A key ingredient for preparing the next generation of K-12 technologies is to harness discovery curiosity in our cohorts - a skill of stimulating diverse exploration and mastery-oriented learning. Students, even the most advanced, often struggle to understand complex foundational ideas and concepts in a discipline, which acts as a barrier to asking productive questions, which often gets in the way of deeper learning. Addressing this issue of just-in-time contextual, conceptual support at scale is a significant educational challenge.

Aim. Using life sciences as an application domain, we will use big data and Artificial Intelligence to create self-guided learning strategies for developing the art of asking productive questions and seeking new hypotheses. Through preliminary work, we have already assembled a community of scientists, educators, programmers, and K-12 members (200 students and 17 teachers) called the BioSenseNetwork that represents 10 diverse districts in Arizona (60% Title-1, and three Native American Reservations). For AFOSR STEM, we will train a Large Language Model (LLM) tutor to gather novel queries from our growing community of learners. The content-rich, context-and-goal-driven autonomous tutor agent will serve as a "guide-on-the-side" for learners to provide primary and alternative solutions and possibilities to explore. This will bridge the gap between general and personalized education, promoting knowledge equity across learners of different ages, genders, and socio-economic backgrounds at no cost as part of our BioSenseNetwork.

Method. Using our technology integration framework TPACK, we prepare online laboratories of interactive molecular visualization that overcome the presentational limitations of textbooks to train participants in foundational life sciences (structure, dynamics, sensing, genetics, photosynthesis, and vaccines), with content aligned to standards. The proposed project will capitalize on our human-autonomy models to simulate a tutor embedded within BioSenseNetwork activities that provide contextual, dyadic interactions with learners to scaffold learning and discovery curiosity.

1) Using mouse-tracking data accrued over thousands of human-computer interactions during an interactive lab and applying nonlinear dynamical systems methods, we will determine the predictability of operations across our network of members. **2)** The quality of these community-wide operations will be classified with Blooms taxonomy, weighing the student, teacher, and advisor responses distinctly. Thus, new queries defining the context of an exercise will be positively reinforced. **3)** A digital tutor will be programmed by augmenting an LLM (e.g. ChatGPT) with our content and exercises supplementing the foundational concepts in life science textbooks. This LLM tutor will guide learners, helping them contextualize visualization tasks addressing unresolved questions and creating new questions for exploration.

Expected outcome. The cyclic process of exchanging questions between an individual and a community through our programmed "digital twin" of a tutor will improve both personal and collective learning. An LLM with knowledge of life sciences (and its contexts) will supplement teachers that will be demonstrated for the Tuba City, Casa Grande and Chandler School District of AZ, and over time, the models will grow in knowledge exchange, connecting students worldwide and inspiring the creation of similar LLMs in other fields.