:** Both papers emphasize that their standardized formats are designed to be expressive and compatible across diverse agent tasks and environments. - **Original:** the design of adp is expressive enough to capture a large variety of tasks, including api/tool use, browsing, coding, software engineering, and general agentic workflows - **Candidate:** this unified format is compatible with various environments and tasks, making our data processing pipeline adaptable to different datasets and scalable to large amounts of data. moreover, the modularized design allows for fine-grained data augmentation and quality verification

Evidence 3 - **Rationale:** Both papers describe converting multiple heterogeneous datasets into a unified standardized format, demonstrating that xLAM[14] implemented a similar unification approach prior to the original paper. - **Original:** we unified a broad collection of 13 existing agent training datasets into adp format, and converted the standardized adp data into training-ready formats for multiple agent frameworks. - **Candidate:** existing agent datasets are collected from diverse environments and designed in various formats... function-calling formats form the basis for how models understand and execute tasks, motivating us to design our unified data format in a function-calling style.

7. A unified framework for unsupervised Reinforcement Learning algorithms

URL: [View paper](#)

Brief Assessment

Unified Reinforcement Learning[18] focuses on unifying unsupervised RL algorithms through successor measures and state abstractions for policy learning, not on standardizing heterogeneous agent training dataset formats or conversion pipelines for supervised fine-tuning.

8. UniTR: A Unified Framework for Joint Representation Learning of Trajectories and Road Networks

URL: [View paper](#)

Brief Assessment

UniTR[19] focuses on joint representation learning for urban spatial-temporal data (road networks and trajectories), not on standardizing agent training datasets or creating interlingua schemas for heterogeneous agent data formats.

9. Agent lumos: Unified and modular training for open-source language agents

URL: [View paper](#)

Brief Assessment

Agent Lumos[20] focuses on a modular agent architecture with planning and grounding modules for task execution, not on standardizing heterogeneous datasets into a unified schema like ADP.

10. You only learn one representation: Unified network for multiple tasks

URL: [View paper](#)

Brief Assessment

YOLOR[15] focuses on unified representations for multi-task learning in computer vision (object detection), not on standardizing heterogeneous agent training datasets or trajectory representations for LLM agents.

Contribution 2: Unified collection of 13 agent training datasets

Description: The authors implemented converters to transform 13 pre-existing agent datasets (covering coding, software engineering, tool use, and web browsing) into the ADP standardized format, and further converted ADP data into training formats for three different agent architectures (OpenHands, SWE-Agent, AgentLab).

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Unifying 3d vision-language understanding via promptable queries

URL: [View paper](#)

Brief Assessment

Promptable Queries[24] focuses on 3D vision-language understanding with scene representations (voxels, point clouds, multi-view images), not on converting agent training datasets across frameworks for coding, software engineering, tool use, and web browsing tasks.

2. AgentOhana: Design Unified Data and Training Pipeline for Effective Agent Learning

URL: [View paper](#)

Prior Art Analysis

AgentOhana[16] demonstrates prior work on aggregating and converting multiple agent datasets into unified training formats. The candidate paper explicitly describes aggregating agent trajectories from distinct environments spanning a wide array of scenarios and creating a generic data loader optimized for agent training. This directly parallels the original paper's contribution of implementing converters for 13 datasets. The candidate's approach of aggregating diverse data sources and unifying them for training purposes establishes that similar work existed before the original paper's submission.

Evidence

Evidence 1 - **Rationale:** Both papers describe aggregating multiple agent datasets and converting them into unified formats for training. AgentOhana[16] aggregates trajectories from distinct environments, which is functionally equivalent to the original paper's collection of 13 datasets from diverse sources. - **Original:** as a first step to demonstrate the practical utility of adp, we implement converters from 13 preexisting datasets into adp, and converters from adp to 3 different agent architectures, demonstrating its generality. - **Candidate:** \textit{agentohana} aggregates agent trajectories from distinct environments, spanning a wide array of scenarios. it meticulously standardizes and unifies these trajectories into a consistent format, streamlining the creation of a generic data loader optimized for agent training.

Evidence 2 - **Rationale:** Both papers describe the process of unifying multiple datasets and converting them into training-ready formats. AgentOhana[16]'s aggregation and standardization of trajectories from distinct environments directly parallels the original paper's unification of 13 datasets. - **Original:** we unified a broad collection of 13 existing agent training datasets into adp format, and converted the standardized adp data into training-ready formats for multiple agent frameworks. - **Candidate:** \textit{agentohana} aggregates agent trajectories from distinct environments, spanning a wide array of scenarios. it meticulously standardizes and unifies these trajectories into a consistent format, streamlining the creation of a generic data loader optimized for agent training.

3. Unified-io: A unified model for vision, language, and multi-modal tasks

URL: [View paper](#)

Brief Assessment

Unified-IO[23] focuses on unifying vision and language tasks through a common token representation for multimodal AI, not on converting agent training datasets across different agent frameworks or standardizing agent trajectory data.

4. Scene transformer: A unified architecture for predicting multiple agent trajectories

URL: [View paper](#)

Brief Assessment

Scene Transformer[25] focuses on motion prediction for autonomous driving agents (vehicles, pedestrians), not on converting diverse agent training datasets into unified formats for LLM agent fine-tuning across coding, software engineering, tool use, and web browsing domains.

5. Towards learning a generalist model for embodied navigation

URL: [View paper](#)

Brief Assessment

Generalist Navigation[21] focuses on embodied navigation in 3D environments (e.g., CVDN, SOON, R2R, REVERIE) rather than converting agent datasets across coding, software engineering, tool use, and web browsing into standardized formats for multiple agent architectures.

6. Agent Data Protocol: Unifying Datasets for Diverse, Effective Fine-tuning of LLM Agents

URL: [View paper](#)

Brief Assessment

Agent Data Protocol[3] is the same paper as the original submission. The candidate text is identical to the original paper's abstract, indicating this is not prior work but the same work being evaluated.

7. Unified world models: Coupling video and action diffusion for pretraining on large robotic datasets

URL: [View paper](#)

Brief Assessment

Unified World Models[22] focuses on robotic imitation learning by combining video and action diffusion for robot policy learning, not on converting agent training datasets across software frameworks or standardizing data formats for LLM agents.

8. Agent lumos: Unified and modular training for open-source language agents

URL: [View paper](#)

Brief Assessment

Agent Lumos[20] converts ground-truth reasoning steps from existing benchmarks into training annotations (around 56k samples), but does not implement converters for 13 pre-existing agent datasets into a standardized format across multiple agent frameworks.

9. Decoding Time Series with LLMs: A Multi-Agent Framework for Cross-Domain Annotation

URL: [View paper](#)

Brief Assessment

Time Series Agents[26] focuses on multi-agent annotation systems for time series data, not on converting agent training datasets into unified formats across frameworks.

10. Graphonomy: Universal human parsing via graph transfer learning

URL: [View paper](#)

Brief Assessment

Graphonomy[27] focuses on human parsing tasks (segmenting body parts and clothing in images) using graph transfer learning across three human parsing datasets. This is fundamentally different from the original paper's work on converting agent training datasets for coding, software engineering, tool use, and web browsing tasks into standardized formats.

Contribution 3: ADP Dataset V1 and empirical validation

Description: The authors created and released the largest publicly available agent training dataset (1.3M trajectories) by unifying data through ADP. Supervised fine-tuning experiments show approximately 20% average improvement over base models and achieve state-of-the-art or near-SOTA results across multiple benchmarks without domain-specific tuning.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. AgentOhana: Design Unified Data and Training Pipeline for Effective Agent Learning

URL: [View paper](#)

Prior Art Analysis

AgentOhana[16] demonstrates prior work on creating unified agent training datasets and validating their effectiveness through model training. The candidate paper describes training a large action model (XLAM-v0.1) on unified data that demonstrates exceptional performance across various benchmarks. This directly parallels the original paper's contribution of creating a unified dataset and showing performance improvements through supervised fine-tuning. The candidate's empirical validation of unified agent training data establishes that similar experimental work existed before the original paper.

Evidence

Evidence 1 - **Rationale:** Both papers describe training models on unified agent data and demonstrating strong performance across benchmarks. AgentOhana[16]'s XLAM-v0.1 model showing exceptional performance across various benchmarks parallels the original paper's empirical validation showing ~20% improvements and SOTA results. - **Original:** we performed supervised finetuning on the unified data, and demonstrated an average performance gain of ~20% over corresponding base models, and delivers state-of-the-art or near-sota performance on standard coding, browsing, tool use, and research benchmarks, without domain-specific tuning. - **Candidate:** leveraging the data unification, our training pipeline maintains equilibrium across different data sources and preserves independent randomness across devices during dataset partitioning and model training. additionally, we present `\textbf{xlam-v0.1}`, a large action model tailored for ai agents, whi...

Evidence 2 - **Rationale:** Both papers describe creating large-scale unified datasets for agent training. AgentOhana[16] aggregates and standardizes agent trajectories from diverse environments, which is the same type of contribution as the original paper's ADP Dataset V1. - **Original:** based on this, we create and release the largest publicly available dataset for agent training, consisting of 1.3m training trajectories, dubbed the adp dataset v1. - **Candidate:** `\textit{agentohana}` aggregates agent trajectories from distinct environments, spanning a wide array of scenarios. it meticulously standardizes and unifies these trajectories into a consistent format

Evidence 3 - **Rationale:** Both papers validate their unified datasets through empirical training experiments showing performance improvements across multiple benchmarks. AgentOhana[16]'s demonstration of exceptional performance across various benchmarks parallels the original paper's validation across coding, browsing, and tool use tasks. - **Original:** our experiments show training agents using adp leads to significant performance improvements across diverse domains, including coding (swe-bench verified), web browsing (webarena), research (gaia), and agentic tool use (agentbench) - **Candidate:** leveraging the data unification, our training pipeline maintains equilibrium across different data sources and preserves independent randomness across devices during dataset partitioning and model training. additionally, we present `\textbf{xlam-v0.1}`, a large action model tailored for ai agents, whi...

2. Agentbank: Towards generalized llm agents via fine-tuning on 50000+ interaction trajectories

URL: [View paper](#)

Brief Assessment

AgentBank[34] focuses on fine-tuning on 50,000+ trajectories with a unified agent interaction format, but the extremely limited context provided (only two sentence fragments) does not contain sufficient technical detail to assess whether it challenges the novelty of creating a 1.3M trajectory dataset through ADP standardization or the specific empirical validation methodology.

3. Agent-flan: Designing data and methods of effective agent tuning for large language models

URL: [View paper](#)

Brief Assessment

Agent-FLAN[28] focuses on designing data and methods for agent tuning but does not provide sufficient detail in the available context to assess whether it challenges the novelty of creating a 1.3M trajectory unified dataset or the specific empirical validation approach using ADP standardization.

4. Agent Data Protocol: Unifying Datasets for Diverse, Effective Fine-tuning of LLM Agents

URL: [View paper](#)

Brief Assessment

Agent Data Protocol[3] is the same paper as the original submission. The candidate text is identical to the original paper's abstract, confirming this is not prior work that could refute novelty claims.

5. VisReason: A Large-Scale Dataset for Visual Chain-of-Thought Reasoning

URL: [View paper](#)

Brief Assessment

VisReason[30] focuses on visual chain-of-thought reasoning for multimodal models, not on unifying diverse agent training datasets for supervised fine-tuning across coding, browsing, and tool-use benchmarks.

6. Training Language Models to Critique With Multi-agent Feedback

URL: [View paper](#)

Brief Assessment

Multi-agent Critique[31] focuses on improving critique ability through multi-agent feedback for supervised fine-tuning and reinforcement learning, not on creating unified agent training datasets across diverse tasks like coding, browsing, and tool use.

7. LLMs Can't Handle Peer Pressure: Crumbling under Multi-Agent Social Interactions

URL: [View paper](#)

Brief Assessment

Peer Pressure[32] focuses on multi-agent social interactions and conformity bias in quiz-style collaboration, not on unified agent training datasets or supervised fine-tuning across diverse benchmarks.

8. MedResearcher-R1: Expert-Level Medical Deep Researcher via A Knowledge-Informed Trajectory Synthesis Framework

URL: [View paper](#)

Brief Assessment

MedResearcher[33] focuses on medical domain-specific agent training with knowledge graph-based synthesis and specialized medical retrieval tools, not on creating unified multi-domain agent training datasets or standardization protocols.

9. Collaborative reasoner: Self-improving social agents with synthetic conversations

URL: [View paper](#)

Brief Assessment

Collaborative Reasoner[29] focuses on multi-turn conversational collaboration between agents for reasoning tasks, not on unified agent training datasets or supervised fine-tuning across diverse benchmarks like the original paper.

10. UrbanKGent: A Unified Large Language Model Agent Framework for Urban Knowledge Graph Construction

URL: [View paper](#)

Brief Assessment

UrbanKGent[35] focuses on urban knowledge graph construction using instruction fine-tuning for domain-specific tasks (relational triplet extraction, knowledge graph completion), not on creating unified agent training datasets across diverse domains like coding, browsing, and tool use.

Appendix: Text Similarity Detection

Textual similarity detection checked 26 papers and found 3 similarity segment(s) across 2 paper(s).

The following **2 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. The agent ohana: Designing unified data and training pipeline for effective agent learning

Detected in: Core Task (sibling), Contribution: contribution_1

△ **Note:** This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

2. Agent Data Protocol: Unifying Datasets for Diverse, Effective Fine-tuning of LLM Agents

Detected in: Core Task (sibling), Contribution: contribution_2, Contribution: contribution_3

△ **Note:** This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] Agent Data Protocol [View paper](#)
- [1] Recent advances in multi-source data fusion for traffic flow prediction: A review [View paper](#)
- [2] The agent ohana: Designing unified data and training pipeline for effective agent learning [View paper](#)
- [3] Agent Data Protocol: Unifying Datasets for Diverse, Effective Fine-tuning of LLM Agents [View paper](#)
- [4] Scaling proprioceptive-visual learning with heterogeneous pre-trained transformers [View paper](#)
- [5] A Framework for Fine-Tuning LLMs using Heterogeneous Feedback [View paper](#)
- [6] ER-RAG: Enhance RAG with ER-Based Unified Modeling of Heterogeneous Data Sources [View paper](#)
- [7] Normalizing flow policies for multi-agent systems [View paper](#)
- [8] Prediction guided decision-making for autonomous driving [View paper](#)
- [9] AlchemistCoder: Harmonizing and Eliciting Code Capability by Hindsight Tuning on Multi-source Data [View paper](#)
- [10] Normalizing Flow Policies for Multi-agent [View paper](#)
- [11] Agentmerge: Enhancing generalization in fine-tuned llm agents [View paper](#)
- [12] Motion tracks: A unified representation for human-robot transfer in few-shot imitation learning [View paper](#)
- [13] Aguis: Unified Pure Vision Agents for Autonomous GUI Interaction [View paper](#)
- [14] xlam: A family of large action models to empower ai agent systems [View paper](#)
- [15] You only learn one representation: Unified network for multiple tasks [View paper](#)
- [16] AgentOhana: Design Unified Data and Training Pipeline for Effective Agent Learning [View paper](#)
- [17] Transferable Tactile Transformers for Representation Learning Across Diverse Sensors and Tasks [View paper](#)
- [18] A unified framework for unsupervised Reinforcement Learning algorithms [View paper](#)
- [19] UniTR: A Unified Framework for Joint Representation Learning of Trajectories and Road Networks [View paper](#)
- [20] Agent lumos: Unified and modular training for open-source language agents [View paper](#)
- [21] Towards learning a generalist model for embodied navigation [View paper](#)
- [22] Unified world models: Coupling video and action diffusion for pretraining on large robotic datasets [View paper](#)
- [23] Unified-io: A unified model for vision, language, and multi-modal tasks [View paper](#)
- [24] Unifying 3d vision-language understanding via promptable queries [View paper](#)
- [25] Scene transformer: A unified architecture for predicting multiple agent trajectories [View paper](#)
- [26] Decoding Time Series with LLMs: A Multi-Agent Framework for Cross-Domain Annotation [View paper](#)
- [27] Graphonomy: Universal human parsing via graph transfer learning [View paper](#)
- [28] Agent-flan: Designing data and methods of effective agent tuning for large language models [View paper](#)
- [29] Collaborative reasoner: Self-improving social agents with synthetic conversations [View paper](#)
- [30] VisReason: A Large-Scale Dataset for Visual Chain-of-Thought Reasoning [View paper](#)
- [31] Training Language Models to Critique With Multi-agent Feedback [View paper](#)
- [32] LLMs Can't Handle Peer Pressure: Crumbling under Multi-Agent Social Interactions [View paper](#)
- [33] MedResearcher-R1: Expert-Level Medical Deep Researcher via A Knowledge-Informed Trajectory Synthesis Framework [View paper](#)
- [34] Agentbank: Towards generalized llm agents via fine-tuning on 50000+ interaction trajectories [View paper](#)
- [35] UrbanKGent: A Unified Large Language Model Agent Framework for Urban Knowledge Graph Construction [View paper](#)