ADVANCING PRECISION HEALTH AND MEDICINE THROUGH

ARTIFICIAL INTELLIGENCE AND BIOMEDICAL DATA

STANFORD DEPARTMENT OF BIOMEDICAL DATA SCIENCE

ANNUAL REPORT 2026

Stanford MEDICINE
School of Medicine

Department of Biomedical Data Science

CONTENTS

"As AI transforms biomedicine, DBDS is helping ensure this era is defined by scientific rigor and collaboration — and never wavers from a focus on improving human health."

LLOYD B. MINOR, MD
Carl and Elizabeth Naumann Dean of the Stanford School of Medicine
Vice President for Medical Affairs, Stanford University

Biomedical data science is at an extraordinary inflection point.

Advances in AI, machine learning, statistics, multimodal data integration, and computational biology are transforming how we analyze biomedical data to understand health and disease at a deep level. DBDS is Stanford Medicine's methodological and collaborative academic hub for biomedical AI and data science. We are shaping a future in which computation on multimodal data drive discovery, strengthen clinical decision-making, and advance precision health and medicine.

Our research mission is defined by innovation in methodology grounded in the use of biomedical data — to solve pressing, real-world challenges. Our faculty and trainees are developing data-driven computational approaches to understand disease, guide care, and improve outcomes.

Increasingly, this work includes advances in multimodal foundation models and agentic AI systems. These newly emergent technologies can integrate diverse patient-specific data, reason through complex and evolving biological and clinical conditions, and support scientific discovery and clinical workflows. At the same time, DBDS is helping ensure that these technologies are developed responsibly — with methods that make them accurate, understandable, reliable, fair, and clinically meaningful.

Especially important to our success is training. DBDS is the academic home to a 40+ year old interdisciplinary training grant, offering MS and PhD degrees in biomedical data science — a field that is more important than ever. Through graduate training, mentorship, interdisciplinary collaboration, and research opportunities, we prepare trainees to move across biomedical fields, develop new methods and systems, and address complex challenges in biology and medicine.

As we look ahead, DBDS remains committed to building and growing a rigorous, collaborative, and responsible scientific community that advances biomedical science, strengthens clinical decision making, and improves human health.

Sylvia K. Plevritis, PhD
William M. Hume Professor in the School of Medicine
Chair, Department of Biomedical Data Science

DBDS BY THE NUMBERS

51 FACULTY

55 PHD STUDENTS

80 MS STUDENTS

16 POSTDOCTORAL FELLOWS

27 RESEARCH STAFF

18 ADMINISTRATIVE STAFF

FACULTY HIGHLIGHTS

Assistant Professor **Roxana Daneshjou** received the Rice 360 Alumni Leadership Award

Assistant Professor **Alex Ioannidis** received funding from the Stanford Discovery Innovation Fund

Assistant Professor **Serena Yeung-Levy** received a 2025 Arc Institute Ignite Award

Associate Professor **Aaron Newman** received an American Association for Cancer Research Trailblazer Cancer Research Award

Associate Professor **Julia Salzman's** NIGMS Maximizing Investigators' Research Award was renewed for another 5 years

Professor **Russ Altman** was appointed a Hoover Institution Fellow (by courtesy)

Professor **Teri E. Klein** and colleagues led the full integration of pharmacogenomics across Stanford Health Care Primary Care

Professor **Mark Musen** was named a Distinguished Fellow of the American College of Medical Informatics

Professor/Chair **Sylvia Plevritis** was elected a 2026 Fellow of the International Society of Computational Biology and received the Faculty Women's Forum Outstanding Leader Award

Professor **Nigam Shah** was inducted into the Association of American Physicians

Professor **Rob Tibshirani** was named the 2025 Myles Hollander Distinguished Lecturer

GRADUATE STUDENTS

15 NIH NLM TRAINEES

6 NATIONAL SCIENCE FOUNDATION SCHOLARSHIPS

5 STANFORD GRADUATE FELLOWSHIPS

5 KNIGHT-HENNESSY SCHOLARS

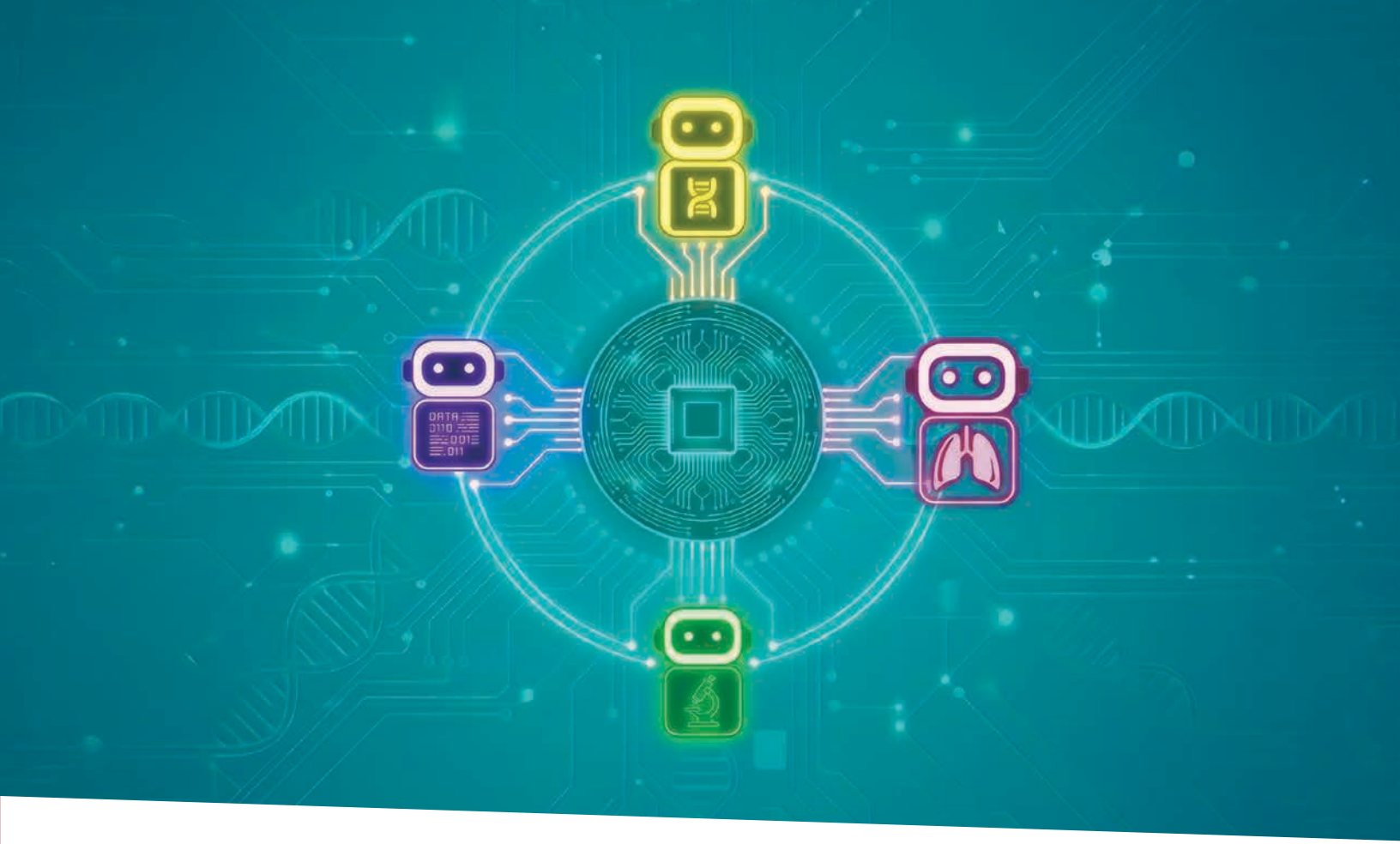
2 STANFORD DATA SCIENCE SCHOLARS

8 WARREN ALPERT SCHOLARS

4 ARC INSTITUTE FELLOWS

1 STANFORD INTERDISCIPLINARY GRADUATE FELLOW

1 MAC3 COMPUTING IN PRECISION HEALTH FELLOW



DBDS is at the Center of AI and Biomedical Discovery

As Stanford Medicine's academic hub for biomedical AI, data-driven discovery, and precision health, DBDS continues to enrich its research footprint — welcoming new faculty and offering expanded training opportunities. DBDS research and training are supported through a mix of federal funding, nonprofit partnerships, and industry collaboration. Below is a quick snapshot of our progress over the past year.

- **Multimodal AI:** Several DBDS faculty developed foundation models that cross all the key clinical modalities — including electronic health records (EHRs), radiology images/reports, pathology slides/reports, and genomic data/reports. These models are improving the ability to predict disease risk, identify treatment response patterns, and help decision making across the continuum of care.
- **AI Agents Speed Clinical Workflows and Scientific Discovery:** DBDS researchers are exploring novel uses of AI agents. Increasingly, AI is evolving from a tool that waits for instructions to an active collaborator that helps researchers and clinicians navigate complex biomedical problems.
- **Inaugural Warren Alpert Computational Biology & AI Symposium:** DBDS hosted a symposium[1] convening researchers and students working at the intersection of AI, computation, and biomedicine. Presentations highlighted a growing shift from AI as a tool for analyzing data to AI systems that can help scientists explore questions, generate ideas, and accelerate discovery.
- **Collaboration & Careers Forum:** The department's fourth annual C&C event[2] brought together students, faculty, clinicians, nonprofits, and industry leaders to learn about emerging opportunities and partnerships in biomedical AI. Attendance rose more than 30% since last year with 234 participants. Company participation grew from 40 to 52 organizations.

Beyond Algorithms: *Training for Real-World Decisions*

Supported for 40+ years by a National Library of Medicine training grant, the DBDS curriculum[3] balances stability with constant adaptation. As AI rapidly transforms medicine and science, DBDS is training students not only to build and use new technologies, but also to evaluate them critically, understand their limitations, and apply them responsibly. Professor Chiara Sabatti, Associate Chair of Education & Training, explains how.

How is the DBDS curriculum designed?
"Our coursework is both foundational and very dynamic. While many of our classes are continually refined to reflect the latest advances in computational methods and biomedical data assets, we also introduce new courses each year in response to emerging areas such as foundation models, generative AI, and new data modalities."

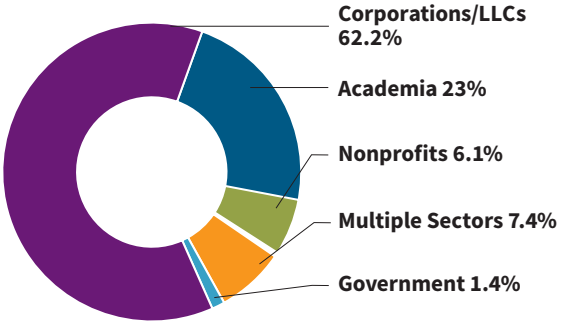
Why is "bilingual" training so important for DBDS students?
"AI can write code, summarize information, and even help analyze data. But it cannot replace scientific judgment. Our students continue to learn fundamental principles that they apply to know when to trust a model, when to question it, and how to interpret findings in the context of human biology, biomedical advances, and better patient care."

How do you keep the curriculum connected as it grows?
"When faculty propose new courses, we work with them to be sure each one fills a meaningful gap. We ask: Who is this for? How does it fit into a pathway? We concentrate on helping students build a coherent foundation that will remain valuable even as AI methods continue to evolve and new computational technologies emerge."

What distinguishes DBDS students from those in programs that combine, rather than integrate, disciplines?
"Our graduates are prepared to lead wherever their curiosity takes them: from developing new methods and advancing biological discovery to improving patient care and shaping the responsible use of AI in society."

"The most valuable skill in the age of AI is knowing how to ask the right questions, evaluate the answers, and translate them into real-world impact."

DBDS CAREERS



Learning Together — *In and Out of the Classroom*

DBDS students are part of a broad, interdisciplinary community that is tightly connected. Small cohorts, shared coursework, and regular touchpoints — from research talks to informal gatherings — build a culture of collaboration and support. The result: scientists who can move fluidly between subject matter domains, bringing technical rigor to meaningful biological and clinical problems.

LEARNING ACROSS FIELDS

Students arrive at DBDS with different backgrounds, but the program quickly expands their abilities. That way, they gain fluency across disciplines while preserving depth in a chosen area. The goal is not to replace specialization, but to strengthen it with a broader set of tools.

EXPANDING OPTIONS

Graduates leave with the ability to work in academia, industry, startups, government, and foundations. Exposure to different paths is built into the program, through career forums, alumni panels, and day-to-day interactions that show how learning translates into real-world impact.

SHARING WORK & PLAY

Students, faculty, and staff work together to build community. Opportunities include monthly student-led events — coffee meetups, dinners, informal gatherings — that bring together students across cohorts and programs. These moments create connections that go beyond labs and often last long after graduation.

LEARNING TO COMMUNICATE

Travel to conferences and weekly seminars enrich science knowledge but also offer students the chance to practice how to communicate their research, which is the lifeblood of a science career in any sector. Over time, these opportunities becomes central to how students develop as scientists.

Students are central to the vitality of our department, our discipline, and our shared future. From organizing events to participating in admissions interviews to recruiting new students at conferences, students play an active role in shaping the community and experiences of those who come after them. DBDS students leave with more than technical expertise. They carry a shared language across disciplines, a network of collaborators, and the ability to move confidently between research, clinical care, and industry to improve human health and well-being.

HELPING AI SPEAK THE LANGUAGE OF MEDICINE

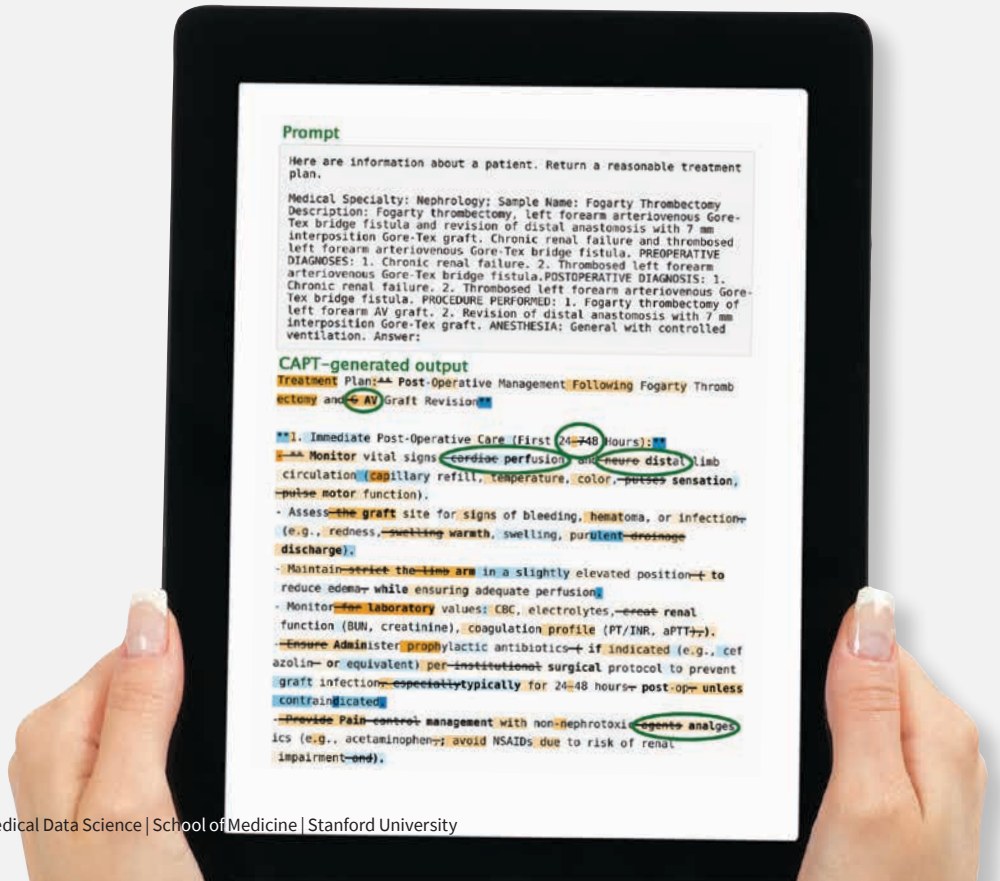
Large language models (LLMs) are widely used for everyday tasks like writing emails, summarizing documents, and answering questions. Their use in medicine is on the rise — but there is good reason for caution. A model that sounds confident but gives incomplete or incorrect information could create risk for patients.

One way to make these tools safer for medicine is to train them with clinical data. But that process is expensive, time-consuming, and difficult to repeat every time a stronger general-purpose AI model is released. The hard choice for healthcare organizations: use the newest AI systems, which may not fully reflect clinical reality, or rely on older medical models that are less technically advanced.

Instead of retraining each new model from scratch, Assistant Professor Emily Alsentzer and her team developed a method[4] to combine newer, more advanced general-domain LLMs with older ones already trained with clinical data. The work is published in *Proceedings of Machine Learning Research.*

LLMs generate their output (text) by predicting the next word in a sentence. Alsentzer's approach, called Cross-Architecture Proxy Tuning (CAPT), works during this process — adding medical context without the need for retraining. At each step, the newer model proposes likely words, while the clinical model helps steer the choice toward more medically specific and clinically useful language. For example, the CAPT system substitutes vague terms like "agents" or "cardiac symptoms" with more precise ones like "analgesic medicines" or "perfusion."

The CAPT approach also avoids "catastrophic forgetting," in which an AI model can lose its knowledge suddenly when it learns something new. Alsentzer's model-merging system was better than either model alone at various clinical tasks, such as identifying important follow-up actions, suggesting treatment plans, and drafting clinical notes.



From Tool to Teammate

Rethinking the Role of AI in Clinical Decision-Making

New research from Associate Professor and Faculty Director for Medical Education in AI Jonathan Chen explores not just whether AI can make good medical decisions, but how it should work with doctors.

Last year, Chen and his team showed that an AI system on its own[5] outperformed doctors (with or without AI help) to make a correct diagnosis from simulated patient cases. Recently, Chen dug deeper to ask a different question: can AI be designed to better support doctors at the moments when they make decisions about patient care?

The researchers first designed a custom AI system that works less like a web search and more like a back-and-forth conversation. It guides both the doctor and the AI system through a case, compares answers, and encourages discussion about differences of opinion. The team tested two ways of using the AI system in a study involving 70 real doctors: as a first or second opinion.

The new results in *NPJ Digital Medicine* suggest that how and when AI is used matters as much as how well it performs.[6] Doctors using the system were more accurate than those using standard, non-AI tools alone.

But timing mattered. AI could guide doctors' thinking when presented as a first opinion or follow their lead when shown second. Doctors were also more likely to engage with the AI system conversationally, using phrases like "Yes, that's a great thought." These patterns reflect basic human tendencies, like relying on first impressions, which could affect how consistent and reliable medical decisions are.

Using Computer Vision to Watch Real-World Medicine

Video recordings of real-world medical scenarios can capture how doctors make decisions on the spot. Watching this footage can yield insight into clinical practice, but doing so by hand is slow and rarely practical for busy hospital settings.

New research from Assistant Professor Serena Yeung-Levy suggests that AI could make video review more practical and useful.

Yeung-Levy and her team analyzed 95 trauma resuscitation videos from a Level I trauma center. The videos recorded events such as patient arrival, acute resuscitation, and bedside processes and procedures like X-rays, ultrasound exams, and putting in intravenous lines. Published in *Annals of Surgery Open*, the researchers report an AI system[7] that breaks up and stitches together these clinically meaningful events into a searchable timeline. Instead of scanning an entire recording, the trauma team could zoom in on specific events.

An expert in computer vision, Yeung-Levy also wondered whether AI could approximate clinical judgment itself. Starting with 319 laparoscopic gallbladder surgery videos collected from hospitals and public datasets, the researchers trained an AI system to judge disease severity[8] by feeding the model evaluations made by surgeons. The results, published in *Surgical Endoscopy*, were mixed. Even though the AI model performed at a level similar to individual doctors, the surgeons themselves agreed with each other only half the time: highlighting how much the scoring system itself relies on individual interpretation.

Although AI systems can help standardize review and make key moments easier to find in long videos, they remain limited by the data they learn from. That means AI will only be useful in clinical care if it is carefully tested, fits into existing workflows, and is measured against a reliable standard.

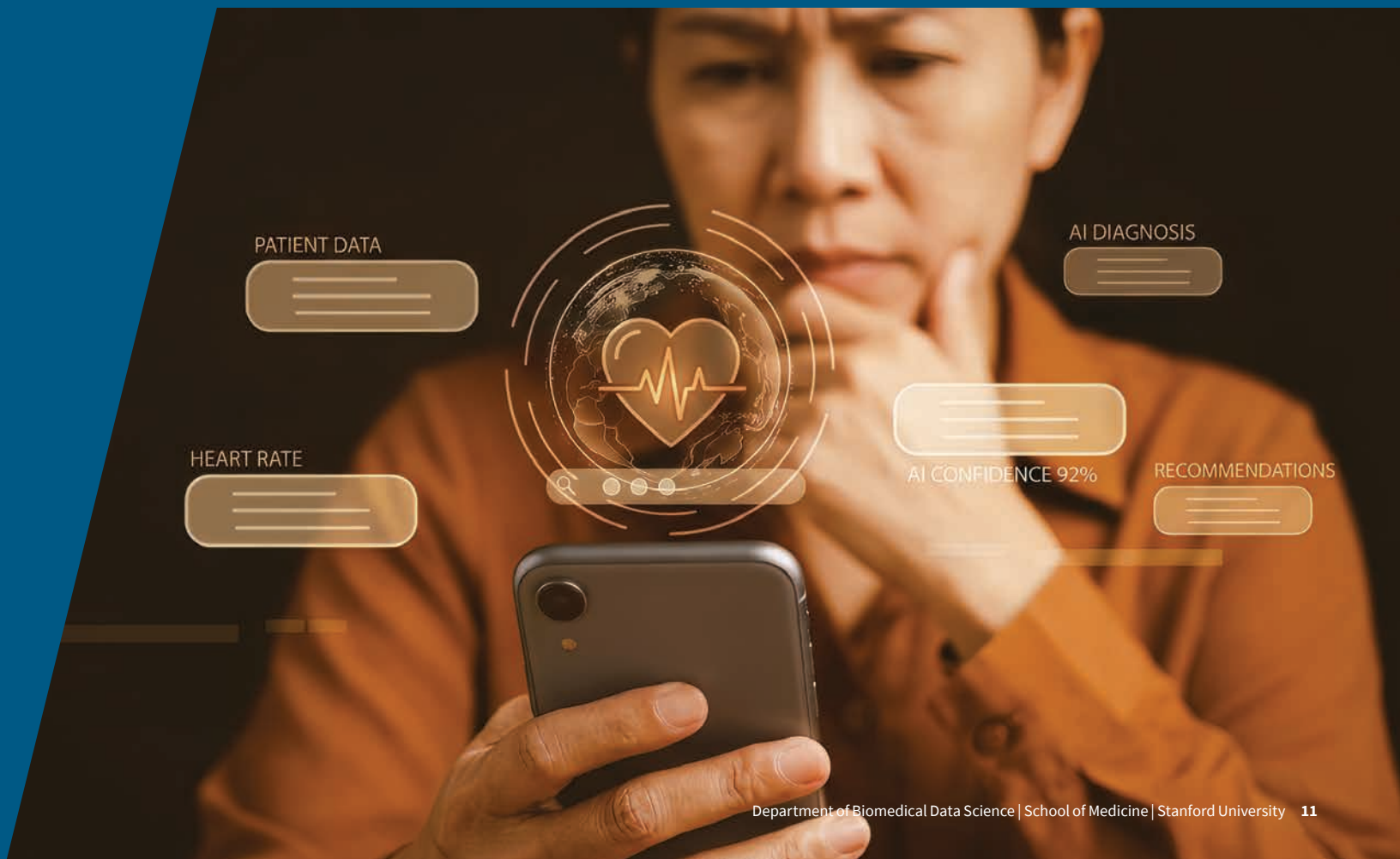
Confidence Without Caution

Since generative AI models emerged in late 2022, AI systems used in chatbots have gotten a lot better at what they do — and for better or for worse many people are using them for health advice. Yet these systems were not originally designed for clinical use, raising an important question: how well do they communicate their shortcomings with human users?

Assistant Professor Roxana Daneshjou and her team looked directly at whether AI models give disclaimers in their answer to a medical question or a requested analysis of a medical image (like a mammogram or an X-ray). An example of a medical disclaimer might be: "I'm not a doctor, but chest pain can have many causes, so you should seek care if symptoms are severe."

Reporting in *NPJ Digital Medicine*, the researchers noticed a clear trend.[9] As models became more advanced (between 2022 and 2025), far fewer medical disclaimers were included in chatbot responses. For text-based medical questions, rates dropped from 26% in 2022 to less than 1% by 2025. The results were almost the same for interpreting medical images. More advanced models were also less likely to include cautionary language, the team learned.

As AI models sound more and more like humans, users may be more likely to trust them as expert guidance, not recognizing their limitations. As these tools increasingly become a part of our everyday lives, being sure they clearly state their limits is essential for keeping people safe and earning public trust.



AI Agents Join the Lab

Over the past decade, single-cell biology has transformed how scientists study gene activity across thousands of individual cells. But single-cell datasets are so large and complex that researchers can usually only sift through a small fraction.

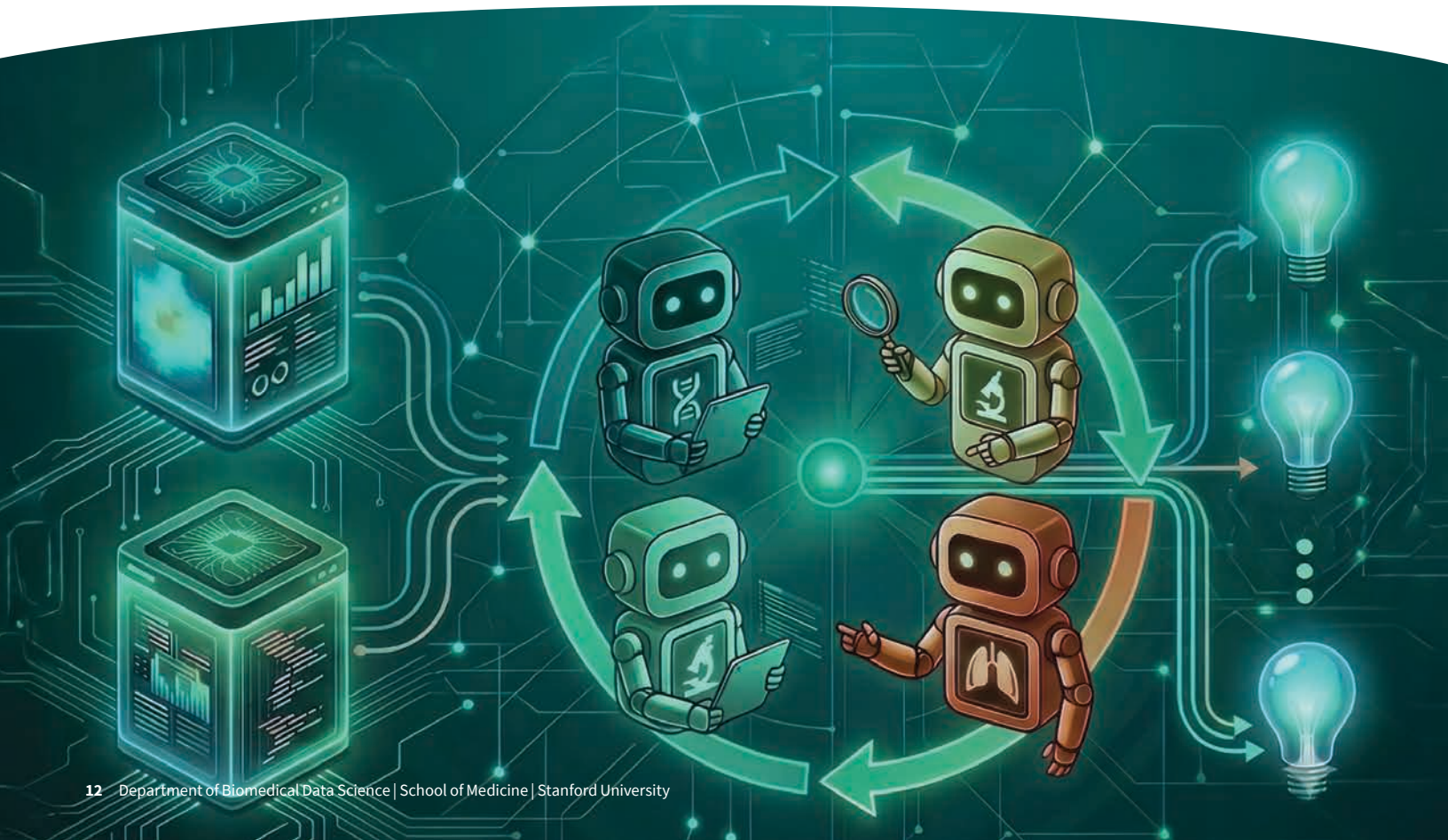
Just a few years ago, datasets like this were underused, but AI is rapidly changing that. Associate Professor James Zou is taking AI to a new level, creating AI agents that can work as part of the research team itself.

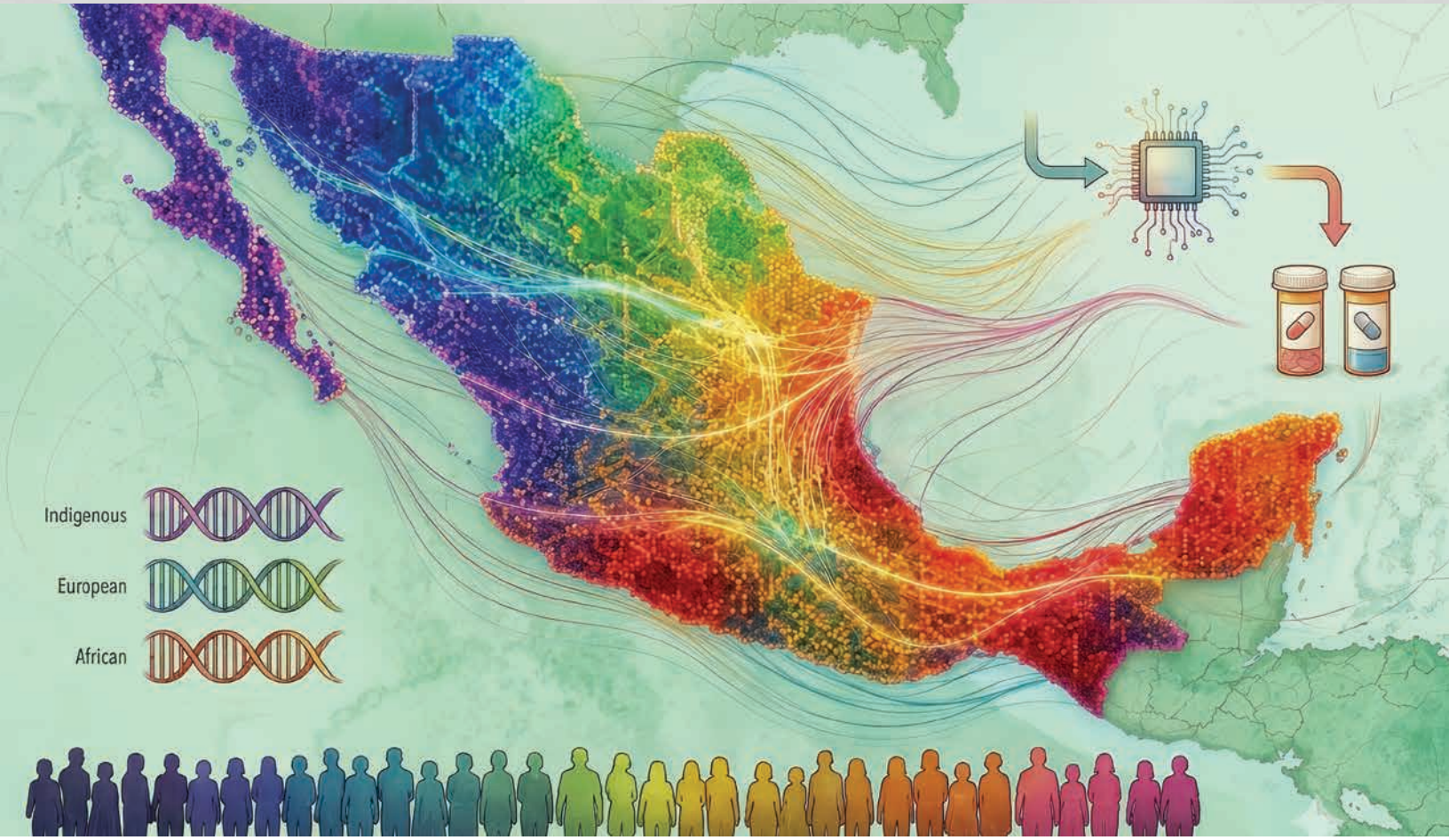
The computational biology agent, called CellVoyager,[10] starts with a text-based AI model (an LLM), then uses a step-by-step process to plan, act, evaluate results, and adjust — over and over. As described in *Nature Methods*, to develop the agent, Zou and his team set up the model to read data, write computer code, run analyses, and view the results. Because the agent can repeatedly test ideas and pursue new leads (and work 24/7), it can search large and complex datasets faster and more thoroughly than is humanly possible.

Zou and his team asked the AI agent to review a published scientific paper along with the data analyzed in that study and then come up with (and test) new scientific questions. The team set a high bar, asking the agent to explore directions not considered by the original researchers.

CellVoyager outperformed leading AI models at predicting which analyses researchers would do. It also made its own discoveries, identifying unexpected immune patterns in COVID-19 and proposing new questions related to aging in the brain.

Because single-cell biology research can generate many possible results (not all meaningful), the team used a human-in-the-loop approach, in which researchers reviewed the agent's work to focus on biologically relevant findings.





When "Average" Isn't Good Enough

For the most part, doctors decide how to treat individual patients based upon research done with large groups of people. These population-level studies generate "average" risk for most health conditions and how people will respond to treatments.

But averages can hide key differences based in genes, because even though people share most of the same genes, the small differences between them can still matter a lot.

In a recent study[11] published in *Nature Medicine*, Assistant Professor Alexander Ioannidis and his team analyzed genetic data from more than 6,000 people across the entire country of Mexico, using advanced data analysis methods to spot hidden patterns. They wanted to better understand these differences, because current medical guidelines often treat Hispanic or Latino populations as a single group.

The scientists learned that genetic differences linked to disease risk and medication response varied a lot among all these people. The results showed clear regional patterns affected by differences in Indigenous, European, and African ancestry — reflecting migration and population history over many generations.

Two striking examples include genetic changes that affect how the body processes medications like opioids used for pain management and statins, drugs used to control high cholesterol. These patterns were common in some regions of Mexico and rare in others — meaning the standard dosing guidelines based on average risk may not apply equally to all people.

Applying what they learned to advance research by others, Ioannidis developed an interactive tool, MexVar. While not yet part of routine clinical care, MexVar allows users to see how genetic variants differ across regions and ancestry groups, which could help guide future testing and treatment.

Tracking How Diseases Change Over Time

Scientists who study common diseases often focus on who is most likely to get the condition. But for people already diagnosed, an equally important question is whether the disease will get worse, and how fast?

In new work[12] posted on *medRxiv*, Assistant Professor Manuel Rivas and his team set out to understand the genetics not only of disease risk, but of disease progression: how conditions like heart disease and diabetes change over time.

The researchers used genetic and health record data from the UK Biobank — a database of health and genetic data from about 500,000 people. They developed a method that combines genetic information with long-term health records to identify genetic factors linked not just to disease risk, but to how fast disease progresses. Rather than asking only who develops a condition, the approach asks how long it takes for people to move from an initial diagnosis to a more serious outcome like heart failure, kidney disease, or surgery.

To do this, Rivas grouped many small genetic effects into a larger, overall readout called a polygenic hazard score (PHS). They learned that a PHS can help identify those people whose disease is more likely to worsen. For example, among people with high cholesterol, some genetic profiles were linked to a much higher chance of needing coronary bypass surgery within years of diagnosis, while others had a lower risk for needing surgery.

Clinical tools like PHS may in time offer a new way for doctors to understand risk more precisely and when to intervene.

HOW GENE NEIGHBORHOODS SHAPE HEALTH

Human DNA is about 99.9% identical from person to person. But the remaining differences — about 4-5 million small changes called genetic variants — are what make people unique. Most studies search for genetic variants that affect one gene at a time.

But biology is a lot more complicated: a single DNA change can affect many nearby genes at the same time. Professor Stephen Montgomery developed a new way to detect these shared effects.

He and his team started with a large public dataset showing gene activity across tissues and looked for genes that tend to turn on or off together in the same region of DNA.

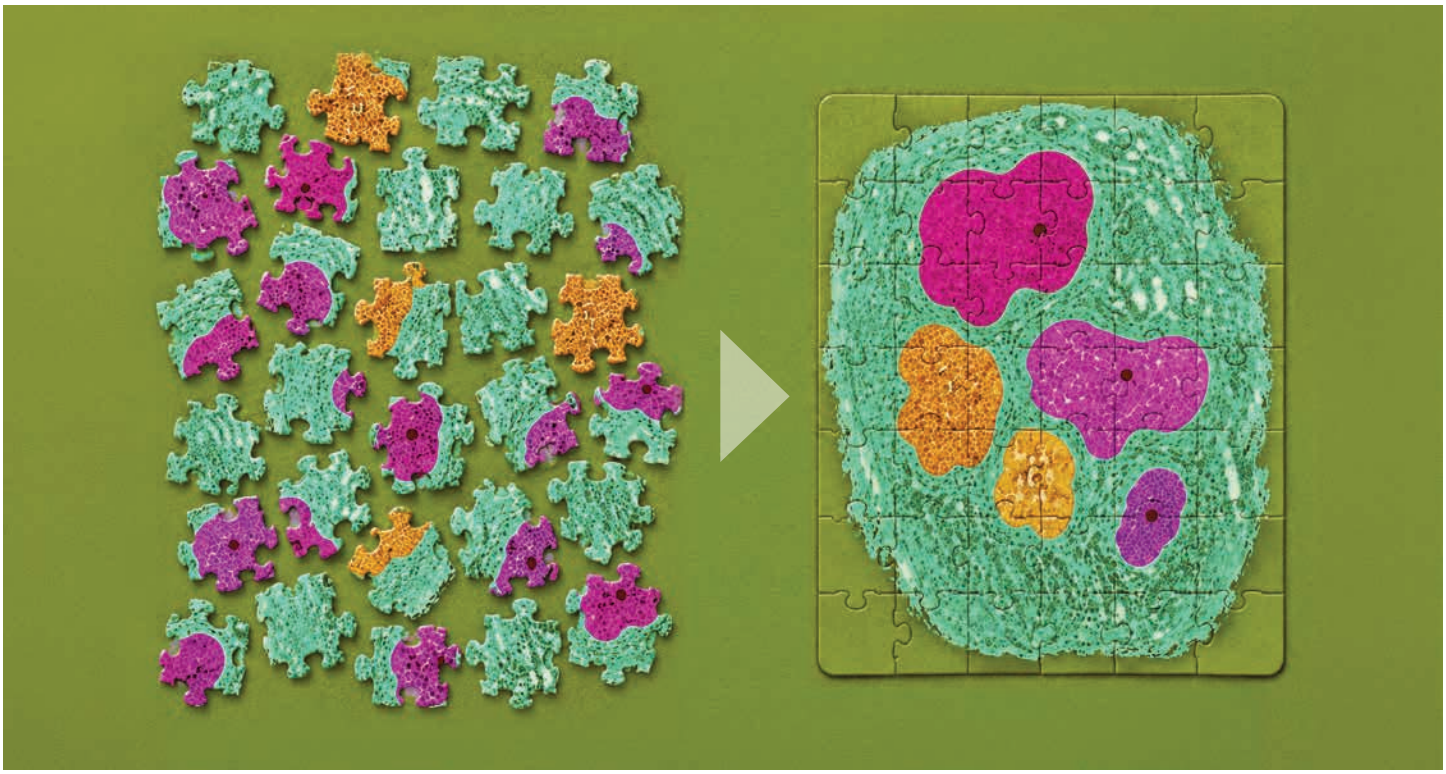
They grouped these "gene neighborhoods" and combined their activity into a single score that reflects how active the group is overall. The researchers then tested whether a single DNA change could change the activity of an entire group of genes up or down, all at once.

Montgomery and his team identified thousands of gene neighborhood effects[13] that standard gene-by-gene analysis methods had missed in the past. The results appear in the *American Journal of Human Genetics.*

Looking for groups of nearby genes that respond together to the same DNA change, the scientists could uncover genetic effects linked to complex traits such as height that standard methods had overlooked. They also found genetic effects linked to conditions such as heart disease and immune disorders, revealing connections that had been missed by traditional gene-by-gene approaches.

By uncovering hard-to-find but influential effects, this work shows how new data and analytical methods can improve how we identify risk and understand disease.

Teaching AI to See
THE BIGGER PICTURE



Today, many pathology slides used to diagnose disease are digitized using high-resolution scanners and viewed on a computer screen. Digitized pathology images have many advantages over traditional glass slides. They can be shared and stored indefinitely and allow side-by-side comparisons. But they contain massive amounts of data — one slide contains hundreds of times more data than even a single smartphone photo. To read them, AI systems divide the image into many small sections and combine the results into an overall interpretation.

But while current AI models are very good at learning about individual cells or small tissue regions, they are not so good at understanding overall layout across millions of cells. Pathologists rely on such patterns to accurately diagnose and treat disease.

To address this shortcoming, Professor Sylvia Plevritis and her postdoctoral fellow Shahira Abousamra and graduate student Asmita Sood developed TopoSlide,[14] which uses AI tools to read digital pathology slides a lot like a human pathologist would. It works by identifying sub-regions of the same tissue pattern — such as those containing tumor, glandular, or immune cells — and then uses topology to learn how they all fit together. That is, do the different regions cluster together, stay apart, or form boundaries around other areas?

TopoSlide was tested on several types of tissue slides from people with cancer. It worked as well or better than top foundation models to identify tissue patterns and gene mutations, as well as to predict patient outcomes. But it was much more efficient: using far fewer images to train the model and not relying on labeling by humans. The work was presented at CVPR 2026, a leading AI/computer vision conference.

TopoSlide is not intended to replace pathologists. Tools like it can help complement doctors by illuminating hard-to-see tissue patterns, pointing out unusual cases, and enabling easier comparisons with similar patients.

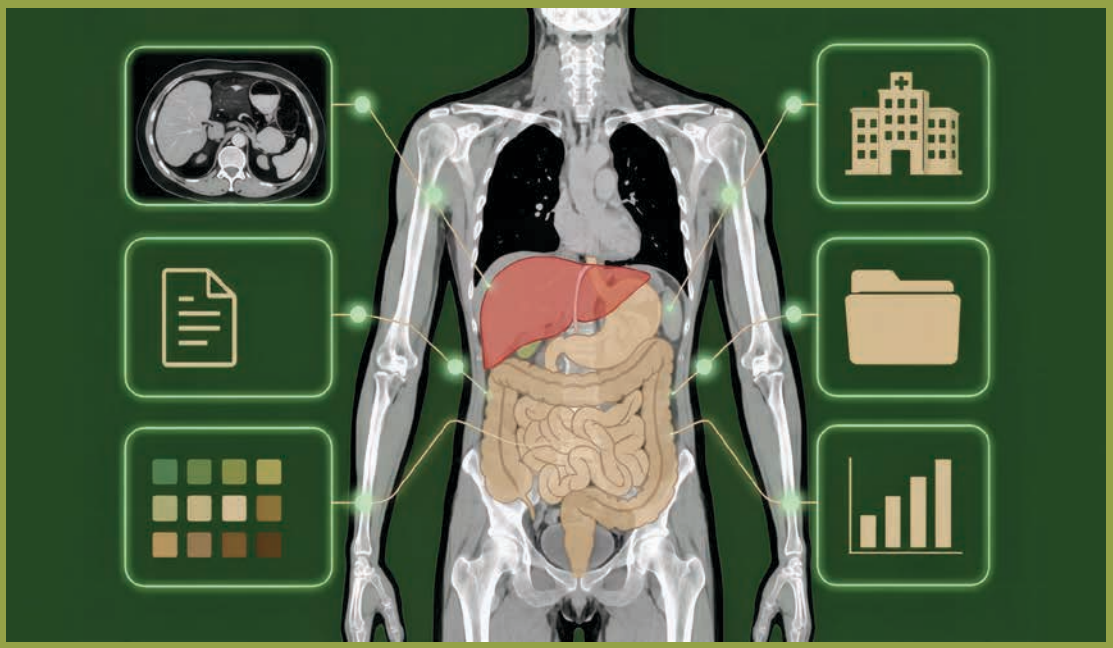
Helping AI Read CT Scans Like a Radiologist

Computed tomography (CT) scans are a core part of medicine — more than 85 million are done each year in the United States. Interpreting these complex 3D images is time-consuming, and a shortage of radiologists means that AI is increasingly being used to help.

Yet most current AI systems analyze CT scans piece by piece, which can miss important details like how large a tumor is. In part, that's because the training of these models relies only on images. A newer type of AI, called a vision-language model (VLM), goes a step further. VLMs link images with the reports doctors write about them, adding clinical context and helping the models handle tasks beyond imaging alone.

Looking to make VLMs even more useful, Associate Professor Akshay Chaudhari and his team developed Merlin,[15] a publicly available 3D VLM designed for abdominal CT scans that analyzes a whole CT all at once, rather than piece by piece. Its training combines images, radiology reports, and diagnosis codes from EHRs, learning from data already captured in routine care rather than requiring experts to label images by hand. The team trained and validated Merlin using more than 55,000 CT scans from multiple hospitals, including millions of images, reports, and diagnosis codes. The work is published in *Nature*.

Instead of building a new AI model for each task, Merlin can be used to identify findings and disease risk, generate draft reports, and even estimate which patients may develop chronic conditions years later. It can also help with practical tasks such as assigning diagnostic codes used in medical records and billing. Across these tasks, Merlin performed as well as or better than specialized models, while using data more efficiently.



PRIMARY FACULTY | DEPARTMENT OF BIOMEDICAL DATA SCIENCE

SYLVIA K. PLEVRITIS, PhD
Professor & Chair

EMILY ALSENTZER, PhD
Assistant Professor

ROXANA DANESHJOU, MD, PhD
Assistant Professor

BARBARA ENGELHARDT, PhD
Professor

JASON FRIES, PhD
Assistant Professor

ALEX IOANNIDIS, PhD, MPhil
Assistant Professor

IAIN JOHNSTONE, PhD
Professor

TERI KLEIN, PhD
Professor & Associate Chair of Faculty Affairs

YING LU, PhD
Professor

AARON NEWMAN, PhD
Associate Professor

JULIA PALACIOS, PhD
Associate Professor

MANUEL RIVAS, DPhil
Assistant Professor

CHIARA SABATTI, PhD
Professor & Associate Chair of Education and Training

JULIA SALZMAN, PhD
Associate Professor

LU TIAN, ScD
Professor & Director of Graduate Studies for Master of Science Programs

ROBERT TIBSHIRANI, PhD
Professor

WING HUNG WONG, PhD
Professor

SERENA YEUNG-LEVY, PhD
Assistant Professor

JAMES ZOU, PhD
Associate Professor

SECONDARY FACULTY | DEPARTMENT OF BIOMEDICAL DATA SCIENCE

NIMA AGHAEEPOUR, PhD
Professor

RUSS B. ALTMAN, MD, PhD
Professor

EUAN A. ASHLEY, MRCP, DPhil
Professor and Associate Dean
of the School of Medicine

AKSHAY CHAUDHARI, PhD
Associate Professor

JONATHAN CHEN, MD, PhD
Associate Professor

MANISHA DESAI, PhD
Professor

OLIVIER GEVAERT, PhD
Associate Professor

TINA HERNANDEZ-BOUSSARD, PhD
Professor

CURTIS LANGLOTZ, MD, PhD
Professor

STEPHEN B. MONTGOMERY, PhD
Professor & Director
of Graduate Admissions

MARK MUSEN, MD, PhD
Professor

NIGAM H. SHAH, MBBS, PhD
Professor

DENNIS WALL, PhD
Professor & Director
of Graduate Studies

JOHN WITTE, PhD
Professor

EMERITUS FACULTY | DEPARTMENT OF BIOMEDICAL DATA SCIENCE

DANIEL BLOCH, PhD

TREVOR HASTIE, PhD

PHILIP W. LAVORI, PhD

DANIEL RUBIN, MD

COURTESY FACULTY

DEPARTMENT OF BIOMEDICAL DATA SCIENCE

EHSAN ADELI, PhD
Assistant Professor

ANDREW GENTLES, PhD
Associate Professor

ZIHUAI HE, PhD
Assistant Professor

JOHN P.A. IOANNIDIS, MD, DSc
Professor

DAVID KIM, MD, PhD
Assistant Professor

STANLEY QI, PhD
Associate Professor

MIRABELA RUSU, PhD
Assistant Professor

ADJUNCT FACULTY

DEPARTMENT OF BIOMEDICAL DATA SCIENCE

FRANCISCO M. DE LA VEGA, DSc
Adjunct Professor

MOHIT KAUSHAL, MD
Adjunct Professor

MATT LUNGREN, MD
Adjunct Professor

VIJAY PANDURANGAN, MSc
Adjunct Professor

MEI-CHIUNG SHIH, PhD
Adjunct Professor

TANVEER SYEDA-MAHMOOD, PhD
Adjunct Professor

Department of Biomedical Data Science, 2025 DBDS Annual Retreat

Help Us Shape the Future of Precision Health and Medicine

The DBDS vision of precision health and medicine integrates data with powerful computational tools to better predict, prevent, and treat disease — delivering the right care to the right person at the right time. Today, AI is accelerating the path to that vision in ways that were hard to imagine only a few years ago, with an unprecedented ability to harness multimodal data across clinical, molecular, imaging, and biological domains.

Realizing AI's potential for healthcare requires careful attention to how these systems are built, trained, and used. Close collaborations between research and medical communities are essential — especially at this pivotal moment when AI models are shaping clinical decision-making and biomedical discovery. For that reason, biomedical data science research is essential for identifying limitations and risk — and for developing rigorous, explainable, and transparent solutions.

Still more complex, and equally exciting, is the emergence of AI agents. These systems can collect information, reason through problems, and perform complex, multi-step tasks. DBDS researchers are helping lead the development of a range of new AI technologies, including AI agents. They are also doing the work needed to ensure these systems are safe, accurate, trustworthy, and clinically meaningful.

We need your help to continue this important work in an increasingly complex biomedical ecosystem. Philanthropic support allows us to recruit exceptional faculty, train the next generation of biomedical data science leaders, and expand access to advanced computing resources needed to drive discovery to improve human health and well-being.

For more information, contact Director of Finance and Administration Ruth Torres (ret2@stanford.edu) or scan the QR code below.

ENDNOTES

1. https://dbds.stanford.edu/next-frontier-in-medicine-ai-that-learns-adapts-and-helps-develop-medical-cures/
2. https://dbds.stanford.edu/reimagining-healthcare-through-ai/
3. Yeh CY, Wall DP, Matthys K, Sabatti C, Palacios JA. Curriculum Design in an Evolving Field: Perspectives on Biomedical Data Science from Stanford. Annu Rev Biomed Data Sci. 2025 Aug;8(1):341-354. PMID: 40203230.
4. Sasha Ronaghi, Chloe Stanwyck, Asad Aali, Amir Ronaghi, Miguel Angel Fuentes Hernandez, Tina Hernandez-Boussard, Emily Alsentzer; Proceedings of the 7th Conference on Health, Inference, and Learning, PMLR 333:354-388
5. Goh E, Gallo R, Hom J, Strong E, Weng Y, Kerman H, Cool JA, Kanjee Z, Parsons AS, Ahuja N, Horvitz E, Yang D, Milstein A, Olson APJ, Rodman A, Chen JH. Large Language Model Influence on Diagnostic Reasoning: A Randomized Clinical Trial. JAMA Netw Open. 2024 Oct 1;7(10):e2440969. PMID: 39466245.
6. Everett SS, Bunning BJ, Jain P, Lopez I, Agarwal A, Desai M, Gallo R, Goh E, Kadiyala VB, Kanjee Z, Koshy JM, Olson A, Rodman A, Schulman K, Strong E, Chen JH, Horvitz E. From tool to teammate in a randomized controlled trial of clinician-AI collaborative workflows for diagnosis. NPJ Digit Med. 2026 Mar 18. PMID: 41851268.
7. Villarreal JA, Heo J, Wang X, Bain A, Succar B, Yao DH, Jopling JK, Yeung-Levy S, Dumas RP. Enhanced Trauma Video Review With Computer Vision: Trauma Resuscitation Phase Segmentation and Procedure Detection. Ann Surg Open. 2025 Dec 1;6(4):e631. PMID: 41451193.
8. Olsen GH, Goodman ED, Aklilu JG, Bartoletti S, Hung KS, Yang JH, Sorenson EC, Jopling JK, Yeung SY, Azagury DE. Using artificial intelligence to model expert panel diagnosis of cholecystitis severity. Surg Endosc. 2025 Oct;39(10):6560-6568. PMID: 40825893.
9. Sharma S, Alaa AM, Daneshjou R. A longitudinal analysis of declining medical safety messaging in generative AI models. NPJ Digit Med. 2025 Oct 2;8(1):592. PMID: 41038984.
10. Alber S, Chen B, Sun E, Isakova A, Wilk AJ, Zou J. CellVoyager: AI CompBio agent generates new insights by autonomously analyzing biological data. Nat Methods. 2026 Apr;23(4):749-759. PMID: 41845065.
11. Barberena-Jonas C, Medina-Muñoz SG, Cedillo-Castelán V, Sepúlveda-Morales T, Gonzaga-Jáuregui C; ENSA Genomics Consortium; García-García L, Ioannidis AG, Moreno-Estrada A. Clinical genetic variation across Hispanic populations in the Mexican Biobank. Nat Med. 2026 Feb;32(2):725-735. PMID: 41566040.
12. Justesen JM, Venkataraman G, Tanigawa Y, Li R, Hastie T, Tibshirani R, Knowles JW, Rivas MA. Genetics of cardiometabolic disease progression. medRxiv [Preprint]. 2025 Feb 3:2025.02.01.25321518. PMID: 39974115.
13. Lawrence KA, Gjorgjieva T, Nachun D, Montgomery SB. Focus on single-gene effects limits discovery and interpretation of complex-trait-associated variants. Am J Hum Genet. 2026 Apr 2;113(4):842-851. PMID: 41875896.
14. Abousamra S, Sood A, Plevritis S. TopoSlide: topologically-informed histopathology whole slide image representation learning. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR);* 2026. p. 13691–13701.
15. Blankemeier L, Kumar A, Cohen JP, Liu J, Liu L, Van Veen D, Gardezi SJS, Yu H, Paschali M, Chen Z, Delbrouck JB, Reis E, Holland R, Truyts C, Bluethgen C, Wu Y, Lian L, Jensen MEK, Ostmeier S, Varma M, Valanarasu JMJ, Fang Z, Huo Z, Nabulsi Z, Ardila D, Weng WH, Junior EA, Ahuja N, Fries J, Shah NH, Zaharchuk G, Willis M, Yala A, Johnston A, Boutin RD, Wentland A, Langlotz CP, Hom J, Gatidis S, Chaudhari AS. Merlin: a computed tomography vision-language foundation model and dataset. Nature. 2026 Apr;652(8112):1318-1328. PMID: 41781626.

Editorial: Alison F. Davis, PhD
Design: Alexander Atkins Design, Inc.

Department of
Biomedical Data Science

Stanford University School of Medicine
Department of Biomedical Data Science
Edwards Building
300 Pasteur Drive
Room R348
Stanford, CA 94305-5464

dbds.stanford.edu