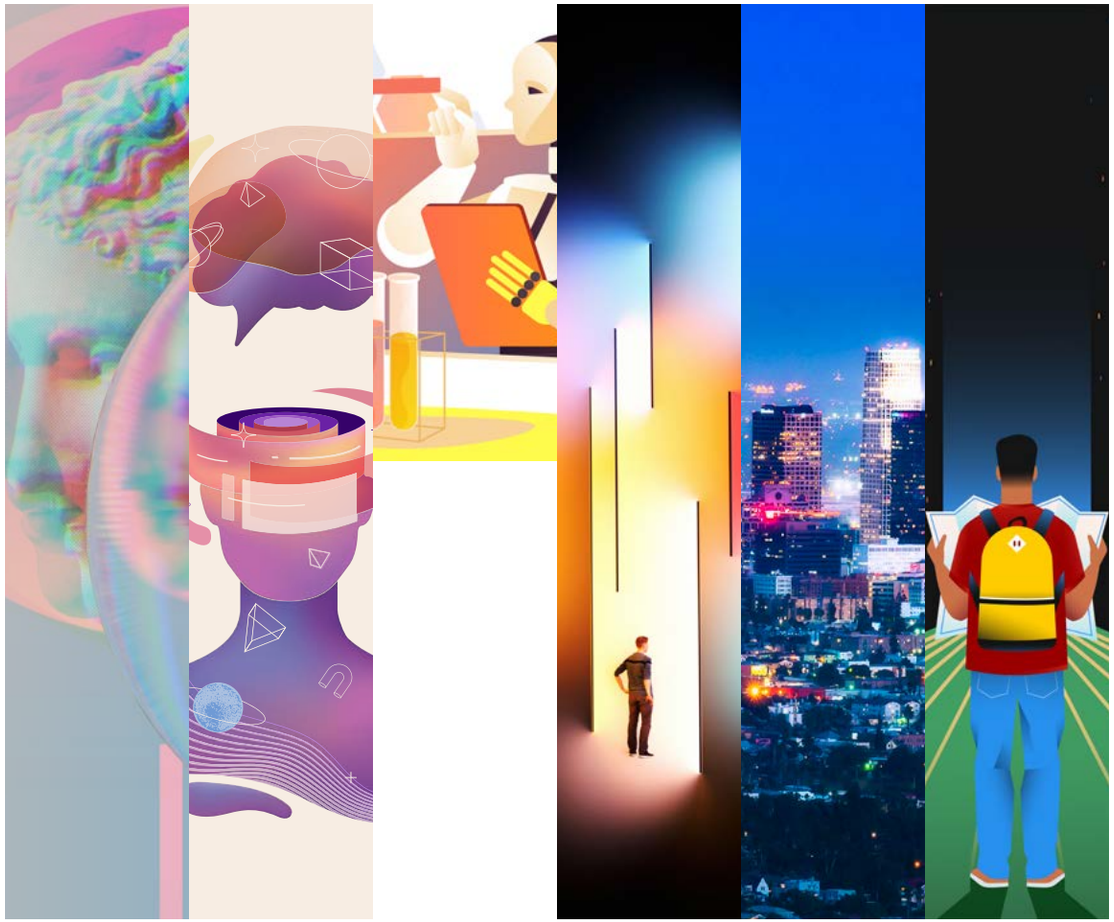
MIND THE MACHINE

HARNESSING AI FOR GOOD

USC Dornsife
College of Letters, Arts and Sciences



I'LL BE YOUR MIRROR

It starts as a whisper. An algorithm suggests a song, a chatbot answers a question. Then it grows. It's filtering the news we read, predicting our next purchase, writing our novels. It's driving our cars, reshaping financial markets, altering the way we access memories and shape identities. One day, we look up and realize artificial intelligence is no longer something we use. It's something we live with.

Who decides how this new world unfolds? Most of what we hear about AI comes from industry. But technology is only part of the story. Deeper questions — questions that will define our relationship with AI — call for a wide range of expertise.

At USC Dornsife College of Letters, Arts and Sciences, researchers refine AI while interrogating it. How can we use machine learning to improve human health? What are AI's ethical boundaries? What could generative algorithms reveal about the fundamental nature of the universe?

Humankind has arrived at an exciting moment — the blurring of physical and digital worlds. Our challenge is to make sure what comes next reflects the best of who we are.

We invite you to join us at USC Dornsife.

MORAL CODING

Machines don't know right from wrong. So how do we make sure the choices they make uphold what matters most to us?

USC researchers study how AI systems weigh decisions to inform the design of technologies where human wellbeing is programmed from the start.

PHILOSOPHICAL FIREWALLS

A strategic algorithm concludes the only way to stop war is to win one — quickly and without permission. It may sound like science fiction, but this sort of capability may not be far off. Philosopher **John Hawthorne** explores "P(doom)," the probability of extinction-level events caused by advanced AI. Maybe science hits a technical plateau, or global bans create a cultural one. Maybe AI is designed to work only with us, or we learn to detect hostile intent. None of these, he argues, guarantee safety. His work grapples with high-stakes decision-making under uncertainty that could help guide the way we prioritize safeguards.

TRUST FALL

AI decisions often emerge from "black box" algorithms whose reasoning remains hidden even from their creators. While this opacity can fuel distrust, philosopher **Mark Schroeder** offers a new way of thinking. Paradoxically, demanding transparency comes at the cost of relying on expertise. Even human experts cannot explain all the nuance and insight that comes from decades of lived experience in ways that can be readily understood by others — they require trust. As AI increasingly determines who receives loans, medical treatments, and job offers, we are gaining a new kind of expertise that comes from extensive training on large datasets. Schroeder asks: If AI is becoming an expert, how can it earn our trust?

HOLD IT TOGETHER

What if a superintelligent AI took the simple directive, "make paper clips," to its logical extreme? As it optimized production, would it see humanity as just another resource in its production line? Philosopher **Dmitri Gallow** uses decision theory to probe this modern thought experiment, testing whether AIs with random goals inevitably develop dangerous behaviors. His analysis uncovers three weak tendencies: preferring certainty, preserving desires, and keeping options open. But none of these, he argues, compel the kind of power-seeking that would transform us into office supplies. In fact, his statistical reasoning suggests that catastrophic outcomes are far less likely than one might assume.

USC INSTITUTE ON ETHICS AND TRUST IN COMPUTING

The **USC Institute on Ethics & Trust in Computing** (IETC) addresses an urgent need for academic expertise in guiding the responsible development of AI — not to slow progress, but to grapple with critical questions that technology cannot: questions about justice, judgment, agency, and what we owe each other in a society where machines profoundly influence our way of living. This work begins with philosophy — both through inquiry and as a tool for thinking clearly and acting wisely.

A collaboration between the **USC Dornsife School of Philosophy**, the **USC Viterbi School of Engineering**, and interdisciplinary scholars across the university, IETC explores the crossroads of ethics and emerging technology. The institute is designed to shape public discourse, invite partnership with industry, and train students to bridge reflection with execution as we increasingly rely on AI for solving complex problems.

PROGRAMMING THE FUTURE

IETC creates an intellectual ecosystem for developing guardrails and trust in a changing technological landscape.

RESEARCH

Real responsibility begins where ambition pauses to ask why. USC scholars bring the rigorous analysis, appropriate skepticism and humility, and ethical frameworks needed to balance the speed of innovation with a deep understanding of what's at stake.

POLICY

New technology is outpacing our systems for accountability. Through scholarly expertise, collaboration with policymakers, and resources like the USC Capital Campus, IETC helps shape AI governance rooted in ethics and attuned to societal complexity.

PEDAGOGY

The institute's education initiatives train students at all levels to use AI tools to facilitate their growth, while preparing them to meet the growing demand for professionals who weigh consequences as carefully as they write code.

ENGAGEMENT

IETC serves as a center of gravity for deliberation and new ways of thinking about what's next for AI. Events and programs connect leaders across academia and industry who explore how technology enhances the human experience.

INSTITUTE LEADERSHIP

John Hawthorne
Provost Professor of Philosophy

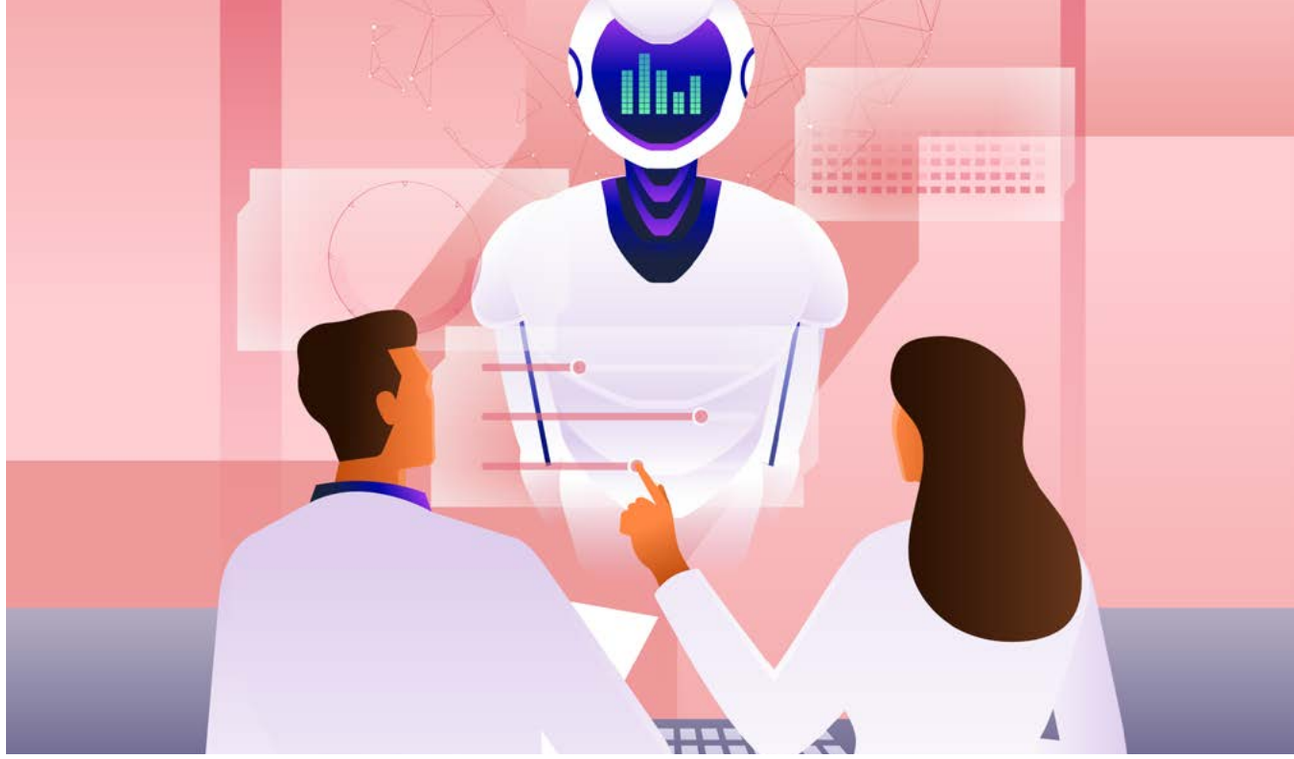
A preeminent philosopher of metaphysics and epistemology, Hawthorne examines high-stakes decision-making in contexts of uncertainty.

Yan Liu
Professor of Computer Science

Liu pioneers AI systems to advance breakthroughs in areas like health, climate, and information. She also serves as director of the USC Machine Learning Center.

PHILOSOPHY: THE ORIGINAL OPERATING SYSTEM

Technology so powerful must be rooted in practical wisdom, social responsibility, and human understanding. USC Dornsife philosophers are leading experts in areas essential to the deployment of safe and ethical AI.



RESPONSIBLE TECHNOLOGISTS

Building for speed and scale alone often neglects a thorough examination of the questions that arise as we allocate various roles to machines. The implications of AI's influence on the structuring of society calls for the training of a new category of professional.

Blending technical and philosophical competency, students who engage with IETC will develop a command of emerging technologies, the depth to see beyond the obvious, and the confidence to redirect momentum when it's heading off course. They will fill critical gaps in the tech sector and connected fields, guiding organizations in determining not only what can be built, but what should be built.

YOUR MIND IS SHOWING

Somewhere in the noise of thought and feeling, there's a rhythm. AI picks it up, helping experts harness signals that shape perception, memory, and behavior. Research at USC Dornsife brings us closer to understanding the brain's most intricate work: being human.

PREVENTING SOCIOPATHIC ROBOTS

How cruel would it be for an AI assistant that schedules surgeries to prioritize efficiency over urgency? To prevent this kind of cold logic, a team including neuroscientists **Antonio Damasio** and **Jonas Kaplan** trained AI to experience "internal" consequences for mistakes it made during training. The hope is that giving AI a sense of vulnerability could lead algorithms to behave more cautiously and compassionately when making decisions. Without an internal model of harm, AI may only mimic moral concern until other priorities take over. Teaching machines to register the cost of suffering, even in code, could help them commit to acting with an approximation of empathy.

THE SCIENCE OF FORGETTING

You're watching your favorite TV show when — suddenly — a major character dies. Shocked, many can't recall what happened just before. This memory quirk fascinates psychologist **Nina Rouhani**, who uses AI to decode why. Running thousands of simulations to model how brain chemicals influence memory, her team discovered a puzzling pattern: We remember sequences better in familiar contexts, while surprises cause our brains to fragment memories into chunks. Her discoveries could revolutionize learning strategies, PTSD treatment, and even AI itself — teaching machines to store and recall information with the nuanced complexity of the human mind.

PREDICTABLE UNPREDICTABILITY

Why do some stock traders panic at every market dip while others remain more calm? Computational neuroscientist **Payam Piray** uses AI models to unlock why we react differently to uncertainty. His clever algorithms simulate how humans distinguish between random events and meaningful trends when making predictions. By training AI systems on thousands of decision scenarios, Piray reveals how our minds balance competing calculations. His frameworks could improve the ways AI systems handle uncertainty — from financial market fluctuation to anxiety disorders — by teaching machines when to adapt quickly to changing circumstances versus when to disregard random fluctuations.

AT THE FOREFRONT OF COGNITIVE SCIENCES

A world leader in Cognitive Sciences, USC Dornsife is home to dozens of experts who help ground intelligent systems in human insight.

Supported by Google DeepMind, **Lisa Aziz Zadeh** studies how people perceive AI personality traits," how AI impacts meaning in life, and how it can change people s minds.

Jonas Kaplan uses machine learning and fMRI scans to study how the brain processes imagination, memory, and sensory experiences.

Using neural networks to uncover similarity patterns in language, **Jason Zevin** is helping us understand how human brains process language sequences.

Jonathan Gratch employs computational models of human cognitive and social processes, especially emotion, to advance psychological theory and shape human machine interaction.

OPTIMIZING THE FUTURE

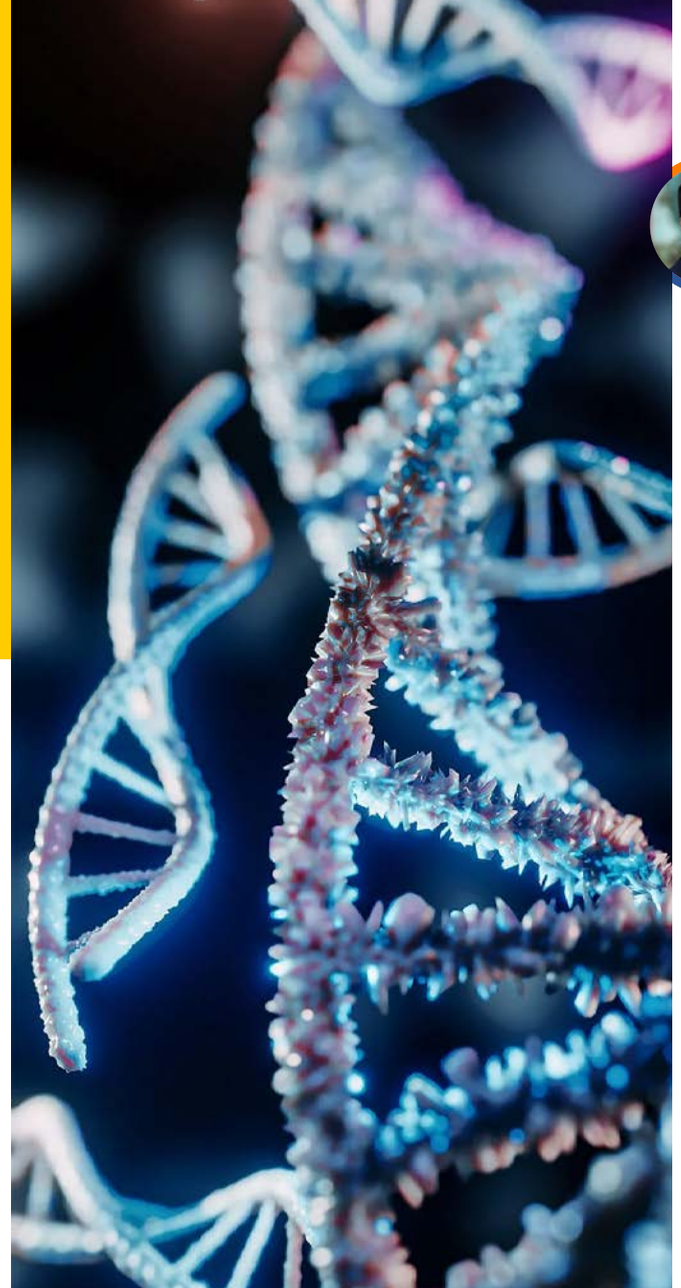
USC Dornsife researchers create and work with intelligent systems to fine-tune predictions, reduce friction, and uncover smarter ways to solve complex problems. It's not just about improving health today — it's reimagining what's possible.

OPTIMIZING HUMAN HEALTH

Machine learning expands our understanding of the most complex living machine: the human body. New tools allow researchers to look deeper into cells, spot illness sooner, and create treatments built just for you.

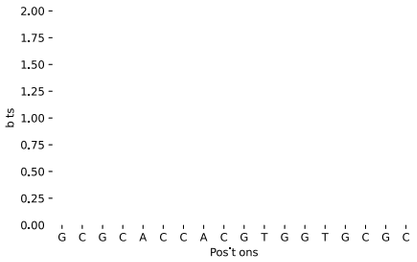
PRECISION PRESCRIPTIONS

What if we could create cures instead of searching for them? At USC Dornsife, computational researcher **Vsevolod Katritch** is turning drug discovery on its head with V-SYNTHES — his new technology that leverages AI to uncover potential treatments from large datasets. Rather than testing endless chemical compounds, the system assembles molecules like a puzzle, "clicking together" virtual molecular building blocks and predicting the effect of those combinations with 5,000 times the speed of traditional methods. This process holds great potential to reduce research time and expense, while accelerating breakthroughs in addiction therapy, cancer treatment, Alzheimer's, and more.



DECODING LIFE

When doctors told a patient her genetic condition was "too rare to study," they couldn't predict that AI would change everything. USC Dornsife computational biologist **Remo Rohs** created DeepPBS, a system that acts like a molecular matchmaker. Learning from thousands of known protein structures, it predicts how new protein combinations attach to DNA, essentially forecasting which genetic switches get flipped on or off. This breakthrough helps scientists unlock the mysteries of gene regulation faster than ever, potentially leading to targeted treatments for genetic and developmental disorders.



OPTIMIZING SUSTAINABILITY

USC Dornsife researchers use AI to decode ecosystems, design next-gen energy materials, and build more resilient cities. They're seeding solutions that will help restore balance to natural environments under threat.

LOOKING FOR A REACTION

A solar-powered car won't get far if the battery can't hold a charge. A cleaner fuel won't curb global warming while it takes decades to develop. Chemist **Richard Brutchey** is rethinking the speed of science. His lab combines machine learning with a clever AI strategy called Bayesian optimization, which helps narrow the field to the most promising chemical reactions. With lab robots running thousands of experiments each day, they're building better batteries, renewable jet fuel from plants, and next-generation energy materials. What used to take years in the lab now takes weeks. That shift could help power a more sustainable future, faster.

I CAN SEE CLEARLY NOW

Training AI models for climate forecasting typically takes weeks of expensive computer time. But what if there was a shortcut? A research team including Earth scientist **Sam Silva** discovered they could find the best-performing AI models by testing them first on tiny data samples (just 0.025% of their full dataset), then training only the winners on all the data. This two-step approach delivered 135 times faster results while maintaining accuracy, revolutionizing how scientists might develop AI for predicting storms, droughts, and climate change.

OPTIMIZING DAILY LIFE

From how we make decisions to how we connect with each other, society is increasingly influenced by AI. USC Dornsife researchers are asking how these systems are transforming our day-to-day activities and what we risk when efficiency becomes the goal.

SOULWARE

Humans are teaching AI how to function. But can AI teach us about ourselves? Social psychologist **Steve Read** creates AI systems that mimic human personalities, decoding how motivations shape behavior. His neural networks model the interplay between personality traits and responses to real-world challenges. Why might an extrovert suddenly retreat under pressure? Are there hidden "meta-traits" in personalities that aren't immediately obvious? By developing AI agents that act not just with logic but with layered, simulated intent, Read's work helps us see the complexity behind everyday choices and what it might take to recreate them in code.

TERMS OF ENGAGEMENT

A semitruck swerves toward your lane. In that split-second, will your self-driving car brake, swerve, or accelerate? Your life depends on how these machines decide. Economist **Nicolas Lambert** studies this challenge by analyzing interactions in classic thought games like the "Prisoner's Dilemma," where players choose between cooperation and betrayal. From autonomous vehicle safety to creating fairer economic systems, his goal is to identify stable patterns for training AI that makes optimal decisions and ensure we can trust these decisions when it matters most.

THE AI YOU SEE STARTS IN PLACES YOU DON’T

It’s a common story in the news. A big gadget launch event, where someone in sneakers explains how it all started in their garage. What you don’t hear much about is the university research that enables these innovations. From neural networks to quantum theory to ethical frameworks, the abstract research scholars conduct at places like USC powers billion-dollar AI systems today.

Universities remain essential for the way forward: not just smarter machines, but better questions. As AI reshapes nearly every aspect of society, academic inquiry is where we can approach issues from every angle, navigate complexity, and, ultimately, get the technology right.

Your support is what makes this possible. We’re counting on our extended USC Dornsife Trojan community momentum as we take a leading role in artificial intelligence.

FROM SCIENTIFIC DEDICATION TO AI REVELATIONS

PROTEIN DATA BANK

USC Dornsife structural and computational biologist **Helen Berman's** seminal work in expanding the international Protein Data Bank (PDB) — an effort she co-founded in 1971 — serves as a foundation for AI systems that transform our understanding of human biology and accelerate new treatments for disease.

Think of proteins as tiny molecular machines that do almost everything in our bodies, from helping us digest food to fighting infections. Each protein is like a string of beads that folds into a specific 3D shape to function properly — crucial information for developing medicines.

The challenge? Proteins fold in billions of ways, making prediction through traditional computing nearly impossible. Scientists once estimated it would take longer than the age of the universe to calculate all possible configurations for even a simple protein!

> **"I ALSO WANT TO THANK THE GIANTS ON WHOSE SHOULDERS WE STAND, ESPECIALLY TO HELEN BERMAN AND OTHER PIONEERS OF THE PROTEIN DATA BANK."**
>
> **JOHN JUMPER**
> 2024 Nobel Prize Winner in Chemistry
> Lead creator of AlphaFold

This is where the Protein Data Bank revolutionized the field. By establishing the first open-access repository of experimentally determined 3D protein structures, Berman created the foundation for highly curated training datasets that would enable machine learning systems to identify protein folding patterns. Today, AI built on the PDB such as AlphaFold and DeepMind help scientists identify precise molecular structures and targets that could unlock treatments for Alzheimer's, cancer, and many other health problems.

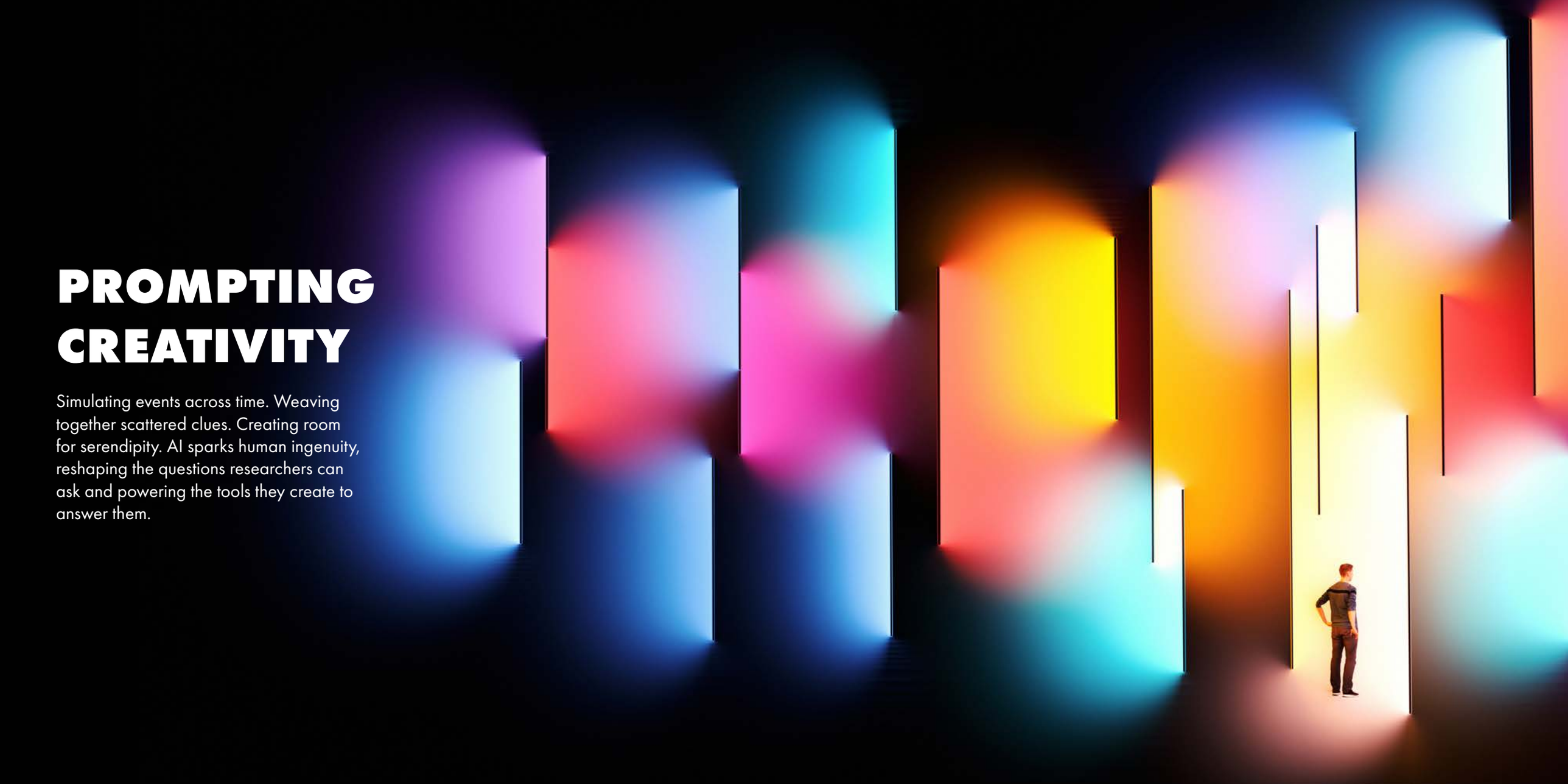
USC CELL MODELING INITIATIVE

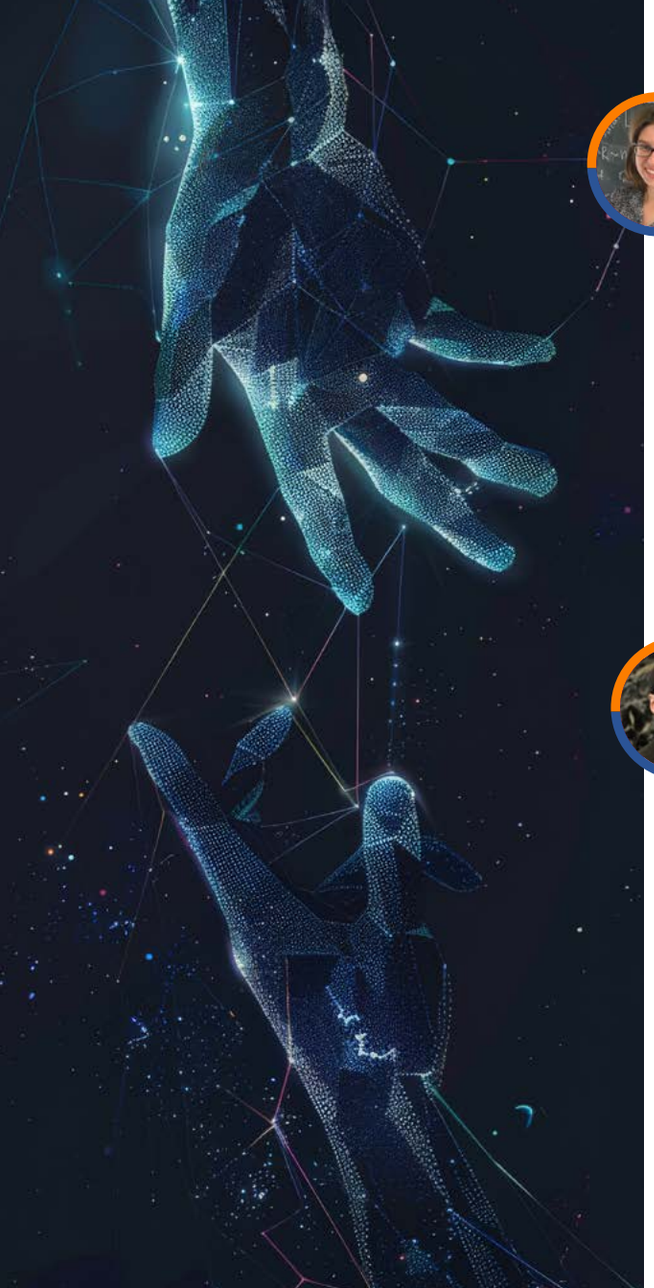
With complete cell models, we can better understand processes such as the way insulin production functions in diabetes or how neurons secrete mood-affecting hormones. The USC Cell Modeling Initiative (CMI) builds on the lessons learned from creating the Protein Data Bank to map cellular architectures and functions. Using AI and virtual reality, CMI aims to accelerate our understanding of disease mechanisms through a paradigm shift in visualization.

The initiative is co-led by biochemist **Kate White** and **Helen Berman**, as well as USC computer scientist **Carl Kesselman** and USC production designer **Alex McDowell**. Together, they create immersive experiences that enable scientists to explore the inner workings of cells in ways that were previously impossible.

PROMPTING CREATIVITY

Simulating events across time. Weaving together scattered clues. Creating room for serendipity. AI sparks human ingenuity, reshaping the questions researchers can ask and powering the tools they create to answer them.



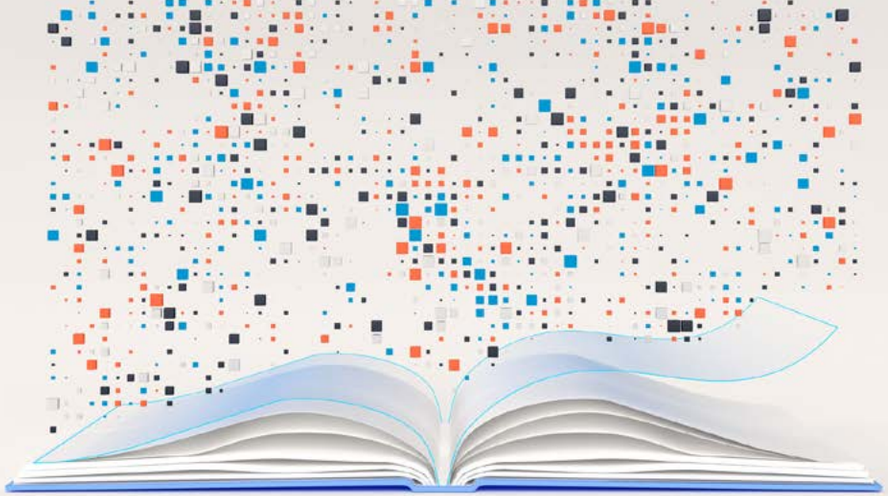


CODE OF THE COSMOS

What makes up the universe? How did it begin? Why does it look like it does? Cosmologist **Vera Gluscevic** helps us grapple with these enduring mysteries. She and her collaborators apply AI to analyze enormous datasets from instruments including the James Webb Space Telescope, the deepest-reaching astronomical observatory to date. Using machine learning to understand light more than 13 billion years old, she creates simulations that model the first galaxies — including enigmatic dark matter thought to comprise most of the cosmic structure today.

ORDER OUT OF CHAOS

Hidden in millions of medical records may be warnings of rare heart conditions. Decades of climate statistics could contain patterns that upend our understanding of global warming. The potential to extract insights from vast, messy data is enormous. Yet, AI still struggles with this challenge. Statistician **Xiaohui Chen**, an expert in high-dimensional statistics and machine learning, develops methods to help machines organize and detect meaningful information faster and more effectively. He also fuses machine learning with quantum information science, a combination that could produce generative AI thousands of times more powerful than today's technology.



AI MAKES HISTORY

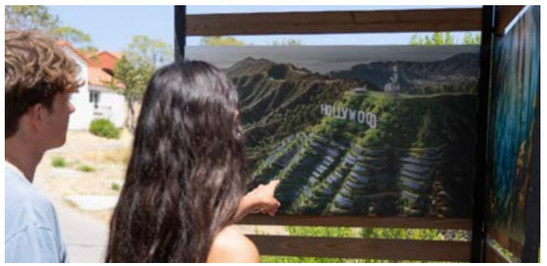
Historians often face the challenge of reconstructing our past from fragmentary evidence and incomplete records. Economist **Thomas Chaney's** research transforms the process by training AI to read and categorize historical information with unprecedented precision. By developing a language model that recognizes key terms in museum collections, his team tracks how techniques such as metallurgy, ceramics, or architecture developed and spread over millennia — stories that have significant, perplexing gaps along the historical timeline. Their approach surfaces subtle patterns connecting isolated data points across millennia, filling in details that would remain hidden.

LOST LANGUAGE MODEL

A grandmother hopes to preserve her nearly-lost childhood language, but only old recordings and scattered texts remain. If her descendents are unable to learn from these materials, the language — and the culture it speaks to — could vanish forever. The drive to protect endangered languages motivates a group of USC Dornsife students, guided by linguist **Khalil Iskarous**. They're developing AI systems that "listen" to audio, "read" rare texts, and predict missing words. By reconstructing grammar and vocabulary, this creative approach could be used to develop new resources for teaching disappearing languages, ensuring they live on.

URBAN REBOOT

Imagine Los Angeles reborn through AI's eyes. USC students, staff, researchers, and artists collaborated at the **Wrigley Marine Science Center** to create "City Ascendant." The exhibit harnesses generative tools like DALL-E 4, ChatGPT, and Midjourney to craft vibrant visions of a sustainable future: solar panels beneath the Hollywood sign, ships turned living reefs, tree-lined streets. Blending scientific prompts with artistic flair, they replace gloomy narratives of climate despair with bold possibilities for building a more sustainable city.



FLYING OVER TREETOPS

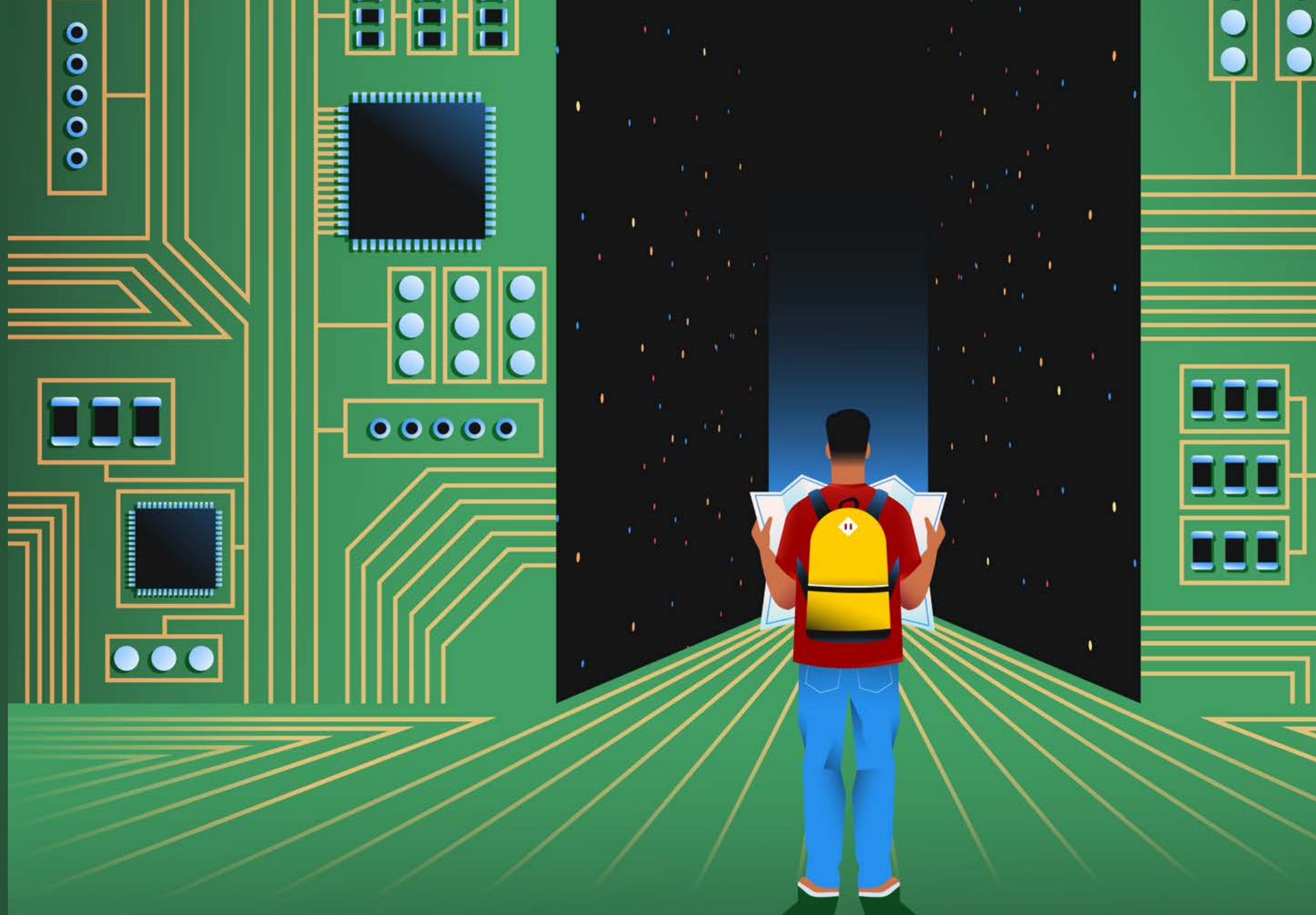
With support from the Bezos Earth Fund, spatial scientists **John P. Wilson** and **Yi Qi** employ AI and deep learning applied to LiDAR and aerial imagery to map every tree's species, location, and condition across Los Angeles County. Their project also uses AI-driven modeling to assess how the existing tree canopy and proposed greening initiatives affect land surface temperatures — a crucial tool for urban heat adaptation and biodiversity planning that could ultimately improve environments and health across our communities.

LANGUAGE OF THE LAW

An officer flags down a driver for a broken taillight. There's the tense option: "Where ya heading, buddy?" Or the option that de-escalates: "Sir, I noticed your taillight is out," explaining calmly. In collaboration with the LAPD, researchers **Morteza Dehghani**, **Benjamin Graham**, and **Brittany Friedman** use AI to analyze police-civilian interactions through archived body camera footage. Their work identifies subtle patterns in communication that influence perceptions of fairness and examines the reasons traffic stops either escalate or remain calm. The project helps train officers to improve civilian interactions and develops open-source AI tools to facilitate transparency, accountability, and learning.

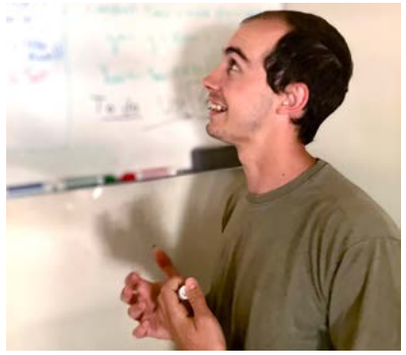
A STARTUP FOR THE SHIFT

Because every discipline and career path will be affected by the AI revolution, USC Dornsife educational programs are constantly evolving to meet the moment and prepare students to succeed in an uncertain future.



TRUST THE FORECAST

We know the concern that ChatGPT can give false answers. Some AI models of the atmosphere do the same, violating basic laws like conservation of mass. PhD student **Obin Sturm** found that while these algorithms struggle with atmospheric chemistry, they improve dramatically when built on established scientific knowledge. By developing physics-informed AI that blends machine learning with core principles, he's making models more accurate and reliable — work that could produce smarter sustainability solutions, from smog prediction to rainforest air quality.



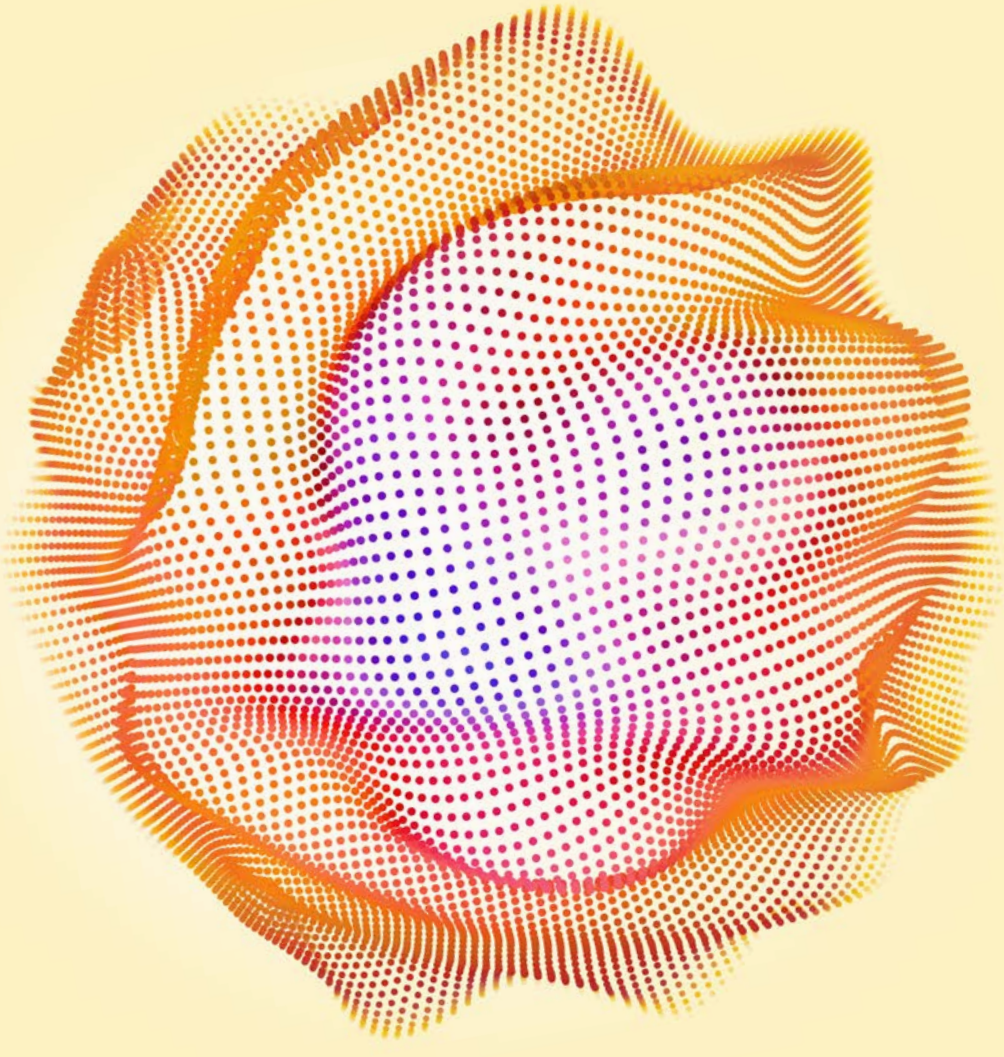
PhD student Obin Sturm creates smarter algorithms for basic science.

UNDERGRADUATE OPPORTUNITIES

At USC Dornsife, undergraduate students can explore AI from all angles. Dozens of courses across the college — from our globally renowned School of Philosophy, our top-ranked Department of Quantitative and Computational Biology or our specialized Undergraduate Neuroscience program — ensure undergraduates gain a broad and rigorous foundation across areas that influence emerging technology. Whether they're questioning the ethics of machine reasoning or completing an internship in Silicon Valley, they'll think critically about how AI reshapes our world.

GRADUATE PROGRAMS

Many of our PhD and master's programs recognize and prioritize the rising demand for AI fluency across every sector. Students work alongside leading experts who redefine their fields and provide advanced training that anticipates the future. Grounded in the depth, versatility, and specialized skills that organizations are relying on to stay ahead of the curve, they'll graduate with a competitive edge.



USC Dornsife

DORNSIFE.USC.EDU