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

Jason Leigh, Mahdi Belcaid, Dylan Kobayashi, Nurit Kirshenbaum, Troy Wooton, Alberto Gonzalez
Laboratory for Advanced Visualization & Applications, Hawaii Data Science Institute, University of Hawaii at Manoa, Honolulu, Hawaii, USA
leighj@hawaii.edu

Luc Renambot, Andrew Johnson, Maxine Brown, Andrew Burks, Krishna Bharadwaj, Arthur Nishimoto, Lance Long
Electronic Visualization Laboratory, University of Illinois at Chicago, USA
renambot@uic.edu

Jason Haga
National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan
jh.haga@aist.go.jp

John Burns, Francis Cristobal, Jared McLean
University of Hawaii at Hilo Hilo, Hawaii, USA
frcristo@hawaii.edu

Roberto Pelayo
University of California, Irvine, Irvine, California,USA, rcpelayo@uci.edu

***Abstract* — SAGE (the Scalable Adaptive Graphics Environment) and its successor SAGE2 (the Scalable *Amplified Group* Environment) are operating systems for managing content across wideband display environments. This paper documents the prevalent usage patterns of SAGE-enabled display walls in support of the e-Science enterprise, based on nearly 15 years of observations of the SAGE community. These patterns will help guide e-Science users and cyberinfrastructure developers on how best to leverage large tiled display walls, and the types of software services that could be provided in the future.**

Keywords—tiled display wall, immersive analytics, visualization, human centered computing, computer supported cooperative work

I. Introduction

SAGE (Scalable Adaptive Graphics Environment) [9] and its successor SAGE2 (Scalable *Amplified Group* Environment) [15, 25] are today's de facto operating systems to manage Big Data on large tiled display walls (generally called wideband display environments), enabling users to collaborate locally and remotely, and to access, display, share and manipulate documents and visualizations [14]. SAGE's ease of use makes it an excellent platform on which to display a variety of related, high-resolution information, enabling collaborators to reach conclusions and make decisions with greater speed, accuracy, comprehension, and confidence.

This paper briefly introduces SAGE's capabilities and user community, and documents prevalent usage patterns in support of the e-Science research, development and training enterprise.

We hope it serves as a guide for new SAGE users to understand how to best leverage SAGE's capabilities to support the various phases of the e-Science enterprise, as well as a means to stimulate cyberinfrastructure researchers and developers to develop new end-user services for use on wideband display environments – which are fast becoming a standard in universities, national laboratories, companies, and non-profit organizations.

II. The e-Science Research, Development and Training Enterprise

Great progress is being made in the development of Workflows and Gateways (such as Jupyter Notebooks and Agave) to support scientific analysis [7, 20]. While these are crucial to e-Scientists, there is also a need to provide broader end-to-end support of the scientific enterprise that involves the "human in the loop."

Figure 1 summarizes the *kinetics* of the scientific research, development and training enterprise based on our years of experience researching, developing and deploying cyberinfrastructure tools for domain scientists.

The Phases include: (C) conceptualization/hypothesis generation, (W) data collection and wrangling, (V) visual analysis, (X) knowledge crystallization, and (P) knowledge presentation. Workflows and Gateways primarily fall under W & V. However, research, development, and training frequently transition between all phases, alone or in groups, particularly in the early phases as a result of many micro-failures before achieving and reporting success in X & P. SAGE is being used to support collaborative work in all of these phases, as we will elaborate in the following sections.

III. SAGE and Its Community

SAGE has become the community standard for cyber-enabled, data-intensive, visual analytics and collaboration. The SAGE project started in 2004, and in 2014 we introduced SAGE2, a web-based system. Since then there have been 5,676 downloads of the SAGE2 server, and approximately 1500 weekly end-users installing the SAGE2 plug-in that is required to connect to the servers. We estimate there have been over 600 administrators in charge of wall deployment in over 16 countries, with an estimated hardware infrastructure investment of $120M. Table 1 lists the 100 most recent installations, where each installation, or site, has many users.

978-1-7281-2451-3/19/$31.00 ©2019 IEEE
DOI 10.1109/eScience.2019.00041

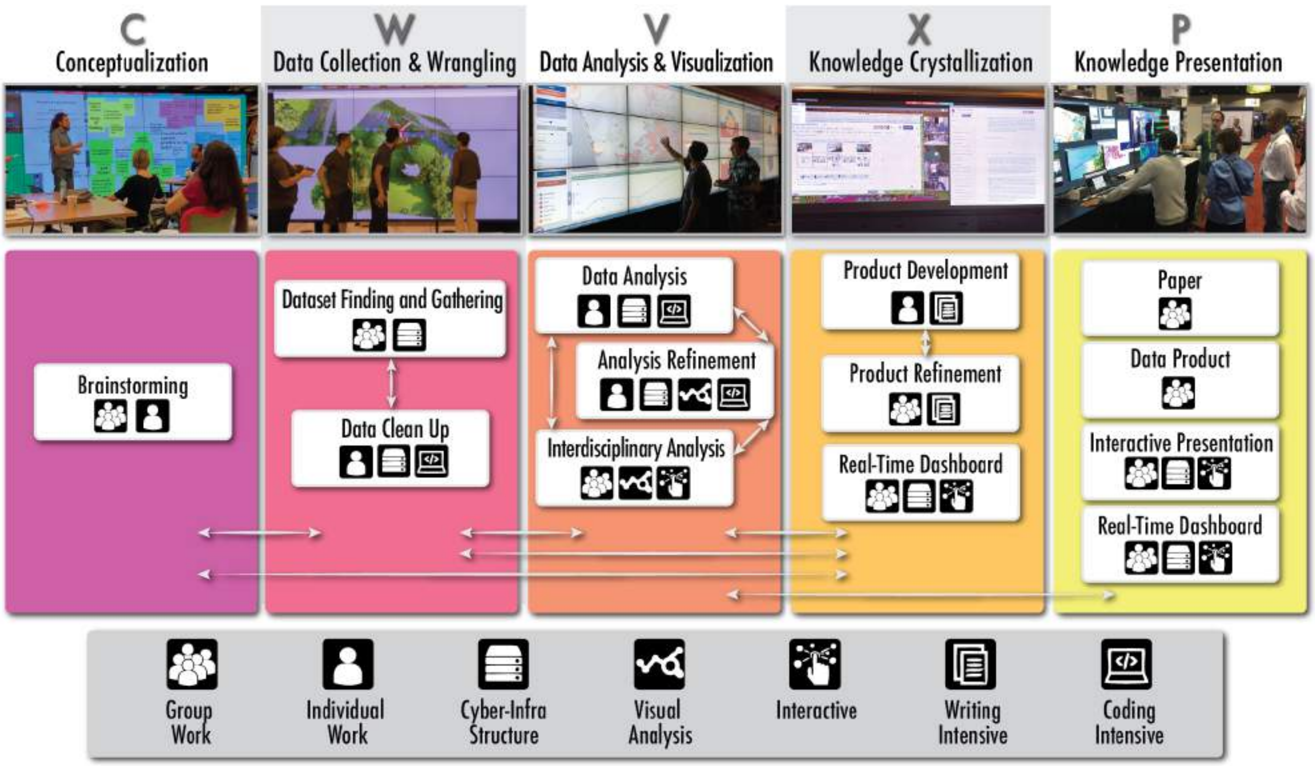

Figure 1: The e-Science Research, Development and Training Enterprise.

TABLE I. 100 RECENT INSTITUTIONS USING SAGE2.

SAGE2 User Sites 2019 (102 = 48 International + 54 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 Paraiba, 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 (2)
- 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 University, 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 Vis 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

THAILAND
- Mahidol University, Faculty of ICT

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
- Hawaii Community College–Palamanui
- 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/UIC
- University of Illinois at Chicago, Communications
- University of Illinois at Chicago, EVL (5)
- University of Illinois at Chicago, 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

Selected SAGE2 Walls

U.S., University of California Merced, Kolligian Library | Australia, Monash University, Immersive Visualization Platform | U.S., NOAA, National Weather Service, Operations Proving Ground

Figure 2: Examples of SAGE2 Walls

A 2017 community survey showed that the institutions were 67% from academia, 11% from government labs, 9% from industry and 13% other. In this worldwide community, 61% of the sites managed one tiled display wall, and 39% managed more than one. 77% of those who responded to our survey said their walls were used several times a week, with 20% using their walls as often as several times a day. Meeting sizes ranged between 2 and 200, with 20 being the average. Furthermore, the disciplines that use it include: 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.

Users treat SAGE display walls as collaboration canvases on which to juxtapose and manipulate a variety of data types and visualizations to help discover underlying trends. For example, users can spread time-series data over an entire display wall, with each visualization depicting a timestep, or they can stream large-scale visualizations from supercomputers and video conferencing systems to the wall to support distance collaboration. These are but some of the features that make SAGE a leading technology for harnessing the data revolution.

SAGE's widespread adoption resulted from rapidly decreasing hardware costs (a SAGE wall can be constructed with as little as $10,000 today compared to $500,000 in 2004); growing numbers of web applications that interoperate with SAGE2 (e.g., Jupyter Notebooks, ParaViewWeb, Google Docs and Hangout, ARCGIS, PubMed, VTK.js); a community-driven App Store [27, 28]; a library of training videos for end users [29]; and, licensing to companies to provide fully installed display walls with 24/7 support.

However, the most important reason for SAGE's rapid adoption is that it empowers users with the ability to work collaboratively with local and remote colleagues on large datasets and data types to come to conclusions and make decisions with greater speed, accuracy, comprehensiveness, and confidence [1, 2, 3, 4, 5, 6, 18, 21, 33, 34, 37].

IV. SAGE Usage Patterns

A. Activities that can Take Place within the Five Patterns

Over the course of the 15 year lifespan of SAGE, we have observed a number of prevalent patterns of its use in the five phases of the e-Science enterprise. The notion of Design Patterns [10] revolutionized Software Engineering by codifying recurring data management and manipulation patterns into highly reusable software architectures. Similarly, the notion of Visualization Thinking Design Patterns (VTDP), first articulated by Colin Ware [36], describes "tools to help with the design of cognitively efficient visualizations. They describe human-machine cognitive processes that are executed when interactive data visualizations are used as cognitive tools." Our SAGE Usage Patterns can be thought of as expanding Ware's concept of VTDP to incorporate Design Patterns for Collaborative Work in Wideband Display Environments [22, 23, 24].

At this point it is worth noting the icons in Figure 1, as they summarize the kinds of activities that occur within SAGE's five prevalent usage patterns – Conceptualization, Data Collection, Data Analysis, Knowledge Crystallization, and Knowledge Presentation. The icons represent: group work, individual work, cyberinfrastructure use, visual analysis, interactivity, intensive writing, and intensive coding.

Group Work: represents work that can benefit from collaborations among group members. E.g., when brainstorming, scientists may sit together and display possible directions for investigation. Or, in the interdisciplinary analysis phase, group members may meet to present individual progress in the project and form an overarching narrative of the cumulative progress, which then fuels further rounds of analysis.

Individual Work: represents work performed by a scientist, independent of their group, either on their personal computer or a display wall. E.g., during product development, an individual may require concentration to write large amounts of text for a publication, which can later be refined in a group setting.

Cyberinfrastructure: represents stages that may require considerable cyber facilities to complete, such as High Performance Computing clusters. Depending on the size of the data collected and the complexity of analysis, scientists need to have advanced infrastructure locally. This icon is included in data gathering and cleaning, analysis and refinement, as well as for presentation modes that are interactive, since all these need to access data and potentially manipulate it.

Visual Analysis: represents a mode of work that involves visualizing the data (i.e., via charts and imagery) to allow a scientist to physically inspect the data for trends and generate hypotheses. E.g., during analysis refinement, a scientist may graph some of the data they are studying to see if there are any pronounced outliers that should be resolved.

Interactivity: represents work that may benefit by providing control over the data being presented to users who wish to interact with it, such as filtering data by some criterion or changing parameter values of an algorithm to better analyze the data. E.g., during the interdisciplinary analysis phase, many scientists may meet with colleagues to discuss their individual work, prompting the colleagues to contribute ad-hoc ideas for future work and testing. For that purpose, a scientist may expose their scripts or filters for the group's use.

Intensive Writing: represents work that requires large amounts of writing, such as generating reports, memos, or publications.

Intensive Coding: represents work that requires large amounts of computational writing, such as, creating scripts using Python or R, to clean and complete the data or analyze it. Coding is an inherently iterative process and can consume time and computational resources.

The following sections explain how these activities are employed within each of SAGE's five usage patterns. Note: while not explicitly depicted in Figure 1, SAGE has two overarching patterns, Collaboration and Security, which we will also elaborate on in the sections below.

B. The Conceptualization Pattern

Early conceptualization of research through freeform brainstorming can leverage SAGE's sticky note tool (digital *Post-It Notes* created in free text or Markdown syntax). Traditional brainstorm meetings often use paper Post-It Notes – not digital Post-It Notes on display walls. The advantages of the latter are manifold. For example: paper Post-It Notes do not adhere well to poster boards, especially when the Post-Its are frequently moved, and after a brainstorming session, notes on paper Post-It Notes need to be recorded – either transcribed into another document or photographed. Digital quick notes, however, do not suffer from these limitations. In SAGE, display walls with electronic quick notes can be saved in their entirety and subsequently recalled. The notes can also be coalesced into a single text document, which affords the opportunity to generate real-time Word Clouds that summarize the topics posted. At location sites with multiple SAGE walls, a "spare" wall can be used as a permanent idea wall that evolves over time.

One particularly effective SAGE-based approach to brainstorming consists of the following steps: (1) a group of participants use their laptops to independently generate quick notes and post them on the SAGE wall; (2) after some time (e.g., 30 minutes), each group member reads all the notes on the wall and writes a short summary or opinion of what has been posted; (3) all members tile their summaries across the wall and briefly read their notes to the entire group; and, (4) all members then openly discuss what they have just heard.

This methodology has the benefit of giving everyone in the room a voice in the process of ideation. The wall enables them to externalize their thoughts at their own pace and voice their opinions about what has been shared. Each note can contain the author's name or be made anonymously, freeing the discussion and enabling shy users to participate equally. This is particularly valuable when the native language of a participant is not the dominant language.

Figure 3: Picture of a LandSAGE brainstorming meeting – a project to leverage advanced cyberinfrastructure to study and mitigate landslides in South East Asia [13]. For most participants, English is not their native language.

Scheduling is another common task in both research and industry, which requires relatively small groups to collaborate and negotiate in order to solve a large logistical problem with many nuances. For scheduling large conferences, the process is traditionally a manual procedure performed using physical tools, such as whiteboards and Post-It Notes, and can take days. The SAGE2 application StickySchedule [8] is a multi-user application for display walls that better enables groups to organize conference schedules. Through authentic use cases with expert feedback from organizers who are heavily involved in large conference scheduling, we validated the usefulness of SAGE and StickySchedule in the collaborative and competitive aspects of conference scheduling, in both collocated and distributed settings.

Figure 4: Group using StickySchedules to schedule a conference program.

Another brainstorming task is project planning. For example, Caterpillar periodically plans and schedules project planning meetings for the design, development, engineering and manufacturing processes of construction equipment with worldwide teams and suppliers. These sessions are extremely complex and currently performed manually with paper and Post-It Notes. We worked with Caterpillar to explore the use of SAGE to automate the process [30, 16].

C. The Data Collection Pattern

Data collection typically occurs in the early stages of an e-Science enterprise. The major aspects are identifying available information, inspecting them for quality (through visualization), and planning for future data gathering expeditions. An excellent example is demonstrated in the University of Hawaii C-MĀIKI project. C-MĀIKI (Center for Microbiome Analysis through Island Knowledge and Investigation) scientists study microbiomes in Hawaiian ahupuaʻa (ridge-to-reef) to better understand the impact of Earth's microbial communities on biosphere sustainability. Given the difficulties of reaching the terrain from which microbiome samples are selected and limited resources, there is a critical need to optimize sampling efforts.

Figure 5: Identifying potential data gathering sites.

C-MĀIKI intends to plan, design and execute experiments over difficult terrains (some of which are reachable only by drones), necessitating the use of SAGE. Scientists will be split into field and lab groups. The lab group will use SAGE to view an interactive high-resolution 3D model of the terrain created from a previous drone flight. The field group will fly a drone and its footage will be streamed to the SAGE wall in real time along with the drone's current location. The lab group will be able to bring up images from past and current flights, assisted by automated image analysis and stitching, and discuss differences and identify points of interest before sending coordinates back to the field group to collect data.

D. The Data Analysis Pattern

The best way to leverage SAGE display walls is to use their large, high-resolution screens to juxtapose and view a variety of data and visualizations simultaneously, without having to swap between windows on laptop displays. Data analysis involves identifying data or pieces of data with relevant information to the topic at hand, noting trends, and evaluating alternatives. In particular, it enables multiple users to pursue multiple investigations simultaneously, enabling them to consider more scenarios in a given time [18].

This usage pattern is particularly effective for collaborators working together using Computational Notebooks. Computational Notebooks are windows to large-scale computational resources and provide users with a medium for exploratory and explanatory analysis, with JupyterLab becoming the de facto standard across industry and research [Perkel18]. However, users often remain isolated from one another when using these notebooks. While SAGE walls can provide users with a high-resolution analysis workspace, the traditional linear notebook format does not fully utilize the afforded "space to think" [1]. By extending JupyterLab to interface with SAGE, users can synchronize analyses in their notebook with visualizations on the SAGE large display wall. This supports parallel, juxtaposed analyses produced using a remote, scalable computational resource pool. The following photo illustrates three data analysts using SAGE in conjunction with JupyterLab. Two users work natively in their computational notebooks, displaying their notebooks and visualizations in the shared group workspace while the third manipulates and semantically organizes the information to further support analysis. The SAGE wall content includes one analyst's shared laptop display, two JupyterLab Notebooks, and five visualizations produced from analyses.

Figure 6: Collaboration over multiple Jupyter notebooks.

Another example of the data analysis pattern is exemplified by neurobiologists who are striving to identify mechanistic biological models of cognitive processes in the brain (in other words, to understand how we think and how our brain works).

Figure 7: Neurobiologists analysing their data.

Collaborations among experts in such diverse fields as neurobiology, neuroanatomy, neuroimaging, computer vision, signal processing, dynamic systems, machine learning, and data visualization will enable them to study diverse data flows from animal subjects and make biologically plausible interpretations of dynamic functional neural circuits involved in decision making. As researchers scale and optimize their workloads, SAGE helps them access, share, compare and discuss dynamic brain models of recorded populations. The following photo shows data streams corresponding to various stages of a data mining pipeline from raw data on the left (video of a neurobiological experiment in which a mouse performs behavioral tasks in virtual reality) to an extracted 3D temporal dynamics map of activated neuronal populations (real-time correlogram of all recorded neurons) in the middle section, to reconstructed live dynamics of functional connectivity of the brain circuits on the right.

E. Knowledge Crystallization Pattern

After data analysis comes knowledge crystallization. While data analysis reveals the most relevant data and methods to support the investigation at hand, knowledge crystallization seeks to solidify the linkages among the disparate pieces of evidence in order to draw conclusions. Given the ever increasing size of data, there is a need to create innovative, interactive environments that facilitate the analysis and interpretation of data. This pattern helps address challenges facing researchers and decision makers from all domains, in the creation of data-intensive workflows that bring together different information.

Figure 8: Juxtaposing rain and flood sensor data in SAGE-RDI.

An example of such a workflow is exemplified in SAGE-RDI (River Disaster Information), a SAGE-based tool to enhance a publicly available website in Japan [26] for monitoring river levels, to enable decision makers to arrive at decisions about potential flood events during heavy rainfall with greater speed and confidence. SAGE-RDI [12] was built to take advantage of wideband display walls by embedding meaning into location [1]. At the center of SAGE-RDI, the Japanese government's web portal is displayed on the SAGE wall surrounded by visualizations showing more detailed information about nearby sensors in a region. Predictable clustering of information is observed to encourage familiarity and hence aid the speed of understanding. Future research will attempt to automate the clustering process by understanding how various placements of information affect user cognitive load on wideband display walls. This layout was useful to determine how an area was doing as opposed to the original format, which only allowed viewing a single sensor at a time without context of where the sensor was located on the map.

Today it is not uncommon for the writing of research proposals and papers to be collaborative endeavors that require the use of shared word processors such as Google Docs or Microsoft Office 365, spreadsheets for budgets and charts, and live video conferences with collaborating partners. The wideband screen estate affords the juxtaposition of all these elements simultaneously so that a more accurate and compelling story can be told. For example, in Figure 1, the picture used for Knowledge Crystallization (X) is the meeting held to discuss the writing of this very paper. Participants in Chicago, Oahu and the Big Island of Hawaii all connected via SAGE using appear.in for video conferencing and were able see and speak with each other and simultaneously view the conference website, and examine prior written documents to identify components that would apply to the new paper. Early PDF drafts of the entire paper can also be tiled across the entire wall between collaborating sites for co-authors to see both the detail in the paper in relation to the whole. In general, as part of knowledge crystallization, notes are used to represent the important or critical knowledge to take away from the information reviewed. In addition to heavy note usage during the Conceptualization phase, notes perform a critical role during the Knowledge Crystallization phase as a medium for all participants to identify and confirm relevant knowledge with each other. This is a capability also built into SAGE-RDI.

Figure 9: A group of graduate students in front of a shared wall discussing notes.

F. The Knowledge Presentation Pattern

The SAGE display wall is an excellent medium for storytelling through the use of images, text, videos and visualizations. The walls are frequently used as part of tours of a facility, such as a research lab. Increasingly, the walls are replacing traditional whiteboard/blackboard classrooms, and frequently they are used for conference and museum demonstrations, which may require them to be portable.

Figure 10: Picture of a portable SAGE wall that ships entirely within a road case. (Designed by collaborators at Calit2-Qualcomm Institute at UC San Diego.)

Perhaps one of the burdens of owning a wideband display wall is that they are often the target of tours and demonstrations for VIPs. Nevertheless, these demonstrations are important to showcase the research output of an organization to funding agencies, to conference attendees, and to the general public (via museum shows). To manage the high demonstration volume these walls typically receive, it is not uncommon for SAGE users to build a second nearby wall specifically dedicated to demonstrations.

In classroom environments, wideband display walls can have a truly transformative impact when used to their full potential. For example, a pathologist and educator at Geisinger, a health care provider, who was previously at University of Illinois at Chicago (UIC), has used SAGE walls for distributed medical Clinical Informatics classes. At UIC, they are used for data visualization classes, and at the University of Hawaii at Manoa and Hilo, they are used to teach distributed Video Game classes as well as Math and Informatics classes.

Using SAGE on large 20- or 30-foot wideband displays at these situations has several advantages, and is similar in many ways to having a ‘classic’ classroom with multiple rolling blackboards or whiteboards, giving more room to show context for the current material, and leaving enough room to make comparisons and linkages without having to flip back and forth through a PowerPoint deck or scroll back and forth through a web browser or film roll on an overhead projector. For example, for the Clinical Informatics class previously taught at UIC, a universal microscopic whole slide image (WSI) viewer for SAGE was developed to combine WSI with medical images and patient care data for medical education [17].

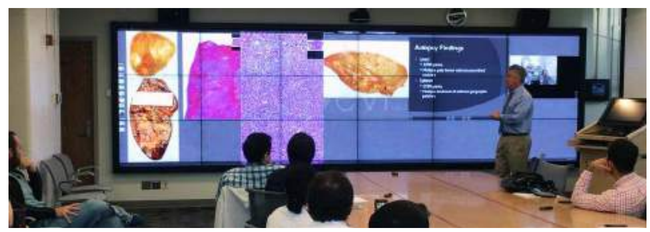

Figure 11: A “Virtual case conference,” held during a UIC fellowship program in Clinical Informatics, displayed high-resolution microscopy images.

At the University of Illinois Hospital, a large interactive visualization of Intensive Care Unit (ICU) data for effective patient care handoffs was developed [35]. One of the significant challenges of patient care transitions in ICUs is the lack of effective support tools for outgoing clinicians to find, filter, organize, and annotate information that can be effectively handed off to the incoming team. Using SAGE, a large display, interactive, multivariate, visual approach, aimed towards supporting clinicians during the transition of care, was developed based on observation sessions and interviews with several biomedical researchers and ICU clinicians. It is designed as a scalable, interactive visual approach that supports both overview and detail views of ICU patient data, as well as anomaly detection, comparison, and annotation of the data. Feedback from domain experts indicates this approach successfully meets the requirements of effective care transitions.

Figure 12: ICU patient data visualization on two different sized SAGE walls.

Using SAGE, multiple potential solutions can be easily compared, and evolution over time is easier to see, as shown in the following data visualization class at the University of Hawaii at Manoa where students were asked to compare their approach to visualizing data of the 1854 London Cholera outbreak.

Figure 13: SAGE2 used for comparing various visualization approaches.

Instructors interact with SAGE on a large display wall through a web browser on their laptop or, via touch if the wall has a touch-screen overlay. During course project presentation days, the classroom can function like a grammar school hallway where the instructor posts students' solutions on the wall at the same time, so everyone sees the multiplicity of solutions and reflects on their solution compared to the solutions of others.

For seminar-style courses, SAGE makes it easy for the students in the audience to add to current conversation by dragging and dropping relevant papers or videos or images or notes or their laptop screen onto the wall, where the wall is large enough that they can upload these additions on the sides of the screen without blocking the current discussion.

Instructors are initially reluctant to give students so much control of a lecture, for fear that the class may descend into chaos. From our experience, students quickly learn that assisting others during their presentation results in similar assistance being reciprocated.

In an Applied Informatics course at the University of Hawaii at Hilo, SAGE walls are used in a course designed to explore innovative tools that facilitate the collection, storage, synthesis, analysis, and visualization of data. This course examines the theory and application of several informatics and analytics tools used in marine science and other disciplines in the natural sciences. The students primarily learn and develop new skills by completing coding-based modules in R and Python. The SAGE wall provides a useful and interactive tool to engage students and improve their learning experience. SAGE allows the instructor to structure the modules so the students can, not only view the instructional template, but also share and explain their results to the class by presenting their material interactively and in real time. Student feedback has been exceptional, as they enjoy having control of the SAGE wall to present material, which encourages them to apply skills learned in the modules. The students can complete exercises and then subsequently mine data online and run new analyses, and ultimately discuss results with the class. The short presentations they give when sharing their results with SAGE helps solidify the concepts covered in class and improves the overall educational experience. Without SAGE, the course would consist of simple lectures with the instructor presenting all the information. This "one-way" transfer of information limits the ability of the students to practice the skills they are learning in class. Engaging the students with SAGE enhances their learning while simultaneously exposing them to the value of interactive data science tools.

In all the scenarios outlined thus far, there is a need to prepare the visual elements ahead of time. SAGE provides the notion of *sessions* which enables all content on the wall and their layout to be recorded for later recall. In addition, a collection of content (PDF files, images, videos, etc.) can be packaged into an archive file (zipped folder) and dropped onto the wall for staging before the meeting or, alternatively, can be opened right away and laid out on the wall.

For example: Scott Sellars used SAGE and resources at UC San Diego's Calit2-Qualcomm Institute, to do a remote presentation from Virginia and California to a SAGE wall set up in the SC18 conference research exhibition. He utilized the Nautilus cluster, a Pacific Research Platform (PRP) "GPU Challenge," to speed up the segmentation process for his weather data [19]. Through SAGE, he was able to simultaneously use videoconferencing, a shared desktop, a shared JupyterLab Notebook, local animations and real-time dashboards in webviews to demonstrate his process and illustrate GPU cluster utilization. SAGE sessions were effective for storing and retrieving these demonstrations.

Figure 14: A screenshot of a Scott Sellar's demonstration at the SC18 conference.

In a museum setting, SAGE has also been heavily utilized by the University of Hawaii at Hilo ʻImiloa Astronomy Center, Hawaiʻi's premier astronomy museum. ʻImiloa's Executive director Kaʻiu Kimura has been supportive of the CyberCANOE and SAGE technologies. SAGE is used for work meetings, public education and outreach by astronomers, and other scientists at nearby observatories. At ʻImiloa's 13th anniversary, the Carnegie Airborne Observatory (CAO) lab used SAGE to showcase their high-resolution land survey scans of areas on the Big Island to the public.

Figure 15: SAGE2 on the ‘Imiloa CyberCANOE used for a public talk by Carnegie Airborne Observatory staff.

The day before the event, the CAO crew uploaded high-resolution photos and videos to SAGE and positioned them as desired across the ‘Imiloa CyberCANOE – a system with an array of three 80-inch displays. The SAGE session was then saved for easy retrieval the following day, reducing the setup overhead. Additionally, SAGE enabled the presenters to screen share additional content from their laptops to the display wall as necessary. The 10-minute presentation was given once an hour for the four hours the event ran. SAGE’s ability to share a variety of rich media items on a single platform made it easier for the speakers to give high-quality educational presentations to the public. Using SAGE with a CyberCANOE is a novel offering at ‘Imiloa, facilitating better patron engagement and information dissemination. The museum has added the use of this system to its daily exhibitions, running two daily presentations with it.

V. SAGE Security Patterns

Security in a typical web platform involves encrypting communications, securing resources and providing authorized access to those secure resources. The SAGE environment, due to its open collaborative model, needs a more nuanced handling of security aspects. The resources in a SAGE environment are all user generated. Hence, security and privacy requirements vary with each resource. Further, most of the resources are generated or shared during collaborative sessions, making the security requirements time sensitive as well. These aspects make SAGE a unique use case for cyberinfrastructure, especially given SAGE’s distributed nature and application diversity. As the first step to address these security requirements, the SAGE team designed an Identity and Access Management (IAM) scheme appropriate for SAGE’s user community, by collaborating with Trusted CI (NSF’s Cybersecurity Center of Excellence). The results were presented to the SAGE community at SC18 [31], which suggested a need to authenticate SAGE2 users, to establish their identities, and to grant appropriate levels of access to resources (computation, storage, networking, workflows). Further, we have to manage data and cyber services in a secure, yet easy, way so these resources can be quickly shared with intended collaborators. The need to distinguish various meeting ‘types,’ such as ‘classroom’ or ‘brainstorming,’ and to set appropriate access levels for participating users based on the type of meeting, was also a concern. Finally, the need to distinguish between physical rooms and virtual meeting spaces was noted.

Work has begun to build this security model starting with the incorporation of OpenID Connect (OIDC) into SAGE. OIDC is an integrated identity and access management platform for science, combining federated identity management (Shibboleth, InCommon) with collaborative organization management (COmanage), which lets users participate in SAGE sessions through a single sign-on mechanism. Our goal is to extend this robust identity management system whereby user credentials from a single sign-on will be used throughout a session to grant access to various resources (i.e. to retrieve a dataset, start a computation, launch a workflow on remote resources).

VI. SAGE Collaboration Patterns

Use patterns that affect a large group of collaborators, such as knowledge crystallization and knowledge presentation, can engage users that are either collocated or remote. SAGE enables bridging the gap between a variety of domain specialties where the users are not necessarily physically available to each other.

When collaborators share the same workspace and have access to the same SAGE display, they benefit from shared context and a large “space to think” [1]. Most of the scientific patterns we reflect on in this paper were described with local collaborators in mind. In addition, SAGE is greatly advantageous for remote collaboration as well.

Collaborators at one SAGE display site will commonly work with remote colleagues who have their own SAGE display wall. Different sites often have different physical configurations: for example, one site may have eight screens arranged in two rows, while another has three large displays side-by-side. SAGE allows collaborators to share content between the sites, which can be screen-sharing sessions, PDF documents, images, videos, notes, websites and other applications. SAGE synchronizes their content while giving each site control over the position and size of the content. Local users’ pointers are shared with the remote sites, so a user can point to an element in a webpage or PDF, letting everyone follow the flow of the discussion.

Since SAGE runs on web technologies, we are able to take advantage of online services, such as video conferencing (BlueJeans, Google Hangout, appear.in, etc.), and online editors (Google Docs, Microsoft Office 360, Overleaf, etc.). The benefits of using these services on top of SAGE are the synchronization of content beyond sharing a screen or a camera feed between locations. For instance, when sharing a PDF document between remote SAGE sites, participants at either location are able to control the page being viewed, the number of pages, and use the SAGE pointer to direct everyone’s attention. This makes sure everyone is literally “on the same page.” Extensions to these collaboration patterns are being explored, such as shared partitions of one’s wall (each content dropped into a designated region of the display wall is automatically shared), whole wall mirroring (every window is shared and scaled to fit the remote site physical screen layout), and adapting to the remote participant ‘on-the-go’ (user with a phone or laptop only) to follow a meeting and watch the remote walls without needing a large display.

VII. Conclusion & Future Work

This paper describes how SAGE's support of the kinetics of the e-Science research, development and training enterprise can be characterized by a number of prevalent usage patterns and tasks that occur within those patterns (Figure 1). These patterns can be used to help e-Science users better leverage SAGE's capabilities as well as understand how to better use wideband displays in general.

Perhaps most evident from these patterns is that they are effective because they fully leverage the screen estate of wideband display environments – enabling users to distribute information on them in a variety of ways that make it easier for them to see the "big picture" in their data. This then raises the question, once the information is laid out across the wall, is that enough? As the scale and complexity of science and its data continue to grow exponentially, simply being able to present all the information will not be enough. Humans will increasingly need an artificially intelligent assistant to assist with collaboration in the e-Science enterprise.

In the past year, the SAGE team has been considering the design of SAGE3. In SAGE3, the "S" in SAGE will stand for *Smart* to emphasize its use of AI. For example, SAGE3 will introduce the notion of SmartSpaces – templates that help teams quickly and intuitively initiate common SAGE usage patterns. SmartSpaces will be shareable across platforms, made portable through containerization, and leverage new back-end resources to enable data provenance, recovery, and consistency. Furthermore, the moment documents or data are deposited on the wall, AI will autonomously begin to act upon them. These capabilities will be enabled in SAGE3's *Foresight Engine*, which will operate under user-defined rule sets and state-of-the-art machine learning theory, grounded in recommendation systems and reinforcement learning algorithms. The module will be scaffolded atop a lightweight micro-services architecture using existing machine-language libraries (e.g. TensorFlow, PyTorch) to provide AI capabilities. Generated insights will be delivered, in accordance with SAGE's human-in-the-loop model, to users who will take over from machine learning to interpret what patterns mean and determine if there is any scientific value. These interactions will be recorded and used to reinforce learning and support/justify future decisions. SAGE's Foresight Engine will bring us one step closer to our ultimate vision of making SAGE a first-class partner in a scientific collaboration [32].

Acknowledgment

SAGE2 and SAGE are trademarks of the University of Illinois Board of Trustees. Major funding for SAGE2, SAGE, and CyberCANOE is provided by U.S. National Science Foundation awards ACI-1441963, 1550126 and CNS-1530873; the Academy for Creative Media System at the University of Hawaii; the United States Department of the Army (SBIR topic A18-032); and, the Asi@Connect project, which is a European Union co-funded project under grant contract ACA 2016-376-562.

References

[1] C. Andrews, A. Endert, and C. North, "Space to think: large high-resolution displays for sensemaking," Proc. of CHI'10, ACM, 2010, pp. 55–64.

[2] R. Ball and C. North, "Analysis of user behavior on high-resolution tiled displays," Human-Computer Interaction-INTERACT 2005, Springer, 2005.

[3] R. Ball, and C. North, "Effects of Tiled High-Resolution Displays on Basic Visualization and Navigation Tasks," CHI '05 extended abstracts on Human Factors in Computing System 2005, ACM, 2005, pp. 1196-1199.

[4] R. Ball, C. North, and D. Bowman, "Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays," Proceedings of the ACM Conference on Human Factors in Computing Systems 2007 (CHI 2007), ACM, 2007, pp. 191-200.

[5] X. Bi and R. Balakrishnan, "Comparing Usage of a Large High-Resolution Display to Single or Dual Desktop Displays for Daily Work," Proceedings of the ACM Conference on Human Factors in Computing Systems 2009 (CHI 2009), ACM, 2009, pp. 1005-1014.

[6] M. Czerwinski, G. Smith, T. Regan, B. Meyers, G. Robertson, and G. Starkweather, "Toward characterizing the productivity benefits of very large displays," Proceedings of Human-Computer Interaction – Interact 2003, pages 9-16. IOS Press - IFIP, 2003

[7] R. S. Dooley and M. R. Hanlon, "Recipes 2.0: Building for today and tomorrow." 5th International Workshop on Science Gateways, IWSG 2013 - Zurich, Switzerland, 2013.

[8] V. Doshi, . Tuteja, K. Bharadwaj, D. Tantillo, T. Marrinan, J. Patton, and G. E. Marai, "StickySchedule: an interactive multi-user application for conference scheduling on large-scale shared display"s. In Proceedings of the 6th ACM International Symposium on Pervasive Displays (PerDis '17). ACM, New York, NY, USA, Article 2, 7 pages, doi:10.1145/3078810.3078817.

[9] B. Jeong, L. Renambot, R. Jagodic, R. Singh, J. Aguilera, A. Johnson, and J. Leigh, "High-Performance Dynamic Graphics Streaming for Scalable Adaptive Graphics Environment," ACM / IEEE Supercomputing 2006.

[10] E. Gamma, R. Helm, R. Johnson, and J. Vlissides, "Design Patterns: Elements of Reusable Object-Oriented Software", Addison-Wesley, 1994.

[11] N. Kawano, R. Theriot, J. Lam, E. Wu, A. Guagliardo, D. Kobayashi, A. Gonzalez, K. Uchida, and J. Leigh, "The Destiny-class CyberCANOE – a surround screen, stereoscopic, cyber-enabled collaboration analysis navigation and observation environment, Journal of Electronic Imaging 2017, doi.org/10.2352/ISSN.2470-1173.2017.3.ERVR-093.

[12] D. Kobayashi, M. Ready, A. Gonzalez Martinez, N. Kirshenbaum, T. Seto-Mook, J. Leigh, and J. Haga. 2018. "Sage River Disaster Information (SageRDI): Demonstrating Application Data Sharing In SAGE2." In Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces (ISS '18). ACM, New York, 33-42, doi:10.1145/3279778.3279798.

[13] landsage.info

[14] J. Leigh, A. Johnson, L. Renambot, T. Peterka, B. Jeong, D. J. Sandin, J. Talandis, R. Jagodic, S. Nam, H. Hur, Y. Sun, "Scalable Resolution Display Walls," Proceedings of the IEEE, Issue 99, May 2012, pp. 1-15, doi:10.1109/JPROC.2012.219160.

[15] T. Marrinan, J. Aurisano, A. Nishimoto, K. Bharadwaj, V. Mateevitsi, L. Renambot, L. Long, A. Johnson, J. Leigh, "SAGE2: A New Approach for Data Intensive Collaboration Using Scalable Resolution Shared Displays," In Proceedings of the 10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing, Miami, FL, October 22, 2014.

[16] T. Massari, "Collaborative Project Planning Using Large Format Displays (thesis)," Master of Science in Computer Science, Graduate College of the University of Illinois at Chicago, Chicago, IL, May 2019, www.evl.uic.edu/entry.php?id=2432.

[17] V. Mateevitsi, T. Patel. J. Leigh, B. Levy, "Re-imagining the microscope in the 21st Century using the Scalable Adaptive Graphics Environment,"

Journal of Pathology Informatics, Vol. 6, Issue 1, 2015, pp. 25, doi:10.4103/2153-3539.158042.

[18] K. Park, L. Renambot, J. Leigh, and A. Johnson, “The Impact of Display-rich Environments for Enhancing Task Parallelism and Group Awareness in Advanced Collaborative Environments,” Workshop on Advanced Collaboration Environments, June 22-24, 2003, Seattle, WA.

[19] pacificresearchplatform.org/nautilus/

[20] J. M. Perkel, "Why Jupyter is data scientists' computational notebook of choice." Nature 563.7729 (2018): 145. doi: 10.1038/d41586-018-07196-1.

[21] K. Reda, A. Febretti, A. Knoll, J. Aurisano, J. Leigh, A. Johnson and M. Hereld, “Visualizing Large, Heterogeneous Data in Hybrid Reality Display Environments.” IEEE Computer Graphics & Applications, IEEE Computer Society, July/August 2013, pp. 38-48, doi:10.1109/MCG.2013.37.

[22] K. Reda, A. Knoll, K. Nomura, M.E. Papka, A.E. Johnson, J. Leigh, “Visualizing Large-Scale Atomistic Simulations in Ultra-Resolution Immersive Environments,” Proceedings of the IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV ’13), IEEE Visualization 2013, Atlanta, GA, October 13-14, 2013, IEEE, pp. 59-65, doi:10.1109/LDAV.2013.6675159.

[23] K. Reda, J. Aurisano, A. Febretti, J. Leigh, A. Johnson, “Visualization Design Patterns for Ultra-Resolution Display Environments,” In Proceedings of the Workshop on Visualization Infrastructure and Systems Technology (VISTech ’13), Denver, CO, November 22, 2013.

[24] Reda, K., Offord, C., Johnson, A., Leigh, J., Expanding the Porthole: Leveraging Large, High-Resolution Displays in Exploratory Visual Analysis, Proceedings of CHI’14 EA on Human Factors on Computing Systems, Toronto, Ontario, Canada, pp. 2047-2052, ACM, April 26, 2014.

[25] L. Renambot, T. Marrinan, J. Aurisano, A. Nishimoto, V. Mateevitsi, K. Bharadwaj, L. Long, A. Johnson, M. Brown, J. Leigh, “SAGE2: a Collaboration Portal for Scalable Resolution Displays,” Future Generation Computer Systems, Volume 54, January 2016, Pages 296–305, doi:10.1016/j.future.2015.05.014.

[26] www.river.go.jp

[27] apps.sagecommons.org

[28] www.sagecommons.org

[29] tinyurl.com/learnsage

[30] tinyurl.com/y6zya3oj

[31] mm.tt/1173069736?t=V1SUCu36mn

[32] Y. Sun, J. Leigh, A. Johnson, S. Lee, “Articulate: A Semi-automated Model for Translating Natural Language Queries into Meaningful Visualizations.” Proceedings of 10th International Symposium on Smart Graphics, Lecture Notes in Computer Science 2010, Banff, Canada, 06/24/2010 - 06/26/2010.

[33] D.S . Tan, J.K. Stefanucci et al., “The Infocockput: providing location and place to aid human memory,” PUI, 2001.

[34] D. Tan, D. Gergle, P. Scupelli, R. Pausch, “With similar visual angles, larger display improve spatial performance.” Proc. of CHI ’03, pp. 217-224.

[35] M. M. Thomas, T. Kannampallil, J. Abraham and G. E. Marai, "Echo: A large display interactive visualization of ICU data for effective care handoffs," *2017 IEEE Workshop on Visual Analytics in Healthcare (VAHC)*, Phoenix, AZ, 2017, pp. 47-54. doi:10.1109/VAHC.2017.8387500.

[36] C. Ware, W. Wright, and N. Pioch, “Visual Thinking Design Patterns.” Proc. of the 19th International Conference on Distributed Multimedia Systems (DMS '13). Aug 8-10, 2013, Brighton, UK.

[37] B. Yost, Y. Haciahmetoglu, C. North, “Beyond visual acuity: the perceptual scalability of information visualizations for large displays,” Proceedings of SIGCHI Conference on Human Factors in Computing Systems (CHI ’07), ACM, 2007, pp. 101-110.