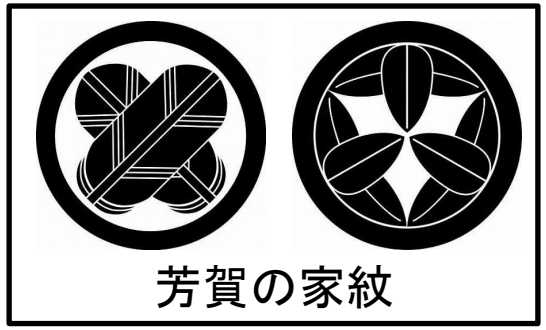
Immersive Visualization and Analytics to Derive Insight from Data

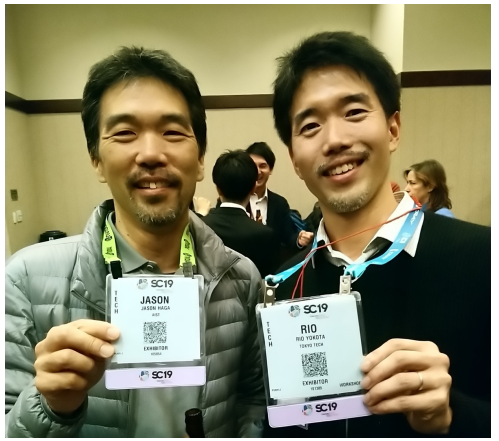
Jason Hideyo HAGA (芳賀ジェイソン英世)
Chief Senior Research Scientist
Continuum Computing Architecture Research Team
SAGE Evangelist
SAGE Development Team

National Institute for Advanced Industrial Science and Technology (AIST)
Tsukuba, Japan

Self Introduction

- Jason Hideyo HAGA
  - Chief Senior Research Scientist
  - SAGE Evangelist
- Started at AIST in 2014
- Moved from America
- Fun fact: Dr. Rio Yotota doppleganger according to Facebook





National Institute for Advanced Industrial Science and Technology (AIST)

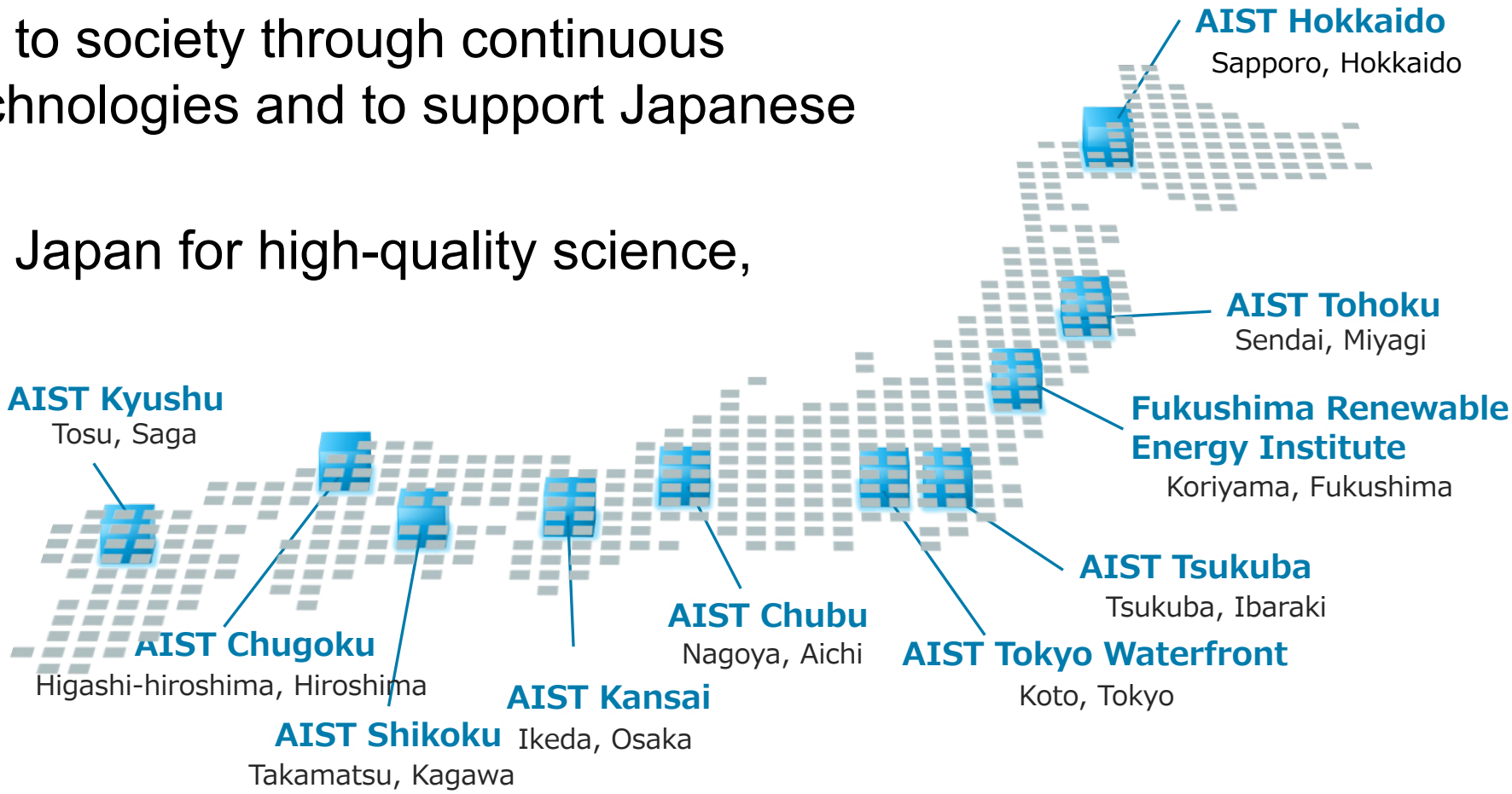
(http://www.aist.go.jp/index_en.html)

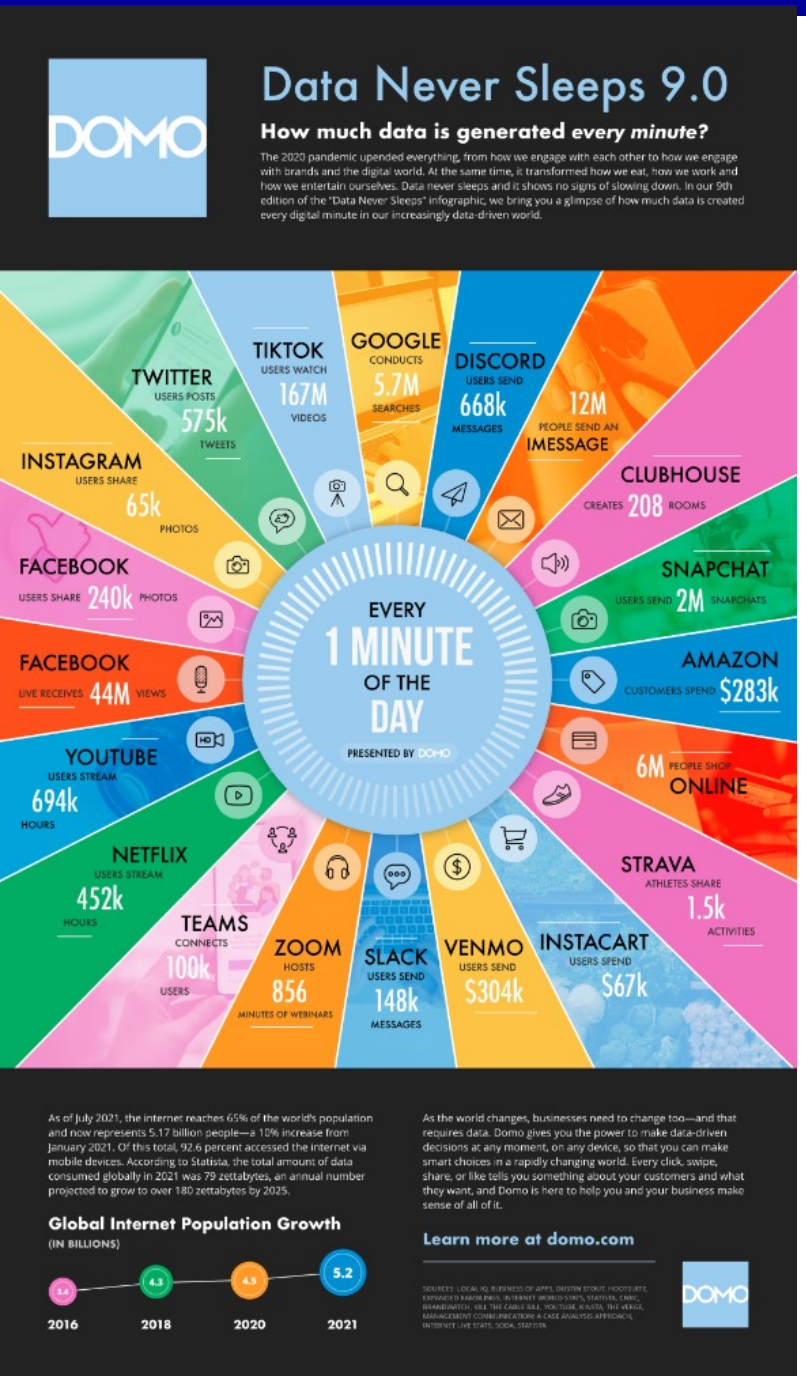
- Mission: Contribute to society through continuous advancement in technologies and to support Japanese industries
- AIST ranked 11th in Japan for high-quality science, Nature Index, 2018

Headquarters at Tsukuba, Japan

10 research bases in Japan

About 2300 researchers, 700 administrative personnel





Data is generated every minute

- In 2020, 1.7MB of data/sec for each person
- By 2020 **44 zettabytes** in digital universe
  - the number of bytes in the digital universe is 40 times bigger than the number of stars in the observable universe
- July 2021, 65% of global population uses the internet

- Data is the new currency
- Data is the new oil

- Drowning in data…
- If I could be in the data…

Data Visualization

Techniques to turn data into pictures to help people come to a clearer understanding and make better decisions.

From J. Leigh

Immersive Visualization/Analytics

Troves of big data are being created and are invaluable to all sectors of society, but deriving insight from multiple data sources is a challenge.

Therefore, we are using an interdisciplinary effort to create interactive environments that immerse the user in data and provide analytic tools.

Research Topics

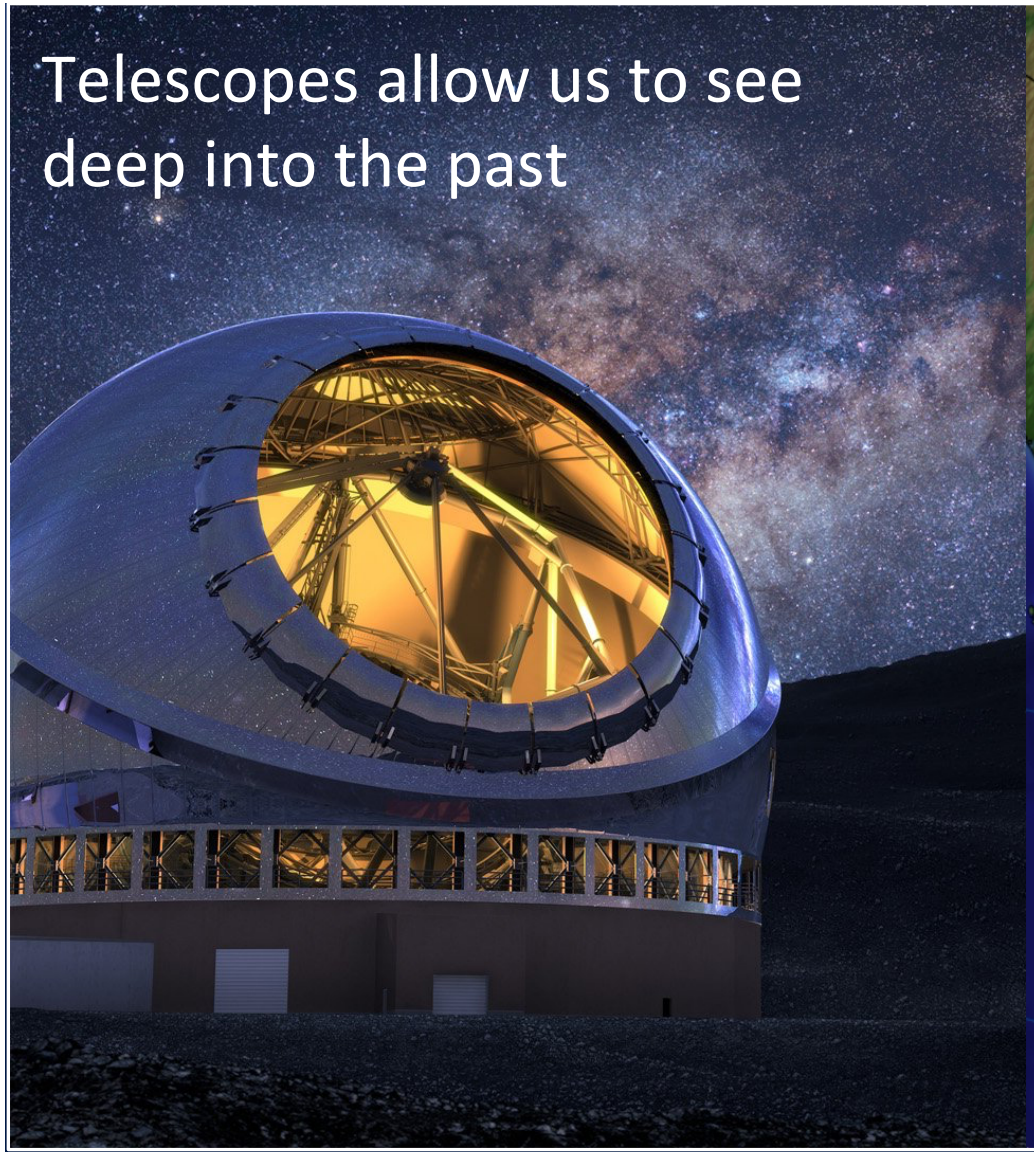
- Application-oriented approach to investigate the usability of new immersive visualization technologies for disaster management
- Explore combinations of 2D and 3D representations of data to reduce cognitive load
- Investigate the broader value of immersive visualization for different domains, because all data needs visualization at some level

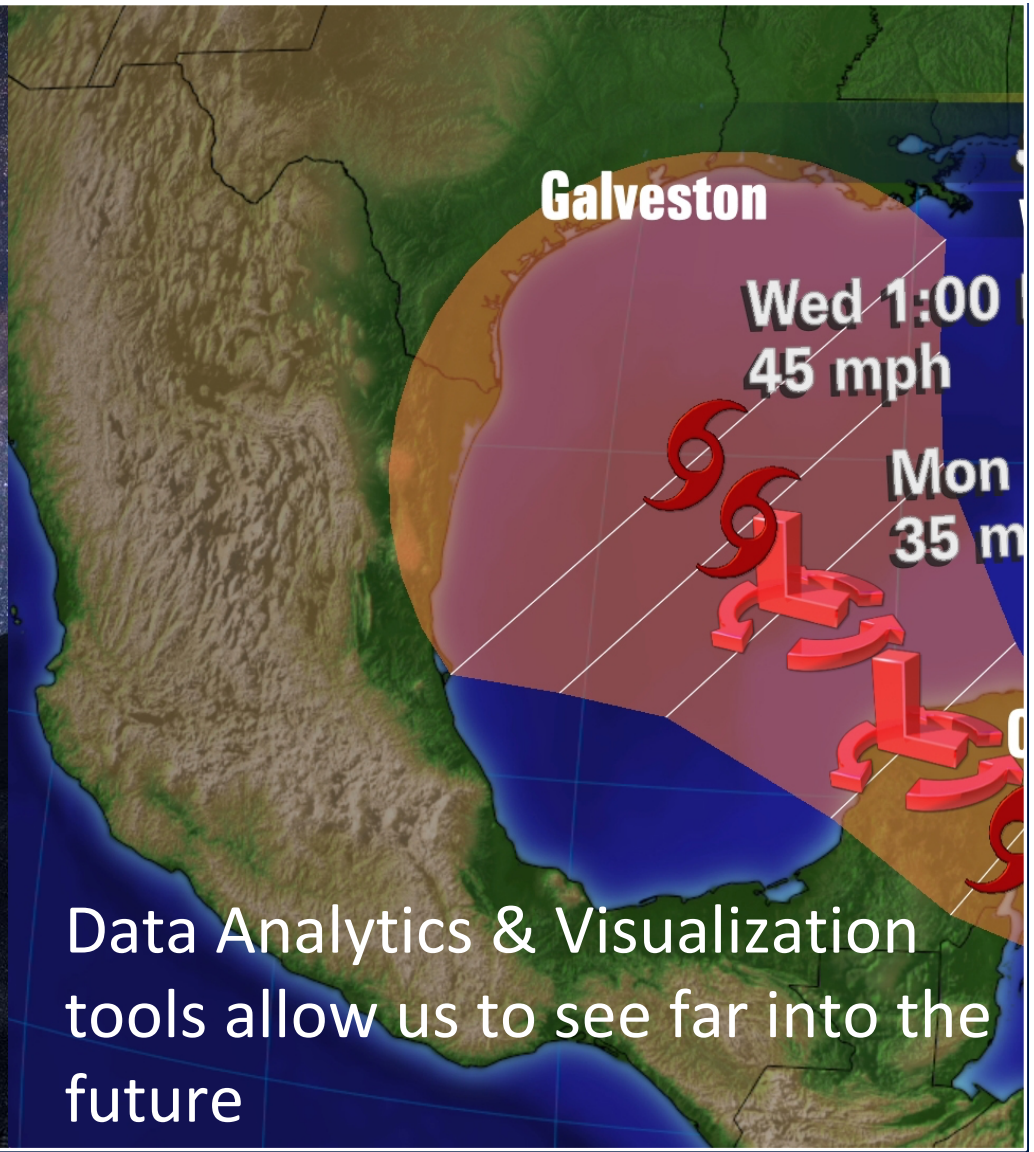
CyberCANOE: Cyber-enabled Collaboration Analysis Navigation & Observation Environment

The Canoe is the Island, The Island is the Canoe

CyberCANOEs allow users to come to decisions with greater speed, accuracy, comprehensiveness & confidence

Lenses for Bringing Data into Focus





Amazing Things Happen When You Put People In Front of Big Walls

- **See detail & context** simultaneously by reducing Window switching [Czerwinski03, Ball05]
- Helps **externalize the working memory of teams** [Andrews10]
- **Increases parallel processing** amongst team members [Park03]
- Improves **location memory** of information [Tan01]
- **Reduces gender performance gap** in navigation tasks in VR [Tan03]
- Helps user **performance keep pace** with increasing data size (perceptual scalability) [Yost07]
- Results in **greater confidence in conclusions** drawn when able to see all info at once [Ball05]
- Users **begin to look for higher level patterns** & relationships (i.e. they start to look for
- the bigger picture) [Reda12]

CyberCANOEs allow users to come to decisions with greater speed, accuracy, comprehensiveness & confidence

SAGE2 User Sites 2018 (100 = 47 International + 53 U.S.)

AUSTRALIA
- Monash University
- RMIT, (VX)Lab
- University of Sunshine Coast, Mechanical Engineering (3)
- University Southern Queensland
- University of Technology, Sydney

BRAZIL
- Bahia School of Medicine and Public Health
- Catholic University of Salvador (UCSal)
- Federal University Paraíba, LAViD
- Federal University of Rio Grande do Sul, PRAV
- Mackenzie University, LabCine
- National Institute of Space Research
- RNP, Rio de Janeiro (2)
- University of Campinas, Cinema
- University of Sao Paulo, LARC
- University of Sao Paulo, LASSU (2)

CANADA
- Ciena Research Labs
- Simon Fraser University, IRMACS

CHINA
- Chinese Academy of Forestry
- Shanghai University
- Tianjin University of Technology

CZECH REPUBLIC
- CESNET and Czech Technical University, SAGElab
- CESNET, Mobile SAGE
- Masaryk University, Cyber Exercise & Research Platform Project
- Masaryk University, Laboratory of Adv. Networking Technologies (2)
- Mavenir, Network Operations Center

FRANCE
- INRIA, ILDA

ITALY and SWITZERLAND
- University Urbino and ETH Zürich

JAPAN
- National Institute of Advanced Industrial Science and Technology (AIST) (2)
- NTT Network Innovation Laboratories, Yokosuka
- Osaka Univerisy, Cyber Media Center

KOREA
- Gwangju Institute of Science & Technology, Networked Computing Systems Lab
- KISTI, KREONET Center

NETHERLANDS
- Air France-KLM, CIO Group Technology Office
- SURFsara, Scientific Visualization Group, Collaboratorium
- University of Amsterdam, SNE

NEW ZEALAND
- REANNZ

SOUTH AFRICA
- University of Cape Town, Informatics and Visualisation Laboratory

TAIWAN
- National Center for High-performance Computing
- National Chung Hsing University
- National Museum of Marine Science and Technology

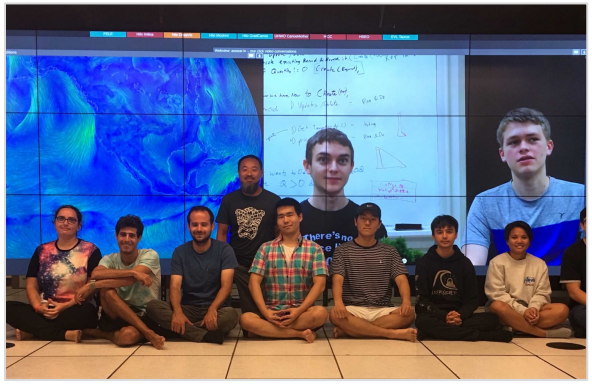
UNITED KINGDOM
- Imperial College London, Data Science Institute

UNITED STATES
- Adler Planetarium
- Argonne National Laboratory, ALCF
- Caterpillar Inc.
- Catherine Cook School
- Chaminade University of Honolulu (2)
- Digital Manufacturing and Design Innovation Institute (DMDII)
- Hawaii State Energy Office
- Honolulu Community College
- Jackson State University, ECE
- Kamehameha Schools
- NASA Marshall Space Flight Center, SPoRT
- NOAA, National Weather Service, OPG
- Northern Illinois University, Computer Science
- Northwestern University, iCAIR
- Stanford University, HIVE
- University of Alaska Fairbanks, DTN
- University of California, Merced, Library
- University of California, San Diego, Calit2-QI
- University of California, Santa Cruz, CITRIS/ Banatao Institute
- University of Chicago, RRC
- University of Florida Gainesville, ACIS
- University of Hawai‘i at Hilo (3)
- University of Hawai‘i at Mānoa, Applied Rsrch Lab
- University of Hawai‘i at Mānoa, Data Science Inst
- University of Hawai‘I Mānoa, HIGP
- University of Hawai‘i at Mānoa, Information Technology Center
- University of Hawai‘I Mānoa, i-LAB
- University of Hawai‘i at Mānoa, LAVA (3)
- University of Hawai‘I at West Oahu, Academy for Creative Media
- University of Illinois at Chicago, ACM/LUG
- University of Illinois at Chicago, Communications
- University of Illinois at Chicago, EVL (5)
- University of Illinois at Chicago, Engineering Maker Space
- University of Illinois at Chicago, Innovation Center
- University of Illinois at Chicago, Learning Sciences
- University of Illinois at Chicago, Ophthalmology
- University of Illinois at Chicago, Pathology (2)
- University of Illinois Urbana-Champaign, NCSA
- University of Maryland, Baltimore County, ARC
- University of Oregon, Library
- University of Pennsylvania, Idea Factory
- University of St. Thomas
- University of Texas, Austin, TACC

SAGE2 User Sites 2021 Examples

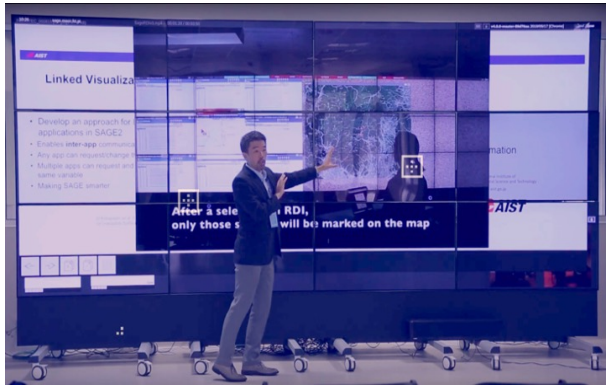
JAPAN, National Institute of Advanced Industrial Science and Technology (AIST)

USA, University of Chicago



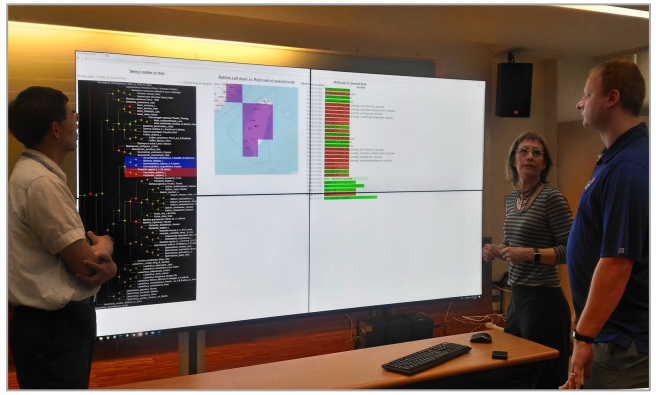
USA, University of Hawai‘i at Mānoa, Data Science Institute

CANADA, University of Toronto



PORTUGAL, INESC-TEC

JAPAN, University of Tokyo



TAIWAN, National Museum of Marine Science and Technology

AUSTRALIA, Monash University

CyberCANOE Hardware

- Lowest - $4,500 (maybe less)
- Low - $7,500
- Med - $12,000
- High - $27,000 (may need special installation)

- Configurations are highly dependent on available space, budget, intended audiences, and usage scenarios

CyberCANOE Hardware

- Lowest End
  - **Three 75-inch LCD TVs**
  - LG Electronics 75UJ6470 75" 4K Ultra HD Smart LED TV ($1,500 X 3 = $4,500)

CyberCANOE Hardware

- Low End
  - **Three 86-inch LCD TVs**
  - Sony XBR85X900F 85-Inch 4K TV (~$2,500 each x3 = $7,500)

CyberCANOE Hardware

- Mid End
  - **8 Small bezel LCDs**
  - Samsung UD55E-B UDE-B (~$1,500 each x8 = $12,000)

CyberCANOE Hardware

- High End
  - **12 or 18 Small bezel LCDs**
  - Samsung UD55E-B UDE-B (~$1,500 each x18 = $27,000)

Additional Equipment

- PC budget is about $3,000-$5,000
  - Recommend SSD in the PC
  - The number of graphics card and type will depend on which of the above configurations you go for
    - 3 displays: NVIDIA 1080 or better
    - 4x2 displays: 2x NVIDIA 1080Ti or better
    - 6x3 displays: 2X NVIDIA 1080 or above. 6 x Matrox TripleHead2Go T2G-DP-MIF DP Edition
- Video Conferencing
  - Logitech Group Conferencing ($1,300)
  - Tripod
- Miscellaneous cables about $500 (Amazon)
- Networking – 100Mb/s and up
- Warranty – 3-year warranty, 4 is better

CyberCANOE Mobility

- Peerless 4x2 Display Stand ~$3,500

- Peerless 4x3 Display Stand ~$4,500

CyberCANOE Software

- Scalable Amplified Group Environment (SAGE) created in 2004 as part of OptIPuter project.
  - Currently in 17 countries. ~4000 users, ~800 sites.
  - In 2017 61% of sites had 1 tiled display. 39% managed more than one. 77% of walls used several times a week.
  - Meeting sizes ranged from 2-200, 20 on average.
  - Disciplines: Archaeology, Architecture, Art, Atmospheric Science, Biology, Chemistry, Civil Engineering, Communications, Computer Science, Education, Geoscience, Health, Library Science, Mathematics, Medical, Meteorology, Network Engineering, Neuroscience, Physics, Psychology, and Statistics.
  - Open-source middleware
- Different from current video conferencing applications
  - WebEx, GotoMeeting, Skype, Hangouts, etc.
  - Provides multiple users with a common operating environment enabling parallel interaction with data
- Smart Amplified Group Environment (SAGE3)
  - $5 million NSF Grant in May 2020
  - Open-source
  - Alpha out now - https://sage3.sagecommons.org/
  - Watch this space!

Evolution of the World from SAGE to SAGE3

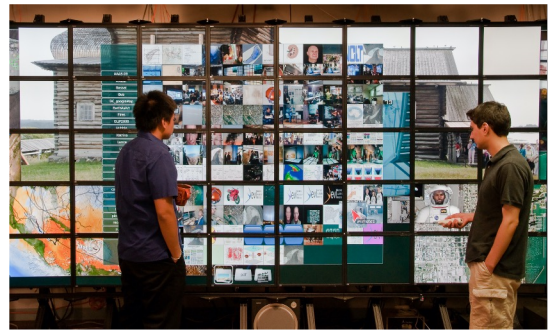
(from J. Leigh)

2004

(C++)

Compute Cluster
$500K-$1M
systems

Grid Computing



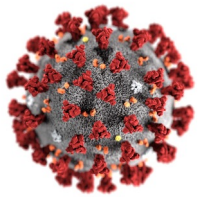
2014

(Javascript)

Single PC
$100K-$300K
systems

Science Portals &
Gateways
Cloud Computing

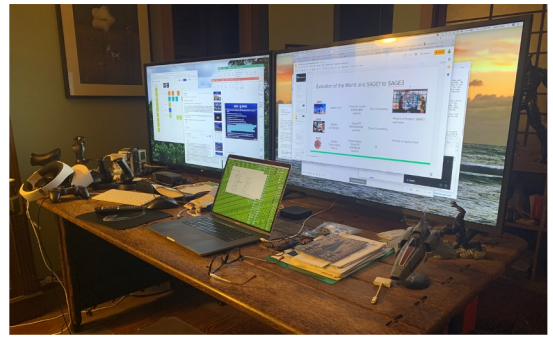
2020



(Typescript & Python)

Single Laptop -
Single PC
$3K-$100K systems

Data Science
Notebooks,
Cloud AI &
Containers
NSF Cyber
Ecosystem

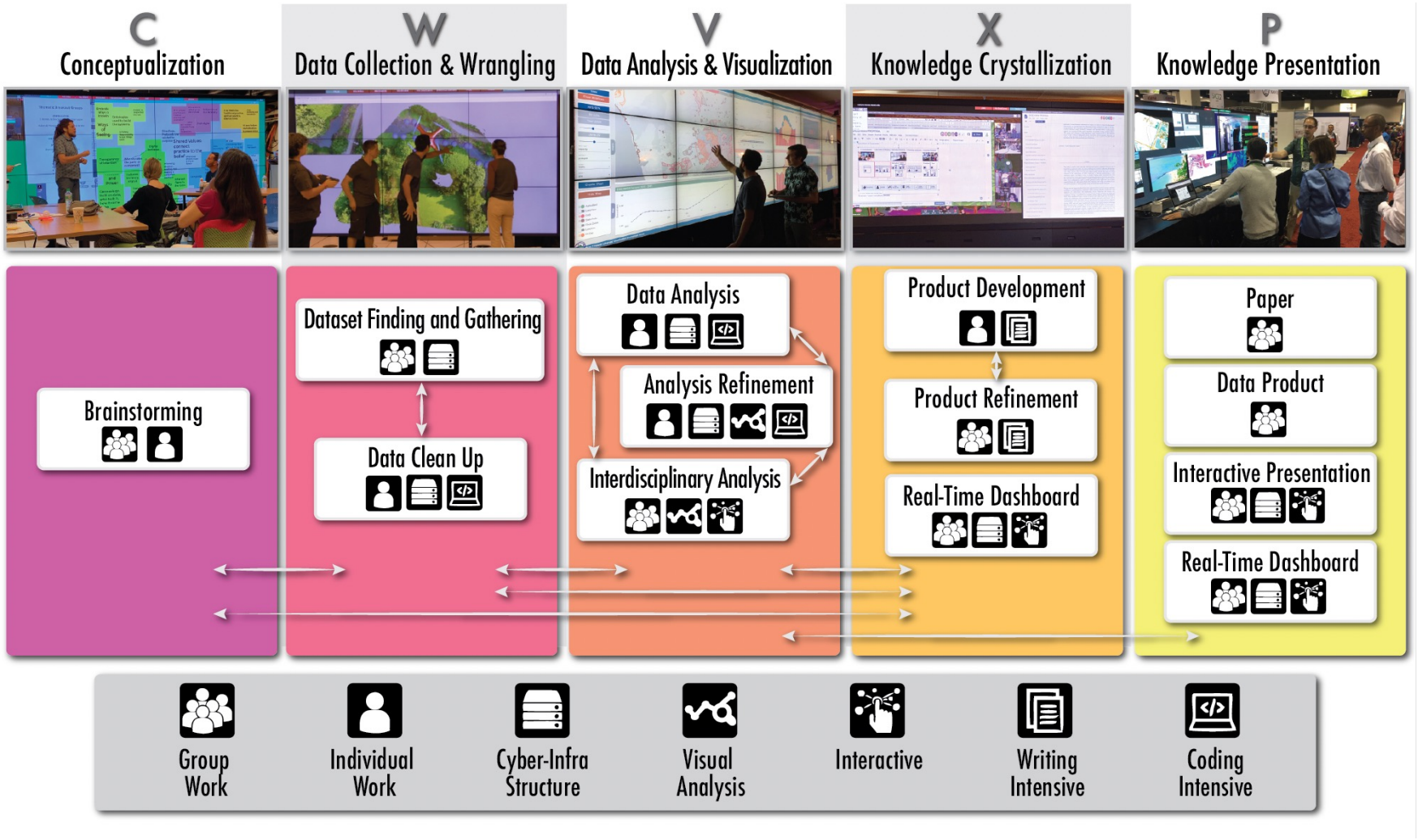


SAGE3 vs. SAGE2

- SAGE3
  - Replicates SAGE2 functionality, but more seamless transition between large walls, working on laptops, and home or office
  - Open-source, uses React js framework
  - Immersive visualizations infused with AI analytics
  - Re-designed user interface for post-pandemic reality
  - Lowering the barrier for non-AI experts to work with models, data, and visualizations
  - SAGE3 source code not released yet

- SAGE3 Introduction
  - https://sage3.sagecommons.org/

Usage Patterns of Wideband Display Environments In e-Science Research, Development and Training



Leigh, et al., eScience 2019, San Diego, CA, USA

SAGE2 Enabled Brainstorming Session

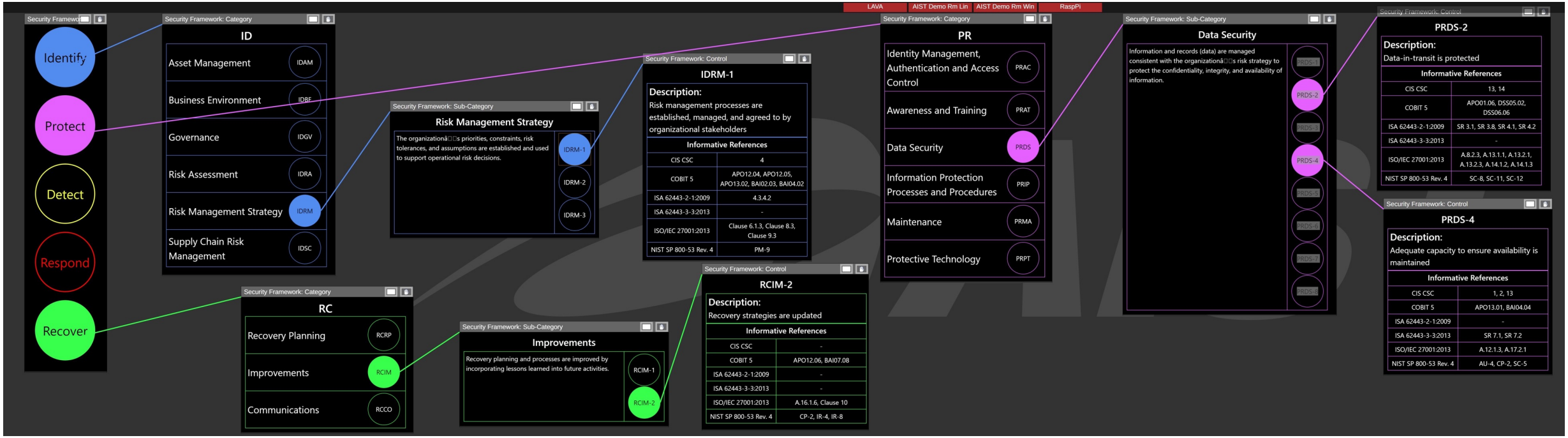
SAGE River Disaster Information

- Challenging to provide detail and context with heterogeneous datasets
- Created a framework for linking visualizations
  - Enables **inter-app** communication through data requests/changes
  - Works as a publish / subscription system
- First step to making SAGE smarter

D Kobayashi, et al. ACM International Conference on Interactive Surfaces and Spaces (ISS '18), pp 33-42, Nov 2018

SAGE RDI Movie Demo

SecuritySAGE

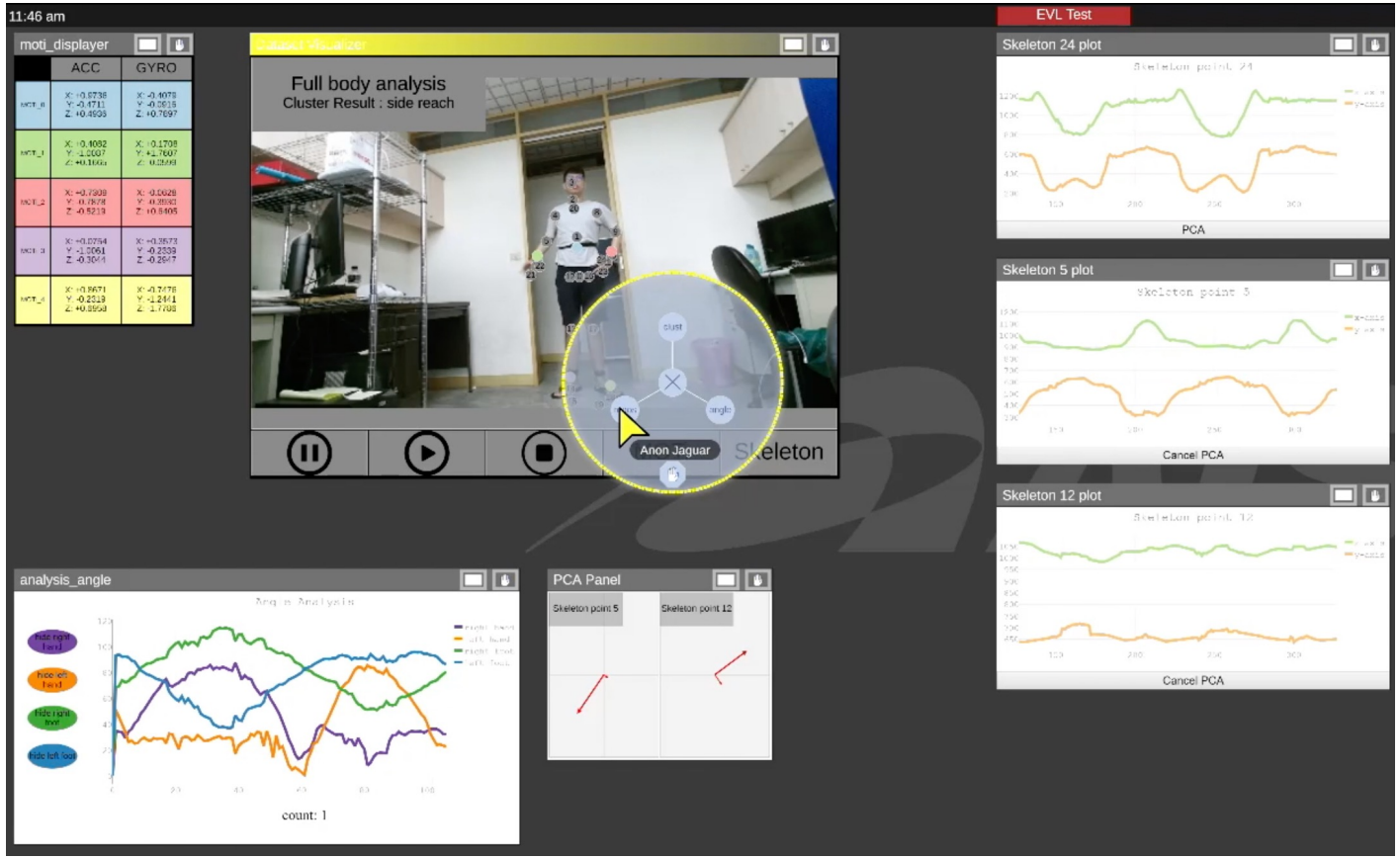


- Cybersecurity standards are large documents used by many corporations
- Challenge to understand and compare large volumes of information in standards
- Created a SAGE2 native application to better enable stakeholders in their decision-making process
  - Based on NIST Cybersecurity Framework
  - Provides a more interactive mode of user interaction with content

SecuritySAGE App Demo

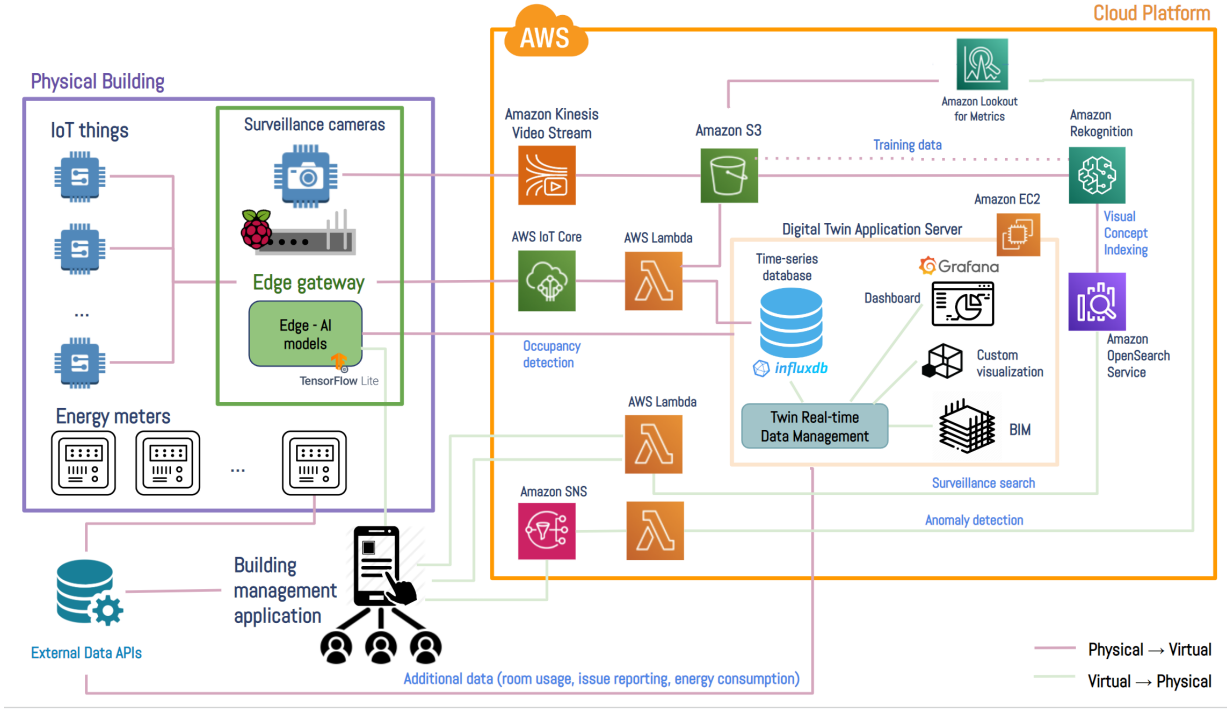
ExerciseSAGE

- Collaboration with NYCU (Dr. Chih-Wei Yi)
- Sports/exercise medicine is a very data intensive domain
- Challenge to integrate these data for analysis
- Analyze movement patterns during exercise
  - Assess level of reproducibility of movement to encourage proper exercise
  - Compare the ability of several subjects for sports teams



ExerciseSAGE Movie Demo

Smart Building Digital Twin



- Collaboration with Thammasat University
- Currently, building models lack real-time information and no exemplars for smart building digital twins that take advantage of large display spaces
- We created an end-to-end prototype SAGE2 app for optimizing building management and utilization

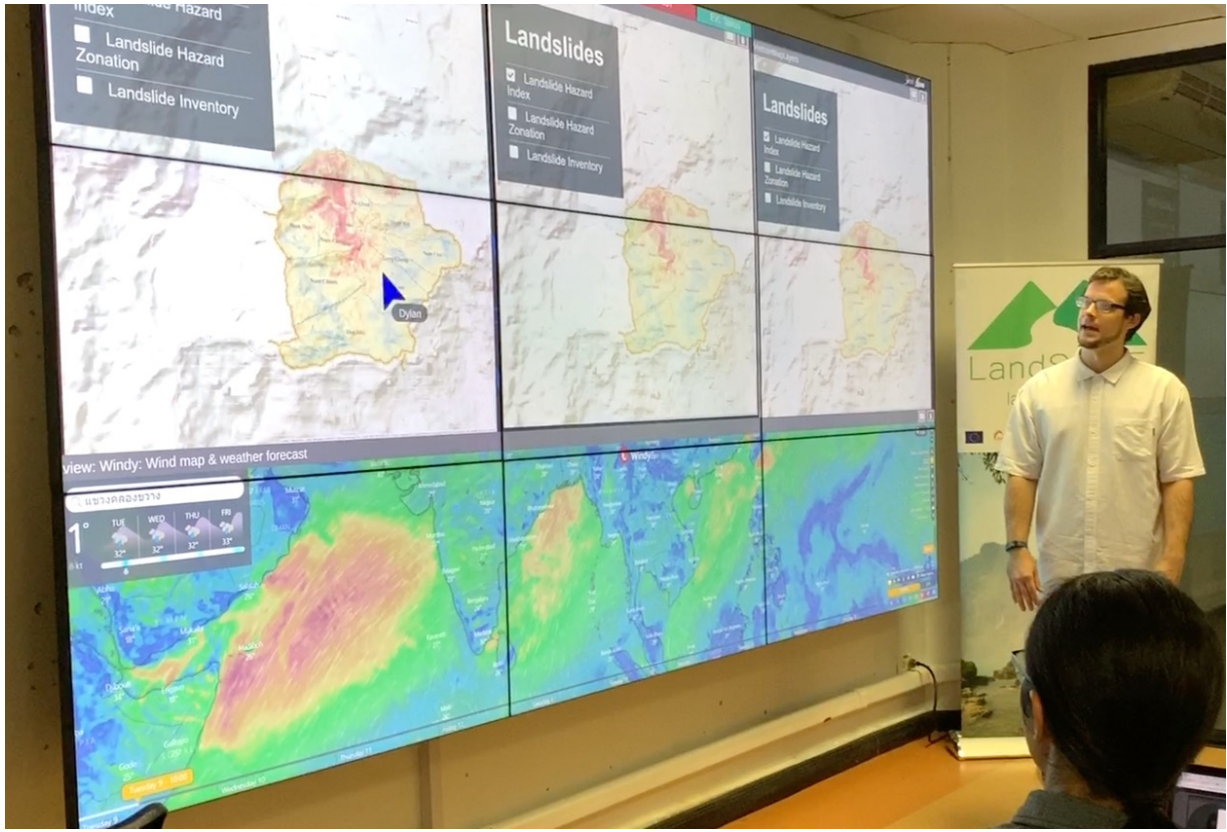
To appear WFIoT 2022, Oct 31-Nov 4, Yokohama, Japan

SAGEBuilding App Demo

Visualization-rich decision support for monitoring and mitigation of natural disasters (landslides, mudflows, and floods) in Southeastern Asia

LandSAGE
landsage.info

- Visualization capacity building that includes construction, deployment and use of CyberCANOE in research and development for landslide decision making and mitigation.
- CyberCANOE is a big data networked collaboration platform that allows teams of researchers to work over distance to make sense of large amounts of complex and disparate data.
- Collaborators from Laos, Cambodia, Vietnam, Thailand, Indonesia, Philippines, Japan and US.

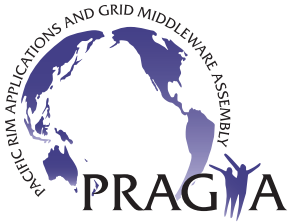


Immersive Visualization and Analytics Interdisciplinary Research Framework

- Form novel partnerships
  - IT researchers
  - Subject matter experts
  - Students
  - Industry
- Share resources and expertise
  - Technologies and infrastructure
  - Data collections (public and private)
  - Develop and deploy prototypes
  - Cultural differences in approaches to data management
- Investigate research questions
  - Data integration and accessibility
  - Application and analysis requirements
  - Usability studies
  - Approaches to user engagement
  - Impact of technology on society

Acknowledgments

AIST International Collaboration Unit
PRAGMA
University of Hawaii, Manoa
Thammasat University
National Yang Ming Chiao Tung University
Asi@Connect (TEIN*CC)

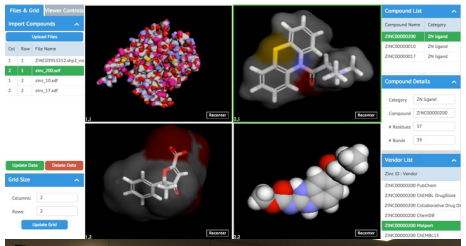


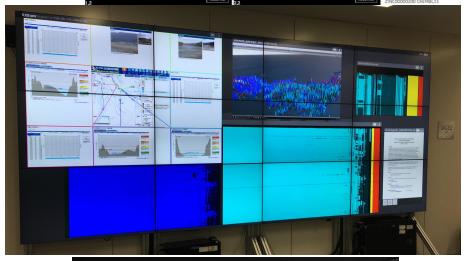
AIST Internship Program

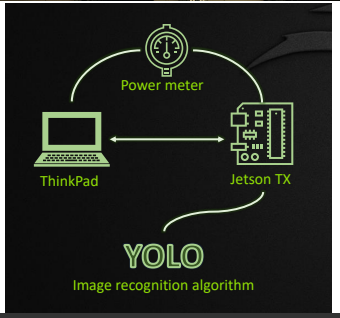
- Science is becoming **increasingly global**, especially with **information technologies**
- It is critical that the **next generation** of scientists be both **technically and culturally prepared** to enter the 21st century workforce
- To address this **challenge**, we are creating an **international hub** for research and cultural training for students
  - Provide real-world research experiences and training for the next generation of scientists
  - Expose students to international cultural awareness when engaged in global science
  - Increase AIST global presence in research and training

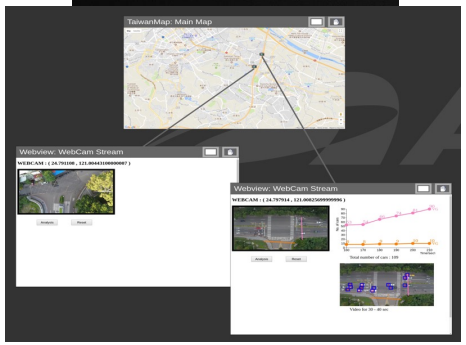
Project Focus Areas

- Immersive Visualization
  - Hydra: A High-Throughput Virtual Screening Data Visualization and Analysis Tool
  - Virtual Reality for Natural Disaster Management
  - Visualization Application for Medical Research Data
  - EddvisAR – Augmented Reality in Web Browser Technology
- Networking for IoT Era
  - Toward Fast and Scalable Key-Value Stores Based on User Space TCP/IP Stack
  - µmq – A Lightweight and Scalable MQTT Broker
  - MQTT RasPI-Cluster with NGINX Load-balancing
- Artificial Intelligence
  - Context Sensitive Fact Extraction With Machine Learning
  - Machine Learning to Support Distributed Query Processing Over SPARQL Endpoints
  - NVIDIA Jetson TX1 and Jetson TX2 comparison on image recognition
  - Integrating vehicle traffic analysis system with machine learning in SAGE2
- Cybersecurity









And many other IT related projects…

Requirements

- Minimum 3.0 GPA or equivalent
- Undergraduate or graduate students
  - Sophomore (2nd year) undergraduates preferred
- Work on project before and after internship period
  - 10 weeks in Japan
  - One year activity ideal
- Intellectual interest in the project
  - Bring your ideas!
- Japanese language is not a requirement

Participating Institutions and Programs

- PRAGMA, USA
- Mahidol University, Thailand
- Thammasat University, Thailand
- University of Hawaii, Manoa, USA
- INESC-TEC, Portugal
- NYCU, Taiwan

Become part of a global network!

Contact Jason Haga (jh.haga@aist.go.jp)

CyberCANOE Software

- Scalable Amplified Group Environment (SAGE) created in 2004
  - SAGE2 rewritten in 2014 leveraging on web browser and cloud technologies
  - Open-source middleware
    - Provides multiple users with a common operating environment
    - Access, display, and share heterogeneous, data intensive information
- Different from current video conferencing applications
  - WebEx, GotoMeeting, Skype, Hangouts, etc.
  - Enables parallel interaction with data

- Smart Amplified Group Environment (SAGE3)
  - $5 million NSF Grant in May 2020
  - Open-source
  - Alpha out now - https://sage3.sagecommons.org/
  - Watch this space!