COMPUTING
edge

- Augmented and Virtual Reality
- Automation
- Information Technology
- Edge Computing

JUNE 2020

www.computer.org

IEEE COMPUTER SOCIETY

IEEE

COMPUTING edge

IEEE COMPUTER SOCIETY computer.org

IEEE COMPUTER SOCIETY IEEE Printed with inks containing soy and/or vegetable oils

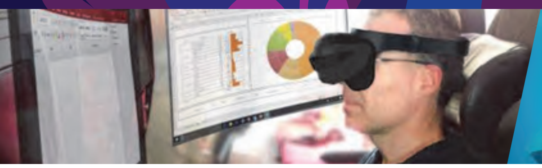
Circulation: *ComputingEdge* (ISSN 2469-7087) is published monthly by the IEEE Computer Society. IEEE Headquarters, Three Park Avenue, 17th Floor, New York, NY 10016-5997; IEEE Computer Society Publications Office, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720; voice +1 714 821 8380; fax +1 714 821 4010; IEEE Computer Society Headquarters, 2001 L Street NW, Suite 700, Washington, DC 20036.

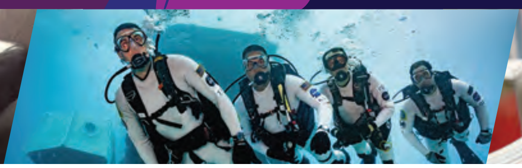
Postmaster: Send address changes to *ComputingEdge*-IEEE Membership Processing Dept., 445 Hoes Lane, Piscataway, NJ 08855. Periodicals Postage Paid at New York, New York, and at additional mailing offices. Printed in USA.

IEEE Computer Society Magazine Editors in Chief

JUNE 2020 · VOLUME 6 · NUMBER 6

COMPUTING edge





8
The Office of the Future: Virtual, Portable, and Global

16
Under Water to Outer Space: Augmented Reality for Astronauts and Beyond

50
Computer Architecture for Orbital Edge Computing

Augmented and Virtual Reality

Automation

Information Technology

Edge Computing

Departments

Magazine Roundup

The IEEE Computer Society's lineup of 12 peer-reviewed technical magazines covers cutting-edge topics ranging from software design and computer graphics to Internet computing and security, from scientific applications and machine intelligence to visualization and microchip design. Here are highlights from recent issues.

Computer

Virtual Avatar-Based Life Coaching for Children with Autism Spectrum Disorder

The rapid development of computer and sensing technologies enables e-coaching systems for individuals with various physical and mental health challenges. In this article from the February 2020 issue of *Computer*, the authors review recent research efforts on using virtual avatar-based computer technologies to improve the social and communication skills of children with autism spectrum disorder.

Computing in Science & Engineering

Jupyter Notebooks as Discovery Mechanisms for Open Science: Citation Practices in the Astronomy Community

Citing data and software is a means to give scholarly credit and to facilitate access to research objects. Citation principles encourage authors to provide full descriptions of objects, with stable links, in their papers. As Jupyter notebooks (JNs) aggregate data, software, and other objects, they may facilitate or hinder citation, credit, and access to data and software. This article from the January/February 2020 issue of *Computing in Science & Engineering* reports on a study of references to JNs in astronomy over a five-year period (2014-2018).

IEEE Annals of the History of Computing

High Noon on the Creative Frontier: Configuring Human and Machine Expertise

In 1960, CBS aired a special entitled "The Thinking Machine," which featured three Western playlets scripted by a computer programmed by MIT researchers. Almost 60 years later, two researchers at Autodesk used a computer program to help design a chair. In this article from the October–December 2019 issue of *IEEE Annals of the History of Computing*, the author links these two seemingly discrete examples of computational creativity to highlight how digital fabrication technologies have served as an important test site for defining human and computational expertise. The author illustrates how concepts of "creativity" and "routine" were produced alongside the concepts of computational creativity during the development of digital fabrication. This dichotomy of "creative" and "routine" is not only used to determine the kinds of tasks that are appropriate for humans and computers to perform within the design and production process, but also used to render invisible the embodied craft knowledge required to substantiate these systems.

IEEE Computer Graphics and Applications

Aggregated Ensemble Views for Deep-Water Asteroid Impact Simulations

Simulation ensembles such as the ones simulating deep-water asteroid impacts have many facets. Their analysis in terms of detecting spatiotemporal patterns, comparing multiple runs, and analyzing the influence of simulation parameters requires aggregation at multiple levels. The authors of this article from the January/February 2020 issue of *IEEE Computer Graphics and Applications* propose respective visual encodings embedded in an interactive visual analysis tool.

2469-7087/20 © 2020 IEEE

IEEE Intelligent Systems

Factual and Counterfactual Explanations for Black Box Decision Making

The rise of sophisticated machine-learning models has brought accurate but obscure decision systems, which hide their logic, thus undermining transparency, trust, and the adoption of artificial intelligence (AI) in socially sensitive and safety-critical contexts. The authors of this article from the November/December 2019 issue of *IEEE Intelligent Systems* introduce a local rule-based explanation method, providing faithful explanations of the decision made by a black box classifier on a specific instance. The proposed method first learns an interpretable, local classifier on a synthetic neighborhood of the instance under investigation, generated by a genetic algorithm. Then, it derives from the interpretable classifier an explanation consisting of a decision rule, explaining the factual reasons of the decision, and a set of counterfactuals, suggesting the changes in the instance features that would lead to a different outcome. Experimental results show that the proposed method outperforms existing approaches in terms of the quality of the explanations and of the accuracy in mimicking the black box.

IEEE Internet Computing

Performance Analysis of Microservice Design Patterns

Microservice-based solutions are currently gaining momentum because they do not have the disadvantages of traditional monolithic architectures. Business interest in microservices is increasing since the microservice architecture brings a lightweight, independent, reuse-oriented, and fast service deployment approach that minimizes infrastructural risks. This approach is at an early stage of its development, and in view of this, it is important to understand the performance of its design patterns. In this article from the November/December 2019 issue of *IEEE Internet Computing*, the authors obtained performance results related to query response time, efficient hardware usage, hosting costs, and packet-loss rate for three microservice design patterns practiced in the software industry.

IEEE micro

High-Quality Fault Resiliency in Fat Trees

Coupling regular topologies with optimized routing algorithms is key in pushing the performance of interconnection networks of supercomputers. In this article from the January/February 2020 issue of *IEEE Micro*, the authors present Dmodc, a fast deterministic routing algorithm for parallel generalized fat trees (PGFTs), which minimizes congestion risk even under massive network degradation caused by equipment failure. Dmodc computes forwarding tables with a closed-form arithmetic formula by relying on a fast preprocessing phase. This allows complete rerouting of networks with tens of thousands of nodes in less than a second. In turn, this greatly helps centralized fabric management react to faults with high-quality routing tables and has no impact on running applications in current and future very large-scale high-performance computing clusters.

IEEE MultiMedia

Modification of Gradient Vector Flow Using Directional Contrast for Salient Object Detection

Scene analysis is a relevant research field for its several applications in the area of computer vision. This article from the October–December 2019 issue of *IEEE MultiMedia* attempts to analyze scene information present in the image by augmenting salient object information with background information. The salient object is initially identified using a method called Minimum Directional Contrast (MDC). The underlying assumption behind using this method for defining salient objects is that salient pixels have higher minimum directional contrast than non-salient pixels. Finding MDC provides us with a raw salient metric. The gradient vector flow (GVF) model of image segmentation inculcates the raw saliency information. The gradient of MDC is calculated and added to the data term of the energy functional of GVF so that the contour formation utilizes not only edge formation but also saliency information. The result gives us not only the salient object but also added background information. Three public datasets are used to evaluate the results. The comparative study of the proposed method for salient object detection with other state-of-the-art methods available in the literature is presented in terms of precision, recall, and F1-Score.

IEEE Pervasive Computing

Design Different: Pen and Paper for Laser Cutting

Interdisciplinary teams and studies need new approaches to design prototypes using tools indistinguishable from the ones they are used to. The authors of this article from the October–December 2019 issue of *IEEE Pervasive Computing* utilize a digital pen and physical paper to build a smart interface for laser-cutters, giving non-technical experienced people the possibility to rapidly, seamlessly, and collaboratively fabricate creative prototypes.

IEEE Security & Privacy

Does Insurance Have a Future in Governing Cybersecurity?

Cyber insurance could achieve public policy goals for cybersecurity using private-sector means. Insurers assess organizational security postures, prescribe security procedures and controls, and provide post-incident services. The authors of this article from the January/February 2020 issue of *IEEE Security & Privacy* evaluate how such mechanisms impact security, identify market dynamics restricting their effectiveness, and sketch out possible futures for cyber insurance as governance.

IEEE Software

Migrating a Software Factory to Design Thinking: Paying Attention to People and Mind-Sets

Design thinking (DT) has found its way into software engineering, promising better requirements elicitation, customer relations, and cohesion within the development team. The authors of this article from the March/April 2020 issue of *IEEE Software* report on Proaction Technologies' migration toward DT and evaluate the process through interviews with employees and clients.

IT Professional

Sending More with Less: Crowdsourcing Integrated Transportation as a New Form of Citywide Passenger–Package Delivery System

Although much effort has been devoted by both academic and industrial communities to improve the efficiency of urban passenger and package flows, current urban transport systems still fail to balance speed and cost. To fill the gap, in this article from the January/February 2020 issue of *IT Professional*, the authors propose a novel form of transport system called crowdsourcing integrated transportation (CIT). It leverages the underused transport capacity, which is generated while delivering passengers to hitchhike packages so that more transportation needs can be met with fewer vehicles and drivers (i.e., sending more with less). They identify the unique features of the new delivery system when comparing to the traditional transport systems and discuss the key research challenges and potential solutions. They implement passenger-occupied taxis as the package carriers and evaluate the effectiveness.

Augmented and Virtual Workplaces

Visual aids on a factory floor. Virtual conference rooms for remote employees. Simulations of customer service scenarios. These are ways that augmented reality (AR) and virtual reality (VR) are entering the workplace. AR and VR technology have the potential to create productive workspaces and provide workers with helpful information in real time. AR- and VR-based training allows workers to learn in realistic, immersive environments and could therefore help reduce accidents and increase adherence to procedures. In this issue of *ComputingEdge*, two articles from *IEEE Computer Graphics and Applications* show how AR and VR are helping people better perform their jobs, whether they work in a traditional office or a less conventional setting.

In "The Office of the Future: Virtual, Portable, and Global," the authors argue that VR workstations will allow people more flexibility in where they work and will make them more productive, because using a VR headset eliminates disturbances and enables consistency and privacy, even in a public place. "Under Water to Outer Space: Augmented Reality for Astronauts and Beyond" describes an AR tool that astronauts-in-training used in an undersea habitat that resembles the International Space Station, showing how AR can help people in extreme environments complete challenging tasks safely and successfully.

Automation can also aid—and sometimes replace—human work. In *Computer*'s "Automated Coding: The Quest to Develop Programs That Write Programs," the author explains the objective of automated programming, its history, its challenges, and why it might be on the verge of becoming a reality. *Computing in Science & Engineering*'s "The March of Kiosks" evaluates how automatic service machines are affecting today's jobs and consumers.

Information technology (IT) jobs are always evolving. "Where the Frontier Thrives: Bricks, Mix, and Zip," from *IEEE Micro*, identifies some factors that determine where technology companies—and therefore technology-sector jobs—grow and thrive. The author of *IT Professional*'s "Working Abroad in a Research Laboratory in the U.S." gives a personal account of his career path in IT.

The final two articles in this *ComputingEdge* issue cover architectures for edge computing. *IEEE Internet Computing*'s "Architectural Considerations for Privacy on the Edge" presents a privacy-protecting edge architecture. Meanwhile, "Computer Architecture for Orbital Edge Computing," from *Computer*, proposes an edge-computing approach to image processing for nanosatellites in low-Earth orbit.

EDITORS: Frank Steinicke, frank.steinicke@uni-hamburg.de
Christian Sandor, chris.sandor@gmail.com

This article originally appeared in Computer Graphics vol. 38, no. 6, 2018

DEPARTMENT: SPATIAL INTERFACES

The Office of the Future: Virtual, Portable, and Global

Jens Grubert, ***Coburg University of Applied Sciences and Arts***
Eyal Ofek and Michel Pahud, ***Microsoft Research Redmond***
Per Ola Kristensson, ***University of Cambridge***

Virtual reality has the potential to change the way we work. We envision the future office worker to be able to work productively everywhere solely using portable standard input devices and immersive head-mounted displays. Virtual reality has the potential to enable this, by allowing users to create working environments of their choice and by relieving them from physical world limitations, such as constrained space or noisy environments. In this paper, we investigate opportunities and challenges for realizing this vision and discuss implications from recent findings of text entry in virtual reality as a core office task.

Much of the hype around virtual reality (VR) has focused on immersive gaming and entertainment, and considerable progress has been made in those directions in recent years. The underpinning thesis in this paper, however, is that recent VR research progress allows us to also reimagine the office work of the future.[1] Raskar *et al.* imagined novel use cases for office work based on projection-based augmented reality, allowing local office workers with remote groups. Immersive head-mounted displays (HMDs) build upon this idea without the need for instrumentation of the environment with projector-camera systems and, hence, enable novel office experiences on the go. VR office based on immersive HMDs open up a novel design space with exciting new opportunities for immersive, flexible, and fluid office work.

Despite the rapid rise of mobile devices such as smartphones and tablets, the traditional workstation and laptop setups still dominate today's office work. Users type text on full-sized physical QWERTY keyboards and use a mouse or trackpad to select and manipulate on-screen objects. Common activities such as typing, editing text, changing the input focus between text fields, switching between windows in an application, and switching between applications use well-established keyboard shortcuts and direct manipulation techniques. Also, in stationary work settings, workers often use multiple screens to create a larger display area. Past research work indicates that large monitors enable more efficient work.[2]

Supporting the above and other typical office activities in a VR environment requires translating the processes of familiar everyday office work practices into efficient and comfortable interaction techniques that simultaneously maximize the advantages posed by VR and minimize its limitations. A further constraint is path dependency: the tendency of users to prefer well-established processes despite being suboptimal in order to minimize learning effort.

A VISION OF VR OFFICE WORK

VR headsets can filter users from the physical world and provide full control of the inputs to their senses, such as visual, auditory, and haptics. This provides several advantages.

Control of the Environment Around Users

Many times, the physical environments surrounding users are clearly suboptimal. The available physical, as well as display space, might be small, and illumination may be less than adequate, resulting in a slew of disturbances all around users. An extreme example might be a person trying to work while sitting in an economy seat on an airplane (Figure 1).

Using VR HMDs, users can work in ideal environments of their liking: wide, well illuminated, private, and with a wide-display area without outside disturbances.

Location-Independent Repeatability of User Experiences

Users who travel frequently might like to keep their familiar work environment constant (for example, the number of monitors, their order and arrangement of the applications around them, the shape of the room, notes on a virtual whiteboard, etc.) even when they are in different places with different physical constraints. This reduces context switching overhead and enables the use of muscle memory by the user during travel: as long as there is access to a table to place a keyboard and a mouse, laptop, or slate, users can carry a large virtual office with them wherever they go.

Virtual displays can recreate a similar arrangement of resources around the user in any location. Even if the recreated VR arrangement may be limited by the physical environment, due to for instance the lack of reachability or real-world haptics, it is possible to identify a VR arrangement that approximates the original one and leverage users' familiarity.

A VR office allows everyday office interactions to transition from locations to temporal events. Interactions can be accessed by temporal events. Instead of a meeting being accessed by the presence in a dedicated meeting room, the meeting can continue from a snapshot of the moment where the last meeting ended: writing still appearing on the whiteboard and all relevant documents being open.

FIGURE 1. A virtual office environment in VR. A user may enjoy a large multidisplay environment and background disturbance reduction, even in challenging environments.

Privacy

Working in public environments exposes the contents of users' screens to unauthorized viewers in her vicinity. Directional visibility filters may lower visibility for people sitting next to the user but do not block all directions, such as people standing behind the user.

HMDs are personal and enable users to work without such privacy implication. Potential access to content can be controlled by the user.

However, privacy is still not fully guaranteed as onlookers, for instance, could observe the user's typing. This opens up interesting research questions on mitigation strategies, such as introducing people around the VR user as avatars or mixed-reality blending of the surroundings with a virtual office.

Relieve Physical World Limitations

The virtual world allows users to do things that are impossible in the physical world. They may move their hands and reach longer distances than their physical hand reach, change their appearance or draw on a whiteboard in front of them while their physical hands are resting on a table, reducing fatigue. Users may travel immediately to a meeting room somewhere else in the world. In the virtual world, there is a potential to equalize differences that may limit users from local resources, distances, or physical capabilities.

CHALLENGES AND OPPORTUNITIES FOR OFFICE WORK IN VR

The above-mentioned vision may hold multiple benefits, yet there are many challenges and technological improvements that need to be addressed to make VR-based office work practical for the public.

Head-Mounted Display Quality

The field-of-view of current HMDs is substantially smaller than a human's field-of-view. The common horizontal field-of-view is around 100°, which is about half of the natural field-of-view, and the vertical view angle is even smaller. Several upcoming HMDs offer 200° horizontal field-of-view and in a few years, we may see HMDs that cover the full field-of-view of the user. Furthermore, the resolution of the HMD display is limited by the need to cover a very large view angle. Currently, this resolution is too low to make it effective for users to read small text, as is possible on a high-resolution monitor. Therefore, current VR applications use larger font sizes, which undermines the VR advantage of a large field-of-view. These limitations will probably be mitigated when new HMDs are introduced.

Another concern is that most of today's HMDs generate 3-D images through stereoscopic image generation (i.e., generating separate 2-D images for the left and right eye) resulting in vergence-accommodation conflicts that can have a negative impact on the user experience and performance in VR.[3] New technologies, such as lightfields or holographic displays, may enable more natural views but have yet to reach consumer product levels.

Most current HMDs are tethered and use external sensors/beacons for tracking, limiting the user to a small volume of operation. This obviously results in a nonmobile VR setup. Although much of the office work might be limited to around a desk area, there is an advantage in allowing users free movement without being restrained by wires, or coverage of room-based sensors. Again, there are already some commercial products that offer inside-out optical tracking, which is independent of environmentally located sensors, as well as wireless transmission of VR content. Further challenges arise from using inertial-based tracking systems in mobile contexts such as cars.[4]

Finally, any error in the tracking of the user's motion or latency in the reaction of the display content to the user motion may increase the risk for the generation of motion sickness.[5] The nauseating feeling rises from a disagreement between the user senses, mainly the visual one and the vestibular system that monitors our balance.

Situational Awareness and Physical Isolation

VR is at one extreme end of the reality–virtuality continuum. This can be beneficial as a user is potentially more immersed in the task at hand and it is plausible this could have positive ancillary effects, such as better concentration and less stress due to the removal of distractions in the environment. On the other hand, VR may also result in a loss of situational awareness and lead to unwanted physical isolation. The current popular applications are entertainment-oriented, and as such, they tend to use the immersive nature of the VR display to replace the user environment with a new one and give the impression of being in a different reality.

The use of VR in a work environment may be a mix use of both reality blocking (removing disturbing elements, having larger screens, etc.) as well as environment representation, enabling manipulation of physical objects, environmental awareness, and communication. Current approaches to move the operating point on the reality–virtuality continuum and use mixed reality to maintain a connection to the physical surroundings ranges from streaming stereo video of the environment to the display (video-based AR) to modeling the environment and representing it in the virtual world.[6] This opens up a rich design space. In this context, an open research question is whether there are any situational awareness or physical isolation issues in VR office work, and if so, how these effects could be quantified and understood in terms of contributing factors. Such investigations can help to identify design principles for future systems.

Fluidity, Flow, and Locus of Control

Users' sense of agency and locus of control is an indicator of usability, as evidenced by its inclusion in user interface guidelines and research on the agency. It is unclear how VR affects users' sense of control of their own actions. Related, flow can be important for office work. It is also unclear whether VR office work is likely to increase or decrease flow.

It is an open question whether effective mitigation strategies that minimize loss of positive VR office work benefits can be identified. It is possible to envision several strands of research, including investigating the relative effects of video-based mixed reality versus optical see-through augmented reality.[7] It may also be interesting to explore minimal interventions in the form of some type of awareness-markers that relate to the physical surroundings that can be subtly introduced in the VR environment. The translation of such awareness-markers to VR need not be graphical, but could also use audio cues or haptic feedback.

Communication Between Users

The need to wear an HMD blocks the view of the user's face from the environment, resulting in loss of an important communication channel between people, although recent research work attempts to recover this channel by methods ranging from virtual avatars representing the users and their facial expressions, internal sensing within the headset, to using prior captured data to better synthesize the view of the user's face. Currently, this is an active field of research.

Typing and Control Efficiency

A key challenge is to minimize the performance gap between ordinary office work, in particular, typing and editing, using a workstation or laptop setup versus a VR setup. Typing is a learned motor skill and recent empirical research has discovered users can be clustered into a small set of different typing styles and type using their own full-sized keyboards at an average rate of 52 words-per-minute, where a word is defined as five consecutive characters including spaces.[8]

In addition to typing text, users also spend considerable effort editing text. This requires interaction techniques that are both fast and precise, which can be challenging if the input is relying on noisy sensor data, such as depth sensing. In contrast, established mice and touchpads provide users with robust control, but at the expense of being 2-D input devices that can be challenging to use for 3-D interaction. Still, text editing on a standard PC is typically conducted using mouse and keyboard and we will indicate later in this paper that text entry in VR can also benefit from standard keyboards.

Furthermore, virtual environments, unbounded by the limitation of a physical world, can introduce new interaction techniques that may prove to be even more efficient than current physical ones. For example, a physical keyboard is limited to lie on a supporting surface such as a table, far away from the display and the edited document. This distance results in large head movement for occasional glancing at the keyboard, slowing down the work, and may generate back and neck pains. In contrast, the virtual keyboard and the user's hands can be remapped from their physical locations to positions closer to the edited documents. Another example involves changing the look and transparency of the user's hands to enable better visibility of the keyboard and the edited document during manipulation (for example, see Figure 3), bottom-left. The design space of such possible alterations of reality is vast.

TEXT ENTRY IN VR

The main focus of our research work so far has been on text entry, as it is fundamental to many tasks ranging from document editing to internet browsing, and a task that has a considerable learning curve (most users are not fluent in touch typing, and still use a various hunt-and-peck and other improvised strategies[7]). In fact, the cost of learning this task has prevented much progress of keyboard technology since the introduction of mechanical typewriters. Among most users, a combination of a traditional keyboard and a large, high-resolution monitor is still the preferred input method for editing longer text documents, working on spreadsheets, or form-filling activities. Given the above observation, we set out to leverage user familiarity with traditional keyboards, and the widespread of such off-the-shelf devices, for work in VR, while using VR freedom of the physical world to improve the user experience.

Initially, it is not obvious that existing physical keyboards or nowadays common touchscreen keyboards

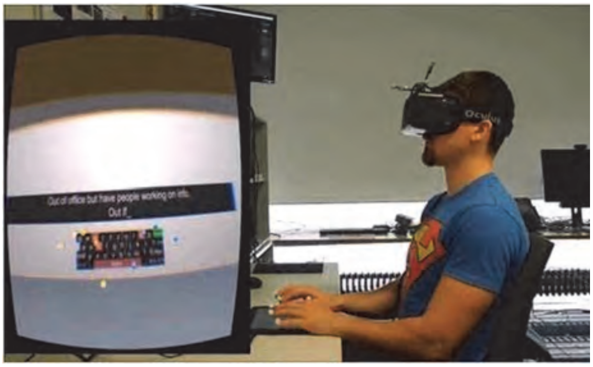


FIGURE 2. Displaying the user's hands in the view direction, rather than at the natural position has the potential to help the user remain focused on the document. It also has little to no impact on typing performance when using a traditional keyboard.

FIGURE 3. Clockwise from top-left: no hands, an inverse-kinematic hand model, a video blending of the user's hands, and fingertips as spheres.

are suitable for typing in VR. The wearable displays block users' view of the real world, including their physical hands and the keyboards, either physical or touchscreen-based, and create a challenge to appropriately represent them in the virtual world. Despite today's VR HMD-limitations, we believe that the ability to control the user's environment, generating virtual displays that are as large as needed, both flat and 3-D, the flexible mapping of the user's interaction space to the virtual space, and the advantage of privacy, may eventually make VR HMDs suitable environments for text entry and document editing. To investigate the potential of today's available hardware, we have studied text entry using standard keyboards (using the QWERTY layout), as described next.

Our user study of the performance of typing on physical and touchscreen keyboards[9] revealed that, while a user's typing speed in a baseline virtual environment is markedly slower than typing in the physical environment, users typed at an average of 60% of their usual typing rate when working in VR. We attribute this loss of speed to two factors: first, the novelty of the setup and user's lack of experience with VR; and second, the limitations of today's VR HMDs (specifically, lower resolution and latency). A key finding, however, is that typing skills transfer seamlessly from the real world to the virtual world.

VR allows the system to situate the keyboard wherever and whenever needed based on context; for example, placing it closer to the document or object of interest, and displaying a graphic representation of the user's hands in relation to the keyboard (in our experiments, we used circles representing the fingertips), see Figure 2. While this eliminates the need to constantly shift attention between the keyboard and document, it may also require the users to reposition their hands while typing. While such repositioning of the keyboards and hands proved to have little impact on typing efficiency with a physical keyboard, it resulted in some degradation of performance on touchscreen keyboards (perhaps due to the change of the direction of the finger motion as they disconnect from the touch surface).

Another freedom VR provides is changing the representation and display of the user's hands in the virtual environment.[10] For example, the user's hands can become translucent in the virtual environment, which might provide an unobstructed view of the keyboard.

We presented users with four different hand representations as they typed in a VR scene (see Figure 3). The first two methods were analogous to traditional input methods; the third and fourth methods used manipulations only possible in VR.

1. A video of the user's hands, which is closest to the natural situation of typing without VR. However, the quality of such video is dependent on the conditions of the physical environment, and it may limit the manipulations that can be

generated in the virtual world, such as movement of the hands in space.
2. A full 3-D model of the users' hands animated according to the tracking of the user's real hands.
3. A minimalistic 3-D model in which most of the users' palms were transparent, and only the users' fingertips were displayed, to maximize the visibility of the keyboard.
4. Only showing the keys being pressed on the keyboard; that is, with hands that are completely transparent.

Surprisingly, the minimalistic model of the transparent hand with only fingertips visible was as easy to use and as efficient as blending a video of the users' hands. Such a model is easy to animate (it only requires sensing of the user's fingertips), and as a 3-D model, it supports a large variety of manipulations in the virtual space. In contrast, the full 3-D model of the hand was not as useful; subtle differences in the model's motions, as well as differences between the look of the model and the actual look of the user's hand, may have generated a dissonance between the user and the model and thereby reduced typing speed and accuracy. In fact, the results of the full 3-D model were as poor as not revealing the hands at all to the user.

BEYOND CURRENT OFFICE TASKS

Text entry and document editing is an important and common task of today's office work, yet many other tasks could potentially benefit from the VR medium. For example, meetings can be independent of distances, travel time, and availability of meeting rooms and their instrumentation. Conversations, recorded by wearable microphones are easier to transcribe and translate, people, objects, and social happenings in virtual spaces can be easier to analyze and describe to people who cannot visually observe the meeting room. Conversations may be mediated to include relevant information or help people challenged in social situations by using the private display of each participant[11] and more (see Figure 4).

Even more exciting might be the opening up of new opportunities that are impossible today, or are limited in their reach. In a VR office, there is practically no importance for the physical location of the users. It may open up jobs for remote people or people with disabilities that were prevented from joining the workforce as equals. VR can enable people literally to see the work from other people's points of view, which may help communication and improve empathy, remote help, education, and reduce misunderstandings and disputes. VR has the potential to better use users' limited attention and mental resources, by minimizing travel and smoothing out transitions between tasks to minimize ramp-up costs, and control external disturbances based on the user's activities and estimated concentration. These are just a few possible future applications and potential benefits. We believe the freedom of the VR world along with very accurate sensing of users' movement, their attention, and behaviors will prove to be fertile ground for more such transformative applications that eventually will reimagine office work as we know it.

FIGURE 4. Private displayed content could support conversations, in particular for people challenged with social interaction.

REFERENCES

1. R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs, "The office of the future: A unified approach to image-based modeling and spatially immersive displays," in *Proc. ACM SIGGRAPH, 1996*, Jul. 1998, pp. 179–188.
2. M. Czerwinski, G. Smith, T. Regan, B. Meyers, G. G. Robertson, and G. K. Starkweather, "Toward characterizing the productivity benefits of very large displays," *Interact*, vol. 3, pp. 9–16, 2003.

3. M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn, “Visual discomfort and visual fatigue of stereoscopic displays: A review,” *J. Imaging Sci. Technol.*, vol. 53, no. 3, pp. 30201-1–30201-14, 2009.
4. P. Hock, S. Benedikter, J. Gugenheimer, and E. Rukzio, Carvr: Enabling in-car virtual reality entertainment,” in *Proc. ACM CHI 2017*, May 2017, pp. 4034–4044.
5. J. Barrett, “Side effects of virtual environments: A review of the literature,” Defense Sci. Technol. Org. Inf. Sci. Lab., Canberra, ACT, Australia, Tech. Rep. DSTO-TR-1419, 2004.
6. D. Lindlbauer, and A. D. Wilson, “Remixed reality: Manipulating space and time in augmented reality,” in *Proc. ACM CHI 2018*, Apr. 2018, Paper 129.
7. M. Krichenbauer, G. Yamamoto, T. Taketomi, C. Sandor, and H. Kato, “Augmented reality versus virtual reality for 3d object manipulation,” *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 2, pp. 1038–1048, Feb. 2018.
8. V. Dhakal, A. M. Feit, P. O. Kristensson, and A. Oulasvirta, “Observations on typing from 136 million keystrokes,” in *Proc. ACM CHI 2018*, Apr. 2018, Paper 646.
9. J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson, “Text entry in immersive head-mounted display-based virtual reality using standard keyboards,” in *Proc. IEEE VR 2018*, 2018, pp. 159–166.
10. J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson, “Effects of hand representations for typing in virtual reality,” in *Proc. IEEE VR 2018*, 2018, pp. 151–158.
11. T. T. Nguyen, D. T. Nguyen, S. T. Iqbal, and E. Ofek, “The known stranger: Supporting conversations between strangers with personalized topic suggestions,” in *Proc. ACM CHI 2015*, Apr. 2015, pp. 555–564.

JENS GRUBERT is an Associate Professor for Human–Computer Interaction in the Internet of Things and Lab Director of the mixedrealitylab, a laboratory for Augmented and Virtual Reality, at Coburg University, Coburg, Germany. He received the Dr. techn degree (2015) with highest distinction from Graz University of Technology, Graz, Austria, the Dipl.-Ing. degree (2009) with highest distinction from Otto-von-Guericke University Magdeburg, Magdeburg, Germany. He is author of more than 50 peer reviewed publications and patents and published a book about Augmented Reality development for Android. His current research interests include interaction with multimodal augmented and virtual reality, body proximate display ecologies, around-device interaction, multidisplay environments, and cross-media interaction. Contact him at jg@jensgrubert.de.

EYAL OFEK is a Senior Researcher with the Redmond lab of Microsoft Research. His research interests include augmented reality, virtual reality, HCI, and haptics. He received the Ph.D degree in computer vision from the Hebrew University of Jerusalem, Jerusalem, Israel. Contact him at eyalofek@microsoft.com.

MICHEL PAHUD received the Ph.D degree in parallel computing from the Swiss Federal Institute of Technology, Lausanne, Switzerland. He won several prestigious awards including the Logitech prize for an innovative industrially-oriented multiprocessors hardware/software project. He joined Microsoft in 2000 to work on many different projects including videoconferencing/networking technologies and research in education. More recently, he has been focusing on human–computer interaction at Microsoft Research. His research includes bimanual interaction, novel form-factors, context sensing, haptics, augmented reality, and virtual reality. Contact him at mpahud@microsoft.com.

PER OLA KRISTENSSON is a University Reader in Interactive Systems Engineering with the Department of Engineering, the University of Cambridge, Cambridge, U.K., and a Fellow of Trinity College, Cambridge, U.K. His research interests include designing intelligent interactive systems that enable people to be more creative, expressive, and satisfied in their daily lives. He is an Associate Editor of *ACM Transactions on Computer-Human Interaction* and *ACM Transactions on Intelligent Interactive Systems*. Contact him at kristensson@acm.org.

Contact department editors Frank Steinicke and Christian Sandor at frank.steinicke@uni-hamburg.de and chris.sandor@gmail.com, respectively.

IEEE COMPUTER SOCIETY

PURPOSE: The IEEE Computer Society is the world's largest association of computing professionals and is the leading provider of technical information in the field.

MEMBERSHIP: Members receive the monthly magazine *Computer*, discounts, and opportunities to serve (all activities are led by volunteer members). Membership is open to all IEEE members, affiliate society members, and others interested in the computer field.

COMPUTER SOCIETY WEBSITE: www.computer.org

OMBUDSMAN: Direct unresolved complaints to ombudsman@computer.org.

CHAPTERS: Regular and student chapters worldwide provide the opportunity to interact with colleagues, hear technical experts, and serve the local professional community.

AVAILABLE INFORMATION: To check membership status, report an address change, or obtain more information on any of the following, email Customer Service at help@computer.org or call +1 714 821 8380 (international) or our toll-free number, +1 800 272 6657 (US):

- Membership applications
- Publications catalog
- Draft standards and order forms
- Technical committee list
- Technical committee application
- Chapter start-up procedures
- Student scholarship information
- Volunteer leaders/staff directory
- IEEE senior member grade application (requires 10 years practice and significant performance in five of those 10)

PUBLICATIONS AND ACTIVITIES

***Computer*:** The flagship publication of the IEEE Computer Society, *Computer* publishes peer-reviewed technical content that covers all aspects of computer science, computer engineering, technology, and applications.

Periodicals: The society publishes 12 magazines and 18 journals. Refer to membership application or request information as noted above.

Conference Proceedings & Books: Conference Publishing Services publishes more than 275 titles every year.

Standards Working Groups: More than 150 groups produce IEEE standards used throughout the world.

Technical Committees: TCs provide professional interaction in more than 30 technical areas and directly influence computer engineering conferences and publications.

Conferences/Education: The society holds about 200 conferences each year and sponsors many educational activities, including computing science accreditation.

Certifications: The society offers three software developer credentials. For more information, visit www.computer.org/certification.

BOARD OF GOVERNORS MEETING

24 – 25 September 2020 in McLean, Virginia, USA

COMPUTER SOCIETY OFFICES

Washington, D.C.: 2001 L St., Ste. 700, Washington, D.C. 20036-4928; **Phone:** +1 202 371 0101; **Fax:** +1 202 728 9614; **Email:** help@computer.org
Los Alamitos: 10662 Los Vaqueros Cir., Los Alamitos, CA 90720; **Phone:** +1 714 821 8380; **Email:** help@computer.org

MEMBERSHIP & PUBLICATION ORDERS

Phone: +1 800 678 4333; Fax: +1 714 821 4641; Email: help@computer.org

IEEE

revised 1 May 2020

EDITORS: Melanie Tory, mtory@tableau.com
Daniel F. Keefe, dfk@umn.edu

This article originally appeared in IEEE Computer Graphics and Applications vol. 40, no. 1, 2020

Under Water to Outer Space: Augmented Reality for Astronauts and Beyond

Benjamin Nuernberger, Robert Tapella, Samuel-Hunter Berndt, So Young Kim, and Sasha Samochina,
Jet Propulsion Laboratory, California Institute of Technology

Augmented reality (AR) has the potential to help astronauts execute procedures in a quicker, more intuitive, and safer way. A key part of realizing these benefits has been the use of an undersea research facility—the Aquarius—that acts as an analog to the International Space Station to a certain extent. In a June 2019 mission, the Aquarius crew successfully executed a complex procedure taking place across four different task areas by using an AR application called ProtoSpace developed at the Jet Propulsion Laboratory. In this article, we share the detailed results of the study, lessons learned, and future work needed to further enable the enhancement of procedure execution through augmented reality.

Augmented reality (AR) has the potential to help astronauts execute procedures in a quicker, more intuitive, and safer way. Traditional instruction methods (such as paper procedures) often introduce a large "cognitive distance" between the informational and physical spaces,[4] inducing a higher mental workload for users when trying to understand and execute procedures. AR instruction methods address this problem by directly overlaying virtual guidance onto the physical world via a head-mounted display, thereby enhancing the user's understanding and execution of the procedure. While the benefits of AR procedural guidance have been shown previously in controlled laboratory settings,[1,4] in this article, we describe a study of using AR for an execution of an actual procedure during a National Aeronautics and Space Administration (NASA) mission.

This study took place during the 23rd NASA Extreme Environment Mission Operations (NEEMO) mission off the coast of Florida in June 2019. During the mission, crew members lived in an undersea habitat (Aquarius), see Figure 1, and were tasked with executing the Sanitation Tank Purge procedure by following AR instructions presented in a Microsoft HoloLens AR device. The goals of the study were to assess the feasibility of using AR instructions for a complex procedure in a NASA mission environment, understand the crew's perceptions of using AR for procedure execution, and to assess various technical and user experience aspects of our AR procedure execution system. Although it was developed specifically for a NASA mission, this study and its results are nonetheless applicable to other application areas, such as industrial and commercial use cases.

There are several challenges that made this study unique and interesting both for space and non-space applications. First, the Sanitation Tank Purge procedure takes place across four different task areas, covering all three main areas inside the undersea Aquarius habitat; thus wayfinding guidance is needed. Along with the close quarters and sometimes cluttered environment of the Aquarius, the fact that the procedure takes place in multiple locations presented an interesting technical hurdle for testing the limits of

DOI No. 10.1109/MCG.2019.2957631
Date of current version 6 January 2020.

FIGURE 1. The Aquarius habitat is 62 feet underwater and 9 km offshore of the Florida Keys; it houses six people (four crew, two habitat technicians) and has approximately 400 square feet of lab and living space.

current AR technology. Second, the sanitation tank purge procedure is a fairly complex task that helped us determine when and where AR is best utilized. Since the crew was executing the actual procedure, habitat technicians monitored the crew during the procedure execution to ensure the task was carried out correctly and safely. Finally, because the undersea Aquarius habitat is an extreme environment that we were unable to physically access to test the system beforehand, the Aquarius is a unique analog to the International Space Station (ISS) that allowed us to verify the readiness of such AR technology for extreme environments such as outer space or in dangerous industrial environments.

BACKGROUND

From a procedure execution perspective, AR attempts to reduce the cognitive distance between the informational and physical spaces.[4,7] With traditional work instructions, the cognitive or informational space is typically presented in paper format or on a tablet display. The difficulty lies in transferring that understanding of how to execute the procedure from the information space onto or into the physical environment. AR attempts to reduce this gap by overlaying the procedural guidance instructions directly onto the physical environment (e.g., by using a head-mounted display). In NEEMO 23, we overlaid virtual text, images, video, attention directors (e.g., arrows and location pins), and 3D animated levers onto and around the physical environment.

Much research has been done to investigate how AR may assist procedure execution in various scenarios.[6,8] In relation to space applications, in 2018, we conducted a controlled user experiment with a mockup of a space flight instrument, the Cold Atom Laboratory (CAL).[1] We compared user performance and feedback between traditional paper instructions and AR instructions for a simple assembly task of mating and demating cables from the instrument. The results were that AR instructions were 19% faster than paper instructions, users reported lower mental and temporal demand (using the NASA-TLX questionnaire[3]), and 14/20 users preferred the AR approach. While the study with CAL was in a controlled laboratory environment, the study presented in this article was in a mission environment, thus providing invaluable observation and feedback from the crew.

NEEMO

The following quote* describes the overall project:

> *"NEEMO—the NASA Extreme Environment Mission Operations project—is a NASA analog mission that sends groups of astronauts, engineers, and scientists to live in Aquarius, the world's only undersea research station, for up to three weeks at a time. Operated by Florida International University (FIU), Aquarius is located 5.6 kilometers (3.5 miles) off Key Largo in the Florida Keys National Marine Sanctuary. It is deployed next to deep coral reefs 62 feet (19 meters) below the surface."*

During typical missions, many science and technology objectives are included, especially in support of ISS and space objectives. Examples of mission objectives include: extravehicular activities; marine science

*https://www.nasa.gov/mission_pages/NEEMO/about_neemo.html

FIGURE 2. Left: Crew diving outside the Aquarius with the habitat technicians inside. Right: Crew, (from left to right) Dr. Csilla Ari D'Agostino, Samantha Cristoforetti, Dr. Shirley Pomponi, Dr. Jessica Watkins.

conducted as a proxy for planetary science concepts and strategies; incapacitated crew member rescue; diver augmented vision display; technology demonstration; medical studies; lunar simulation.

In 2015, the Jet Propulsion Laboratory (JPL) participated in NEEMO 20 by testing Project Sidekick[†] that focused on the AR remote collaboration aspect of procedural execution; Project Sidekick was tested using the Microsoft HoloLens on the ISS in 2016 with Astronaut Scott Kelly. For NEEMO 23, our focus was on testing virtual instructions without the need of remote expert assistance.

The NEEMO 23 crew (see Figure 2) were: Astronaut Samantha Cristoforetti, European Space Agency; Astronaut Candidate Dr. Jessica Watkins, NASA; Dr. Csilla Ari D'Agostino, University of South Florida; and Dr. Shirley Pomponi, Florida Atlantic University. Habitat technicians were Mark Hulsbeck and Thomas Horn.

ProtoSpace

The system we used in NEEMO 23 was built upon an AR collaborative CAD visualization tool called ProtoSpace[‡] that was developed at JPL. Used by JPL and the Johns Hopkins Applied Physics Lab,[2] ProtoSpace is regularly utilized by engineers and scientists to collaboratively discuss early designs of CAD models, plan and rehearse assembly procedures, facilitate cross-team collaborations, and communicate engineering design to the public. Whereas ProtoSpace has typically been used more in the early stages of a product lifecycle, our study in NEEMO 23 focused on a daily maintenance procedure, and thus the AR system presented in this article could be seen as being used for the later stages of a product lifecycle.

STUDY OVERVIEW

The main task of the study was to test our AR procedure execution system for a daily maintenance procedure during the NEEMO mission. Figure 3 shows our system being used inside the Aquarius habitat.

The Sanitation Tank Purge procedure was chosen since it involves maneuvering several levers in four different task areas in the three main areas of the Aquarius habitat: the Main Lock, the Entry Lock, and the Wet Porch (see Figure 4). While some steps are straightforward, several are more complex and time-sensitive, involving pressurizing tanks while monitoring gauges to ensure they are not over pressurized. Previously, training and instructional guidance on how to execute this procedure took place either with paper or PowerPoint instructions.

We had the following four main goals in this study.

1. Assess the feasibility of using AR instructions for this daily NEEMO procedure.
2. Qualitatively compare to other methods (e.g., paper instructions).
3. Understand the crew's perceptions of using AR to engage in everyday tasks.

[†]https://youtu.be/DGoV9mTic4I
[‡]https://youtu.be/dD0FoH8M1EM

FIGURE 3. Left: Dr. Csilla Ari D'Agostino executing the sanitation tank purge procedure while using our AR procedural guidance system. Right: Screenshot of virtual instructions seen by a user of the system.

4. Assess technical and design aspects of our AR procedure system: virtual-to-physical alignment, task instructions, location guidance, and attention directors.

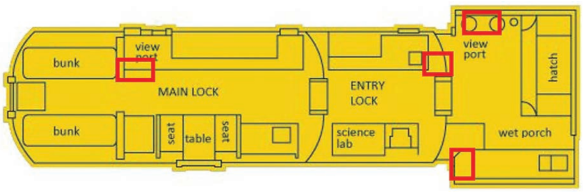


FIGURE 4. Top-down view of the Aquarius Habitat layout with red highlights labeling the task areas.

AR INSTRUCTION INTERFACE

Design and implementation of the AR instructions followed an iterative approach with several rounds of user testing. Since we could not physically access the Aquarius Habitat in developing the system, we created 3D-printed physical mockups of the task areas inside the Aquarius habitat (see Figure 5). We used the Microsoft HoloLens AR head-mounted display and AprilTags[5] to achieve alignment of virtual instructions to the physical world. AprilTags were only scanned in a system setup step and additionally on demand when the HoloLens' spatial anchoring failed to identify their locations.

The system's AR instructions, shown in Figure 6, consisted of the following: A floating virtual menu describing the step number, task and location areas, task instructions, and buttons for changing steps of

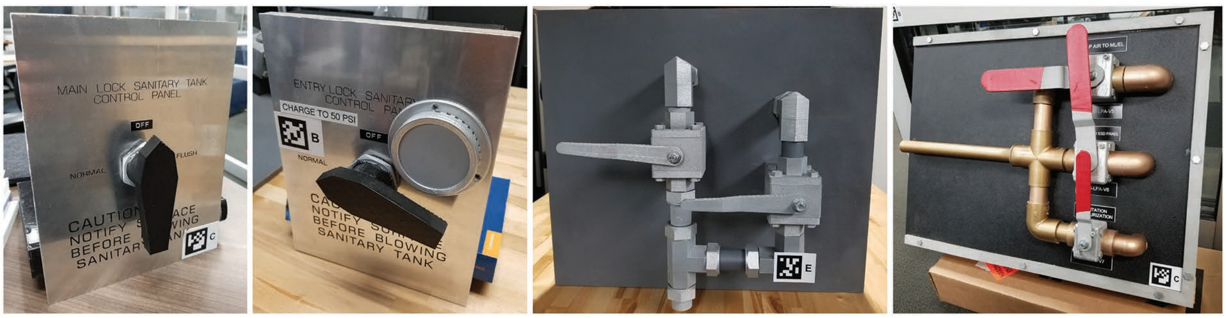


FIGURE 5. Physical mockups (of the four tasks areas) used during development of the AR instructions. These mockups proved to be crucial during iterating on the design and implementation of the system. Also notice the AprilTags[5] used for virtual-to-physical alignment.

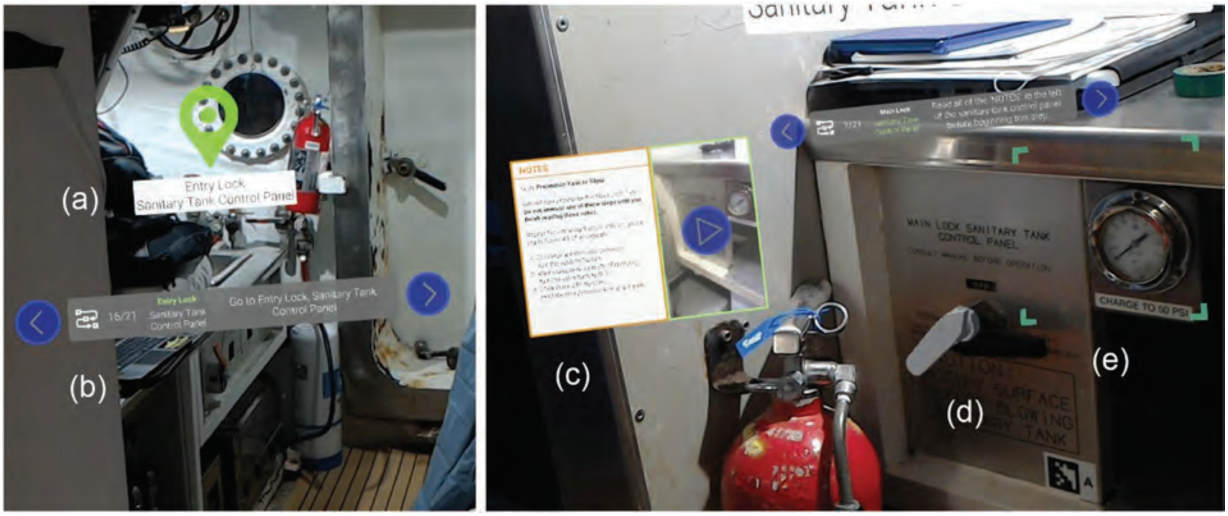


FIGURE 6. Various AR instructions provided to the user. (a) Location guidance. (b) Floating virtual menu. (c) Detailed notes with video. (d) 3D animated levers. (e) Task area guidance.

the procedure; wayfinding guidance to the task area (arrows, green pins, location text); more focused spatial guides at the task area (arrows, bounding box corners); additional instructional guidance through text, images, video, and 3D animated levers. Users could switch steps by using the HoloLens air-tap gesture or by issuing a voice command ("NEEMO next" or "NEEMO previous" to go forwards and backwards through the procedure, respectively).

During preliminary designs, we tried using simple arrows to indicate which way to turn a lever; the motivation behind this was that simple arrows would possibly be easier to author in future procedures to which we would want to apply AR. However, during user testing, we found that such arrows were actually ambiguous to users, perhaps due to the different types of levers found in the Aquarius (see Figure 5). Thus, we instead opted to more often use the more complex but less ambiguous animated 3D levers.

MISSION WEEK

Shortly before the mission began, we found out that a late-breaking update to the procedure instructions would be needed. Since the crew was actually going to be executing the procedure, it was decided that due to this late-breaking change, they would first do a "walkthrough" of the procedure using the AR system with guidance by one of the habitat technicians. After the walkthrough, the crew would actually execute the procedure by again using the AR system, but this time with only minimal help on an as-needed basis from the technician. On June 13, 2019, the NEEMO 23 mission started, and on four separate days each of the four crew members executed the procedure.

Task completion time and speech usage were both recorded on the headsets, and at least one observer watched the crew remotely on shore through a video feed of the habitat. After completing the task, each crew member completed a questionnaire. Finally, after the mission ended on June 21, the crew was interviewed briefly regarding their experience with the system.

Overall Results

All crew members successfully executed the procedure, thus demonstrating the feasibility of using AR instructions for these types of tasks.

Crew members noted that AR greatly helped in spatial wayfinding and task understanding. However, because our system only allowed users to go forward and backwards one step at a time, they noted that paper procedures were more easily navigable and searchable, which may be important for time-sensitive steps. The crew also had mixed feedback on using AR for everyday tasks, with some noting ergonomic and usability issues with HoloLens in general and also with our AR application specifically. They mentioned that time-critical steps should be labeled as such and it should be harder to accidentally exit from those steps in the application. Note that the crew had been trained on using the HoloLens and our application about a month before the mission, and they noted that the preparation was adequate; yet additional refresher training could have alleviated any usability issues.

Finally, the virtual-to-physical alignment and HoloLens tracking mostly worked in spite of the close-quarters and dynamic environment of the Aquarius. Still, the HoloLens could not locate the spatial anchors placed at some of the AprilTags for the last two crew members. This most likely occurred due to walking between the rooms in the Aquarius (i.e., through thick submarine-style hatch doorways) and due to the cluttered Wet Porch area (see Figure 7) where equipment was regularly being moved around.

FIGURE 7. Dynamic wet porch area with diving equipment caused trouble for the HoloLens tracking.

Quantitative Results

Figure 8 shows the task execution times. As previously mentioned, the crew first did a walkthrough of the procedure and then completed the actual physical task. On average, the walkthrough took 15.4 min, the execution took 23.8 min, for a total of 39.2 min. The third crew member's execution took substantially longer than others because she had to realign the virtual to the physical by scanning AprilTags; note that the crew was only briefly trained on how to achieve the alignment, which would explain why her time to align took several minutes.

For voice command usage, Figure 9 (left) shows that speech was used most of the time for switching steps

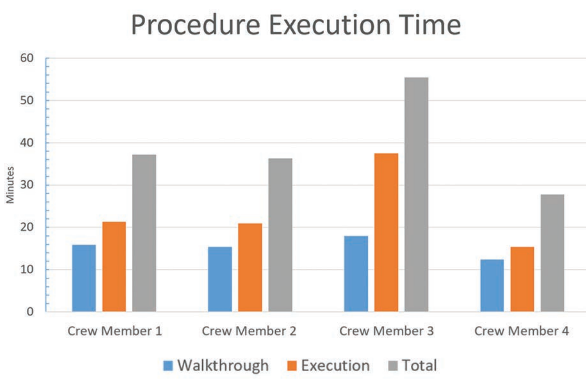
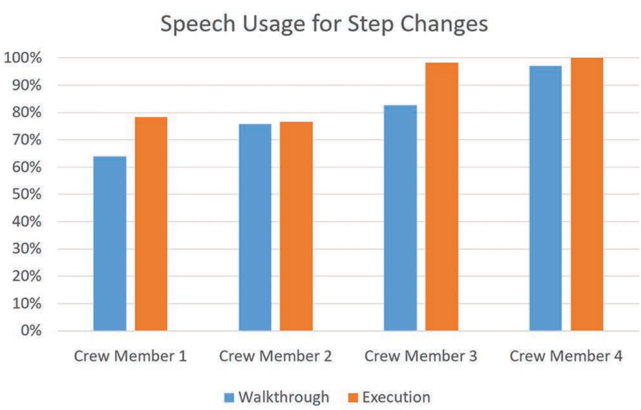


FIGURE 8. Execution time for both the walkthrough and execution of the procedure.



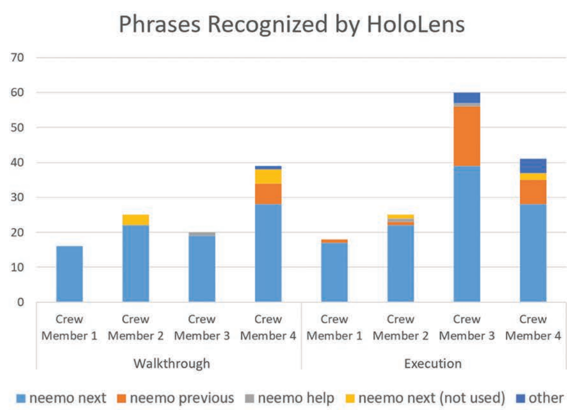


FIGURE 9. Left: Percentage of step changes where speech was used to change the step compared to using air-tap gestures. Right: Number of phrases recognized by HoloLens in our application.

(the other option was an air-tap hand gesture). The fourth crew member exclusively used speech during her procedure execution.

It is also interesting to see which phrases were recognized by the HoloLens. Figure 9 (right) shows the various phrases recognized by HoloLens during both the walkthrough and the execution for each crew member. Note that three phrases were not applicable in this procedure (noted by "other" in the figure); in the future, such voice commands should be disabled if possible.

We also note that "neemo next (not used)" is referring to when that phrase was recognized but the user had just changed steps within a couple of seconds (we set a time threshold for switching steps to avoid users accidentally skipping steps if two voice commands or two air-taps were issued immediately after one another) and therefore that second "neemo next" did not change the step.

Finally, we note that the third and fourth crew members used "neemo previous" several times, either indicating that the system accidentally recognized that phrase or that the crew needed to go backwards in the procedure. For the latter, we should design the system so that it is easier to navigate backwards in a procedure if this is common.

Crew Quotes

Finally, notable quotes include:

> *"The 3D renderings showing how to actuate the valves are great and are much more efficient than words. The virtual direction guidance within the space is also extremely helpful. The one aspect that could be improved is that time-critical steps should be clearly labeled as such (with the appropriate level of urgency), and the next step(s) should be available for view prior to starting the time-critical steps. One simple solution may be to have related steps appear in action groups rather than always one step at a time."*

> *"Unfortunately, the headset is still quite bulky/heavy/awkward to wear for a long period of time, and quickly removes a lot of the efficiencies and advantages of using the system."*

Regarding where a system like this would be most helpful:

> *"Long procedures when not time critical. E.g., Putting together equipment;" "Unfamiliar users that need to be guided to locations and/or specific components. If this is to be used for critical operations, safety features to make sure that steps are not missed/skipped should be embedded. For example with a reference to the step number. A couple of times I had the feeling that it jumped ahead (maybe because I had to repeat NEEMO next and eventually it reacted to it multiple times)."*

CONCLUSION

Enhancing procedure execution through AR has enormous potential. In this article, we showcased some of those potential benefits in an extreme mission environment, outside of a controlled laboratory setting. The NEEMO Aquarius habitat provided various challenges—close quarters, high humidity, dynamically changing areas—all working to push the limits of current AR technologies. The HoloLens tracking functioned fairly well yet had some trouble relocating the task areas in the Wet Porch (see Figure 7). All crew participants were able to successfully execute the procedure, showing the feasibility of using AR for such daily maintenance routines. On the other hand, there is still room for improvement, specifically in easy navigability of all the procedure steps in the user interface.

In retrospect, several important lessons were learned. First, as with developing any user interface, iteration in design and implementation was crucial; by using physical mockups, we were able to iteratively alleviate usability issues to ultimately make the study a success. Second, extreme environments will often have unforeseen changes needing to be made with a quick turnaround; since we used an XML file separate from the HoloLens application to define the procedure, we were able to quickly make needed changes without having to rebuild our application.

In the long term, in order for AR autonomous assistance to be successfully used by astronauts and in other industry settings, AR instructions will need to be authored in an easy way. Currently, it is easier to use traditional methods to author a procedure (e.g., writing a document) than to create an AR version of those same instructions. Having the ability to easily create

AR instructions will be key for ultimately realizing the benefits of AR in procedure execution.

ACKNOWLEDGMENTS

This work was supported by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). Early stages of this work were done in collaboration with L. Wang and C. Russell at the Johnson Space Center. The authors would like to thank Rob Ray for creating the physical mockups of the task areas. Special thanks to NEEMO 23 Mission Director Marc Reagan as well as the entire NEEMO 23 team for a successful mission.

REFERENCES

1. A. M. Braly, B. Nuernberger, and S. Y. Kim, “Augmented reality improves procedural work on an International Space Station science instrument,” *Human Factors*, vol. 61, no. 6, pp. 866–878, 2019.
2. D. J. Hanhe, F. Ruiz, N. W. DeMatt, and A. R. Fagan, “Risk reduction of integration and testing operations using augmented reality on Parker Solar Probe,” in *Proc. 30th Space Simul. Conf.*, Institute of Environmental Sciences and Technology, Annapolis, MD, Nov. 2018. [Online]. Available: http://www.spacesimcon.org /conference.html
3. S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” in *Advances in Psychology* (Human mental workload), vol. 52. Oxford, England: North-Holland, 1988, pp. 139–183.
4. 4.S. J. Henderson and S. K. Feiner, “Augmented reality in the psychomotor phase of a procedural task,” in *Proc. IEEE Int. Symp. Mixed Augmented Reality*, 2011, pp. 191–200.
5. E. Olson, “AprilTag: A robust and flexible visual fiducial system,” in *Proc. IEEE Int. Conf. Robot. Autom.*, 2011, pp. 3400–3407.
6. R. Palmarini, J. A. Erkoyuncu, R. Roy, and H. Torab-mostaedi, “A systematic review of augmented reality applications in maintenance,” *Robot. Comput.-Integr. Manuf.*, vol. 49, pp. 215–228, Feb. 2018.
7. M. E. Porter and J. E. Heppelmann, “A manager’s guide to augmented reality,” *Harvard Bus. Rev.*, vol. 95, no. 6, pp. 45–57, 2017.
8. X. Wang, S. K. Ong, and A. Y. C. Nee, “A comprehensive survey of augmented reality assembly research,” *Adv. Manuf.*, vol. 4, no. 1, pp. 1–22, 2016.

BENJAMIN NUERNBERGER is currently an AR Researcher and Developer at the Jet Propulsion Laboratory, Pasadena, CA, USA. He received a Ph.D. in computer science from the University of California, Santa Barbara, CA, USA, and he currently serves as the Technical Lead of ProtoSpace. Contact him at benjamin.nuernberger@jpl.nasa.gov.

ROBERT TAPELLA is currently a UX Designer and User Researcher at the Jet Propulsion Laboratory, Pasadena, CA, USA. He has more than a decade of experience in product development and commercialization in science industries. Contact him at robert.k.tapella@jpl.nasa.gov.

SAMUEL-HUNTER BERNDT is currently an Immersive Software Systems Engineer at the Jet Propulsion Laboratory, Pasadena, CA, USA, who serves as a Developer and Scrum Master for ProtoSpace. He received a master’s degree in computer science from Michigan State University, East Lansing, MI, USA. Contact him at samuel-hunter.berndt@jpl.nasa.gov.

SO YOUNG KIM is currently a Senior Lead Designer at the Jet Propulsion Laboratory, Pasadena, CA, USA. She has been leading Human-Centered Design efforts on projects such as designing AR systems for space mission applications and products for model-based systems engineering for spacecraft design. Her specialty is Human-System Integration and, in particular, Human-Autonomy Teaming. She received a Ph.D. in aerospace engineering from the Georgia Institute of Technology, Atlanta, GA, USA. Contact her at soyoung @jpl.nasa.gov.

SASHA SAMOCHINA is currently the Deputy Manager of the Ops Lab and the Project Lead of ProtoSpace at the Jet Propulsion Laboratory, Pasadena, CA, USA. She is a Creative Technologist with a BFA from School of the Art Institute of Chicago. She leads innovative work to help scientists, engineers, and astronauts transform their workflow, and leverages that technology to educate and inspire the public to engage in STEAM. Contact her at sasha.samochina @jpl.nasa.gov.

Contact department editors Melanie Tory at mtory@tableau .com and Daniel F. Keefe at dfk@umn.edu.

EDITOR: Mark Campbell, Trace3; MCampbell@trace3.com

COLUMN: IT INNOVATION

This article originally appeared in
Computer
vol. 53, no. 2, 2020

Automated Coding: The Quest to Develop Programs That Write Programs

Mark Campbell, *Trace3*

Computer scientists have long pondered the possibility of crafting systems capable of creating programs directly from human intent. Recent developments in neural networks are closing the gap between fuzzy user objectives and concrete automated code generation.

Today, "computer-generated" is a mainstream adjective for many forms of media from music to art, literature, voice, conversation, images, animation, sensor data, and even deep-fake videos. One of the last digital-automation frontiers is the automatic synthesis of complete programs or applications. Recent developments, however, are closing the gap between human-generated concepts and concrete code generation.

Automated coding has meant a variety of behaviors over the evolutionary lifetime of computers, but the core objective of automated coding is the generation of machine-executable code from high-level design. This seemingly straightforward task has proven exceedingly tricky. For example, when translating concepts from English to French, a 95% accuracy rate is acceptable to get a tourist to the Eiffel Tower. However, translating concepts from English to a coronary surgical robot's control code requires considerably more precision, especially if you are the patient. Today, we see the advent of complex systems that leverage a variety of technologies to enable automated coding with impressive accuracy.

The quest for automated programming began before digital computers even existed. During World War II, Alan Turing and Doug Michie spoke of advanced computing machines created for "solving problems by means of searching through the space of possible solutions, guided by rule-of-thumb principles."[1] This is what we currently call *heuristic search*. With the emergence and maturation of computability theory through the early to mid-20th century, modern programming languages evolved from paper to raw machine code, to symbolic assembly code, to higher-level formalism, allowing for ever more abstract human concepts to be expressed and translated into operational commands by assemblers, compilers, and interpreters.

Although most computer scientists would not instinctively consider a compiler as an autonomous programmer, its role in automatically generating binary machine code from human intelligible instructions is indeed automated coding. When Fortran was released in 1957, it was called *The Fortran Automatic Coding System*, and its stated goal was "to enable the programmer to specify a numerical procedure using a concise language like that of mathematics and obtain automatically from this specification an efficient [IBM]704 program to carry out the procedure."[2] The Fortran team even pondered the benefits of a computer that could "code problems for itself and produce as good [of] programs as human coders (but without the errors)"[2]—certainly a lofty goal in a pre-Sputnik era.

If compilers, interpreters, and other high- to low-level code translators can be called the first wave of automated coding, then the second wave began with the advent of model-driven development (MDD) tools. MDD applications generate high-level code from human concepts represented in textual or graphical

DOI No. 10.1109/MC.2019.2957958
Date of current version: 12 February 2020

models. In 1982, the Massachusetts Institute of Technology's (MIT's) Artificial Intelligence Lab released The Programmer's Apprentice, which developed programmer intent into a language-independent "plan representation" of the intended program. From this, a knowledge-based editor and coder would produce a program from other plan fragments plus a library of standard algorithms and data structures.[3] MathWorks extended model-based design with the 1984 release of Simulink, a graphical programming environment, to take a systems designer's graphical models through design, simulation, analysis, and ultimately production-quality C and hardware description language code generation for myriad dynamic systems.[4]

Model-based automated development progressed through the 1990s and 2000s with the emergence and maturation of model-centric engineering methods, like the Unified Modeling Language (UML), which allows engineers to develop structure, behavior, and interaction models using graphical and textual notation. UML can not only generate class, function, and data directly from these models, but it can also conversely import existing or modified code back into the model in a process called *round-trip engineering*. Despite UML being a leading design paradigm and a prolific code generator, it does have its limits. "Code generation was not front and center when creating UML," said Grady Booch, codeveloper of UML, when I interviewed him recently. "The continuous sync and regeneration of round-trip engineering can become pragmatically very difficult," he pointed out.[5]

While MDD was proficient at generating the general structure of a program, it left the low-level detailed programming to the developer. Essentially, the skeleton of the program was in place but not the flesh. Sketching emerged as one approach to span this gap. With sketching, a developer describes the desired implementation strategy using partial programs written in a sketch language. This sketch is then fed into a synthesizer, which derives the missing details and produces a working implementation that satisfies the defined criteria.[6] Sketch-based program synthesis remained mostly a theoretical approach until recent advancements in deep-learning, neural-network technologies enabled a wave of new solutions to automate the coding of both the overarching program structure and the low-level program details—both the bones and the guts.

Inductive program synthesis (IPS), exemplified by DeepCoder, trains a neural network on a large repository of existing inputs, source code, and output examples. Once trained, DeepCoder uses a gradient descent to search for the source code that best fits the given input and output sets.[7] In 2017, a Microsoft and MIT team exhibited RobustFill, a recurrent neural network that generated low-level program code from input and output examples even when the input was not pristine.[8]

Building on IPS, search-based synthesis techniques were developed such as the Bayou Project at Rice University. Funded by DARPA and Google, Bayou uses neural sketch learning and a specialized neural network called a *Gaussian encoder–decoder* trained to associate programming tasks with high-level patterns found in a vast real-world body of open source code. A user need only provide a sketch consisting of a draft program and a query describing the intent of the overall programming task. From these, Bayou infers a probability distribution across myriad code fragments and application programming interfaces on which it has been trained, giving higher weight to those more likely to produce the requested behavior. Using this technique, Bayou often predicts the entire program body given just a few keywords or prompts.[9]

Yet another promising solution is SketchAdapt, a collaboration of MIT's Computer Science and Artificial Intelligence Laboratory and Center for Brains, Minds and Machines. SketchAdapt is a symbiotic intertwining of two neural networks to tackle both high-level structure and low-level detail. Like Bayou, SketchAdapt starts from a user-provided sketch. This sketch is fed to a neural network that recalls the best structures that match the sketch's intent. From this, the network generates as much code as possible, skipping over any areas it does not know how to code. These "holes" are then handed over to a Bayou-like search-based synthesis engine to generate the missing low-level code.[10] In a recent interview, cocreator Dr. Armando Solar-Lezama commented on SketchAdapt's "ability to learn from a large corpus of programs to create top-down intuition, if you will."[11]

Although today's SketchAdapt is limited to small solutions, its combination of top-down intuition and semantic search show very promising results. Recently SketchAdapt outperformed all other programs in translating mathematic problems from English to code, making it every seventh grader's dream. While still in its infancy, SketchAdapt's hybrid model of structural pattern matching and symbolic reasoning demonstrates one of today's most inviting avenues of program synthesis exploration.

However, today's applications are composed of more than just code-based components. Artificial-intelligence (AI) models are now found at the core of a growing number of applications, and several efforts are underway to automate AI-model generation from sketches. Like the ouroboros eating its own tail, automated machine learning (AutoML) is a process that uses neural networks to generate neural networks. The emerging field of neural architecture search (NAS) allows systems like Google's Cloud AutoML to search a compendium of neural network architectures and select the design with the highest expected performance that satisfies the user sketch, thus enabling even inexperienced users to construct complex custom ML models.[12]

Programmatic source code, whether generated automatically or by humans, does not live in a vacuum; it must be inspected, tested, and documented. Not surprisingly, many technologies used to generate programs can also inspect source code statically or create test harnesses and test data to verify the program's behavior dynamically.[13] CodeLingo translates the whole software stack into a graph that is searched for patterns indicative of defects, incorrect coding style, refactoring opportunities, and optimizations. It can even develop embedded comments and external documentation.[14] DeepCode (not to be confused with DeepCoder mentioned earlier) has developed an AI-powered code-assessment tool that, unlike traditional code-review solutions, not only detects syntax mistakes but also reports deviations from the inferred intent of the code based on its analysis of millions of open-source examples and thousands of programming concepts. In essence, DeepCode reverse-engineers a sketch of the program and compares its actual implementation to this derived sketch.[15]

It is reasonable to predict that AutoML techniques, like NAS, will soon be spliced into hybrid systems, like SketchAdapt. Such a bones, guts, and brain approach would be able to take simplistic user sketches and automatically develop a high-level programmatic structure fleshed out with AI models and search-based program generation to fill in the functional gaps. For such a system to emerge, existing methods, such as NAS, search-based synthesis, and sketch interpretation, will certainly need to mature and incorporate further advancements. But the evolution of these techniques is already well underway. The flywheel of innovation will also undoubtedly continue to spin off additional novel techniques to be added to the automated-coding pipeline.

Automated coding has been a goal of software engineers for more than 80 years, and recent developments are bringing us tantalizingly closer to the realization of that dream. Automated coding may never fully replace programmers, but as the automation tool kit deepens, human fingers will create an ever-shrinking slice of the coding pie.

REFERENCES

1. X.-S. Yang, Ed., *Artificial Intelligence, Evolutionary Computing and Metaheuristics: In the Footsteps of Alan Turing*. Heidelberg, Germany: Springer-Verlag, 2013.
2. J. Backus et al., "The FORTRAN automatic coding system," in *Proc. Western Joint Computer Con, Los Angeles*, 1957. doi: 10.1145/1455567.1455599.
3. R. E. Handsaker. 1982. "Code generation in the programmer's apprentice." *MIT Artificial Intelligence Laboratory Working Papers, WP-233. MIT Artificial Intelligence Lab.*, May 1982.

4. J. Krizan, L. Ertl, M. Bradac, M. Jasansky, and A. Andreev, "Automatic code generation from Matlab/ Simulink for critical applications," in *Proc. 2014 IEEE 27th Canadian Con Electrical and Computer Engineering*. doi: 10.1109/CCECE.2014.6901058.
5. M. Campbell., "*Model-driven development interview with G. Booch*," private communication, Nov. 11, 2019.
6. A. Solar-Lezama, "The sketching approach to program synthesis," in *Proc. Asian Symp. Programming Languages and Systems*, 2009, pp. 4–13.
7. M. Balog, A. L. Gaunt, M. Brockschmidt, S. Nowozin, and D. Tarlow, "DeepCoder: Learning to write programs," in *Proc. Int. Conf. Learning Representations* 2017, Toulon, France.
8. J. Devlin, J. Uesato, S. Bhupatiraju, R. Singh, P. Kohli, and A.-r. Mohamed, "RobustFill: Neural program learning under noisy I/O," in *ICML'17 Proc. 34th Int. Conf. Machine Learning*, vol. 70, Aug. 6–11, 2017, pp. 990–998.
9. V. Murali, L. Qi, S. Chaudhuri, and C. Jermaine, "Neural sketch learning for conditional program generation," in *Proc. ICLR* 2018, Houston, TX.
10. K. Martineau, "*Toward artificial intelligence that learns to write code*," MIT News, June 14, 2019.
11. M. Campbell, "*Interviewee: Automated coding interview with A. Solar-Lezama*," private communication, Oct. 31, 2019.
12. X. He, K. Zhao, and X. Chu, "*AutoML: A survey of the state-of-the-art*," Dept. Science, Hong Kong Baptist Univ., Hong Kong, 2019.
13. Y. Liang, Q. Liu, J. Xu, and D. W. Wang, "The recent development of automated programming assessment," in *Proc. 2009 Int. Con Computational Intelligence and Software Engineering*, Wuhan, China. doi: 10.1109/ CISE.2009.5365307.
14. R. D. Caballar, "*New Zealand startup seeks to automate (most) code review*," IEEE Spectrum, May 31, 2019.
15. N. Lomas, "*DeepCode gets $4M to feed its AI-powered code review tool*," TechCrunch, Aug. 6, 2019.

MARK CAMPBELL is the chief innovation officer for Trace3. Contact him at MCampbell@trace3.com.

This article originally appeared in
computing in SCIENCE & ENGINEERING
vol. 20, no. 6, 2018

The March of Kiosks

Charles Day, *The American Institute of Physics*

The word "kiosk" is one of the few in the English language to come from Turkish. When I was a nipper in Wales, it connoted a roofed, gazebo-like structure that housed a tiny shop or booth. Now in the U.S. and elsewhere, it connotes a freestanding automatic service machine.

Although the new definition of kiosk is somewhat flexible, the first kiosk-like machine to become widely deployed was the ATM. The card-activated cash dispenser made its debut on June 27, 1967, at a North London branch of Barclays Bank. I began using ATMs in the early 1980s when I was an undergraduate in London.

Despite the proliferation of kiosks at airports and elsewhere, I had not thought much about the machines since then until one of my friends, a historian of science, posted a one-sentence rant on Facebook earlier this year: "I hate kiosks!" Although he did not elaborate, I could guess at his frustration. The last time I used a self-checkout kiosk at the grocery store, I placed a bunch of green onions on the scale. The attached label said "green onions." But when I scrolled through the kiosk's touchscreen menu, I could find neither "green onions" nor "onions, green." The vegetables were listed in the database as "scallions."

The ability to look up large numbers of items, albeit imperfectly, illustrates the power of kiosks. They tap into large databases and compute in real time. They are also attached to sensors such as laser scanners, piezoelectric scales, and touch screens. None of those technologies is new. Kiosks in grocery stores lagged ATMs on high streets because of business conditions. In the District of Columbia, where I live, and its surrounding suburbs, installing kiosks in grocery stores became cheaper than employing human tellers about five years ago.

The most recent expansion of kiosks in D.C. has been into fast food. This past September on NPR's *Morning Edition*, business reporter Ally Schweitzer recounted her experience using a kiosk at a McDonald's restaurant in Arlington, Virginia. Surprisingly, saving money was not the motivation for installing the kiosks. Rather, customer convenience was paramount. The restaurant redeployed its humans to deliver food to tables and to people who ordered takeout.

How much further will kiosks encroach? Hotels retain humans for check-ins but use kiosks or other automatic methods for checkouts. The choice suggests that humans remain favored over machines for making guests feel welcome. But for transactions that do not require warmth and other human qualities, the kiosks' advance will continue. However, whether we grow to hate them, as my friend does, could depend on how much interaction with human strangers we really want.

CHARLES DAY is *Physics Today's* editor in chief. The views in this column are his own and not necessarily those of either *Physics Today* or its publisher, the American Institute of Physics. Contact him at cday@aip.org.

This article originally appeared in IEEE micro vol. 39, no. 2, 2019

Where the Frontier Thrives: Bricks, Mix, and Zip

Shane Greenstein, *Harvard Business School*

A visitor to Manchester England in 1840 knew they stood at a major cloth manufacturing center, and that it was changing the world. Pittsburgh in 1892? A frontier steel manufacturing center, and (again) changing the world. Detroit in 1925? Frontier automaking, also changing the world.

Where are the equivalent frontier locations today for information technology? The San Francisco Bay Area and New York City readily come to mind. So do other cities in North America. Austin, Seattle, and Toronto also produce leaders. Around the globe? Tel Aviv, Seoul, Bangalore, and several cities in China.

It is not just an academic question. The answer has important consequences for where economic growth takes place. This topic has played itself out in the last decade, for example. Two additional technical options became available to users, broadband and smartphones. They diffused. Leading suppliers emerged in a few places, and, largely not in most places. We are in the midst of watching a third new deployment, machine learning. Which firms will lead? Where? Not clear yet.

Many business development directors from many places claim that they are one or two innovative products away from becoming a world-changing center. How should an objective observer sort between a reasonable claim and a ridiculous one? There is shorthand for the economics of the answer: bricks, mix, and zip.

READINESS

Avoid a common narrow conception of this topic. Many observers focus on the features of a winning firm. What do they see? Technically facile organizations who are trying to solve big user problems, and meet common user needs, and at tremendous scale, and with processes that approach—or beat—world standards.

Step back and take in a wide view. Why do we NOT see those firms everywhere? Why do only a few places give birth to them?

The answer starts with an obvious observation: People with high aspirations sometimes live in locations that do not nurture their ambitions. Some stick it out, and mostly fail. Others leave and go to places that raise their chances of success, and, again, mostly fail. That yields the first point: this topic is all about raising chances, and any one success has many causes. It also stresses that success is rare.

More to the point, it illustrates an irony of virtual economics, namely, the virtual follows the corporal. The corporal assets that support bricks and mortar can make a place more or less attractive.

Consider the three types of IT assets a human can touch and see. One type carries data. Think of lines under the ground, antennas for feeding smartphones, switches routing the traffic, and thousands of rights-of-way permits to allow them. Another type consists of other large assets for storage and computing and data exchanges. Think of data centers and content delivery networks, and all that goes into operating those, such as large scale ac and also inexpensive and reliable electrical supply. The third type consists of privately owned IT hardware for business. Think of servers, local area networks, WiFi routers, and gazillions of miles of cable in walls.

Entrepreneurship is a symptom of the place that nurtures high aspiration. Some locations will spring a disproportionate share of great outcomes from entrepreneurs. Recognize that as an effect, not a cause.

By itself, bricks do not make a place attractive. It also needs a mix of appropriate software for the frontier hardware. That is so obvious, we usually take it for granted. One type of software comes from open

DOI No. 10.1109/MM.2019.2897516
Date of current version 15 March 2019.

sources and industry consortia, while another comes from private suppliers. The hardware that touches infrastructure and carriers involves important components from the Internet Society, the World Wide Web Consortium, and the Apache Foundation, and involves suppliers affiliated with Ngnix and Linux, which involves a combination of unpriced and priced services. As one gets closer to privately owned hardware, and behind firewalls, the software tends to be middleware, and more private too—either from third parties or written internally. That comes from Microsoft, Oracle, SAP, IBM, and AWS, and a huge number of others.

That software does not get there on its own. A mix of labor installs, maintains, modifies, and operates it. A vibrant location contains a wide mix of experienced and new labor—programmers who can program in C++ and java script, and those who can resolve issues with load balancing and manipulate visualization tools for internal reports, and more. While the presence of open source software, per se', does not seem to define these work forces, such expertise does correlate with the presence of this mix of skills, which is an essential ingredient for world changing aspirations. The symptoms are readily visible too: Not long ago all the vibrant places contained an abundance of programmers in Hadoop, and, with this most recent wave, familiarity with Tensor Flow.

A different dimension also matters. Mix fosters a variety of viewpoints. Variety is a strength in uncertain technical settings. It fuels different outlooks about new opportunities. Great areas contain a range of firms, drawn from a range of industries, comprised of a range of sizes, who hire a variety of IT employees for their input supply and for their distribution networks.

What is zip? A vibrant location gets an extra zip from the presence of professionals with know-how, whose business concerns the act of translating innovation into commercially valuable services. This comes from, again, a mix of specialists who work with the STEM fields—lawyers with a penchant for giving sage intellectual property advice, engineers with managerial experience taking an enterprise from 50 to 500 employees, experts on how to finance investment without increasing the risks, marketing advisors with knowledge about running online ad campaigns, and so on. Many local firms with similar needs share this expertise. It gives vibrant areas an extra ability to push many firms into commercializing frontiers, and with fewer frictions.

One other dimension of zip is ambition, though it is somewhat elusive to define. Usually distributed among an elite class of experienced managers and entrepreneurs, great places also leave room for new arrivals, and accept them readily. You might reasonably say: ambitious about what? To which I say: these people are rare, and you know them when you meet them. They talk about changing the world, and they do not sound naive. They can describe in detail the activities that will accomplish those goals.

ENTREPRENEURSHIP IS A SYMPTOM OF THE PLACE THAT NURTURES HIGH ASPIRATION. SOME LOCATIONS WILL SPRING A DISPROPORTIONATE SHARE OF GREAT OUTCOMES FROM ENTREPRENEURS. RECOGNIZE THAT AS AN EFFECT, NOT A CAUSE.

In my experience, these aspirants—got a better label?—can thrive in large firms or startups. It is not the size. It is the environment. Such people happily give up a few dollars to work with other dreamers on a meaningful development project. They seek the world-changing efforts, and those are in world-changing locations.

Entrepreneurship is a symptom of the place that nurtures high aspiration. Some locations will spring a disproportionate share of great outcomes from entrepreneurs. Recognize that as an effect, not a cause.

OVERVIEW

Today high aspirations live in a crowd. That is unusual by historical comparison. Historians describe the way U.S. Steel dominated Pittsburgh, and how a small set of auto firms dominated Detroit. Nobody today talks about one firm dominating a vibrant city. Maybe the closest comparable situation today is the way Microsoft and Amazon dominate Seattle, which is notable for its rarity. Today the ownership structure for the supply chain for IT lies in multiple hands in many locations. The location developed the factors in response to many forces, not because a single firm asked for it.

The opposite also holds. The mix has to be nearly present for independent reasons, before the aspirations become common. That means small regions are at an inherent disadvantage because they simply cannot get big enough to have that mix. To be sure, it does not prevent world-changing aspirations from thriving in cities such as Austin and Tel Aviv. It does mean smaller places are out of the running.

Because this is a global phenomenon, it is important to finish with remarks about rules and regulations and norms. Those matter because hardware and software have to work together in private businesses, where businesses make investments with the intention to see a return. Things break and patches emerge for code. An investor must have a reasonable assurance that their operations will work, and that their business suppliers will not jerk them around (too much). Everybody has to play by the (almost) same rules.

Anybody who travels widely knows of countries where that does not happen. There are countries where courts do not redress disputes, government regulators take actions on their own sweet time and without regard to consequence, or do not take action unless bribed, and laws are about as worthless as an old cell phone. It is possible to get something done in such places, but the frictions deter plenty, and, more to the point, just as with mediocre infrastructure, it can motivate those with great aspirations to migrate.

As we watch the next wave unfold on a global level, those rules and regulations take center stage. There is an important concern today that government actions will help or deter commercialization of machine learning. This concern takes two flavors. One concern focuses on the effects of repressive censorship and lack of privacy, such as in China, or all-encompassing biometric identification programs, such as in India. Another stresses the effects of regulatory intervention, such as Europe's attempts to protect privacy, or Australia's attempts to foster security by requiring back doors in devices. Most interesting, though it shapes the global standing and competitive position of a locality, the economic interests of most localities have only a minor role in national policy for these issues.

SHANE GREENSTEIN is a Professor at the Harvard Business School, Boston, MA, USA. Contact him at sgreen-stein@hbs.edu.

This article originally appeared in IT Professional vol. 21, no. 2, 2019

Working Abroad in a Research Laboratory in the U.S.

Diego Angel Masini, *Universidad Nacional de La Plata*

Many computer science students dream of the possibility of working abroad, usually at one of the big companies that define the course in technology, innovating and creating the most significant technology breakthroughs. I wasn't the exception. As a student at the Facultad de Informática, Universidad Nacional de La Plata, I always thought about working as a Software Developer in Silicon Valley for a couple of years. In my mind, accomplishing such a goal would provide me with sufficient insights to better perform in a company (or even start a new one) when returning to my home country. I figured I needed some time until I managed to achieve my goal, but I didn't actively pursue it. Looking back, it was a pretty straightforward path; however, at the moment, I did not know what the result would be.

FIRST STEPS IN THE INDUSTRY

In my third year at the university, I managed to get an internship at a small company in Buenos Aires. My responsibilities were to assist the Database Administrator (DBA) since the previous intern had quit for a different job. The job was challenging and very interesting for me; at that moment, the leading company's product was a multidatabase Enterprise Resource Planner, and we were in charge of maintaining four different Database Management Systems (DBMS) trying to survive to the differences between Transact SQL, PL/SQL, and standard SQL.

The DBA was a great guy, eager to share his knowledge with me and open to suggestions. Every question from me, even the ones that were easy to answer, opened the door to an explanation of the inner workings of a DBMS. This training lasted two weeks, after which the DBA left the company, leaving me taking over his role due to the training I had during my first two weeks. It was the first time I verified the saying: "Luck is what happens when preparation meets opportunity."

During my time as a DBA, I was able to apply what I had learned at the University. I was also free to explore and implement new ideas; the only requirement was that these ideas materialized in solutions that simplified my work and the development team work. Due to the volume of work, and because I had no one to help me, I had to learn how to search for the most cost-effective solution, trading-off between the best (but usually impossible to achieve in mortal time) option and the not-quite-best (but feasible in human terms) option.

MOVING TO A TECHNOLOGY TRANSFER PROJECT

The second time, I confirmed the truth behind "Luck is what happens when preparation meets opportunity" was when I decided to start collaborating as an aid in one of the courses I have enjoyed the most at the University: Object Oriented Programming. I had just finished an advanced course in software development. I was eager to share what I had learned and

DOI No. 10.1109/MITP.2019.2898489
Date of current version 27 March 2019.

collaborating as an aid seemed like the natural choice for that. I managed to accommodate my working schedule as a DBA with this new task, and everything went fine for the first couple of months.

After a while, it was clear that having a full-time job and collaborating with the University course was not that particularly compatible with studying to finish my career. I needed to drop something, and since I also needed to work to sustain myself, leaving the collaboration was the logical choice. I went to talk with one of the teachers of the course to explain the situation. He listened carefully, and after I finished, he offered to let me join a technology transfer project coordinated by the University. Being a project lead by the University, it had a couple of perks useful for me: It would allow me to continue collaborating with the Object Oriented Programming course, make progress with my studies and would be in a new domain: casino slot machines. It was a triple win!

I accepted the offer, communicated my departure from the DBA role, trained a replacement, and I started working as part of the Quality Control team, designing and implementing tools to test casino slot machine software and its backend. The work also required some level of manual testing when the interaction with the mechanical parts of the slot machine needed to be tested. I will always remember a rapid feed coin device we assembled out of wood to introduce coins in the slot machine at a velocity no gamer could ever achieve.

During the time I spent doing this job, I learned a lot about the importance of the quality of the software (and the hardware in this particular case) for a company. How much effort from the development team (and from the quality control team) can be saved if errors get caught early in the development process? How do we deal with bugs in production when you have little control over the deployment conditions and no way to distribute online upgrades? Often, upgrading a slot machine involved sending a technician with the new software in a digital medium (e.g., a compact flash) to the casino and do the installation on site. The software industry often underestimates the value of investing in quality control and quality assurance early in the development process.

Eventually, I moved from the Quality Control team to the Backend Development team as a Software Developer. The new role allowed me to experiment with unit testing and Test Driven Development, techniques the Backend team was already exploring and implementing these ideas in an attempt to reduce the number of bugs arriving at the Quality Control team. It was the first time I worked in an agile environment using Scrum.

The team aimed to excel, always collaborating with all the members of the project, assisting newcomers (like me) in the learning process and encouraging experimentation and the proposal of new ideas. It was also the first time I saw the importance of defining an explicit lifecycle model for the software, being able to do reproducible builds, run automated tests, integration tests, keeping an eye on any technical debt we might have, and making it visible for everyone. I incorporated a lot of new ideas and concepts regarding what are the implications of building software.

I am grateful for having the chance to work for several years with such a fantastic team. We were assigned to work on several challenging projects, each one allowing us to research, propose new ideas, and experiment with technology. It was the beginning of a trend for me. Being curious, trying to learn as much as possible, and apply all the knowledge gained to solve problems, combined with the freedom to implement and verify the feasibility of many crazy ideas, created a positive feedback loop, making the whole process very satisfying.

However, the project ended, and it was time to move on.

LANDING AT IBM

After the end of the technology transfer project, I started working at a travel and tourism company. My work implied doing a mixture between architect, backend, and frontend development for one of the products of its online site. Once again, I was lucky enough to apply all my knowledge and crazy ideas (and also introduce some new crazier ones). The team was terrific; I enjoyed the cozy mood at the office and the company's culture. Still, I wanted to do a different type of work: one I could not name at the moment. I wanted to experiment with technology but not to build a product. I wanted to create something that other people use to create technology. I had already developed frameworks, source code generation tools, and resource processors to generate game templates. I wanted to do something different.

I mentioned this to a couple of friends, and then I forgot about the issue. Several months later, one of my friends reminded me of our conversation; he offered to send my resume to IBM: the company he was working for at the time. The position to cover was as a software engineer at a local research team, working with remote groups located in the U.S. at IBM's research laboratories.

In retrospect, it was crucial for me to express what I wanted to do, even if I could not put a name to it and even when I was comfortable with the environment of my job. After the interviews where I had to solve some coding problems and discuss architectural solutions for a hypothetical system, I got hired by IBM Argentina to work as a Software Engineering Researcher, which is a fancy way to say I was going to implement software based on research done by researchers at a research facility in the U.S.

At that moment, I could not realize that the knowledge I had gained during the previous years would be of vital importance for my new role. The team I worked with motivated its members to pursue other challenges besides the ones related to software development, like conducting my own research to write papers or to draft patents on new ideas and defend them. Finally, I had found a place where I was able to explore new topics, apply the acquired knowledge to concrete problems, and, at the same time, research to generate new ideas and new technology.

My first project was in the field of information security, designing the second generation system of a solution for the broadcast encryption problem: how to deliver content over a broadcast channel in a way that only specific users can decrypt and access the content.

It was a challenging year where I had to learn several cryptography concepts and start to think about the security aspects of a system, the types of possible attacks to the software, how to prevent them, how to reduce the attack surface, and so on. These were things that I had never considered before, but somehow complemented what I already knew. After a year working in the project, I got invited to relocate to the U.S. to work on site at IBM's Almaden Research Center (ARC) in San Jose, CA. It was 2015, and I had finally achieved my original goal of working abroad for a big software company, with the plus of going to a research laboratory.

I spent two years working at ARC, and during those years, I had the pleasure to work with many brilliant people on topics that ranged from information security, identity management, cryptography, Blockchain, Internet of Things, embedded systems, and machine learning. I collaborated on several patents and in the design of a couple of research projects. I was also able to coordinate the work of two interns and to present directly to clients some of the work we were doing at the time. I learned about several new technologies and new soft skills. It was a rich experience, from every possible angle. However, the most valuable lessons I learned while being in San Jose were the following:

One, there were many opportunities; it was just a matter of being aware of them. I always wanted to work abroad, but I did not have the chance until I landed at IBM. It was a combination of the decisions I took professionally and seizing opportunities. It was also the result of many significant efforts done by several people across my life, starting with my parents, my teachers, my friends, and my coworkers. People were always open to share their knowledge and to give me enough space to try and err.

Two, sharing and asking is always the best strategy. If you want to do something and you do not know if it is possible, just ask. However, do not ask for permission; instead, ask how. If you know something other people do not, share it, and you might be surprised how the insights from other people can enhance what you already know.

THE ROAD AHEAD

Upon my return to Argentina in 2017, I still needed to finish my degree thesis. It made sense for me to take advantage of what I have learned while in the U.S. Given that most of the work I did at ARC was on Blockchain, I decided to combine two aspects I find interesting: Identity management and Blockchain, or stating it differently, how to manage identity in a distributed system without relying on a single central authority.

Identity management on Blockchain is quite a challenge, and, in part, that is the reason I enjoy working on it. Blockchain is a buzzword that is resonating a lot these days. Part of it is just noise and marketing chit-chat, but there is real value and real challenges in the technology. I still do not know if Blockchain technologies will succeed as we know them today, but many of the advances done in this field today can have a significant impact in the technology we are going to see in the near future.

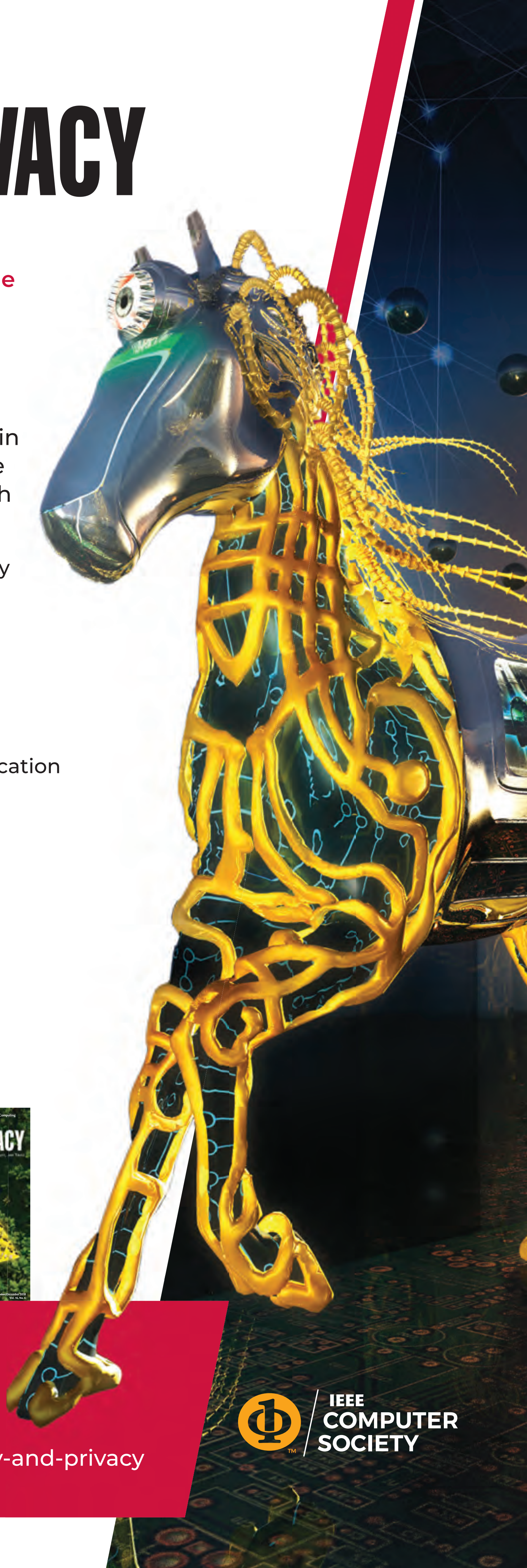
Early this year, I moved from IBM to a medium-size startup where I can continue learning and working as a Researcher. I am focused on applying all the learning from my experience in the U.S. to improve existing solutions and design new ones in the Blockchain space. My goal is to contribute to the technology, making it more natural to adopt. I do not know where I will be in the next five years, but I like to think I will still be doing what I love most: learning.

REFERENCES

1. D. Naor, M. Naor, and J. Lotspiech, "Revocation and tracing schemes for stateless receivers," in *Proc. 21st Annu. Int. Cryptol. Conf. Adv. Cryptol.*, 2001, pp. 41–62, Available: http://www.wisdom.weizmann.ac.il/~naor/PAPERS/2nl.html
2. Encyclopedia of Theoretical and Applied Cryptography, "Broadcast encryption," Available: http://cryptowiki.net/index.php?title = Broadcast_encryption, Sep. 2009.
3. Advanced Access Content System, "Introduction and Common Cryptographic Elements Book," Available: https://www.aacsla.com/specifications/AACS_Spec_Common_FINAL_0951.pdf
4. S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Available: https://bitcoin.org/bitcoin.pdf, Oct. 2008.
5. D. Puthal, N. Malik, S. P. Mohanty, E. Kougianos, and G. Das, "Everything you wanted to know about the blockchain," *IEEE Consum. Electron. Mag.*, vol. 7, no. 4, pp. 6–14, Jul. 2018. Available: http://www.smohanty.org/Publications_Journals/2018/Mohanty_IEEE-CEM_2018-Jul_Blockchain.pdf

DIEGO ANGEL MASINI has been working for more than 11 years in computer science as a Software Engineer and as a Systems Administrator. In 2014, he started working for IBM Argentina in Information Security on the design and implementation of the Advanced Access Content System 2 for Ultra HD Blu-ray. In 2015, he relocated to the U.S. to be part of the Information Security Research team with IBM's Almaden Research Center, where he was part of the IBM's Blockchain initiative that lead to the open source project Hyperledger Fabric. During that time, he mostly focused on private Blockchain networks and Identity Management. In 2018, he returned to Argentina and joined the RSK Labs team, where he is part of the research team as a Senior Researcher, designing new features and scalability solutions for public Blockchain networks. He received the Graduate degree in computer science from Universidad Nacional de La Plata, La Plata, Argentina. Contact him at: diego.masini@gmail.com.


E-Currency and Fairness • Ransomware Defense • A National Cybersecurity Policy
SECURITY & PRIVACY
SECURITY & PRIVACY
SECURITY & PRIVACY
SECURITY & PRIVACY
Privacy and Automated Airport Screening



EDITOR: Schahram Dustdar, dustdar@dsg.tuwien.ac.at

This article originally appeared in IEEE Internet Computing vol. 23, no. 4, 2019

DEPARTMENT: INTERNET OF THINGS, PEOPLE, AND PROCESSES

Architectural Considerations for Privacy on the Edge

Christos Tsigkanos, Cosmin Avasalcai, and Schahram Dustdar, *TU Vienna*

Novel pervasive systems integrate technologies and paradigms, such as mobile and cloud computing and Internet of Things, where systems are composed of heterogeneous infrastructures and services. Privacy emerges as a first-class design goal throughout such systems' development lifecycle, and suggests its management to occur architecturally at the network edge, closer to end-users as the privacy stakeholders. We discuss concerns emerging from privacy requirements and how they pertain to contemporary pervasive systems, and we distill architectural considerations highlighting privacy protection mechanisms and tactics for edge computing.

"And about whatever I may see or hear in treatment, in the life of human beings—things that should not ever be blurted out outside—I will remain silent, holding such things to be unutterable;" Article 8 of the Hippocratic Oath provides a strong metaphor for engineering privacy-aware—by design and by default—systems.[1] Hippocrates talk about treatment of possibly sensitive medical information by a healthcare provider. Current more than ever, Hippocrates provides us the foundation of privacy-aware data management: a system may collect, use for some intended purpose, but not misuse or disclose private information. Such an ancient principle is particularly relevant in the increasingly integrated and pervasive computing environments of today, where mobile, cloud, and Internet of Things (IoT) converge inducing systems that collect, process, and disseminate information, which may be sensitive. Traditionally, organizations have been viewed as trusted custodians of information; however, data breaches, misuse, or malicious use of the sensitive information can harm privacy of the individuals.

10.1109/MIC.2019.2935800
Date of current version 7 October 2019.

Especially relevant in today's integrated world, comprehensive privacy mechanisms are essential for the widespread uptake and acceptance of the systems we engineer, as the ever-increasing number of devices collecting (possibly sensitive) data and interacting with the physical environment poses privacy risks. Regularly acknowledged in contemporary legal and regulatory frameworks, privacy emerges as a first-class design goal throughout an application's development lifecycle. Pervasive systems are particularly sensitive to this; resource-constrained IoT devices, cloud offloading, and generally heterogeneous software components operate within diverse administrative domains.[2,3] At the same time, the volume of the data generated by devices close to end-users grows exponentially. Resource-constrained IoT devices limit the techniques that can be used to deliver efficient and effective privacy-preserving schemes.

We argue that the drive toward edge computing can help pervasive systems engineering tackle privacy threats. We elaborate the role of edge computing, by outlining concerns emerging from privacy requirements and how they pertain to contemporary pervasive systems and discuss software architecture considerations to meet privacy needs. The edge computing paradigm can help preserve privacy and protect users: first, by establishing privacy controls at a layer close to data-producing end-users and subjects, second, by minimizing the need to transmit sensitive data to the cloud for analysis, and third, by offering opportunities for stronger privacy with respect to the data collection and identification through an edge-centric anonymization. Since user-facing software components are the ones that generate and act on sensitive data, empowering the edge appears to be a reasonable decision. However, considerable engineering challenges arise to support this. We make the case that the decentralization inherent in the edge computing paradigm yields significant benefits for privacy.

In the following, we start from a societal perception of privacy—how privacy has come to be treated in legal terms[4]—and we adapt such concepts for contemporary pervasive and ubiquitous systems. Subsequently, we discuss architectural considerations that highlight privacy protection mechanisms and tactics for edge computing.

PRIVACY CONCERNS IN PERVASIVE SYSTEMS

Privacy concerns that we identify are based on a loose interpretation of the legal privacy taxonomical framework designed by Solove,[4] and adapted for contemporary ubiquitous systems; we treat them as building blocks of privacy requirements. Motivated by architectural considerations, we, however, propose a different grouping and ignore ones that are software-architecture-irrelevant (e.g., interrogation as an activity that may violate privacy). We follow the data lineage within IoT systems to discuss privacy concepts; data are collected, processed, or analyzed in edge or cloud computers and possibly disseminated, closing the loop with a possible control flow back to user-facing devices. In the following, for each privacy concern we identify, we discuss challenges, opportunities, and emerging solutions within the edge-based systems.

Data Collection and Identification

Devices that collect, store, and send information from various surrounding sources are ubiquitous in modern environments and an integral part of the IoT paradigm. Information collection—even if no information is revealed publicly—can be problematic with respect to privacy laws and regulations. Requirements may require not only data within an application to remain locally close to where it is sourced, but also for all mechanisms managing it to respect different legal or administrative frameworks (e.g., the EU General Data Protection Regulation versus the California Consumer Privacy Act) and user preferences. A further challenge is that the data often traverses through computational resources of diverse domains; enabling applications to operate across them points to another facet where data producers within an edge-based system require control over data exchange. Furthermore, collected data may be used to identify individuals. To this end, anonymization methods have been developed to counteract the possibility of linking attacks and achieving *k*-anonymity over data.[5] While such processes are well understood for the static datasets, they become challenging on dynamic and streaming data typically found in contemporary ubiquitous and IoT applications.

We advocate that because the edge is closer to data sources and users, there is not only obvious latency advantages, but also an opportunity for stronger privacy with respect to data collection and identification. This includes anonymization as well as operation within administrative domains; both suggest building appropriate data handling logic inside the edge-based software components. To counter identification threats while collecting data, anonymization can occur at the edge before transmission to the cloud. Anonymization facilities[6] need to be developed for streaming data and perhaps able to be deployed on resource-constrained edge devices, e.g., within a user's home. Finally, the risk of (re-)identification of anonymized data with machine learning (ML) methods spotting patterns needs to be acknowledged; possible mitigations can include use of diversification techniques or feeding of fake transactions.

Aggregation and Inference

Information processing in the context of privacy, involves various ways of combining data together and linking it to individual people to whom it relates. Vast amounts of personal data about individuals stored at different commercial vendors and organizations is the norm, which in combination can pose privacy risks. Furthermore, we identify information aggregation and inference as particularly relevant privacy aspects due to the widespread adoption of data-producing IoT devices and the increasing technological advances in (and need for) inference using artificial intelligence (AI) and ML methods. We treat aggregation and inference activities as similar from a technological perspective, as they similarly concern multidimensional data. For example, patterns within smart meter readings processed by an energy operator for analytics or energy efficiency purposes can reveal the occupancy of a residence.

Complex processing and inference is typically performed on the cloud and is dominated by training deep ML models requiring heavy compute capacity. The recent trend is on moving the inference part of the AI workflow close to end-devices. This may be desirable for nonfunctional requirements like security, cost or latency, but can have a positive impact on privacy as well, as user data are kept constrained to an edge device. Novel approaches, such as federated learning[7]—where user-facing devices learn a shared prediction model in a collaborative manner while keeping all the training data on the device—can hinder aggregation and inference attacks that presuppose centralized, organizationally-curated data repositories usually on the cloud.

Secondary Use, Insecurity and Exclusion

Following the data lineage within IoT systems, after data have been sourced from mobile devices, sensors, or users within some organization, privacy issues arise with use, storage, and manipulation of collected data. This privacy aspect concerns issues arising from an organization's maintenance and use of collected data. Insecurity—from a legal perspective—involves carelessness in protecting stored information from leaks and improper access. Secondary use concerns information collected for one purpose, used for a different purpose without the data subject's consent. This reflects a common principle in information privacy, where for all data collected and processed, there should be a stated purpose; usage of data for another purpose than the one it was intended for must be prevented. Exclusion concerns the failure to allow the data subject to know about the data that others have about her and participate in its handling and use. Such aspects have been reflected also in recent regulatory frameworks, such as EU/GDPR (Art. 6 – Lawfulness of processing), making compliance mandatory. The typical example lies within the healthcare domain, where there is an organic abundance of sensitive data collection for diagnostic or other medical purposes. First, data may be improperly stored leading to data leaks. Secondly, they may be shared with third parties for some other originally unintended use (such as insurance companies or research institutions). Finally, fulfilling requests by a subject (e.g., a concerned patient) for what data have been collected and with whom it was shared may be difficult due to improper data handling processes.

Information privacy research has long developed privacy models and frameworks to ensure compliance. P-RBAC[8] is able to capture roles and permissions, actions on data, conditions, and obligations that arise in privacy requirements, whereas Contextual Integrity's model of communicating agents[9] shows suitability for streaming data. However, integration in engineering processes and architectures within privacy-compliant ubiquitous systems have to be investigated. In edge computing architectures, with user-facing devices handling (possibly sensitive) data and interacting with the physical environment, the edge device is by definition located within the administrative domain of its local IoT devices—one can take that as the devices being in the same privacy scope. Thus, the edge can be treated as a first-class entity regarding privacy and data management, and can ensure that the data flows between the edge and other external components (i.e., other edge nodes, the cloud etc) always respect defined privacy policies in the system.[10]

Decision and Boundaries

Privacy does not always involve information. Harms may come from invasive acts that disturb an individual's personal boundaries and tranquility. A manifestation of this are the legal protections of the privacy of the home, protecting it from trespass and external nuisances (for hundreds of years).[4] Invasive acts may harm privacy, and—to borrow from cyber-physical systems terminology—are a form *actuation*. The interplay between computational and physical aspects must be additionally considered.[11] For example, a virtual assistant making actuation decisions, such as enabling indoor surveillance cameras when occupants are home, can violate privacy boundaries of subjects.

The drive to decentralization points to putting trust and security in the hands of users. Edge computers operated by and within reach of users (e.g., virtual assistants in their home), and potentially invasive (e.g., in charge of controlling window blinds or cameras) should operate transparently and in an accountable manner. Putting security in the hands of users can be a double-edged sword, but empowering users to make their own decisions about control actions, devices and networks they own is desirable from a privacy perspective.

Appropriation and Distortion

Traditionally, organizations have been viewed as trusted custodians of information; however, data breaches or malicious use of sensitive information can harm privacy of individuals. Generally, appropriation involves the use of the data subject's identity to serve the aims and interests of another, such as the deliberate use of someone else's personal data in context of identity theft. Distortion consists of the inappropriate dissemination of false or misleading information about individuals, for example misuse of personal data acquired by a corporation in a marketing campaign.[4]

Failures of the various data custodians, the increased prevalence of sensitive data nowadays, and the mistrust placed on institutional organizations to manage them calls for new paradigms. Security concepts, such as identity management and digital signatures can be leveraged in a decentralized manner for sensitive data security and management at the edge of networks. Personal data for instance, can be signed before reaching data stores, rendering source verification, and lineage tracking possible. Solutions made possible with blockchain technology can be further used, with validation that occurs in edge nodes outside of control of organizations, but within the control—and trust sphere—of users.

EDGE ARCHITECTURES FOR PRIVACY-PROTECTION

In edge-based systems, the edge software component is by definition (by virtue of deployment) located close or within the administrative domain of the local end-user. Software components operated by the end-user and hosted on, e.g., mobile or IoT devices interact with the edge. As such, one can take that the edge and user components are architecturally in the same privacy scope, and trust exists between them. Sensitive data flows between the edge and external components (i.e. other edge nodes, or the cloud) should always respect privacy policies, over which the user should have complete control. As such, edge components must enforce privacy policies when data are shared, processed, and collected.

To distill the previous into an architecture, we loosely follow the conventional description structure of ISO/IEC/IEEE 42010—a software architecture view, architecture description and privacy as a cross-cutting perspective. We take the concerns of the previous section to be the functional viewpoint of a privacy-protecting edge architecture, showing scenarios describing how various architectural elements and views should address privacy.

Architecture Description

A privacy-protecting platform must accommodate the decentralized nature of today's infrastructures, where sensitive data are distributed between multiple devices in the network, and apply privacy protection close to where data are collected. Such a platform may be composed of three entities: first, the subject or user that has total control of his/her private data, second, the privacy-supporting software components representing facilities used to enforce privacy measures, and third, the organizational infrastructure, typically centrally located in the cloud (see Figure 1).

The cloud layer consists of multiple data servers, located at remote (to the data sources) locations, where organizations collect data for further processing. The cloud may be in a different administrative domain compared to the origin of data—privacy controls should be enforced before data are sent to the cloud. We advocate that the role of this architectural layer is overall *privacy compliance* with requirements, laws and regulations—the burdens of specification and implementation move to lower architecture layers. The edge layer is populated with geo-located distributed edge devices that may or may not belong on the same administrative domain themselves. Such computing devices have different characteristics—from resource-constrained energy-optimizing devices to powerful gateways and edge servers where

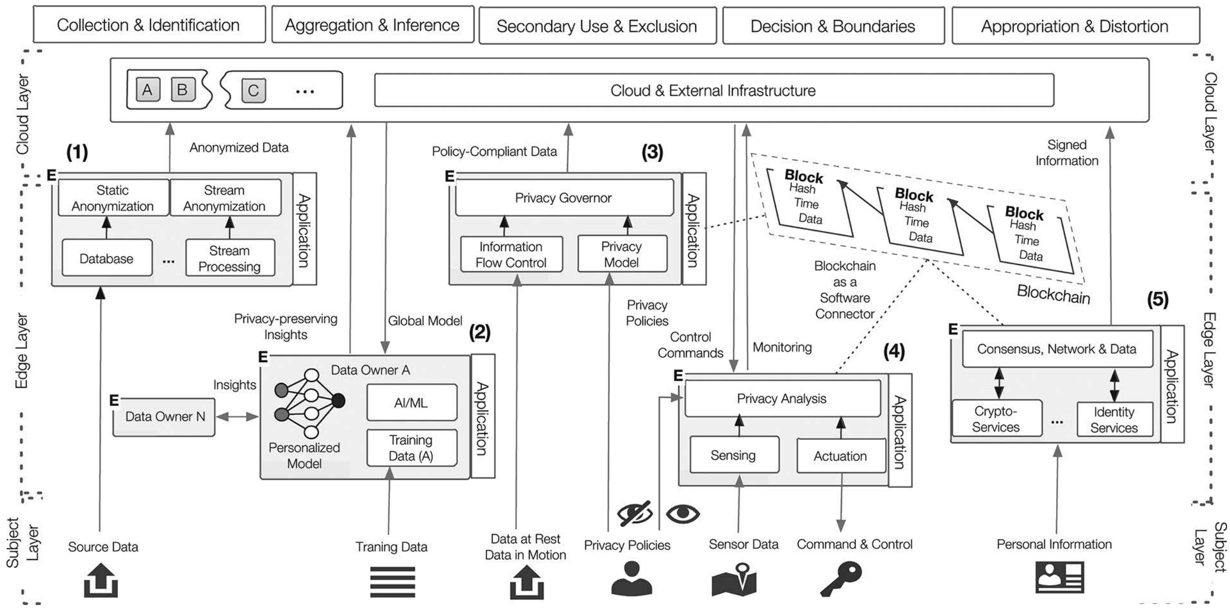


FIGURE 1. Combined dataflow and deployment architectural diagram showing different architectural manifestations for the edge as a first-class entity for privacy protection.

computation can be offloaded. The role of the edge layer is *privacy implementation*—measures and techniques are deployed in edge nodes as privacy-protecting software components, and act as intermediary monitoring and enforcement facilities. Finally, the subject layer concerns data sources and *privacy specification*. Data are generated by different user-facing devices as primary sources. Specification of privacy requirements and control of the edge facilities employed to satisfy them lies similarly within control of the user.

Software Architecture View

The view of edge entity internals consists of two layers: first, the application layer, fulfilling user goals and second, the edge support layer, facilitating privacy governance. The former implements business logic and is application specific. The latter has a supportive role to applications; this includes provisioning, configuration management, data storage, and event processing, but also privacy governance facilities: anonymization, policy enforcement, and control. Such capabilities may be exposed to applications.

In Figure 1, edge components (middle layer) act as intermediaries between privacy subjects and primary data sources (lower layer), and the organizational, external infrastructure (top layer). Privacy concepts are shown within a horizontal division, with the architecture outline of the respective edge component. Within each, dataflow and typical deployment of software components is illustrated, highlighting various capabilities. Such supportive privacy components can be accessed and used by end-user applications.

Privacy Perspective

We treat the privacy perspective as a cross-cutting concern, i.e., capabilities that cut across architectural layers (from privacy subjects to organizations and cloud infrastructures), and outline how privacy protection mechanisms can be employed at the edge.

Data collection and identification threats concern sensitive data occurring at the subject layer, which are used to uniquely pinpoint a subject. Anonymization techniques[5] can be used to counter privacy threats from multidimensional data, performed at the edge before data reaches an organizational cloud infrastructure [see diagram (1) of Figure 1]. *k*-Anonymity and *k*-indistinguishability have been historically employed for data-at-rest—contemporary ubiquitous systems, however, are often characterized by streaming data. We advocate development and deployment (and oversight) of such techniques at (or by) the edge node.[6]

Aggregation and inference threats are about using diverse data and advanced methods to identify subjects by combining information, typically with AI techniques. Recent developments on federated learning have shown that ML applications can be engineered in a decentralized manner, with training data not leaving the privacy sphere of the subject.[7] The evocative term *edge AI* has emerged, where deployment of AI/ML techniques occurs at edge nodes. In such a case [see diagram (2) of Figure 1], personalized models reside and are trained by subject data at nodes, whereas insights are shared with other data owners. The key idea is that insights shared can preserve privacy, as personal data do not leave the scope of the node.

Secondary use, insecurity, and exclusion threats can be tackled by using privacy models to control how, what and for which purpose data are collected, shared, and processed. Given specification of privacy policies capturing subject preferences (or privacy laws and regulations), a privacy governor can operate on incoming subject data, ensuring compliance [see diagram (3) of Figure 1]. Furthermore, if the subject has control of the edge node, privacy policies can by-design dictate how data leaves the scope.

Decision and boundaries threats can emerge by control actions that come from outside a subject's privacy scope. Control actions are usually coupled with sensing (e.g., a virtual assistant opens the window blinds when daylight is detected). Privacy analysis wrapping such actions can occur at the edge [e.g., at the subject's home network, see diagram (4) of Figure 1], forbidding or allowing command, and control based on subject's explicit preferences.

Appropriation and distortion refers to deceptive use of a subject's personal information. Cryptographic advances have enabled widespread schemes to counter such trust issues. Digital signatures can verify authenticity or source of messages or documents, and symmetrically, can also provide nonrepudiation. Such measures can be employed to validate that the sensitive information has not been altered, and that it belongs to the subject that generated it. Moreover, recent developments have shown the benefits of blockchains for trust management without central authorities or servers. A blockchain deployment in an edge-to-edge network where nodes collectively adhere to a protocol for communication and validation, can support such functions in a decentralized manner and ensure that personal information is not appropriated by others.[12] Functionality essentially entails recording transactions between subject parties or organizational entities in a verifiable and permanent manner. We suggest that the edge can host such supportive functions [see diagram (5) of Figure 1] and can make them available to end-user applications, by considering a blockchain as a software connector for identity, trust and verification facilities.

EMERGING RESEARCH AGENDA

Contemporary pervasive systems integrate multiple technologies and paradigms in systems that are composed of heterogeneous infrastructures and services. Privacy emerges as a first-class design goal throughout such systems' application development lifecycle, and suggests its management to occur architecturally at the network edge, closer to the end-devices. We discussed aspects emerging from privacy requirements and how software architecture considerations pertain to edge-enabled systems, and highlighted privacy protection mechanisms and tactics for edge computing.

As future work, we identify providing a complete reference architecture that engineers and organizations can use for documenting viewpoints as per ISO/IEC/IEEE 42010. Qualitative aspects and other nonfunctional requirements, such as performance often are in conflict with the privacy protection, as it adds processing overhead. Such design tradeoffs need to be carefully considered, as, e.g., timeliness of data or events can be critical in edge-based systems. We ignored network and communication issues, which must be treated as well in operationalizations of the privacy-protecting architectures discussed. Finally, identification and employment of specific technological options for the refinement of the architectural components discussed is highly desired.

ACKNOWLEDGMENTS

This work was supported in part by the Research Cluster "Smart Communities and Technologies (Smart CT)" at TU Vienna, and in part by the EU H2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 764785, FORA–Fog Computing for Robotics and Industrial Automation.

REFERENCES

1. K. Barker, et al., "A data privacy taxonomy," in *Proc. Brit. Nat. Conf. Databases*, 2009, pp. 42–54.
2. J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle, "Privacy in the internet of things: Threats and challenges," *Secur. Commun. Netw.*, vol. 7, no. 12, pp. 2728–2742, 2014.
3. C. Tsigkanos, S. Nastic, and S. Dustdar, "Towards resilient internet of things: Vision, challenges, and research roadmap," in *Proc. 39th IEEE Int. Conf. Distrib. Comput. Syst.*, Jul. 2019, pp. 1754–1764.
4. D. J. Solove, "A taxonomy of privacy," *Univ. Pennsylvania Law Rev.*, vol. 154, no. 3, 2006, Art. no. 477.
5. C. C. Aggarwal, "On *k*-anonymity and the curse of dimensionality," in *Proc. 31st Int. Conf. Very Large Databases*, 2005, pp. 901–909.
6. J. Cao, B. Carminati, E. Ferrari, and K.-L. Tan, "CASTLE: Continuously anonymizing data streams," *IEEE Trans. Dependable Secure Comput.*, vol. 8, no. 3, pp. 337–352, May/Jun. 2011.
7. K. Bonawitz, et al., "Towards federated learning at scale: System design," 2019, *arXiv:1902.01046*.
8. Q. Ni, et al., "Privacy-aware role-based access control," *ACM Trans. Inf. Syst. Secur.*, vol. 13, no. 3, 2010, Art. no. 24.
9. A. Barth, A. Datta, J. C. Mitchell, and H. Nissenbaum, "Privacy and contextual integrity: Framework and applications," in *Proc. IEEE Symp. Secur. Privacy*, 2006, Art. no. 15.
10. N. Li, C. Tsigkanos, Z. Jin, S. Dustdar, Z. Hu, and C. Ghezzi, "POET: Privacy on the edge with bidirectional data transformations," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, Mar. 2019, pp. 1–10.
11. C. Tsigkanos, L. Pasquale, C. Ghezzi, and B. Nuseibeh, "On the interplay between cyber and physical spaces for adaptive security," *IEEE Trans. Dependable Sec. Comput.*, vol. 15, no. 3, pp. 466–480, May/Jun. 2018.
12. M. A. Khan and K. Salah, "Iot security: Review, blockchain solutions, and open challenges," *Future Gener. Comput. Syst.*, vol. 82, pp. 395–411, 2018.

CHRISTOS TSIGKANOS is a Researcher with the Distibuted Systems Group at TU Vienna. Previously, he was a Postdoctoral Researcher at Politecnico di Milano, where he received the Ph.D. defending a thesis entitled "Modeling and Verification of Evolving Cyber-Physical Spaces." His advisor was Prof. Carlo Ghezzi. His research interests lie in the intersection of dependable systems and formal aspects of software engineering, and include security and privacy in distributed, self-adaptive and cyber-physical systems, requirements engineering and formal verification. He received the B.Sc. degree in computer science from the University of Athens, Greece, and the M.Sc. degree in software engineering from the University of Amsterdam. Contact him at: christos.tsigkanos@tuwien.ac.at.

COSMIN AVASALCAI is a Research Scientist and a Ph.D. student working under the supervision of Prof. Schahram Dustdar with the Distibuted Systems Group, TU Vienna. Before joining the Distributed Systems Group, he worked for two and a half years as a research assistant with the Technical University of Denmark (DTU), developing modeling and decision making tools for the automotive industry for mixed-criticality applications in dynamic and changeable real-time environments and the oil and gas industry for structural monitoring. His current activities include quality of service-aware resource provisioning in edge computing within the Fog Computing for Robotics and Industrial Automation European Training Network project. After finishing the bachelor studies from Technical University "Gheorghe Asachi" Iasi, Romania, in 2015, he received the master's degree in computer science and engineering from DTU in 2017. Contact him at: c.avasalcai@dsg.tuwien.ac.at.

SCHAHRAM DUSTDAR is a Professor of Computer Science with the Distributed Systems Group, TU Vienna, Austria, and an IEEE Fellow. He was an Honorary Professor of Information Systems with the University of Groningen, The Netherlands from 2004–2010, a Visiting Professor with the University of Sevilla, Spain from 2016–2017, and a Visiting Professor with the University of California at Berkeley, USA in 2017. He is an elected member of the Academia Europaea, where he is the Chairman of the Informatics Section. He was the recipient of the ACM Distinguished Scientist Award in 2009, the IBM Faculty Award in 2012, and the IEEE TCSVC Outstanding Leadership Award for outstanding leadership in services computing in 2018. He is the Co-Editor-in-Chief for the ACM *Transactions on Internet of Things* and the Editor-in-Chief for *Computing* (Springer). He is also an Associate Editor for the IEEE Transactions on Services Computing, the IEEE Transactions on Cloud Computing, the ACM *Transactions on the Web*, and the ACM *Transactions on Internet Technology*. He is on the Editorial Board of *IEEE Internet Computing* and the *Computer Magazine*. Contact him at: dustdar@dsg.tuwien.ac.at.

EDITOR: Ron Vetter, University of North Carolina Wilmington, vetterr@uncw.edu

DEPARTMENT: SPOTLIGHT ON TRANSACTIONS

This article originally appeared in
Computer
vol. 53, no. 4, 2020

Computer Architecture for Orbital Edge Computing

Daniel J. Sorin, *Duke University*

This installment of Computer*'s series highlighting the work published in IEEE Computer Society journals comes from* IEEE Computer Architecture Letters.

Major innovations in computer architecture are often driven by changes in how and where computers are deployed. There are dramatic differences between the architectures of stationary desktops, mobile laptops and smartphones, and data centers, and these differences reflect different software workloads and constraints.

In "Orbital Edge Computing,"[1] authors Bradley Denby and Brandon Lucia introduce architects to a new deployment scenario: processors for clusters of nanosatellites that have been launched into low-Earth orbit.

Many nanosatellites are used for Earth imaging and observation, and they are equipped with sensors (cameras) and a processor. The primary challenge to overcome is that communication between a nanosatellite and Earth is slow and unreliable, and increasingly large constellations of nanosatellites are overwhelming the capabilities of centralized terrestrial processing. Results show that simply communicating all of the raw data from the nanosatellites is infeasible.

The proposed solution is to adopt an edge processing paradigm, in which the onboard computer processes the large amount of camera data and sends the smaller amount of postprocessed data, but the onboard computer is tightly constrained by the limited amount of power that can be harnessed with solar panels. The article shows how being deployed to orbit constrains camera design, which, in turn, constrains the satellite's volume, surface area, and ability to compute onboard. To further reduce the required downlink bandwidth, the authors propose three techniques. First, they show that, with their tiled image processing (Figure 1), it is critical to choose a good size for the tiles when trading

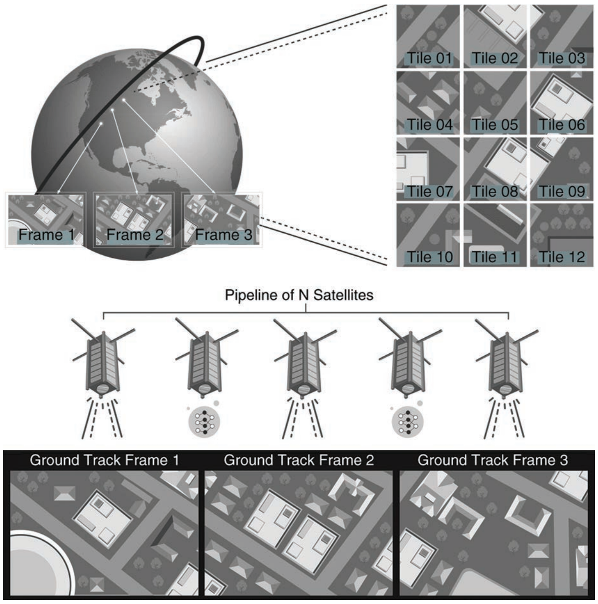


FIGURE 1. In a computational nanosatellite pipeline, nanosatellites collect frames along the orbital ground track. Together as an orbital edge computing system, the nanosatellites decompose the frames into tiles and distribute the work of processing tiles across the entire constellation without the need to communicate to Earth. (Modified from [1].)

DOI No. 10.1109/MC.2020.2969673
Date of current version: 9 April 2020

bandwidth for image resolution. Second, they show how their edge computing approach can enable them to discard entire tiles because, for example, the tiles are redundant or obscured by cloud cover. Third, they use the classic architectural approach of pipelining to divvy up the tiles among the nanosatellites. All of these approaches, taken together, vastly reduce the demand on precious downlink bandwidth.

Although nanosatellites are not new, their consideration by computer architects is, and it is exciting to think about exploring this vast design space from an architect's perspective. Similar to how mobile computing and data centers drove new processor designs, orbital edge computing could also lead to dramatically new architectures.

REFERENCE

1. B. Denby and B. Lucia, "Orbital edge computing: Machine inference in space," *IEEE Comput. Archit. Lett.*, vol. 18, no. 1, pp. 59–62, 1 Jan.–June 2019. doi: 10.1109/LCA.2019.2907539.

DANIEL J. SORIN is a professor of electrical and computer engineering at Duke University. He currently serves as the editor-in-chief of *IEEE Computer Architecture Letters*. Contact him at sorin@ee.duke.edu.


stay connected.

A global pioneer in technology innovation, OPPO provides cutting-edge devices and services to empower users in an intelligently connected future.

The OPPO brand was first registered in 2001 and OPPO was officially incorporated in 2004. To date, OPPO has expanded its business presence to more than 40 countries and regions, with over 400,000 sales outlets worldwide. With nine Intelligent Manufacturing Centers, six Research Divisions, four R&D Centers and a London-based Global Design Center, OPPO has more than 40,000 employees around the world. OPPO was awarded as "Most Respected Chinese Company" by in 2018 for the third consecutive year and was also included on Boston Consulting Group' s" 2018 BCG Global Challengers" list. Also in 2018, OPPO received the Canstar Blue "Most Satisfied Customers" award for the second year in a row. In 2019, OPPO earned a spot on the "Top 50 Chinese Global Brands" ranking by BrandZTM, the leading global authority on brand power.

Position	Job Discription	Location
Vice President and Chief Scientist (OPPO Research Center)	Research strategy and talent growth to enable the company to hold a leading position. A natural collaborator, working with other senior leadership team members within the company, and with academia (including joint labs), industry partners, and government outside the company.	Shenzhen, China
General Manager of HW R&D Center	Responsible for the overall management and operation of OPPO HW R&D Center; Guide the HW technical planning for the next 3-5 years.	Shenzhen/Dongguan, China
General Manager of Imaging Technology Center	Responsible for imaging technology and industry insight analysis, lead the research direction and technical planning of core image technologies.	Shenzhen, China
Multimedia Algorithm Expert (Imaging, Video, AR/CV, Display,and AI)	Look for self-motivated, experienced, and creative algorithm expert to research and explore the cutting-edge technologies and algorithms in the multimedia areas (at least one of the following areas: imaging, video, audio, computer vision, and AI)	Palo Alto/San Diego Shenzhen/Beijing/Shanghai, China
Architect of Machine Learning Infrastructure	Responsible for designing and building the next-generation of unified ML platform for vision, language, speech, multi-modal context understanding, and large-scale recommendation models	Shenzhen/Beijing, China
Architect of Recommendation & Search Platform	Responsible for the design and development of real-time, scalable, AI-powered search engines and recommendation systems, particularly for streaming feeds, photos and short vidoes	Shenzhen/Beijing, China
Pr. Research Engineer – Image/ Video Multi-modal Learning	Leverag multidisciplinary approach to develop highly scalable video/image vertical search engine and tools	Palo Alto, CA Shenzhen,China
Pr. Research Engineer – Image/ Video Enhancement	Develop cutting-edge deep learning algorithms for image（especially super-resolution）/video enhancement	Palo Alto, CA Shenzhen, China
Pr. Research Engineer – Scene Text detection and recognition	Develop, evaluate, improve, optimize, and deploy our advanced OCR models.	Palo Alto, CA
Principal Mobile Software Architect	Provide technical leadership and expertise of the whole Android software platform. Lead and perform the design, development, and maintenance of applications and frameworks for Android systems.	Palo Alto, CA Shenzhen/Chengdu/Nanjing /Shanghai, China
Pr. SLAM|3D Sensing Engineer	Research and develop advanced real-time robotic SLAM solutions and practical solutions for 3D scene understanding	Palo Alto, CA Shenzhen, China
Pr. Engineer – 3D Vision and AR system	Analyze and optimize computer vision and deep learning algorithms on mobile/embedded devices	Palo Alto, CA Beijing/Shanghai/Shenzhen, China
Pr. Engineer/Manager – Camera Algorithm	Develop camera algorithms and use cases with ISP, computer vision and AI for smartphone platforms.	Palo Alto, CA Shenzhen, China
Pr. Speech or NLP research scientist	Advanced asr、tts、nmt、dialog、kg、nlp algorithm research	Beijing/Shenzhen, China
Pr. Research Scientist – Privacy Protection Computing	Research and develop federated learning, differential privacy and SMC techniques for privacy protection computation platform	Shenzhen, China
Pr. Research Engineer – Machine Learning Algorithm	Research and develop advanced machine learning solutions for business and decision intelligence	Shenzhen, China
Principal Research Scientist – Neural Architecture Search	Search the best neural network architecture balancing accuracy and latency, including building search platform, designing search space and search strategy, and perforamce estimation strategy	Beijing/Shanghai/Shenzhen, China
Principal Research Scientist – Computer Vision	Research and develop advanced computer vision and state-of-the-art deep learning techniques，such as Image Retrieval，Image Recognition，Object Detection，and optimize algorithms in combination with real business scenarios	Beijing/Shanghai/Shenzhen, China
Principal Research Scientist – Human-centered Computing	Design and develop systems which automatically analyze faces and human activies, and create media approaches (audio, video, etc.) to faclitate communications between people and machine, people and people	Beijing/Shanghai/Shenzhen, China
Principal Research Scientist – Visual Generation	Design and develop systems which generate synthetic visual information using GAN or rendering from graphical model, or a combination of both	Beijing/Shanghai/Shenzhen, China

Link: https://www.linkedin.com/company/oppo/jobs/

IEEE Internet Computing

Special Issue on
Toward Fully Autonomous and Networked Vehicles

Deadlines:

Submissions due: 10 September 2020

Publication issue: March/April 2021

From automobiles and aircraft to ships and scooters, vehicles of all shapes and sizes are increasingly autonomous. As autonomous vehicles mature and become widely adopted, networked systems composed of multiple autonomous vehicles could represent the next technological evolution. Fleets of autonomous and networked vehicles could yield novel features, execution efficiencies, and synergistic capabilities otherwise unachievable with standalone designs. Imagine a traveler leaving her home in Seattle to vacation in Tokyo and the fleet of heterogeneous vehicles planning, coordinating, and executing her travel: A car picks her up at home and delivers her to an airport, a scooter gets her from the airport curbside to the plane, an unmanned aircraft flies her to Tokyo, and an autonomous bus takes her to a hotel. The fleet of vehicles can also adjust plans at runtime to improve efficiency (for example, by minimizing carbon and energy footprints) or to provide reliability guarantees in the presence of failures. To realize this type of system, we need research contributions toward autonomous and networked vehicles. This special issue calls for research on various issues and solutions that can enable autonomous and networked vehicles.

Topics of interest include:

- Deployment experiences with multiple autonomous vehicles
- Testbeds and benchmarks to evaluate and compare research frameworks
- Assessing energy, availability, data consistency, reliability, and cost of autonomous vehicles
- Secure-by-design approaches to connected autonomous vehicles
- Storing and sharing data sensed by autonomous vehicles
- Computer architecture issues for autonomous vehicles
- Resource management for autonomous vehicles

Submission guidelines:

www.computer.org/ic/author-information

Contact the guest editor(s) at

ic2-2021@computer.org

Guest editor(s):

- Christopher Stewart, The Ohio State University
- Deepak Vasisht, University of Illinois Urbana-Champaign
- Weisong Shi, Wayne State University

Conference Calendar

IEEE Computer Society conferences are valuable forums for learning on broad and dynamically shifting topics from within the computing profession. With over 200 conferences featuring leading experts and thought leaders, we have an event that is right for you. Questions? Contact conferences@computer.org.

JULY

6 July

- ICME (IEEE Int'l Conf. on Multimedia and Expo), London, UK

13 July

- COMPSAC (IEEE Computer Software and Applications Conf.), Madrid, Spain

28 July

- CBMS (IEEE Int'l Symposium on Computer-Based Medical Systems), Rochester, USA

AUGUST

1 August

- JCDL (ACM/IEEE Joint Conf. on Digital Libraries), Xi'an, China

14 August

- SmartIoT (IEEE Int'l Conf. on Smart Internet of Things), Beijing, China

16 August

- Hot Chips, Palo Alto, USA

17 August

- ACSOS (IEEE Int'l Conf. on Autonomic Computing and Self-Organizing Systems), Washington, DC, USA

19 August

- RTCSA (IEEE Int'l Conf. on Embedded and Real-Time Computing Systems and Applications), Gangnueng, South Korea

24 August

- FiCloud (Int'l Conf. on Future Internet of Things and Cloud), Rome, Italy

31 August

- RE (IEEE Int'l Requirements Eng. Conf.), Zurich, Switzerland

SEPTEMBER

7 September

- EuroS&P (IEEE European Symposium on Security & Privacy), Genova, Italy

14 September

- CLUSTER (IEEE Int'l Conf. on Cluster Computing), Kobe, Japan

21 September

- ASE (IEEE/ACM Int'l Conf. on Automated Software Eng.), Melbourne, Australia

24 September

- BigMM (IEEE Int'l Conf. on Multimedia Big Data), New Delhi, India

27 September

- ICSME (IEEE Int'l Conf. on Software Maintenance and Evolution), Adelaide, Australia

28 September

- SecDev (IEEE Secure Development), Atlanta, USA

OCTOBER

3 October

- PACT (Int'l Conf. on Parallel Architectures and Compilation Techniques), Atlanta, USA

5 October

- EDOC (IEEE Int'l Enterprise Distributed Object Computing Conf.), Eindhoven, the Netherlands
- ICSE (IEEE/ACM Int'l Conf. on Software Eng.), Seoul, South Korea

12 October

- ISSRE (IEEE Int'l Symposium on Software Reliability Eng.), Coimbra, Portugal

16 October

- ICEBE (IEEE Int'l Conf. on e-Business Eng.), Guangzhou, China

17 October

- MICRO (IEEE/ACM Int'l Symposium on Microarchitecture), Athens, Greece

18 October

- ICCD (IEEE Int'l Conf. on Computer Design), Hartford, USA
- MODELS (ACM/IEEE Int'l Conf. on Model-Driven Eng. Languages and Systems), Montreal, Canada

19 October

- DFT (IEEE Int'l Symposium on

Defect and Fault Tolerance in VLSI and Nanotechnology Systems), Frascati, Italy

21 October
- FIE (IEEE Frontiers in Education Conf.), Uppsala, Sweden

25 October
- VIS (IEEE Visualization Conf.), Salt Lake City, USA

NOVEMBER

2 November
- ICFEC (IEEE Int'l Conf. on Fog and Edge Computing), Melbourne, Australia

6 November
- CCEM (IEEE Int'l Conf. on Cloud Computing in Emerging Markets), Bengaluru, India
- SmartCloud (IEEE Int'l Conf. on Smart Cloud), Washington, DC, USA

8 November
- NAS (IEEE Int'l Conf. on Networking, Architecture, and Storage), Riverside, USA

9 November
- ICTAI (IEEE Int'l Conf. on Tools with Artificial Intelligence), Baltimore, USA
- IRC (IEEE Int'l Conf. on Robotic Computing), Taichung, Taiwan
- ISMVL (IEEE Int'l Symposium on Multiple-Valued Logic), Miyazaki, Japan

11 November:
- SEC (IEEE/ACM Symposium on Edge Computing), San Jose, USA

15 November
- SC, Atlanta, USA

16 November
- FOCS (IEEE Symposium on Foundations of Computer Science), Durham, USA
- LCN (IEEE Conf. on Local Computer Networks), Sydney, Australia

30 November
- ICHI (IEEE Int'l Conf. on Healthcare Informatics), Oldenburg, Germany

DECEMBER

2 December
- ICDCS (IEEE Int'l Conf. on Distributed Computing Systems), Singapore
- ISM (IEEE Int'l Symposium on Multimedia), Naples, Italy

6 December
- HOST (IEEE Int'l Symposium on Hardware-Oriented Security and Trust), San Jose, USA

7 December
- UCC (IEEE/ACM Int'l Conf. on Utility and Cloud Computing), Leicester, UK

9 December
- CC (IEEE Int'l Conf. on Conversational Computing), Irvine, USA
- AIKE (IEEE Int'l Conf. on Artificial Intelligence and Knowledge Eng.), Irvine, USA

14 December
- CloudCom (IEEE Int'l Conf. on Cloud Computing Technology and Science), Bangkok, Thailand

16 December
- BIBM (IEEE Int'l Conf. on Bioinformatics and Biomedicine), Seoul, South Korea
- HiPC (IEEE Int'l Conf. on High-Performance Computing, Data, and Analytics), Pune, India