

Title:

BacterAI: Making AI accessible to scientists across disciplines

Authors:

Yesenia V. Juarez-Vargas^{1,2}, Joonhoon Kim², Andrew Frank², Bram Stone²

¹Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

²Biological Science Division, Pacific Northwest National Laboratory (PNNL), Richland, WA

Abstract:

BacterAI is an automated science platform developed by the Jensen lab at the University of Michigan to map microbial metabolisms and accelerate discovery in microbiology. At PNNL, BacterAI's purpose is to address the highly combinatorial challenge by predicting microbial growth across diverse nutrient and stressor conditions. By applying reinforcement learning, BacterAI learns directly from experimental data to simulate and predict microbial responses in growth, allowing scientists to explore the effect of carbon sources or stress substrates at different concentrations. The platform integrates reinforcement learning, genome-informed metabolic modeling, and transfer learning to extend predictions across different environments.

The BacterAI workflow consists of different components such as experiment request, data generation, feature extraction, predictive model, and simulation. My main goal was to make BacterAI accessible to scientists in different disciplines. This goal was divided into three sections starting with simulation notebooks to assess the main classes and functions that make microbial growth predictions and experimental simulations for two different machine learning methods: neural networks and Gaussian process regression. We created a dashboard that would simplify file management with input and output files of the platform and would allow scientists to visualize their round. Finally, a genome informed model was implemented, with the overall goal to create a function that would model OD600 from experimental data, this would help leverage genomic information to establish a model that can simulate or predict microbial growth with different nutrient sources. The product of these tasks made advanced AI phenotyping capabilities accessible at PNNL.