Task 1 — Open datasets → PhysiCell "nanobots with pheromones"

A. Curate open datasets (what they give you & why)

- **BraTS (2018–2021/2020 link)**: multi-modal MRI (T1, T1Gd, T2, FLAIR) with expert tumor labels (enhancing tumor/edema/necrosis). Use to extract 3-D tumor geometry & initialize cell densities. Perelman School of Medicine

- **TCGA-GBM / TCGA-LGG on TCIA**: MRI matched to genomics; use for additional patient geometries + genotype/phenotype calibration. The Cancer Imaging Archive (TCIA)+1

- **IVY-GAP (images + RNA-seq)**: glioblastoma anatomic regions and transcriptomics—good for spatial heterogeneity (hypoxia, stem-like niches → different cell phenotypes). The Cancer Imaging Archive (TCIA)glioblastoma.alleninstitute.orgNCBI

- **UCSF-PDGM**: standardized 3T pre-op brain-tumor MRI cohort (diffuse gliomas); consistent acquisition for robust geometry building. The Cancer Imaging Archive (TCIA)

- **QIN-BRAIN-DSC-MRI**: perfusion (DSC) MRI with ROIs—use to set vascularity/perfusion fields. The Cancer Imaging Archive (TCIA)

- **REMBRANDT**: clinical + omics (for phenotype priors / growth & treatment response heuristics). The Cancer Imaging Archive (TCIA)Nature

B. Build the simulation pipeline

1. **Voxelize geometry:** Convert BraTS labels to a voxel grid; map tumor core/edema to initial **cell populations** and **ECM/porosity** fields.

2. **Microenvironment:** Use **BioFVM** in PhysiCell for oxygen, drug, pheromones (2 scalar fields: *trail* and *alarm*). Diffusion/uptake rates can be initialized from EVONANO's ranges (e.g., D ~ 10^{-7} cm²/s for small NP) and then tuned. PLOSNature

3. **Agents:**

    - **Tumor cells** with hypoxia-dependent phenotypes (PhysiCell built-ins).

    - **Nanobots** as motile agents with chemotaxis toward (i) drug-need gradients (low O_2 / high tumor density) and (ii) **pheromone trails** from peers; bots **secrete trail**, **emit alarm** near

obstacles or non-response zones.

- Optional **PhysiBoSS 2.0** to plug intracellular Boolean logic for drug response inside each cell agent. PMC

4. **LLM hierarchy:**
   - **Queen (planner):** ingests low-res global state (downsampled tumor heatmap + aggregate toxicity & kill metrics) every N steps; outputs **policy knobs** (e.g., bot swarm split ratios, secretion rates, exploration bias, drift toward vasculature).
   - **Workers (bots):** follow simple rules (chemotaxis + pheromones) with small stochasticity; Queen updates their parameters, not each action—keeps sim fast.
5. **Control/baselines:** Compare to (a) no-pheromone bots, (b) fixed-policy bots, (c) EVONANO-style optimized nanoparticles/drug schedules. Nature
6. **High-throughput search:** Wrap with **EMEWS** to sweep policy knobs (and LLM prompt variables) efficiently; PhysiCell+EMEWS demonstrated large parameter sweeps for tumor-immune systems. BioMed Central
7. **Performance:** Start with 2-D slices (fast iterate) → 3-D. Consider **OpenACC/GPU** BioFVM acceleration later for big runs. arXiv

Why PhysiCell? It's a 3-D off-lattice ABM coupled to BioFVM; exactly the substrate/agent combo we need for "nanobots + pheromones + drug diffusion." PLOS

C. "Most efficient" way to simulate "LLM-powered nanobots"

- **Re-use EVONANO ideas**: it already couples PhysiCell with ML to optimize nanoparticle design & dosing—excellent scaffold for our bots/drug behavior and for setting physically plausible ranges. Nature
- **Keep the LLM at the *episodic* level** (planner every K steps), let simple heuristics run at each timestep; this preserves PhysiCell speed.
- **Use stochastic multi-seed batches** (PhysiCell is stochastic; >5 seeds per setting recommended) and only save sparse checkpoints to avoid I/O bottlenecks. BioMed Central
- **Prototype visually in PhysiCell Studio** to validate rules quickly before HPC runs. PMC

Task 2 — Sepolia testnet as a trust + continual-learning surface

Principle: training stays off-chain; **provenance, reputation, and pointers** go on-chain.

Minimal on-chain design (Sepolia, chainId 11155111):

- **ExperienceRegistry.sol**
  - `submitExperience(bytes32 runHash, string ipfsCid, bytes32 dataHash, uint256 score)` → emits `ExperienceSubmitted(runHash, ipfsCid, dataHash, score, msg.sender)`
  - Store only hashes + CIDs (BraTS-derived data is de-identified, but still keep payloads off-chain). ChainList
- **Reputation/Staking.sol**
  - Contributors stake; evaluators (or automated validators) can **attest** quality via **EAS**; low-quality or unverifiable results can be **slashed**. Ethereum Attestation ServiceEasscan
- **Schemas (EAS):**
  - `SimulationResult(hash runHash, string model, string dataset, uint32 seeds, float meanKill, float tox, string cid)`
  - `Evaluation(hash runHash, float reproducibilityScore, string notes)`
    Attestations written on Sepolia via EAS SDK; the contract trusts only results with adequate attestations/stake. QuickNode

Off-chain loop (what your agents actually do):

1. Run N PhysiCell episodes with a proposed policy (from Queen LLM).
2. Compress logs + metrics; pin to **IPFS** (get CID).
3. `submitExperience` on Sepolia with the CID + hashes; request EAS attestations from designated validators or from an auto-validator bot that re-runs a *subset* (reproducibility check).

4. Queen LLM **queries the contract** (via The Graph or direct RPC) to fetch top-attested runs relevant to the current patient geometry; updates policy priors.

This gives you a *shared, tamper-evident memory* of "what worked under which conditions," and a way for swarms to **trust** external knowledge while improving. (You can add meta-tx relayers so bots don't manage keys/gas.) OpenZeppelin Docs

---

Alternatives you might not have considered

- **PhysiBoSS 2.0** for intracellular signaling: couple Boolean networks for drug response per cell—richer biology without crushing performance. PMC
- **Federated results**: keep different labs' simulations off-chain, **only attest** summaries on Sepolia (EAS) to avoid data movement. Ethereum Attestation Service
- **Patient-specific meshes**: for cases with longitudinal MRIs, build per-patient geometries and compare policies across time (QIN-DSC perfusion helps initialize vasculature). The Cancer Imaging Archive (TCIA)

TL;DR (practical summary)

- Use **BraTS/TCIA** cohorts to build 3-D brain-tumor geometries.
- Implement **nanobots + two pheromone fields** in **PhysiCell**; let a **Queen LLM** adjust high-level swarm parameters while workers follow chemotaxis/pheromone rules.
- Reuse **EMEWS/EVONANO** patterns for efficient parameter sweeps and ML-guided optimization.
- Log every episode to **IPFS** and **attest** results on **Sepolia** (EAS + staking) to create a shared, trustworthy memory your agents can learn from.