Facilities, Existing Equipment, and Other Resources

The University of Chicago is a private non-profit institution located on the ethnically-diverse South Side of Chicago that has been a center of advanced learning and research since its inception in 1892. The University of Chicago is comprised of four graduate Divisions (Biological Sciences, Physical Sciences, Social Sciences, and Humanities), six professional schools (Chicago Booth School of Business, Divinity School, Harris School of Public Policy Studies, Law School, Pritzker School of Medicine, and School of Social Service Administration), the Graham School of General Studies, and the undergraduate College. The University has a unique history of organizing around research questions that cross disciplines rather than operating within disciplinary boundaries. The extent to which this strategy reflects University of Chicago is illustrated by its numerous interdisciplinary Committees, Centers and Institutes (described below). The University of Chicago maintains its commitment to scholarship, teaching, and research through its more than 2100 faculty members and a student population of approximately 15,600 with nearly 2/3 engaged in advanced research and professional study. Through the years, 86 Nobel Laureates (8 are current faculty), 44 members of the National Academy of Sciences, 169 members of the American Academy of Arts and Sciences, and 14 recipients of the National Medal of Science have been associated with the University as students, teachers or research investigators. The University of Chicago is ranked among the world's top universities by a number of criteria, including the amount of federal research funding received (despite a size much smaller than many of its academic peers). This spirit of discovery, innovation and public service provides a robust foundation for success.

Computational Facilities: The principal investigator has access to extensive computational facilities available at the University of Chicago to carry out the tasks described.

Access to Clinical Data for AI-enabled Analytics: The ZeD lab (overseen by Professor Chattopadhyay) is housed within the Department of Medicine at the University of Chicago, and has access to the full range of high end computing resources offered by the University of Chicago. In addition, Prof. Chattppadyay's laboratory has access to the HIPAA compliant clinical data warehouse maintained by the Biological Sciences Division as detailed below:

Research Computing Center (RCC) at the University of Chicago: The Research Computing Center (RCC) at the University of Chicago is a hub for researchers across various fields, offering advanced computing resources, storage systems, and a wide range of software packages to facilitate high-performance computing. This document provides an extensive overview of the RCC's facilities and capabilities, focusing on the Midway2 supercomputing cluster, storage and backup systems, software, and networking infrastructure.

Midway2 Supercomputing Cluster: Midway2 is the second-generation high-performance computing cluster at RCC, replacing the first-generation cluster, Midway, which was decommissioned in 2019. Midway2 comprises a large pool of servers, software, and storage that researchers can utilize to enhance the efficiency and scale of their computational science. The RCC provides resources for distributed computing and shared memory computing, as well as emerging technologies like accelerators and big-data systems.

University of Chicago researchers can access RCC resources for free. To extend RCC resources with additional storage and computation, researchers can refer to the Cluster Partnership Program.

Cluster Computing Resources: The RCC maintains three pools of servers for parallel and distributed high-performance computing:

1) Tightly-coupled nodes: Ideal for tightly coupled parallel computing tasks, these nodes on Midway2 have a fully non-blocking FDR and EDR Infiniband, providing up to 100Gbps interconnect.
2) Loosely-coupled nodes: Similar to the tightly-coupled nodes but connected with 40Gbps GigE instead of Infiniband, these nodes are best suited for distributed tasks.
3) Shared memory nodes: These nodes contain much larger main memories (up to 1 TB) and are ideal for memory-bound tasks.

RCC maintains Intel Broadwell (28 cores @ 2.4 GHz with 64 GB memory per node) and Intel Skylake (40 cores @ 2.4 GHz with 96 GB memory per node) CPU architectures.

RCC also maintains specialty nodes, such as large shared memory nodes with up to 1TB of memory per node and 16, 28, or 32 Intel CPU cores. At the time of writing, Midway2 contains a total of 16,016 cores across 572 nodes and 2.2 PB of storage.

Emerging Technology: RCC's emerging technology resources allow researchers to be at the cutting edge of scientific computing. These include:

1) Hadoop: A framework for large-scale data processing based on Google's paper and initially developed at Yahoo. Researchers can experiment with RCC's Hadoop infrastructure to become familiar with big data techniques.
2) GPU Computing: Scientific computing on graphics cards can unlock even greater amounts of parallelism from code. GPU nodes on Midway each have four Nvidia K80 accelerator cards and are integrated into the Infiniband network.

Storage and Backup: RCC hosts and maintains various storage systems, including persistent high-capacity storage that can be shared among a research group or remain private to each individual user, and high-performance storage for temporarily staging and quickly accessing data.

Persistent and High-Capacity Storage:

1) Home Directory: Each RCC user has a home directory for storing small, frequently used items such as source code, binaries, and scripts. The home directory is only accessible by its owner and is suitable for storing files that do not need to be shared with others. Data in the home directory can be accessed from Midway and remotely via different protocols.
2) Project Space: Principal investigators can request a project space for their research group. These directories are used for longer-term storage of data/files shared by members of a research group/project and are accessible from all RCC compute systems and remotely.

High-Performance Scratch Space: ScratchSpace: Hosted on RCC's high-performance storage system, scratch space is intended for staging data required/generated by computational processes running on the cluster. Unlike home and project directories, scratch space is neither snapshotted nor backed up and may be periodically purged. Users are responsible for ensuring any important data in the scratch space is replicated in a location providing persistent storage, such as project or home directories.

Backup and Data Recovery:

1) Filesystem Snapshots: Automated snapshots of home, project, and cds directories are available for recovering data in case of accidental file deletion or other problems. Typically, 7 daily and 4 weekly snapshots are available. However, RCC may reduce the number of snapshots during periods of high space usage.
2) Tape Backup: Nightly backups to a tape machine located in a different data center safeguard against hardware failure or disasters. These backups are intended for disaster recovery only, and users should rely on filesystem snapshots for regular data recovery. Users should also avoid using special characters in filenames, as they are not supported by the backup system.

Software: RCC installs, configures, and maintains hundreds of software packages on the Midway cluster. Some of the key software packages include:

1) Compilers: GNU and Intel C/C++/Fortran compiler suites, Nvidia's CUDA compiler for GPU computing, and compilers for Java, Julia, Go, and Haskell.
2) Interactive programming environments: Open-source and commercial programming environments like Python, MATLAB, Mathematica, Stata, and R. These environments often include pre-installed libraries and packages.
3) Data processing tools: Programs for dealing with large-scale data formats (HDF5 and NetCDF), data-movement programs (Globus), and database software (PostgreSQL and Hadoop).
4) Numerical libraries: Intel's Math Kernel Library (MKL) and OpenBLAS, the GNU Scientific Library (GSL), and FFTW (a Fourier transform library).
5) Community codes: Commonly-used scientific software such as LAMMPS, Gromacs, YT, lfrit, QIIME, and genetic analysis programs like SAMtools, Bowtie, BLAST, GATK, PLINK, and TopHat.

RCC can build and install open-source software upon request and help with negotiating licensing agreements, purchasing commercial software, or migrating purchased commercial software to RCC systems.

Networking: RCC's Midway supercomputing cluster is connected to the University of Chicago network backbone through a 10 Gbps network uplink. The university network connects to Internet2 at 10 Gbps and has 10 Gbps connections to other commercial networks. All Midway file transfer nodes and login nodes uplink through the Midway switch at 20 Gbps to the University of Chicago campus network backbone.

The research networks at the University of Chicago are deployed at two campus core distribution points, connecting via two 10-Gigabit Ethernet circuits to MREN and the CIC OmniPop. The connectivity provides the university with flexibility and capacity to connect to other institutions and share research data and resources.

The University of Chicago has built a Network infrastructure to establish a Science DMZ, which is distinct from the general-purpose campus network and purpose-built for data-intensive science. The Science DMZ includes support for virtual circuits, software-defined networking, and 100 gigabit Ethernet. RCC compute resources are now connected with a 40 Gbps Ethernet connection to the UChicago Science DMZ, and tests are being performed to ensure proper network traffic segregation.

Tape Backups. Backups are performed on a nightly basis to a tape machine located in a different data center than the main storage system. These backups are meant to safeguard against events such as hardware failure or disasters that could result in the complete loss of RCC's primary data center.

Data Sharing. All data in RCC's storage environment is accessible through a wide range of tools and protocols. Because RCC provides centralized infrastructure, all resources are accessible by multiple users simultaneously, which makes RCC's storage system ideal for sharing data among your research group members. Additionally, data access and restriction levels can be put in place on an extremely granular level.

Data Security & Management. The HIPAA compliant security of the Research Computing Center's storage infrastructure, protected by two-factor authentication, gives users peace of mind that their data is stored, managed, and protected by HPC professionals. Midway's file management system allows researchers to control access to their data. RCC has the ability to develop data access portals for different labs and groups.

The Institute for Molecular Engineering at the University of Chicago house a vibrant research community of multidisciplinary scientists that regularly collaborates to make significant scientific contributions. UChicago also features several translational resources, such as the Human Tissue Research Center, and the Transgenic Animal Center.

The University of Chicago Comprehensive Cancer Center (UCCCC): One of only two NCI-designated Comprehensive Cancer Centers in Illinois, the UCCCC has a reputation for excellence and innovation and a commitment to address cancer through clinical and basic science cancer research and training, clinical cancer care, and expertise in population research. UCCCC researchers have access to a comprehensive set of shared technologies with the University of Chicago Biological Sciences Division (BSD), including 13 Core facilities. The UCCCC offers a wealth of intellectual, technological, and financial resources to pursue a comprehensive, collaborative research program involving more than 215 renowned scientists and clinicians.

Translational and collaborative research: The University of Chicago's strong physical sciences division, including my home department of Chemistry, is located in direct proximity to the medical school and hospital system. Indeed, I chose to start my independent career here at UChicago specifically so that I could develop a group whose work could impact human health. I have now witnessed firsthand the benefits of this proximity and have developed several strategic collaborations with clinicians and clinical researchers to develop new technologies. The University devotes substantial resources to translational research, which provides a clear path to move from the bench to the clinic. This includes the University of Chicago Innovation Exchange (https://innovation.uchicago.edu), which provides seed money and expertise to translate basic science discoveries into commercial ventures and to foster collaborations between the basic science divisions, medical school, and national labs. Therefore, the University of Chicago is an exceptional location to pioneer paradigm-shifting biomedical technologies.

University of Iowa, Facilities and Other Resources: **Laboratory Space.** Dr. Manicassamy's Laboratory is housed in the Department of Microbiology & Immunology on the second floor of the Bowen Science Building (BSB; Core 400) at The University of Iowa. This state of the art facility is comprised of 1,700 square feet of newly renovated laboratory space. In addition to the infrastructure available in the Department of Microbiology & Immunology, Dr. Manicassamy's research is supported by excellent core facilities at the University of Iowa, including Next Generation Sequencing, Bioinformatics Core, Flow Cytometry, Central Microscopy Research Facility, Small Animal Imaging, DNA sequencing. Dr. Manicassamy has full-time administrative support through the Department of Microbiology & Immunology.

Office Space. Dr. Manicassamy has 200 square feet of separate office space adjunct to the laboratory in BSB. PC and Mac computers, computational network, laser printers, color printers, and scanners are available. Manuscripts and desktop publishing of papers can be prepared in several offices available to scientists working in BSB.

Scientific Environment. The University of Iowa has a highly collaborative research community with several leading Virologists and Immunologists, including Drs. Stanley Perlman (Coronaviruses pathogenesis/host responses), Mark Stinski (Herpes Virus), John Harty (T cell responses to Infection), Kevin Legge (dendritic cell-T cell responses to Influenza virus), Steven Varga (Host responses to RSV), Gail Bishop (Viral Immunology), Wendy Maury (Filovirus entry), Richard Roller (Molecular Herpes virology), Jack Stapleton (HIV/HCV pathogenesis), and Hillel Haim (HIV evolution/pathogenesis). Weekly Microbiology seminar series, and the Virology and Immunology journal clubs provide excellent opportunities for students and postdocs for scientific interactions. In addition, research program in pulmonary biology provide an excellent scientific forum for discussion and collaboration. The Levitt Center for Viral Pathogenesis provides funding for seminar speakers and student travel and is another venue for interactions with research interests related to the project. Moreover, the students and postdocs are supported by several NIH T32 training grants.

Animal Care Unit (ACU) at The University of Iowa is in full compliance with all NIH guidelines and regulations pertaining to the care and use of experimental animals (PHS Assurance no. A3021-01). The ACU has enjoyed accreditation from the American Association of Accreditation of Laboratory Animal Care since 1994. The ACU maintains centralized animal housing facilities, which are staffed by highly trained individuals to provide husbandry and research support services. In addition to providing daily animal care, all ordering and receipt of animals, quarantine and health monitoring is performed by the ACU. The ACU also provides research support services, such as anesthesia and surgical support, rodent breeding assistance, diagnostic laboratory services, and investigator training. The ACU veterinarians are faculty members of The University of Iowa. They are available to assist investigators and their staff with all aspects of their animal research activities.

Biosafety Facility. The select agent containment laboratories are housed on the 5 rd floor of the Carver Biomedical Research Building (CBRB) and on the 4 th floor of the Medical Laboratories (ML) Building, University of Iowa, Iowa City, IA. BSL3 facilities are CDC certified and is managed by Ms. Dana Reis (Director). Dr. Manicassamy has one BSL3 suite with one class IIb biosafety cabinet and one Animal BSL3 suite available for work on highly pathogenic influenza viruses and coronaviruses.

Major Equipment in Dr. Manicassamy's laboratory::

1) **Biosafety cabinets:** Sterile environment for handling infectious materials, safeguarding researcher and samples.
2) **Tissue culture incubators:** Optimal temperature, humidity, and gas conditions for cell and tissue growth.
3) **PCR machines and real-time PCR machines:** Amplify specific DNA sequences, monitor amplification process in real-time.
4) **Table top centrifuges:** Separate liquid sample components based on density.
5) **Microscopes:** Visualize microscopic organisms, cells, and minute structures.
6) **-80 freezers and -20 freezers:** Long-term storage and preservation of biological samples.
7) **Bacterial incubators:** Optimal conditions for bacterial growth and development.
8) **Bacterial shakers:** Facilitate aeration and mixing of bacterial cultures.
9) **Thermal cyclers:** Precise temperature control for DNA amplification and molecular biology applications.
10) **UV-Spectrophotometers:** Measure light absorbance, determine concentration and purity of biomolecules.
11) **Fluorescence microscope:** Visualize and analyze samples using fluorescence, study cellular structures and biomolecular interactions.
12) **Gel doc unit:** Capture and document images of DNA, RNA, and protein samples in gel electrophoresis experiments.

University of Iowa Department of Microbiology and Immunology Resources:

1) **3 flow cytometers:** Measure and analyze physical and chemical properties of cells or particles in fluid.
2) **Fluorescence plate readers:** Measure fluorescence intensity in microplate format, high-throughput analysis of cellular and biochemical events.
3) **Zeiss inverted fluorescent and confocal microscopes:** High-resolution imaging of biological samples, visualize living cells and tissues.
4) **qRT-PCR thermocyclers:** Quantify RNA transcripts in real-time, provide insights into gene expres-

sion.

5) **Fuji CCD imaging system:** Capture high-resolution images of fluorescent and chemiluminescent samples, sensitive and accurate detection of biomolecules.
6) **High-speed and ultracentrifuges with varied rotors:** Rapid separation of samples based on size, shape, and density, accommodate various rotor types.
7) **Typhoon imaging system:** Detect, quantify, and analyze proteins, nucleic acids, and biomolecules in gels, membranes, and microplates.
8) **ELISA plate readers:** Measure absorbance of enzyme-linked immunosorbent assays (ELISA), quantify proteins, peptides, and hormones.
9) **TopCount:** High-throughput quantification of radioactivity in samples, study biochemical and cellular processes.
10) **LiCor imaging system:** Sensitive and accurate detection of fluorescent and chemiluminescent samples in various formats.
11) **Darkroom facilities:** Controlled processing of light-sensitive materials, such as photographic films and imaging plates.
12) **Cold and warm rooms:** Temperature-controlled spaces for storage and experiments requiring specific temperature conditions.