

Project Proposal: A “Focused Research Organization” to Build the Foodome Project for the Future of Nutrition

Our current knowledge of the biochemical compounds in food is incredibly limited, but existing databases of MassSpec scans contain massive amounts of untapped, unannotated information about food ingredients. A project to leverage these databases with the tools of data mining, AI, and high-throughput measurement will systematically unveil the chemical composition of all food ingredients and revolutionize our understanding of food and health.

[Learn more about FROs, and see our full library of FRO project proposals here.](#)

Problem Statement

Diet is the single biggest determinant of health over which we have direct control. An unhealthy diet poses more risk to morbidity than alcohol, tobacco, drug use, and unsafe sex combined. Indeed, our diet exposes us to thousands of food molecules, many of which are known to play an important role in multiple diseases including coronary heart disease, cancer, stroke, and diabetes. Despite the demonstrated and complex role of diet on health, nutrition science remains focused on molecules that serve as energy sources such as sugars, fats, and vitamins, leaving most disease-causing compounds uncatalogued and invisible to researchers and health care professionals. Further, our current understanding of the way food affects health is limited to nutritional guidelines that rely on a panel of 150 essential micro- and macro-nutrients in our diet. This is a tiny fraction of the more than 130,000 compounds known to be present in food, hence limiting our ability to unveil the health implications of our diet.

Project Concept

The Foodome project aims to unveil this “dark matter of nutrition” by creating an open-access high-resolution compendium of food compounds through a strategy that combines Big Data, ML/AI, and experimental techniques, implemented by a focused cross-disciplinary team, motivated to bring transformative change and maximize public benefits.

In the past five years, [BarabásiLab](#) has curated the largest library of compounds in food, consisting of more than 135,000 biochemicals linked to 3,500 foods. While the number of biochemicals is exceptional, the coverage is highly uneven, sparse, and largely unquantified. Yet, information about the missing biochemicals is carried by the unannotated MassSpec peaks available for each MassSpec scan of food ingredients. Because chemicals are invisible to the one-chemical-one-peak tools employed today, we have designed a strategy that relies on data mining, AI, and high-throughput measurements to resolve them: we plan to collect the more than 3,000,000 MassSpec scans already available in databases, and mine the full scientific literature to collect knowledge on food composition. We also plan to take advantage of the

increasing number of annotated genomes to infer their chemical makeup. These data will serve as input for a ML/AI platform designed to learn associations between biochemical structures and the ingredients' phylogenetic position, helping us systematically unveil the chemical composition of all food ingredients.

What is a Focused Research Organization?

Focused Research Organizations (FROs) are time-limited mission-focused research teams organized like a startup to tackle a specific mid-scale science or technology challenge. FRO projects seek to produce transformative new tools, technologies, processes, or datasets that serve as public goods, creating new capabilities for the research community with the goal of accelerating scientific and technological progress more broadly. Crucially, FRO projects are those that often fall between the cracks left by existing research funding sources due to conflicting incentives, processes, mission, or culture. There are likely a large range of project concepts for which agencies could leverage FRO-style entities to achieve their mission and advance scientific progress.

This project is suited for a FRO-style approach because the Foodome platform and knowledge base will address problems in health science beyond the competence of any single academic group or start-up. The project started in the academic environment involving groups at Northeastern University, Harvard Medical School, and Tufts Medical School, but typical academic researchers and institutions are motivated by short term publication strategies and unable to devote the years needed to develop a public resource. Federal nutrition research funding exists, but is fragmented, and normal funding channels are generally unable to offer sustained support for a project of this size. With VC funding, we were able to move the project to a startup environment to standardize the toolset and develop key technologies, but company management decided that the Foodome platform's timeline is too far from the market. Based on these experiences, an FRO appears to be the best framework to accomplish the vision of Foodome. The project enters a field limited by technological stagnation, and will fundamentally change our understanding of health and disease, impacting multiple fields and industries.

How This Project Will Benefit Scientific Progress

A high-resolution knowledgebase on the composition of food will revolutionize our ability to explore the role of each food-borne molecule in human health, impacting multiple fields: 1) It will be transformative for health care, changing our ability to prevent and control disease. 2) It will aid the development of healthier, more nutritious, and biochemically balanced foods. 3) It will facilitate the development of novel pharmaceuticals. 4) By improving MassSpec annotations, it will provide a more accurate biochemical descriptions of any sample, empowering diagnosis, and detection. 5) It will unlock innovations in personalized nutrition and precision medicine, allowing clinicians to offer precision advice to a patient on how to use diet to prevent and manage disease.

Key Contacts

Author Albert-László Barabási, Northeastern and Harvard Medical School, alb@neu.edu

Referrers Matt Hourihan, Federation of American Scientists, mhourihan@fas.org
Alice Wu, Federation of American Scientists, awu@fas.org