MAKE-IT

Advancing National Security through Fundamental Research

MAKE - IT

THE NEED AND OPPORTUNITY

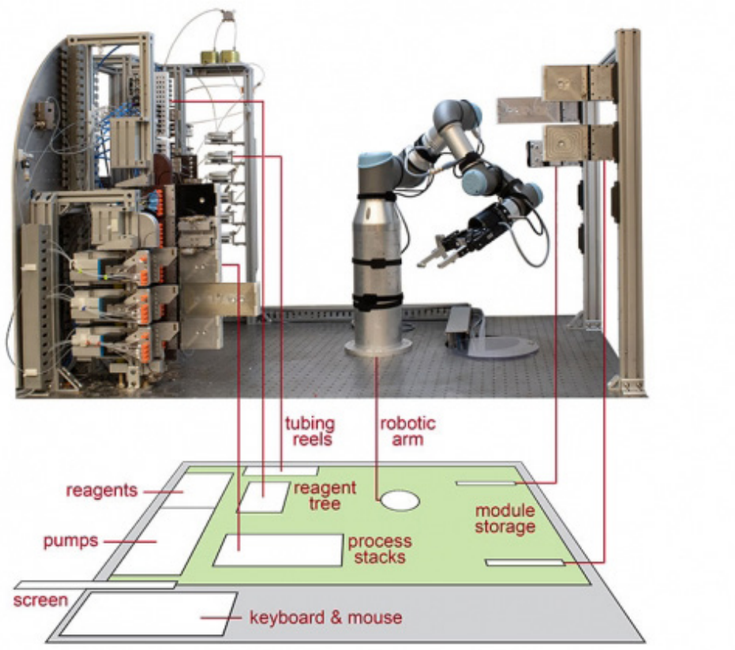
Chemical innovation is central to developing cutting-edge technologies for the military. Research chemists design and make new molecules that could enable a slew of next-generation military products. These include novel propellants for spacecraft engines; new pharmaceuticals and medicines for troops in the field; lighter and longer-lasting batteries and fuel cells; advanced adhesives, coatings, and paints; and less expensive explosives for the military that are safer for troops to handle.

The challenge today, however, is that designing and producing new molecules with desired properties is slow and requires the expert intuition and experience of PhD-level researchers, who can spend weeks to months developing and optimizing a synthetic route (the "recipe") to a single molecule. Chemists then test that molecule to determine if it has the desired properties, and, more often than not, need to go through more iterations before finding success. This cycle continues as chemists iterate designs and synthetic protocols to develop more efficient production methods and to deliver the new, more capable materials that can further the cause of national security. Adding to this innovation challenge is that the chemical design space is so vast that it's impossible for humans to fully explore all the possibilities. The number of potential molecules is on the order of 10^{60}, and there might be many different possible synthetic routes to making each one.

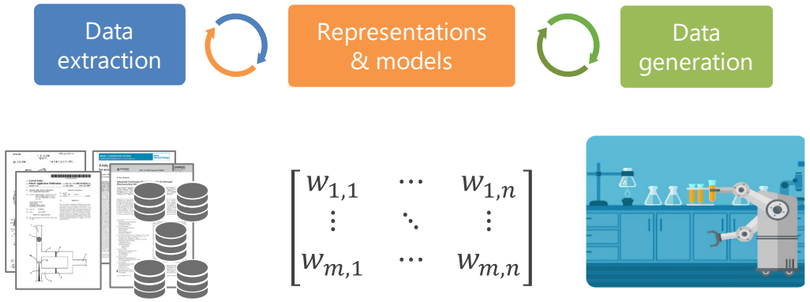
THE DARPA SOLUTION

The DARPA Make-It program was initiated in 2015 to address these challenges, with the ultimate goal of automating chemical synthesis – both in terms of front-end recipe optimization and the back-end hardware to carry out synthetic protocols. Toward that end, Make-It performers have been adapting fundamental advances in artificial intelligence (AI) in fields centered on image and textual data for chemical production. They have also established cutting-edge capabilities to deliver new computational tools that enable exploration – even without advanced training in chemistry – of an immense space of possible molecules much more rapidly and efficiently than has been possible before.

In one approach that has been moving the technology forward, researchers at the Massachusetts Institute of Technology (MIT) have been developing machine-learning techniques that can exploit existing chemical reactivity data for designing synthetic routes. During their involvement in the program from 2015 to 2019, the MIT team addressed a gauntlet of critical challenges in adapting and implementing machine-learning methods to handle chemical reactivity data.



MIT's automated synthesis system combines an AI-guided synthesis planner that helps chemists hone molecule-making procedures and a robotic platform that configures hardware to automatically carry out the prescribed reactions.



The Accelerated Molecular Discovery program aims to build on the successes of Make-It to open new AI-driven pathways of discovery within the vast space of possible molecular structures.

The MIT team quickly learned, for one, that existing reactivity data was insufficient and unreliable. To compensate for some of these shortfalls, the team developed means to generate negative data computationally to help synthesis-planning algorithms identify dead-end synthesis routes to rule out. They were then able to integrate the "negative data" into a new reaction-screening capability that Make-It performers at Purdue University had been developing.

Researchers in the program also had to contend with the unsurprising reality that data in the chemical sciences is presented in formats for human consumption, not for algorithmic consumption. To address this, the MIT team developed digital approaches to extracting information from the text and figures in published chemical literature. For this task, the researchers adapted advances in fields such as natural language processing and computer vision for the purpose of making chemistry data readily available to their machine-learning algorithms for automating planning and execution of chemical syntheses. At the crux of this advance were new digital representations that capture reaction-relevant features of textual information in technical chemistry papers.

THE IMPACT

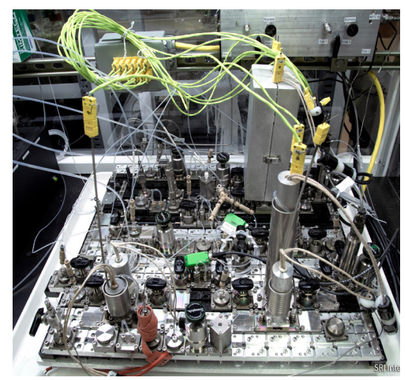
The MIT performers made rapid progress in understanding how to apply machine learning to the goal of developing an unprecedented capability for designing new synthetic routes for useful chemicals. Over the course of four years, they layered their new methods into a software tool that now incorporates features such as prediction of reaction pathways, retrosynthetic analysis, and reaction-condition suggestion. The predictive power of the tool is now on par with the performance of doctorate-level chemists, a goal previously regarded as unobtainable without decades of research.

Researchers and organizations in academic, government, and industry sectors have been embracing an open-source software tool based on guiding principles and common datasets that emerged in the Make-It program. Now, those in molecular sciences anywhere can utilize the software developed by the MIT performers to boost their own efficiency and productivity. IBM, for one, created and offers its IBM RXN tool, which is a free, web-based utility for predicting chemical reactions.

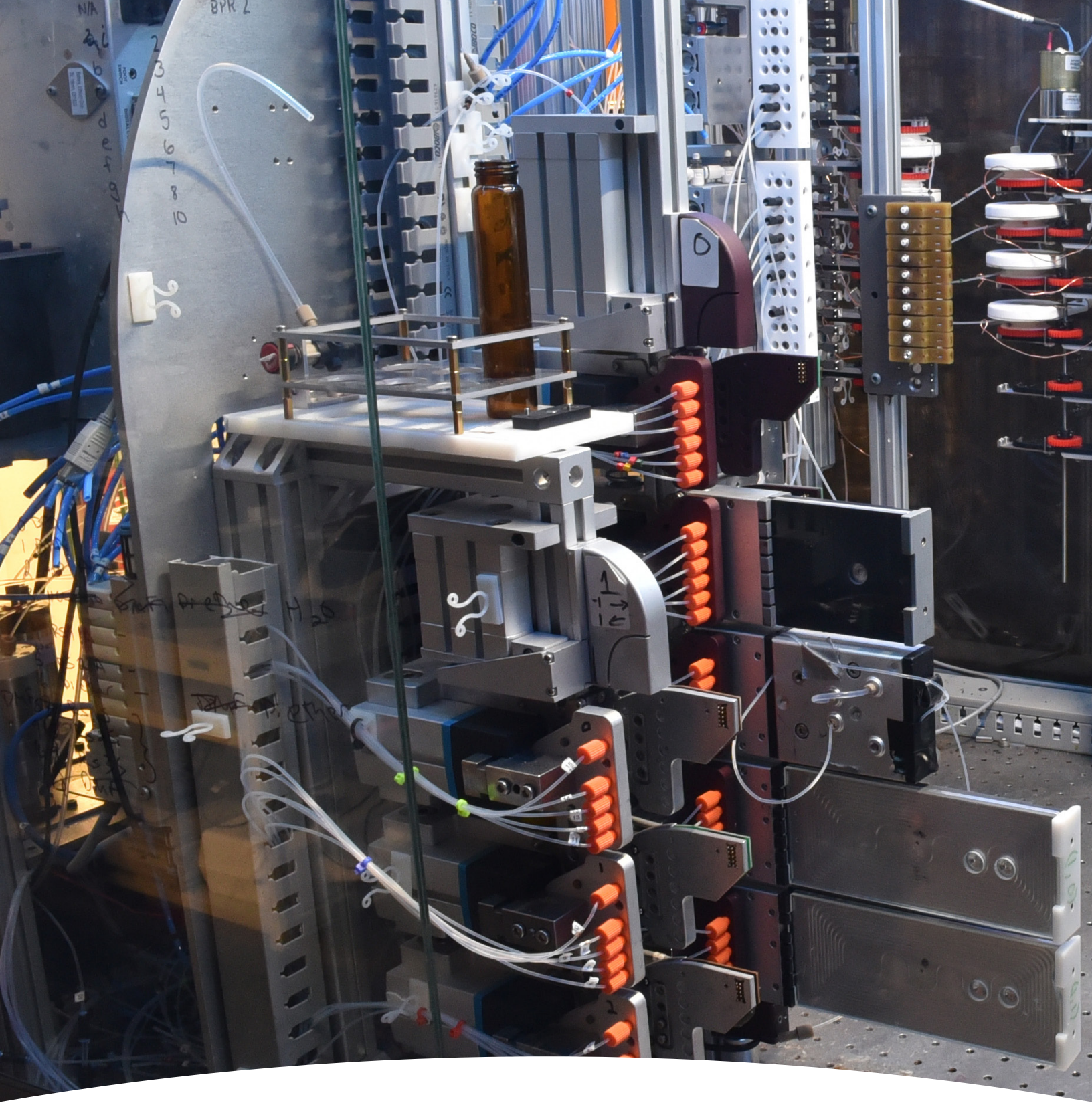
Government organizations including the Army Research Laboratory, Lawrence Livermore National Laboratory, and the National Institutes of Health's National Center for Advancing Translational Sciences also are applying and evaluating the MIT software for their specific needs and interests – energetic materials development, forensic analysis, and human health, respectively. Meanwhile, in follow-on work, MIT has developed a consortium with 13 pharmaceutical companies, which are summoning their own datasets in an effort to develop new machine-learning tools for accelerating development of new medicines, from cancer therapeutics to antibiotics to countermeasures to chemical, biological, and radiological threats.

LOOKING AHEAD

MIT's work in the Make-It program was explicitly focused on chemical reactivity and small molecule production. DARPA, in its Accelerated Molecular Discovery program, is now building on its chemistry-centric, machine-learning investments in chemical synthesis to speed the rate at which researchers can identify new useful chemical structures. This work will build yet a stronger foundation for applying AI to chemical discovery and production in ways that can bolster national defense while opening new opportunities throughout the chemical enterprise.



With funding from DARPA's Make-It program, SRI International developed this AutoSyn module of its SynFini platform, which can produce gram-scale quantities of chemicals with automated and AI-informed synthesis procedures.



DARPA

www.darpa.mil