



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/243170/>

Version: Accepted Version

Article:

Datta, B., Buehler, M.J., Chow, Y. et al. (2026) Artificial intelligence for food innovation. Nature Food. ISSN: 2662-1355

<https://doi.org/10.1038/s43016-026-01380-7>

This is an author produced version of an article published in Nature Food, made available via the University of Leeds Research Outputs Policy under the terms of the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Artificial Intelligence for Food Innovation

Bianca Datta, Markus J. Buehler, Yvonne Chow, Kristina Gligorić, Dan Jurafsky ,
David L. Kaplan, Rodrigo Ledesma-Amaro, Giorgia Del Missier, Lisa Neidhardt,
Karim Pichara, Benjamin Sanchez-Lengeling, Miek Schlangen, Skyler R. St. Pierre,
Ilias Tagkopoulos, Anna Thomas, Nik Watson, Ellen Kuhl

corresponding author: ekuhl@stanford.edu;

Abstract

Global food systems must deliver nutritious, sustainable foods while sharply reducing environmental impact. Yet, food innovation remains slow, empirical, and fragmented. Artificial intelligence (AI) offers a transformative path to link molecular composition to functional performance, connect chemical structure to sensory outcomes, and accelerate cross-disciplinary innovation across the production pipeline. While broadly applicable to food systems, we focus on sustainable proteins—plant-based, fermentation-derived, and cultivated—as a high-impact testbed for AI-driven closed-loop design. We review the applications, opportunities, and challenges of AI for Food as an emerging discipline that integrates ingredient design, formulation development, fermentation and production, texture analysis, sensory science, manufacturing, and recipe generation. We identify four priorities: advancing scientific machine learning with embedded domain priors, treating food as a programmable biomaterial, building self-driving laboratories for automated discovery, and developing deep reasoning models that integrate nutrition and sustainability. Integrating AI responsibly into the food innovation cycle can accelerate the transition to sustainable food systems and establish a predictive, design-driven science of food for human and planetary health.

Editor’s summary

Artificial intelligence is transforming how new foods are designed, produced, and experienced. While AI has the potential to link molecular composition to functional and sensory performance and accelerate food innovation, challenges related to data quality, transparency, and governance remain. Addressing these requires interdisciplinary collaboration, open and inclusive data ecosystems, and AI systems that augment rather than replace human expertise.

Food is among the most powerful determinants of human health, planetary sustainability, and cultural identity [1]; yet, research and development for food remains slow, fragmented, and driven by trial-and-error rather than rational design. Radical, urgent, and creative food solutions are needed to combat the current climate crisis [2]. Notably, across farmland, livestock, and production, the food sector generates 35% of all global greenhouse gas emissions [3], and contributes heavily to biodiversity loss and public health challenges [4]. To meet climate goals and feed a growing population, we need alternative sources of protein that are delicious, affordable, scalable, and nutritionally compelling [5]. Today,

however, the scientific bottlenecks in texture, flavor, scalability, and cost pose barriers to adoption [6]. Artificial intelligence (AI) has reshaped domains, from protein design and drug discovery to materials engineering [7]. AI is now poised to revitalize how we grow and make food—from optimizing plant-based meats to generating climate-conscious recipes [8]. Enabling protein diversification will have profound implications, and AI can substantially accelerate scientific progress if applied responsibly and towards the correct challenges [9]. But the success of AI is by no means guaranteed. When deployed in a scatter-shot way, AI can generate low-quality outputs,

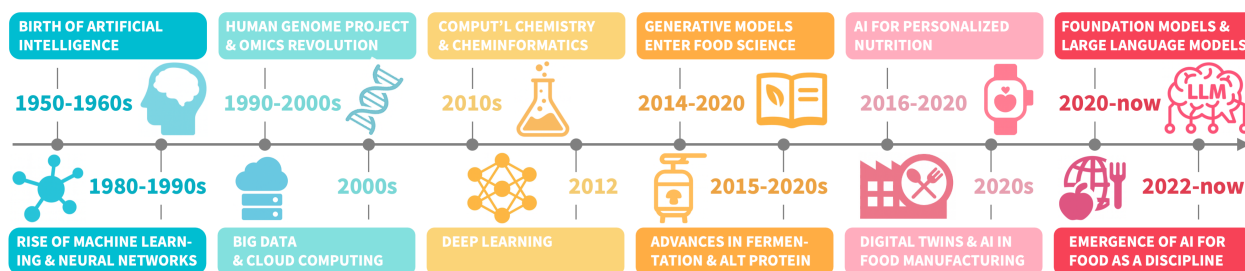


Fig. 1 Historical milestones in AI for Food. The birth of AI in the 1950s–1960s, marked by Alan Turing’s concepts of computing machinery and intelligence, established the foundations of machine intelligence. The rise of machine learning and neural networks in the 1980s–1990s, driven by backpropagation and statistical learning, shifted AI from rule-based systems to data-driven approaches. In parallel, the human genome project and the omics revolution in the 1990s–2000s, enabled by high-throughput sequencing, created a culture of large-scale biological and nutritional data. The expansion of big data and cloud computing in the 2000s further empowered the analysis of massive ingredient, supply chain, and consumer preference datasets. Since the beginning of the 21st century, advances in computational chemistry and cheminformatics in the 2010s have enabled the prediction of flavor molecules, aroma profiles, and chemical interactions. Breakthroughs in deep learning, exemplified by AlexNet in 2012, catalyzed applications in food recognition and image-based dietary tracking. Since 2014, generative models have expanded into food design, flavor and recipe generation, and ingredient discovery, with companies such as NotCo applying AI to plant-based food development. Since 2015, advanced fermentation and alternative protein models have started to include AI-driven strain engineering and process optimization to accelerate precision fermentation and industrial adoption. The current decade is witnessing the integration of wearable and microbiome data for AI-driven personalized nutrition since 2016, the emergence of digital twins for real-time monitoring and optimization of food manufacturing processes, and the rise of foundation models and large language models that leverage unstructured data for formulation and recipe innovation since the 2020s. Together, these advances are fueling the emergence of AI for Food as a distinct scientific discipline.

increase environmental burdens, or undermine consumer trust, aspects that are especially sensitive in food [10]. The goal is thus not to deploy AI indiscriminately. Instead, we need to partner with AI where it measurably reduces time, cost, and uncertainty for problems that matter [11]. Responsible application—focused on the right challenges—is essential. In this review, we identify areas across the food system where we can constructively engage with AI: we provide background on the current state of the art, outline new opportunities, and discuss the challenges that AI is facing in the quest to accelerate the future of food.

The intellectual and technological trajectory that has shaped to the emergence of *AI for Food* as a discipline spans more than seven decades, weaving together advances in AI, food science, and biotechnology (Fig. 1). From the birth of symbolic AI in the 1950s [12], and the rise of machine learning and neural networks in the 1980s–90s [13], to the deep learning breakthroughs in the 2010s [14], and the advent of generative models [15] and foundation models [16] in the 2020s, AI has progressively expanded its ability to model, design, and predict complex systems. In parallel, developments in genomics [17], computational chemistry [18], fermentation, and personalized nutrition [19] have created unprecedented data resources and biological insights directly relevant to food design. Their convergence has given rise to a new field in which AI is now central to ingredient discovery, recipe formulation, manufacturing optimization, and consumer personalization [11].

Designing food materials poses unique challenges, as their properties emerge from multiscale interactions among proteins, carbohydrates, lipids, and processing conditions [20]. Yet, AI is shifting food innovation from trial-and-error to data-driven discovery. Modern models link chemical composition to sensory outcomes and enable targeted formulation and rapid iteration. Machine-learning frameworks predict consumer preferences from chemical-sensory panels [21], while graph neural networks map molecular structure to odor quality [22]. Advances in protein modeling and molecular learning accelerate functional ingredient discovery [23, 24], and precision fermentation provides scalable routes to novel proteins and textures [25]. Together, these advances address long-standing challenges—engineering egg-free emulsions, meltable dairy alternatives, or fibrous plant-based meats—and have translated into market-ready products from companies such as NotCo, Perfect Day, and Climax Foods [26]. This transition from conceptual promise to industrial deployment underscores AI’s growing role as a driver of culinary innovation and sustainability.

We focus on the post-harvest stages and follow an AI-powered cycle through ingredient design, formulation development, fermentation and production, texture analysis, sensory properties, manufacturing, and recipe generation (Fig. 2). In the following sections, we highlight how AI-powered approaches, from generative models to digital twins, drive innovation across these steps.

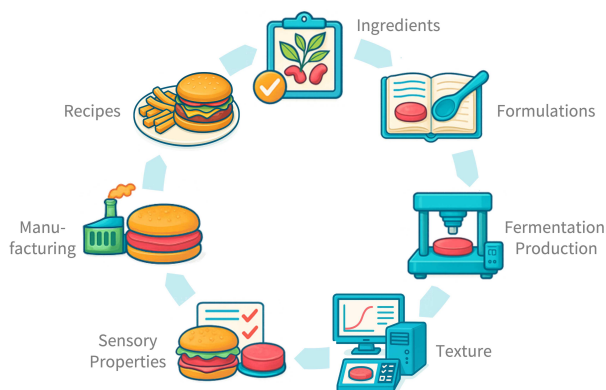


Fig. 2 The AI-powered production cycle. Seven key stages where artificial intelligence can accelerate the time to market: ingredient design, formulation development, fermentation and production, texture analysis, sensory properties, manufacturing, and recipe generation. The example of a plant-based burger illustrates the tight integration of food science, biotechnology, materials science, behavior science, and artificial intelligence to guide the development of plant-based burgers from raw ingredients to commercial products.

Results

Ingredients: Driving functionality-first design.

Novel food ingredients must balance sustainability, nutrition, sensory appeal, and functionality. Functionality includes water-holding, emulsification, gelation, viscosity, and stability that shape processing and structure. Many ingredients aim to replicate animal-derived functions, such as foaming in egg replacers; yet, matching performance remains challenging. Structure–function relationships link molecular architecture to activity. In food, macromolecular features of proteins, lipids, and fibers encode functionality. This raises the question to which extent AI can bridge this scale gap and translate molecular features into techno-functional predictions for a functionality-first design–build–test–learn cycle.

The success of animal-free ingredients depends on functionality and the ability to meet formulation and consumer requirements [28]. These traits arise from molecular features, environmental conditions such as pH and temperature, and interactions with fats, fibers, and solutes. Proteins, lipids, and fibers provide core functionality, with proteins most versatile. The most promising alternative proteins today are wheat gluten and soy; they can form fibrous textures under extrusion and shear [28]. New sources include legumes, oilseeds, grains, algae, and leaves [29]. Processing strongly shapes functionality: dry fractionation preserves native structures, while wet fractionation produces isolates that control solubility and gelation [30]. Lipids and fibers

add complementary features, from stabilizing interfaces to modulating viscosity and lubrication. Since plant heterogeneity limits predictability, developers modify ingredients chemically, physically, or enzymatically. The field is now shifting toward functionality-enriched fractions for more consistent, cost-effective performance [29].

AI and informatics enable functionality-driven ingredient design by linking molecular and mesoscopic features to macroscopic outcomes under processing constraints. Classification models support discovery; for example, a QSAR model identified glycyrrhizin as a natural emulsifier [31]. Regression models predict properties such as pectin viscosity to reduce cost and accelerate formulation [32]. Machine learning can predict techno-functional traits across crops with $R^2 > 0.80$ [33], while AI models can capture plant protein rheology with $R^2 > 0.99$ [34]. Deep learning increasingly links functionality to sensory design [35]. A recent study optimized starch-based gels for saltiness and strength using e-tongue and oral-processing data [36]. Multi-objective optimization can predict solubility, gel strength, and emulsification towards scalable formulation strategies [33]. Together, these advances support closed-loop design–build–test–learn cycles. High-throughput screening, active learning, and digital twins operationalize models, while generative approaches target multiple criteria and bridge lab-to-plate gaps.

Ingredient design still struggles with consistent performance across food matrices and key challenges remain. First, models often fail to generalize across ingredients and conditions. Second, nonlinear interactions in multicomponent systems complicate prediction and require hybrid physics-informed approaches. Third, sensory preferences vary across cultures, which demands diverse datasets. Fourth, models must link molecular features such as sequence, charge, and hydrophobicity to functionality in a transparent and scalable way. Addressing these challenges requires open, collaborative tools and an AI-driven ecosystem that predicts techno-functional performance from molecular features.

Formulations: Mapping flavor and recipe spaces.

Ingredient functionality sets the foundation, but effective food design depends on how those ingredients interact in complex systems; formulation brings these components together. Meeting the dual mandate of sustainability and sensory fidelity requires moving beyond slow trial-and-error formulation development. AI offers a compelling approach to uncover hidden patterns that connect ingredients, formulations, production, texture, and sensory properties, and propose candidate formulations that are predicted to meet nutritional, functional,

or sensorial targets [37–39]. Today, AI practitioners can draw on a broad collection of public resources including Recipe1M+ [40], FlavorGraph [41], FlavorDB [42], USDA FoodData Central, and FAO/INFOODS, as well as domain-specific databases such as Phenol-Explorer, LipidBank, and the Volatile Compounds in Food database. Inputs include molecular descriptors, nutritional and functional properties, recipe structures, and process sequences; outputs are complete candidate formulations aligned with target criteria. AI-powered approaches combine latent-space learning through autoencoders with conditional generation, constraint-aware optimization, and expert-in-the-loop refinement (Fig. 3A). But can they truly capture the complex, non-linear interactions between ingredients, processes, and perception—even under sparse and noisy data—and unlock a new generation of food formulations?

Computational gastronomy has long shown that recipes and flavor chemistry form structured networks that we can mine for predictive patterns. The Flavor Network formalized ingredient compatibility via shared volatiles [43], while FlavorGraph integrated one million recipes with $\sim 1,500$ flavor molecules to generate graph embeddings for pairing prediction and ingredient clustering [41]. These studies have successfully shown that high-dimensional, multimodal food representations can capture the complexity of culinary design. In alternative proteins, industrial disclosures from NotCo patents [37–39] illustrate hybrid generative optimization pipelines. These systems embed ingredients and recipes in latent spaces using autoencoders and recurrent networks, condition sampling on feature vectors of target product profiles, refine candidates with constraint solvers such as sparse regression, and iteratively incorporate expert sensory feedback. This architecture mirrors multimodal generative trends in other fields and enables analogue discovery. It can identify non-obvious ingredient substitutions, while maintaining process feasibility. Other companies, including Climax Foods and Eat Just, deploy similar proprietary approaches, while academic research continues to build on Recipe1M+ [40], FlavorGraph [41], and FlavorDB [42]. Together, these efforts highlight a state of the art that blends representation learning with constraint-aware generation to accelerate formulation discovery.

Generative AI expands opportunities across protein diversification. For example, it can design plant-based cheeses that melt and stretch by identifying proteins that act like casein, the key driver of structure and meltability in animal cheese. It can discover ingredient blends that replicate meat

textures and pinpoint volatile combinations that evoke familiar flavors [44]. And it can improve for plant-based meats from sensory panel feedback and match expert predictions of panel rankings [45]. Large-scale screening lets scientists focus on a few promising prototypes instead of hundreds of trials, which shortens the development cycle. In the near future, these tools will be able to tailor formulations to individual nutritional needs and taste preferences and enable personalized food design.

But significant challenges remain before AI can transform formulation discovery at scale. High-quality sensory datasets are limited and often fail to generalize across different food types. Even with detailed chemical and nutritional data, predicting human perception remains an open problem. Modeling ingredient performance under varying heat, shear, or storage conditions is incomplete. Cost, supply chain, allergen control, and clean-label requirements must be built into models as hard constraints. Trust among regulators, industry, and consumers will depend on transparent, auditable design processes. The near-term path will be hybrid, where AI proposes and humans refine, toward transparent formulation platforms that integrate constraints and predict performance at scale.

Fermentation: Enabling smart bioreactors. Molecular and formulation design define the building blocks, but realizing performance requires precise control at the cellular and microbial scale. Precision fermentation—where engineered microorganisms are programmed to produce functional proteins such as casein or heme—has become central to protein diversification and sustainable food production. Its success depends on tightly coupled variables, strain stability, metabolic fluxes, nutrient composition, and bioreactor dynamics. AI has the potential to fuse these heterogeneous data with mechanistic insight to enable predictive, scalable, and cost-effective fermentation and production to address critical bottlenecks such as yield optimization and resource efficiency (Fig. 3C).

Optimizing fermentation requires a deep understanding of bioprocesses to maximize yield and efficiency while maintaining homogeneity, sterility, and scalability [46]. Kinetic bioprocess models remain widely used but are constrained by steady-state assumptions, and metabolic flux models for engineered strains often face computational bottlenecks. Experimental optimization through orthogonal design [47] and response surface methodology [48] remains the industry standard, while multiscale feature engineering is still limited [49]. Machine learning has proven valuable for modeling nonlinear fermentation dynamics [50] and for real-time monitoring of anaerobic digestion

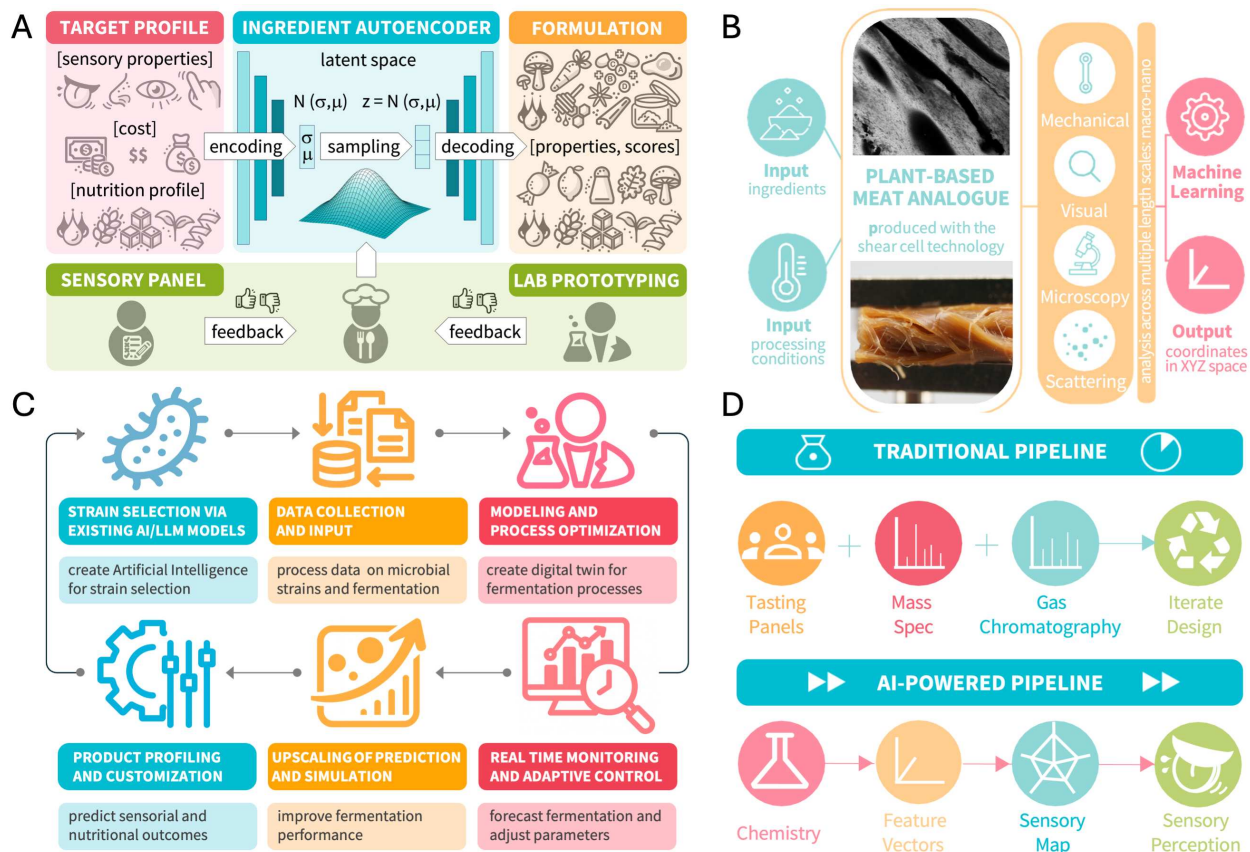


Fig. 3 AI for food. (A) AI for Formulations. An iterative workflow to discover new formulations encodes target profiles such as sensory properties, cost, and nutritional profiles; generates formulation sets with deep learning; predicts properties and scores; and integrates laboratory prototyping and sensory panels in continuous feedback loops. (B) AI for Texture. Machine learning integrates multimodal data, from ingredients to process conditions, to optimize plant-based meat analogues, from microstructural appearance to macrostructural texture. Future machine learning tools could help reverse engineer ingredients and process conditions to achieve a desired appearance and texture. (C) AI for Fermentation and Production. Artificial intelligence and large language models can accelerate the fermentation and production cycle during microbial strain selection, data collection, process optimization, real time monitoring and adaptive control, upscaling of prediction and simulation, and product profiling and customization. (D) AI for Sensory Properties. Optimizing the sensory perception of food is traditionally a costly, time-consuming process involving expert tasting panels and instrumental analysis, including mass spectrometry and gas chromatography. In contrast, an AI-powered pipeline has the potential to predict sensory perceptions from only the chemical or molecular make-up of food, eliminating the need for costly human panels and reducing the time to iterate product formulations.

[51]. For example, deep learning soft sensors that combine convolutional neural networks with bidirectional long short-term memory architectures have achieved R^2 values up to 0.98 in predicting key anaerobic digestion state variables from real-time pH, electrical conductivity, and oxidation–reduction potential measurements [51]. While AI and reinforcement learning have been successfully applied to sustainable protein fermentation [52] and co-culture optimization [53], large-scale adoption of AI, and especially generative AI, remains limited.

AI can shift fermentation and production from reactive optimization to predictive management, with customized strains and self-optimizing processes. Machine learning models can optimize strain design, nutrient feeds, and bioreactor control, and

deep learning approaches can capture non-linear interactions better than traditional regression [54]. Automation and AI-enhanced imaging [55], soft sensors [56], and spectroscopy [52] can improve monitoring and adaptive control. Algorithms such as support vector machines [57] and fuzzy inference systems [58] support robust predictions, while evolutionary optimization [59] aids process tuning. Hybrid mechanistic-AI models [60] and explainable AI frameworks [61] can improve transparency and decision-making. Training LLMs on unstructured bioprocess data—including text, images, and sensor logs—could support knowledge retrieval, troubleshooting, and experimental design and bridge biology, chemistry, engineering, and design through

actionable insights, like predicting protein structures or cell morphology [62].

To realize this potential, the field must overcome the key barriers of data scarcity, data integration, and model transferability. Generating large experimental datasets remains labor intensive, but scaled-down bioreactors and small-data algorithms now offer new possibilities. At the same time, robust biosecurity and ethical frameworks will be critical to safely employ generative AI. Integration with genome-scale metabolic models and digital twins of bioreactor performance could create predictive platforms for design and scale-up. The collective goal is an integrated fermentation platform that links strain design, bioreactor dynamics, and real-time sensing to support predictive scale-up and reliable process optimization.

Texture: Learning structure–perception maps.

Texture is a key quality attribute that shapes consumer acceptance. It emerges from coupled physical, mechanical, and sensory properties across scales—from molecular structure to oral processing—which complicates quantification and prediction. We distinguish between physical texture, measured through mechanical and rheological tests, and sensory texture, perceived through attributes such as chewiness or fibrousness. This distinction links measurable material behavior to human perception and defines how AI can connect mechanics with sensory response. Current approaches rely on costly sensory panels and labor-intensive tests such as cutting, tension, compression, shear, and rheology. The core challenge is to map intuitive descriptors like crispiness or chewiness onto measurable parameters that guide product design (Fig. 3B). This raises the question of how AI can integrate multimodal datasets to predict and engineer consumer-preferred textures and accelerate the adoption of sustainable proteins.

Every ingredient and product possesses a characteristic texture, which we can engineer through processing techniques such as extrusion, shear cell structuring, or 3D printing to mimic animal-based foods. Recent advances show the promise of machine learning for decoding texture–structure relationships: Constitutive neural networks can map mechanical signatures of plant-based meat alternatives onto interpretable and generalizable physics-based models [63] that achieve R^2 values from 0.92 to 0.99 for commercially available plant-based meats [64]. Autoencoders can predict sensory texture attributes from rheological data and advance data analysis beyond traditional correlation methods [35]. Classifiers such as support vector machines and neural networks can identify snack freshness from mechanical and acoustic signals with up to

92% accuracy [65]. Researchers are using machine learning to predict viscosity, gelation, and foaming of plant proteins [33, 66]. Meanwhile, texture analysis from images is rapidly growing. Transfer learning with models like VGGNet, ResNet, and DenseNet allows us to predict the cooking times for fish with potential extension to plant-based foods [67], while ML-driven microscopic image analysis can help us identify droplets in colloidal systems to link microstructure to texture [68]. Beyond academic applications, commercial platforms already harness AI for texture optimization: For example, NotCo [37–39] and Climax Foods successfully deploy AI to mine ingredient databases, combine plant ingredients, and design formulations that replicate animal textures.

AI accelerates prototyping by predicting textural outcomes before physical trials and reduces time- and resource-intensive trial-and-error cycles. In addition, ingredient design and process optimization can balance target textures with environmental impact, for example, to reduce greenhouse gas emissions. Another important application is personalization, for example designing softer foods for elderly populations to improve swallowing safety. LLMs can also convert consumer feedback such as too chewy or not crispy enough [45] into quantifiable parameters for product adjustment, and bridge the gap between perception and engineering [69]. To address data sparsity, LLMs can generate synthetic training data by pairing realistic texture descriptors with mechanical properties. Ideally, AI can integrate data across multiple length scales, from molecular interactions to macroscopic structure, to capture the hierarchical origins of texture. Inversely, future machine learning tools could help reverse engineer ingredients and process conditions to achieve a desired texture. More broadly, generative design tools, similar to those used for inorganic materials [7], hold the potential to propose ingredient–process combinations that meet desired textural constraints.

Significant obstacles remain before AI-driven texture engineering will truly become routine. The biggest challenge is the lack of large, open datasets that couple textural, sensory, and processing data. Industrial data sharing is rare and data fusion remains technically challenging, as mechanical, rheological, visual, and compositional data are inherently multimodal and often unstructured. The ultimate vision would be an AI system for inverse design—with a specific target texture as input and optimal ingredients and processes as output—that transforms texture development into a predictive, model-guided science.

Sensory: Learning perceptual representations.

Ultimately, food is not a material design challenge,

but a sensory experience. Aroma, flavor, and texture are the strongest determinants of consumer acceptance. The central challenge is to map measurable physical and chemical properties of ingredients to the multidimensional space of human perception. Machine learning provides a framework to learn this mapping and enables both the prediction of sensory profiles and the computational design of novel ingredients with desired characteristics (Fig. 3D). But how far are current AI models from accurately predicting human sensory perception and from generating novel flavor and aroma experiences?

Traditional sensory science has relied on human panels for perceptual evaluation and on analytical instruments such as gas chromatography-mass spectrometry for chemical composition. Panels remain the gold standard, but are slow and expensive, while instruments offer precision without perceptual insight. Early machine learning approaches combined biomimetic sensors, such as E-noses, with classifiers [70], but required heavy feature engineering and were limited to narrow classification tasks [71]. The field has since shifted toward deep representation learning [72], which learns general-purpose sensory maps directly from data. The Principal Odor Map exemplifies this trend: a graph neural network organizes molecules into a vector space where distance encodes perceptual similarity with human-level accuracy in odor labeling: On a prospective validation set of 323 novel odorants, the model’s predictions closely matched a trained panel mean with per-molecule correlations approaching $R \approx 0.6$ in representative cases [22]. Researchers now employ attention-based architectures to extend this mapping to molecular mixtures [73]. Algorithmic tools for predictive sensory modeling are maturing rapidly, but progress is constrained by the scarcity and cost of high-quality sensory data.

Learned sensory representations open powerful opportunities for food innovation. High-throughput virtual screening of molecular libraries can identify novel flavor compounds for plant-based meats and accelerate product development [11]. Integrating predictive models with low-cost, miniaturized sensors enables real-time quality monitoring in manufacturing and packaging [74], for example with wireless freshness indicators that reduce food waste [75]. These tools also provide systematic frameworks for biomimicry to predict how modifications of ingredients or processes alter the final sensory profile. Still, a feedback loop remains essential: computational predictions must be continuously validated and refined with real-world experiments to ensure consumer relevance.

Today, the bottleneck is no longer algorithmic capacity but high-quality data acquisition [76]. Deep learning models require large cross-scale datasets, yet sensory panels cannot generate data at scale to capture the vast chemical space of food [45]. The path forward lies in closing the loop between modeling and experiment. Generative models such as transformers and diffusion architectures are algorithmically ready, but their success depends on new data pipelines. Developing robust, low-cost hardware and embedding it into automated platforms will be critical to realize the full potential of AI for sensory science. The ultimate vision is a self-driving sensory lab that integrates chemical synthesis, sensory characterization, and machine learning to autonomously generate, predict, and validate flavor, aroma, and texture perception across diverse food systems.

Manufacturing: Bridging lab and industrial scale.

Even the best foods will fail if they cannot be produced consistently and affordably. Manufacturing convincing meat analogues is inherently complex because technical goals such as texture, flavor, appearance, and shelf life interact non-linearly with environmental and economic constraints. Relevant datasets span composition, structure, rheology, thermal behavior, processing steps such as extrusion or fermentation, as well as telemetry, sensory, and consumer data. Through digital twins, computer vision, and other modes, AI is emerging as a key technology to integrate these multimodal datasets and streamline process optimization. Typical model inputs include formulation and process profiles, while outputs can range from texture, color, and flavor to cost and environmental impact. But can these models deliver reliable, scalable, and cost-effective production paths without compromising sensory quality or sustainability goals?

AI and ML span across the manufacturing pipeline, from formulation and process design to in-line quality control. High-moisture extrusion produces fibrous structures that mimic muscle, but outcomes depend on moisture, temperature, and screw configuration. Bayesian optimization explores this space more efficiently than trial-and-error methods [78]. In a pilot study, it identified optimal parameters in ten trials versus fifteen for response-surface methodology and reduced validation errors from 61% to below 15% [79]. Digital twins support soft sensing and control, but most applications remain at pilot scale [81]. Data fragmentation persists across proprietary logs and repositories with inconsistent metadata. To address these limitations, FAIR data principles [82] and LLM-based

extraction [83] have successfully employed knowledge graphs to link formulation, process, structure, and sensory performance.

AI is opening multiple opportunities for advancing alternative protein manufacturing. Engineers can use machine learning and Bayesian optimization to tune high-moisture extrusion parameters to target fibrousness, yield, and energy efficiency, while digital twins evaluate what-if simulations for scale-up and real-time sensing [77, 78, 80]. Food scientists can use informatics pipelines to rank plant proteins and design blends at scale using rheology–functionality maps that allow ingredient discovery well beyond commodity sources [79, 82]. LLMs further act as formulation copilots, transform unstructured texts into structured knowledge graphs, and suggest experiments to resolve uncertainty [82]. At the same time, computer vision and hyperspectral analytics can advance in-line quality assurance by linking color and structure to sensory proxies to enable continuous assessment during production [80]. Together, these approaches accelerate prototyping, reduce waste, and bridge lab and industrial manufacturing.

Scaling these tools requires high-quality, shareable datasets that link formulations, processes, structures, and sensory outcomes with standardized metadata. Proprietary and fragmented data currently hinder model transfer and benchmarking and underscore the need for domain-specific FAIR implementations, shared testbeds, and community baselines [81]. Ingredient variability remains a crucial bottleneck and robust hybrid models are needed to achieve reliable process control [77, 78]. Industrial adoption also demands capital investment, workforce training, and validation on real production lines since most food digital twins have not been tested at scale [80]. Shared data standards and LLM-enabled knowledge bases can bridge lab and industrial scale and enable consistent, cost-efficient manufacturing at or below price parity.

Recipes: Personalizing food experiences. Beyond production, food preparation shapes consumer experience. Individuals and foodservice operations must balance enjoyment with health, cost, convenience, and allergens while advancing broader goals such as human and planetary health. LLMs offer strong potential because they draw on large recipe datasets and integrate sources such as Recipes1M, USDA FoodData Central, the HESTIA climate database, and Food.com reviews. The key challenge is to embed human knowledge so that models generalize across diverse users, dietary needs, and cultural traditions.

Recent research evaluates and refines LLMs for recipe generation, revision, and preference prediction. Progress relies on fine-tuning, retrieval-augmented generation, and integration with optimization tools. Benchmarks on Recipe1M show that off-the-shelf models produce unreliable outputs [84]. Fine-tuning improves performance [85] and supports revision tasks [86]. A generative model trained on 2,216 curated burger recipes explored 8.9×10^{43} candidate formulations and identified Pareto-efficient recipes that reduced environmental impact without compromising taste [44]. In blinded restaurant surveys, some AI-generated burgers outperformed the BigMac benchmark in flavor, texture, and overall liking [87]. Other work demonstrates retrieval-augmented generation for recipes [88] and sustainable menu design [89]. Some LLMs now match expert food scientists in preference modeling [89]. Combined with optimization algorithms, these models can generate recipes that meet nutritional and sustainability targets.

The ability to generate recipes under explicit constraints creates opportunities across individual and institutional contexts. For individuals, fine-tuned LLMs could power AI health coaches that teach users how to modify meals to be more sustainable, nutritious, or culturally appropriate, while respecting restrictions such as allergies or religious requirements. Chefs and students could practice recipe generation in educational applications, receiving AI-driven feedback to refine their skills [90]. In institutional settings, LLMs could support foodservice operations in offices, universities, or hospitals, tailoring descriptions, ingredients, and preparation methods under strict constraints. This is particularly valuable in programs such as “Food as Medicine,” which deliver cost- and time-sensitive meals for patients with chronic conditions after hospital discharge [91]. In all these scenarios, AI could accelerate constrained recipe generation, broaden access to healthy and sustainable foods, and scale public health interventions that rely on tailored meal planning.

Despite this promise, important challenges remain. Off-the-shelf LLMs perform unevenly across cultures, fail to capture diverse preferences, and produce numerically inconsistent or unsafe outputs. They lack molecular-level understanding and have limited knowledge of food safety. Training models from scratch requires substantial resources. These limitations introduce risks of impractical or unsafe recipes and increase environmental cost. Mitigation strategies include integrating external datasets, using open-source pre-trained models, and coupling LLMs with optimization tools. Beyond technical concerns, systems must preserve user agency and

support human expertise. A more human-centered approach would allow LLMs to assist with drafting and revision, while users adapt recipes to their needs, preferences, and cultures.

Conclusions

AI is beginning to transform food science across the entire innovation cycle, from ingredient design, formulation development, and fermentation and production to texture analysis, sensory properties, manufacturing, and recipe generation (Fig. 2). Yet, the field’s true potential not only lies in optimizing today’s processes, but also in charting a bold trajectory for the future of food as a programmable, predictive science. Building on recent advances, we envision four major areas that define the frontier of AI for sustainable food production:

Future materials for food innovation. Treating food as a programmable biomaterial creates opportunities to design ingredients from first principles: texture, functionality, stability, and sensory perception emerge from multiscale interactions among proteins, carbohydrates, lipids, and processing conditions [20]. This perspective enables the design of edible films and gels with controlled nutrient release and plant-based meat analogues with tailored mechanical and sensory performance. Advances in protein design, including AlphaFold and diffusion-based generative models [92], point toward de novo edible proteins with tunable functionality. Early work in cellular agriculture, such as engineered fat and muscle cells with modified composition [93], demonstrates this potential, while microbial and fungi-based fermentation offer complementary routes. Mycelium fermentation can upcycle agricultural side-streams such as brewers’ spent grain or fruit pomace into protein- and fiber-rich ingredients with distinctive textures and clear links to circularity [94]. A key opportunity lies in integrating molecular design with tunable processing, including fermentation and upcycling, to build predictive platforms that generate new food materials with targeted properties and reduced waste.

Scientific machine learning for food systems. Scientific Machine Learning (SciML) [95] hardwires physical, chemical, and biological principles directly into AI architectures and provides a foundation for trustworthy and scalable food innovation. Physics-Informed Neural Networks (PINNs) [96] enforce conservation laws in data-sparse settings [97] such as fermentation and extrusion. Constitutive Artificial Neural Networks (CANNs) [98] embed mechanical laws to discover physically interpretable models of food texture and rheology [64]. Sparse

Identification of Nonlinear Dynamics (SINDy) [99] discovers compact governing equations from experimental data and can reveal the kinetics of ingredient interactions or microbial growth. The opportunity ahead is to couple physics-based ML with data-driven models to achieve interpretable, generalizable predictions under sparse data and enable mechanism-aware food design.

Automation and self-driving labs. Empirical trial-and-error methods based on rheology, sensory panels, and compositional analysis have long shaped food science, but they remain slow, costly, and hard to scale. Generative AI, robotics, and high-throughput experimentation enable self-driving labs that design, test, and refine food materials within a closed-loop design–build–test–learn cycle, accelerating systematic exploration across scales. Recent advances in automated model discovery with custom-designed constitutive neural networks, show how AI can autonomously identify the best models, parameters, and even experimental protocols directly from data [98]—a method already successfully applied to compare plant-based and animal meats [64], deli products [100], and fungi-based steak [94]. Agentic AI systems such as Sparks [101], BioDiscoveryAgent [102], and the Virtual Lab [103] demonstrate the capacity to generate hypotheses, perform simulations, and report results without human intervention. In food, such systems could integrate genome-scale metabolic models, digital twins of bioreactors, and in-line quality sensors [62, 104]. A fully automated discovery pipeline that links synthesis, characterization, and modeling would substantially accelerate innovation in sustainable, personalized, and scalable food production.

Deep reasoning models. AI models with reasoning capabilities will become central to self-driving lab infrastructure (Fig. 4). The emerging discipline of AI for Food will benefit from rapid progress in reasoning models and agentic systems that design and execute tasks with external tools [105]. Current generative models propose molecules and recipes, but they rely on static datasets and struggle to assess feasibility, manufacturability, or emergent properties. Closing this gap requires systems that generate protein sequences, simulate folding and mechanics, and predict sensory outcomes such as flavor and texture within an interpretable, self-improving loop. Deep reasoning models define the next frontier beyond pattern recognition, and instead hypothesize, simulate, and generate data. They integrate protein sequences, structural mechanics, metabolic fluxes, and sensory responses to predict how molecular composition

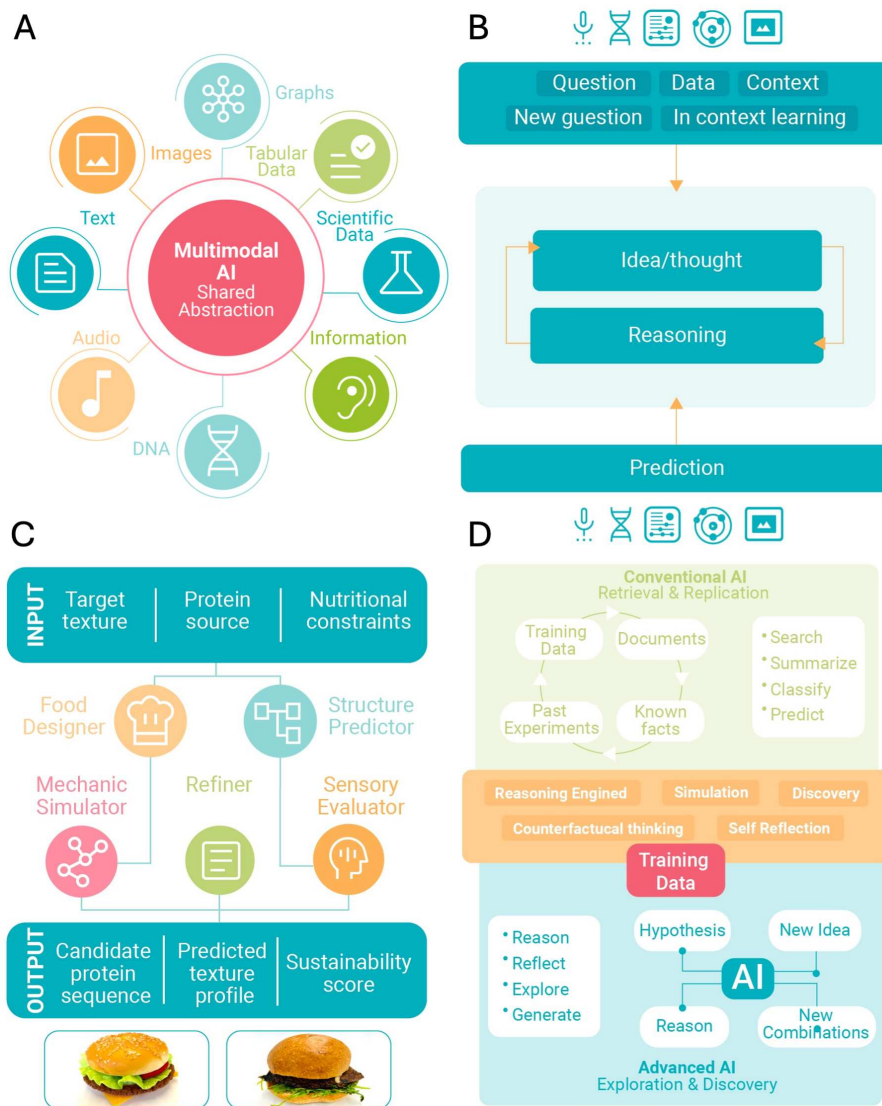


Fig. 4 Next-generation AI systems for scientific discovery in food. (A) Multimodal AI integrates diverse data types such as text, images, graphs, and molecular structures; (B) first-principles AI reasoning models perform iterative thinking and hypothesis generation; (C) a possible workflow applies these capabilities to edible protein design and predicts structure, mechanics, and sensory outcomes by explicitly modeling disparate concepts and their relationships; and (D) the transition from conventional retrieval-based AI to reasoning-enabled, self-reflective AI enables exploration, discovery, and ultimately, the generation of novel, functional foods.

translates into taste, texture, and stability [20]. They can reduce the simulation-to-reality gap by validating predictions against experiments and updating models in real time. Such approaches can guide experiment selection [63], integrate results into finite element simulations of oral processing [106], and predict sensory perception from physical principles [64]. By embedding sustainability constraints, cultural diversity, and personalized health data, these systems can design foods that meet functional, nutritional, and societal goals. The ultimate goal is a closed-loop, reasoning-driven platform that designs, tests, and refines food materials with predictive accuracy across scales.

Limitations and risks. This review uses sustainable proteins as a testbed for AI-enabled food systems and focuses on post-harvest product development, from ingredients to recipes. It does not cover upstream agriculture, supply chains, packaging, safety testing, or retail analytics, domains that also adopt AI and merit dedicated reviews. Benchmarking remains limited by heterogeneous datasets, proprietary silos, and weak metadata standards, which constrain reproducibility and global representation. AI integration also introduces technical and societal risks, including limited transferability, domain shift, and bias from sparse data. In food systems, these risks carry direct safety implications:

inconsistent outputs may yield unsafe formulations, opaque optimization may hide allergen or nutritional trade-offs, and automated processes raise safety and biosecurity concerns. Responsible use of AI requires transparent benchmarks, physics-informed constraints, human-in-the-loop validation, regulatory alignment, and inclusive data design. AI must augment expert oversight to ensure safe, scalable, and sustainable innovation.

Outlook. Taken together, these four areas define a path toward a predictive and design-driven science of food. AI can link molecular composition to structure, processing, and perception and integrate these levels into closed-loop systems that design, test, and refine foods across scales. From programmable biomaterials and physics-informed models to self-driving laboratories and reasoning-based AI, the field now has the foundations to move from empirical optimization to mechanism-aware design. Realizing this vision will require interoperable data standards, rigorous benchmarking, and platforms that integrate experiments, simulations, and learning. Progress will also depend on human-centered design that embeds cultural context, nutrition, and safety into AI systems. If these challenges are addressed, AI will not only accelerate food innovation, but establish a new paradigm in which food design becomes predictive, scalable, and aligned with both human and planetary health.

Acknowledgments We thank Dr. Heideh Fattaey for stimulating discussions on the future of food, Ryan Yow for helping to design Figure 3C, and Goran Atanasovski for illustrating Figures 3 to 4. This project was supported by the USDA Grant FA9550-23-1-0606 to DLK, by the BBSRC Grant BB/Y008510/1, the ERC Grant DEUSBIO-949080, and the Bezos Earth Fund through the Bezos Centre for Sustainable Protein BCSP/IC/001 to RLA, by Novo Nordisk Foundation grant NNF23OC0085919 to MS, by the NSF Graduate Research Fellowship to SRSP, by the UK National Alternative Protein Innovation Centre NAPIC, an Innovation and Knowledge Centre funded by the Biotechnology and Biological Sciences Research Council BBSRC and Innovate UK Grant BB/Z516119/1 to NW, and by the Food@Stanford Snack Grant, the Stanford SDSS Accelerator Grant, the NSF CMMI Award 2320933, and the ERC Advanced Grant 101141626 to EK.

Conflicts of interest The authors declare no competing interests.

Author information Bianca Datta, Good Food Institute, Washington, USA; Markus J. Buehler,

Massachusetts Institute of Technology, Cambridge, USA; Yvonne Chow; Singapore Institute of Food & Biotechnology Innovation SIFBI and Agency for Science, Technology and Research A*STAR, Nanos, Singapore; Kristina Gligoric, Johns Hopkins University, Baltimore, USA; Dan Jurafsky, Stanford University, Stanford, USA; David L. Kaplan, Tufts University, Medford, USA; Rodrigo Ledesma-Amaro, Giorgia Del Missier, Lisa Neidhardt; Bezos Centre for Sustainable Protein; Microbial Food Hub; Centre for Engineering Biology, Imperial College London, London, UK; Karim Pichara, NotCo, San Francisco, USA; Benjamin Sanchez-Lengeling, University of Toronto, Toronto, Canada; Miek Schlangen, University of Southern Denmark, Odense, Denmark; Skyler R. St. Pierre, Stanford University, Stanford, USA; Ilias Tagkopoulos, USDA/NIFA AI Institute for Next-Generation Food Systems, University of California, Davis, USA; Anna Thomas, Stanford University, Stanford, USA; Nik Watson, University of Leeds, Leeds, UK and National Alternative Protein Innovation Centre NAPIC, UK; Ellen Kuhl, Stanford University, Stanford, USA

Authors' contributions BD and EK designed the layout, BD, MJB, YC, KG, DJ, DLK, RLA, GDM, LN, KP, BSL, MS, SRSP, IT, AT, NW, and EK wrote and edited the paper.

References

- [1] Lappé, F.M. Diet for a Small Planet. Ballantine Books, New York (1971).
- [2] Clark, M. A., et al. Global food system emissions could preclude achieving the 1.5 and 2C climate change targets. *Science*. **370**, 705-708 (2020).
- [3] Xu, X. et al. Global greenhouse gas emissions from animal-based foods are twice those of plant-based foods. *Nature Food*. **2**, 724-732 (2021).
This study shows that animal-based foods produce roughly twice the greenhouse gases of plant-based foods. It underscores the environmental benefits of plant and microbial proteins central to sustainable food design.
- [4] Keesing, F. Diet for a small footprint. *Proc. Nat. Acad. Sci.* **119**, e2204241119 (2022).
- [5] Friedrich, B. Transforming a 12,000-year-old technology. *Nature Food* **3** 807-808 (2022).
- [6] Barabasi, A.L. et al. The unmapped chemical complexity of our diet. *Nature Food* **1** 33-37 (2020).
This perspective reveals the vast and largely unexplored chemical diversity in everyday foods. It

- highlights the need for computational and AI-driven mapping of food molecules to understand nutrition, flavor, and health interactions.*
- [7] Zeni, C. et al. A generative model for inorganic materials design. *Nature*. **639** 624–632 (2025).
 - [8] Al-Sarayreh, M. et al. Inverse design and AI/Deep generative networks in food design: A comprehensive review. *Trends Food Sci. Tech.* **138** 215–228 (2023).
 - [9] Datta, A. et al. Computer-aided food engineering. *Nature Food*. **3** 894–904 (2022).
This review outlines how computation, modeling, and automation can transform food engineering. It establishes a framework for integrating AI, digital twins, and process optimization into food design and manufacturing.
 - [10] King, A. Four ways to power-up AI for drug discovery. *Nature*. doi: 10.1038/d41586-025-00602-5 (2025).
 - [11] Kuhl, E. AI for food: accelerating and democratizing discovery and innovation. *npj Sci. Food*. **9**, 82 (2025).
 - [12] Turing, A. M. Computing machinery and intelligence. *Mind*. **LIX**, 433–460 (1950).
 - [13] Rumelhart, D. E. et al. Learning representations by back-propagating errors. *Nature*. **323**, 533–536 (1986).
 - [14] Krizhevsky, A. et al. ImageNet classification with deep convolutional neural networks. *Comm. ACM* **60**, 84–90 (2017).
 - [15] Goodfellow, I. et al. Generative adversarial networks. in: *Advances in Neural Information Processing Systems (NeurIPS)*. **27**, 2672–2680 (2014).
This work presents the GAN framework for generating realistic synthetic data. It underpins AI applications in food imaging, formulation design, and sensory data augmentation.
 - [16] Vaswani, A. et al. Attention is all you need. in: *Advances in Neural Information Processing Systems (NeurIPS)*. **30**, 5998–6008 (2017).
This landmark paper introduces the transformer architecture of today’s large language models. It enables multimodal AI systems for recipe generation, ingredient substitution, and scientific discovery.
 - [17] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
 - [18] Gilmer, J. et al. Neural message passing for quantum chemistry. in: *Proc. 34th Int. Conf. Machine Learning (ICML)*. **34**, 1263–1272 (2017).
This paper introduces message-passing neural networks for predicting molecular properties from graph structures. It lays the foundation for AI models in food chemistry and ingredient design.
 - [19] Zeevi, D. et al. Personalized nutrition by prediction of glycemic responses. *Cell* **163**, 1079–1094 (2015).
 - [20] Gordon, E. B. et al. Biomaterials in cellular agriculture and plant-based foods for the future. *Nat. Rev. Materials* **10**, 500–518 (2025).
 - [21] Schreurs, M. et al. Predicting and improving complex beer flavor through machine learning. *Nat. Comm.* **15**, 2368 (2024).
 - [22] Lee, B.-K. et al. A principal odor map unifies diverse tasks in olfactory perception. *Science*. **381**, 999–1006 (2023).
This work creates a continuous map linking molecular structure to human odor perception. It enables AI models to predict and generate new flavor and aroma profiles for food design.
 - [23] Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. **596**, 583–589 (2021).
This breakthrough model achieves near-experimental accuracy in protein structure prediction. It inspires the design of functional food proteins and enzymes for alternative protein production.
 - [24] Butler, K. T. et al. Machine learning for molecular and materials science. *Nature*. **559**, 547–555 (2018).
This review outlines how machine learning accelerates discovery in chemistry and materials science. It provides key methodologies directly transferable to AI-based formulation and texture engineering in food.
 - [25] Graham, A. E. & Ledesma-Amaro, R. The microbial food revolution. *Nat. Comm.* **14**, 2231 (2023).
 - [26] Lurie-Luke, E. Alternative protein sources: science-powered startups to fuel food innovation. *Nat. Comm.* **15**, 4425 (2024).
 - [27] Loveday, S. M. Food proteins: Technological, nutritional, and sustainability attributes of traditional and emerging proteins. *Ann. Rev. Food Sci. Tech.* **10**, 311–339 (2019).

- [28] Kyriakopoulou, K. et al. Plant-based meat analogues. in: *Sustainable Meat Production and Processing*, Academic Press, 103–126 (2019).
- [29] Fasolin, L. H. et al. Emergent food proteins—Towards sustainability, health and innovation. *Foods*. **10**, 600 (2021).
- [30] Day, L. Proteins from land plants—Potential resources for human nutrition and food security. *Trends Food Sci. & Tech.* **32**, 25–42 (2013).
- [31] Liu, M.-Q. et al. Digging natural emulsifiers based on machine learning and exploration their performance for stabilizing dairy products. *SSRN*. <https://ssrn.com/abstract=4994781> (2025).
- [32] Siejak, P. et al. The prediction of pectin viscosity using machine learning based on physical characteristics—case study: Aglupectin HS-MR. *Sustainability*. **16**, 5877 (2024).
- [33] Lie-Piang, A. et al. Quantifying techno-functional properties of ingredients from multiple crops using machine learning. *Curr. Res. Food Sci.* **7**, 100601 (2023).
- [34] Yilmaz, M. T. et al. Explainable AI-driven evaluation of plant protein rheology using tree-based and Gaussian process machine learning models. *Ain Shams Eng. J.* **16**, 103565 (2025).
- [35] Kraessig, P. M. et al. Sensory-biased autoencoder enables prediction of texture perception from food rheology. *Food Res. Int.* **205**, 116007 (2025).
- [36] Meng, Y. et al. Advancing salt reduction technologies: AI-assisted structural design of starch-based emulsion gel systems for next-generation low sodium food formulations. *Trends Food Sci. & Tech.* doi:10.1016/j.tifs.2025.105234 (2025).
- [37] NotCo. Latent space method of generating food formulas. *US 10,915,818*. Feb 9, 2021.
- [38] NotCo. Neural network method of generating food formulas. *US 10,957,424*. Mar 23, 2021.
- [39] NotCo. Systems and methods to mimic target food items using artificial intelligence. *US 11,164,478*. Nov 2, 2021.
- [40] Martin, J. et al. Recipe1M+: A dataset for learning cross-modal embeddings for cooking recipes and food images. *IEEE Trans. Pattern Anal. Mach. Intell.* **43** (2021).
- [41] Park, D. et al. FlavorGraph: a large-scale food-chemical graph for generating food representations and recommending food pairings. *Sci. Rep.* **11**, 931 (2021).
- [42] Garg, N. et al. FlavorDB: a database of flavor molecules. *Nucl. Acids Res.* **46**, D1210–D1216 (2018).
- [43] Ahn, Y.-Y. et al. Flavor network and the principles of food pairing. *Sci. Rep.* **1**, 196 (2011).
- [44] Tac, V. et al. Generative artificial intelligence creates delicious, sustainable, and nutritious burgers. *arXiv* doi:10.48550/arXiv.2602.03092 (2026).
This study uses generative AI to create delicious, sustainable, and nutritious burgers that outperformed the classical Big Mac in a blinded restaurant survey.
- [45] van den Bedem, S.D. et al. Open-source benchmarking of plant-based and animal meats. *arXiv*. doi:10.48550/arXiv/2603.03370 (2026).
- [46] Eastham, J. L. & Leman, A. R. Precision fermentation for food proteins: ingredient innovations, bioprocess considerations, and outlook—a mini-review. *Curr. Op. Food Sci.* **58**, 101194 (2024).
- [47] Feng, X. et al. Enhanced lipid production by *Chlorella pyrenoidosa* through magnetic field pretreatment of wastewater and treatment of microalgae-wastewater culture solution: Magnetic field treatment modes and conditions. *Bioresour. Technol.* **206**, 123102 (2020).
- [48] Sharma, D. et al. Response surface methodology and artificial neural network modelling for enhancing maturity parameters during vermicomposting of floral waste. *Bioresour. Technol.* **324**, 124672 (2021).
- [49] Liao X. et al. Artificial Intelligence: A solution to involution of Design-Build-Test-Learn cycle *Curr. Op. Biotechnol.* **75**, 102712 (2022).
- [50] Liu, L. et al. Modeling and optimization of microbial hyaluronic acid production by *Streptococcus zooepidemicus* using radial basis function neural network coupling quantum-behaved particle swarm optimization algorithm. *Biotechnol. Prog.* **25**, 1819–1825 (2009).
- [51] Jia, R. et al. Exploration of deep learning models for real-time monitoring of state and performance of anaerobic digestion with online sensors. *Bioresour. Technol.* **363**, 127908 (2022).
- [52] Cheng, Y. et al. Artificial intelligence technologies in bioprocess: Opportunities and challenges. *Biores. Tech.* **369**, 128451 (2023).
- [53] Treloar, N.J. et al. Deep reinforcement learning for the control of microbial co-cultures in bioreactors. *PLoS Comput. Biol.* **16**, e1007783 (2020).

- [54] Singhal, A. et al. Pretreatment of *Leucaena leucocephala* wood by acidified glycerol: optimization, severity index and correlation analysis. *Bioresour. Technol.* **265**, 214–223 (2018).
- [55] Austerjost, J. et al. A machine vision approach for bioreactor foam sensing. *SLAS. Technol.* **26**, 408–414 (2021).
- [56] Butean, A. et al. A review of artificial intelligence applications for biorefineries and bioprocessing: From data-driven processes to optimization strategies and real-time control. *Processes.* **13**, 2544 (2025).
- [57] Alejo, L. et al. Effluent composition prediction of a two-stage anaerobic digestion process: machine learning and stoichiometry techniques. *Environ. Sci. Pollut. Res. Int.* **25**, 21149–21163 (2018).
- [58] Akinade, O. O. & Oyedele, L. O. Integrating construction supply chains within a circular economy: An ANFIS-based waste analytics system (A-WAS). *J. Clean. Prod.* **229**, 863–873 (2019).
- [59] Peng, J. et al. Time-dependent fermentation control strategies for enhancing synthesis of marine bacteriocin 1701 using artificial neural network and genetic algorithm. *Bioresour. Technol.* **138**, 345–352 (2013).
- [60] Sharma, D. & Singh, K. AI-enhanced bioprocess technologies: machine learning implementations from upstream to downstream operations. *World J. Microbio. Biotech.* **41**, 278 (2025).
- [61] Holzinger A. et al. AI for Life: Trends in Artificial Intelligence for biotechnology. *New Biotech.* 16–24 (2023).
- [62] Ghafarollahi, A. & Buehler, M. J. ProtAgents: protein discovery via large language model multi-agent collaborations combining physics and machine learning. *Digit. Discovery.* **3**, 1389–1409 (2024).
- [63] St. Pierre, S. R. et al. Discovering the mechanics of artificial and real meat. *Comp. Meth. Appl. Mech. Eng.* **415**, 116236 (2023).
- [64] St. Pierre, S. R. et al. The mechanical and sensory signature of plant-based and animal meat. *npj Sci. Food.* **8**, 94 (2024).
This study quantitatively links mechanical texture measurements with human sensory perception in plant-based and animal meats. It provides critical benchmark data for AI models to predict texture and mouthfeel from physical properties.
- [65] Sanahuja, M. et al. Classification of puffed snacks freshness based on crispiness-related mechanical and acoustical properties. *J. Food Eng.* **226**, 53–64 (2018).
- [66] Dahl, J. F. et al. Predicting rheological parameters of food biopolymer mixtures using machine learning. *Food Hydrocoll.* **160**, 110786 (2025).
- [67] Zhang, Y. et al. Construction of an intelligent recognition system for the cooking doneness of deep-fried golden pompano (*Trachinotus ovatus*) based on deep learning. *Food Biosci.* **68**, 106644 (2025).
- [68] Saalbrink, J. et al. Quantifying microscopic droplets in colloidal systems through machine learning-based image analysis. *Food Hydrocoll.* **166**, 111301 (2025).
- [69] Dunne, R. A. et al. Texture profile analysis and rheology of plant-based and animal meat. *Food Res. Int.* **205**, 115876 (2025).
- [70] Munekata, P. E. S. et al. Applications of electronic nose, electronic eye and electronic tongue in quality, safety and shelf life of meat and meat products: A review. *Sensors.* **23**, 672 (2023).
- [71] Karakaya, D. et al. Electronic nose and its applications: A survey. *Int. J. Auto. Comp.* **17**, 179–209 (2019).
- [72] Bengio, Y. et al. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Machine Intell.* **35**, 1798–1828 (2012).
- [73] Tom, G. et al. From molecules to mixtures: Learning representations of olfactory mixture similarity using inductive biases. *arXiv* doi:10.48550/arXiv.2501.16271 (2025).
- [74] Istif, E. et al. Miniaturized wireless sensor enables real-time monitoring of food spoilage. *Nature Food.* **4**, 427–436 (2023).
- [75] Watson, N. J. et al. Intelligent sensors for sustainable food and drink manufacturing. *Sustain. Food Syst.* **5**, 642786 (2021).
- [76] Rombach, R. et al. High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 10684–10695 (2022).
- [77] Watson, N. & Rafiq, S. NAPIC Partner Engagement Workshop Report. <https://doi.org/10.48785/100/320>. accessed 15 Aug. 2025.
- [78] Dinali, M. et al. Fibrous structure in plant-based meat: High-moisture extrusion. *J. Food Process Preserv.* **40**, 2940–2968 (2024).

- [79] Jiang, Y. et al. Plant-based protein extrusion optimization: Comparison between machine learning and conventional experimental design. *Curr. Res. Food Sci.* **11**, 101157 (2025).
- [80] Reducing the Price of Alternative Proteins. <https://gfi.org/reducing-the-price-of-alternative-proteins>. accessed 15 Aug. 2025.
- [81] Abdurrahman, E. E. M. et al. Digital Twin applications in the food industry: a review. *Front. Sustain. Food Syst.* **9**, 1538375 (2025).
- [82] Top, J. et al. Cultivating FAIR principles for agri-food data. *Comput. & Electron. Agric.* **196**, 106909 (2022).
- [83] Bölücü, N. et al. An evaluation of Large Language Models for supplementing a food extrusion dataset. *Foods*. **14**, 1355 (2025).
- [84] Mohbat, F. & Mohammed, J. Z. Llava-chef: A multi-modal generative model for food recipes. *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. (2024).
- [85] Li, P. et al. Cheffusion: Multimodal foundation model integrating recipe and food image generation. *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management* (2024).
- [86] Senath, T. et al. Large Language Models for ingredient substitution in food recipes using supervised fine-tuning and direct preference optimization. *arXiv* doi:10.48550/arXiv.2412.04922 (2024).
- [87] Tac, V. & Kuhl, E. Generative AI for material design: A mechanics perspective from burgers to matter. *arXiv* doi:10.48550/arXiv.2604.03409 (2026).
- [88] Liu, G. et al. Retrieval augmented recipe generation. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (2025).
- [89] Thomas, A. T. et al. What can large language models do for sustainable food? *International Conference on Machine Learning*. (2025).
- [90] Yang, D. et al. Social skill training with large language models. *arXiv*. doi:10.48550/arXiv.2404.04204 (2024).
- [91] Rosas, L. G. et al. The effectiveness of Recipe4Health: a quasi-experimental evaluation. *Am. J. Prevent. Med.* **68**, 377-390 (2025).
- [92] Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. **630**, 493-500 (2024).
- [93] Yuen, J. S. K. et al. Aggregating in vitro-grown adipocytes to produce macroscale cell-cultured fat tissue with tunable lipid compositions for food applications. *eLife*. **12**, e82120 (2023).
- [94] Vervenne, T. et al. Probing mycelium mechanics and taste: The moist and fibrous signature of fungi steak. *Acta Biomat.* **202**, 341-351 (2025).
- [95] Rackauckas, C. et al. Universal differential equations for scientific machine learning. *arXiv* doi:10.48550/arXiv.2001.04385 (2020).
- [96] Raissi, M. et al. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comp. Phys.* **378**, 686-707 (2019).
- [97] Karniadakis, G. E., et al. Physics-informed machine learning. *Nat. Rev. Phys.* **3**, 422-440 (2021).
- [98] Linka, K. & Kuhl, E. A new family of Constitutive Artificial Neural Networks towards automated model discovery. *Comp. Meth. Appl. Mech. Eng.* **403**, 115731 (2023).
- [99] Brunton, S. L., et al. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Nat. Acad. Sci.* **113**, 3932-3937 (2016).
- [100] St. Pierre, S. R. et al. Biaxial testing and sensory texture evaluation of plant-based and animal deli meat. *Curr. Res. Food Sci.* **10**, 101080 (2025).
- [101] Ghafarollahi, A. & Buehler, M. J. Sparks: Multi-agent artificial intelligence model discovers protein design principles. *arXiv*. doi:10.48550/arXiv.2504.19017 (2025).
- [102] Roohani, Y., et al. BioDiscoveryAgent: An AI agent for designing genetic perturbation experiments. *International Conference on Learning Representations*. *arXiv*. doi:10.48550/arXiv.2405.17631 (2024).
- [103] Swanson, K., et al. The Virtual Lab of AI agents designs new SARS-CoV-2 nanobodies. *Nature*. doi:10.1038/s41586-025-09442-9 (2025).
- [104] Buehler, M. J. Accelerating scientific discovery with generative knowledge extraction, graph-based representation, and multimodal intelligent graph reasoning *Mach. Learn.: Sci. Technol.* **5**, 035083 (2024).
- [105] Boiko, D. A., et al. Autonomous chemical research with large language models. *Nature* **624**, 570-578

(2023).

This paper demonstrates how large language models can autonomously plan, execute, and interpret chemical experiments. It marks a major step toward self-driving discovery systems, directly inspiring similar AI-driven pipelines for food formulation and ingredient design.

- [106] Peirlinck M, et al. On automated model discovery and a universal material subroutine for hyperelastic materials. *Comp. Meth. Appl. Mech. Eng.* **418**, 116534 (2024).