

FACILITIES AND OTHER RESOURCES

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Research Laboratory of Electronics, RLE. Offices are equipped with personal computers for each person, as well as large monitors for programming and visualization of medical images. Access to a Linux server with 24GB of RAM and servers with Tesla graphics processors housed in a cooled basement of building 36 on the MIT campus. A Matlab license is also in place for all laboratory computers.

Computer Science and Artificial Intelligence Laboratory, CSAIL Laboratory: The MIT Computer Science and Artificial Intelligence Laboratory (CSAIL) is located in the MIT Stata Center, a truly unique building designed by world-famous architect Frank Gehry. The Center occupies approximately 713,000 gross square feet, and is located in the new gateway to MIT campus at 32 Vassar Street.

Environment. CSAIL has long been a leader in the fields of Artificial Intelligence, Cognitive Science, and Computer Science. CSAIL is a home to 110 principal investigators, who include both MIT faculty and research staff; their numbers include seven current or former MacArthur fellows and eight Turing award winners. CSAIL consistently ranks near or at the top of undergraduate and graduate Computer Science programs in the world.

Institute for Data Science and Society (IDSS). The IDSS is a recently founded epicenter of statistical modeling and machine learning at MIT, and with the MGH/ BWH Clinical Data Science Center (CDSC) explores applications of machine learning in medical image computing and radiology.

COMPUTING

Stata Center's main data network was designed and implemented by CSAIL's full-time staff of network and computing system managers, known collectively as T!G (the Infrastructure group). It consists of a state-of-the-art 10 gigabit, single-mode fiber backbone, Cisco Catalyst series switches, a fault-tolerant network topology, and 10/100/1000 Ethernet service via CAT6 copper cable to the desktop. T!G also maintains and supports a variety of centralized computing resources for the CSAIL community, including email service, web servers, DNS, DHCP, and many other common enterprise services. An 802.11g wireless lan is available throughout virtually all of the occupied space in the building.

The computer vision group, in addition to having access to the resources described above, also maintains an extensive software library. This includes a range of generic public domain and commercial software: different compiler, debugger and software library packages (among others, GNU and Sun proprietary), Netscape, Matlab, Mathematica, and different security related packages such as S-key (one time passwords for access of the lab from outside), ssh and PGP (encryption). Graphics and image processing software packages include XGL, XIL, Open GL, Open Inventor, Data Explorer, Analyze, XV, IslandWrite, IslandPaint, and IslandDraw, and Adobe Photoshop for Unix systems.

CSAIL investigators also have access to a high performance cluster at the Massachusetts Green High Performance Computing Center (MGHPCC), the state of the art computing facility in Holyoke, Massachusetts. This resource provides peak capacity of 9,000 - 18,000 cores, peak memory of 36TB and a high-speed working storage of 1.5PB. The system is dedicated to life sciences research, and is jointly operated by a consortium of research groups from five universities and their associated partners. The system enables individual research participants to scale to the entire system for bursts of computation, but is operated so that all participants have balanced access on average. In addition to the resources available at CSAIL, Golland's group maintains a 300- core cluster for computationally demanding processing. The group also owns 300TB of secure storage space that is maintained by the Infrastructure group. A recent addition to the computing cluster includes servers with close to 100 GPUs to enable acceleration of machine learning algorithms. The computing resources available to the PI will be instrumental for the proposed research, which requires extensive computing power, memory and storage for manipulating large data sets.