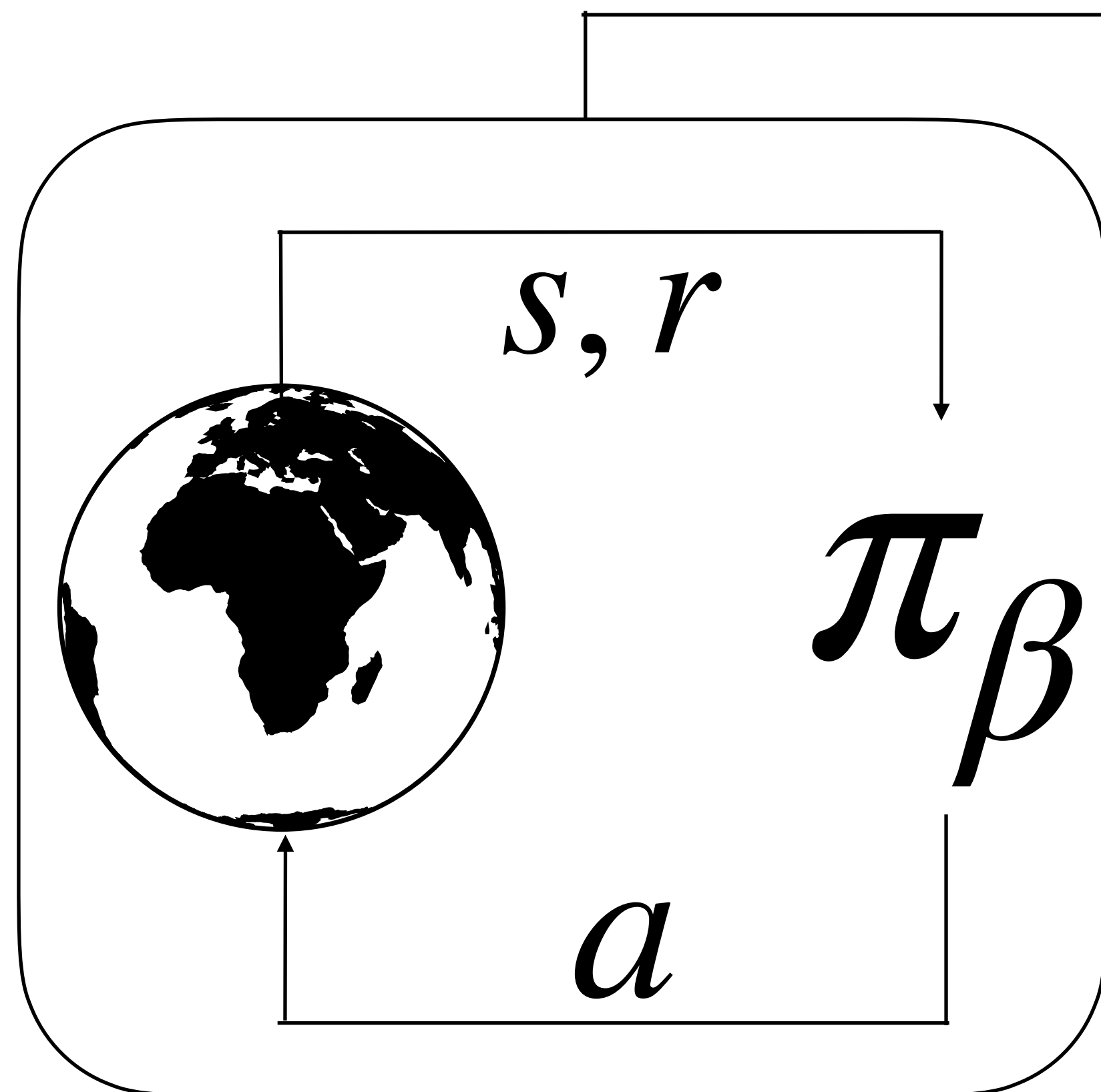
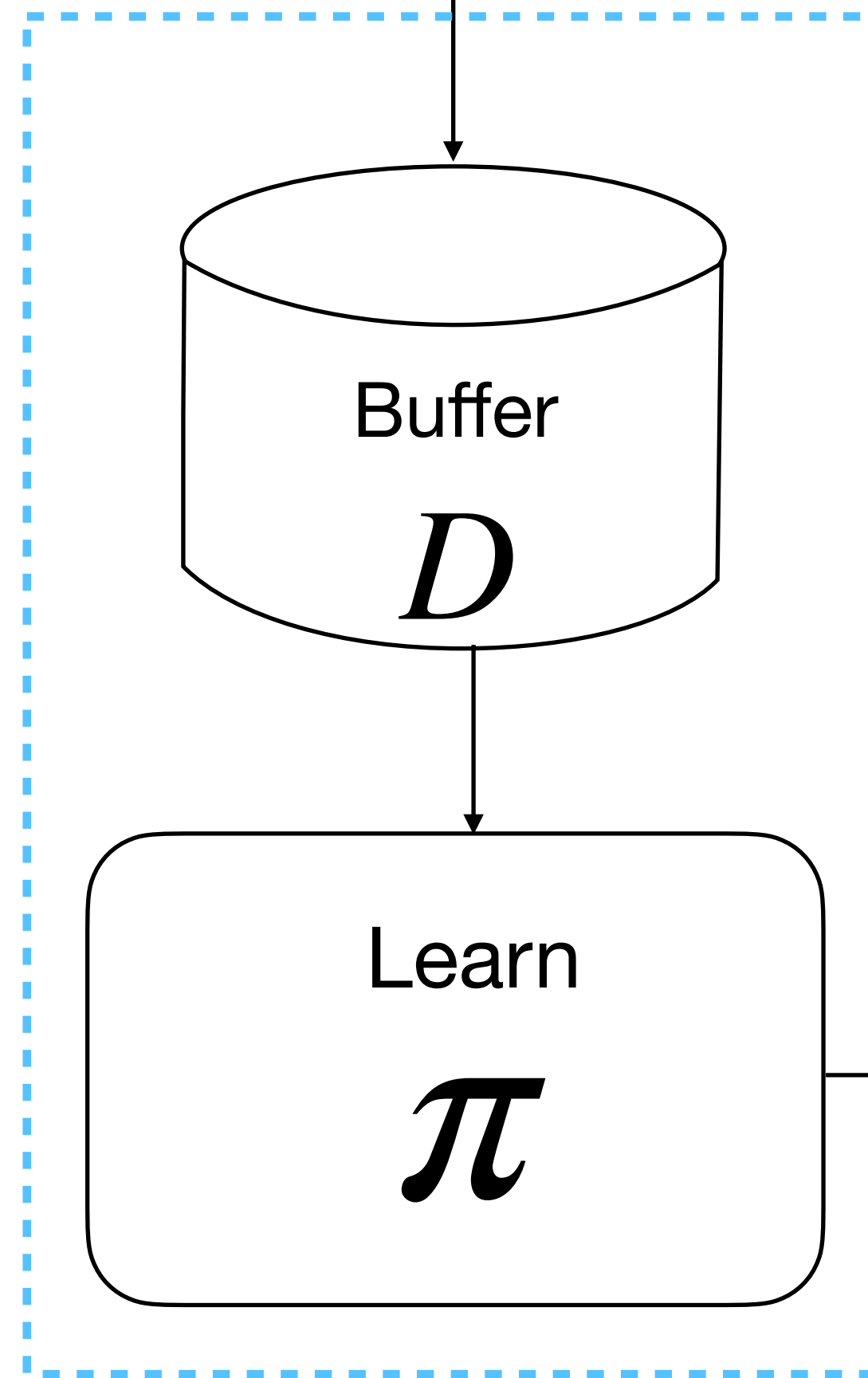


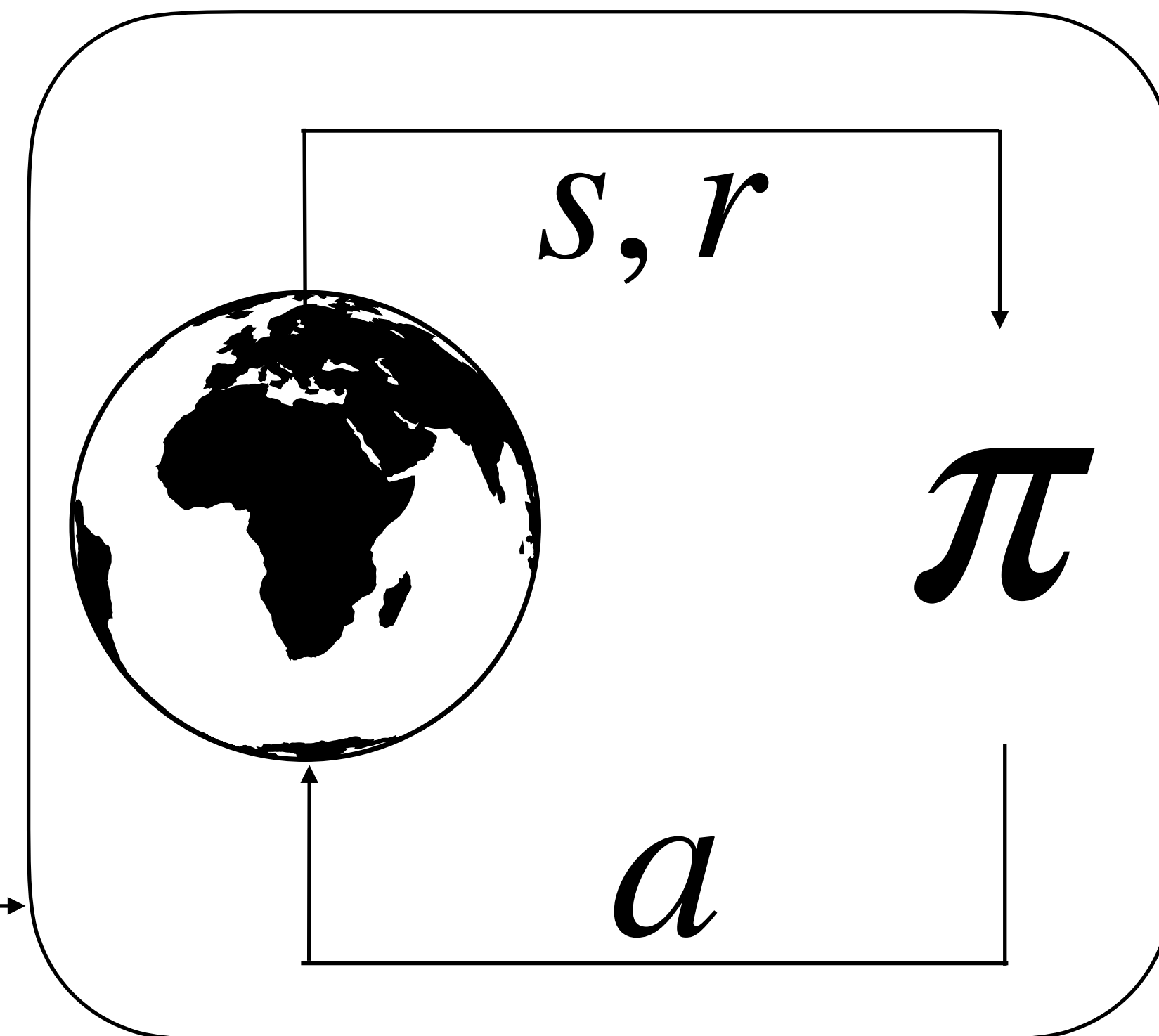
Rollout data $\{(s_i, a_i, s'_i, r_i)\}$



Data is collected **once** using
arbitrary policies



Training Phase



Deployment