

Extreme-Scale Search...when Billions aren't Enough

Chiharu Konda

OpenEye, Cadence Molecular Sciences

OpenEye's Extreme Scale Search (ESS), deployed on the Orion cloud molecular design platform, is engineered to search ultra-large chemical spaces, ranging from billions to trillions of molecules, for hit identification in drug discovery. ESS leverages reaction informatics to construct 2D and 3D synthon libraries of synthetically accessible molecules derived from over 7 million building blocks, enabling the creation of a virtual space for trillions of compounds. By operating in an unenumerated form, ESS avoids the prohibitive costs of full enumeration. Instead, ESS applies novel query strategies and shape-based scoring of ROCS and FastROCS to efficiently identify high-quality hits.

ESS integrates multi-armed bandit models, Bayesian active learning, and reaction-driven sampling to prioritize and explore diverse, chemically tractable space. It supports both ligand-based and pocket-derived queries (via SZMAP), enabling the discovery of structurally novel hits that are synthetically feasible.

In collaboration with a pharmaceutical partner on their drug discovery projects, ESS has already yielded approximately 150 new synthesizable compounds not found in vendor catalogs, demonstrating its real-application impact. For medicinal chemists, ESS provides a powerful and fast approach to find novel synthesizable molecules. By uniting cheminformatics, cloud-scale computation, and AI-guided search, ESS redefines early-stage drug discovery by enabling efficient exploration of trillion-scale chemical space.