Journal of Materials Education Vol. 44 (3-4) 57-66 (2022)

COMPUTATIONAL MATERIALS DESIGN (CMD®) WORKSHOP: A HANDS-ON WORKSHOP OF FIRST-PRINCIPLES CALCULATIONS ONLINE

Masaaki Geshi

R^3 Institute for Newly-Emerging Science Design, Osaka University, 1-2 Machikaneyama, Toyonaka, Osaka, Japan 560-0043 *geshi.masaaki.insd@osaka-u.ac.jp*

ABSTRACT

Operation and results of Computational Materials Design (CMD®) Workshop held annually at Osaka University are reported from the perspective of an online format. The first CMD® workshop was held in 2002 with the goal of publicizing and disseminating first-principles calculation software that is essential for advanced research in materials science. In the past, participants gathered at venues such as those on campus at Osaka University and followed instructions directly using computers provided by the organizers connected to computer clusters. However, due to the coronavirus pandemic, it was necessary to change to an online format starting with the 37th Workshop in September 2020. Since the CMD® workshop required participants to connect to the Osaka University's computer clusters for hands-on training, they had to prepare in advance to connect to the computer clusters using their own computers through the Internet. For example, it was necessary to prepare a Linux environment running the X Window System for all participants. Organizers needed to instruct the participants to prepare their own computing environment for the hands-on training. The organizers needed to make thorough preparations prior to the workshop, and also make smart workarounds during the workshop to help the participants gain full benefit from the workshop. This paper presents and describes solutions to the problems inherent in holding this type of specialized hands-on workshop online, which focuses on both knowledge and real-life practice of advanced computational techniques for cutting-edge materials research.

Keywords: *hands-on training, first-principles calculations, online*

INTRODUCTION

The coronavirus pandemic that overtook the world by surprise in 2020 has forced various events to go online. While the real pleasure of deepening discussions in person is hindered, many feel that the savings in travel expenses and time are major advantages. It is also true that even if the spread of the infection is diminishing It is also true that even if people are learning to live with the virus and go back to previous lifestyles, it is still desirable for a considerable number of events to remain in the online arena. The Computational Materials Design (CMD®)

Workshop[1], which has been conducted since 2002 to publicize and disseminate first-principles calculations, also changed to the online format beginning from the 37th workshop held in September 2020. Since then, it has successfully completed four workshops up to the 40th workshop held in February 2022. The workshop is five days long and focuses on practical hands-on training, and offers six courses (Beginners Course, Advanced Course, Supercomputer Course, Spintronics Design Course, Materials Informatics Course, and Expert Course) according to the proficiency level and objectives of the participants. The Beginners course is conducted by lecturers and tutors who help participants with little or no Linux experience to learn first-principles calculations by helping them familiarize with the commands. Other courses offer practical in-depth hands-on training, focusing on topics in which the respective code package specializes. More specific details are presented below.

In the CMD® workshop, participants are trained to perform first-principles calculations in the same way one would actually do research. They copy software from the developer's directory on a Linux platform. The participants decompress and extract it with the tar command. They run the samples already placed in the directory according to the developer's instructions. The samples are prepared for typical and simple materials such as Si and Fe, so that the participants can modify the input files and calculate their desired materials during the free hands-on times provided each day during the workshop. In higher level courses, the participants in those courses can participate with a specific theme and be trained to build a calculation model together with the developers. For the participants in the beginners course, tutors provide support in operating Linux commands. The number of participants in the beginners course has recently ranged from 30 to sometimes more than 50. The total number of participants in the entire CMD® workshop has ranged from 50 to 70 or more.

There are several similar workshops in Europe[2-6]. These cover several days of hands-on training around a particular software and how to use it. Some of these are conducted through online methods. Unlike these, the CMD® workshop is one of the few workshops in the world where more than one method or software can be learned during the same period of time. In particular, in the Beginners Course, participants can learn the Korringa–Kohn–Rostoker (KKR) method, the pseudopotential method, and the full-potential linearized augmented plane-wave (FLAPW) method based on the Density Functional Theory (DFT)[7,8] developed by Japanese researchers. This allows participants to compare the advantages and disadvantages of each method. These three methods are the main methods for DFT calculations, and thus allow the participants to learn the basics of DFT calculations.

The significance of the CMD® Workshop lies in the fact that it envisions to use first-principles calculations not only to analyze physical phenomena, but also to design new materials and new functions using calculations originating from first-principles. The central concept of the CMD® workshop is to verify new materials created in this way through verification experiments, and to promote development through mutual feedback between computation and experimentation. This concept where computation and experimentation go hand-in-hand is widely used and has influenced many national projects in Japan. That is, each theme consists of a computational group and an experimental group[9]. The members of the CMD® Workshop have also contributed immensely to such projects. Thus, the participants are graduate students and researchers from materials science laboratories, from both computational and experimental groups. Approximately 25% of the participants are from the industry. Participants from experimental groups and/or industry are often at a stage where they cannot readily make an educated choice on a specific software to use for their research purposes. Therefore, for them it is extremely important to be able to compare and learn the characteristics of each method and

underlying concepts to make a smooth entrance to computational science. In this sense, the CMD® workshop is educational- and research-oriented with a substantial hands-on training component in comparison to the online workshops on various software from around the world[2-6]. In actual research, there is an immense need to approach the essence of a phenomenon from various approaches, rather than looking at it solely based on what can be done with computational methods; a need the CMD® workshop aims to facilitate.

Prior to the pandemic, the CMD® workshop was conducted on-campus, where most of the participants used computers prepared by the organizers. Hence, they could travel to the workshop basically empty-handed. Conducting the workshop online meant that the participants must first prepare their own Linux environment, which raises the hurdle for participation considerably. At first, there was concern that the number of participants would decrease significantly, but surprisingly, the number of participants reached a record high. This is likely due to the fact that many were encouraged by the elimination of travel time and expenses. In the post-conference survey, many participants expressed a desire to have the workshop online in the future as well. The organizers are also considering the use of the online format as a more sustainable method due to the overall benefits it offers.

However, successfully conducting the hands-on-training workshop without on-site presence was not easy. A great deal of preparation was required beforehand to ensure that the hands-on training would run smoothly without in-person instruction and guidance. The next section will describe how this was accomplished.

PREPARATION OF THE WORKSHOP

The preparation for the online workshop began with considerations on how the lectures will be delivered from Osaka University via some type of web conferencing system and how the participants will connect to the Osaka University's computer cluster for the hands-on training. There were several candidates for the web conferencing system at the beginning of April 2020, but Cisco's Webex was chosen due to security issues and the number of connections per account for the system contracted by the university. Since it was also a practical training course for the education of working adults participating from companies, there was a situation where it was impossible to employ tools that were prohibited by the companies. Therefore, it was not a simple matter of just using tools that were available at the organizer.

When hands-on training is conducted online, questions and answers and hands-on training support must be handled online real-time. In particular, the Beginners course typically has between 30 and 50 participants, making it difficult to deal with each participant individually in a dialogue. Since the chat function of the web conferencing system was not sufficient, Slack was employed. This allowed the participants to paste a snapshot of what was happening when they enter a command in the terminal into Slack and ask a question, even if they cannot explain in detail the situation they are having trouble with. Other participants who were also in the same situation could see the communications between the lecturers or tutors and the participant, and they could use this as a reference to solve similar problems.

The participants unfamiliar with Linux often have trouble keeping up with operations when command lines disappear from the display and become invisible. To tackle this, a list of the command operations was posted on Slack to help participants facing the same situation.

The use of Slack was not only for questions and answers, but also to allow all participants to paste some of the results of the exercises so that the lecturers could confirm that they are following the instructions and learning. Furthermore, by assigning different parameters to each participant and having them perform calculations, various results that depict a broader picture of a material were derived and viewed within a limited amount of time. These also provided evidence that the participants have

performed the hands-on training, and helped in evaluating their progress. Another advantage was that the workspace contents and communications remained available after the workshop, so that participants can review it at a later time.

Next, it was necessary to connect each participant's computer to the computer cluster at Osaka University in order to actually perform the hands-on training. The participants were required to "be able to prepare their own computer with Linux environment including the X Window System" and "be able to prepare a public key for public key authentication". In the hands-on training, the electronic density of states and band dispersion curves are generally calculated and displayed as they are on the computer cluster at Osaka University using gnuplot, gs, etc. Therefore, having the above environment available was essential. It is possible to transfer the calculated data to their own PCs and plot using familiar software, such as Excel. However, even if the lecturers output the data in formats readable by such software, each participant must install software for the file transfer. In addition, the organizer will not be able to provide support if the participants ask about what to do after the data has been transferred. It is ultimately less troublesome to complete the process within the PC cluster. Since this may prove to be a major hurdle for beginners, a manual for preparing this environment was designed and made available on our website. The recommended method was to install software that prepares a virtual environment and install Linux on it, such as Ubuntu[10]. There are many other ways to setup the environment but with certain shortcomings. In the case of Windows software, it often depends on the settings of individual PCs, and the organizer may not be able to deal with problems related to these settings or may not be able to accurately comprehend the varying situations. For this reason, the organizer does not support the use of these software. It is noted that the Windows Subsystem for Linux (WSL) and WSL2 are improving, and Mac has a Linux environment by default, but XQuartz must be installed for X Window System.

The preparation of the public key also required that they be created in the ED25519 format, which is currently considered the most secure with the passphrase requirement. Since this can also be difficult for beginners, a manual was prepared and was made available online.

Even if one prepares the Linux environment, trouble can surely occur if it is not tested in the same environment as the real one where hands-on training takes place. Therefore, all participants were advised to connect to the computer cluster before the day of the workshop and test a few instructions. Everyone was required to run xeyes[11] on the computer cluster to verify that the X Window System was working properly, and to take and send a snapshot including the prompts. A sample of this is shown in Figure 1. This screen shows the login screen of the Linux distribution Ubuntu, with the terminal launched for command-line operations. Following the prompt displayed in green, the desired command can be typed. The circled area at the top is the command to connect to the PC cluster at Osaka University, which changes to the circled prompt below after a successful login. In this state, we can verify that X Window System is forwarding correctly by launching xeyes, which is a graphical display of two eyes that follow the movements of the mouse. This is often used to verify that X Window System is working properly. The prompt shows that this is being done by connecting to the computer cluster, and the fact that xeyes can be displayed shows that the X Window System is working properly.
Performing the above tests significantly reduced the chances of problems on the day of the workshop. However, there were still a few participants who could not connect to the computer cluster. This was because the network environment on the day of the test was different from that of the day of the workshop. The condition for participation is to have a "stable network environment," but the beginners often do not fully understand their network environment or the Internet itself. In one instance, a participant did the test at home but participated from his office on the day of the workshop and was unable to connect because of the company firewall.

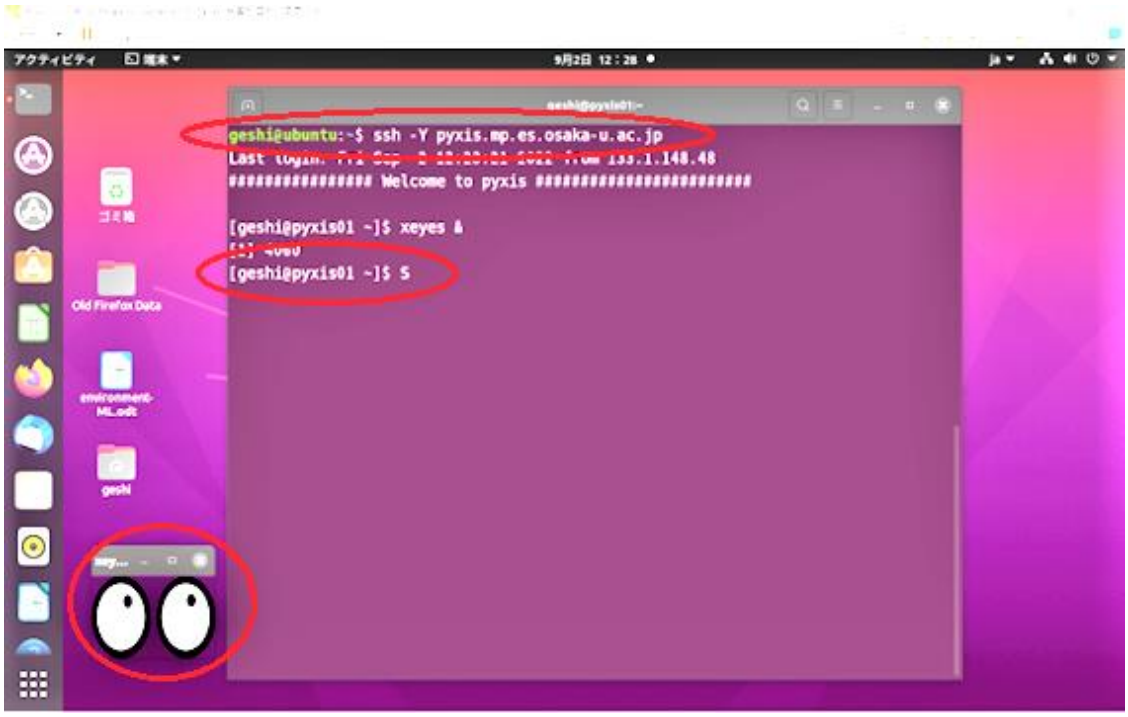

Figure 1: A sample snapshot of xeyes running, connected to the computer cluster at Osaka University. The areas circled in red show that the prompts are in correct state and that xeyes is running.

The recommended environment for online hands-on training was a single computer with two displays. For example, one laptop and one additional monitor. Figure 2 shows a laptop computer connected to an external monitor. The laptop screen in the foreground shows the web conference screen. Following the lecturers' instructions, the calculations are performed using the screen of the Linux environment displayed on the extended monitor. On the right side of the Linux environment, Slack, a chat tool used for discussions based on questions and answers, is displayed on the browser. Participants can direct questions to the lecturers and tutors via Slack. In addition, by pasting the screen shots where some trouble occurred onto Slack, participants can send more detailed information and ask specific questions. Also, commands given by the lecturers can be copied and pasted onto the terminal of the Linux environment on the left, reducing typographical errors. It is also possible to make the screen of the virtual machine full screen and launch Slack inside Linux, which is usually a better way of running the different tools used here. If two devices are used, it is better to run Linux and Slack on the same PC, and watch the web conference on another PC or tablet. This is because when the lecturers give instructions on Slack, there is a high chance for the instructions to require code to be copied and executed by pasting as they are.

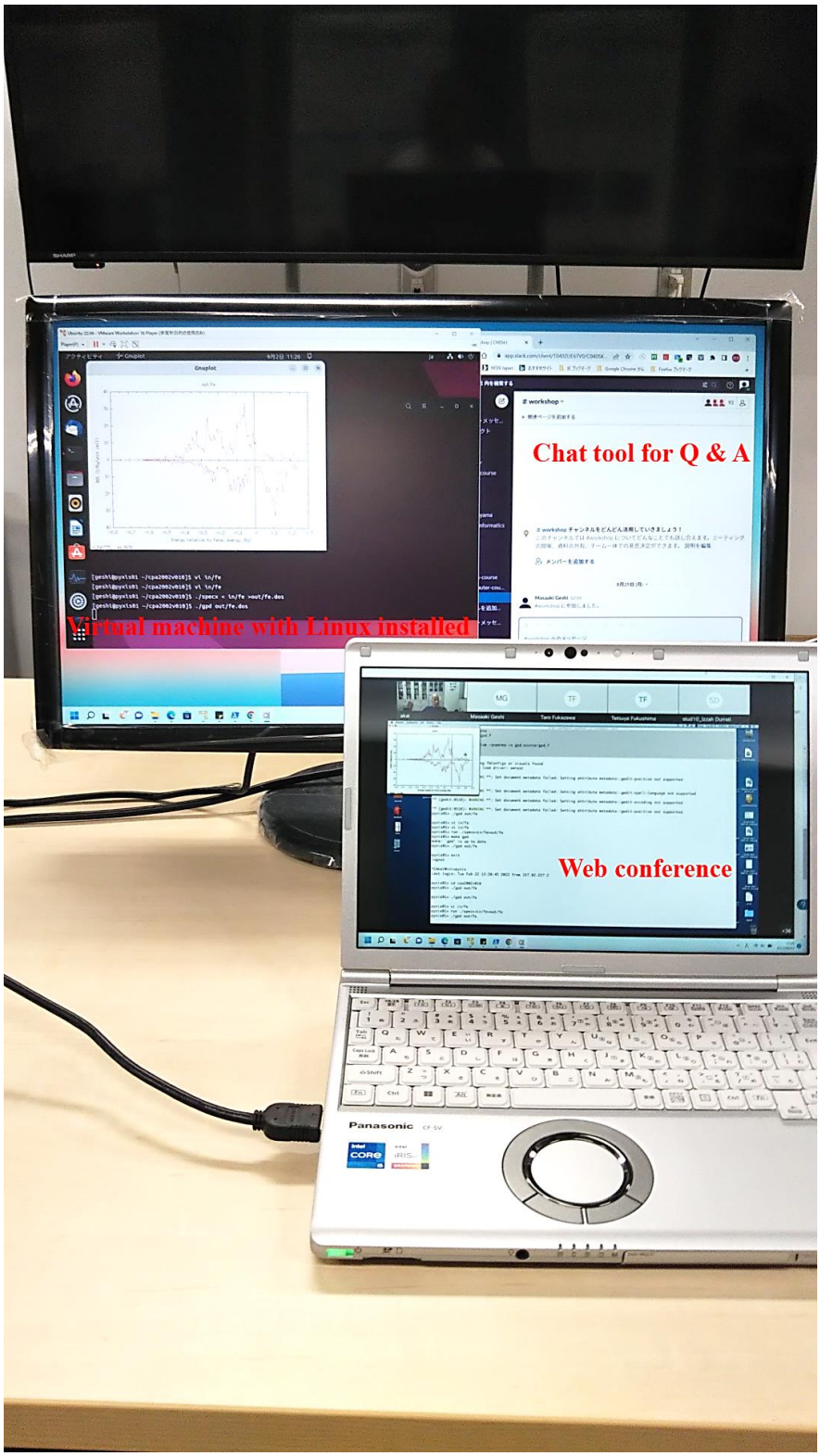

Figure 2: The recommended environment for online participation; one computer with an additional external monitor.

ONLINE-SPECIFIC ISSUES

We will discuss some of the problems that occurred during the hands-on training, especially, those that were unique to an online setting. One was originating from the communication state of the network, and the other was due to the fact that the workshop participants were not on-site.

Editing files in an editor also had problems specific to an online setting. When emacs or other software is launched as is, it appears in a separate window on each user's computing display, which itself depends on the state of communication. There were cases where participants were unable to keep up with hands-on training due to the time the editor took to start up. It took tens of seconds to start emacs. When using emacs, one can solve this problem by using the "-nw" option to display it in the terminal and start it. Sometimes the editor "vi" and "emacs" are too difficult for beginners, and "gedit" is used. However, it does not have an option like "-nw" in emacs, so another similar editor called "nano" can be used instead. Many participants in the Beginners course have difficulty keeping up with the hands-on training unless the lecturers pay attention to these aspects of the course. Alternatively, one can leave emacs once and select the file that needs editing each time. Although such solutions are available, since the participants in the Beginners Course were unfamiliar with Linux, coming to such solutions immediately proved to be difficult. The participants of the Beginners Course were easily confused when there was a difference in operation between lecturers in the fast network environment and the participant in the less fast network environment.

Another difficulty with online hands-on training for a large number of people was making sure that they were really learning and being trained. The CMD® workshop does not permit participation only for listening to the lectures. In addition, it is unacceptable for them not to be performing hands-on training, as it is used for evaluation of the participants because the CMD® workshop is an accredited course of Osaka university, and some attend to attain credits for their degree course. For these reasons, the files in all participants' directories should be checked, and the lecturers' instructions should be followed to copy the files and ensure that the calculations have been performed. This was partially supported by having the participants paste the results of their calculations into Slack, and participants who were only listening and not performing could not respond. In this case, pasting the results into Slack served as a kind of assignment submission.

Although in-person on-site instruction would, of course, be more effective in terms of education, online hands-on training is the effective method for both the organizer and the participants to learn without incurring costs and significant travels times. Hence, for the CMD® workshop, the online format will continue to be the effective method moving forward. Although we have focused on the case of the Beginners Course, the same pre-test was administered to all participants in order to minimize problems during the limited time of the workshop. Thus, the pre-tests were conducted for 75 or more participants. In addition, there were about ten tutors and more than 20 lecturers in charge of the hands-on training, which meant that nearly 100 user accounts had to be managed each time. The workload for those in charge of managing participants, networks, and computers was significant. Extreme care and effort were necessary not only in preparing in advance, but also in monitoring the entire workshop.

For the on-site workshop, computers were provided by the organizer. Therefore, the participants were able to attend the workshop basically empty-handed. This was helpful for participants. However, it was not easy for the participants to continue calculations after the workshop because they could not prepare the Linux environment by themselves. In the case of the online workshop, preparing the Linux environment by themselves on their local

computers proved to be very helpful when performing and continuing calculations on their own after the workshop. Within the scope of educational use, our computer clusters are available to those who wish to use them, and more people are continuing to practice than before. The good thing about online hands-on training is that once a significant hurdle is overcome for online participation, it is then easier to continue with the calculations long-term.

Some people may think that it is sufficient to install software locally on each participant's PC, and some online workshops do this. However, it is often impossible to deal with problems that occur in each participant's unique environment, and the organizer may end up wasting time on problems that cannot be solved by the workshop organizer. After all, having the lecturers and participants do the same thing on the same computer is the most efficient way to get hands-on training.

CONCLUDING REMARKS

In summary, the example of online CMD® workshop, which focuses on computational hands-on training of first-principles calculations was presented together with inherent problems of the online format and devised solutions . It is highlighted that it is extremely important for the organizer of an online workshop of this nature to take various detailed preparations into consideration to minimize problems on the day of the workshop. If the support of many technicians and ample funding are available, the technical details of holding the event online can be left to the experts, but if this is not the case, it becomes necessary to undertake what was described here.

The concept of first-principles-calculation-driven materials design that was launched 20 years ago is a groundbreaking one. The author would like to pay tribute to the core members, Prof. Hisazumi Akai, Prof. Hiroshi Katayama-Yoshida, and Prof. Hideaki Kasai. Over the past 20 years, with the help of this provision more than 1,900 participants have learned about first-principles methods, of which about 25% was from the industry. The implementation of first-principles calculations or computational science in the industry is still in its infancy, and it is expected that many more participants will join in the future.

This workshop was partially financially supported by MEXT under the “Program for Promoting Researches on the Supercomputer Fugaku” (DPMSD), “Program for Promoting Researches on the Supercomputer Fugaku (Fugaku Battery & Fuel Cell Project), and “Program for Promoting Researches on the Supercomputer Fugaku” (Quantum-Theory-Based Multiscale Simulations toward the Development of Next-Generation Energy-Saving Semiconductor Devices). Each member of the CMD® workshop is involved in a separate project, and the author is directly involved in one of them. This workshop is also partially supported by the Center for Spintronics Research Networks (CSRN).

ACKNOWLEDGEMENTS

The author would like to thank all the lecturers, tutors, and administrative staff for their cooperation in organizing the CMD® workshop. The author attended MateriApps LIVE! at the Institute for Solid State Physics of the University of Tokyo[12], which was held online before the online CMD® workshop , and used it as a reference for the Slack response protocols and other activities. The activities of the author for the CMD® workshop are partially financially supported by MEXT under the “Program for Promoting Researches on the Supercomputer Fugaku” (DPMSD, Grant No. JPMXP1020200307).

REFERENCES

1. https://cmdworkshop.sakura.ne.jp/index.html.
2. http://susi.theochem.tuwien.ac.at/events/index.html
3. https://www.vasp.at/
4. https://www.quantum-espresso.org/schools-workshops-tutorials/
5. https://greenalmtutorial.mcl.at/index.php
6. https://www.trex-coe.eu/qmcworkshop-agenda
7. P. Hohenberg and W. Kohn, *Phys. Rev*. **136,** B864 (1964).
8. W. Kohn and L. J. Sham, *Phys. Rev.* **140,** A1133 (1965).
9. https://elements-strategy.jp/en/
10. https://ubuntu.com
11. “xeyes” is software that demonstrates the X Window System, with eyeballs appearing and chasing the mouse.
12. http://cmsi.github.io/MateriAppsLive/

This page intentionally left blank