Figure 1 - Cell Trajectory Recovery, Imputation, and Perturbation to Uncover Temporal and Spatial Dynamics using Multi-Agent Reinforcement Learning (CellTRIP)
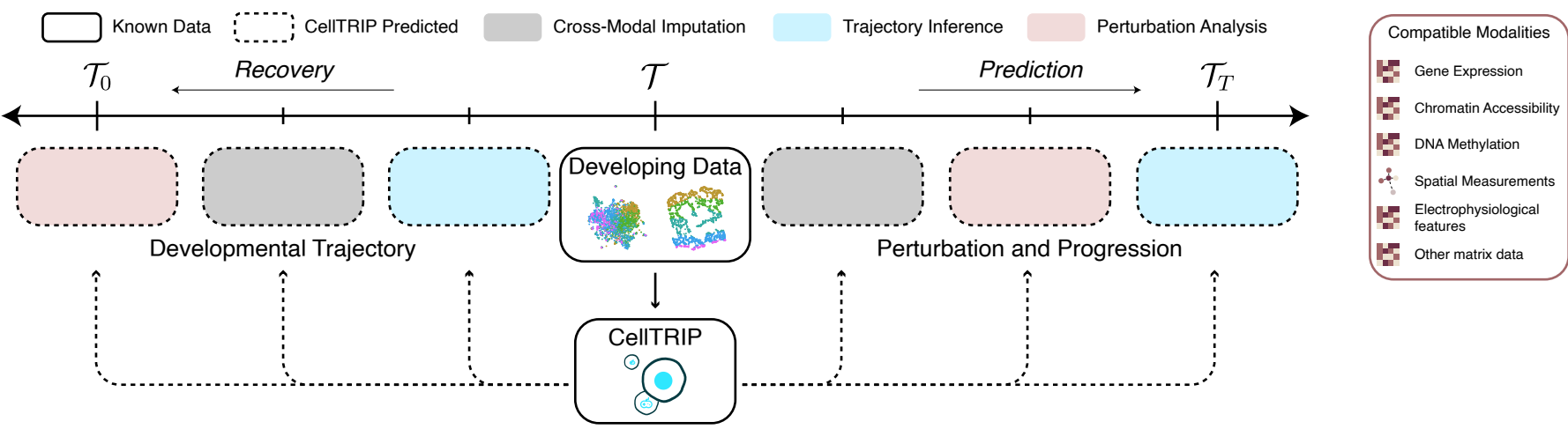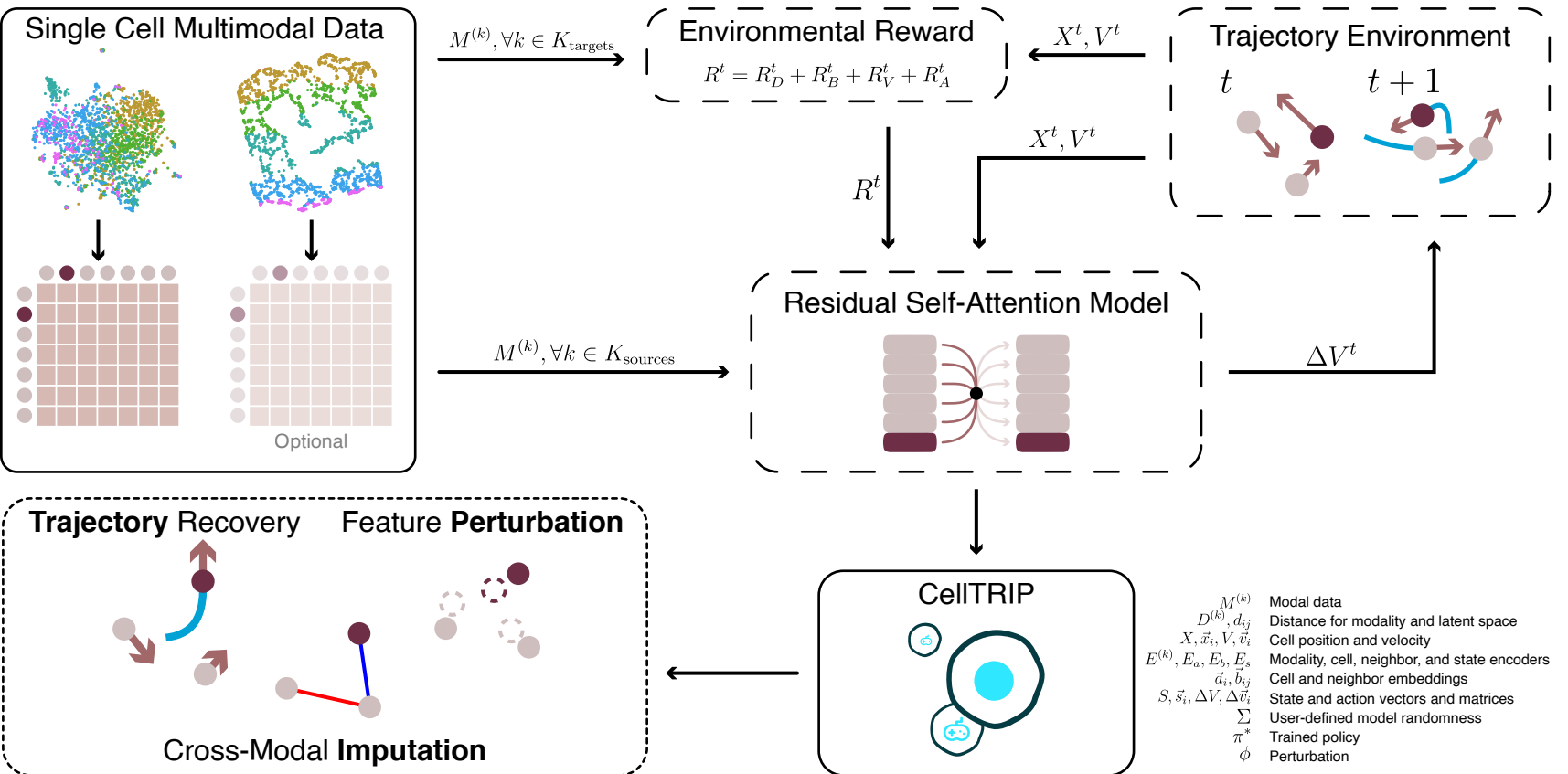
Known Data   CellTRIP Predicted   Cross-Modal Imputation   Trajectory Inference   Perturbation Analysis

Compatible Modalities
- Gene Expression
- Chromatin Accessibility
- DNA Methylation
- Spatial Measurements
- Electrophysiological features
- Other matrix data

$\mathcal{T}_0$ *Recovery* $\mathcal{T}$ *Prediction* $\mathcal{T}_T$

Developing Data

Developmental Trajectory   Perturbation and Progression

CellTRIP

Figure 2 - CellTRIP Model Training

Single Cell Multimodal Data

$M^{(k)}, \forall k \in K_{\text{targets}}$

Environmental Reward
$R^t = R_D^t + R_B^t + R_V^t + R_A^t$

X^t, V^t

Trajectory Environment
t $t+1$

X^t, V^t

R^t

Residual Self-Attention Model

$M^{(k)}, \forall k \in K_{\text{sources}}$

ΔV^t

Optional

Trajectory Recovery   Feature Perturbation

Cross-Modal Imputation

CellTRIP

$M^{(k)}$	Modal data
$D^{(k)}, d_{ij}$	Distance for modality and latent space
$X, \vec{x}_i, V, \vec{v}_i$	Cell position and velocity
$E^{(k)}, E_a, E_b, E_s$	Modality, cell, neighbor, and state encoders
$\vec{a}_i, \vec{b}_{ij}$	Cell and neighbor embeddings
$S, \vec{s}_i, \Delta V, \Delta \vec{v}_i$	State and action vectors and matrices
Σ	User-defined model randomness
π^*	Trained policy
ϕ	Perturbation

Supp. Figure 1 - CellTRIP Model Training Details

Inter-Cell Distances

Optional

$M^{(1)}, M^{(2)}, \ldots, M^{(k)}$

Modal features (e.g. genes, spatial coordinates)

Cells

i $D_{ij}^{(1)}$ j $D_{ij}^{(k)}$

Environmental Reward

$\vec{v}_i$
$\vec{x}_i$
D_{ij}
$\vec{x}_j$ $\vec{v}_j$

Boundary Penalty
$$[R_B^t]_i = \begin{cases} -1 & \max(\vec{x}_i) = 1 \\ 0 & \text{else} \end{cases},$$

Velocity Penalty
$$[R_V^t]_i = -||\vec{v}_i^t||^2,$$

Action Penalty
$$[R_A^t]_i = -||\Delta\vec{v}_i^t||^2,$$

Distance Reward
$$[R_D^{t+1}]_i = \delta_i^{t+1} - \delta_i^t, \quad \delta_i^t = \sum_{\text{targets}} \left\| \vec{d}_i^{(k)} - \vec{d}_i^t \right\|^2,$$

Total Reward
$$R^t = R_D^t + R_B^t + R_V^t + R_A^t.$$

$M^{(k)}$	Modal data
$D^{(k)}, d_{ij}$	Distance for modality and latent space
$X, \vec{x}_i, V, \vec{v}_i$	Cell position and velocity
$E^{(k)}, E_a, E_b, E_s$	Modality, cell, neighbor, and state encoders
$\vec{a}_i, \vec{b}_{ij}$	Cell and neighbor embeddings
$S, \vec{s}_i, \Delta V, \Delta \vec{v}_i$	State and action vectors and matrices
Σ	User-defined model randomness
π^*	Trained policy
ϕ	Perturbation

Trajectory Environment

Random Initialization

$t = t_0$ $t = t_1$ $t = t_2$

$\vec{v}_i$
$\vec{x}_i$
D_{ij}
$\vec{x}_j$ $\vec{v}_j$

$\vec{v}_i + \Delta\vec{v}_i$

Repeat

Residual Self-Attention Model

$$\vec{a}_i = E_a\left(\left[\vec{x}_i, \vec{v}_i, E^{(1)}(\vec{m}_i^{(1)}), \ldots, E^{(k)}(\vec{m}_i^{(k)})\right]\right)$$

$$\vec{b}_{ij} = E_b\left(\left[\vec{x}_j, \vec{v}_j, E^{(1)}(\vec{m}_j^{(1)}), \ldots, E^{(k)}(\vec{m}_j^{(k)}), a_i\right]\right)$$

Residual Self-Attention

$\Delta\vec{v}_i$

$\Delta\vec{v}_i \sim \mathcal{N}(E_s(\vec{s}_i), \Sigma)$

$\vec{s}_i$

Mean