Help Page

This BRAINCELL-AID platform[1] provides a comprehensive resource for mouse brain cell cluster annotations, integrating marker gene sets and large language model (LLM) for literature-grounded interpretations. The platform enables users to explore 5,322 brain cell clusters, examine multiple marker gene definitions and access Gene Ontology and PubMed–linked evidence. This platform supports transparent annotation workflows and community-driven refinement of cluster-level biological interpretations.

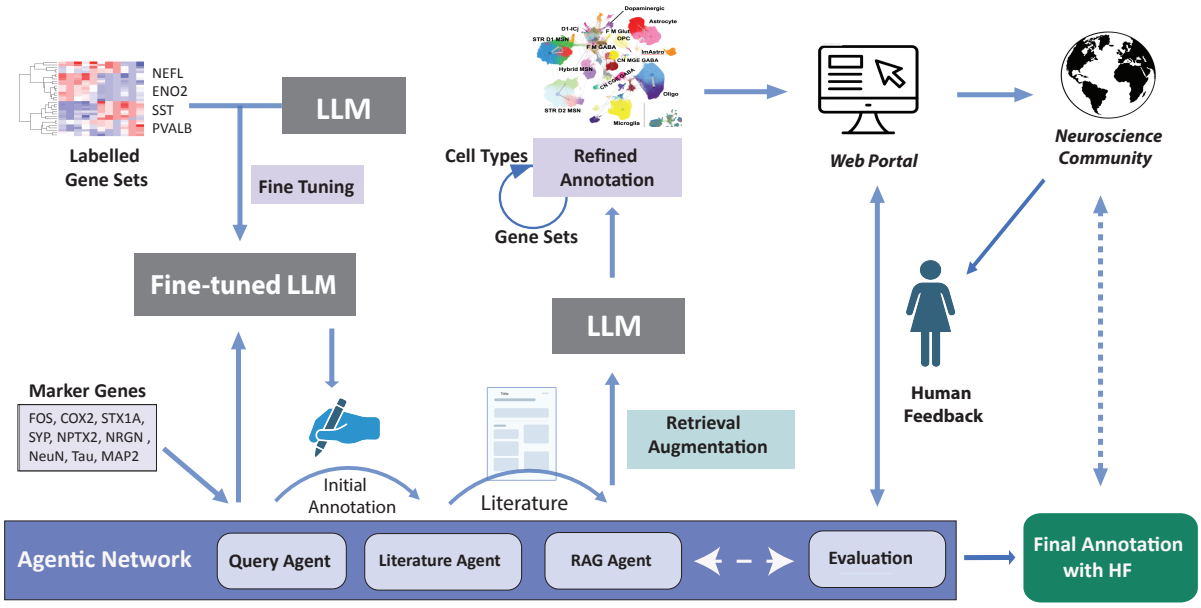

Figure 1. Agentic network workflow and goal overview for comprehensive brain cell type annotation.

The agentic workflow of BRAINCELL-AID, detailing the components and workflow that an input marker gene set will undergo to obtain the final annotation.

Home Page

Brain Cell Marker Gene Set List

The **List** covers four maker gene sets for each of the 5,322 cell clusters from the adult mouse brain cell atlas with their hierarchical information. The four marker gene sets are: cluster_combo, merfish, TF, and Top20 (top 20 most differentially expressed genes). Each row includes a **view** button that links to the detailed annotation page for that marker gene set.

The table includes the following fields:

Cluster ID: Unique identifier for each brain cell cluster.

Marker Type: Type of marker gene set.

Marker Genes: Specific marker genes for each cluster.

Super Type: Supertype label of the cluster.

Class Label: Class label of the cluster.

NT Type label: Most dominant neurotransmitter type of the cluster.

LLMs Annotation: Comprehensive annotations for each gene set.

Search Function

The platform features a robust search tool for quickly locating specific clusters or marker genes. The main search box covers all fields in the **Brain Cell Marker Gene Set List** Table.

The Advanced Search provides:

- Six boxes corresponding to the fields in the **Brain Cell Marker Gene Set List** Table (Cluster ID, Marker Type, Marker Genes, Super Type, Class Label, NT Type Label).
- Three additional boxes:
    - **Annotation**: Search for all annotations from cluster information pages.
    - **GO**: Search for specific Gene Ontology (GO) terms.
    - **PubMed**: Search for related PubMed literature.

Brain Cell Cluster Information Page

Each marker gene set information page provides detailed annotations and linked references via the **View** button in the **Brain Cell Marker Gene Set List** Table. Users can return to the main page anytime by clicking the **Home** button on the top banner.

Each Brain Cluster Information Page contains three main sections:

1. **Mouse Brain Cell Cluster Information** - An overview card of the mouse brain cell cluster that also links to **Allen Brain Knowledge Platform**.
2. **Cell Cluster Annotation Summary** - A comprehensive cell type summary generated by LLM and agentic AI for a cluster summarized from the cluster's four marker gene sets if applicable. For each cluster ID, the summary remains the same across all four marker gene sets. Detailed Summary and rationale are provided as well.
3. **Marker Gene Set** - The detailed annotations for a marker gene set.
    - Type: Marker type
    - GPTON Annotation: Including the initial GPTON annotation[2] and top 5 mapped GO terms.
    - Annotation refined by RAG (Retrieval-Augmented Generation):

- **TopPM**: Annotation refined using the most relevant PubMed references **to the initial annotation**, including refined annotations, relevant PubMed references and top five mapped GO terms.
- **TopGene**: Annotation refined using the most relevant PubMed references **for each individual gene**, including refined annotations, relevant PubMed references and top five mapped GO terms.

Users can propose improved or alternative annotations by selecting **Annotate this Cluster** under the **Brain Cell Cluster Information** section. Submissions are timestamped and stored as parallel records, which remain accessible to the community.

Related Publications

For methodological details and full descriptions of the data and annotation framework, users are encouraged to review and cite the following publications.

1. Li, R., Chen, W., Li, Z., Muñoz-Castañeda, R., Li, J., Maurya, N.S., Solanki, A., He, H., Xing, H., Ramlakhan, M., et al. (2025). A Brain Cell Type Resource Created by Large Language Models and a Multi-Agent AI System for Collaborative Community Annotation. ArXiv abs/2510.17064.
2. Li, R., Chen, W., Li, J., Xing, H., Xu, H., Li, Z., and Zheng, W.J. (2024). GPTON: Generative Pre-trained Transformers enhanced with Ontology Narration for accurate annotation of biological data. arXiv preprint arXiv:2410.10899.