COMPUTING edge

- Machine Learning
- Software Development
- Autonomous Vehicles
- Education

JANUARY 2023

www.computer.org

IEEE COMPUTER SOCIETY

IEEE

IEEE Computer Society Has You Covered!

WORLD-CLASS CONFERENCES — Stay ahead of the curve by attending one of our 210 globally recognized conferences.

DIGITAL LIBRARY — Easily access over 800k articles covering world-class peer-reviewed content in the IEEE Computer Society Digital Library.

CALLS FOR PAPERS — Discover opportunities to write and present your ground-breaking accomplishments.

EDUCATION — Strengthen your resume with the IEEE Computer Society Course Catalog and its range of offerings.

ADVANCE YOUR CAREER — Search the new positions posted in the IEEE Computer Society Jobs Board.

NETWORK — Make connections that count by participating in local Region, Section, and Chapter activities.

Explore all of the member benefits at www.computer.org today!

IEEE COMPUTER SOCIETY

IEEE

COMPUTING edge

IEEE COMPUTER SOCIETY computer.org

STAFF

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

Computer
Jeff Voas, *NIST*

Computing in Science & Engineering
Lorena A. Barba, *George Washington University*

IEEE Annals of the History of Computing
Ramesh Subramanian, *Quinnipiac University and Yale Law School*

IEEE Computer Graphics and Applications
André Stork, *Fraunhofer IGD and TU Darmstadt*

IEEE Intelligent Systems
San Murugesan, *Western Sydney University (Interim EIC)*

IEEE Internet Computing
George Pallis, *University of Cyprus*

IEEE Micro
Lizy Kurian John, *University of Texas at Austin*

IEEE MultiMedia
Balakrishnan Prabhakaran, *University of Texas at Dallas*

IEEE Pervasive Computing
Fahim Kawsar, *Nokia Bell Labs and University of Glasgow*

IEEE Security & Privacy
Sean Peisert, *Lawrence Berkeley National Laboratory and University of California, Davis*

IEEE Software
Ipek Ozkaya, *Software Engineering Institute*

IT Professional
Charalampos Z. Patrikakis, *University of West Attica*

JANUARY 2023 • VOLUME 9 • NUMBER 1

COMPUTING edge

Machine Learning

Software Development

Autonomous Vehicles

Education

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

CUF-Links: Continuous and Ubiquitous FAIRness Linkages for Reproducible Research

The ability to share data is critical to reproducible research, yet data sharing is often limited because issues of findability, accessibility, interoperability, and reusability—the FAIR principles—are not integrated into every step of the scientific process. Read more in this article from the August 2022 issue of *Computer*.

Computing in Science & Engineering

Information-Theoretic Exploration of Multivariate Time-Varying Image Databases

Modern scientific simulations produce very large datasets, making interactive exploration of such data computationally prohibitive. An increasingly common data reduction technique is to store visualizations and other data extracts in a database. The Cinema project is one such approach, storing visualizations in an image database for post hoc exploration and interactive image-based analysis. This article from the May/June 2022 issue of *Computing in Science & Engineering* focuses on developing efficient algorithms that can quantify various types of multivariate dependencies existing within multivariable datasets. It applies specific mutual information measures for the quantification of salient regions from multivariate image data.

IEEE Annals of the History of Computing

What/Whom is the Brazilian University for? The Case of the Computing Projects Laboratory

This article from the July–September 2022 issue of *IEEE Annals of the History of Computing* discusses the social function of Brazilian universities from the standpoint of the history of computing in Brazil, revisiting the dissolution of the Computing Projects Laboratory (LPC), a lab of the Pontifical Catholic University of Rio de Janeiro (PUC-RJ). LPC was responsible in the mid-1970s for the pioneering development of the G-10 minicomputer basic software, a machine designed and developed in Brazil in the context of the so-called "Market Reserve" policy, which aimed at developing local computer production. The authors highlight the narratives that position the source of the LPC crisis in the tensions between a supposed division of "pure science" versus "applied science."

IEEE Computer Graphics and Applications

News Globe: Visualization of Geolocalized News Articles

When exploring articles on online news portals, navigation is mostly limited to the most recent ones. The spatial context and the history of topics are not immediately accessible. To support readers in the exploration or research of articles in large datasets, the authors of this article from the July/August 2022 issue of *IEEE Computer Graphics and Applications* developed an interactive 3D globe visualization. They worked with datasets from multiple online news portals containing up to 45,000 articles. Using agglomerative hierarchical clustering, they represent the referenced locations of news articles on a globe with different levels of detail.

IEEE Intelligent Systems

Maximizing Fairness in Deep Neural Networks via Mode Connectivity

With frequent reports of biased outcomes of AI systems, fairness rightfully becomes an active area of current ML research. However, while progress has been made on theoretical analysis and formulation of fairness as constraints on error probabilities, our ability to design and train modern deep learning models that reach the targeted fairness goals in practice is still limited. In this *IEEE Intelligent Systems* May/June 2022 article, the authors focus on an interesting yet common fairness setting, where multiple samples are collected from each individual, and the goal is to maximally reduce performance disparity among individuals while maintaining overall model performance.

IEEE Internet Computing

A Vision for Leveraging the Concept of Digital Twins to Support the Provision of Personalized Cancer Care

Exploring the opportunity for applying digital twins in the healthcare context is an emerging research area that has the potential to support more personalized care. A recognized aspect in cancer care is the need for more personalized treatment planning to complement the recent advances in precision medicine. In this article from *IEEE Internet Computing*'s September/October 2022 issue, the authors present a classification of digital twins into Grey Box, Surrogate, and Black Box models using systems and mathematical modeling theory. They then explore one possible approach: a Black Box classification for incorporating the use of digital twins in the context of personalized uterine cancer care.

IEEE micro

Performance Left on the Table: An Evaluation of Compiler Autovectorization for RISC-V

Next-generation length-agnostic vector instruction set architecture (ISA) designs, the RISC-V vector extension, and ARM's scalable vector extension enable software portability across hardware implementations with different vector engines. While traditional, fixed-length single-instruction–multiple-data ISA instructions, such as Intel AVX and ARM Neon, enjoy mature compiler support for automatic vectorization, compiler support is still emerging for these length-agnostic ISAs. This September/October 2022 *IEEE Micro* article studies the compiler shortcomings that constitute the gap in autovectorization capabilities between length-agnostic and fixed-length architectures. The authors examine LLVM's support for both the RISC-V vector extension and traditional vector ISAs.

IEEE MultiMedia

Privacy-Preserving Image Classification Using an Isotropic Network

In this article from *IEEE MultiMedia*'s April–June 2022 issue, the authors propose a privacy-preserving image classification method that uses encrypted images and an isotropic network, such as the vision transformer. The proposed method allows us not only to apply images without visual information to deep neural networks for both training and testing, but also to maintain a high classification accuracy. In addition, compressible encrypted images, called encryption-then-compression (EtC) images, can be used for both training and testing without any adaptation network. Previously, to classify EtC images, an adaptation network was required before a classification network, so methods with an adaptation network have only been tested on small images.

IEEE Pervasive Computing

Breaking Age Barriers With Automatic Voice-Based Depression Detection

Most existing voice-based depression datasets comprise speakers younger than 60, and variations in speech due to age and depression are not well understood. In this article from *IEEE Pervasive Computing*'s April–June 2022 issue, which uses Patient Health Questionnaires for depression severity ground-truth, automatic depression detection is explored using acoustic-based prosodic, spectral, landmark, and voice quality features derived from smartphone recordings from 152 speakers in four age ranges. An age-dependent modeling paradigm for voice-based depression detection is proposed and evaluated. Results show that age-dependent models improve voice-based automatic depression classification accuracy with up to 10% absolute gains when compared with an age-agnostic model.

IEEE Security & Privacy

Data Privacy and Trustworthy Machine Learning

The privacy risks of machine learning models are a major concern when training them on sensitive and personal data. The authors of this *IEEE Security & Privacy* article from the September/October 2022 issue discuss the trade-offs between data privacy and the remaining goals of trustworthy machine learning (fairness, robustness, and explainability).

IEEE Software

The Advantages of Maintaining a Multitask, Project-Specific Bot: An Experience Report

Bots are becoming a popular method for automating everyday tasks in many software projects, thanks to the availability of many off-the-shelf task-specific bots that teams can quickly adopt. The authors of this article from the September/October 2022 issue of *IEEE Software* argue that an alternative approach deserving more attention is to develop a multitask project-specific bot, because it strikes a good balance between productivity and adaptability.

IT Professional

SAME: The Design Space for Seamless Automotive Multimodal Experience

Multimodal user interfaces for automotive applications, such as voice, gesture, and head-up displays, have significantly enriched the user experience. From a brief literature review, the authors of this May/June 2022 *IT Professional* article discuss the concept of seamlessness in autonomous vehicles and propose a new design space of multimodal user interfaces within the upgrade of driving automated systems and human multitasking across devices.

Editor's Note

Can We Trust Machine Learning?

We entrust machine learning (ML) with decisions that have significant impacts in our society. But are we certain that ML systems are behaving responsibly? Can we trust ML techniques to be secure and to deliver explainable and fair results? This January 2023 issue of *ComputingEdge* explores various aspects of ML trustworthiness in real-world applications.

"Trustworthy Machine Learning," from *IEEE Intelligent Systems*, describes a four-layer architecture to support scalable trustworthy ML. In *IEEE Internet Computing*'s "Sustainable and Trustworthy Edge Machine Learning," the author proposes approaches for overcoming the hyper-heterogeneity and high failure rate of Internet of Things sensors to achieve trustworthy edge ML.

Quality matters in all types of software, and developers need strategies for designing high-quality programs. The authors of "Nudging Software Developers Toward Secure Code," from *IEEE Security & Privacy*, identify bad habits like reusing code from the Internet and suggest ways to change behavioral patterns. *IEEE Software*'s "Decision-Making Principles for Better Software Design Decisions" presents a systematic approach to decision-making that emphasizes facts, contexts, risks, constraints, and priorities.

Next, two articles from *Computer* discuss current issues in autonomous vehicles. In "The Importance of Interoperability in Functional Safety Standards," the authors argue that recent standardization activities for automated vehicles are creating information-exchange challenges. They summarize two new standards that focus on interoperability. "Pushing the Limits of Autonomy for Enabling the Next Generation of Space Robotics Exploration Missions" predicts a paradigm shift from fragile, remote-operated missions to fully autonomous and resilient robotic systems.

Finally, this *ComputingEdge* issue features two articles on education. *IEEE Computer Graphics and Applications*' "A Recipe of Capabilities for Pursuing Expertise in Data Visualization: A Practitioner's Perspective" details the skills needed for competence in data visualization. "Building a Culture of Computing in the Sciences Using Images as Data Within a Community of Practice," from *Computing in Science & Engineering*, showcases a program to improve the computational skills of college students majoring in biology or chemistry.

EDITOR: Longbing Cao, University of Technology, Sydney, Australia

COLUMN: AI FOCUS

This article originally appeared in
Intelligent Systems
vol. 37, no. 1, 2022

Trustworthy Machine Learning

Bhavani Thuraisingham, *The University of Texas at Dallas, Richardson, TX, 75080, USA*

Machine learning (ML) techniques have numerous applications in many fields, including healthcare, medicine, finance, marketing, and cyber security. For example, ML techniques are being applied to determine whether to give a loan to a customer or whether the computing system has been attacked. However, the ML techniques themselves may be subject to attacks and may discriminate when determining who should get the loan. Therefore, the ML techniques have to be secure, ensure privacy of the individuals, incorporate fairness and be accurate. Such collection of ML techniques has come to be known as trustworthy machine learning (trustworthy ML). This article describes an architecture to support scalable trustworthy ML and describes the features that have to be incorporated into the ML techniques to ensure that they are trustworthy.

The collection, storage, manipulation, analysis, and retention of massive amounts of data have resulted in new technologies including the development of novel machine learning (ML) techniques. It is now possible to analyze massive amounts of data and extract useful nuggets with ML. Therefore, ML techniques are being applied to analyze the massive amounts of data in every field such as healthcare, finance, retail, manufacturing, marketing, and security. However, the collection and manipulation of this data has also resulted in serious security and privacy considerations. Various regulations are being proposed to handle big data so that the privacy of the individuals is not violated via ML. Furthermore, the massive amounts of data being stored and the associated ML techniques to analyze the data may also be vulnerable to cyber-attacks.

ML techniques are being integrated to solve many of the security challenges. For example, ML techniques are being applied to solve security problems such as malware analysis and insider threat detection. However, there is also a major concern that the ML techniques themselves could be attacked. Therefore, the ML techniques are being adapted to handle adversarial attacks. This area is known as adversarial ML. In addition, privacy of the individuals may also be violated through these ML techniques as it is now possible to gather and analyze vast amounts of data and extract sensitive information about the individuals. Therefore, privacy-enhanced ML techniques are being developed to ensure the privacy of individuals and at the same time extract useful results from the data.

The integration of ML and cyber security has many dimensions. In addition to applying ML for cyber security problems, as well as detecting and preventing the attacks to the ML techniques and ensuring the privacy of individuals, there are also other aspects: 1) the ML techniques are being applied for detecting fake news as well as handling cyber bullying; 2) the ML techniques themselves have to be fair and not biased; 3) the ML techniques have to be fault tolerant; 4) the ML techniques have to ensure the integrity of the data. The ML techniques that provide the above features have come to be known as trustworthy ML. Ensuring that the ML techniques possess all of the features such as fairness, privacy, security, integrity, and fault tolerance is a huge challenge. What may be desirable is for them to be adaptive depending on the applications. That is, for some applications, the ML techniques may have to ensure privacy (e.g., healthcare) while for some other applications they have to be fair (e.g., approving bank loans). As we make more progress with the research and development of the ML tools, one can expect to develop ML techniques that are secure, private, accurate, fair, and fault tolerant.

Trustworthy ML has several applications including in the Internet of Things and Social Media. First, with the advent of the web, computing systems are now being used in every aspect of our lives from mobile phones to smart homes to autonomous vehicles. It is now possible to collect, store, manage, and analyze vast amounts of sensor data emanating from numerous devices and sensors including from various transportation systems. Such

1541-1672 © 2022 IEEE
Digital Object Identifier 10.1109/MIS.2022.3152946
Date of current version 11 April 2022.

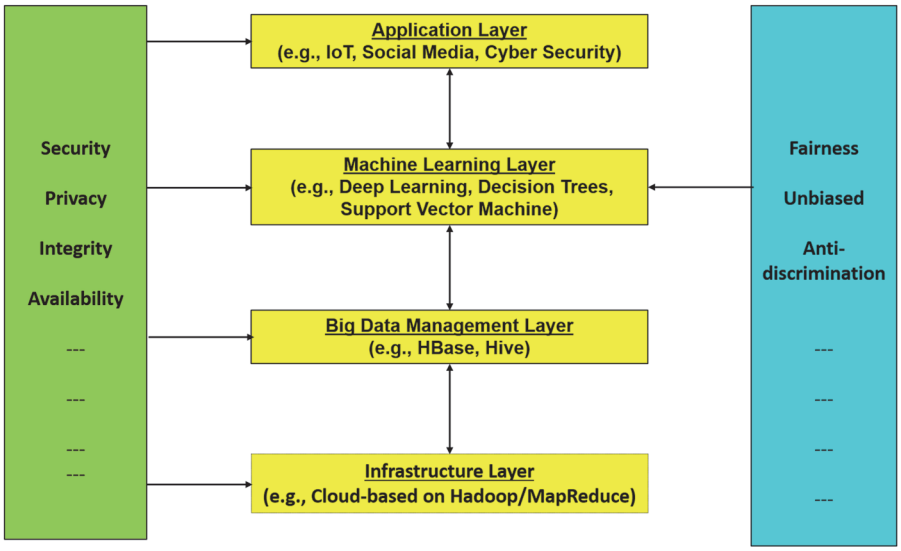

FIGURE 1. Architecture to support trustworthy ML.

systems collectively are known as the Internet of Transportation systems, which are essentially the Internet of Things for Transportation, where multiple autonomous transportation systems are connected through the web and coordinate their activities. However, security and privacy for the Internet of Transportation and the infrastructures that support it is a challenge. Due to the large volumes of heterogeneous data being collected from numerous devices, the traditional cyber security techniques such as encryption are not efficient to secure the Internet of Transportation. Some physics-based solutions being developed are showing promise. More recently, the developments in ML are also being examined for securing the Internet of Things and Transportation systems and their supporting infrastructures with the ultimate goal of developing smart cities. Second, trustworthy ML techniques also have applications in social media analytics including fake news detection and handling cyber bullying. In addition, the social media systems are being integrated with transportation systems resulting in a new area called vehicular social media systems. ML techniques are being applied to vehicular social media systems to provide the best driving experience.

While there are articles on integrating ML and Cyber Security,[1] work on trustworthy ML as we have defined it has only just begun. That is, it is only recently that work has begun on examining all aspects of ML techniques that are integrated with security, privacy, fairness, bias, fake new detection, cyber bullying, integrity, and fault tolerance. This article will provide a fairly comprehensive overview of trustworthy ML. The organization of this article is as follows. The "Architecture to Support Trustworthy Machine Learning" section describes an architecture to support Trustworthy ML. Trustworthy ML concepts such as security, privacy, fairness, and integrity are discussed in the "Concepts in Trustworthy ML" section. The "Directions" section describes some directions for trustworthy ML.

ARCHITECTURE TO SUPPORT TRUSTWORTHY MACHINE LEARNING

Figure 1 illustrates a four-layer architecture to support trustworthy ML. The bottom layer is the infrastructure layer based on high-performance computing technologies. One such technology is the cloud. The cloud provides virtualization as well as the storage (e.g., the Hadoop/MapReduce framework) The other layers are essentially hosted on the cloud. The data layer provides the data management services for applications that include ML. Since the amount of data needed by the applications such as social media and the Internet of Things may be massive, we assume that this data layer is essentially the big data layer. This layer may utilize systems such as Hive, HBase, and Cassandra to manage the data. This data layer interfaces to the infrastructure layer and is processed by saying the Hadoop/MapReduce framework.

The ML layer is the layer that carries out both learning and analytics. This layer may include several ML technologies that utilize support vector machines, decision trees, and neural networks. This layer interacts with the Big Data layer to extract the data needed for the learning process. More sophisticated learning algorithms such as deep learning may also be utilized at this layer. The highest layer is the applications layer that requires the results from ML to carry out various tasks. For example, with respect to social media applications,

ML algorithms may provide the analytics capabilities such as determining the leader of a group or the preferences of a customer. One important application of ML for social media is fake news detection. Here, the algorithms are trained with fake news and then tested with different pieces of news articles to determine whether the news is fake or not.[2] The other applications include Internet of Things (e.g., Internet of Transportation) that give advice to the drivers regarding the best tours to take. Cyber security is also an application for ML. For example, the cyber security data could be analyzed by the ML algorithms to determine when and where the attacks occurred and to predict future attacks.[1]

What makes this architecture trustworthy are the features that augment each layer. For example, features such as security, private, and integrity cut across all the layers. That is, security at the infrastructure layer would provide security for the Hadoop/MapReduce framework. Security at the big data layer would provide secure features such as access control for big data systems such as HBase. Security at the ML layer would ensure that the attacks to the ML algorithms are detected. Finally, security at the applications layer would provide fine-grained access control for social media systems. Also, additional features may be unique to a particular layer. That is, at the ML layer, additional features such as fairness could be included. The various concepts in trustworthy ML such as security, privacy, integrity, and fairness will be discussed in the "Concepts in Trustworthy ML" section.

CONCEPTS IN TRUSTWORTHY ML

For ML techniques to be trustworthy they must be secure, ensure privacy, provide fairness, and have high integrity among others. We will examine each of these features.

Security: First, the ML techniques must only access the data they are authorized to do so to carry out the tasks (e.g., learning). Numerous types of access control techniques have been proposed in the literature for various types of data management systems. While ML has been explored to develop access control policies, controlling access to the data by the ML systems has not received much attention.[3] However, the areas of adversarial ML have been studied extensively. The ML techniques are prone to cyber-attacks. The adversary tries to learn as much as possible about the data we use to train the ML models as well as the ML models we have developed. Therefore, we have to learn the behavior of the adversary and develop appropriate solutions such as adversarial support vector machines.[4] Another proposed approach to detect cyber-attacks is to test the ML techniques thoroughly and in many situations be able to apply formal verification techniques. Appropriate formal specifications, verification, and testing are being investigated to ensure that the ML software does not contain malicious code. Some novel approaches being examined for specifying, verifying, and testing ML techniques are discussed in the Verified AI project, led by Sanit Seshia at University of California, Berkeley;[5] the Verified AI project also describes formal specifications of deep neural networks[5] and they include identifying properties of interest for ML models and systems and specifying these properties in appropriate formal languages.

Privacy: Privacy-preserving (also referred to as privacy-aware) ML (PPML) has been studied for the past 20 years.[6] ML algorithms learn from the data, build the models and make predictions. These predictions could be highly sensitive or private. Furthermore, the data used for learning may be sensitive. Therefore, the goal of PPML techniques is to ensure privacy while at the same time make accurate predictions. There are tradeoffs between privacy and accuracy. Therefore, the amount of privacy that is desired often depends on the user/applications. For some applications, one may need, say, 70% privacy and 30% accuracy while for some others one may need 30% privacy and 70% accuracy. So the question is, how do you make an ML technique such as decision trees a PPML technique? The idea is to incorporate privacy metrics at various stages of the learning process. For example, when the ML algorithm is learning from the data, the data may be perturbed (or randomized) depending on the privacy metric so that sensitive values are not divulged. Similarly, the privacy metrics may even applied to the model development processes. Finally, the privacy metric may be incorporated into the prediction process so that predictions are made in a privacy-sensitive manner.[7] It is however important to define the variables that have to be protected such as personally identifiable information.

Fairness: Fairness/bias in ML is becoming a critical consideration. ML techniques are being used to determine who to give bank loans to, who to admit to college, and how to ration healthcare. The outcome of the ML techniques could enable a physician to make life or death decision and also affect the advancement of human beings. Therefore, the ML techniques must be fair and unbiased. It is important that the ML techniques do not discriminate against various individuals based on, say, gender or race. An excellent discussion of fairness in ML is provided by Caton and Haas.[8] The authors state that fairness is with respect to some variable such as gender, race, religion, and age. The challenge is that often there is no legal definition of sensitive variables. Therefore, the application needs to determine what the fairness variables are. Fairness can be addressed at different stages. One is at the data level, where the data may be biased.

Therefore a fairness metric is incorporated into the data processing stage. The second stage is at the model development level. Here again, one needs to understand how the ML model is developed and incorporate a fairness metric. The third stage is postprocessing where a fairness metric is incorporated when predictions are made after the model is developed. In some ways, fairness is treated in a manner similar to privacy. While fairness is a critical consideration, it is also an extremely challenging problem to solve. There are many research efforts now on this topic and we believe that tremendous progress will be made in the near future.

Integrity: The final feature we discuss is integrity. Integrity means many things from the accuracy of the data to the integrity of the process to the correctness to the algorithms. Accuracy of the data, and in this case the training data, may be tainted through malicious corruption due to faults. This means we need to determine the provenance of the data.[9] Where did the data come from? Who modified the data? Integrity is also critical during the model development process. If the model is incorrect then even the most accurate data would not matter. The accuracy of the model depends on the features extracted. This is one of the most critical tasks in ML. Which features and how many features do we extract? Too few features or too many features may not give an accurate model. Therefore, experience in ML model development is necessary to extract the correct number of features. Another challenge is to ensure that the ML software meets the specification. As in the discussion for security, formal methods may be used to ensure that the ML software is correct. Another aspect to consider is what happens if the ML algorithms fail due to faulty software? The literature in fault tolerant and dependable computing needs to be examined to handle fault tolerance for ML techniques. Finally, is it possible to develop real-time ML algorithms? More recent work on ensemble-based ML models for classifying big data streams provides some insights into the problem.[1]

DIRECTIONS

The challenge we are faced with is how do you integrate all of these features to provide ML techniques that are truly trustworthy? It is hard enough to develop ML techniques that are privacy-preserving or fair. The question is, how do you handle both? Even worse what happens if the ML techniques are attacked? Finally, how can we ensure that the training data is accurate. One solution is to incorporate quality of service into the trustworthy ML algorithms. That is, for certain applications privacy may be more important while for other applications (e.g., healthcare), we may need to focus on fairness (e.g., finance). There are some complex problems that researchers and developers have to focus on if we are to develop ML techniques that incorporate many of the features we have discussed in this section.

Another challenge is to develop the architecture we have illustrated in Figure 1. That is, how do we integrate the cloud with big data management to support trustworthy ML? Furthermore, how can the trustworthy ML algorithms solve problems such as analyzing IoT data and detecting fake news in social media? Lot of research has been carried out on using the cloud to carry out analytics including for cyber security problems such as malware analysis and insider threat detection. It is critical that the cloud, the big data management, and the ML techniques used to carry out analytics are secure. We believe that we need secure infrastructure, data managers, and ML systems to solve some of the challenging problems in various applications we are faced with today. An excellent resource on trustworthy ML can be found in https://www.trustworthyml.org/.[10]

REFERENCES

1. B. Thuraisingham, P. Pallabi, M. Masud, and L. Khan, *Big Data Analytics With Applications in Insider Threat Detection*. Boca Raton, FL, USA: CRC Press, 2017.
2. K. Shu, A. Silva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *SIGKDD Explorations*, vol. 19, pp. 22–36, 2017.
3. NIST, *Machine Learning for Access Control Policy Verification*, NISTIR, vol. 8360, 2021. [Online]. Available: https://csrc.nist.gov/News/2021/nistir-8360-published
4. Y. Zhou, M. Kantarcioglu, B. M. Thuraisingham, and B. Xi, "Adversarial support vector machine learning," *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2012, pp. 1059–1067.
5. S. Seshia, "Verified AI Project." University of California, Berkeley. [Online]. Available: https://berkeleylearnverify.github.io/VerifiedAIWebsite/
6. R. Agrawal and R. Srikant, "Privacy-Preserving data mining," in *Proc. ACM SIGMOD*, 2000, pp. 439–450.
7. L. Liu, M. Kantarcioglu, and B. M. Thuraisingham, "Privacy preserving decision tree mining from perturbed data," in *Proc. 42nd Hawaii Int. Conf. System Sci.*, 2009, pp. 1–10.
8. S. Caton and C. Haas, "Fairness in ML: A survey," 2020, *CoRR abs/2010.04053*.
9. T. Cadenhead, M. Kantarcioglu, and B. Thuraisingham, "A framework for policies over provenance," *TaPP*, Crete, Greece, 2011.
10. Trustworthy ML Organization, Trustworthy ML Initiative. Accessed: Mar. 08, 2022. [Online]. Available: https://www.trustworthyml.org/

EDITOR: Ewa Deelman, deelman@isi.edu

This article originally appeared in IEEE Internet Computing vol. 25, no. 5, 2021

DEPARTMENT: VIEW FROM THE CLOUD

Sustainable and Trustworthy Edge Machine Learning

Ivona Brandic, *Vienna University of Technology, 1040 Vienna, Austria*

Nowadays, our world is driven by complex, large scale, yet tactile information systems requiring various degrees of trustworthiness. Trustworthiness of the systems always comes with costs. The traditional and rather costly way to understand the behavior of large scale systems is to develop powerful mathematical abstractions that allow us to condense these behaviors and to reason about them at a very abstract level. In our FWF funded project Rucon, we introduce an orthogonal, data driven, and probabilistic concept to model and reason uncertainty of the systems. In Rucon, deliberated system failures are tolerated due to the benefits of the costs and sustainability. Rucon's approach targets large scale near real-time systems like live video analytics, streaming, vehicular applications, and smart city information systems.

The ongoing digital transformation is disruptively changing all aspects of our lives. The sectors immediately affected and revolutionized by the IoT are healthcare (smart medical devices), manufacturing (smart factories), energy (smart power grids) as well as urban development and transportation (smart buildings, cities and vehicles). One of the direct impacts of the increased digitalization is the nearly exponential increase of energy demand to process and store data. Already, data centers represent an estimated 1% of global electricity demand. One of the most worrying models predicts that electricity use by ICT could exceed 20% of the global total by the time a child born today reaches her teens.[9]

To meet the demands of the ongoing digitization efforts, a new generation of information systems is emerging with latencies less than 100 ms or even less than 10 ms what is called nowadays "tactile internet," addressing upcoming data driven business applications like virtual reality, telemedicine, smart cities, or self-driving cars. The trend in transforming traditional backend applications to "tactile internet" applications is also affecting the High Performance Computing (HPC) area. In the area of HPC, we have the concept of "Extreme Data" that follows the "Big Data" problem, by addressing massive amounts of information that must be processed and analyzed in near-real time through the utilization of Exascale systems. In the past decade, Cloud Data Centers and Supercomputers have been envisioned as the essential computing architectures for enabling the next generation of extreme data applications. However, in recent years we experienced the rise of near-real time extreme data systems. These applications process complex data intensive workflows with strict latency requirements. An example of an extreme data application is the early earthquake alert based on the analysis of thousands of sensors from smart phones.[7] In both cases (commercial and HPC computing), we evidence the paradigm shift not only to "tactile internet" but also in the type of emerging applications, where traditional simulation and optimization applications are replaced or enhanced with data intensive machine learning (ML) applications.

When considering such huge, complex, and geographically distributed systems, it is not intuitively clear where the most common performance bottlenecks are, or which parts of the systems are the most inefficient in terms of the energy consumption. In case of the tactile internet, both the hardware and the application can contribute equally to the inefficiency and high failure rates.

1089-7801 © 2021 IEEE
Digital Object Identifier 10.1109/MIC.2021.3104383
Date of current version 29 September 2021.

DATA INTENSIVE APPLICATIONS

Data intensive applications in the context of tactile internet differ fundamentally from traditional ML applications trained and executed in centralized and highly controlled environments like Cloud Data Centers.[14] Many geographically distributed ML applications, e.g., Apple Siri are entirely based on cloud computing. Such applications do not function if the network is unavailable. Also many of the existing intelligent applications generally adopt centralized data management, where users upload their data to a central cloud based data center. However, with the ever increasing volumes of data, which has been generated and collected by billions of mobile users and IoT devices it is estimated that Internet traffic is reaching 235.7 Exabytes per month in 2021, up from 73.1 Exabytes per month in 2016.[2]

HYPERHETEROGENEOUS HARDWARE

Hardware architectures and networks utilized for IoT systems and tactile internet differ radically from all other well known systems. Since IoT devices (e.g., sensors) are rather tiny and not capable of running complex computation, more powerful nodes in the vicinity of IoT devices are necessary to process and store data and ensure low latency; this is called "the Edge." The concept, when the local version of the ML model is deployed at the Edge, is called *Edge Machine Learning (EML)*.[1,10,11] An important application area for EML is (near) real-time object detection, as it is currently often impossible to run the video inference without GPU on board.[5] Another application area is to enhance 5G with computational facilities.[6] The resource landscape is becoming more and more heterogeneous with Edge nodes that can significantly vary in their shape, size, and computational power, resulting in the so-called hyperheterogeneity. Edge nodes might range from the so-called μ-Data Centers[8] consisting of several servers to simple Raspberry Pis. Network connections might vary as well ranging from wifi to LTE or 4G.

CHALLENGES IN TERMS OF SUSTAINABILITY AND TRUSTWORTHINESS

Hyper heterogeneity and geographical distribution of Edge systems make it difficult to manage the competing priorities like sustainability and trustworthiness. Even worse, a high degree of geographical distribution very often results in intermittent connectivity that prevents us from utilizing well-known sustainability and trustworthiness concepts from Cloud Computing or other types of distributed systems. A typical concept to achieve sustainable systems in Clouds is to shut down virtual machines (VMs) in case of low workload. Edge systems very often rely on event driven microservices that do not allow management of resources at the granularity level of VMs. A similar example for the lack of trustworthiness is the requested availability of Edge systems. In Cloud systems, we can achieve high availability by using sophisticated backend scheduling and load balancing algorithms. Sophisticated scheduling and load balancing are difficult at the Edge due to resource scarcity.

In our FWF Rucon (Runtime Control on Multi-Clouds) project,[a] we developed in the last five years the fundamentals for sustainable and trustworthy Edge Machine Learning systems with two major goals:

1) *ML for the Edge:* The first goal is to develop ML based methods for the sustainable and trustworthy operation of Edge nodes, regardless of the applications being executed at the particular Edge node.
2) *ML at the Edge:* The second goal of the Rucon project is to develop methods for the sustainable and trustworthy execution of geographically distributed ML applications (e.g., streaming apps).

In order to make decisions on the huge amounts of data in a relatively short time frame, the whole Rucon architecture was developed in a probabilistic manner capable of dealing with hyper heterogeneity, geographical distribution, intermittent connectivity, and low availability of the nodes. The centerpiece of the Rucon architecture are as follows:

- The novel method for the fault tolerance and trustworthiness based on the Dynamic Bayesian Networks (addressing ML for the Edge).
- A sustainable model management approach based on the Reinforcement Learning (RL) for geographically distributed ML (addressing ML at the Edge).
- A sustainable and trustworthy method for data quality management for failure prone IoT devices (addressing both ML at the Edge and ML for the Edge).

TRUSTWORTHINESS IN RUCON: FAULT TOLERANCE

Edge computing is prone to failures as it trades reliability against other QoS properties such as low latency and geographical prevalence. Failures on the Edge happen much more often than in other large scale systems

[a]http://rucon.ec.tuwien.ac.at

FIGURE 1. (a) Object detection. (b) Traffic light with the camera and Raspberry Pi. (c) App with the alert about the object in the dead corner.

due to geographical dispersion, ad hoc deployment, and rudimentary support systems (e.g., lack of diesel generators to compensate for power outages). Software services that run on Edge infrastructures must rely on failure resilience techniques for uninterrupted delivery. Due to the lack of other support systems, this has to happen at the software layer. Edge nodes are usually deployed in urban areas with space restriction and using low cost devices like well known Raspberry Pis. Figure 1(a)–(c) depicts such a real-life system installed to collect traces and data samples for our experimental smart traffic light system.[b]

We developed an app that visualizes objects on the smartphone that appear in vehicles' dead corners, thus preventing severe accidents. To detect objects [Figure 1 (a)], we used cameras inside a traffic light [Figure 1(b)], together with a Raspberry Pi equipped with convolutional networks for object detection. Once an object is detected in the dead corner, the message is broadcast to all vehicles of interest while the detected objects are visualized on the app as shown in Figure 1(c).[13] As can be seen on the pictures, we used low cost devices suitable for mass rollout in smart cities. However, these devices can easily fail and require sophisticated failure tolerance mechanisms. A well-known approach to counteract low availability is the utilization of geographically distributed replicas for the deployed services. In case of the failure of a service, the workload is redistributed to the standby replica. Standby replicas, however, should not fail concurrently.

In *Rucon*, we developed a novel fault-tolerance mechanism for the redundant service deployment that minimizes the cost (e.g., in terms of the number of redundant services) while preventing joint failures of the replicas. Spatiotemporal dependencies of failures appear very frequently in Edge systems. Reasons might be a network failure affecting multiple servers in the same physical/virtual network or a power outage affecting multiple servers in the same grid or multiple servers deployed in hostile locations failing due to environmental interference. Neglected spatiotemporal dependencies can lead to the so-called cascading failures, and in general to catastrophic effects for the overall reliability of systems. In *Rucon*, we detect spatiotemporal failure dependencies among Edge servers to improve the failure resilience of services with minimum possible redundancy by applying the dynamic Bayesian networks (DBNs). In this architecture, dependence learning occurs in a resource-rich environment such as the cloud based on the received past failure traces of the system. Trained DBNs are then used to perform the inference about the joint failure probability of the random servers.[3]

In our approach, we learn the spatiotemporal dependencies between Edge server failures and combine them with the topological information to incorporate link failures. Eventually, we infer the probability that a certain set of servers fails or disconnects concurrently during service runtime. Our experimental results show that after eliminating the noise and by analyzing randomly large scale failure traces of Edge datasets of various applications (e.g., Skype supernodes), there is a significant amount of spatiotemporal failures. We developed multiple dependence- and topology-aware deployment algorithms that minimize either failure probability or redundancy cost. Experimental results show that we can reduce the service downtime by several orders of magnitude compared to the baseline while preserving the requested latency. The utilization of the deployment

[b]http://intrasafed.ec.tuwien.ac.at

algorithms that consider the joint failure probability can further decrease the redundancy loss up to 50% compared to the baseline. We consider our spatiotemporal failure dependence approach as the first step toward trustworthy edge machine learning.

SUSTAINABILITY IN RUCON: MODEL DISTRIBUTION FOR EDGE MACHINE LEARNING

In a distributed setting, ML models are usually trained in a large scale data center. Afterward (a reduced) version of the model is distributed to the Edge to perform inference in the vicinity of the end users and thus achieve low latency. When deploying ML models over geographically distributed Edge nodes several problems arise, in particular in nonstationary environments when the data distribution changes. Due to environmental changes models that have been learned, trained, and distributed to Edge nodes might become inaccurate and in the worst case no longer valid. In traditional data centers, nonstationarity is solved using so called online learning, where models are trained in batches as new data arrives. Applying online learning in a geographically distributed setting bears several problems in terms of sustainability, where distributed ML models can be independently trained and periodically synchronized through a centralized parameter server. Too frequent updates would result in a heavy message exchange and could lead to bandwidth problems and eventually bad sustainability of the whole system. Less frequent model updates can result in poor performance of the ML models at the Edge.

This dilemma might occur in many geographically distributed streaming applications, which are very common in the area of the "tactile internet." One such example is the e-vehicle application managing recharging intervals of electrical cars. Changing environmental conditions (wind, sun, water, etc.,) cause volatile availability of the electricity, which has to be matched with user requests for short queuing intervals at the charging stations. In this setup, cars communicate with roadside units (RSU) to be updated about the current situation at the charging stations, i.e., to receive the freshest version of the ML model. On the other hand, RSUs are used to collect data from the cars to accurately predict the needs of car fleets on the road.

With our staleness control mechanism proposed in *Rucon*, we tackle the concept drift issues in Edge data analytics to minimize its accuracy loss of the distributed ML without losing its timeliness benefits. We propose an efficient model synchronization mechanism for distributed and stateful data analytics. Our RL-based algorithm learns over time the connectivity patterns of the cars and the most suitable intervals for the distribution of newly collected data in the form of model updates at the RSU. The data are distributed from the car to the RSU on one hand, and also from the parameter server to the RSU in the form of the updated models. Since we use online RL, the algorithm has a low computational overhead, automatically adapts to changes, and does not require additional data monitoring contributing to the sustainability of Edge Machine Learning. Our thorough evaluation shows that we are able to save up to 90% of the updates while having the same quality of the model compared to the fully synchronous oracle approach.[4] Reduced number of model updates directly increases the energy efficiency of the geographically distributed ML applications.

SUSTAINABILITY AND TRUSTWORTHINESS IN RUCON: DATA QUALITY MANAGEMENT

Edge computing is usually utilized to collect and process data from the sensors and other IoT devices that have a very high failure rate. Missing or invalid data may appear very often on the IoT systems due to monitoring system failures, data packet loss, or sensor aging. Consequently, near-real-time decisions are often based on limited and incomplete data. Low data quality might significantly impact accuracy of the decision-making processes on a large scale, e.g., in large scale cloud data centers that use collected, aggregated, and processed IoT data.

Rucon's approach to counteract the low data quality is a generic mechanism for recovery of multiple gaps in incomplete datasets, using multiple recovery techniques. To ensure outliers removal, detection, and forecasting of each gap, we use different techniques addressing different dataset characteristics. The autoregressive integrated moving average (ARIMA) method can be used, if data contain stationary characteristics, such as trend stationarity. The Exponential Smoothing method (ETS) can be used for short-term seasonal series or with multiple complex seasonality. Another feature of *Rucon* is a sustainable Edge data management that achieves a tradeoff between the amount of data stored at the Edge and high accuracy for predictive analytics. Our data quality approach facilitates adaptive storage management mechanisms for reducing the amount of data stored at the Edge, keeping only the data necessary for predictive analytics.[12] We utilize data clustering techniques where we detect stable accuracy clusters. Those clusters can be used as a border between relevant and irrelevant data for the accurate near-real time analytics. Once identified, irrelevant data can be released and thus data storage is optimized on the resource scarce Edge nodes. Data quality management addresses both, sustainability of the Edge storage but also trustworthiness of processed data.

DISCUSSION AND OUTLOOK

In *Rucon*, we have developed the first fundamental approaches for achieving sustainable and trustworthy Edge Machine Learning focusing on the current challenges like hyper heterogeneity and high failure rate of IoT sensors. However, the demand for the sustainable and trustworthy Edge systems will significantly increase in the future as discussed next.

Arbitrary resources. Nowadays, we usually have dedicated Edge nodes, for example, installed at road side units or in combination with 5G antennas. Usually, 5G antennas are equipped with additional servers to process the workload on demand in the vicinity of the end users. For moving objects like drones and scooters, it is even harder to facilitate efficient resource usage as they have intermittent connectivity and very frequent handovers between the Edge nodes. Building stationary Edge nodes, where drones fly every now and then, is highly inefficient. The future challenge is to develop sustainable and trustworthy Edge systems even in case of arbitrary and/or opportunistic computing. In both cases, the idea is to incentivize people to share their resources if there is a high demand for them. The computation could be dynamically offloaded and dynamically migrated on already existing but idle systems (e.g., idle laptop, idle server) in a secure way, if the required middleware for the management and charging of such systems is installed. This paradigm will only succeed, if the resources are trustworthy. On the other hand, arbitrary and/or opportunistic computing is the basic concept of the sharing economy and can create many other benefits in terms of sustainability in the long run.

ACKNOWLEDGMENTS

This work was supported in part by the Austrian Science Fund (FWF Y 904 START-Programm 201) and in part by the City of Vienna (5G Use Case Challenge InTraSafEd 5G).

REFERENCES

1. Okanohara, Daisuke, *et al.,* "Machine learning with model filtering and model mixing for edge devices in a heterogeneous environment," U.S. Patent No. 10 387 794, Google Patent, 0217387A1/en, 2016. [Online]. Available: https://patents.google.com/patent/US
2. CISCO, *VNI Complete Forecast Highlights*, 2016. [Online]. Available: https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecasthighlights/pdf/Global_2021_Forecast_Highlights.pdf
3. A. Aral and I. Brandic, "Learning spatiotemporal failure dependencies for resilient edge computing services," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 7, pp. 1578–1590, Jul. 2021.
4. A. Aral, M. Erol-Kantarci, and I. Brandic, "Staleness control for edge data analytics," *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 4, no. 2, pp. 38:1–38:24, 2020.
5. L. Cavigelli, P. Degen, and L. Benini, "CBinfer: Change-based inference for convolutional neural networks on video data," in *Proc. 11th Int. Conf. Distrib. Smart Cameras*, Stanford, CA, USA, 2017, pp. 1–8.
6. Y. C. Hu *et al.*, "Mobile edge computing: A key technology towards 5G," *ETSI White Paper*, vol. 11, no. 11, pp. 1–16, 2015.
7. K. Fauvel *et al.*, " A distributed multi-sensor machine learning approach to earthquake early warning," in *Proc. 34th AAAI Conf. Artif. Intell./32nd Innov. Appl. Artif. Intell. Conf./ 10th AAAI Symp. Educ. Adv. Artif. Intell.*, 2020, pp. 403–411.
8. A. G. Greenberg, J. R. Hamilton, D. A. Maltz, and P. Patel, "The cost of a cloud: Research problems in data center networks," *Comput. Commun. Rev.*, vol. 39, no. 1, pp 68–73, 2009.
9. N. Jones, "How to stop data centres from gobbling up the world's electricity," *Nature*, vol. 561, no. 7722, pp. 163–167, Sep. 2018. [Online]. Available: https://www.nature.com/articles/d41586-018-06610-y
10. H. Li, K. Ota, and M. Dong, "Learning IoT in edge: Deep learning for the internet of things with edge computing," *IEEE Netw.*, vol. 32, no. 1, pp. 96–101, Jan./Feb. 2018.
11. D. Liu, G. Zhu, J. Zhang, and K. Huang, "Data-importance aware user scheduling for communication-efficient edge machine learning," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 1, pp. 265–278, Mar. 2021.
12. I. Lujic, V. De Maio, and I. Brandic, "Resilient edge data management framework," *IEEE Trans. Serv. Comput.*, vol. 13, no. 4, pp. 663–674, Jul./Aug. 2020.
13. I. Lujic, V. De Maio, K. Pollhammer, I. Bodrozic, J. Lasic, and I. Brandic, "Increasing traffic safety with real-time edge analytics and 5G," in *Proc. 4th Int. Workshop Edge Syst., Analytics Network.*, 2021, pp. 19–24.
14. M. Zaharia *et al.*, "Apache spark: A unified engine for big data processing," *Commun. ACM*, vol. 59, no. 11, pp. 56–65, 2016.

IVONA BRANDIC is a Full Professor for High Performance Computing Systems at the Vienna University of Technology. In 2015 she was awarded the FWF START prize, the highest Austrian award for early career researchers. She received the Ph.D. degree in 2007 and the Distinguished Young Scientist Award in 2011, both from the Vienna University of Technology. Her main research interests are runtime management of large scale distributed systems, Cloud Computing, energy efficiency, QoS and autonomic computing. Contact her at ivona.brandic@tuwien.ac.at.

IEEE COMPUTER SOCIETY

www.computer.org

PURPOSE: Engaging professionals from all areas of computing, the IEEE Computer Society sets the standard for education and engagement that fuels global technological advancement. Through conferences, publications, and programs, IEEE CS empowers, guides, and shapes the future of its members, and the greater industry, enabling new opportunities to better serve our world.

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, 19 journals.

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

Conferences/Education: The society holds more than 215 conferences each year and sponsors many educational activities, including computing science accreditation.

Certifications: The society offers three software developer credentials.

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: +1 800 272 6657; **Fax:** +1 714 821 4641;
Email: help@computer.org

EXECUTIVE COMMITTEE

President:	William D. Gropp
President-Elect:	Nita Patel
Past President:	Forrest Shull
First VP:	Riccardo Mariani
Second VP:	David S. Ebert
Secretary:	Jyotika Athavale
Treasurer:	Michela Taufer
VP, Membership & Geographic Activities:	Andre Oboler
VP, Professional & Educational Activities:	Hironori Washizaki
VP, Publications:	David S. Ebert
VP, Standards Activities:	Annette Reilly
VP, Technical & Conference Activities:	Grace Lewis
2021–2022 IEEE Division VIII Director:	Christina M. Schober
2022–2023 IEEE Division V Director:	Cecilia Metra
2022 IEEE Division VIII Director-Elect:	Leila De Floriani

BOARD OF GOVERNORS

Term Expiring 2022:
Nils Aschenbruck, Ernesto Cuadros-Vargas, David S. Ebert, Grace Lewis, Hironori Washizaki, Stefano Zanero

Term Expiring 2023:
Jyotika Athavale, Terry Benzel, Takako Hashimoto, Irene Pazos Viana, Annette Reilly, Deborah Silver

Term Expiring 2024:
Saurabh Bagchi, Charles (Chuck) Hansen, Carlos E. Jimenez-Gomez, Daniel S. Katz, Shixia Liu, Cyril Onwubiko

EXECUTIVE STAFF

Executive Director:	Melissa Russell
Director, Governance & Associate Executive Director:	Anne Marie Kelly
Director, Conference Operations:	Silvia Ceballos
Director, Information Technology & Services:	Sumit Kacker
Director, Marketing & Sales:	Michelle Tubb
Director, Membership Development:	Eric Berkowitz
Director, Periodicals & Special Projects:	Robin Baldwin

IEEE BOARD OF DIRECTORS

President & CEO:	K. J. Ray Liu
President-Elect:	Saifur Rahman
Director & Secretary:	John W. Walz
Director & Treasurer:	Mary Ellen Randall
Past President:	Susan "Kathy" Land
Director & VP, Educational Activities:	Stephen M. Phillips
Director & VP, Publication Services & Products:	Lawrence O. Hall
Director & VP, Member & Geographic Activities:	David A. Koehler
Director & President, Standards Association:	James E. Matthews
Director & VP, Technical Activities:	Bruno Meyer
Director & President, IEEE-US:	Deborah M. Cooper

BOARD OF GOVERNORS MEETING

TBA

IEEE

revised 27 October 2022

EDITORS: Jens Grossklags, jens.grossklags@in.tum.de
Heather Richter Lipford, richter@uncc.edu
Jessica Staddon, jessica.staddon@gmail.com

This article originally appeared in
IEEE SECURITY & PRIVACY
vol. 20, no. 2, 2022

DEPARTMENT: SOCIOTECHNICAL SECURITY AND PRIVACY

Nudging Software Developers Toward Secure Code

Felix Fischer and Jens Grossklags, *Technical University of Munich*

The prevalence of insecure code is one of the main challenges security experts are trying to solve. We study behavioral patterns among developers which largely contribute to insecure software—googling and reusing code from the Web—and apply nudge theory to harness these behaviors and help developers write more secure code.

Programming is not only a highly difficult task; today it has become utterly complex. There is a vast and quickly growing amount of languages and application programming interfaces. Developers need to be flexible and willing to learn how to apply them in a very short time, and, to deal with this sometimes overwhelming task, they search online for help. Very often they find ready-to-use code examples or open source software that solves the problem at hand. The reuse of these resources provides a very efficient and effective way out. However, it becomes problematic if very popular resources provide solutions that are flawed security wise. Many solutions end up in production software used by billions of people. Some introduce critical vulnerabilities that can be exploited by attackers, for instance, to steal credentials or credit card data or to compromise a device.

We believe that we cannot keep developers from reusing content from the web as this behavior seems to be deeply rooted today. Therefore, we have opted for a different approach that harnesses this observation and tries to guide developers toward content on the web that is safe to reuse. We borrowed this idea from nudge theory, which is a concept from behavioral science and economics. It does not expect people to change their behavior but redesigns things in a way such that common behaviors lead to better outcomes. We redesigned two very fundamental resources—Google Search and Stack Overflow—such that developers can find advice that is helpful and secure.

INSECURE CODING ADVICE ON THE WEB

Stack Overflow is one of the most popular resources. It is a Q&A site that provides helpful advice for almost any coding problem. However, in Fischer et al. 2017,[2] we showed that Stack Overflow provides a large amount of highly vulnerable code examples. Many of them were reused in production code; 15% of apps available on Google Play contained at least one of those insecure snippets.

Even though Stack Overflow provides countless secure code examples that are safe to apply in code, we found that these were hardly reused. In Chen et al.,[5] we compared the popularity of secure and insecure code from Stack Overflow among users. We did this by relying on Stack Overflow's own voting system, which provides a community-given score for each post. Interestingly, insecure code had significantly more upvotes and was more often duplicated across discussion threads or indicated as the top answer. We also found that highly trusted Stack Overflow users—users with a particularly high reputation score—posted insecure code. In other words, all of the very meaningful indicators on Stack Overflow were pointing in the wrong direction security wise.

Digital Object Identifier 10.1109/MSEC.2022.3142337
Date of current version: 21 March 2022

While Stack Overflow is part of most developers' journey through the web, they typically begin with Google Search. They type in a query and usually follow one of the top-ranked results. Depending on Google's ranking algorithm, developers end up on webpages that provide either secure or insecure advice. Therefore, we wanted to know whether top-ranked results are biased toward secure or insecure code and whether this has a direct effect on software security.

In an online study, we asked 192 developers to solve several programming tasks.[4] They were instructed to use Google Search to find help online. Afterward, we analyzed the distribution of secure and insecure advice among the top search results of all participants. The chance to receive at least one insecure result among the top three ranks was 23%—more than twice as high as for secure code. Developers who clicked on one of those links ended up on a Stack Overflow page that provides insecure code in the top answer of the discussion thread.

In summary, not only are Stack Overflow's own content indicators often misleading, but Google Search's ranking algorithm is too. The two fundamental web mechanics that developers rely on to find information on the Internet are inadvertently promoting insecure content.

NUDGE THEORY

The paternalistic way to solve this problem is to urge developers not to use Stack Overflow or even Google Search to look for help online but rather advocate for established resources that are safe. Of course, we do not expect this idea to be fruitful. Several studies explored alternatives, such as formal documentation, books, simplified programming interfaces, and code analysis tools.[1] Even though they do help in improving code security, developers still struggle to get functional solutions out of them. In this regard, the web seems to provide better options. Since functional code is the developers' primary goal, it seems unrealistic to convince developers not to use popular web resources. Behavioral science underpins this assumption: changing people's behavior is very hard! Richard Thaler—one of the inventors of nudge theory—framed it the following way: "First, never underestimate the power of inertia. Second, that power can be harnessed."

Nudge theory attempts to design around people's default behavior in a way that leads to better outcomes for the individual and society as a whole. People do not need to change; the surrounding "choice architecture" is changed. We build upon this theory and rely on the observation that developers often make the easy choice. Copying and pasting code examples from the web is as simple as it gets. By ensuring that people reuse secure examples instead of insecure ones, we can keep this level of convenience. Developers do not need to find alternatives to Google Search and Stack Overflow. We designed several nudges that help them to make safe choices. We applied the following nudges in our work.[6]

The *simplification* nudge has been applied to reduce the complexity of measures related to education, health, finance, and employment. Undue complexity reduces the benefits of measures, causes confusion, and deters participation. We implemented this nudge by moving security advice to already-existing and well-established resources that are being used by almost all developers.

Warnings are nudges that are already deployed in user communication of security issues on the web, for instance, if users visit a malicious webpage. It has been shown that warnings are much more effective if they provide *recommendations* that help people out of a potentially dangerous situation. We designed security warnings for insecure code examples on Stack Overflow. They inform developers why the examples were marked as being insecure and what risks could result from the reuse of the code. Below each warning, we provided an ordered list of recommended Stack Overflow posts that offer a very similar but secure example. In the best case, developers only have to make one additional click to find a functional and secure solution.

Reminders can have a significant impact; however, timing greatly matters. Therefore, whenever we identified a copy attempt of insecure code on Stack Overflow, we showed a reminder nudge that warns the user once more and displays recommendations.

STACK OVERFLOW

We integrated these nudges on Stack Overflow and performed a developer study.[3] Participants were divided into two condition groups. The treatment

group used a modified Stack Overflow version that applied nudging, while the control group used the original Stack Overflow. Both had to solve several security-related programming tasks where we afterward evaluated the security and functionality of the submitted solutions.

The treatment group submitted more secure solutions than the control group with statistical significance. Both groups achieved the same high level of functional solutions, which meant that our nudging interventions did not interfere with the usability of Stack Overflow. This was also a very important result since less functional solutions in the treatment group would result in developers being drawn away from the website. We were not able to isolate a specific nudge being responsible for the effects. It was rather a combination of the displayed warnings, recommendations, and reminders.

GOOGLE SEARCH

The most effective nudge from the literature is the so-called *default* nudge.[6] It automatically preselects the most beneficial choice by default, and people only need to take action if they disagree. Popular examples are automatic enrollments in programs, including education, health, and savings.

A web search generally tries to optimize its ranking in a way that presents the user with the most relevant results. People want to immediately find the information they desire within the top ranks. It is the same for software developers. When searching for code examples, we found that they usually click on one of the top three links. Currently, there is a much higher chance to find insecure code among those results.

From Healthy Food to Secure Code

We approached this problem with an approach similar to the so-called *healthy food nudge*. It has been observed that people usually buy food that is presented at eye level in grocery stores. That means, to nudge people toward eating healthy, one should place healthy food at eye level.

We implemented this nudge in Google Search by putting relevant and secure results "at eye level." In other words, we modified the search ranking in a way that it moves secure and relevant advice to the top three ranks in the results. Developers would then be presented with a secure and relevant choice by default. Since we simultaneously down-ranked insecure results, it becomes even more unlikely that developers will click on one of them.

Ranking Signals

To rerank webpages based on security and relevance, we had to find signals first that sufficiently informed about these properties. In Fischer et al. 2019,[3] we developed a deep learning model that is able to predict whether a Java code example on Stack Overflow is insecure or not. We applied this model to determine the security signal for Stack Overflow pages that discussed questions related to Java. Further helpful tools are publicly available to obtain security signals for different programming languages. For example, LGTM performs large-scale analyses on several popular open source websites, such as GitHub, GitLab, and Bitbucket. It is able to detect the most dangerous known vulnerabilities.

To find relevant results, we tried three different approaches. First, we simply relied on Google Search to find relevant results. Since it is the most popular search engine among software developers, we expected it to perform well in this task. Second, we developed an additional method that identifies the most relevant code examples for a set of given use cases, such as encrypting a message or establishing a secure communication channel. Even though the approach was largely automated, it required manual labeling of a small sample and was also restricted to a programming language and specific use cases. Third, we relied on Stack Overflow's voting system as a signal to identify helpful examples. Both signals—security and relevance—were used to update the ranking algorithm of a custom Google Search engine.

Developer Study

We tested the updated Google Search in comparison to the original Google Search in another online study where developers had to write code to solve several programming tasks.[4] We divided the 218 participants into two groups. The control group was provided with a search bar that used original Google Search. The treatment group used the updated Google Search engine, which applied security-based reranking. Our hypothesis was that the more the treatment group

used our modified search engine, the more functional and secure code they would submit in comparison to the control group.

After we evaluated the results from the study, we found that participants in the treatment group submitted more functional and secure solutions than the control group—with statistical significance—the more they used the modified search engine. This showed that the reranking had a significant positive effect on the security and functionality of the written code.

We performed an in-depth analysis of the retrieved and clicked results. We found that 83% of the results received by participants in the treatment group were secure, while 46% of the results were highly relevant to the query. In contrast, in the control group, 68% of the results were insecure. A similar distribution was also present in the clicks made by our participants. Sixty-seven percent of the clicked results were secure in the treatment group—among those 26% highly relevant—while the control group predominantly clicked on insecure results with 84% of clicks made. These results provide a much clearer picture of the causal chain: a higher usage of search engines, up-ranked relevant and secure results, clicks predominantly made on the top three results, and the reuse of code examples found on the related webpages ultimately led to more functional and secure code.

TRANSPARENCY VERSUS UNOBTRUSIVENESS

Both interventions—on Stack Overflow and Google Search—follow the design principles given by nudge theory. They try to make it as easy and convenient for developers to engage in better security decision making. They achieve this exactly by not interfering with established behavior, such as Googling or copying and pasting code examples. They do not try to restrict any options but rather harness the status quo and lead to better outcomes.

Both approaches do not require developers to be aware of them to use them. Developers do not need to download, install, or learn how to use these methods. They do not have to cope with incomplete or unhelpful documentation or gain the advanced skills that are sometimes required to use security tools such as code analysis.

However, both approaches differ in certain aspects. Warnings and recommendations on Stack Overflow allow developers to make informed decisions on whether or not to reuse insecure code. Security-based reranking of Google Search results provides more secure options by default, without user awareness. On the one hand, the Google Search intervention leads people to stay more or less uninformed about which results are secure and which are insecure and why. On the other hand, developers do not have to pay attention to and follow security warnings, indicators, or recommendations that are often difficult to understand. Moreover, people quickly become habituated to these kinds of interventions. This happens once they disagree with a warning or find recommendations unhelpful.

With the Google Search intervention, developers do not need to evaluate whether vulnerabilities reported by code analysis tools are false positives. Moreover, there are no disruptive effects on the main programming task. The intervention remains completely invisible and does not require anything from the user. Therefore, typical human factors that need to be addressed in the field of usable security may not have any negative effects on security in this approach.

Following the defense-in-depth principle, a combined approach might provide the ideal solution. While Google Search includes code security as a signal in ranking, websites, such as Stack Overflow and GitHub inform and educate their user base about insecure content. This works best if all players are part of the game. Alternatively, a scenario that does not rely on Google and other webpages would be one where companies and institutions run our interventions internally on top of Stack Overflow and Google Search.

Based on the results of our studies, we believe that designing security interventions for developers—as well as for end users—must consider behavioral aspects. In our work, observed behavior formed the basis upon which we designed our interventions. It puts people at the center of the design and dramatically shifts responsibilities away from developers who may be laymen in security toward experts in security and beyond. This way of designing security interventions shows that there is a potential for fixing important security issues in code on a very large scale. The urgency to take action is high as the problem is otherwise much likely to worsen.

REFERENCES

1. Y. Acar, M. Backes, S. Fahl, D. Kim, M. Mazurek, and C. Stransky, "You get where you're looking for: The impact of information sources on code security," in *Proc. 2016 IEEE Symp. Security Privacy (S&P)*, pp. 289–305, doi: 10.1109/SP.2016.25.
2. F. Fischer et al., "Stack overflow considered harmful? The impact of Copy&Paste on Android application security," in *Proc. 2017 IEEE Symp. Security Privacy (S&P)*, pp. 121–136, doi: 10.1109/SP.2017.31.
3. F. Fischer et al., "Stack overflow considered helpful! Deep learning security nudges towards stronger cryptography," in *Proc. 28th USENIX Security Symp. (USENIX Security)*, 2019, pp. 339–356.
4. F. Fischer, Y. Stachelscheid, and J. Grossklags, "The effect of Google search on software security: Unobtrusive security interventions via content re-ranking," in *Proc. 28th ACM Conf. Comput. Commun. Security (CCS)*, 2021, pp. 3070–3084, doi: 10.1145/3460120.3484763.
5. M. Chen, F. Fischer, N. Meng, X. Wang, and J. Grossklags, "How reliable is the crowdsourced knowledge of security implementation?" in *Proc. 41st ACM/IEEE Int. Conf. Softw. Eng. (ICSE)*, 2019, pp. 536–547, doi: 10.1109/ICSE.2019.00065.
6. C. R. Sunstein, "Nudging: A very short guide," in *The Handbook of Privacy Studies*. Amsterdam, The Netherlands: Amsterdam Univ. Press, 2018, pp. 173–180.

FELIX FISCHER is a Ph.D. student at the Technical University of Munich, Munich, 80333, Germany. He is also a senior researcher at Avast. His research studies include the interaction of people with information security and privacy technologies. His most recent publications focus on software engineers struggling with getting cryptography right and explore machine learning as a tool for usable security and privacy. Fischer received a Diplom in mathematics (with a focus on computer science) from Leibniz University Hannover, Germany, in 2014. Contact him at flx.fischer@tum.de.

JENS GROSSKLAGS is a professor of Cyber Trust in the Department of Informatics at the Technical University of Munich, Munich, 80333, Germany. His research and teaching activities focus on interdisciplinary challenges in the areas of security, privacy, and technology policy. Grossklags received a Ph.D. in information management and systems from the University of California, Berkeley. He is a Senior Member of IEEE. Contact him at jens.grossklags@in.tum.de.

IEEE COMPUTER SOCIETY JOBS BOARD

Evolving Career Opportunities Need Your Skills

Explore new options—upload your resume today

www.computer.org/jobs

Changes in the marketplace shift demands for vital skills and talent. The **IEEE Computer Society Jobs Board** is a valuable resource tool to keep job seekers up to date on the dynamic career opportunities offered by employers.

Take advantage of these special resources for job seekers:

- JOB ALERTS
- TEMPLATES
- WEBINARS
- CAREER ADVICE
- RESUMES VIEWED BY TOP EMPLOYERS

No matter what your career level, the IEEE Computer Society Jobs Board keeps you connected to workplace trends and exciting career prospects.

IEEE COMPUTER SOCIETY

IEEE

EDITOR: George Fairbanks, gf@georgefairbanks.com

DEPARTMENT: THE PRAGMATIC DESIGNER

This article originally appeared in
IEEE Software
vol. 38, no. 6, 2021

Decision-Making Principles for Better Software Design Decisions

Antony Tang and Rick Kazman

FROM THE EDITOR

Our own decisions are perfectly rational, but this article has great advice for our teammates, who are not as lucky. Still, it might not hurt us to glance at it first, you know, so we can help them be more rational.—*George Fairbanks*

Making decisions, particularly where the decision space is complex, is hard. Few people are innately good at it, and learning how to make good decisions often takes a lifetime of hard-won experience from making mistakes and suboptimal decisions. In this article, we propose a systematic approach to software design decision making (DM). We break DM down into nine principles that can be taught, learned, and practiced. Each principle addresses one DM aspect that focuses on a specific type of information used in making and evaluating decisions.

DM is a process to determine a course of action; the resulting judgment could be reached by a jolt of intuition, conscious reasoning, or something in between. Good DM is hard to teach and—as evidenced by the ever-growing mountains of technical debt and failures in software—difficult to achieve in practice.

One reason for these challenges is that software is created in greatly varying environments. There are differences in the types of requirements, organizations, stakeholders, technologies, quality requirements, contexts, time and budget pressures, and so on. Facing a plethora of information that includes unknowns (and, often, unknown unknowns), designers must make decisions, even in cases where their experience and knowledge do not offer obvious solutions. The process of DM is, therefore, often unsystematic, and designers are left to make decisions solely based on their experience, intuition, or gut feeling.

Imagine that we give a set of requirements to two developers and ask them to create a system. Is it likely that they will produce identical solutions? Almost certainly not. Will one of the solutions be better than the other? Quite possibly. The decisions made and resulting solutions created by these two developers may be different, and those variations may have important ramifications for system quality. Both developers likely believe that they made good choices and would champion their own design. Should we believe them?

Our judgments are often subject to cognitive biases[1] and limitations (or bounded rationality[2]). As humans, our brain's processing is not always logical; we are subject to influences, such as anchoring bias, attention bias, sunk-cost fallacy, satisficing, confirmation bias, and the Dunning–Kruger effect. These biases affect the quality of our decisions, often in negative ways. We need ways to train ourselves to think logically and systematically when making decisions to counter biases.

Digital Object Identifier 10.1109/MS.2021.3102358
Date of current version: 22 October 2021

 Published by the IEEE Computer Society

SOFTWARE DESIGN DM PRINCIPLES

How can we tell if a design decision is well made? Rather than trying to answer this question directly, we approach this problem by examining decision considerations and reasoning. We outline nine DM principles (the "9Ps") and describe what they mean, how they work, and what we can do to check them. Designers can learn the principles and ask each other reflective questions (RQs) to check their design reasoning. There is evidence to suggest that asking RQs during design helps designers improve design dialogue and reason better.[3,4]

P1: Use Facts

Facts and evidence are the foundations of logical decisions. Incorrect information and unknown facts (or incomplete truths) lead to invalid conclusions. Hearsay can play a part in requirement gathering and technology design. Business analysts may guess instead of verifying user requirements; designers may hear praise for a new technology instead of testing it themselves.

For example, a colleague of ours chose a NoSQL database in a key portion of the reservations system he was designing because he felt it was a good fit; he had not done any prototyping or analysis but made this decision based on hearsay, experience, and gut feelings. Was this the right choice? Time will tell. To check facts, he could have asked, "What evidence supports that NoSQL would satisfy the lifecycle of the system?"

There are simple generic RQs to ask: "Do we have all of the facts?" "What evidence do we have to support this information?" "Are the information sources trustworthy?" When we cannot have all of the facts, we make assumptions.

P2: Check Assumptions

In the absence of facts, we make assumptions to continue with design. For instance, we may not know if a technology can perform adequately until we prototype the software. In making assumptions, we judge the chance, or probability if we want precision, that the assumptions would hold. This is an example of an explicit assumption that is made knowingly.

Explicit assumptions can sometimes be checked and validated to improve certainties and even establish facts through prototyping, pilot testing, or sensitivity analysis. Implicit assumptions, on the other hand, are made unconsciously. For example,

> *IMAGINE THAT WE GIVE A SET OF REQUIREMENTS TO TWO DEVELOPERS AND ASK THEM TO CREATE A SYSTEM. IS IT LIKELY THAT THEY WILL PRODUCE IDENTICAL SOLUTIONS? ALMOST CERTAINLY NOT.*

if we build our application as a set of microservices with node.js using the existing three-tier client–server architecture (which contains a monolithic database tier) without considering compatibility, then this implicit assumption would not be checked, which can create risks. RQs can be used to check assumptions: "Is this an assumption or a fact?" "Have we made or missed any assumptions?" "How certain are we about this assumption?"

P3: Explore Contexts

Contexts are conditions that influence software decisions. There are many contextual factors, such as development resources, financial pressures, legal obligations, industry norms, user expectations, and past decisions. For example, we want to implement a scalable and highly reliable database system, but our budget is limited. The budget is not a system requirement, but it affects our decision on database license procurement.

Some contexts will end up being constraints on design. Design contexts shape our decisions implicitly, and they are often diverse external factors that are not necessarily technology related. Exploring contextual factors can broaden our design considerations. To check that we have considered contexts, we may ask the following: "What are the contexts that could influence X?" "Have I missed any contexts?" "Does the team have experience in implementing X?"

P4: Anticipate Risks

A risk is the possibility of an undesirable outcome. A documented risk contains an estimate of the size and probability of the loss. There are many risks that a designer needs to estimate, such as extreme spikes of demand and security attacks. Anticipating and quantifying risks is the process of exploring the unknowns and estimating the possibility of risks occurring as well as, if they occur, their impacts.

This is challenging, but designers may use techniques such as the spiral model or risk- and cost-driven architecture to decide what is acceptable and how risks can be mitigated. RQs can help designers identify risks: "What are the potential undesirable outcomes?" "Is there a chance that X would not work?"

P5: Assign Priorities

Priorities quantify the relative importance of choices, such as which requirement to implement or solution to use. If we can afford to implement only one of the two requirements, which one is more important? Prioritization is required when the things that we desire are competing for the same limited resource, such as time, money, developer skills, CPU, memory, or network bandwidth.

Some of these are contextual factors that add constraints. To sort out our priorities, we can ask the following: "Which requirement is more important?" "What can we do without?" "What should we use this resource for?"

P6: Define the Time Horizon

The time horizon defines the time period relevant to a decision and its effects. Risks, benefits, costs, needs, and impacts can change over time, and we want to anticipate how they evolve. For example, we might estimate that the system processing load will reach 85% capacity in three years. Defining the time horizon allows designers to explicitly state and evaluate the pros and cons of actions (and nonactions) in terms of their short- and long-term impacts.

Without explicitly considering the time horizon and reasoning with it, long-term implications may be undermined, or short-term needs may be ignored. RQs can be asked about the time horizon: "What would be affected in the short and long term if I decide on X?" "What needs to be considered in different time horizons for X?"

P7: Generate Multiple Solution Options

Some designers accept the first solution they find without considering further options. If the architect is experienced, and the problem is well understood and low risk, this may be ideal. However, in more challenging contexts, a single solution may be risky; the first solution is not necessarily the best, especially when a designer is inexperienced or facing an unfamiliar situation.

This behavior may be due to anchoring bias—a refusal to let go of the first idea. Generating multiple solution options helps a designer broaden choices and stimulate creativity. RQs can help to broaden solution ideas: "Are there other solutions to this problem?" "Can I find a better solution than X?"

P8: Design Around Constraints

Constraints are limitations that set the boundaries of what a solution cannot do. They may come from requirements, contexts, technologies, or the existing design. For instance, a CPU can compute only W instructions per second, the budget of the project is $X, the number of concurrent users supported by a software license is Y, platform Z doesn't support a certain protocol, developers have no experience with some technology, and so forth.

In software development, we often find connected sets of constraints: if we choose component A, we must also use component B. When there are no apparent solutions, designers must work around constraints and introduce novel solutions, relax parameters, or manipulate the context. A designer can check constraints by asking the following: "If I choose X, how would it affect the solution?" "Are there any constraints that could impede this solution?"

P9: Weigh Pros and Cons

Pros and cons represent the arguments for and against each of the choices in a selection. Weighing pros and cons is a tradeoff evaluation[5] that takes place when there is more than one choice to consider. The evaluation of the pros and cons, quantitatively or qualitatively, allows designers to decide what to take and give up. A quantitative evaluation can be based on measurable elements, such as costs, benefits, priorities, immediacy (i.e., the time horizon), and risks.

However, some pros and cons cannot be easily quantified. Consider the navigation menu design of a mobile app: how can one quantify the pros and cons of a hamburger menu versus a set of tabs? In this case, qualitative arguments, such as the ease of access and learning as well as the effort to implement, can be marshalled. Weighing pros and cons offers designers the chance to think about relative benefits and drawbacks and whom they affect. To check tradeoffs, one can ask these RQs: "Are there more relative benefits in solution X than Y?" "Is the tradeoffs evaluation reasonable?"

EXAMPLE OF DM PRINCIPLES IN PRACTICE

An equipment manufacturer wants to bring a new model to market. A new sensor and sensor software have not been fully tested (P1), and the design is not according to the industry standard (P4). The rush is due to the time to market and competition (P3). Designers have the choice to redesign and delay delivery (P6) or compromise system safety (P2).

With several solution options (P7), the DM considerations revolve around the evaluation of P4—the probability and negative impacts of the software failing. The time to (P6) and contractual obligation of (P1, P3, and P8) delivery creates the tradeoffs (P9) between meeting the deadline and hoping that the chosen solution is good enough (P2) versus missing the deadline and performing thorough testing to increase confidence in the quality (P5).

Let us apply this example to two different equipment manufacturers: an aircraft company making a safety-critical system (P3) and a coffee machine maker developing a sensor for warming coffee (P3). The risks (P4), priority (P5), time horizon (P6), options (P7), constraints (P8), and considerations for the sensor software would be totally different. The DM considerations and responsibilities of the software teams would also be vastly different, and so would the outcome.

APPLYING DM PRINCIPLES

The application of the 9Ps can be incorporated into everyday software development practices, such as Scrum meetings and architecture evaluation. For instance, during Scrum retrospectives and architecture evaluations, designers can focus on the relevant principles and use RQs to check them: "What assumptions have been considered?" "Are they factual and realistic?" "What are the contexts for this requirement?" "What constraints were imposed by the contexts?"

These RQs can help to tease out different aspects and the reasoning of decisions. If a designer is fixated on an idea without giving good reasons, it might be a symptom of anchoring bias. Asking DM principle-based RQs can help to clarify the thinking, and this may rectify such a bias.

DM principles provide the basic perspectives for logical reasoning. For example, Swift code compiles on iOS but not Android (P1 and P8). If we want portability over both platforms (P1 and P3), we will not choose Swift. This is deductive reasoning. Reasoning can help to avoid biases and fallacies, such as the appeal to force (the authority says so), appeal to people (emotional arguments), or accident fallacy (applying general rules to specific cases).

Designers can also use inductive reasoning, such as analogies, to make decisions. Someone who is familiar with one solution may decide that it can be applied to another system (P2). Such analogical reasoning can be problematic if the contexts (P3) of the two systems are significantly different, as in the example presented, falling into the trap of the fallacy of weak induction and hasty generalization.[6]

Good designers are made, not born. Learning to use DM principles is like learning martial arts. Karate and Taekwondo students repeatedly practice the basic stances and patterns until mastery is achieved and those motions become natural. Similarly, DM principles can be used to train basic DM skills, like learning the techniques of martial arts.

This learning, however, should not be confused with what one might do in reality. In a real fight, reactions and responses must be fluid. In a real design exercise, a designer needs to quickly choose the appropriate principles to use and reason with them.

Design is context dependent. There is no prescribed, optimal order for which DM principle to use. By practicing these principles, in software development and everyday decisions, DM will become more intuitive. Master Yoda would say, "Your facts, contexts, and assumptions gather; your time, priority, and constraints measure; your risks and tradeoffs control." May the force be with your designs.

We have introduced the 9Ps to help designers make better and more predictable decisions. Cognitive biases and limitations are difficult to overcome, but training developers to use DM principles can create good reasoning habits. DM principles are not a panacea; they complement but cannot replace the other important attributes that software developers need to possess: domain and technical knowledge, creativity, foresight to imagine possible futures, logical thinking, openness to new information and a willingness to correct mistakes, and a quality mindset to achieve the best outcomes.

REFERENCES

1. D. Kahneman, *Thinking, Fast and Slow*. Baltimore, MD: Penguin, 2011.
2. H. A. Simon, *The Sciences of the Artificial*, 3rd ed. Cambridge, MA: MIT Press, 1996.
3. M. Razavian, A. Tang, R. Capilla, and P. Lago, "In two minds: How reflections influence software design thinking," *J. Software: Evolution Process*, vol. 28, no. 6, pp. 394–426, 2016. doi: 10.1002/smr.1776.
4. A. Tang, F. Bex, C. Schriek, and J. M. E. M. van der Werf, "Improving software design reasoning—A reminder card approach," *J. Syst. Softw.*, vol. 144, pp. 22–40, 2018. doi: 10.1016/j.jss.2018.05.019.
5. R. Kazman, M. Klein, M. Barbacci, T. Longstaff, H. Lipson, and J. Carriere, "The architecture tradeoff analysis method," in *Proc. 4th IEEE Int. Conf. Eng. Complex Comput. Syst. (ICECCS '98)*, 1998, pp. 68–78.
6. P. J. Hurley, *A Concise Introduction to Logic*. Belmont, CA: Thomson Wadsworth, 2006.

ANTONY TANG is an adjunct professor at Swinburne University of Technology, Melbourne, Victoria, 3106, Australia, and Vrije Universiteit Amsterdam, Amsterdam, 1081 HV, The Netherlands. Contact him at https://orcid.org/0000-0002-3574-3977 or atang@swin.edu.au.

RICK KAZMAN is a professor at the University of Hawaii, Honolulu, Hawaii, 96822, USA and a visiting researcher at the Software Engineering Institute of Carnegie Mellon University, Pittsburgh, Pennsylvania, USA. Contact him at https://orcid.org/0000-0003-0392-2783 or kazman@hawaii.edu.

Get Published in the New *IEEE Open Journal of the Computer Society*

Submit a paper today to the premier new open access journal in computing and information technology.

Your research will benefit from the IEEE marketing launch and 5 million unique monthly users of the IEEE *Xplore*® Digital Library. Plus, this journal is fully open and compliant with funder mandates, including Plan S.

Submit your paper today!
Visit www.computer.org/oj to learn more.

IEEE COMPUTER SOCIETY | IEEE

EDITOR: Riccardo Mariani, NVIDIA, rmariani@nvidia.com

DEPARTMENT: STANDARDS

This article originally appeared in
Computer
vol. 54, no. 3, 2021

The Importance of Interoperability in Functional Safety Standards

Riccardo Mariani, *NVIDIA*
Nir Maor, *Qualcomm*
Jyotika Athavale, *NVIDIA*
Kevin Gay, *Aurora*

The increase in standardization activities for automated vehicles is creating interoperability and information-exchange challenges for methodologies, models, and architectures. IEEE is addressing these issues through two standardization projects in functional safety: P2846 and P2851.

In the past few years, there has been a flourish of standardization activities related to functional safety, for example, in the case of automated vehicles (AVs). These standards, published by a few development organizations, such as IEEE, the International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC), and UL, have provided requirements for specification, development, and testing of safety-critical elements at different levels [vehicle, system, software (SW), hardware, and integrated circuits (ICs)].

IN THE PAST FEW YEARS, THERE HAS BEEN A FLOURISH OF STANDARDIZATION ACTIVITIES RELATED TO FUNCTIONAL SAFETY, FOR EXAMPLE, IN THE CASE OF AUTOMATED VEHICLES.

However, these requirements typically are very general, and so the resulting detailed implementation (in terms of methodologies, models, and architectures) can vary to the point that exchanging and combining the work products across the supply chain becomes extremely difficult. For example, a general requirement to provide a failure modes and effects analysis (FMEA) for an IC could cause, in the absence of specific guidelines, different silicon providers to produce disparate FMEAs with varying levels of abstractions and different assumptions. The deviations could be so large that the user of those FMEAs [the Tier 1 or original equipment manufacturer (OEM)] could expend significant effort to combine them for the system-level FMEA. Similarly, a general requirement to specify a safe-driving policy for automated driving could cause, in the absence of specific models, various OEMs or Tier 1s to produce implementations with such large differences that interoperability and verifiability versus common criteria (such as regulations) could be difficult or even impossible to achieve.

This challenge is becoming so critical that, in January 2020, the IEEE Computer Society (CS) decided to start a couple of standardization activities to address specific aspects related to the interoperability of functional safety standards: IEEE P2846[1] (sponsored by the IEEE Vehicular Technology Society and cosponsored by the CS) and P2851[2] (sponsored by the CS). The following paragraphs provide a status on the activity of the two projects after a year of development.

Digital Object Identifier 10.1109/MC.2021.3050453
Date of current version: 12 March 2021

 Published by the IEEE Computer Society

IEEE P2846

Reasonable and foreseeable assumptions play a critical role in the safety-related models used in automated driving systems (ADSs); however, the current body of industry consensus standards does not address how they are identified or establish a minimum set that AV developers should utilize. With that in mind, IEEE P2846 was created with the goal of identifying the minimum set of reasonable assumptions used in foreseeable scenarios to be considered for road vehicles in the development of safety-related models. While the specific values of the assumptions used in an ADS may be specified by regulation or selected by the ADS developer, the minimum set used within safety-related models is common to all ADS developers, regardless of what model is being used.

The IEEE P2846 Working Group (WG) is currently composed of 30 member organizations that encompass government agencies, research institutes, AV developers, OEMs, and Tier 1 suppliers. The WG has representatives from all over the globe, including Europe, the United States, Israel, Japan, and China, and it is led by Intel (chair), Waymo (cochair), and Aurora (secretary).

While the COVID-19 pandemic has impacted the WG's ability to meet in person, overall the group has made great strides in developing this standard during these challenging times. Over the past year, the WG utilized a set of five task forces operating in parallel to develop specific sections of the draft standard, which were assembled to create the draft standard. The WG also dedicated an entire week in November to virtual meetings to review and resolve hundreds of comments on the first complete draft standard submitted by the member entities.

The core content produced by the task forces that currently comprises the draft standard is organized into three major sections. First, the "Scenarios and Assumptions" section identifies a set of scenarios covering safety-relevant driving situations that an AV may encounter in operations on public roads and, within each scenario, the minimum set of assumptions that shall be considered to increase driving safety. As Figure 1 illustrates, the minimum set of reasonably foreseeable assumptions defined by this standard includes properties of other road users, such as velocity v, heading h, rate of change of the heading angle h', braking capabilities β, and response times ρ.

Next, the "Common Attributes of Suitable Safety-Related Models" section identifies a summary set of recommended attributes for safety-related models used in the dynamic driving task. To arrive at this list, the WG conducted a literature review of contributed safety-related models, including sources on responsibility-sensitive safety,[3] the Safety Force Field,[4] rule books,[5] and others. Finally, the "Verification Methods for Assumptions Used in Safety-Related Models" section identifies techniques, such as various design and testing processes, that can be used to verify whether the implementation of a safety-related model conforms to the minimum set of required reasonably foreseeable assumptions defined in the standard.

The third draft of the standard is currently going through a final round of updates before it is shared via liaison agreements with the Society of Automotive Engineers and ISO for the first set of external reviews. The WG is targeting the second quarter of 2021 to submit the standard for balloting with the Vehicular Technology/Intelligent Transportation Systems Standards Committee and to simultaneously initiate a 60-day public review period. The goal is for this standard to be officially published by the end of 2021.

IEEE P2851

The IEEE P2851 goal is to provide an exchangeable and interoperable format for safety analysis and verification activities to facilitate intellectual property (IP) and system-on-chip (SoC) providers to deliver results to safety-critical system integrators in a consistent way and also enable interoperability among tools provided

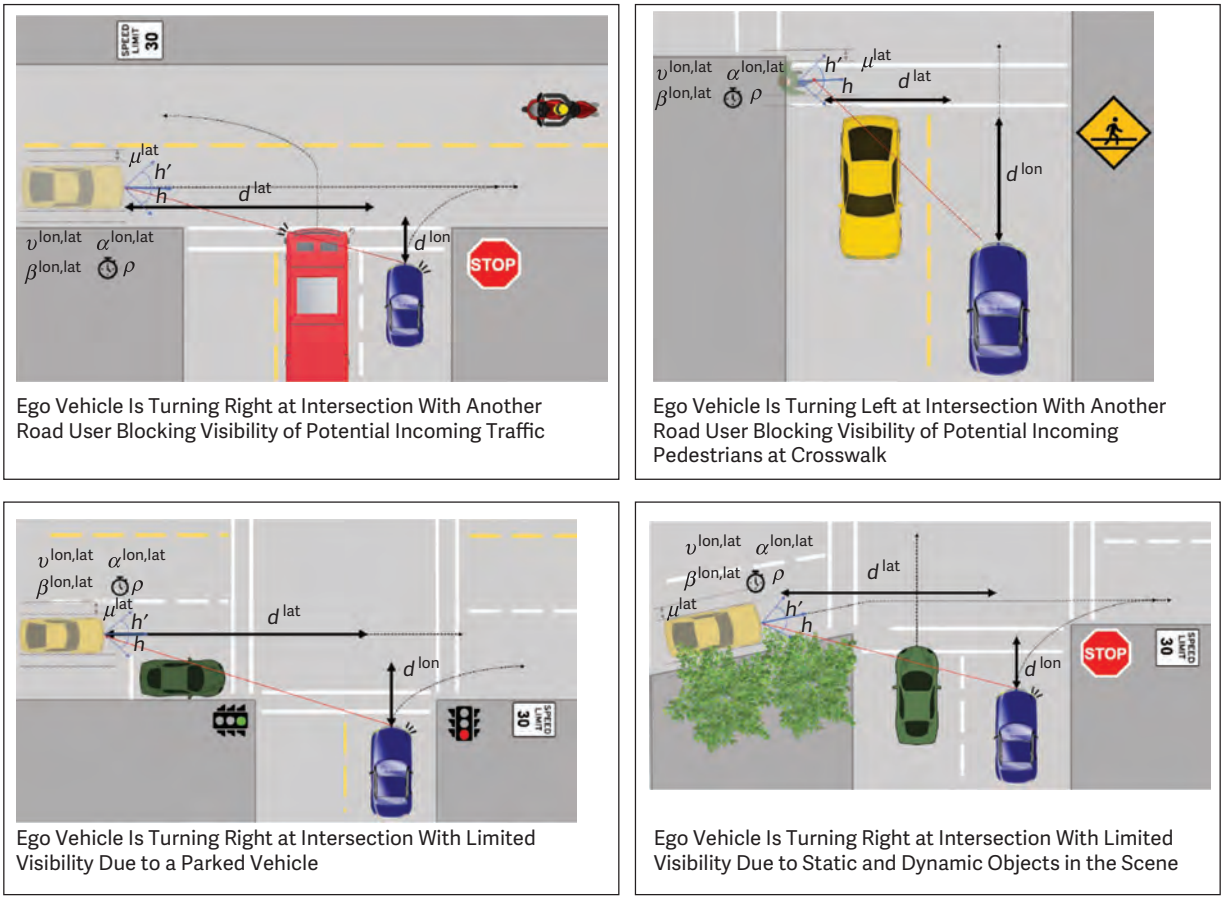

FIGURE 1. A scenario of an intersection with occlusions from IEEE P2846-D2.

by electronic design automation (EDA) tool vendors. IEEE P2851's initial scope was IPs and SoCs, but it has been extended to include items, systems, and SW as well. Artificial intelligence is also a key part of the activity. The WG has representatives from 31 entities, including all of the major IP/SoC providers, EDA vendors, Tier 1s, and OEMs. It is led by NVIDIA (chair and secretary) and Qualcomm (cochair).

The development of IPs and SoCs for safety-critical applications is emerging rapidly because of the growth of applications such as automated driving and robotics. Standards such as ISO 26262 (automotive),[6] IEC 61508 (industrial),[7] and many others require IP and SoC providers to execute safety analysis and related verification activities and deliver results to system integrators. EDA vendors are starting to provide tools to automate activities; however, at this time, there is no common language or format to provide the results. For that reason, the safety-critical community is demanding a solution to accelerate the safety engineering process while reducing risks and costs.

IEEE P2851 will define a data format with which results of safety analyses and related safety verification activities executed for IPs, SoCs, and mixed-signal ICs can be exchanged and made available to system integrators. The format will define languages, data fields, and parameters with which the results of the analyses and verifications can be represented in a technologically independent way. IEEE P2851 will provide a common ground for EDA, SoC, and IP vendors to develop tools, SoCs, and IPs for safety-critical applications.

The end goal is for IEEE P2851 to become a family of standards (P2851.1, P2851.2, P2851.3, and so on) covering broader functional safety topics, such as system- and SW-level safety analyses and formal/semi-formal representations of assumption of use, and also extending to adjacent domains, such as cybersecurity analyses and related verification methodologies.

IEEE P2851 defines a dependability landscape based on an overall product dependability lifecycle (PDL), as represented in Figure 2. The word *dependability* has been selected to cover the broad spectrum of functional safety, safety of the intended functionality (SOTIF),[8] cybersecurity, and other characteristics, such as reliability, maintainability, and real time.

The landscape is represented based on a V-model, as shown in Figure 3. Each level (item, system, component, and part/unit) includes one or more activities belonging to phases of the PDL. Activities are connected with intra- or interlevel interfaces. Each activity can be linked to methodologies and tools to be executed.

Currently, the IEEE P2851 WG members are working on the landscape use-case activities within six subgroups: Automotive Functional Safety, Artificial Intelligence, Avionics, Security, Industrial/Medical/Robotics, and SOTIF. By the end of March 2021, the WG is scheduled to publish a white paper based on the first version of the landscape document, describing the lifecycle activities and related needs of methodologies and tools. By 2021 year's end, the plan is to release a first draft of the standard, and by the end of 2022, a final version of the standard will be published.

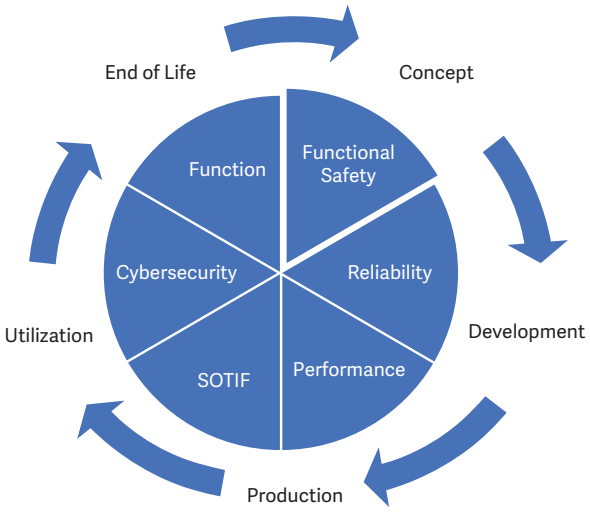

FIGURE 2. A representation of the IEEE P2851 PDL. SOTIF: safety of the intended functionality.

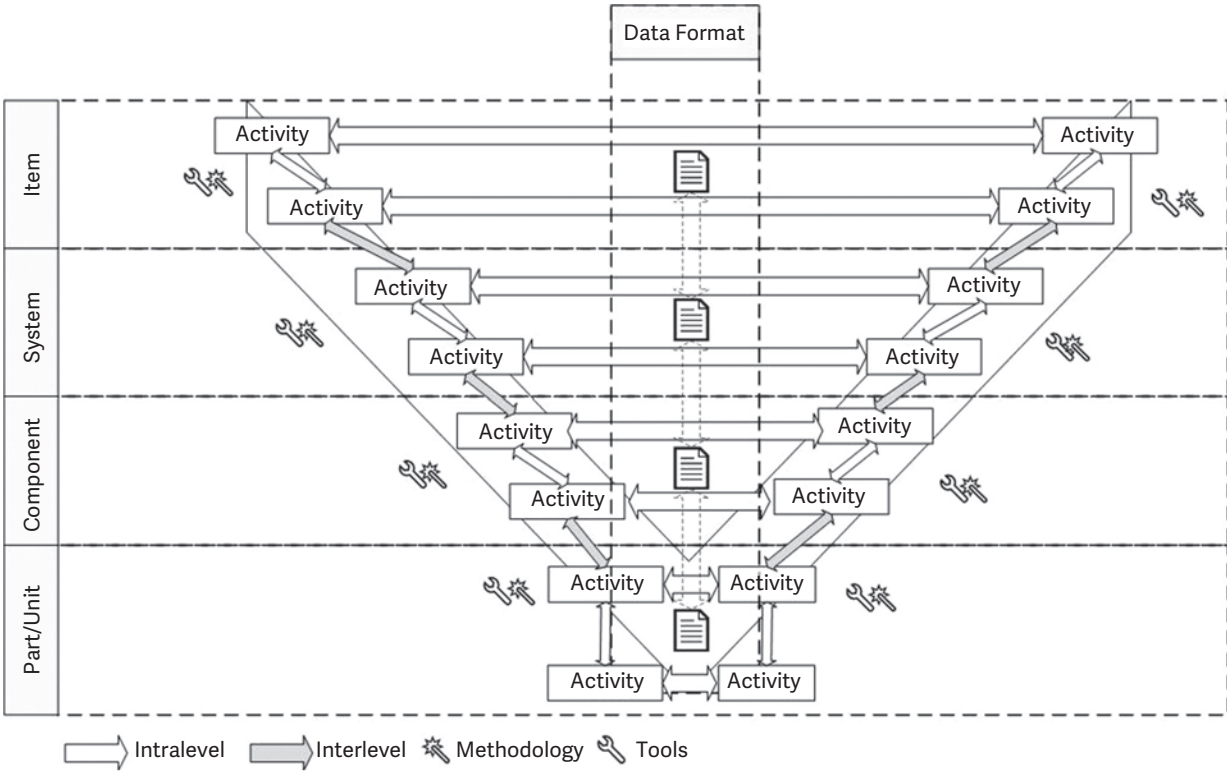

FIGURE 3. An IEEE P2851 landscape representation.

REFERENCES

1. "*Assumptions for models in safety-related automated vehicle behavior*," IEEE Standards Association, Piscataway, NJ. Accessed Feb. 2021. [Online]. Available: https://sagroups.ieee.org/2846/
2. "*Exchange/interoperability format for safety analysis and safety verification of IP, SoC and mixed signal ICs*," IEEE Standards Association, Piscataway, NJ. Accessed Feb. 2021. [Online]. Available: https://sagroups.ieee.org/2851/
3. S. Shalev-Shwartz, S. Shammah, and A. Shashua, "On a formal model of safe and scalable self-driving cars," 2017. [Online]. Available: https://arxiv.org/abs/1708.06374
4. D. Nistér, H.-L. Lee, J. Ng, and Y. Wang. "*The safety force field*." NVIDIA.com. https://www.nvidia.com/content/dam/en-zz/Solutions/self-driving-cars/safety-force-field/the-safety-force-field.pdf (accessed Feb. 2021).
5. A. Censi et al., "Liability ethics, and culture-aware behavior specification using rulebooks," 2019. [Online]. Available: https://arxiv.org/abs/1902.09355
6. *Road Vehicles—Functional Safety*, ISO 26262, 2018.
7. *Functional Safety of Electrical/electronic/Programmable Electronic Safety-Related Systems*, IEC 61508, 2010.
8. *Road Vehicles—Safety of the Intended Functionality*, ISO/PAS 21448, 2019.

RICCARDO MARIANI is the vice president of industry safety at NVIDIA, Santa Clara, California, 95051, USA. He is the 2021 IEEE Computer Society first vice president and chair of IEEE P2851. Contact him at rmariani@nvidia.com.

NIR MAOR is a senior director of technology at Qualcomm Technologies, San Diego, California, 92121, USA. He is the IEEE P2851 vice chair. Contact him at nmaor@qti.qualcomm.com.

JYOTIKA ATHAVALE is a senior functional safety architect at NVIDIA, Santa Clara, California, 95051, USA. She also serves on the IEEE Computer Society (CS) Board of Governors and the core team of the CS Special Technical Community on Reliable, Safe, Secure and Time-Deterministic Intelligent Systems. Contact her at jathavale@nvidia.com.

KEVIN GAY is a senior program manager at Aurora, Pittsburgh, Pennsylvania, 15201, USA. He is the IEEE P2846 secretary. Contact him at kgay@aurora.tech.

IEEE CAI

2023 IEEE Conference on Artificial Intelligence

Hyatt Regency
Santa Clara, California, USA
5–6 June 2023

NEW EVENT

CALL FOR PAPERS

POSTERS | WORKSHOPS | PANEL PROPOSALS

IEEE CAI is a new conference and exhibition with an emphasis on the applications of AI and key AI verticals that impact industrial technology applications and innovations.

IMPORTANT DATES

Workshop & Panels proposal deadline: 1 Feb 2023
Poster paper submission deadline: 19 Mar 2023
Acceptance notifications & reviewers' comments: 30 Apr 2023
Final reviewed submission deadline: 21 May 2023

IEEE CAI seeks original, high-quality proposals describing the research and results that contribute to advancements in the following AI applications and verticals:

AI IN HEALTHCARE/LIFE SCIENCES

Explores the need for improved decision-making to assist medical practitioners as well as additional medical issues including personnel allocation and scheduling, automated sensing, improved medical devices and manufacturing processes, and supply chain optimization.

INDUSTRIAL AI

Addresses robust cyber-security; the effective use of Digital Twins, and Comprehensive Prognostics and Health Management for industrial assets in Aviation, Oil & Gas, and Power Generation, to reduce/eliminate unplanned maintenance events.

AI IN TRANSPORTATION/AEROSPACE

Focus includes the optimal design of aerospace systems using AI; large- and small-scale transportation systems; deep learning for autonomous driving; and autonomous decision-making in long-term space flight management.

AI FOR EARTH SYSTEMS DECISION SUPPORT

Examining the use of AI to help with the analysis of complex events and their effect on the future of agriculture, power generation, and other environmental sectors of critical importance to humanity.

AI IN ENERGY

AI tools have a noticeable impact on energy sectors and can be utilized in Power load forecasting, Power generation forecast of renewable energy to improve agility and resilience; Smart Grid Control, and Power Network Security Protection.

ETHICAL & SOCIETAL IMPLICATIONS OF AI

We will cover Explainability and Interpretability, to enhance understanding, trust, and informed decision making; Transparency in the AI data, system, and business; Privacy, Governance Risk, and Compliance; Robustness, Security/Resilience to attacks; and Fairness/Bias.

Access all submission details: https://cai.ieee.org/2023/authors

Papers accepted by IEEE CAI will be submitted to the IEEE Xplore® Digital Library. Selected high-quality papers will be further invited for submission to a journal special issue.

COMMITTEE INFO

General Co-Chair: Piero Bonissone
General Co-Chair: Gary Fogel

PLATINUM SUPPORTER

QUANTINUUM

IEEE Computational Intelligence Society

IEEE COMPUTER SOCIETY

IEEE SMC Systems, Man, and Cybernetics Society

IEEE Signal Processing Society

IEEE

EDITOR: Dimitrios Serpanos, CTI DIOPHANTUS and University of Patras, serpanos@computer.org

This article originally appeared in Computer vol. 54, no. 11, 2021

DEPARTMENT: CYBER-PHYSICAL SYSTEMS

Pushing the Limits of Autonomy for Enabling the Next Generation of Space Robotics Exploration Missions

George Nikolakopoulos, *Luleå University of Technology*

Ali Agha, *NASA's Jet Propulsion Laboratory and California Institute of Technology*

The never-ending human curiosity about exploring the universe is pushing the limits of robotic autonomy from remote-controlled to fully autonomous systems characterized by advanced learning, cognition, and reasoning for operating in completely unknown and unstructured environments.

The latest NASA Mars 2020 mission initiated a radical paradigm shift in research toward autonomy as a fundamental enabler for the next generation of fully autonomous space robotics missions. The current space directions are toward the return of humans on the lunar surface and Mars and the corresponding future colonization of these planets. Thus, even if the majority of operations on *Perseverance* and *Ingenuity* are still remotely controlled, there is common agreement in all of the related space road maps[1] as well as in the upcoming ARTEMIS program[2] that the future of space exploration and colonization will be realized by empowering the self-sustainability and resiliency of embodied autonomous control-oriented cyberphysical systems (CPSs). The overarching long-term goal is to answer the fundamental question of humanity: "Is or was there life beyond Earth?"

The answer to this question lies below the surface of planetary bodies in our solar system since subterranean (SubT) voids are the most probable places to find signs of life (both extinct and extant), and they are one of the main candidates for future colonization beyond Earth. The exploration of such SubT environments of other planetary bodies, especially on the moon and Mars, is currently the focus of space autonomous missions, motivated by the potential scientific returns of in situ resource utilization and future human exploration.

Unlike planetary surfaces, which provide disturbed and perturbed surface environments, deep SubT voids and their surroundings, as the ones presented in Figure 1, provide access to unaltered materials such as the geologic history of water remains and other volatile compounds on the planet. Thus, future autonomous planetary missions will focus on subsurface explorations, considering these areas to have the maximum potential for the discovery of sources of life or precious materials for planetary settlements.[3]

Major trends are emerging in the field of next-generation autonomous space robotics technologies. These advances will kick-start a paradigm shift in the space industry—from the current remote-operated missions to fully autonomous ones as well as from fragile to resilient—as the only way to bring humanity closer to inhabiting other planets in our solar system.

These technological breakthroughs are introducing radical novel electromechanical robotic architectures. For the first time ever, these are able to unify multimodalities in robotic locomotion while

Digital Object Identifier 10.1109/MC.2021.3055914
Date of current version: 22 October 2021

 Published by the IEEE Computer Society

(a) (b) (c)

FIGURE 1. Examples of a (a) rocky planetary surface, (b) skylight, and (c) lava tubes system.

creating an embodiment of autonomy, enabled by increased durability, self-sensing, self-healing, and self-reconfiguration in mechanical and CPS aspects as well as the next-generation trusted autonomy framework for space robotics.

In these directions, the need for intelligent reconfiguration and adaptation in all of the aspects of the CPS while retaining the hard real-time characteristics of the control loops is more evident than ever. Autonomous CPSs, fully integrated with embedded intelligence systems for onboard fast and multidimensional data processing in parallel computational architectures and on special-purpose redundant processing circuits, are needed to create the proper resilient baseline to support the extreme demands for autonomy.

To this end, for the last three years (2018–2021), researchers on the CoSTAR team[4] have been participating in the prestigious DARPA Subterranean Challenge.[5] The group's aim is to develop fully autonomous robotic systems to explore subsurface voids,[6] with a dual focus on planetary exploration and terrestrial applications in search and rescue, the mining industry, and artificial intelligence/autonomy in extreme environments.

Additionally, new dominant research trends are emerging toward resilient autonomy for space robotics, with special attention to the exploration of planetary voids as well as hostile, extreme, unknown, and fully unstructured environments. All of these share the same backbone that combines three design principles for embodying resiliency in autonomous space systems: robustness, redundancy, and resourcefulness. These characteristics are reflected in both the software and hardware design, as, for the first time ever, it is desired that the adaptation and reconfiguration of the CPS's architectures be in full orchestration with the hardware reconfigurations and multimodality operating concepts that unify locomotion modalities in ground and aerial vehicles with embodied intelligence.

Toward the next generation of autonomous robots for space exploration, typical examples include the following:

- the demanding and challenging unification of SPOT with a microaerial vehicle (MAV), as indicated in Figure 2(a)
- multisensorial sensor fusion for enhancing perception and localization in fully autonomous exploration missions in Figure 2(b)
- the novel exploration rapidly exploring random trees algorithm,[7] which considers the distance; predicted dynamical model-based actuation; and information gain for multiple possible exploration locations, such as frontiers, and

(a) (b)

FIGURE 2. (a) An MAV integrated on top of the SPOT legged robot and (b) unified multisensorial fused navigation for unknown and unstructured space environments.

selects optimal trajectories for an exploration mission

- multisession mapping methodologies, including loop closure approaches and online point cloud-based synchronous collaborative map-merging for multirobot mapping/exploration applications, as depicted in Figure 3.

FUTURE ALGORITHMIC RESEARCH CHALLENGES

Toward this ground-breaking vision and the overall quest for enhanced autonomy, hypermodality robotic systems would provide a full unification among ground and air vehicles while enabling a truly all-terrain mobility. This hardware advancement needs to further challenge recent trends in the core of the software-embedded intelligent systems and, specifically, toward the resilient autonomy framework for intelligence embodiment. The following sections describe characteristic research challenges.

Next-generation sensing and perception schemes for hypermotility space robotics systems

The backbone for autonomous planetary surface and subsurface exploration is to build up the right level of scene perception. This allows planners to perceive the robot ego motion as well as encode the necessary spatial awareness.

Thus, research in this direction will address collaborative multisensorial and hypermodality-based fused localization. It should focus on resiliency, robustness, and safety as well as mechanisms to handle spatial and temporal variances of multiple pose estimators, which will constitute a continuous and reliable state-estimation source, irrespective of the hypermodal robot formation (for example, unified operation or decoupled in different multiple-modalities operations).

Moreover, multidimensional collaborative environmental perception and reasoning architectures need to be further advanced to contribute to the establishment of a unified back end for each robot modality. This will act in a decentralized mode for joint map and trajectory refinement, under specific communication policies, within the context of scene awareness and map building.

Mobility techniques for hypermodality robotic systems

This research direction will introduce a novel frontier generation framework that is able to increase the volume coverage at the minimum exploration time while respecting traversability and perception constraints, establishing a mechanism based on the hypermodality properties of the platform. The decision-making part of the exploration module will introduce novel information-gain metrics while considering risk and uncertainty bounds.

The framework will, at the same time, optimize the navigation and next-feasible-waypoint generation for the unified platform but also for each modality, considering the related information gain, power consumption, faults, modality-specific dynamics, and area traversability. Unifying the obstacle-free path planning, mapping effort, risk awareness, and considerations of the multidimensional motion primitives and kinematics of each robotic module under one navigation framework will establish a resilient system that will allow collaborative and distributed exploration of 3D SubT environments.

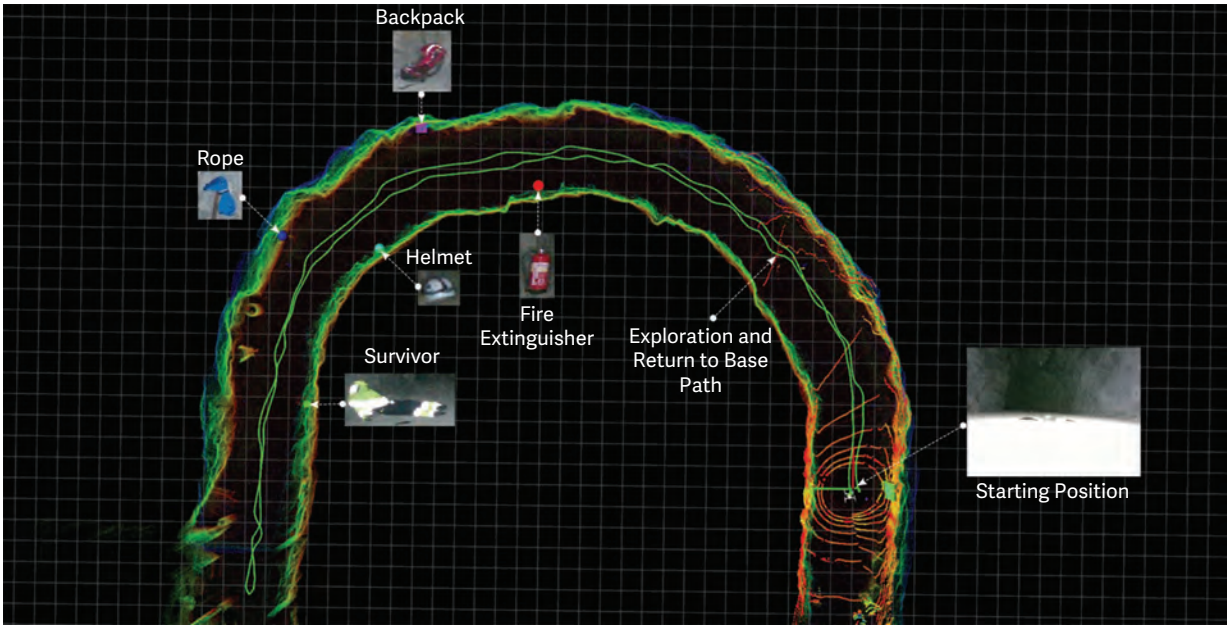

FIGURE 3. A visualization of a SubT search-and-rescue mission with an MAV, including artifact detection.

Distributed CPS-based autonomy for hypermodality robots

The SubT voids could consist of unknown extreme terrain features, such as rocky, granular, or sandy terrains; flat or high-slope areas; and so on. Thus, future mission planners will provide high-level tools to enable individual agents of the hypermodal platform to transverse such extreme terrains while considering the kinematic constraints of each individual component.

In this research direction, autonomous mission planners will take immediate actions in unforeseen contingencies, requiring a response in real time for the successful execution of the mission. The overall framework should be adaptive to the number of agents/modalities as well as seamlessly scalable, and it will be used in different layers of information sharing as well as for multiple sets of unified robots while considering different peer-to-peer communication configurations.

Even though the present article mostly focuses on planetary SubT exploration, such technological progress in the field of autonomy will be useful for other applications since it has the capability to empower other aspects of space advancements, such as on-orbit satellite operations and the related Made in Space movement[8] of the global space community.

Next-generation robotic space exploration demands more flexible, resilient, and self-sustaining high-level autonomy frameworks to address the challenges raised by classical ground loop control mission architectures, which are not feasible due to long-range communication delays or total communication losses. The future evolution of next-generation autonomous systems will be an important step in overcoming many of the current technological limitations in our never-ending journey toward the exploration of other worlds for evidence of life and to boldly go where no human has gone before.

REFERENCES

1. J. A. Starek et al., "Spacecraft autonomy challenges for next-generation space missions," in *Advances in Control System Technology for Aerospace Applications*, E. Feron, Ed. Berlin: Springer-Verlag, 2016, pp. 1–48. doi: 10.1007/978-3-662-47694-9_1.
2. "*The Artemis plan: NASA's lunar exploration program overview*," National Aeronautics and Space Administration, 2020.
3. V. Stamenković et al., "The next frontier for planetary and human exploration," *Nature Astron.*, vol. 3, no. 2, pp. 116–120, 2019. doi: 10.1038/s41550-018-0676-9.
4. "*The CoSTAR Team*," NASA. https://costar.jpl.nasa.gov/
5. "*Subterranean challenge*," DARPA. https://www

.subtchallenge.com/
6. A. Agha et al., "*Nebula: Quest for robotic autonomy in challenging environments; team COSTAR at the DARPA subterranean challenge*," Field Robot., 2021.
7. B. Lindqvist, A.-a. Agha-mohammadi, G. Nikolakopoulos "*Exploration-RRT: A multi-objective path planning and exploration framework for unknown and unstructured environments*," 2021, arXiv:2104.03724.
8. S. Patane, E. R. Joyce, M. P. Snyder, and P. Shestople "Archinaut: In-space manufacturing and assembly for next-generation space habitats," in *Proc. AIAA SPACE Astronaut. Forum Expo.*, 2017. doi: 10.2514/6.2017-5227.

GEORGE NIKOLAKOPOULOS is a chair professor in robotics and artificial intelligence and head of the robotics team at the Department of Computer, Electrical, and Space Engineering of Luleå University of Technology, Luleå, SE-97187, Sweden. Contact him at geonik@ltu.se.

ALI AGHA is the group leader for NASA Jet Propulsion Lab's (JPL's) Aerial Mobility Group and a technologist at JPL and Caltech's Center for Autonomous Systems and Technologies. Contact him at aliagha@jpl.nasa.gov.

IEEE COMPUTER SOCIETY

Publications Seek 2024 Editors in Chief

Application Deadline: 1 March 2023

The IEEE Computer Society (IEEE CS) seeks applicants for the position of editor in chief (EIC) for the following publications:

- *IEEE Internet Computing* magazine
- *IEEE Micro* magazine
- *IEEE Software* magazine
- *IEEE Transactions on Computers*
- *IEEE Transactions on Privacy* (new!)

EIC terms begin 1 January 2024. Candidates for any IEEE CS EIC position must be IEEE members in good standing with clear employer/institutional support and should possess a good understanding of all aspects of the publication's field. EICs must be able to attract a diverse group of talented and respected experts to their editorial board. Candidates must demonstrate the managerial skills necessary to process manuscripts through the editorial cycle in a timely fashion. Candidates with significant prior publications experience are preferred.

For more information, please go to computer.org/press-room/2022-news/ieee-computer-society-publications-seek-applications-for-2024-editors-in-chief.

Submit your application today!

IEEE COMPUTER SOCIETY

IEEE

EDITORS: Beatriz Sousa Santos, bss@ua.pt
Ginger Alford, alfordg@smu.edu

This article originally appeared in
IEEE Computer Graphics and Applications
vol. 41, no. 1, 2021

DEPARTMENT: EDUCATION

A Recipe of Capabilities for Pursuing Expertise in Data Visualization: A Practitioner's Perspective

Andy Kirk, *Visualising Data, Ltd., Leeds, U.K.*

Data visualization is hard to master because of the inherent complexities that characterize the challenge of facilitating understanding. Competence with data visualization is gaining in recognition as an essential capability and thus fostering the necessary skills is paramount to prepare students for their future professional activity in this field; yet, it is a challenge for educators to design programs that cover all facets. This article presents a framework that profiles the range of different capability "ingredients" which form the recipe of expertise in data visualization, from the point of view of an experienced practitioner.

During the COVID-19 global outbreak, the visual communication of data has been shown to be vital for governments, health professionals, and the media to inform a concerned public. Through broadcasts on TV, coverage on the web, opinions shared on social media, or articles from the traditional print media, one cannot have missed the plethora of charts and graphics used to explain and inform.

It is an enormous challenge to facilitate such understanding to the diverse audiences associated with this unprecedented context. There are wide variations in levels of subject knowledge, as well as data literacy. There are those who are receptive, those who are ambivalent, and those who are fundamentally skeptical. The data used is technical, with nuances in the basis of quantities and how categories are determined. The information presented is uncertain and, at times, incomplete. Sometimes it is about presenting headlines; other times it is more about the details. Sometimes it needs to be statistical and authoritative; sometimes it must be human and emotive.

Data visualization is hard to do well.[1] This is not because it is especially complicated, rather it is because of the complexities that characterize this fundamentally philosophical activity, concerned with the exchange of understanding. There are lots of small decisions involved with making selections from an array of possible options. The choices one makes are rarely objective, instead reliant on reasonable—and occasionally rapid—subjectivity. Each decision made has a consequence on shaping the next decision.

It is perhaps unsurprising therefore, that competence with data visualization is gaining in recognition as an essential capability. Just as computer literacy is seen as a necessity, expected of everyone to be able to fully participate in the modern workplace, in education, and in society, the demand for literacy in communicating data follows a similar, if belated, trajectory. As such, fostering the necessary skills is paramount to prepare students for their future professional activity in this field; yet, it may be challenging for educators to facilitate. The intention of this article is to present a framework that profiles the range of different capability "ingredients" which collectively form the recipe of expertise in data visualization.

SEVEN HATS OF DATA VISUALIZATION

The *Seven Hats of Data Visualization* offers a breakdown of different aptitudes, blending together elements of knowledge, skills, behaviors, and habits. This framework has been shaped from the author's

Digital Object Identifier 10.1109/MCG.2020.3034737
Date of current version 12 January 2021.

 Published by the IEEE Computer Society

experiences working in the data visualization field for over a decade helping individuals and organizations attain this capability through commercial and academic teaching, as well as consultancy activities. The seven hats attempt to reflect the reality of the diverse technical, analytical, editorial, and creative dimensions which shape this field. For instance, with almost all visualization work being computer-generated, skills in graphics applications and programming libraries are hugely desirable, yet they represent just one aspect of what it takes to form fully accomplished data visualization expertise. Not all detailed ingredients presented are of equal weighting and they won't always be necessarily relevant for all visualization tasks or projects undertaken.

THE DIRECTOR: The coordinator, overseeing the project.

- Initiates and leads on gathering and understanding requirements.
- Identifies and establishes the project's contextual freedoms and constraints.
- Defines the purpose of the project based on the desired outcome.
- Manages progress through the workflow process.
- The primary decision-maker, often needing to judge tradeoffs.
- Pays strong attention to detail.
- Gets things done: Checks, tests, finishes tasks.

THE COMMUNICATOR: The broker, concerned about people.

- Helps to define and represent the perspective of the audience.
- A good listener with the humility to defer to subject matter experts.
- Has a "thick skin:" needs patience, empathy and diplomacy.
- A confident communicator with specialists and nonspecialists.
- Possesses strong copy-editing abilities with written communications.
- Manages expectations and presents possibilities.
- Launches and promotes the final solution.

THE JOURNALIST: The reporter, pursuing the enquiry.

- Driven by a desire to help others achieve heightened understanding.
- Generates curiosities that fuel the process.
- Has an instinct for research and discovery.
- Is a quick learner with the capacity to develop knowledge in new subjects.
- Is able to identify what content is most relevant for the audience.
- Defines the editorial decision about what data to include and exclude.
- Defines the editorial decision about what data to emphasize and de-emphasize.

THE ANALYST: The wrangler, handling data work.

- Has strong data and statistical literacy.
- Possesses technical skills to obtain data from multiple sources.
- Demonstrates robust statistical and data ethical standards.
- Capable of efficiently transforming and preparing data for its purpose.
- Is able to conduct advanced exploratory data analysis.
- Experienced with data modeling techniques.

THE SCIENTIST: The thinker, providing scientific rigor.

- Brings a strong research mindset to the process.
- Understands the science of visual perception.
- Understands the ethics of visual commun-ication.
- Understands the influence of human factors.
- Verifies/validates the integrity of all data and design choices.
- Has good judgment to know when to follow rules and when to bend rules.
- Undertakes reflective evaluation and critique.

THE DESIGNER: The conceiver, driving creative direction.

- Establishes a coherent creative strategy.
- Harnesses ideas and inspiration from other work and subject areas.
- Has a flair for sketching and illustration.
- Understands the principles of user interface design.
- Is fluent with the full array of possible visualization design options.
- Demonstrates thoroughness and persistence for refining.
- Has a relentless drive to keep experimenting and innovating.

THE TECHNOLOGIST: The developer, constructing the solution.

- Possesses a repertoire of software and programming capabilities.
- Has an appetite to acquire new technical capabilities.
- Has a diligence and discipline to focused on delivering a solution.
- Can automate otherwise manually intensive processes.
- Has the discerning eye to avoid feature creep.
- Works on the prototyping and development of the solution.
- Undertakes pre- and postlaunch testing, evaluation, and support.

ASSESSING CAPABILITIES

This framework should be seen as an ideal potential set of aptitudes one needs to attain or demonstrate in order to accomplish complete expertise.

Its purpose is primarily to operate as a checklist of the capabilities required, to help individuals assess themselves and identify where their strengths and weaknesses lie. On another level it can also be seen as the repertoire of multidisciplinary capabilities demonstrated across a team or collaborative group to ensure equity and balance in the blending together of talents. For organizations looking to hire or develop talent, or construct teams, this framework helps to isolate the skills they seek in people. It may also help them to more surgically identify areas for training and development of staff. For university departments designing programs in data visualization, or related fields like information design, it offers a broad view of the range of teaching branches they need to consider offering.

When starting a journey learning data visualization, the prospect can be overwhelming - "where to begin!?"—and, for some, quite intimidating—"how will I ever achieve that kind of quality!?" One needs to start somewhere and that should always be from the position of recognizing the existing strengths one may possess.

One of the primary motivations for constructing this framework was to dispel any notion of this subject being reserved for uniquely talented individuals alone.

A flair for graphic design, computer science, or journalistic instinct is inevitably advantageous, but the more important starting point is to recognize that data visualization is fundamentally human-centered. Data is about people or phenomena that affect people. Visualizations are generated by people and are for people. The most important skills center around this theme, such as listening to other peoples' requirements, anticipating the curiosities that will be most relevant amongst an audience, caring about the language we use to explain insights.

Judging the effectiveness of a visualization is occasionally tangible, but more often than not it is intangible. Rarely is there a directly measurable nor witnessable moment that can be defined as evidence of success.

There are, of course, some relatively free quantitative measures that are available for digital projects, including web-based measures such as visitor counts and social media metrics (likes, retweets, mentions). These, at least, provide a surface indicator of success in terms of the project's apparent appeal and spread.

If the intention behind a visualization is to inform people, to further the debate about a subject, or to establish one's reputation or voice of authority, then those are hard outcomes to measure. One approach can be to invert the measure of effectiveness on and seek out evidence of tangible ineffectiveness. For example, there may be significant reputation-based impacts should decisions be made on inaccurate, misleading, or inaccessible visual information. One should aspire also to collect reliable qualitative feedback, even if this can, at times, be rather expensive, in effort terms, to secure. Some options include:

- capturing anecdotal evidence from comments submitted on a site, opinions attributed to tweets or other social media descriptors, feedback shared in emails or in person;
- informal feedback through polls or short surveys;
- formal case studies which might offer more structured interviews and observations about documented effects;
- experiments with controlled tasks/conditions and tracked performance measures.[3]

A personal assessment of one's contribution to a project is important for development. The best way to learn is by considering the aspects that were enjoyable and/or successful and identifying the things were unenjoyable and/or less successful. Look back over each project experience and consider the following.

- Was the solution satisfactory? If yes, why; if no, why and what could have been done differently?
- In a different context, what other design solutions might have been considered?
- Were there any skill or knowledge shortcomings that restricted progress and/or solution?

FIGURE 1. Episodes from the first season of "Explore Explain," a podcast and video series based on conversations with visualizers exploring the design story behind a visualization or a series of related works (https://www.visualisingdata.com /podcast/).

FIGURE 2. Sample pages extracted from "The Seinfeld Chronicles," a visualization project published in 2020, designed by Andy Kirk.[2]

- Are there aspects of this project that could be reused or adapted for use in other projects? For instance, ideas that did not make the final cut but could be given new life in other challenges?
- How well was time utilized? Were there any activities that felt inefficiently managed, either through being overly drawn out or uncomfortably rushed?

DEVELOPING CAPABILITIES

The challenge for educators is clearly not limited to just designing a broad curriculum to cover the necessary topics. This needs to be bolstered by sophisticated activities and innovative experiences that offer students the opportunity to encounter a portfolio of challenges. There are many learning strategies to consider.

Self-Directed Learning. The critical discourse and creative energy of data visualization is inevitably found online. Frequent exposure to the dedicated visualization blogs, the websites of news media, portfolios of studios, and creative agencies will ensure awareness of the latest techniques, case studies, and process narratives. Constantly challenge one's design "eye" by evaluating the work of others ("what would I do differently?") is an especially valuable exercise. Listen to podcasts and sign-up for webinars to immerse yourself in contemporary topics and latest discourse.

Practice, practice, practice. The journey from good to great, as with anything, involves hard work, plenty of learning, lots of mistakes, but, most importantly, a relentless appetite to gain experience. Identifying personal "passion" projects offer the chance to experiment without consequence. These are great opportunities to try out new techniques and deliberately expose oneself to resolving different challenges than might otherwise be encountered. Learning reflectively about the ideas, problems, solutions, and rationale behind all decisions will optimize the enduring value gained.

Collaborating With Others: There are advantages to pursuing data visualization solutions collaboratively, bringing together different talents and perspectives to a shared challenge. The best functioning visualization team will offer a collective blend of competencies across all seven hats. Crucially, when setting up group activities, it is important to avoid skewing the sensibilities towards one dominant talent. Success will be hard to achieve if a team comprises a dominance in technologists or a concentration of "ideas" people whose work never progresses past the sketchbook. The right blend is required for any team.

Networking: On social media (especially Twitter and LinkedIn) we see an active and welcoming community of practitioners that is always willing to share and help. Attending conferences and local meetups will offer direct learning opportunities as well as communities to of contributors and enthusiasts to build networks.

Looking beyond: Exposing oneself to the practices of other related creative and communication fields can be enlightening. Fields like video game design, graphic design, architecture, and even cartoon illustration are fields with unique contexts: the need for video games to be sophisticated but fun and accessible; for graphic design to grab attention and convey identity; for architecture to balance utility with beauty within the constraints of costs and materials; for cartoons to be humorous and topical but also self-explanatory within a single frame.

CONCLUSION

This article has proposed a practitioner's view of the range of different capabilities that shape the contemporary visualizer. The Seven Hats, as profiled, offer educators and students alike a potential menu of the necessary learning topics and practical experiences required to flourish in this popular contemporary subject area. This may be useful, for instance, in determining the curricular for teaching content, as well as helping to shape the assessment activities for individual and group course assignments, as well as final visualization projects.

Developing mastery across the full range of capabilities presented is probably unachievable, such are the multidisciplinary demands: even amongst the most celebrated visualizers in the field today, few will genuinely possess expertise in all areas. But it offers a guide to help assess existing levels of competence and provides the shape of a wish list for development in areas of recognized shortcoming that can inform course development and project experience beneficial to students' wholistic development.

Developing expertise in data visualization will take time. The people who flourish in this field are able to harness their strengths and address their weaknesses. It requires motivation for ongoing efforts to learn, apply, reflect, and repeat again.

REFERENCES

1. T. Munzner, *Visualization Analysis and Design*. Natick, MA, USA: AK Peters, 2014.
2. A. Kirk, *Data Visualization: A Handbook for Data Driven Design* 2nd ed. Thousand Oaks, CA, USA: Sage, 2019.
3. S. Carpendale, "Evaluating information visualizations," in *Information Visualization, Human-Centered Issues and Perspectives*, A. Karren, *et al.* Eds. Berlin, Germany: Springer, 2008, pp. 19–45.

ANDY KIRK is the founder of Visualising Data, Ltd., Leeds, U.K. He is a data visualization design consultant, educator, and author. Contact him at andy@visualisingdata.com.

Contact department editor Beatriz Sousa Santos at bss@ua.pt or department editor Ginger Alford at alfordg@smu.edu.

EDITORS: Sharon Broude Geva, sgeva@umich.edu
Dirk Colbry, colbrydi@msu.edu

This article originally appeared in
Computing in Science & Engineering
vol. 23, no. 5, 2021

DEPARTMENT: EDUCATION

Building a Culture of Computing in the Sciences Using Images as Data Within a Community of Practice

Tessa Durham Brooks, *Doane University, Crete, NE, 68333, USA*
Raychelle Burks, *American University, Washington, DC, 20016, USA*
Mark Meysenburg, *Erin Doyle , and Chris Huber, Doane University, Crete, NE, 68333, USA*

The digital imaging and vision applications in science (DIVAS) program was built to improve the computational self-efficacy and skill of first- and second-year college students majoring in biological and chemical sciences. Our three-year pilot study showed that the program could be successful in both fronts. The scholars, faculty, and staff who participated formed a community of practice that became the heart of the DIVAS program. Through this community, we expanded access to the image processing workshop in collaboration with The Carpentries, supported faculty and secondary educators in developing computing modules for their classrooms, and created and staffed a "writing center for computing" on the host campus. Overall, the DIVAS program has sparked a local computing culture. DIVAS interventions and resources are freely available for adoption by other institutions. We hope to grow the community in a way that builds student access and opportunities and supports educators in the process.

Datasets in natural, physical, and computer sciences are now so massive that their analysis, interpretation, and visualization often cannot keep up with the rate at which they can be routinely produced.[1] This data explosion means scientists across disciplines must work together to build tools, models, and visualizations capable of exploring that massive, complex, multifaceted landscape. This cooperation will lead to the development of new tools that will allow us to observe aspects of our universe that have previously been hidden or beyond our reach. These new tools have the potential to transform the foundations of how we understand our universe analogous to the development of the microscope centuries ago.[2]

Education and training within the natural and physical sciences do not sufficiently equip students with the necessary experience and skills to meet emerging data processing and analysis needs and experts across scientific disciplines are raising this issue.[3–6] To this end, we created the Digital Imaging and Vision Analysis in Science (DIVAS) project.

DIVAS secured funds in 2016 from the National Science Foundation's Improving Undergraduate STEM Education division.[7] The program aims to engage novice learners (mostly first- and second-year college students) from biological and chemical sciences in computational thinking and coding using image data as a "hook." Images are widely used for diagnostics, phenotyping, and analytical measurements. They are also easy to obtain and novice learners understand them. When students analyze image data they engage in all of the aspects of computational thinking including recognition of a problem, analysis of solutions, design of a solution, and implementation and testing.

Of note, we saw the computational confidence and skills improve in students not normally well represented in computer science courses. Our pilot study

1521-9615 © 2021 IEEE
Digital Object Identifier 10.1109/MCSE.2021.3108617
Date of current version 23 September 2021.

FIGURE 1. Erin Doyle presents the challenge of tracking the progress of a titration to coding workshop participants.

indicated that the scaffolding of interventions that make up the program and the context of a community of practice were important elements contributing to the gains we observed.[8] The pilot study motivated us to see how broadly these results can be replicated. In this report, we discuss the key elements of the DIVAS project. We will also discuss what we have learned about fostering self-efficacy in computing more broadly across the natural and physical sciences.

DIVAS PROJECT OVERVIEW

The DIVAS program targets first-year students enrolled in introductory biology and chemistry courses. The program name, its leaders, and the recruiting strategies used resulted in a high percentage of individuals who identify as women joining the program (76% of all program scholars).[8] The program begins in the spring with a one-credit seminar and continues five to eight weeks into the following summer. The program ends with a capstone seminar the following fall or spring. From the first DIVAS seminar, students join a community of practice that will support them throughout the rest of the program. In that seminar, scholars learn about images as a form of data; do basic coding; meet professionals who use coding in their work; and explore working environments where coding occurs.

The summer begins with an intensive, one-week coding workshop, which is also open to community members. The workshop devotes two days on basic Python operations, code version control using git, and bash shell functions. The following three days focus on image processing using Scikit-image libraries. We seat participants in pairs to encourage them to cooperate in their work. To increase engagement, participants are presented with a problem at the beginning of each portion of the workshop. By the end of the workshop, everyone has written scripts to count the number of colonies on a plate and to track the progression of a titration over time using a handful of basic image processing functions (see Figure 1).

Following the coding workshop, there are two two-week sessions of pair programming. We give the group a common challenge at the beginning of each session. Scholars discuss the problem, then together decide on the group's goals for the session. We teach scholars how to do pair programming, then divide them up to devise, and implement their own strategy for solving the problem. Each day, the group holds a stand-up meeting to report on progress and discuss daily goals. At the end of each week, scholars push their code to a common repository and the teams print out and annotate each other's code for review. In the first week, code reviews help teams set their goals for the following week. In the second week, reviews help them wrap up their code. In the second two-week session, we reassign pairs and repeat the process.

The program ends with another one-credit seminar in which scholars revisit the DIVAS repository to clean up and annotate code, and make adjustments to the coding workshops as needed based on feedback. They also meet with incoming scholars to welcome them to the community and offer support. Scholars also learn about parallelization and gain experience with parallel programming. Parallel programming is relatively straightforward to apply with the image processing scripts they have already written.

PILOT STUDY OUTCOMES

The DIVAS pipeline was tested on three cohorts of up to six scholars over three years. Seventeen scholars participated, 14 of which attended Doane University, a private liberal arts college in Nebraska. The other three scholars came from our partner campus at St. Edward's University, a private liberal arts college as well as a minority-serving institution, in Austin, TX, USA. Most scholars identified as women (76%) and were in their first year of college (82%) majoring in biological or chemical fields (82%). Overall, we saw self-efficacy in computing increase by 34% on average as well as statistically significant growth in all the computational thinking skills we measured (recognize the problem, analyze solutions, design a solution, implement a solution).[8] Self-efficacy grew even while interest in pursuing careers using computational skills did not and coding tasks became more challenging. Our previous publication includes details of the pilot study,

including resources and information about each of the DIVAS program interventions.[8]

DISSEMINATION OF THE IMAGE PROCESSING WORKSHOP

Image processing has become routine in studies of atomic, molecular, and cellular dynamics, those that associate genomic elements with phenotypes of interest, in breeding programs, and in a variety of monitoring and modeling in fields such as agriculture, ecology, and drug development. This increased demand within the scientific community for image processing skills has led us to turn the image processing elements of our workshop into a Data Carpentry lesson.[9,10] Data Carpentry supports community-driven development of domain-specific lessons to meet research training needs.

The lesson is still in the early adoption process. The workshop originally used OpenCV libraries. Community members converted it to using Scikit image libraries, which are much easier to implement across a range of platforms and environments. The lesson has been tested at three research institutions in the United States and Germany. The lesson assumes basic knowledge of Python, git, and bash and covers the basics of image processing, including image representation; creating histograms; blurring and thresholding; drawing and masking; edge detection; and object segmentation using connected components analysis. The two challenges that DIVAS scholars work on in this portion of the workshop are also available in this lesson. DIVAS project investigator Mark Meysenburg currently maintains the Data Carpentry lesson. As others in the Carpentries community use it and additional needs are identified we hope to see this lesson adapt to meet those needs.

COMMUNITY BUILDING

The Computing Center for the Liberal Arts: An important consequence of the pilot study was the creation of a broader community of students with computing skills on the Doane University campus. No longer siloed into specific departments and programs, students who once may not have interacted with each other academically were now connected through common interests and skills.

The DIVAS team recognized that scholars were broadening their community of practice to include peers who needed to build their own computational skill, as well as peers with more expert knowledge that could provide support. To help build this community further, the DIVAS team created a "writing center for computing" at Doane called the Computing Center for the Liberal Arts (CCLA). The CCLA is a place for anyone within the Doane community to get feedback and help with any computing project, from setting up an Excel spreadsheet to research using Doane's supercomputer, Onyx.

Using funding from the National Science Foundation, the team hired a center director whose duties included packaging CCLA program elements for broad implementation in a "computing center in a box".[11] Like writing centers across the nation, the CCLA is fueled by peer-to-peer support. Peer instructors develop short tutorials and guides on common computing needs, provide support via a Slack workspace and email address, and provide face-to-face support in the learning commons. The CCLA is now in its third year. A quarter of DIVAS scholars have served as peer consultants and one former scholar is the current training manager for 11 new consultants. Individuals majoring in six different disciplines have utilized its services. The center is growing steadily as the campus community becomes more familiar with its goal. The CCLA team has reached out to the humanities and social sciences departments to support their computing needs as well.

Introductory Biology and Chemistry Courses: The DIVAS team has brought computational thinking through image analysis into introductory biology and chemistry courses in two ways. First, we designed an image processing module for general chemistry students to investigate the hydrophobicity of materials by measuring the contact angle of a drop of water upon them. Students take pictures of the water droplets using their smartphones and then use ImageJ, an open-source and widely used image analysis and processing tool,[12] to manually measure the angle of incidence from each image. This module has now been used for over five years.

Second, in an inquiry-based introductory biology course, mostly first-year students used a Google Colab notebook written by former DIVAS scholars to analyze images from a system to measure bacterial movement toward molecules (chemotaxis). A drop of agarose containing a test molecule or saline (control) is added to the center of each well in a six-well plate. The solid agarose droplet is then surrounded by stained *E. coli* in saline. Images are taken over time using flatbed scanners. Students first use ImageJ to develop their own strategies for measuring the change in cloudiness around the agarose plug over time (an indication that chemotaxis has occurred). Completely novice coders then review the code in the Colab notebook by drawing out a

flowchart of what they see the code doing. When students were unsure of how the code was functioning, they tried changing part of it to see what effect it had. As a group, the class added clarifying annotations and create a full code map. From there, students used this code and modified parameters as needed to measure the change in cloudiness around the agarose plugs in their own well plates. This intervention is now in its second year.

Teacher Training Workshop: Finally, Nebraska high school teachers used another Google Colab notebook written by former DIVAS scholars last summer. Scholars wrote this code to measure the height and density of invasive grass growing in pots. The grass system provides a variety of angles to engage students in inquiry-based learning and to learn about asking questions, designing experiments, and analyzing and representing data. Teachers explored the pros and cons of manual measurements. They then studied the image-based approach to taking similar measurements by creating their own code maps as they worked through the Colab notebook. As teachers worked through the code, they identified other ways to take the same measurements and how to test the code to confirm it was functioning. Most of these teachers had never coded. Nevertheless, within a couple of hours, they learned to use the notebook. Further, they could evaluate it for use in the classroom both as a measurement tool and as a source of inquiry-based curriculum.

BROADENING THE ALLIANCE

Over the three years of the pilot study, we have learned a lot about how to support self-efficacy in computing early in a student's college career. The pilot data and anecdotal experiences alone might inspire other schools to try similar approaches. However, we still need to test the DIVAS program interventions across institutional types and different student populations. Broader implementation of all or part of DIVAS program elements will also help us identify its most critical parts. Teams wanting to implement a more streamlined program could implement just those elements with success. A broader DIVAS alliance will provide additional opportunities for students to collaborate, build their skills, and strengthen the community of practice. We also hope to expand the teacher training workshops, including both secondary and undergraduate educators. These workshops would use the training of trainers (ToT) model, empowering educators to integrate DIVAS interventions into their classrooms and research labs.

DISCUSSION

We have long needed to find ways to infuse computational thinking, coding, and the use of scientific software into natural and physical science undergraduate education. Our experience with the DIVAS project, our pilot study, and the additional opportunities it fostered suggest that this can be done in many environments friendly to a community of practice approach. We have seen how a community can change the way students view computing from a specialized, esoteric skill to a set of tools anyone can learn to use. We have seen that novice learners can learn to use computational tools to solve problems relevant to their disciplines, gaining confidence in computational skills, and highly desirable workforce skills. As DIVAS program elements are adopted at other institutions, we will see this impact more clearly.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (NSF) Improving Undergraduate STEM Education under Grant 1608754, the CyberTraining (1974094) and EPSCoR Research Infrastructure Improvement (RII) Track-1 (1557417) programs.

REFERENCES

1. A. Labrinidis and H. V. Jagadish, "Challenges and opportunities with big data," *Proc. VLDB Endowment*, vol. 5, no. 12, pp. 2032–2033, 2012, doi: 10.14778/2367502.2367572.
2. L. Poppick, "Let us now praise the invention of the microscope," *Smithsonian Magazine*, 2017. [Online]. Available: https://www.smithsonianmag.com/science-nature/what-we-owe-to-the-invention-microscope-180962725/
3. F. Lau, L. Katona, J. M. Rosen, and C. E. Koop, "Computer science: The third pillar of medical education," *Creative Educ.*, vol. 3, no. 6, pp. 807–810, 2012, doi: 10.4236/ce.2012.326120.
4. Y. Lu, G. Deng, and Z. Shuai, "Future directions of chemical theory and computation," *Pure Appl. Chem.*, 2021, doi: 10.1515/pac-2020-1006.
5. B. Skuse, "The third pillar," *Phys. World*, vol. 32, no. 3, pp. 40–43, 2019, doi: 10.1088/2058-7058/32/3/33.
6. P. Pevzner and R. Shamir, "Computing has changed biology-biology education must catch up," *Science*, vol. 325, no. 5940, pp. 541–542, 2009, doi: 10.1126/science.1173876.

7. T. D. Brooks, R. Burks, M. Meysenburg, and S. Dworak, "Award abstract # 1608754, Doane DIVAS: Digital Imaging and Vision Applications in Science." Accessed: 2020. [Online]. Available: https://www.nsf.gov/awardsearch/showAward?AWD_ID=1608754&HistoricalAwards=false/
8. T. D. Brooks, R. Burks, E. Doyle, M. Meysenburg, and T. Frey, "Digital imaging and vision analysis in science project improves the self-efficacy and skill of undergraduate students in computational work," *PLOS One*, vol. 16, no. 5, 2021, Art. no. e0241946, doi: 10.1371/journal.pone.0241946.
9. Data Carpentry. Accessed: 2021. [Online]. Available: https://datacarpentry.org
10. "Image processing with Python," The Carpentries. Accessed: 2021. [Online]. Available: https://datacarpentry.org/image-processing/
11. M. Meysenburg and M. Carpenter, "Award Abstract # 1974094, CyberTraining: Pilot: Institutional Cyberinfrastructure Training Center in a Box." Accessed: 2021. [Online]. Available: https://nsf.gov/awardsearch/showAward?AWD_ID=1924094&HistoricalAwards=false/
12. W.S. Rasband, "ImageJ," U.S. National Institutes of Health, Bethesda, MD, USA, 1997–2018. [Online]. Available: https://imagej.nih.gov/ij/

TESSA DURHAM BROOKS is an Associate Professor of biology at Doane University, Crete, NE, USA. She teaches courses in introductory biology, physiology, and a third-year seminar on making meaning. Her teaching focus is to promote a sense of belonging in order to enhance learning in inquiry-based and flipped classroom environments. She and her undergraduate team explore phenotypic responses of plants to environmental stimuli and the effects of stress in early development on later growth phenotypes. She has an interest in developing infrastructure and academic experiences that promote computational and quantitative self-efficacy of undergraduate students in the natural sciences. Dr. Brooks received the bachelor's degree in biochemistry from the University of Nebraska and the Ph.D. degree in cell and molecular biology from the University of Wisconsin. Contact her at tessa.durhambrooks@doane.edu.

RAYCHELLE BURKS is an Associate Professor of chemistry at American University, Washington, DC, USA. Her lab research team is focused on the development of colorimetric and luminescent sensing systems with integrated image and chemometric analysis for forensic applications. She is on the leadership team of the Digital Imaging and Vision Applications in Science (DIVAS) Project, and DIVAS Scholars research advisor. Beyond the bench, she is a popular science communicator appearing regularly on TV, radio, podcasts, and print outlets. Central to her research, teaching, and service is the central tenet that equitable, diverse, and inclusive practices both respect people and produce scientific outcomes of greater integrity. She is a member of several local, national, and international committees, task forces, and projects focused on social justice and STEM. Contact her at burks@american.edu.

MARK MEYSENBURG is a Professor of computing at Doane University, Crete, NE, USA. He teaches the university's programming sequence, networking, and cybersecurity courses. His research interests are varied and eclectic, including evolutionary computation, machine learning, robotics, and computer vision. He also teaches an annual first-year seminar course, using intense role-playing games to teach the history of science. Contact him at mark.meysenburg@doane.edu.

ERIN DOYLE is an Associate Professor of biology at Doane University, Crete, NE, USA. Her teaching focuses on helping students understand the importance of mathematics, computer science, and statistics to modern biology, and supporting positive student experiences in these areas. She teaches courses in introductory biology as well as upper-level electives in genetics, bioinformatics, and computational biology. Students in her undergraduate research lab use computational approaches to generate experimentally testable hypotheses and use experimental results to develop and refine computational models of biological processes. Recent projects in her lab have focused on the identification of disease susceptibility genes in rice plants and functional characterization of bacteriophage genes. Dr. Doyle received the bachelor's degree in applied mathematics from the University of Tulsa and the Ph.D. degree in bioinformatics and computational biology from Iowa State University. Contact her at erin.doyle@doane.edu.

CHRIS HUBER is an Assistant Professor of chemistry at Doane University, Crete, NE, USA. He teaches courses in general chemistry, analytical chemistry, and physical chemistry. He centers his teaching style on delivering course material using an array of pedagogical strategies to match students' different learning styles. His scholarly work centers on developing sensitive spectroscopic sensors capable of detecting ultra-low (sub ppm) concentrations of blood toxins. He received the bachelor's degree in chemistry from the University of Wisconsin-La Crosse, and the M.S. and Ph.D. degrees in chemistry from the University of Minnesota. Contact him at chris.huber@doane.edu.

IEEE

COMPUTER ARCHITECTURE LETTERS

***IEEE Computer Architecture Letters* is a forum for fast publication of new, high-quality ideas in the form of short, critically refereed technical papers.** Submissions are accepted on a continuing basis and letters will be published shortly after acceptance in IEEE Xplore and in the Computer Society Digital Library.

Submissions are welcomed on any topic in computer architecture, especially:

- Microprocessor and multiprocessor systems
- Microarchitecture and ILP processors
- Workload characterization
- Performance evaluation and simulation techniques
- Interactions with compilers and operating systems
- Interconnection network architectures
- Memory and cache systems
- Power and thermal issues at the architectural level
- I/O architectures and techniques
- Independent validation of previously published results
- Analysis of unsuccessful techniques
- Domain-specific processor architecture (embedded, graphics, network)
- High-availability architectures
- Reconfigurable computer architectures

www.computer.org/cal

Join the IEEE Computer Society for subscription discounts today!

www.computer.org/product/journals/cal

IEEE COMPUTER SOCIETY

IEEE

IEEE

SECURITY & PRIVACY

IEEE Security & Privacy **is a bimonthly magazine communicating advances in security, privacy, and dependability** in a way that is useful to a broad section of the professional community.

The magazine provides articles with both a practical and research bent by the top thinkers in the field of security and privacy, along with case studies, surveys, tutorials, columns, and in-depth interviews. Topics include:

- Internet, software, hardware, and systems security
- Legal and ethical issues and privacy concerns
- Privacy-enhancing technologies
- Data analytics for security and privacy
- Usable security
- Integrated security design methods
- Security of critical infrastructures
- Pedagogical and curricular issues in security education
- Security issues in wireless and mobile networks
- Real-world cryptography
- Emerging technologies, operational resilience, and edge computing
- Cybercrime and forensics, and much more

www.computer.org/security

Join the IEEE Computer Society for subscription discounts today!

www.computer.org/product/magazines/security-and-privacy

IEEE COMPUTER SOCIETY

IEEE

Keep up with the latest IEEE Computer Society publications and activities wherever you are.

Follow us:

@ComputerSociety

facebook.com/IEEEComputerSociety

IEEE Computer Society

youtube.com/ieeecomputersociety

instagram.com/ieee_computer_society

IEEE COMPUTER SOCIETY

IEEE

IEEE TRANSACTIONS ON

COMPUTERS

Call for Papers: *IEEE Transactions on Computers*

Publish your work in the IEEE Computer Society's flagship journal, *IEEE Transactions on Computers* (*TC*). *TC* is a monthly publication with a wide distribution to researchers, industry professionals, and educators in the computing field.

TC seeks original research contributions on areas of current computing interest, including the following topics:

- Computer architecture
- Software systems
- Mobile and embedded systems
- Security and reliability
- Machine learning
- Quantum computing

All accepted manuscripts are automatically considered for the monthly featured paper and annual Best Paper Award.

Learn about calls for papers and submission details at
www.computer.org/tc.

IEEE COMPUTER SOCIETY

IEEE

SUBMIT TODAY

IEEE TRANSACTIONS ON SUSTAINABLE COMPUTING

SCOPE

The *IEEE Transactions on Sustainable Computing* (*T-SUSC*) is a peer-reviewed journal devoted to publishing high-quality papers that explore the different aspects of sustainable computing. The notion of sustainability is one of the core areas in computing today and can cover a wide range of problem domains and technologies ranging from software to hardware designs to application domains. Sustainability (e.g., energy efficiency, natural resources preservation, using multiple energy sources) is needed in computing devices and infrastructure and has grown to be a major limitation to usability and performance.

Contributions to *T-SUSC* must address sustainability problems in different computing and information processing environments and technologies, and at different levels of the computational process. These problems can be related to information processing, integration, utilization, aggregation, and generation. Solutions for these problems can call upon a wide range of algorithmic and computational frameworks, such as optimization, machine learning, dynamical systems, prediction and control, decision support systems, meta-heuristics, and game-theory to name a few.

T-SUSC covers pure research and applications within novel scope related to sustainable computing, such as computational devices, storage organization, data transfer, software and information processing, and efficient algorithmic information distribution/processing. Articles dealing with hardware/software implementations, new architectures, modeling and simulation, mathematical models and designs that target sustainable computing problems are encouraged.

SUBSCRIBE AND SUBMIT

For more information on paper submission, featured articles, calls for papers, and subscription links visit:

www.computer.org/tsusc

IEEE COMPUTER SOCIETY

IEEE ComSoc IEEE Communications Society

CEDA IEEE Council on Electronic Design Automation

IEEE

IEEE Internet Computing

IEEE Internet Computing delivers novel content from academic and industry experts on the latest developments and key trends in Internet technologies and applications.

Written by and for both users and developers, the bimonthly magazine covers a wide range of topics, including:

- Applications
- Architectures
- Big data analytics
- Cloud and edge computing
- Information management
- Middleware
- Security and privacy
- Standards
- And much more

In addition to peer-reviewed articles, *IEEE Internet Computing* features industry reports, surveys, tutorials, columns, and news.

www.computer.org/internet

Join the IEEE Computer Society for subscription discounts today!

www.computer.org/product/magazines/internet-computing

IEEE COMPUTER SOCIETY

IEEE

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.

FEBRUARY

1 February
- ICSC (IEEE Int'l Conf. on Semantic Computing), Laguna Hills, USA

8 February
- SaTML (IEEE Conf. on Secure and Trustworthy Machine Learning), San Francisco, USA

13 February
- BigComp (IEEE Int'l Conf. on Big Data and Smart Computing), Jeju, South Korea

20 February
- ICNC (Int'l Conf. on Computing, Networking and Communications), Honolulu, USA

25 February
- HPCA (IEEE Int'l Symposium on High-Performance Computer Architecture), Montreal, Canada

MARCH

13 March
- ICSA (IEEE Int'l Conf. on Software Architecture), L'Aquila, Italy
- PerCom (IEEE Int'l Conf. on Pervasive Computing and Communications), Atlanta, USA

15 March
- ISADS (IEEE Int'l Symposium on Autonomous Decentralized System), Mexico City, Mexico

17 March
- DMIST (Int'l Conf. on Digital Management, Information Systems and Technologies), Shenyang, China

19 March
- ICST (IEEE Conf. on Software Testing, Verification and Validation), Dublin, Ireland

21 March
- SANER (IEEE Int'l Conf. on Software Analysis, Evolution and Reengineering), Taipa, Macao

25 March
- VR (IEEE Conf. Virtual Reality and 3D User Interfaces), Shanghai, China

APRIL

3 April
- ICDE (IEEE Int'l Conf. on Data Eng.) Anaheim, USA

14 April
- IPEC (Asia-Pacific Conf. on Image Processing, Electronics and Computers) Dalian, China

17 April
- DATE (Design, Automation & Test in Europe Conf.) Antwerp, Belgium

18 April
- PacificVis (IEEE Pacific Visualization Symposium) Seoul, Korea (South)

19 April
- COOL CHIPS (IEEE Symposium in Low-Power and High-Speed Chips) Tokyo, Japan

23 April
- ISPASS (IEEE Int'l Symposium on Performance Analysis of Systems and Software) Raleigh, USA

24 April
- VTS (IEEE VLSI Test Symposium) San Diego, USA

MAY

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

7 May
- FCCM (IEEE Int'l Symposium on Field-Programmable Custom Computing Machines) Los Angeles, USA

14 May
- ICSE (IEEE/ACM Int'l Conference on Software Engineering) Melbourne, Australia

15 May
- IPDPS (IEEE Int'l Parallel and Distributed Processing Symposium) St. Petersburg, USA

17 May
- MOST (IEEE Int'l Conf. on Mobility, Operations, Services and Technologies) Detroit, USA

21 May

- SP (IEEE Symposium on Security and Privacy) San Francisco, USA

22 May

- ISMVL (IEEE Int'l Symposium on Multiple-Valued Logic) Matsue, Japan

29 May

- SERA (IEEE/ACIS Int'l Conf. on Software Engineering Research, Management and Applications) Orlando, USA

JUNE

7 June

- CAI (IEEE Conf. on Artificial Intelligence) Santa Clara, USA

10 June

- ICHI (IEEE Int'l Conf. on Healthcare Informatics) Houston, USA

12 June

- WoWMoM (IEEE Int'l Symposium on a World of Wireless, Mobile and Multimedia Networks) Boston, USA

17 June

- CVPR (IEEE/CVF Conf. on Computer Vision and Pattern Recognition) Virtual Conf.

21 June

- CHASE (IEEE/ACM Conf. on Connected Health: Applications, Systems and Eng Technologies) Orlando, Florida, USA

22 Jun

- CBMS (IEEE Int'l Symposium on Computer-Based Medical Systems) L'Aquila, Italy

23 June

- ICIS (IEEE/ACIS Int'l Conf. on Computer and Information Science) Wuxi, China

24 June

- LICS (ACM/IEEE Symposium on Logic in Computer Science) Boston, USA

26 June

- BSC (Biennial Symposium on Communications) Montreal, Canada
- COMPSAC (IEEE Computers, Software, and Applications Conf.) Torino, Italy
- JCDL (ACM/IEEE Joint Conf. on Digital Libraries) Santa Fe, USA

28 June

- WETICE (IEEE Int'l Conf. on Enabling Technologies: Infrastructure for Collaborative Enterprises) Paris, France

JULY

1 July

- SSE (IEEE Int'l Conf. on Software Services Eng) Chicago, USA

3 July

- EuroS&P (IEEE European Symposium on Security and Privacy) Delft, Netherlands

10 July

- CSF (IEEE Computer Security Foundations Symposium) Dubrovnik, Croatia
- ICALT (IEEE Int'l Conf. on Advanced Learning Technologies) Orem, USA
- ICME (IEEE Int'l Conf. on Multimedia and Expo) Brisbane, Australia

18 July

- ICDCS (IEEE Int'l Conf. on Distributed Computing Systems) Hong Kong, Hong Kong
- SCC (IEEE Space Computing Conf.) Pasadena, USA
- SMC-IT (IEEE Int'l Conf. on Space Mission Challenges for Information Technology) Pasadena, USA

19 July

- ASAP (IEEE Int'l Conf. on Application-specific Systems, Architectures and Processors) Porto, Portugal

AUGUST

1 August

- IRI (IEEE Int'l Conf. on Information Reuse and Integration for Data Science), Bellevue, WA, USA

SEPTEMBER

4 September

- RE (IEEE Int'l Requirements Eng. Conf.), Hannover, Germany

IEEE COMPUTER SOCIETY D&I FUND

Drive Diversity & Inclusion in Computing

Supporting projects and programs that positively impact diversity, equity, and inclusion throughout the computing community.

Do you have a great idea for new programs that will positively impact diversity, equity, and inclusion throughout the computing community?

The IEEE Computer Society Diversity & Inclusion Committee seeks proposals for projects, programs, and events that further its mission. New programs that deliver education, outreach, and support, including, but not limited to, mentoring programs at conferences, panel discussions, and webinars, are welcomed.

Help propel the Computer Society's D&I programs—submit a proposal today!

https://bit.ly/CS-Diversity-CFP

Donations to the IEEE Computer Society D&I Fund are welcome!

IEEE COMPUTER SOCIETY

IEEE Foundation