Computer Science and Artificial Intelligence Laboratory

The Computer Science and Artificial Intelligence Laboratory (CSAIL) is focused on developing the architectures and infrastructures of tomorrow's information technology, and on creating innovations that will yield long-term improvements in how people live and work. Lab members conduct research in almost all aspects of computer science, including artificial intelligence, the theory of computation, systems, machine learning, computer graphics, and the explorations of revolutionary new computational methods for advancing healthcare, manufacturing, energy, and human productivity.

CSAIL researchers focus on finding innovative ways to make systems and machines operate faster, better, safer, easier, and more efficiently for the benefit of humanity. Our projects fall into three areas of inquiry:

- Artificial intelligence (AI): We seek to understand and develop both living and artificial systems capable of intelligent reasoning, perception, and behavior.
- Systems: We seek to discover new principles, models, metrics, and tools of both hardware and software-based computer systems.
- Theory: We seek to understand the mathematics of computation and its wide-ranging, real-world consequences.

CSAIL has a long history of technological innovations that have affected how people interact and do business. CSAIL is known as the incubator for some of the greatest technological advances of the past 30 years that were true life-changers, including the Internet, personal computing, mobile computing, open-source software, microprocessors, robotic surgery, and social networking.

CSAIL's current research addresses some of the grand challenges of the 21st century, including developing personalized learning, securing cyberspace, advancing health informatics, reverse engineering the brain, enhancing virtual reality, developing tools for scientific discovery, improving urban infrastructure, and ensuring the health of our environment. Computing is central to solving these challenges and CSAIL contributes to making computing more capable by addressing fundamental algorithmic and systems questions at the core of computing, and broadening the scope of computing to address the important social challenges that confront us. Key CSAIL initiatives currently underway include tackling the challenges of big data, developing new models for wireless and mobile systems, securing computers and the cloud against cyber attacks, rethinking the field of artificial intelligence, and developing the next generation of robots. Advanced software-based medical instrumentation and medical informatics systems to aid clinical decision making is being investigated. Advancements in biological research are also under way, including developments in the field of computational biology and the application of machine learning to the interpretation of complete genomes and understanding gene regulation.

CSAIL research is sponsored by a large number of diverse sources, from US government contracts to the private sector. US government sponsors include: the Air Force Research

Laboratory and the Air Force Office of Scientific Research; the Army Research Office; the Defense Advanced Research Project Agency; Department of Defense Research and Engineering; the Food and Drug Administration; the US Department of Education; the Department of Energy; the Intelligence Advanced Research Projects Activity; the National Institutes of Health; the National Institute of Justice; the National Science Foundation; the Navy (including the Office of Naval Research, and Naval Air Systems Command); and the Space and Naval Warfare Systems Center. US and international nonfederal sponsors include: Accenture, LLP; Boeing; BMW of North America, LLC; Ford Motor Co.; Foxconn Technology Group; Intel Corporation; Jaguar Land Rover Limited; Lockheed Martin Advanced Technology Laboratories; Microelectronics Advanced Research Corporation; Mitsubishi Electric Corporation; Nissan Motor Company, Ltd.; Nippon Electric Company; Nippon Telegraph and Telephone Corporation; Northrop Grumman Corporation; Ping An Technology; Qatar Computing Research Institute; Quanta Computer, Inc.; Samsung Electronics; Siemens; Toyota Research Institute; and Wistron Corporation. Other organizations sponsoring research include: Aarhus University; Battelle Memorial Institute; Delta Electronics Foundation; DSO National Laboratories; Epoch Foundation; Hong Kong University of Science and Technology; IBM; Industrial Technology Research Institute; Nanyang Technical University; Pfizer Inc.; and the Singapore-MIT Alliance.

Research Projects

Within CSAIL we have many single- and multi-investigator projects, as well as a number of virtual centers and large-scale projects. The large-scale projects and collaborations include the following:

Toyota-CSAIL Joint Research Center

Toyota established a collaborative research center with CSAIL and Stanford in 2015 to further the development of autonomous vehicle technologies, with the goal of reducing traffic casualties and potentially developing a vehicle incapable of causing a vehicular accident.

Today, a car crash occurs on average every five seconds in the United States. Globally, road traffic injuries are the eighth leading cause of death, with about 1.24 million lives lost every year. In addition to this terrible human cost, these crashes take an enormous economic toll. The National Highway Traffic Safety Administration has calculated the economic cost in the United States at about $277 billion per year. Putting a dent in these numbers is an enormous challenge—and it's one that is motivating the research of the Toyota-CSAIL Joint Research Center, which was kicked off in September 2015. This new center works in collaboration with the newly formed Toyota Research Institute led by Dr. Gill Pratt.

Imagine if your car could tell you were having a bad day, and turned on your favorite album to improve your mood. What if your car could talk to your refrigerator, figure out that you're low on milk, and suggest where to stop on your way home? Or if your car knew that you forgot to call your parents yesterday and issued a gentle reminder on the way home? And making that call was easy because you could turn the driving over to the car on a boring stretch of highway. These are just a few of the possibilities when

we bring together cars and computer science, and they are motivating the research at the Toyota-CSAIL Joint Research Center.

The objective of the Toyota-CSAIL program is to advance AI and Robotics research, develop a safe and intelligent car, and improve mobility and transportation by advancing the science of autonomy and machine intelligence. The CSAIL researchers are working on (1) new tools for collecting and analyzing navigation data with the objective to learn from humans; (2) perception and decision-making systems for safe navigation; (3) systems that can handle difficult driving situations such as congestion, high-speed driving, and inclement weather; (4) predictive models that can anticipate the behavior of humans and vehicles; and (5) intelligent user interfaces.

More specifically the projects and principal investigators (PIs) that are currently active in the Toyota-CSAIL Joint Research Center are:

- Geordi: A Driver's Assistant for Risk-Bounded Maneuvering (PI: Brian Williams)
- Driver-Friendly Bilateral Control for Suppressing Traffic Instabilities (PI: Berthold Horn)
- Using Vision and Language to Read Minds (PI: Nickolas Roy; Co-PI: Boris Katz)
- Uhura: A Driver's Personal Coach for Managing Risk (PI: Brian Williams)
- Predicting a Driver's State-of-Mind (PI: Antonio Torralba; Co-PI: Wojciech Matusik)
- Exploring the World of High Definition Touch (PI: Ted Adelson; Co-PI: John Leonard)
- Formal Verification Meets Big Data Intelligence to Address the Trillion Miles Challenge (PI: Armando Solar-Lezama)
- The Car Can Explain! (PI: Gerald Sussman; Co-PIs: Daniel Weitzner, Hal Abelson, and Lalana Kagal)
- Crossing the Vision-Language Boundary for Contextual Human-Vehicle Interaction (PI: Jim Glass; Co-PI: Antonio Torralba)
- Analysis by Synthesis Revisited: Visual Scene Understanding by Integrating Probabilistic Programs and Deep Learning (PI: Joshua Tenenbaum)
- Wi-Fi-Based Obstacle Detection for Robot Navigation (PI: Dina Katabi; Co-PI: Daniela Rus)
- Drinking from the Visual Firehose: High-Frame-Rate, High-Resolution Computer Vision for Autonomous and Assisted Driving (PI: Saman Amarasinghe; Co-PIs: John Leonard, and Fredo Durand)
- Decision Making for Parallel Autonomy in Clutter (PI: Daniela Rus; Co-PI: Sertac Karaman)
- A Parallel Autonomous Driving System (PI: John Leonard; Co-PIs: Sertac Karaman and Daniela Rus)

- Uncovering the Pain Points in Driving (PI: Ruth Rosenholtz; Co-PIs: Fredo Durand, William Freeman, Aude Oliva, and Antonio Torralba)
- Simulation and Verification for Vision-in-the-Loop Control (PI: Fredo Durand)
- Tools and Data to Revolutionize Driving (PI: John Leonard; Co-PI: Daniela Rus)

Wistron-CSAIL Research Collaboration

Good health—both mental and physical—is one of the most pressing social and economic issues of the day. A healthier population makes for a happier society and a more productive economy. Today, people are surrounded by an explosion of sophisticated and increasingly affordable information devices, from laptop computers, e-book readers, and smart glasses to mobile phones, smart watches, and health trackers. We monitor stock prices, weather forecasts, and traffic patterns through websites and apps, share our thoughts and experiences through emails, Facebook, and Twitter, and increasingly learn within online communities. These technologies open so many new opportunities for improving how we live, work, and play. But how do they empower us and at what costs? Recent studies show that we consume 11 to 14 hours of technology each day. And this often involves multitasking, which in turn retrains our brains, reduces concentration, and increases stress (e.g., studies show that the brains of heavy technology users show similar patterns to those who suffer from substance addiction). Finding ways to reduce stress and technology's negative impact on a workforce is critical to our future well-being.

The multiyear research program between Wistron and CSAIL focuses on rethinking how we compute and communicate in the digital age to ensure that: (1) health and well-being are at the core of our lives; and (2) our use of technology accelerates this objective. Some of the questions we pose and the answers we seek include: How to design the next generation of computers and communication systems to minimize our body's exposure to electromagnetic radiation? How should we rethink computer and communication architectures for sustainability? How to develop systems that deliver appropriate lighting? How to develop systems that reengineer email? How to develop algorithms that can help with information overload? How to use computing and communication in support of individual and community well-being? How to build computer and communication systems that are friendlier to our environment?

Our vision is to develop new computing and communication hardware and software platforms and supporting algorithms for modeling, controlling, and making decisions that will bring wellness to our use of technology. One thrust of this program focuses broadly on the computer and communication platforms. The second thrust focuses on using these novel platforms to promote healthier living. More specifically, the four projects that are currently active in the Wistron-CSAIL Research Collaboration are:

- Individual Prediction and Interpretation of Risk: Predicting Trajectories of Chronic Disease and Recovery (PIs: Polina Golland and Peter Szolovits)
- Interpretable Predictive Models from Machine Learning (PI: Cynthia Rudin)

- Personally Authored Wellness Applications (PIs: David Karger and Daniel Jackson)
- Smart Homes that Monitor Breathing, Heart Rate, and Life Quality (PI: Dina Katabi)

Ping An Research Collaboration

Ping An is a leading integrated financial services firm and supports CSAIL research in natural language processing, speech recognition, and big data analytics.

Ping An Insurance (Group) Company of China, Ltd., was established in 1988 in Shekou, Shenzhen. The company began in insurance but today has developed into a personal financial services group with three core businesses: insurance, banking and investment, and Internet financing. Ping An strives to become a world-leading personal financial services provider and employs approximately 798,000 life insurance sales agents and 246,000 full-time employees. Both Ping An Life and Ping An Property and Casualty rank as the second largest insurance companies in China by premium income in their respective sectors.

Many organizations today struggle to make intelligent use of all the data they have collected and how to structure their systems most efficiently. In addition, companies dealing with their customer base are continually looking for ways to not only automate but improve the customer experience.

To help achieve their goal at excelling in the financial services space, Ping An sought to partner with CSAIL to address certain technical challenges faced by its business. The Ping An research collaborative began in 2014. Three research projects were identified.

Semantic Summarization for Financial Data (Professor Daniela Rus): A problem is proposed whereby there is a large set of database tables that contain data regarding customers, but tools do not yet exist for the required number of fields and records. To address this problem, we will build on our experience with summarizing data from mobile sensors, such as smart phones, which are playing an increasingly important role in our lives, and can be the source of very large and useful data about the users carrying them. Our goal is to develop systems and algorithms that take large data streams and convert them into semantic summaries.

Speech Recognition (Professor James Glass): This project focuses on Chinese speech recognition methods for customer support. A speech recognition capability would be useful to process the millions of existing Ping An human to human, customer to agent telephone-based call-center communications. There are also potential scenarios for future human to computer spoken interactions via mobile devices as well. There is tremendous potential for speech technology benefits via spoken interfaces but in the scope of this three-year project, we will focus on current spoken language systems and automatic speech recognition research investigating the use of deep, neural network-based methods for multilingual speech recognition.

Speaker Verification (Professor James Glass): Speaker verification is the task of processing a speech recording and deciding whether or not it belongs to a putative speaker. For customer telephony applications, speaker verification is useful as a complementary verification method to existing approaches such as caller ID, PIN numbers, and so on. Speaker verification is closely related to the problem of speaker identification (or recognition) which seeks to determine the identity of a recording from a large candidate pool. There are several potential uses for speaker verification technology for call-center customer support. The majority of these capabilities are directly relevant to the millions of existing human to human, customer to agent telephone-based call-center communications, though there are potential uses for future human to computer spoken interactions as well. Speaker verification technology complements existing methods for verifying customer identity. This project will focus on speaker divarication to automatically separate customer and agent speech turns during a single recorded dialogue.

Quanta/Qmulus Project

In 2005, CSAIL started a 10-year, $45.5 million joint research collaboration called T-Party with Quanta Computer, Inc., of Taiwan. The goal of the first five years was to: (1) develop the next generation of platforms for computing and communication beyond personal computers; (2) create new systems for the development and seamless delivery of information services in a world of smart devices and sensors; and (3) move from a device-centric perspective to a human-centric one. Some of the projects pursued by 15 PIs included:

- Virtualized Computation Platform: We focused on the development of secure and reliable computation and storage.
- T-Net: We developed a direct, secure, authorized, and authenticated access to personal mobile devices.
- Just Play: We designed distributed systems automatically constructed from collections of disaggregated devices.
- Natural Interactions: We focused on the use of human language as a central ingredient in a multimodal interface (combining speech and gesture) for native users and in real-world environments where a traditional graphical user interface is not practical.

The project was renamed Qmulus in 2010 to reflect the shift in our research emphasis from mobile computing to cloud computing. In the second five years, we focused on four broad areas of research.

- Cloud Technologies: The emphasis of this line of research is in hardware and software security. We built a secure processor that can be used to maintain privacy in big data applications. We also developed a prototype database machine called BlueDBM. In software, our research focuses on three aspects: cloud security, multicore, and cloud infrastructure. Finally, we focused on a number of database-related applications.
- Multimedia: We developed and improved human language technologies for natural human–computer interactions and scientific discovery, focusing on

health and wellness-related applications. We demonstrated low-power chip implementation for speech recognition and for voice activity detection. Finally, we improved our multiview autostereoscopic systems capable of demonstrating 3D content viewing without the need for users to wear glasses.

- Healthcare: We developed methods of using computational photography and motion magnification for medical applications. In cloud-based medicine and healthcare, we developed new tools for visualizing large data sets to support interactive knowledge discovery, and analyzed large datasets for risk stratification.
- Education Technology: We developed Pentimento, a flexible content creation and viewing system. We used crowd computing techniques for education. We demonstrated through crowd-sourcing techniques that linking various courseware—lecture, slides, and textbooks, can lead to better learning in massive open online course (MOOC) platforms. Finally, we developed tools to implement the "see one, do one, teach one" pedagogy for online teaching.

To commemorate this fruitful collaboration, we created a set of videos highlighting our accomplishments.

Qatar Computing Research Institute

In 2012, CSAIL signed a seven-year, $35 million research agreement with the Qatar Computing Research Institute to collaborate on a wide-range of research topics in computer science. Currently, the collaboration includes the following eight projects:

- Arabic Speech and Language Processing: This project aims to develop advanced speech and language processing technologies that will support natural interaction via spoken language.
- Content-Adaptive Video Retargeting: This project aims to develop a complete system for delivering high-quality stereoscopic broadcast video. We focus on the real-time video of sporting events, soccer in particular.
- Crosscloud: This project addresses an important and fundamental question related to a decentralized architecture for social networks: How can users get the full benefits of social software even when their friends and colleagues use different vendors' software and different social network providers?
- Database Management: The project began by extending our data integration framework, starting with interactive data curation. It has since expanded to include various aspects of data management, including elasticity in transactional database management systems and graph processing.
- Understanding Health Habits from Social Media Pictures: The goal of the project is to understand the food habits from social media images, thus potentially leading to prediction of population-level health statistics such as obesity, diabetes, and so on.
- Understanding and Developing for Cultural Identities across Platforms: This project aims to develop computational (AI) approaches to model and study identity, self-representation, and related phenomena. This will lead to the

development of value-driven design principles and best practices for virtual identities within a Qatari context.

- A Vertically-Integrated Approach to Resource-Efficient Shared Computing: The goal of this proposal is to investigate and develop integrated intra and internode resource management techniques that provide both near-peak utilization and guaranteed high performance in shared environments.
- Video Magnification and Video Comparison for Sports: This project aims to develop motion magnification and comparison techniques for sports applications and for laparoscopic surgery.

See our website for videos describing the program, as well as each project.

Foxconn Basic Research in Interactive Computing

Hon Hai/Foxconn Technology Group, the largest contract manufacturer of electronics worldwide, is working with CSAIL. The research agreement is currently funding six major research areas in several areas of computer science, ranging from networking, human–computer interactions, computer graphics and vision, to theory.

The Basic Research in Interactive Computing project is a nine-year, $8.7 million research collaboration sponsored by the Hon Hai/Foxconn Technology Group. This research collaboration has completed phases one and two, and began phase three in January 2014. The research is predicated on the belief that computers and information devices are fast becoming interactive: they interact with other computers, with their environments, and above all, with humans. Each form of interaction adds a new dimension to the challenge of modeling and understanding the behavior of computer systems as well as the task of building and using these systems.

During the current year, the collaboration supported five PIs in the following research areas:

- Advanced Hand-Tracking and Gesture-Based Interaction (Professor Randall Davis)
- Reducing the Energy Consumption of Cellular Wireless Interfaces (Professor Hari Balakrishnan)
- Bringing Software Defined Networks to Wireless LANs (Professor Dina Katabi)
- Multicore Software Technology for Video-Conferencing Systems (Professor Charles Leiserson)
- iDiary: Activity Modeling Using Audio, Video, and GPS Streams from Phones (Professor Daniela Rus)

Internet Policy Research Initiative: Enhancing Public Trust in our Connected World

The mission of the Internet Policy Research Initiative (IPRI) is to work with policy makers and technologists to increase the trustworthiness and effectiveness of interconnected digital systems. We accomplish this with targeted engineering and

public policy research, various educational programs geared toward students and policy makers, and outreach programs to build policy communities that facilitate communication, education, and information exchange. Our work during the AY2016 period is summarized below:

Communication and information networks are a fundamental infrastructure for our increasingly digital economy and society. Technologists and policy makers both play key roles in supporting this transition, yet they approach issues from different perspectives. This can lead to not-fully-informed policy making or misdirected research efforts. There is a pressing need to bridge the gap between technical and policy communities because of society's reliance on this critical infrastructure.

Research—Cross-disciplinary research synergy

Opening a new field of study involves identifying key research questions, determining appropriate methodologies, and then doing the research in a cross-disciplinary and collaborative manner. Our first challenge was to identify key research questions, select appropriate methodologies to guide our work, and establish patterns of cross-disciplinary collaboration.

IPRI's core research can be broken down into four categories: cybersecurity, privacy, networks, and the Internet experience.

Cybersecurity

- Anticipating Security Externalities from Internet of Things Architectures (PI: David Clark)
- Net.info: A Provider-to-Subscriber Secure Communication Channel (PI: David Clark)
- Foundations of Internet Security (PI: David Clark)
- Using Natural Language Processing to Understand What Institutions Mean By Cybersecurity and What Topics the Term Covers (PI: David Clark)

Privacy

- Eliciting Expectations: Consumer Privacy Online (PI: Daniel Weitzner)
- Privacy-Relevant Information Leakage from Wi-Fi Computing Devices (PI: Daniel Weitzner)
- Discrimination in Predictive Algorithms (PI: Lalana Kagal)
- Privacy Evaluation: Metrics and Measurement for Privacy Evaluation (PI: Karen Sollins)

Networks

- Improving Performance and Cost of Content Delivery in a Hyperconnected World (PI: David Clark)

- What Affects the Performance and Reliability of the Public Internet? (PI: David Clark)
- Localizing and Characterizing Congestion in the Internet (PI: David Clark)
- Measuring Internet Performance (PI: David Clark)
- Future Internet Architecture (PI: David Clark)
- Disclosure and Transparency Policies for an Open Internet (PI: David Clark)
- Spectrum License Design, Sharing, and Exclusion Rights (PI: David Clark)
- Performance Expectation for Gigabit Broadband Services (PI: David Clark)
- Location-Independent Systems Evaluation: Multidimensional Evaluation Framework for Clean-Slate Internet Architectures (PI: Karen Sollins)

Internet Experience

- Social Linked Data Re-Decentralizing the Web (PI: Tim Berners-Lee)
- The Car Can Explain! (PI: Gerald Sussman; Co-PIs: Daniel Weitzner, Hal Abelson, and Lalana Kagal)
- Investigating the Growth and DoD Application of Internet of Things Technologies (PI: David Clark)
- Accountable Information Systems for Distributed Information Sharing Environments (PI: Lalana Kagal; Co-PIs: Hal Abelson and Daniel Weitzner)
- Foundations of Internet Policy (PI: David Clark)
- MIT App Inventor (PI: Hal Abelson)
- Punya (PI: Lalana Kagal)
- Accountable Cloud Computing (PI: Lalana Kagal)
- AppInventor, the Internet of Things, and Bluetooth Low Energy Devices (PI: Hal Abelson)
- Analyzing Trajectories of Informal Learning with MIT App Inventor (PI: Hal Abelson)

Education—Developing educational pathways

IPRI is working on creating educational pathways for students interested in Internet and cyberpolicy here at MIT. During AY2016, instructors and professors in our group taught the following related subjects.

- Fall 2015
  - 6.805/STS085/STS487 Foundations of Internet Policy
  - 17.310/ESD103/STS482 Science Technology, and Public Policy
  - 17.445/17.446 International Relations Theory

- Spring 2016
  - 6.S898 Cybersecurity Policy
  - 6.S978 Privacy Legislation: Law and Technology

Outreach—Impact on public policy dialogue

Engagement with policymakers, industrial partners, and civil society organizations is central to tuning our research agenda to ensure that our work makes a contribution where needed. We engaged with policy makers both here in the United States and abroad. During the past year, we've had a series of high-level policy makers come to MIT to meet with IPRI.

- Meetings at MIT
  - Robert Hannigan, Government Communications Headquarters (GCHQ), United Kingdom
  - Andrus Ansip, Vice President of the European Commission, European Union
  - Penny Pritzker, US Secretary of Commerce, United States
  - Mignon Clyburn, FCC Commissioner, United States
  - Angelo Cardani, Autorità per le Garanzie nelle Comunicazioni (AGCOM), Italy
  - Fátima Barros, Autoridade Nacional de Comunicações (ANACOM), Portugal
  - Henk Don, Authority for Consumers and Markets (ACM), Netherlands
  - Ömer Fatih Sayan, Information and Communication Technologies Authority, Turkey
  - Sébastien Soriano, Autorité de Régulation des Communications Électroniques et des Postes (ARCEP), France
  - Wilhelm Eschweiler, Bundesnetzagentur (BNetzA), Denmark

- Testimony
  - Daniel Weitzner – Congressional Testimony – April 19, 2016

Intel Science and Technology Center in Big Data

Professors Samuel Madden and Michael Stonebraker continue to head the Intel Science and Technology Center in Big Data, based at CSAIL. Professors Madden and Stonebraker are leading a team of 20 researchers (from Brown, MIT, Portland State University, University of Tennessee, and University of Washington) in a project to build new software tools to process and manage massive amounts of data. Specifically, the center is focused on new data management systems and new computer architectures that together can help users process data that exceeds scale, rate, or sophistication of data processing that existing systems provide. The center is developing new technologies in a number of areas including, data-intensive scalable computing,

machine learning, computer architecture, and domain sciences (genomics, medicine, oceanography, imaging, and remote sensing).

Last year, the center was renewed for an additional two years, until September 2017. During the most recent year, the center has focused on developing a capstone project centered around applying big data technologies to two application areas: (1) medical data processing including imagery, text, genomics, medical records, labs, and physiological signals; and (2) oceanography data, including various measures of ocean chemistry as well as genetic surveys of microscopic sea life. The centerpiece of this capstone project is a new software architecture for combining multiple data processing systems (each suited to a particular type of data, such as text), together in a new platform called BigDAWG.

World Wide Web Consortium

The World Wide Web Consortium (W3C) was founded at MIT in 1994 by the inventor of the web, Tim Berners-Lee. W3C is responsible for developing and maintaining the standards that make the web work and for ensuring the long-term growth of the web. Over four hundred member organizations, including most of the world's leading technology companies, are working to enhance the capabilities of web documents and create the Open Web Platform for application development, available across a wide range of devices, enabling everyone on the planet to collaborate and share data and information.

In recent years, a great many factors (people, devices, bandwidth, policy decisions, etc.) have extended the reach of the web into society. Video, social networking tools, user-generated content, location-based services, and web access from mobile devices are transforming many industries, including mobile, television, publishing, automotive, entertainment, gaming, and advertising. This transformation has led to greater demands on the W3C and other organizations to build robust technology that meets society's needs, in areas such as privacy, security, accessibility, and multilingual content.

Core Technology Focus

W3C standards define the Open Web Platform for application development that has the unprecedented potential to enable developers to build rich interactive experiences, powered by vast data stores that are available on any device. Although the boundaries of the platform continue to evolve, industry leaders speak in unison about how HTML5 (published in October 2014) is the cornerstone for this platform. But the full strength of the platform relies on many more technologies that the W3C and its partners are creating, including cascading style sheets, scalable vector graphics, web open font format, real-time communications, the Semantic Web stack, extensible markup language, and a variety of application programming interfaces. The platform continues to grow, and in turn, the W3C community is growing to meet the demand.

With the completion of HTML5, there are many new areas of focus. Publicly noted security breaches have resulted in unprecedented attention to fixing cybersecurity. The growth of e-commerce has focused new attention on standardizing payment and e-commerce approaches. And with the Internet of Things arriving, our Web of Things

project aims to address semantic interoperability to prevent IoT from driving silos at the application level. The demand is also driving W3C to expand its agenda and the size of its community. W3C launched Community and Business Groups in 2011. After five years over 7,000 people participate. By making it easier for people to participate, W3C has increased the relevance and quality of its work and brought more innovators to the table for pre-standards and standards track work.

Industry impact and broadening the set of participants

In recent years, web technology is not only used by consumers and companies for information sharing, but increasingly the web is the delivery mechanism for companies to deliver their services. Examples include telecommunications (where web access is a key service), entertainment (which is increasingly delivered over the web), publishing (EPUB 3.1, the e-book standard builds on the Open Web Platform), and retail and financial services (both impacted with an increase of payments on the web). This has caused a diversification in the membership of W3C, and also has enriched the technical agenda to address new technical issues that arise.

Centers and Initiatives

CSAIL Alliance Program

The CSAIL Alliance Program (CAP) was approved by the CSAIL Executive Committee in September 2013 as a broader extension of CSAIL's Industry Affiliates Program (IAP). Building upon the strong success of IAP, CAP is a membership-based program that serves as a portal for industry, governmental organizations, and other institutions to engage with CSAIL. CAP's mission is to provide a proactive and comprehensive approach to developing strong connections with all that CSAIL has to offer. CAP has 60 members representing a variety of industry sectors and regions, including North America, South America, Europe, and Asia. Key benefits of the program include:

- keeping abreast of the latest research at CSAIL,
- identify opportunities to foster and explore new research collaborations, and
- engaging with CSAIL students for recruiting.

Members of CAP are given access to new technologies and ideas as they move from laboratory to marketplace. Throughout the year, CAP hosts events, seminars, meetings, and activities to help connect member companies with the lab. To facilitate collaboration and communication with industry partners, CAP maintains a members-only website, produces research briefings, and publishes a CSAIL Student Profile Book.

CAP's flagship event is the annual meeting held each May. The event is a two-and-a-half-day conference that showcases CSAIL's research and students as well as serves as a forum for members to network, connect, and learn from each other. All activities are designed to help members to connect with CSAIL researchers in an effort to create value for their organization.

bigdata@CSAIL

The goal of the MIT Big Data Initiative at CSAIL (bigdata@CSAIL), a multiyear effort launched in May 2012, is to identify and develop new technologies needed to solve the next generation data challenges that will require the ability to scale well beyond what today's computing platforms, algorithms, and methods can provide. We want to enable people to truly leverage big data by developing tools and platforms that are reusable, scalable, and easy to deploy across multiple application domains.

Our approach includes two important aspects. First, we work closely with industry and government to provide real-world applications and drive impact. Promoting in-depth interactions between academia, researchers, industry, and government is a key goal. Second, we believe the solution to big data is fundamentally multidisciplinary. Our team includes faculty and researchers across many related technology areas including algorithms, architecture, data management, machine learning, privacy, security, user interfaces, and visualization. In addition, the bigdata@CSAIL team also includes domain experts in finance, medicine, smart infrastructure, education, and science. The initiative focuses on four broad research themes: Computational Platforms, Scalable Algorithms, Privacy and Security, and Big Data Applications.

Bigdata@CSAIL is supported by a group of sponsor companies including: British Telecommunications, EMC, Huawei, Intel, Microsoft, Quanta, Shell, and Thomson Reuters. The initiative is in year four of its five-year planned run and is in the process of transitioning to new a initiative addressing artificial intelligence and systems.

cybersecurity@CSAIL

Cybersystems cover communications, banking, data processing, purchasing, power and energy infrastructure, transportation, and defense—nearly every aspect of our lives. Consequently, cyberattacks have become more frequent, and more devastating. The present weaknesses in both hardware and software continue to threaten not only confidentiality of private data and the integrity of data at large, but also the availability of the critical operating systems organizations use to support internal operations, manage assets, secure logistics, sales, and even personnel.

Through cybersecurity@CSAIL, we are not just designing technology for specific tasks, but working toward solutions for the whole security spectrum. We approach security from all sides: programming languages, software verification, computer architecture, cryptography, systems, and policy. Our goal is to create security by default and remove program error as a source of vulnerability. We are designing new theoretical and practical foundations of secure computing that integrate security in the design process.

Our objective is to design protocols to make cyberattacks more difficult, retain function despite such attacks, and allow systems to recover quickly after an attack. Cybersecurity@CSAIL intends to maintain an interdisciplinary focus that brings together thought leaders from industry and government with MIT faculty, researchers, and students conducting research across the security spectrum in hardware, software, encryption, and theory specifically addressing the challenges of ensuring operating

system security, secure code, hardware designs for optimal security, defense tools, securing the cloud, multiparty protocols, and usability of encrypted data.

Cybersecurity@CSAIL is an industry consortia model launched in March 2015. Our industry partners provide valuable perspective on the challenges faced across several industry verticals and include: Boeing, BAE Systems, Raytheon, BBVA, BP, Akamai, and VISA.

MIT Center for Wireless Networks and Mobile Computing

MIT Center for Wireless Networks and Mobile Computing (wireless@MIT) is an interdisciplinary center to develop the next generation of wireless networks and mobile devices. Headquartered at CSAIL, the center creates a focal point for wireless research at MIT and addresses some of the most important challenges facing the wireless and mobile computing fields. The goal of wireless@MIT is to develop next-generation wireless network technologies and mobile systems.

The features of the center are:

- Access the research and connect with MIT professors and their groups.
- Early access to the latest emerging wireless technologies.
- A seat at the table on influencing and impacting policies and next-generation technologies, providing neutral ground where companies, academics, and government representatives can discuss the future of the wireless industry and mobile applications.

Over the past year, the center has contributed multiple innovations that enable high-network performance and introduced new architectures and services for the Internet of Things, mobile health, low-energy sensors, big data communication and processing, and innovative radio hardware. These contributions are from multiple faculty members and their students, including Professors Dina Katabi, Hari Balakrishnan, Li-Shiuan Peh, Arvind, and others. The center's work includes contributions from over 50 MIT faculty, research staff, and graduate students. A key feature of the center is that it encourages collaborations between MIT faculty in various departments and labs. In particular, the center has faculty members from Electrical Engineering and Computer Science (EECS), Mechanical Engineering, and the Sloan School of Management. Its members also span multiple labs including CSAIL, the Laboratory for Information and Decision Systems, the Microsystems Technology Laboratories, and the Research Laboratory of Electronics. The center has eight members, Amazon, Cisco, Intel, Mediate, Microsoft, ST Microelectronics, Telefonica, and Google.

entrepreneurship@CSAIL

Entrepreneurship@CSAIL enables students to develop a technology-based entrepreneurial project. The focus is on building an implementation of ideas, technical progress, and the process of establishing a technology based start-up. As part of this program Tom Leighton, CEO of Akamai Technologies, provided an inspiring talk to students, and Professor David Gifford continued to provide mentoring to students that were interested in starting companies.

Research Highlights

In addition to the large-scale collaborative projects and research centers, numerous individual and multi-investigator projects are under way. A sampling of the work is highlighted below:

Program Languages and Verification

The mission of Adam Chlipala's group is to reimagine the programming process by integrating machine-checked mathematical proof at every stage. Chlipala believes that machine-checked proof will come to play the same engineering role played today by technologies such as static type systems. Designing and implementing complex systems is difficult, and our main tools are ways to decompose systems into simpler components, connected via well-defined interfaces. Formal specifications are the ultimate in flexible interfaces, and machine-checked proof allows the group to establish that realistic systems obey even the most precise interfaces.

This technology promises the twin benefits of more correct systems, developed at lower cost, by replacing debugging with proof of conformance to requirements. The group's work spans from gate-level hardware designs; to the low-level code of operating systems, runtime systems, and efficient low-level data structures; to correctness of individual high-level programs and of the compilers used on them; to the highest-level relational specifications that freely mix standard programming constructions with mathematical notation. All of the pieces are designed to connect together, to obtain end-to-end guarantees that systems at the level of hardware gates obey high-level specifications. The group also constructs proofs only for actual code that they are able to execute and benchmark, rather than just with idealized models.

One sample project at the high level is Fiat, a system that lets the programmer determine what a program *ought* to do, not *how* to do it. The Fiat compiler then takes charge of generating efficient low-level code automatically, in the process also building a rigorous proof that the code meets the programmer's requirements. Fiat is our take on how to realize the full promise of domain-specific languages.

Instead of defining a new language for each problem domain, the Fiat programmer works in the universal language of math. Fiat plugins explain how to implement different mathematical theories efficiently with code. Example theories implemented so far include SQL-style relational data management, parsing text via context-free grammars, and parsing binary wire formats by inverting bit-level encoding functions into decoders.

Another project at a lower level is Bedrock, a platform for verified multilanguage programming. The core of Bedrock is an object-file format containing not just the traditional assembly code, but also formal specifications and proofs. The group has developed compilers, linkers, and program-analysis tools that connect to this object-file format from a variety of different programming languages. All tools generate proofs as needed, so that in the end the group can link libraries compiled from many different languages, verified against specifications in those languages, to produce a final, multilanguage executable with its own first-principles correctness proof. Fiat is

connected to Bedrock via an extensible pipeline, where programmers teach the system how to realize high-level concepts with efficient assembly code, in a way that generates proofs witnessing its own soundness.

Aging in Place by Leveraging the Internet of Things Innovations

As life expectancy continues to dramatically increase, eldercare has become a huge challenge for modern societies. As a result, societies are suffering from an aging challenge. Every year in the United States, 2.5 million elders are treated in emergency rooms for falls, which costs the healthcare system $34 billion per year. There are 5 million people who have Alzheimer's disease and need continuous care. Seniors also suffer from depression and other chronic diseases, such as chronic obstructive pulmonary disease and congestive heart failure. Many health problems can be addressed if detected and treated early enough. Even if the condition cannot be cured, early detection can significantly improve health outcomes.

Professor Dina Katabi is working on an Internet of Things system for in-home monitoring of seniors' health and wellness. Emerald is a device that uses wireless signals to monitor elderly falls, mobility, breathing, and heart rate without requiring the person to carry or wear any sensors on their body.

Emerald monitors people without body contact by analyzing the wireless signals in the environment. Emerald can also monitor gait, which is known as the sixth vital sign and is a robust predictor for hospitalization and functional status. A video of Katabi's demo to President Barak Obama is available online.

Autonomous Mobile Robots

Nicholas Roy's research goal is to develop autonomous mobile robots that can carry out complex, long-duration missions with speed and agility in populated environments, and with little prior knowledge of the world. Roy's group develops novel representations and algorithms that allow robots to understand the state of the world, make good decisions about how to move through the world, and learn from experience as they interact with people. While navigation in complex environments may at first glance seem to be a very different problem for human–robot interaction, the group has a developed models and algorithms for learning, approximate inference, and efficient planning under uncertainty that apply equally across a range of domains.

One major thrust of the group's work this year has been navigation in complex, populated environments, with a particular emphasis on unmanned air vehicles (UAVs). Most unmanned vehicles rely heavily on GPS and prior maps for navigation, but GPS is frequently degraded in urban environments and unavailable indoors. Onboard sensing can be used to provide navigation solutions, but a key challenge for a highly dynamic vehicle such as a UAV is that the system must be able to estimate its position and velocity with sufficient accuracy and low enough latency to allow stable control. Additionally, the vehicle must be able to sense the environment around it and build a model that allows safe trajectories be to identified and followed. Finally, the sensing and computation required by the models and algorithms must be designed to match the size, weight, and power limitations imposed by the UAV. The group has shown that

with appropriate models built specifically for high-speed flight, they can address these challenges in a variety of ways. This year they demonstrated high-speed navigation in GPS-denied environments on a ground vehicle by learning predictive models of motion.

Roy's group has developed Bayesian nonparametric techniques that learn models of sensor data and environmental structure that allow very efficient (< 30 ms) predictions of temporally extended (e.g., > 6 m in length) trajectories from instantaneous sensor measurements, and allow a ground vehicle to travel faster than 13 m/s, more than double what could be achieved with a state-of-the-art planner. They have also transitioned their work to UAVs using cameras to efficiently learn models of environment geometry and plan collision-free trajectories through complex environments.

Another major thrust of their work has been in human–robot collaboration, to enable robots to work effectively as part of a team alongside humans in carrying out complex missions. A major driver of the cost of operating most mobile robots in the air and on the ground is the manpower required to support the control of the robot. By developing the models and algorithms that allow robots to communicate naturally with people through existing interaction modalities such as spoken language, we can dramatically reduce the amount of effort required by human teammates to support the robot, and dramatically improve the capabilities of the robot. By developing representations and algorithms that bridge the semantic gap between humans and robots, we can allow robots to interpret natural language speech from human teammates and carry out complex instructions. This year, the group developed the first symbol grounding model to allow robots to learn abstract concepts from data (e.g., the middle of a row of pallets) without needing to learn a new representation for every type of object (e.g., the learned models translate seamlessly from a row of pallets to a row of tools). The group's approach is the first to take advantage of semantic features of the language to factor the symbol grounding graph, substantially accelerating the inference process and allowing the robot to reason over arbitrary numbers of objects and entities in approximately linear. This efficiency matches the intuition that abstraction allows efficient reasoning about large sets of things.

Computational Connectomics

Understanding the structure and function of the nervous system is an exceptionally complex task: the system consists of thousands of cells connected to thousands of other cells in microscopic networks that extend over large volumes and exhibit a seemingly endless variety of behaviors. Nir Shavit's work focuses on mapping such networks at the level of synaptic connections, and understanding the mathematical relation of their connectivity and geometry to function, with the hope that they will help us unravel the mystery of thought.

Shavit's work, in collaboration with colleagues, is directed at creating a Brain Scope. This scope will enable the group to view a brain's connectome, the fine connected structure of its neural circuits. Their approach is to use electron microscopy on fine slices of brain and then use high-performance computing systems to automatically reconstruct the structure latent within them. Their technology for viewing connectomes is still in its infancy: the computational problems in converting the images into connectivity maps

are immense. And yet, there is great hope that the group will be able to do so in coming years through the development of new machine-learning algorithms and parallel-computing paradigms. Shavit likes to think of the field of connectomics as being in a similar stage to that of astronomy centuries ago, whereby using the limited data from their homemade telescopes, researchers such as Galileo managed to view for the first time the craters of Earth's moon as well as the many moons of Jupiter. In the years since Galileo, telescopes provided data that repeatedly reached beyond known theoretical frameworks, allowing us to unravel the mysteries of the universe. Shavit's hopes for the work the group is doing are similar: to provide neurobiologists with a tool to generate currently unobtainable connectivity data that will move us beyond known theories of how the brain works.

Shavit's group works not only on extracting microscopic connectivity and functional data, but also on using this data to build new mathematical models explaining how neural tissues compute. Their modeling spans the connectomics gamut from the behavior of individual neurons in exiguous circuits to collections of neurons in increasingly complex networks. They collaborate with neurobiologists to design experiments based on the group's theoretical models, and work extensively to analyze the resulting data in order to confirm or disprove the group's theoretical predictions.

Computation Structures

Daniel Sanchez's work is centered on computer architecture and computer systems, with a focus on highly-parallel processors. The imminent demise of Moore's law is transforming computing. Scientists, entrepreneurs, and society at large have become accustomed to rapid growth in computational power. Computer architects now face the daunting challenge of maintaining this growth without relying on progress in fabrication technology. Toward this end, systems in all domains, from smartphones to supercomputers, are becoming increasingly parallel and heterogeneous, featuring many simple and specialized cores to improve efficiency. But parallelism and specialization are not enough to make future systems practical. Specifically, systems face three key challenges. First, without reducing data movement, memory accesses and communication will dominate energy consumption. Second, current multicores can exploit only a fraction of the parallelism available in programs, limiting the utility of additional cores. Third, without new abstractions and simple programming models, these systems will be incredibly difficult to understand and use.

To tackle these challenges, Sanchez's research focuses on four main areas. First, Sanchez and his students are investigating scalable memory hierarchies that minimize data movement. Their contributions span hardware and software, including a new memory hierarchy that automatically adapts its structure to the needs of each program, approaching the performance of the best application-specific hierarchy; algorithms to jointly place data and computation; and techniques to eliminate pathologies in caches and predict their behavior. Second, they are designing a new multicore architecture that exploits all types of parallelism available in programs, and features a new programming model that makes writing parallel programs almost as easy as writing sequential ones. Using this architecture, they have demonstrated near-linear scalability of algorithms that are often considered sequential. Third, they are designing techniques to provide

strict performance guarantees in multicores, allowing full isolation among applications sharing the system, improving system utilization, efficiency, and ease of use. Fourth, their research on scalable simulation has improved the performance of architectural simulators by several orders of magnitude, making it practical to study thousand-core systems.

Two common themes underlie their research contributions and differentiate them from related work. First, their contributions often span the hardware–software boundary. By taking a cross-layer approach that combines the strengths of both hardware and software, they can achieve gains that neither hardware-only nor software-only techniques can provide. Second, they place a strong emphasis on designing architectural components that can be characterized with simple analytical models, and use them to understand and drive the design of novel techniques. Given the surging complexity and diversity of modern systems, using components that behave according to simple models will be key to make systems easy to understand and use.

Laboratory Sponsored Activities

CSAIL Outreach

CSAIL's Hour of Code: CSAIL hosts an annual presentation and demo fair in conjunction with the global Hour of Code movement, inviting local high-school students to learn more about a wide array of computer science research. In the past two years approximately 300 STEM students from schools in the greater Boston area attended, with CSAIL receiving support from high-profile public figures such as author John Green and musician will.i.am.

Reddit "Ask Me Anything": CSAIL regularly encourages the online community to submit questions about computer science and academia in a series of Reddit "Ask Me Anything" (AMA) sessions involving the lab's researchers. In the past two years the AMA sessions have spurred approximately 6,000 comments and questions, as well as more than 200,000 page views.

Middle East Education through Tech: CSAIL has been a long-time supporter of the Middle East Education through Technology (MEET) program, an innovative educational initiative aimed at creating a common professional language between young Israelis and Palestinians. MEET enables its participants to acquire advanced technological and leadership tools while empowering them to create positive social change within their own communities. Many MIT students volunteer to teach MEET summer courses at the Hebrew University in Jerusalem. CSAIL continues to provide financial support for the program.

Dertouzos Distinguished Lecture Series

The Dertouzos Lecture Series has been a tradition since 1976, featuring some of the most influential thinkers in computer science. On October 14, 2015, Michael Stonebraker gave a presentation during the AY2016 Dertouzos Distinguished Lecture Series titled "The Land Sharks are on the Squawk Box."

Research Highlights Competition

On February 3, 2016, CSAIL hosted a second research highlights party for CSAIL graduate students and postdoctoral associates. This event allowed 21 students to create and present refined, accessible presentations of their research while members of the CSAIL community were able to enter drawings for raffle prizes by attending their colleagues' talks. It was a fun way to build community and learn entirely new things about research at CSAIL.

Organizational Changes

Professor Daniela Rus has continued in her role as director of CSAIL. The director's duties include developing and implementing strategies designed to keep CSAIL growing and evolving, fund raising, determining laboratory policies, and examining promotion cases.

CSAIL's leadership team includes two associate directors and the executive cabinet. These leaders assist the director with her duties and the positions are appointed by the laboratory's director. Professors Daniel Jackson and Polina Golland continued their roles as the associate directors during AY2016.

Additionally, the CSAIL executive cabinet meets twice per month to review and advise the director on policy, processes, and activities within the laboratory. Members named to the executive cabinet included, Saman Amarasinghe, Regina Barzilay, John Costanza, Randall Davis, Srinivas Devadas, John Fisher, David Gifford, Polina Golland, Daniel Jackson, Martin Rinard, Daniela Rus, Nir Shavit, Karen Shirer, Howard Shrobe, Nickolai Zeldovich, and Victor Zue.

Victor Zue continued in his role of director of international relations managing the engagments and oversight of various important CSAIL international contracts and internationl contract negotations. John Costanza continued as the assistant director for infrastructure, overseeing information technology, infrastructure and user support, building operations, and communications. Karen Shirer continued in her role as assistant director for administration, overseeing finance and human resources. Lori Glover continued in her role as managing director of the CSAIL Alliance Program.

Srinivas Devadas remained the space czar, overseeing the space committee and managing the allocation of space within CSAIL. The space committee also implements improvements to the facilities that will increase the quality of the environment for the laboratory's faculty, staff, and students. The space committee also includes the two assistant directors, Jack Costanza and Karen Shirer as well as John Guttag.

Awards and Honors

Our faculty and staff have achieved many awards including the following:

- Hal Abelson: Classics Award Information Technology, Multimedia Educational Resource for Learning and Online Teaching
- Anant Agarwal: Harold W. McGraw, Jr. Prize in Education, McGraw Hill Education

- Bonnie Berger: Elected American Institute for Medical and Biological Fellows, Institute for Medical and Biological Engineering, Fellow, Institute for Medical and Biological Engineering
- Costs Daskalakis: Research and Development Award, Vatican Giuseppe Sciacca Foundation
- Erik Demaine: Rare Craft Fellowship Award
- Polina Golland: Faculty Research Innovation Fellowship, MIT EECS
- Piotr Indyk: Fellow, Association for Computing Machinery
- Piotr Indyk: Best Paper, International Conference on Machine Learning
- Stefanie Jegelka: Deutscher Mustererkennungspreis, German Pattern Recognition Society, Research Award, Google, CAREER Award, National Science Foundation
- Charles Leiserson: Fellow, Institute of Electrical and Electronics Engineers, Member, National Academy of Engineering
- Andrew Lo: Richard J. Davis Ethics Legal Regulatory Insight Award, Investment Management Consultants Association
- Nancy Lynch: Member, National Academy of Sciences
- Aleksander Madry: CAREER Award, National Science Foundation, Research Fellowship, The Sloan Foundation
- Silvio Micali: Great Immigrants The Pride of America, Carnegie Corporation of New York
- Rob Miller: Faculty Research Innovation Fellowship, MIT EECS
- Una-May O'Reilly: Peter Jackson Best Innovative Application Paper Award, International Association for AI and Law
- Jerry Saltzer: National Cyber Security Hall of Fame Award, National Cyber Security Hall of Fame
- Daniel Weitzner: JD Falk Award, Messaging Malware Mobile Anti-Abuse Working Group
- Patrick Winston: Bloomberg Business Week ranked Winston's course on artificial intelligence as one of the five best computer science classes in the United States
- Peter Yu: Second Place in Amazon Picking Challenge
- Victor Zue: Fellow, American Association for the Advancement of Science

Group Award

- World Wide Web Consortium, Emmy, National Academy of Television Arts and Sciences

Staff Awards

- Martin Demaine, Rare Craft Fellowship Award, American Craft Council

Key Statistics for Academic Year 2016

Faculty: 96 (17% women)
Research staff: 42 (31% women)
Administration, technical, and support staff: 75 (55% women)
Postdocs: 87 (15% women)
Visitors: 65 (17% women)
Paid Undergraduate Research Opportunities Program participants: 121 (35% women)
Master of engineering students: 78 (28% women)
Graduate students: 381 (19% women)

Daniela Rus
Director, Computer Science and Artificial Intelligence Laboratory