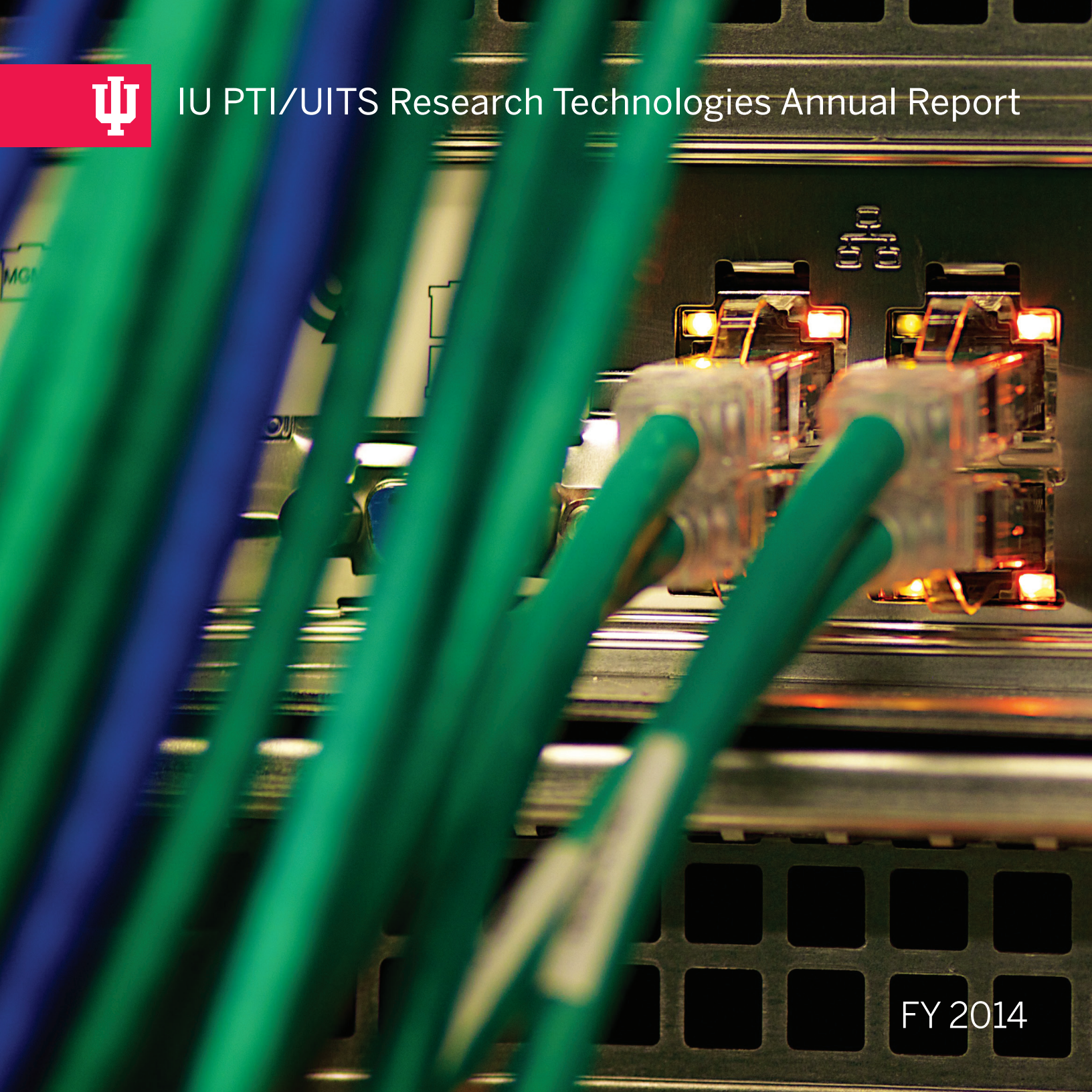
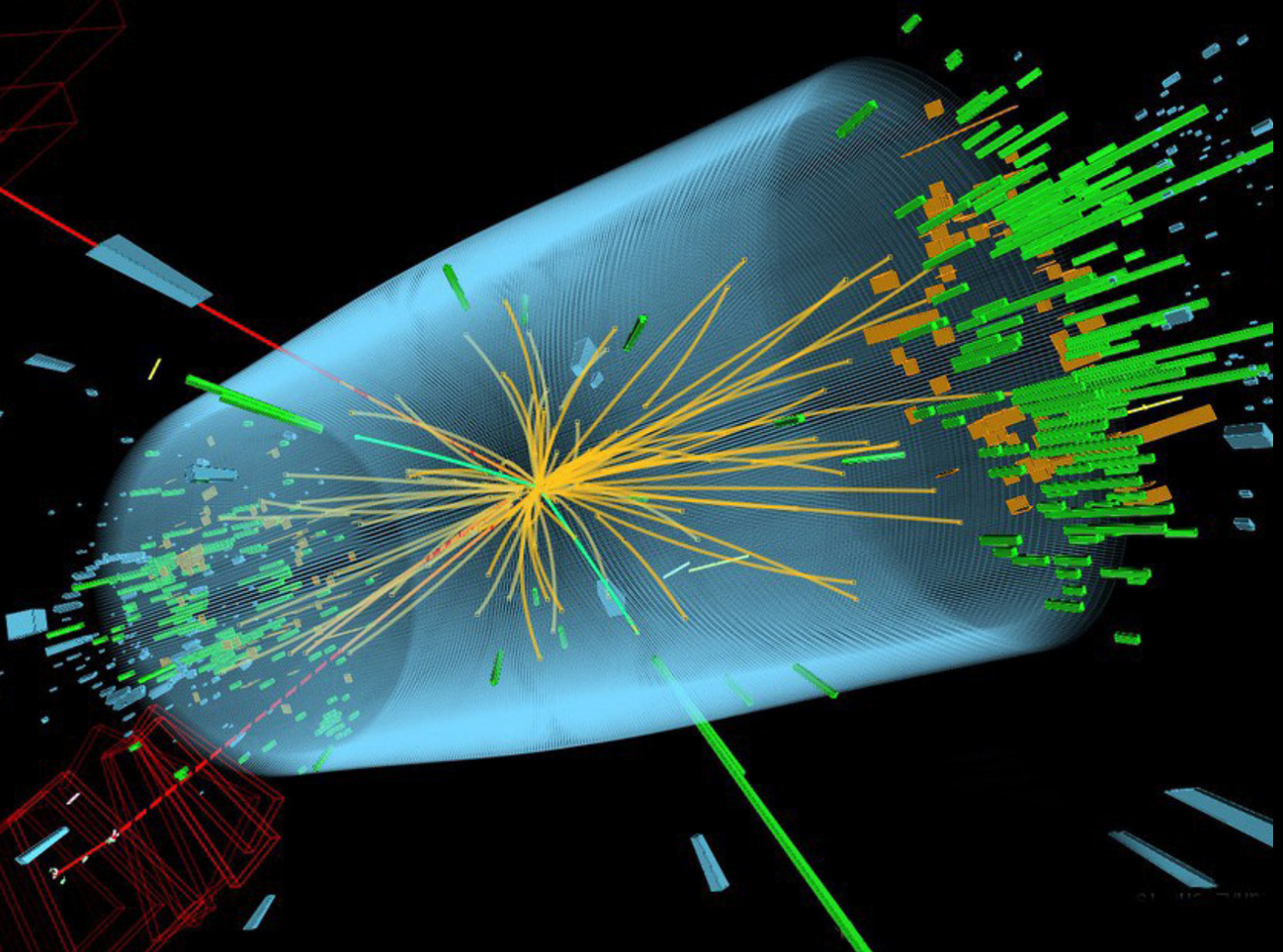


IU PTI/UITS Research Technologies Annual Report

FY 2014

Depiction of a Higgs boson event from collisions between protons in the Large Hadron Collider. IU physicists and information technology personnel contributed to the success of the Higgs boson search.

We have so much to CELEBRATE!



Our scholars used nearly

156M

CORE HOURS

on IU's systems

Our clients were awarded

39% of **$476M**

in IU annual grants

Open access to world-class resources, open doors to a growing community

At PTI/Research Technologies, we believe technology can be a major driver of new and expanded possibilities, from discovering elemental particles like the Higgs boson to understanding the causes of Alzheimer's disease.

IU's technology is so advanced that we use an advanced term to refer to it: Cyberinfrastructure. This includes massive supercomputers, large and fast disk storage systems, and reliable archival storage systems, all in service to one mission — to enable the IU community to improve the quality of life in the state of Indiana and the world by expanding their research, scholarly, and creative capabilities through novel innovations in IT, informatics, and cyberinfrastructure systems.

In this Fiscal Year 2014 (FY2014) report we outline IU community accomplishments related to several IU Bicentennial Strategic Plan goals and ongoing principles of excellence. The report is intentionally concise and focused on overarching facts and figures. (If you're a technical person who wants more detail, view the longer version on Scholarworks: *go.iu.edu/tCq*)

Researchers representing a total of **110 disciplines** *and sub-disciplines are using Big Red II.*

Accelerating discovery

A commitment to student success

More undergraduates use research and analytical software licensed for use at Indiana University than any other PTI/Research Technologies service.

A total of 20 research-oriented software packages are purchased by the university and resold at a discount to students and departments. Research Analytics also provides no-cost access to these titles within Student Technology Center (STC) labs and through IUanyWare.

Titles like SPSS, Mathematica, and MATLAB are used hundreds of thousands of times in the STC labs and IUanyWare each fiscal year. A large percentage of this usage comes from students. Statistical packages like SAS and SPSS are an integral part of their coursework and education.

Graduate and undergraduate students make extensive use of IU's research cyberinfrastructure.

This is part of the IU "open access" policy for advanced cyber-infrastructure resources, put in place in 1955 by Marshall Wrubel, the first permanently appointed director of the IU Research Computing Center. At the time, "research computing" was the prevalent term for what we now call cyberinfrastructure.

Type of system	Undergraduate student users	Graduate student users	Total student users
Supercomputers and computational systems	1,245	1,536	**2,781**
Advanced storage systems	146	690	**836**
Advanced visualization systems*	84	46	**130**
RT Github code repository	1,317	888	**2,205**

* An additional 1,319 students ranging from middle school to graduate school participated in demonstrations of IU's advanced visualization systems.

"... as far as I am concerned I have everything I need through our infrastructure ... ***I realize that I am very lucky."*** **— IU faculty member**

A community of scholars

PTI and the Research Technologies support collaborative networks of scholars both at and beyond IU. At IU, collaborative networks extend across campuses and across research, scholarly, and creative disciplines.

Network	Topic	IU faculty members involved	US faculty members outside IU	Total researchers in collaborative network
Pervasive Technology Institute	Informatics and information technology	6	0	**6**
Indiana CTSI (Clinical and Translational Sciences Institute)	Medical research – translational and clinical research designed to create new medical treatments	3,915	2,782	**6,697**
CIFASD Collaborative Initiative	Research on Fetal Alcohol Spectrum Disorder (Fetal Alcohol Syndrome)	7	53	**60**
National Gene Vector Biorepository (NGVB)	Gene therapy research	10	152	**162**
Alzheimer's Disease Neuroimaging Initiative (ADNI)	Alzheimer's disease causes	50	6,030	**6,080**
Open Science Grid	Physics and biology	30	24,579	**24,609**

To expand these networks, we host or assist with statewide or national workshops like:

- The IEEE Cluster 13 Conference (cluster computing)
- Harness the Power of GPGPU (general purpose computing on graphics processing units)
- The Polar Technology Conference (remote power systems in polar environments)

And we maintain leadership positions in international organizations such as:

- The IU Bioinformatics Clinic (integrative genomics training)
- AAMC GIR Steering Committee (academic medical center IT leadership)
- AAMC GIR Leadership Institute (training IT leaders for medical centers)
- Cray and IBM user groups (officers in both organizations)
- SPEC international computer performance analysis standards organization
- IEEE/ACM Annual International Supercomputing Conference
- SPEXXA — the German Priority Programme for Software for Exascale Computing

Scholary Data Archive

The Scholarly Data Archive has nearly 42 PB of overall tape for storing and accessing research data, with automatic offsite copies for disaster recovery.

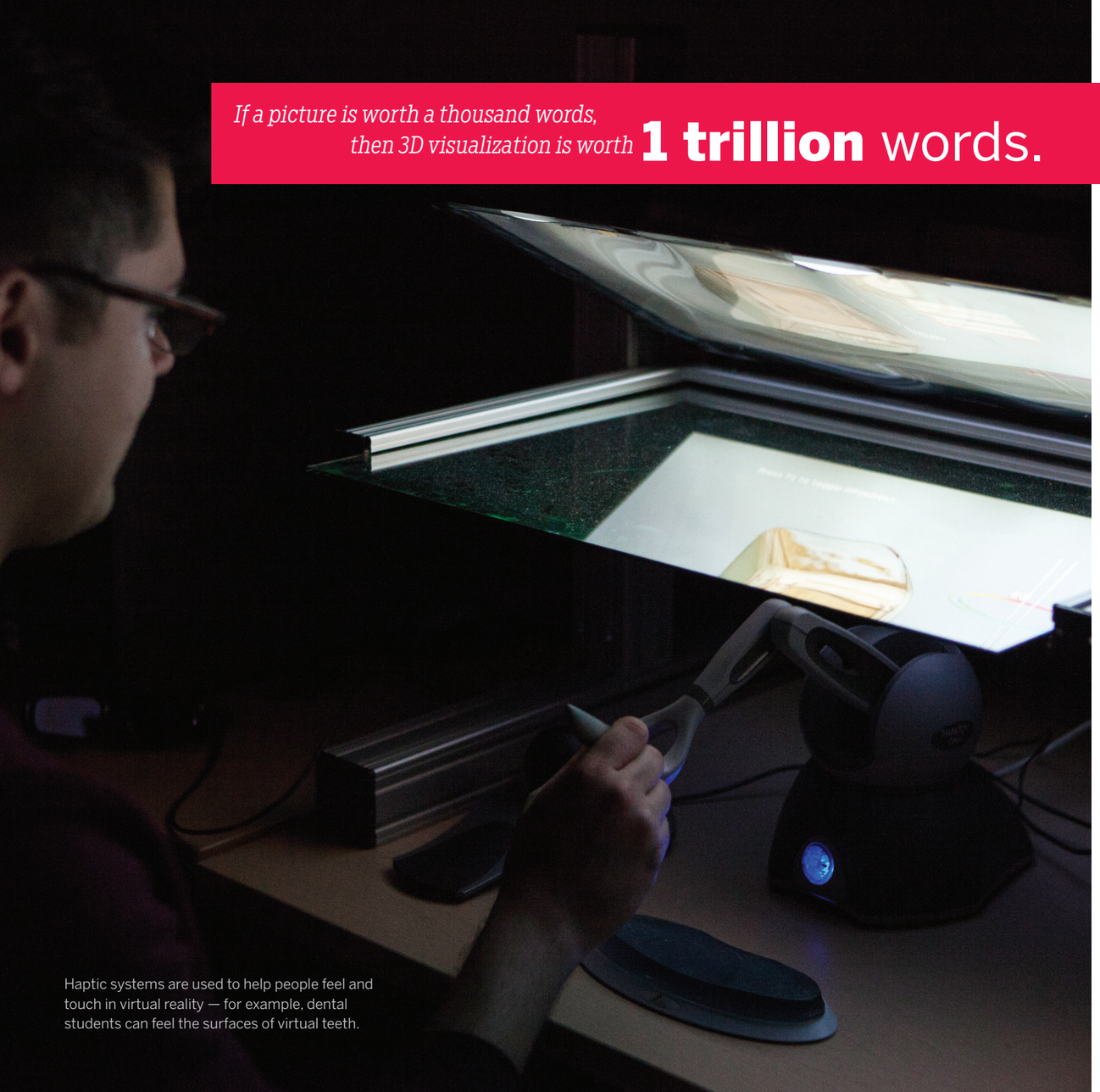
Resources that expand what's possible

IU was the first US institution operating non-classified supercomputers and storage systems to have such supercomputers aligned with HIPAA, making it possible to analyze sensitive, protected personal health information securely. To the best of our knowledge, we're still one of just two public institutions in the US with this capability. For researchers, this means secure access to important diagnostic data; for patients, it means sensitive health information benefits from IU's highest levels of protection.

Another unique aspect of IU's research storage systems is that data stored in the Scientific Data Archive is copied in duplicate — once to a tape library in Indianapolis, and a second time to one in Bloomington. IU was the first US university research computing center to provide this capability.

Data redundancy is key for the School of Medicine and other Clinical Affairs schools. Together, they take advantage of nearly 12.9M core hours on IU supercomputers, which represents 9.53% of total usage. They also take advantage of nearly 614,000 GB of disk storage, which represents 30.35% of total usage.

If a picture is worth a thousand words, then 3D visualization is worth **1 trillion** words.



Haptic systems are used to help people feel and touch in virtual reality — for example, dental students can feel the surfaces of virtual teeth.

Visualization System	Purpose	Recent major activities
Virtual Reality Theater at IUPUI	Reconfigurable "cube" that provides highly immersive, interactive experiences for groups up to 25	Virtual reviews of architectural and lighting designs, exploration of large scale data sets, virtual walkthroughs of re-creations of archeological treasures, and immersion in creative artistic and gaming worlds
Visualization and Collaboration Theater at IU Bloomington	Advanced video conferencing, presentations, and 3D data explorations by groups up to 60	
Science on a Sphere	Six-foot sphere illuminated by multiple high-definition projectors to display a wide range of oceanographic, atmospheric, astronomical, political, and economic data	Visualizations of atmospheric particulate dispersion, mapping of international student populations, student digital humanities projects, and exhibits of student work in digital photography, design, and virtual environments
IQ-Series • Ultra-high resolution IQ-walls • Multi-touch IQ-tables • Hybrid IQ-tilt multi-touch system • Semi-immersive • IQ-stations • IQ-touch system	Bring the essential benefits of high-end facilities directly to labs, classrooms, studios, and galleries across IU	Digitized Audubon portfolio for the Lilly Library, treasures from the Mather's Museum for its 50th anniversary, a collection of rare maps in the Wells Library, and a "scrollable" version of the original 120-foot long manuscript for Jack Kerouac's *On The Road*.

Computation System	Capacity	Purpose
Big Red II	**1,000.3** TFLOPS **21,824** cores	• Large-scale computation • Parallel approaches
Karst	**93.2** TFLOPS **4,096** cores	• High-end, data-intensive apps • Condominium cluster environments
Quarry	**26.11** TFLOPS **2,960** cores	• Uniprocessor and single node jobs • (replaced by new Karst system)
Mason	**4.29** TFLOPS **512** cores	• Computational work • Large memory apps

Storage System	Capacity	Purpose
RFS	**80 TB** (60 TB usable) OpenAFS 420 TB (336 TB usable) GPFS file system	Group collaboration — even at multiple institutions — via file sharing
Data Capacitor II	**5 PB** (3.75 PB usable) Lustre file system with 48 GB/s max I/O	Massive storage capacity and high-speed I/O for big data
DC-WAN	**1.47 PB** (1.1 PB usable) Lustre storage system with 40 GB/s max I/O	Remote data access as if it was stored locally
SDA	**15 PB** tape (plus 0.8 PB disk, 0.6 PB usable) HPSS file system	Research data access and storage with highly reliable disaster protection

Big Red II

Big Red II is IU's biggest and fastest supercomputer. It's capable of one petaFLOPS of calculations — that's a thousand trillion mathematical operations per second.

As of dedication in 2013, Big Red II was the first supercomputer to surpass 1 petaFLOPS and be owned by any US university, paid for by that university, and set aside for that university's exclusive use. As of the end of FY 2014, 393 IU researchers, representing a total of 110 disciplines and sub-disciplines, were using Big Red II. During FY 2014, the IU research community clocked a total of 134,910,214 CPU hours.

Big Red II is roughly the size of **12** industrial refrigerators.

One of the critical purposes of a massive supercomputer is to allow dozens to hundreds to thousands of computer processor cores — the individual computing unit that does calculations — to work together to solve the most challenging data analysis and enable the most sophisticated simulations. A 1 petaFLOPS computer system is capable of performing one quadrillion (10^{15}) floating-point operations per second. To match what it can do in just one second, you would have to perform one calculation every second for 31,688,765 years — without taking a break.

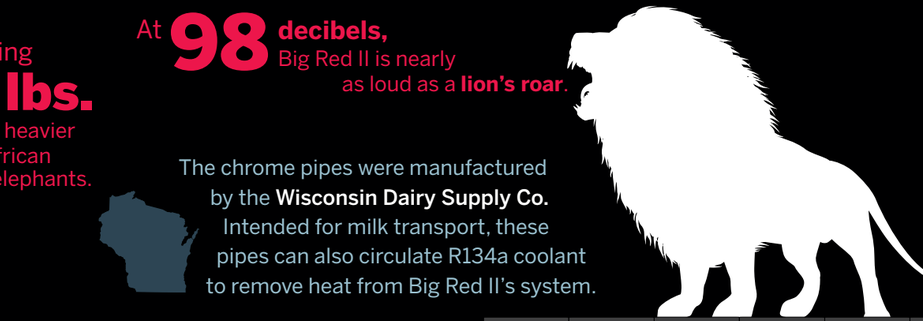
Big Red II removes heat from its system at a rate of **1.2M** BTU/h.

In this time, a bathtub of freezing water would boil in about **4 minutes.**

Weighing **30,000 lbs.**

Big Red II is heavier than **2** African elephants.

At **98 decibels,** Big Red II is nearly as loud as a **lion's roar**.



The chrome pipes were manufactured by the **Wisconsin Dairy Supply Co.** Intended for milk transport, these pipes can also circulate R134a coolant to remove heat from Big Red II's system.

"Data sets of unprecedented scope can facilitate new discoveries regarding the brain, genome, disease and therapies but computational power has become a major bottleneck to scientific progress. To analyze the entire human genome in relation to longitudinal changes on brain MRI and PET scans in over 800 individuals, ***we need significant computing power.*** *"* **— Andy Saykin**

Catalyzing research

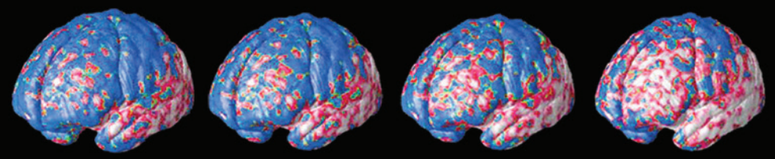
Genetics: Finding the sources of Alzheimer's disease

Nearly 44 million people worldwide may be living with Alzheimer's disease or other dementias, according to the BrightFocus Foundation. Millions of people watch as the disease first robs loved ones of their short-term memory, then long-term memory, and even their personality.

Since 2005, the Alzheimer's Disease Neuroimaging Initiative (ADNI) has been studying genetic causes of Alzheimer's and environmental factors that contribute to the disease. ADNI's immediate goal is to find ways to make early diagnoses of Alzheimer's. Its ultimate goals are to be able to treat and finally prevent this terrible disease.

Dr. Andrew Saykin, Raymond C. Beeler Professor of Radiology and director of the IU Center for Neuroimaging, is using the advanced cyberinfrastructure of PTI/Research Technologies to untangle the causes of Alzheimer's disease. Saykin's current ADNI studies required assembling the entire genomes of 818 study volunteers.

Thanks to Big Red II, Saykin was able to sequence these genomes in roughly eight months — using up to one Petabyte of data storage. The 818 assembled genomes are allowing Saykin and colleagues to relate the genetic sequences of healthy individuals and Alzheimer's victims to their genes. They can then use brain scans and behavioral data to track the disease's progress.



Scans over an 18 month period show healthy brain activity (red and blue areas) and rapidly spreading areas of cell death (gray areas) in someone with Alzheimer's disease.

"Many of us spent decades designing and building the ATLAS experiment. We always hoped that the results of this effort would be significant insights into the workings of our universe, but ***the discovery of the Higgs is an extraordinary outcome****. It is the most significant discovery of my career."* **— Harold Ogren**

Large Hadron Collider

Physics: Nobel prize, anyone?

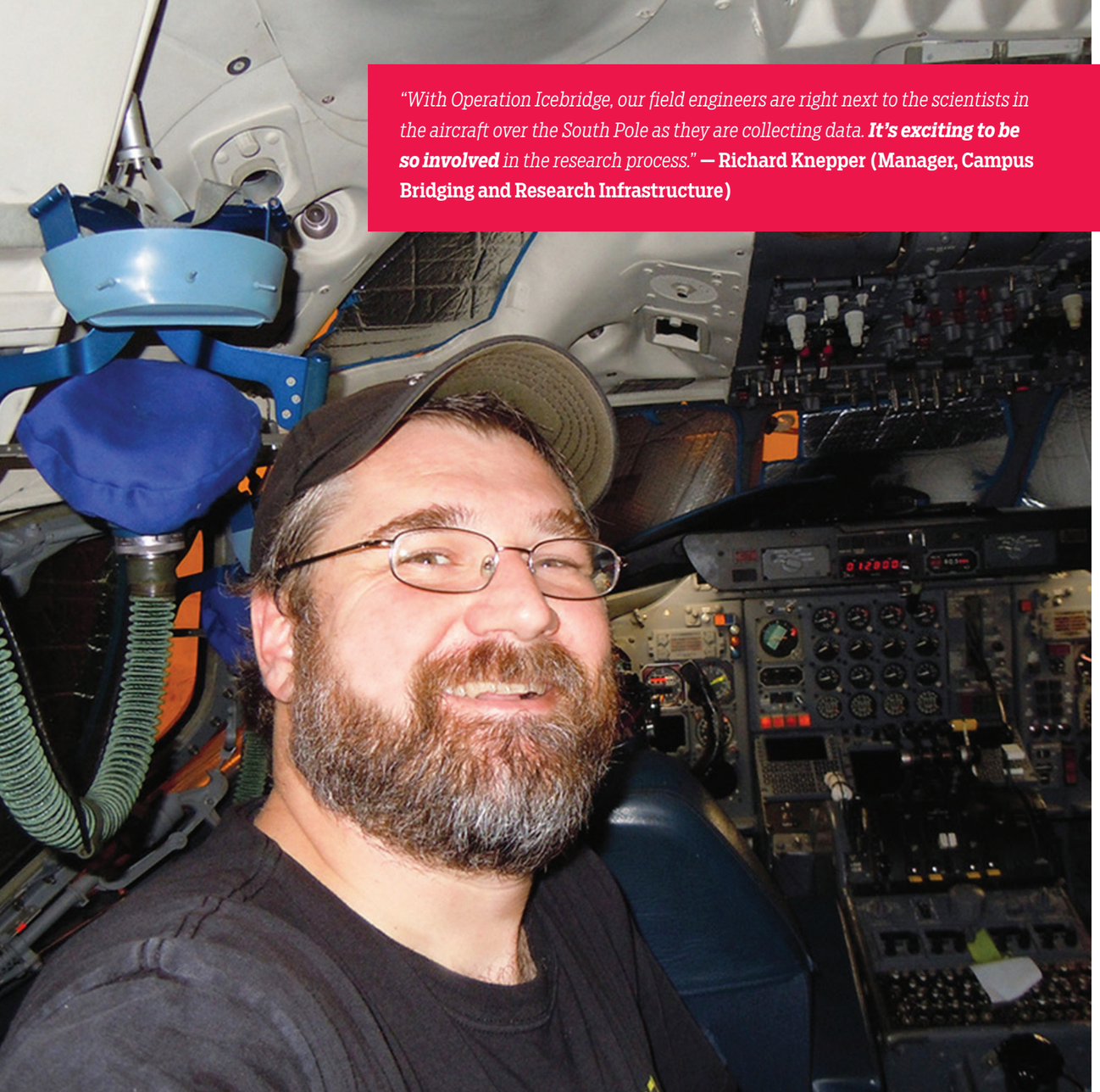
Why does matter have mass? Why does stuff hold together, rather than just flying apart into space? These questions have puzzled researchers for millennia. In 1964, Peter Higgs and others proposed a theory to explain mass, and predicted the existence of a certain type of particle that came to be referred to as the Higgs boson.

Physicists and information technologists at IU, experimenting for decades to confirm this theory, were among thousands of researchers who in 1984 decided to build a massive physics experiment facility called the Large Hadron Collider (LHC). As one of the largest scientific collaborations in history, the LHC aimed to collide subatomic particles together in ways that would either prove or refute the existence of the Higgs boson. Finding it meant confirming Higgs' theory of why matter has mass (and the fact that matter has mass is the reason we don't just fly apart into space).

Construction on the LHC began in 1998. IU's Harold Ogren, Professor of Physics at IU and a CERN Fellow, led development of a key component of the ATLAS detector — the Transition Radiation Tracker (TRT) — at Swain Hall on Third Street in Bloomington. The TRT, a barrel chamber that uses 52,544 tiny tubes to provide continuous tracking, provides measurements in individual tubes and electron identification based on transition radiation from fibers interleaved between the tubes.

Scientists recognized early on that one of the critical challenges in finding evidence of the Higgs boson would be analyzing the data produced by the LHC. To do this, the US Department of Energy (an LHC sponsor) created an international grid of computers to work together to analyze the data. This grid, called the Open Science Grid (OSG), has been in operation since 2005 — and is supported, monitored, and operated by staff of PTI/Research Technologies 7 x 24 x 365. The OSG provided billions of compute hours to sift through LHC data and find evidence of the Higgs boson.

Thanks to data produced by the LHC and analyzed by the OSG, the international physics community announced the discovery of the Higgs boson in 2012. Peter Higgs and François Englert were awarded the Nobel Prize in physics in 2013 as a result of this confirmation of their theory.



"With Operation Icebridge, our field engineers are right next to the scientists in the aircraft over the South Pole as they are collecting data. ***It's exciting to be so involved*** *in the research process."* **— Richard Knepper (Manager, Campus Bridging and Research Infrastructure)**

Climate: Getting polar data and keeping it safe

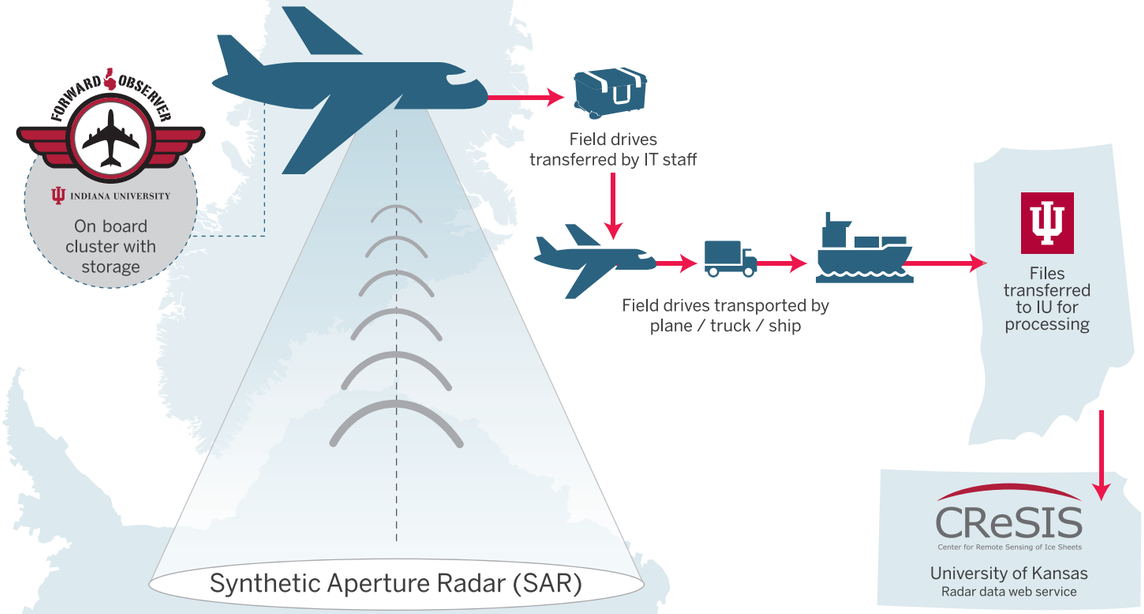
The science on global warming is clear: The Earth is warming up faster than before, and people are contributing to accelerated climate change globally. How much is the Earth changing and how fast? NASA's Operation IceBridge is working to determine year-to-year changes in polar ice.

Operation IceBridge uses sophisticated radar systems to measure the thickness of polar ice sheets from planes that fly over the Antarctic and Arctic glaciers.

PTI/Research Technologies helps Operation IceBridge collect data, keep it safe, and get it home to the US where it can be analyzed thoroughly. Twice a year, PTI/Research Technologies staff don their parkas and fly to the ends of the Earth to take care of this important data in person.

RT has even developed a flying supercomputer — the Forward Observer — that lets Operation IceBridge scientists analyze data in real time, as planes are flying over polar glaciers. This is a challenging, expensive, and dangerous mission, so it's important to capture high quality data and record it right to disk. On landing, data is copied two or three times over so nothing is lost between the Earth's poles and the IU supercomputer clusters enabling a closer look by NASA.

There is no simple answer to the issue of global warming. By measuring changes in polar ice sheets, PTI/Research Technologies and NASA are helping us better understand what's happening to the Earth, the consequences, and changes that may minimize negative effects.



IQ-wall at the Scholars' Commons

"A community is defined in large part by its commons — by the ideas its people hold dear and the resources they cherish and preserve. The Scholars' Commons ***embodies so much of what we as a community cherish****. It's easy to imagine this space supporting and nurturing our next Nobel Laureate."* **— IU Provost and Executive Vice President Lauren Robel**

IQ-wall at the Mathers Museum

Humanities: Creating interactive explorations

Research Technologies and its Advanced Visualization Lab's IQ systems bring the essential benefits of high-end facilities directly to labs, classrooms, studios, and galleries across IU. IQ-walls at the Mathers Museum of World Cultures and the Wells Library display collections of historic photographs, maps, and sheet music, as well as interactive information visualizations, media screenings, and group edit-a-thons for Wikipedia.

55-inch multi-touch IQ-tables have brought to life a digitized Audubon portfolio for the Lilly Library, treasures from the Mather's Museum for its 50th anniversary, a collection of rare maps in the Wells Library, and a "scrollable" version of the original 120-foot long manuscript for Jack Kerouac's "On The Road."

"You start out with empty space—infinite black space—and ***you have an opportunity to build a world*** *. The technical challenge is to build ... a world in which the visitor feels sufficiently immersed and willing to engage."* **— Margaret Dolinsky**

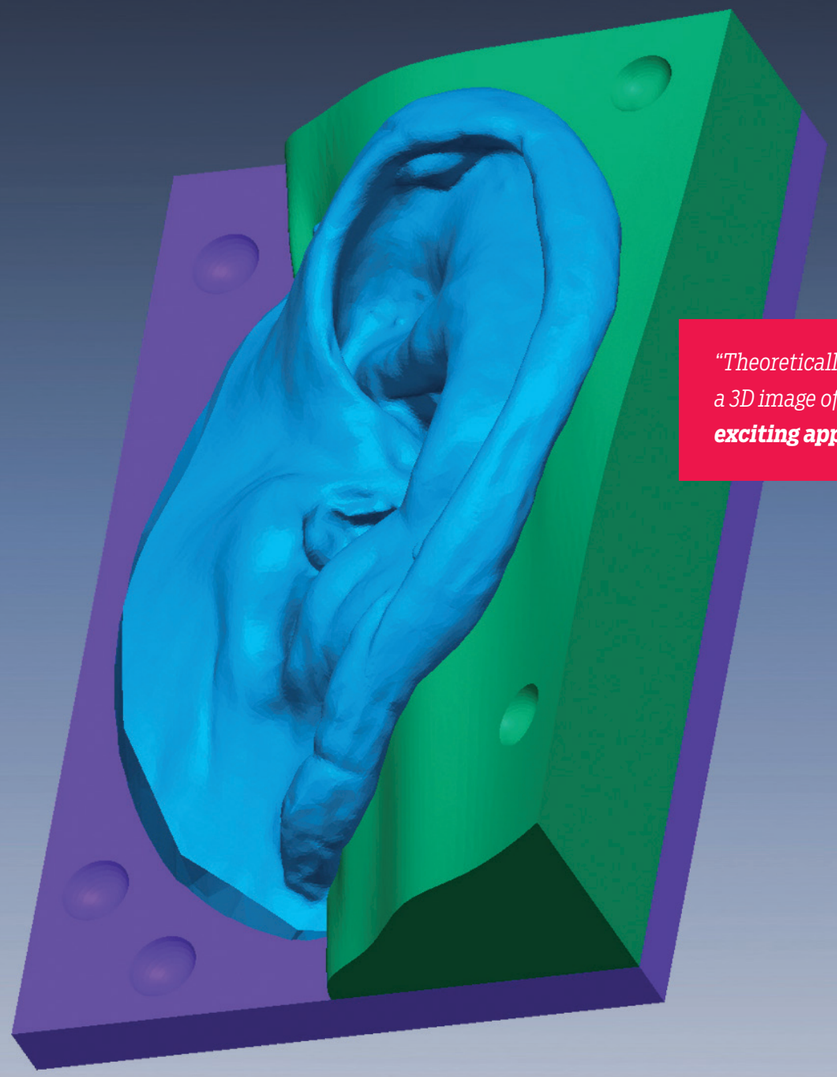
Dolinsky's *Hardwired* features faces that converse with you. The crowd of wire line drawings surround two colorful faces that conceal electronics. *Hardwired* "listens" to onlookers and engages in a musical conversation modeled after the character of the sounds or speech they hear.

Arts: Painting without the restrictions of reality

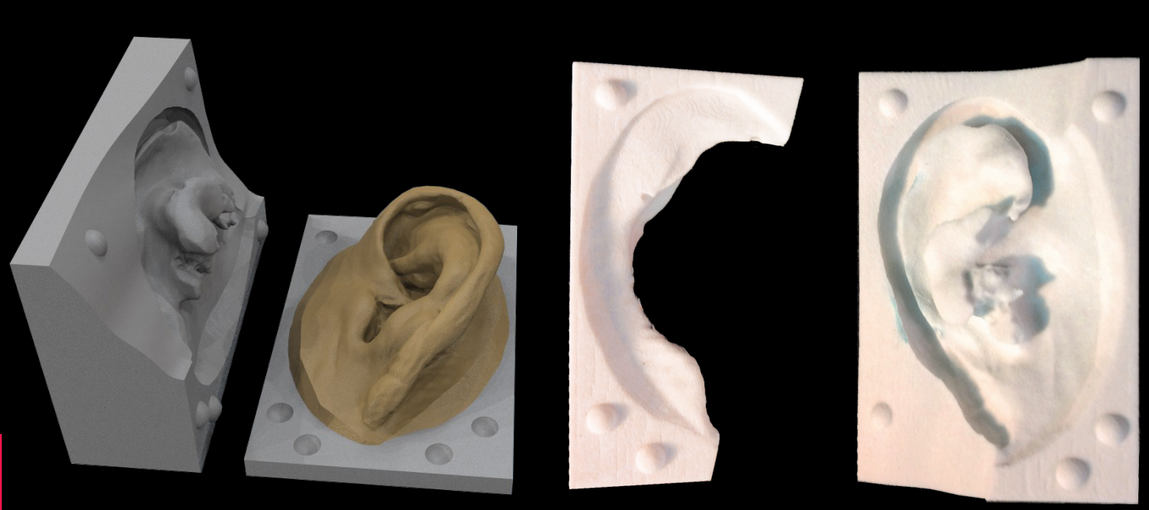
What would an artist do if given the chance to paint a work of art without the limits of reality? Could they create a piece that changes as people interact with it? The Advanced Visualization Lab (AVL) — part of PTI/Research Technologies — lets IU student artists do just that.

Students in Associate Professor Margaret Dolinsky's classes in the Henry Radford Hope School of Fine Arts study and apply the techniques and technology behind immersive virtual reality art. Dolinsky's students master software tools like 3D Studio/Max and Maya (tools likely to be found in Hollywood special-effects and animation studios) in order to translate their visions into virtual realities. AVL staff share expertise with state-of-the-art 3D gaming engines like Unity3D, helping students create interactive 3D environments ranging from fictional worlds to historical visualizations — on devices ranging from smartphones to sophisticated virtual reality displays.

Draft rendering of prosthetic mold.



"Theoretically you can print anything you have a 3D image of with a 3D printer. There are a lot of ***exciting applications****."* **— Victor Soon, Ph.D.**



Left: Final rendering of prosthetic mold. Right: 3D output printed from 3D renderings.

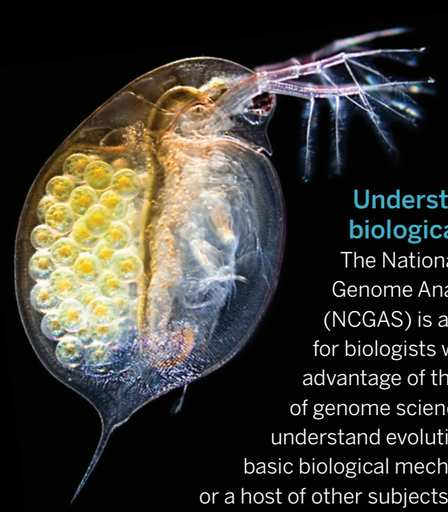
Broadening perspectives

Embracing 3D printing and visualization

Most people are familiar with the benefits of 3D imaging techniques like MRI and CT scans for diagnosing and treating internal problems and ailments. When coupled with other advanced analysis, visualization, and interface technologies, these imaging methods take on even greater value for improved health care research and delivery. Research Technologies (RT) visualization and analysis specialists have collaborated on several projects that highlight this potential.

- Current shoulder replacement techniques can affect mobility after surgery. A collaboration between IU School of Medicine researchers and RT visualization specialists uses CT scans and 3D visualization software to create a custom, per-patient guide. This improves placement of the prosthetic — and the patient's chances of a successful replacement.
- Tooth cavities are still a major problem for dental health. Dr. Masatoshi Ando in the IU School of Dentistry seeks to better understand the demineralization process that leads to cavities, and to evaluate new re-mineralization techniques. RT data analysis experts help Dr. Ando process the nearly 100,000 images a typical multi-sample, multi-week study requires.
- The sense of touch is essential to learning proper techniques for dental procedures like injections, examinations, and surgeries. Researchers in the IU School of Dentistry partnered with the Advanced Visualization Lab to explore applications of more realistic simulators that provide a sensation of force and touch.

"In the future we want to study and assemble clinical and socio-demographic risk factors, along with our blood tests, to ***increase our ability to predict risk****."*
— Dr. Alexander Niculescu



Understanding biological processes

The National Center for Genome Analysis Support (NCGAS) is a key resource for biologists who want to take advantage of the huge potential of genome science to better understand evolution, environment, basic biological mechanisms, diseases, or a host of other subjects that ultimately benefit our society.

First funded by the National Science Foundation in 2011, the center has grown to support National Institutes of Health and US Department of Agriculture projects — as well as a wide range of publicly funded research projects from across the nation. NCGAS sets itself apart by providing a complete genome analysis service for both experienced technical experts and scholars grappling with the complexities of DNA and RNA analysis for the first time.

Improving the state and nation's health

The Indiana Clinical and Translational Sciences Institute (CTSI) translates research discoveries into clinical trials and new patient treatments in Indiana and beyond. Ultimately, CTSI's goal is to accelerate the process of taking medicine from the lab bench to a patient's bedside.

The CTSI is a statewide partnership among Indiana University, Purdue University, the University of Notre Dame, IU Health, state agencies, and a number of public and private entities. Ultimately, it represents a coordinated effort to ensure medical research has a clear focus on directly improving health through both practice and policy. The CTSI hub, an online community resource, is used by more than 6,000 researchers and clinicians statewide. The cyberinfrastructure and research needs of Indiana CTSI are met by the Research Technologies Advanced Biomedical IT Core.

Discovery and validation of blood biomarkers for suicide

Currently, clinicians do not have a reliable test to detect the inclination or propensity towards suicide. To decide if someone requires immediate help and care, they rely on indirect tests that are error prone and subjective — and mainly based on of patient-reported symptoms and feelings. Developing an objective diagnostic tool would be useful in predicting and tracking suicidal tendencies, and could perhaps help clinicians treat patients earlier.

IU Associate Professor of Psychiatry Alexander Niculescu has been researching mood disorders over the past two decades, and suicide more recently. With the benefit of software expertise from the Advanced Biomedical IT Core, he was able to develop a scoring algorithm to identify biomarkers for suicide. Ultimately, this led to Niculescu's crucial discovery that SAT1 mRNA levels may serve as a biomarker for suicide when used in conjunction with other tests.

PTI/RT maintains leadership positions in a number of international organizations, through which we play a role in steering the directions of international research communities and collaborate with peers from many other countries.

Partnerships for resource sharing

Science gateways provide access to web-based portals, tools, applications, and data collections to further science — and help save time and money. Through them, researchers can access sophisticated resources, including applications running on a supercomputer, remote instruments (like a telescope or electron microscope), and curated data collections.

The NSF awarded a $5M, five-year grant to Indiana University, the San Diego Supercomputer Center, and the University of Texas Health Science Center at San Antonio to develop the Science Gateway Platform (SciGaP). The platform brings together hosted and cloud services, shielding researchers from the complexities of resources so they can focus on what counts — and collaborate and share results along the way.

Organization	Focus
HathiTrust Research Center	Provides tools for analysis of large-scale texts, and a portal from which to operate text analysis apps using the nearly 3M HathiTrust volumes in the public domain
Indiana CTSI Hub	The online portal for the CTSI, and one of the most widely used data access resources within the IU clinical and translational research community
Indiana Spatial Data Portal	Provides access to more than 30 terabytes of Indiana geospatial data; most datasets are publicly available for download and have no usage restrictions
ODI Pipeline, Portal, and Archive (ODI-PPA)	A comprehensive online solution that provides computational and visualization capabilities for scientists handling complex astronomical datasets
Cyberinfrastructure Gateway	An online portal designed to centralize information about and access to IU's advanced cyberinfrastructure
QuakeSim Science Gateway	Gives users simple access to sophisticated radar datasets, global positioning system time series data, and modeling tools for comparing earthquake fault models
UltraScan Science Gateway	Allows biophysicists to perform data analysis on analytical ultracentrifuge experiments, uncovering properties of molecules in a solution
NCGAS Galaxy Portals	Interfaces for genomics researchers to create and execute their own workflows, and to access genome assembly, annotation, alignment, and other apps for building workflows
BioDrugScreen XSEDE Science Gateway	A computational drug design and recovery resource for ranking molecules docked to the human proteome
IUScholarWorks	Operated by IU Libraries to serve as the university's primary, persistent digital repository

"We're thrilled that the kids are able to learn so much at this camp. They take away an idea of ***what high performance computing means****."* **—Robert Ping**

Towards innovation

Responsible stewardship

Part of the mission of PTI/Research Technologies is to educate and inform the community where we live and work, encouraging a culture of exploration.

The Ready, Set, Robots! summer camp helps us reach K-12 students and encourage early STEM interest. In addition, RT participates in multiple summer activities with other schools and departments to teach children that technology and math can be fun.

Our annual participation in state STEM events like Celebrate Science Indiana and Conner Prairie Curiosity Fair strengthens our commitment to helping society appreciate how everyday technology works — and its capacity for life changing results.

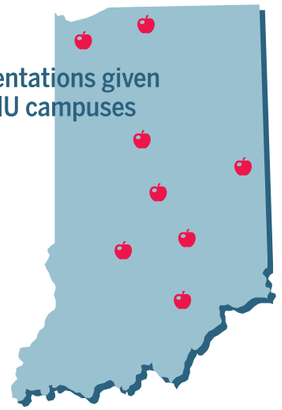
Campus initiatives like the Herron School of Fine Arts "Think It Make It" lab combine design, engineering, and technology for innovative product development — while new spaces like the Social Science Research Center bring together infrastructure, technology, and support to stimulate research.

At the same time, projects like the Energy Aware Lab (EAL) hold tremendous potential for sustainable and energy-efficient computing. The lab's completion in the IU Bloomington Data Center is a significant step forward in facilitating research into understanding power versus computation and into tuning the parameters to achieve the most efficient results.

9969 Total audience members

1,101 K-12

Presentations given at all **8** IU campuses



190 Education, Outreach, and Training events

24 Workshops

61 Talks, seminars, and presentations

105 Campus visits, tours, and online web training

Ready, Set, Robots!

Cyberinfrastructure Building
2709 E. 10th Street
Bloomington, IN 47408

Informatics and Communications Technology Complex
535 W. Michigan Street
Indianapolis, IN 46202

Communication Services Building
2715 E. 10th St.
Bloomington, IN 47408

researchtech@iu.edu