

COMPUTING edge

- Hardware
- Bioinformatics
- Security and Privacy
- Blockchain

MAY 2024

www.computer.org

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.*

DONATE TODAY!



IEEE
COMPUTER
SOCIETY

IEEE Foundation

STAFF

Editor

Lucy Holden

Periodicals Portfolio Senior Managers

Carrie Clark and Kimberly Sperka

Director, Periodicals and Special Projects

Robin Baldwin

Production & Design Artist

Carmen Flores-Garvey

Periodicals Operations Project Specialists

Priscilla An and Christine Shaughnessy

Senior Advertising Coordinator

Debbie Sims

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.

Editorial: Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author's or firm's opinion. Inclusion in *ComputingEdge* does not necessarily constitute endorsement by the IEEE or the Computer Society. All submissions are subject to editing for style, clarity, and space.

Reuse Rights and Reprint Permissions: Educational or personal use of this material is permitted without fee, provided such use: 1) is not made for profit; 2) includes this notice and a full citation to the original work on the first page of the copy; and 3) does not imply IEEE endorsement of any third-party products or services. Authors and their companies are permitted to post the accepted version of IEEE-copyrighted material on their own Web servers without permission, provided that the IEEE copyright notice and a full citation to the original work appear on the first screen of the posted copy. An accepted manuscript is a version which has been revised by the author to incorporate review suggestions, but not the published version with copy-editing, proofreading, and formatting added by IEEE. For more information, please go to: http://www.ieee.org/publications_standards/publications/rights/paperversionpolicy.html. Permission to reprint/republish this material for commercial, advertising, or promotional purposes or for creating new collective works for resale or redistribution must be obtained from IEEE by writing to the IEEE Intellectual Property Rights Office, 445 Hoes Lane, Piscataway, NJ 08854-4141 or pubs-permissions@ieee.org. Copyright © 2024 IEEE. All rights reserved.

Abstracting and Library Use: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Unsubscribe: If you no longer wish to receive this *ComputingEdge* mailing, please email IEEE Computer Society Customer Service at help@computer.org and type "unsubscribe *ComputingEdge*" in your subject line.

IEEE prohibits discrimination, harassment, and bullying. For more information, visit www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

IEEE Computer Society Magazine Editors in Chief

Computer

Jeff Voas, *NIST*

Computing in Science & Engineering

İlkay Altıntaş, *University of California, San Diego (Interim EIC)*

IEEE Annals of the History of Computing

David Hemmendinger, *Union College (Interim EIC)*

IEEE Computer Graphics and Applications

André Stork, *Fraunhofer IGD and TU Darmstadt*

IEEE Intelligent Systems

San Murugesan, *Western Sydney University*

IEEE Internet Computing

Weisong Shi, *University of Delaware*

IEEE Micro

Hsien-Hsin Sean Lee, *Intel Corporation*

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

Sigrid Eldh, *Ericsson*

IT Professional

Charalampos Z. Patrikakis, *University of West Attica*

MAY 2024 • VOLUME 10 • NUMBER 5

COMPUTING
edge



18

Changing
Aesthetics in
Biomolecular
Graphics

26

The DNA Data
Storage Model

42

Scams, Frauds,
and Crimes in the
Nonfungible
Token Market

Hardware

8 Scarcity and Global Insecurity: The Semiconductor Shortage

JEFFREY VOAS, NIR KSHETRI, AND JOANNA F. DEFRANCO

14 Interview With Ronnie Chatterji, Coordinator for the Creating Helpful Incentives to Produce Semiconductors and Science Act

SHANE GREENSTEIN

Bioinformatics

18 Changing Aesthetics in Biomolecular Graphics

LAURA A. GARRISON, DAVID S. GOODSSELL, AND STEFAN BRUCKNER

26 The DNA Data Storage Model

DAVE LANDSMAN AND KARIN STRAUSS

Security and Privacy

35 Pervasive Healthcare: Privacy and Security in Data Annotation

EMMA L. TONKIN AND KRISTINA YORDANOVA

40 Privacy in the Era of 5G, IoT, Big Data, and Machine Learning

ELISA BERTINO

Blockchain

42 Scams, Frauds, and Crimes in the Nonfungible Token Market

NIR KSHETRI

48 Cryptographic–Biometric Self-Sovereign Personal Identities

DORON DRUSINSKY

Departments

4 Magazine Roundup

7 Editor's Note: The Risks and Rewards of Technology Dependencies

65 Conference Calendar

Subscribe to *ComputingEdge* for free at
www.computer.org/computingedge



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

Cui Bono? Software Professionals Should Always Ask "Who Benefits?"

Computing professionals have important responsibilities linked to ethics. The authors of this February 2024 *Computer* article review key codes of ethics for computing professionals, delving into the important guidance described in each. However, they note that by omission or commission, computing professionals do not consistently account for the impacts of their work on a broad range of stakeholders.

Computing

The Intelligence Advanced Research Projects Activity's Advanced Graph Intelligent Logical Computing Environment Program: Reinventing Computing

This July/August 2023 *Computing in Science and Engineering* article describes the Intelligence Advanced Research Projects Activity's Advanced Graph Intelligent Logical Computing Environment program, the first step toward catalyzing a computing revolution by

pioneering new hardware and software co-designs tailored for data handling and movement.

IEEE Annals

Literature and Artificial Intelligence

This July–September 2023 *IEEE Annals of the History of Computing* article gives readers a "literary" perspective on the history of artificial intelligence. On one hand, it is a literary-historical retrospective of the decades since the Second World War, but it also emphasizes the basic probabilistic trait of artificial intelligence, which becomes recognizable especially against the background of current technological developments. The space of literature is the horizon of possibility that accompanies what is real for man in early modern times and to which his activity is directed.

IEEE Computer Graphics

Rendering the Bluish Appearance of Snow: When Light Transmission Matters

This article in the January/February 2024 issue of *IEEE Computer*

Graphics and Applications examines the distinct bluish colorations observed when light is transmitted through snow, and the authors present a method for the predictive rendering of this phenomenon, taking into account the variability of snow's physical and morphological characteristics.

IEEE Intelligent Systems

Unjustified Sample Sizes and Generalizations in Explainable Artificial Intelligence Research: Principles for More Inclusive User Studies

The authors of this November/December 2023 *IEEE Intelligent Systems* article analyzed explainable AI (XAI) user studies ($n = 220$) published between 2012 and 2022, to explore the suitability of the sample sizes for these studies vis-à-vis the conclusions. Where there are methodological problems in sample sizes, it can impede evaluations of whether XAI systems implement the explainability called for in ethical frameworks. The authors outline principles for more inclusive XAI user studies.



IEEE Internet Computing

Measuring the Energy of Smartphone Communications in the Edge-Cloud Continuum: Approaches, Challenges, and a Case Study

As computational resources are placed at different points in the edge-cloud continuum, not only is the responsiveness on the client side affected, so too is the amount of energy spent during communications. The authors of this November/December 2023 *IEEE Internet Computing* article summarize the main approaches used to estimate smartphones' energy consumption and the main difficulties such approaches typically encounter. A case study illustrates how such approaches can be put into practice.

IEEE micro

Addressing the Gap Between Training Data and Deployed Environment by On-Device Learning

The accuracy of tiny machine learning applications is often affected by various environmental factors, such as noises, location/calibration of sensors, and time-related changes. This November/December 2023 *IEEE Micro* article introduces a neural network based on-device learning (ODL) approach

to address this issue by retraining in deployed environments. The authors' approach relies on semisupervised sequential training of multiple neural networks tailored for low-end edge devices.

IEEE MultiMedia

A Novel Learning Dictionary for Sparse Coding-Based Key Point Detection

The rotational-invariant dictionary in the sparse coding-based key point detector (SCK) is manually generated using a time-consuming process of selecting a good seed dictionary and combining multiple versions of its rotated atoms. In this October–December 2023 *IEEE MultiMedia* article, the authors describe automating this process using a novel duplet autoencoder structure, in which the weights between the input and the hidden layers are designed to embed a rotational-invariant dictionary.

IEEE pervasive computing

Considering Wearable Health Tracking Devices and Pandemic Preparedness for Universities

The authors of this article, in *IEEE Pervasive Computing's*

October–December 2023 issue, examine the results of a year-long in-the-wild study in which 35 participants at a university wore Oura Rings, which are worn on the finger and are used to track sleep and physical activity. After an orientation, the group of study participants wore their rings with no restrictions or minimum wearing requirements. By retroactively looking at how participants used the rings for monitoring their health, the authors were able to identify successful strategies and potential problems with employing these types of wearables for health monitoring in universities.

IEEE SECURITY & PRIVACY

Shockvertising, Malware, and a Lack of Accountability: Exploring Consumer Risks of Virtual Reality Advertisements and Marketing Experiences

The authors of this *IEEE Security & Privacy* article, which was published in the January/February 2024 issue, provide evidence and discuss how many companies increasingly use virtual reality (VR) for their advertising campaigns. This begs the question, *What risks does VR advertising pose for consumers?* The authors describe and analyze VR

marketing experiences to identify risks and discuss opportunities to address those and future risks in VR advertising.

IEEE Software

Software Engineering Education for Technical Engineering Degrees: A Comparison With the Needs of Robotics Software Engineering Education

In this November/December 2023 *IEEE Software* article, the author compares software engineering education for traditional computer science and software engineering degree programs with the needs of robotics software engineering, concluding that technical engineering degrees need to emphasize social aspects of software

engineering, group work, and weigh advantages and disadvantages between different solution options.

IT Professional

Identifying Networked Patterns in Memecoin Twitter Accounts Using Exponential Random Graph Modeling

This November/December 2023 *IT Professional* article investigates the structure of memecoin communication on Twitter. The authors found that Dogecoin and ShibaInu served as information dissemination hubs and they explored the limitations of these relationships, noting that memecoins can use Twitter to disseminate information designed to increase their profitability. 🌍

**Join the IEEE
Computer
Society**
computer.org/join

IEEE COMPUTER SOCIETY Call for Papers

Write for the IEEE Computer Society's authoritative computing publications and conferences.

GET PUBLISHED
www.computer.org/cfp





Editor's Note

The Risks and Rewards of Technology Dependencies

Embedding more and more new technology into our society comes with both risks and rewards. Our increasing dependency on evolving technology boosts the economy, yet also makes us vulnerable to supply chain and security risks. This issue of *ComputingEdge* explores the rewards of technological advances, including DNA storage, improved biomolecular graphics and proof of identity, and data collection. The articles also delve into the risks of some of these advances, including fraud, security and privacy issues, and hardware shortage.

The global semiconductor shortage forced governments to reconsider access and production. *IT Professional's* "Scarcity and Global Insecurity: The Semiconductor Shortage" explores the global implications of the shortage of semiconductors and individual nation state security implications. In the *IEEE Micro* article,

"Interview With Ronnie Chatterji, Coordinator for the Creating Helpful Incentives to Produce Semiconductors and Science Act," the author interviews Ronnie Chatterji about the expected impact of the CHIPS and Science Act on the semiconductor industry.

The biotechnology industry is progressing toward exciting new frontiers. The authors of "Changing Aesthetics in Biomolecular Graphics," from *IEEE Computer Graphics and Applications*, discuss the evolving landscape and future of biomolecular imagery. *Computer's* "The DNA Data Storage Model" outlines the advantages of using DNA storage as an archival storage solution.

Although the advancement of data collection tools improves many important systems, it often comes at the expense of personal privacy. In the *IEEE Pervasive Computing* article, "Pervasive Healthcare: Privacy and Security in Data

Annotation," the authors acknowledge the privacy risks underlying pervasive healthcare monitoring and identify possible solutions. "Privacy in the Era of 5G, IoT, Big Data, and Machine Learning," from *IEEE Security & Privacy*, reveals the complications underlying the exchange of private personal data for better collective security.

Blockchain technology introduces potential solutions for security problems although it also poses security risks. *Computer* article "Scams, Frauds, and Crimes in the Nonfungible Token Market" presents an investigation into the rise of cybercrime in the nonfungible (NFT) market. In "Cryptographic-Biometric Self-Sovereign Personal Identities," from *Computing in Science & Engineering*, the author proposes using self-sovereign identities (SSIs), which use blockchain technology, to help enable private and secure proof of identity. 🌐

Scarcity and Global Insecurity: The Semiconductor Shortage

Jeffrey Voas , IEEE Fellow

Nir Kshetri , University of North Carolina, Greensboro, NC, 27599, USA

Joanna F. DeFranco , Penn State Great Valley: School of Graduate Professional Studies,
Malvern, PA, 19355, USA

Scarcity is not considered an “ility” such as reliability and performance are, however, we see scarcity’s growing impact on quality of life, fear, and trust. Nevertheless, scarcity is measurable like reliability and performance. Examples of scarcity are easy to find such as the recent Colonial pipeline ransomware attack that created gasoline hoarding (<https://www.cnet.com/news/colonial-pipeline-ceo-tells-senate-decision-to-pay-hackers-was-made-quickly/>) and in the early days of Covid-19 which caused toilet paper and cleaning product stockpiling.¹

Scarcity has always been a global concern, particularly when Middle East conflicts caused gasoline shortages, rationing, and price spikes (<https://theapopkavoice.com/middle-east-conflict-causing-higher-fuel-prices/>). In the Western U.S. today, the main water reservoir needed by many states is sinking to its lowest level on record (<https://www.reuters.com/world/us/hover-dam-reservoir-hits-record-low-sign-extreme-western-us-drought-2021-06-10/>). Who knows how this will play out—we may have states suing other states for their water supply (<https://wsabc.ca/texas-suing-new-mexico-in-water-war/>).

In this article, we focus on *semiconductors*; they are becoming harder to attain. Semiconductors (also referred to as chips) are embedded in nearly everything today. Any scarcity of semiconductors has global implications as well as individual nation state security implications.² Here, we explore the story behind this shortage.

Most industries rely on uninterrupted access to semiconductors. According to Goldman Sachs, 169 industries in the U.S. use semiconductors in their

products. For example, a typical car uses between 50 and 150 semiconductors (<https://www.cnn.com/2021/04/29/business/chip-shortages-smartphones-consumer-goods/index.html>) and a modern car can use up to 3,000 (<https://www.nytimes.com/2021/04/23/business/auto-semiconductors-general-motors-mercedes.html>). In addition, globally, semiconductors are the fourth most traded product after crude oil, refined oil, and cars (<https://www.fierceelectronics.com/electronics/chip-sales-up-15-as-leaders-focus-government-subsidies>).

In addition to cars, the 2021 semiconductor shortage has also impacted the production of products such as phones, entertainment consoles, and TVs (<https://interestingengineering.com/what-global-shortage-of-computer-chips-means-for-you>). The current shortfall has been especially pronounced in “less-advanced” chips because the world’s biggest semiconductor producers have focused on “cutting-edge” chips that offer higher profit-margins.³ Due to the ubiquitous application of semiconductors, the current shortage has affected most economic sectors. It is reported that kitchen appliances such as microwaves, refrigerators, and washing machines that are controlled by less advanced processors are increasingly difficult to find (<https://www.zdnet.com/article/the-global-chip-shortage-is-a-bigger-problem-than-everyone-realised-and-it-will-go-on-for-longer-too/>). It is predicted that a recovery from this shortage will not be seen before the second quarter of 2022 (<https://www.gartner.com/en/newsroom/press-releases/2021-05-12-gartner-says-global-chip-shortage-expected-to-persist-until-second-quarter-of-2022>).

As a result, companies relying on access to semiconductors have reduced production capacities and delayed new product launches. For example, due to limited supplies of specific chips, Apple predicted a sales loss of between \$3–\$4 billion in Q2 2021. (<https://www.siliconrepublic.com/machines/global-chip-shortage-eu-apple>). The shortage has also resulted in car

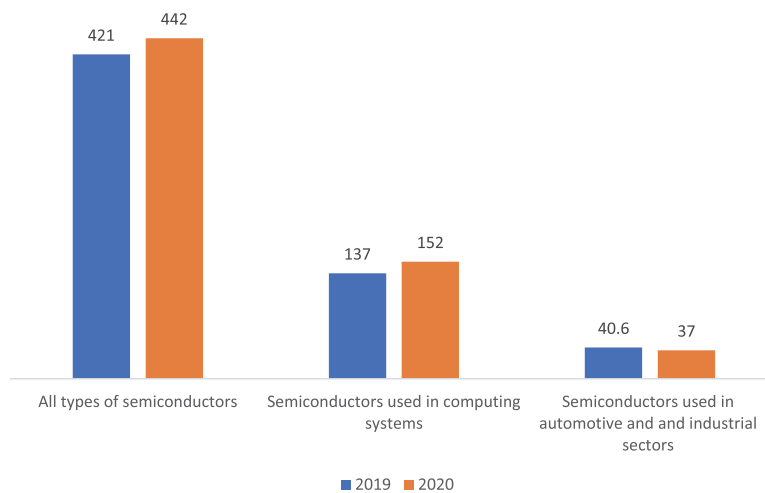


FIGURE 1. Worldwide semiconductor revenues in 2019 and 2020 (\$, billions). Data source: “Worldwide Semiconductor Revenue Grew 5.4% in 2020 Despite COVID-19 and Further Growth is Forecast in 2021, According to IDC” February 2, 2021 <https://www.idc.com/getdoc.jsp?containerId=prUS47424221>.

manufacturers using their limited supply of chips for their more profitable models (<https://www.wsj.com/articles/global-chip-shortage-set-to-worsen-for-car-makers-11619708393>). It is estimated that this shortage of chips will impact 964,000 vehicles in 2021.⁴ The automobile industry’s lost revenue for 2021 is predicted to reach \$61 billion.⁵

FACTORS

The semiconductor shortage is the result of converging factors. The demand for microprocessors was already growing before the COVID-19 pandemic to support the development of new markets such as 5G, self-driving vehicles, artificial intelligence, and the Internet of Things.

During the early weeks of the COVID-19 lockdown, automobile plants worldwide were forced to shut down and sales were drastically reduced. Consequently, the automobile industry reduced semiconductor purchases.⁶ However, the lockdown facilitated a growth demand of PCs due to the increase of online education and those working from home, however, this demand was unfulfilled—keeping PC growth in the single digits (<https://tinyurl.com/5ctef29x>). Figure 1 shows the rapid demand increase for microprocessors used in these products. When the demand for cars rebounded, automobile manufacturers discovered that the semiconductor manufacturers had readjusted their productions to fulfill the orders from those other industries that experienced a boom during the pandemic.⁶

Natural factors and disasters such as weather and fire worsened the situation. In March 2021, Japan’s Renesas Semiconductor Manufacturing Co. Ltd. had fire damage (<https://www.wsj.com/articles/global-chip-shortage-set-to-worsen-for-car-makers-11619708393>). Renesas produces about one-third of microcontroller chips embedded in cars globally (<https://auto.economictimes.indiatimes.com/news/auto-components/rene-sas-sees-17-bln-yen-sales-revenue-loss-in-q2-due-to-fire/82292225>). Likewise, Texas-based semiconductor manufacturing facilities were forced to shut down due to a cold weather outbreak in February 2021.⁶

Another limiting factor is that semiconductor production requires a lot of water. Taichung Taiwan, a production hub of the world’s largest semiconductor company Taiwan Semiconductor Manufacturing Company (TSMC), experienced a severe drought in 2021 which also worsened the shortage. Companies in the city were required to reduce water usage by 15%. TSMC started transporting water using tanker trucks from other parts of the country. However, each truck only carries 20 tons of water. TSMC uses about 200,000 tons of water each day (<https://asia.nikkei.com/Business/Tech/Semiconductors/Taiwan-drought-at-most-critical-phase-for-chip-sector>).

The semiconductor shortage in the U.S. and other countries has been further exacerbated by Chinese firms stockpiling chips. For instance, China’s integrated circuit (IC) imports in Q1 2021 increased by more than a third compared to Q1 2020. China’s increased stockpiling of semiconductors may lead to further sanctions against Chinese tech companies

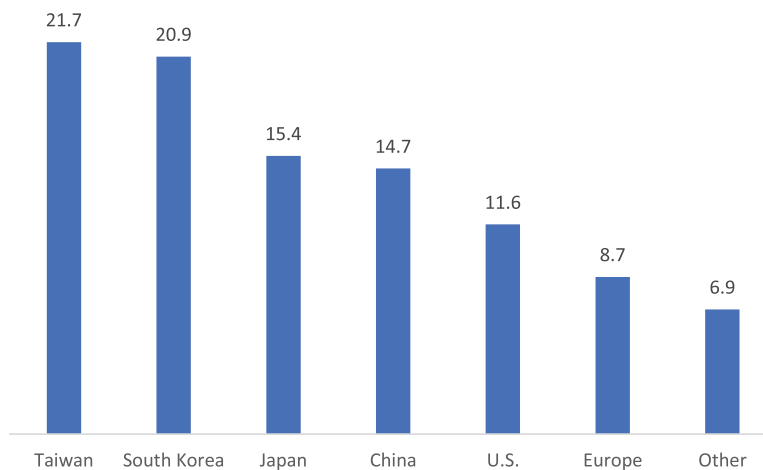


FIGURE 2. Shares of major economies in global semiconductor manufacturing capacity (2020). Data source: Boston Consulting Group, Semiconductor Industry Association.⁹

(<https://www.zdnet.com/article/the-global-chip-shortage-is-a-bigger-problem-than-everyone-realised-and-it-will-go-on-for-longer-too/>). While the Chinese government's long-term plan has been to expand its indigenous semiconductor industry, the country's technology companies have viewed increased chip import as the better option under current conditions (<https://www.theburnin.com/industry/china-wants-327b-domestic-semiconductor-sector-2021-02-15/>).

GLOBAL PRODUCTION NETWORKS

Global production networks (GPNs), an economic model that coordinates the interconnection of stakeholders in a particular industry, are becoming increasingly common across many industries.⁷ In the semiconductor industry, GPNs became more widespread in the late 1980s after the business model among semiconductor designers moved toward outsourced manufacturing. In this fabless model (i.e., outsourcing the fabrication of the chips), a company designs and sells the hardware and semiconductor chips but relies on chip-making factories known as *foundries* to manufacture the chips.

East Asia has emerged as the epicenter of fabless manufacturing. TSMC has been credited for pioneering the "foundry and fabless" model. According to Trendforce, TSMC and Samsung have foundry market shares of 55% and 18%, respectively.⁸ About three-quarters of the global semiconductor manufacturing capacity, as well as key suppliers of key materials, are in Asia (Figure 2).

East Asia's dominance is even more pronounced in the manufacturing of advanced semiconductor

devices. Currently, 100% of the world's highly advanced logic semiconductor (below ten nanometers) manufacturing capacity is in two Asian economies: Taiwan: 92%, South Korea 8%.¹⁰ In 2020, Samsung and TSMC introduced five-nanometer chips. They plan to introduce the first three-nanometer chips in 2022.¹¹

Dominant manufacturers of semiconductors such as TSMC and Samsung rely heavily on equipment and machinery supplied by semiconductor capital equipment vendors (semicap). Among the top five global semicap companies, three are in the U.S., and Japan and the Netherlands each have one (see Figure 3).

It would not be easy to undo existing chip GPNs and the current fabless model. According to a study conducted by Boston Consulting Group (BCG) and Semiconductor Industry Association (SIA), it would cost \$1 trillion in upfront investment to establish a fully self-sufficient local supply chain in each region to meet current chip demands. Such investments would lead to a 35% to 65% increase in chip prices.¹²

THE WEST RESPONDS

In addition to semiconductor usage in a wide range of products, such as household appliances, computers, phones, and cars, semiconductors are also embedded in military equipment. As already mentioned, Taiwan is the key supplier to the world. If Taiwanese foundries were ever fully shut down, replacing their production would take three years and a \$350 billion investment in other economies to build back a sufficient

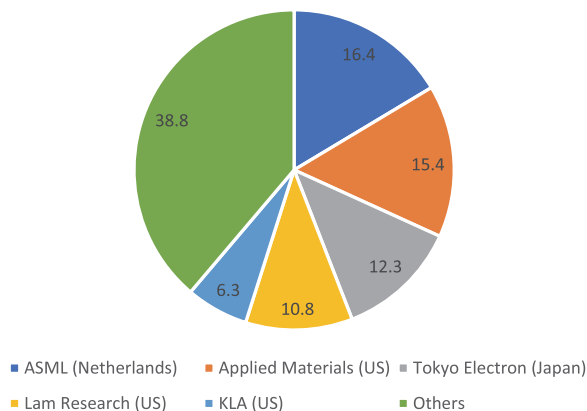


FIGURE 3. Top semicap companies and their market shares (2020). Data source: "Semiconductor wafer front end (WFE) equipment market share worldwide from 2018 to 2020, by supplier," <https://www.statista.com/statistics/267392/market-share-of-semiconductor-equipment-manufacturers/>.

capacity.¹² The economic, military, and national security implications of the current reliance on Taiwan are striking. Finally, East Asia's exposure to high seismic activity further increases risk to global supplies (<https://www.semiconductor-digest.com/2021/05/13/the-chip-shortage-wake-up-call/>).

Addressing the semiconductor shortage has become a priority for policy makers worldwide. However, a single semiconductor fabrication plant costs \$10–\$20 billion to build (<https://www.eastasiaforum.org/2021/02/22/china-chases-semiconductor-self-sufficiency/>). Some form of public support is critical to accomplish this. One important lesson from Taiwan and South Korea is that government support played a key role in the growth of their industry.¹⁰

Western economies are formulating similar strategies to those that worked in Asia. The European Commission developed a 10-year strategy for the development of its semiconductor industry.¹³ In April 2021, Intel's CEO met with two EU commissioners to develop a strategy to make Europe more competitive in chip manufacturing (<https://www.cnbc.com/2021/05/07/chip-shortage-is-starting-to-have-major-real-world-consequences.htm>). The EU has also shown determination to reduce its dependence on the U.S. and Asia. The goal is to manufacture one-fifth of the world's semiconductors by 2030.¹³ The EU is planning to spend more than \$150 billion to develop advanced technologies including chips and artificial intelligence (<https://www.wsj.com/articles/eu-seeks-to-double>

[share-of-world-chip-market-by-2030-in-digital-sovereignty-drive-11615305395](https://www.wsj.com/articles/eu-seeks-to-double-share-of-world-chip-market-by-2030-in-digital-sovereignty-drive-11615305395)).

The U.S. government has realized that the global semiconductor shortage has significant national security implications.¹⁴ The U.S. is taking legislative and policy measures for strengthening domestic semiconductor manufacturing. In February 2021, an executive order was signed, which involves assessing potential risks in semiconductor supply chains. A bill known as the Creating Helpful Incentives to Produce Semiconductors for America Act (CHIPS Act) (H.R.7178), introduced in 2020, aims to provide incentives to enable advanced R&D in the semiconductor industry and securer supply chains (<https://www.congress.gov/bill/116th-congress/house-bill/7178>).

The U.S. government has also taken measures to support domestic manufacturing of semiconductors through subsidies and other incentives. For instance, \$50 billion has been designated for semiconductor manufacturing and research as part of the President Biden's expansive infrastructure proposal.

U.S. and foreign semiconductor manufacturers have been responding to the favorable policy environment. In February 2021, the U.S. firm Intel announced a plan to spend \$20 billion to build two chip factories. This is likely to reduce the current reliance on foreign semiconductor foundries such as TSMC and Samsung.⁹

Major foreign semiconductor manufacturers are also entering the U.S. market. In 2020, TSMC announced a plan to spend \$12 billion to build a semiconductor plant in Arizona, which is expected to be completed in 2024. Samsung was reported to be considering Texas and Arizona for a new logic-chip facility; this facility would likely be the most advanced in this category in the U.S. Both companies are planning to take advantage of government subsidies to help cover the costs of setting up these fabrication plants in the U.S.⁸ Critics, however, argue that the planned \$50 billion investment in the U.S. semiconductor industry over multiple years is not sufficient given that TSMC alone is planning to spend \$100 billion over the next three years (<https://www.marketplace.org/2021/04/21/shortage-of-semiconductors-is-a-security-risk-and-what-the-u-s-can-do-about-it/>).

CONCLUSION

Scarcity engenders stockpiling, fear, and distrust. Global semiconductor supply chains are vulnerable to a wide array of factors such as fires, weather events, and political tension. This industry's economic and

national security implications are evident—chips are embedded in all types of products including those used by the military. And Western countries have realized the importance of reducing their reliance on Asian suppliers.

Building a new semiconductor manufacturing facility is slow and expensive. However, strong, local, semicap companies provide an advantage to European and the U.S. semiconductor industries that countries in Asia may not be able to compete with. Fortunately, there may be an opportunity using the strong ties between the manufacturing facilities and semicap companies to eventually boost semiconductor production. 🌐

DISCLAIMER

The authors are completely responsible for the content in this article. The opinions expressed here are their own.

REFERENCES

1. J. Voas and N. Kshetri, "Scarcity," *IEEE Comput.*, vol. 54, no. 1, pp. 26–28, 2021.
2. N. Kshetri and J. Voas, "Where's the silicon?," *IEEE Comput.*, vol. 54, no. 8, pp. 11–12, Aug. 2021.
3. E. Jeong and D. Strumpf, "Why the chip shortage is so hard to overcome," *Wall Street J.*, Apr. 19, 2021. [Online]. Available: <https://www.wsj.com/articles/why-the-chip-shortage-is-so-hard-to-overcome-11618844905>
4. B. Klayman, "GM hit by chip shortage, to cut production at four plants," Feb. 3, 2021. [Online]. Available: <https://www.reuters.com/article/us-gm-semiconductors-exclusive/gm-hit-by-chip-shortage-to-cut-production-at-four-plants-idUSKBN2A32LL>
5. K. Nicholas, "Carmakers face \$61 billion sales hit from pandemic chip shortage," Jan. 27, 2021. [Online]. Available: <https://www.bloomberg.com/news/articles/2021-01-27/covid-pandemic-slows-down-chipmakers-causes-car-shortage>
6. J. Whalen, "Chip shortage spreads, hurting sales at Apple and Samsung," Apr. 29, 2021. [Online]. Available: <https://www.washingtonpost.com/technology/2021/04/29/apple-caterpillar-chip-shortage/>
7. G. Gereffi et al., "Introduction: Global commodity chains," in *Commodity Chains and Global Capitalism*, G. Gereffi and M. Korzeniewicz, Eds., Westport, CT, USA: Praeger, 1994, pp. 1–14.
8. A. Kharpal, "How Asia came to dominate chipmaking and what the U.S. wants to do about it," *CNBC*, 2021. [Online]. Available: <https://www.cnbc.com/2021/04/12/us-semiconductor-policy-looks-to-cut-out-china-secure-supply-chain.html>
9. S. Kim, "South Korea and Taiwan's Chip power rattles the U.S. and China," *Bloomberg*, 2021. [Online]. Available: <https://www.bloomberg.com/news/articles/2021-03-03/chip-shortage-taiwan-south-korea-s-manufacturing-lead-worries-u-s-china>
10. Semiconductor Industry Association, "SIA Urges U.S. Government Action to Strengthen America's Semiconductor Supply Chain, Monday," Apr. 5, 2021. [Online]. Available: <https://www.semiconductors.org/sia-urges-u-s-government-action-to-strengthen-americas-semiconductor-supply-chain/>
11. R. Sharma, "Pound for pound, Taiwan is the most important place in the world," *The New York Times*, 2020. [Online]. Available: <https://www.nytimes.com/2020/12/14/opinion/taiwan-computer-chips.html>
12. A. Varas et al., "Strengthening the global semiconductor supply chain in an uncertain era," Apr. 2021. [Online]. Available: https://www.semiconductors.org/wp-content/uploads/2021/04/SIA-BCG-Report_Strengthening-the-Global-Semiconductor-Supply-Chain_April-2021.pdf
13. J. Keane, "The EU sets out plan to build 20pc of the world's semiconductors," Mar. 9, 2021. [Online]. Available: <https://www.siliconrepublic.com/machines/semiconductors-manufacturing-eu-2030>
14. T. Kaplan, "Amid a chip shortage, the White House gathers business leaders to discuss supplies," Apr. 12, 2021. [Online]. Available: <https://www.nytimes.com/2021/04/12/business/semiconductor-chip-shortage.html>

JEFFREY VOAS is the Editor-in-Chief of *Computer*. He is a Fellow of IEEE. Contact him at j.voas@ieee.org.

NIR KSHETRI is a Professor with the Bryan School of Business and Economics, the University of North Carolina at Greensboro, Greensboro, NC, USA. Contact him at nbkshetr@uncg.edu.

JOANNA F. DEFRANCO is an Associate Professor of software engineering with the Penn State Great Valley: School of Graduate Professional Studies, Malvern, PA, USA. Contact her at jfd104@psu.edu.



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:** +1 714 816 2121;
Email: help@computer.org

EXECUTIVE COMMITTEE

President:	Jyotika Athavale
President-Elect:	Hironori Washizaki
Past President:	Nita Patel
First VP:	Grace A. Lewis
Second VP:	Nils Aschenbruck
Secretary:	Mrinal Karvir
Treasurer:	Darren Galpin
VP, Membership & Geographic Activities:	Kwabena Boateng
VP, Professional & Educational Activities:	Cyril Onwubiko
Interim VP, Publications:	Jaideep Vaidya
VP, Standards Activities:	Edward Au
VP, Technical & Conference Activities:	Terry Benzel
2023–2024 IEEE Division VIII Director:	Leila De Floriani
2024–2025 IEEE Division V Director:	Christina M. Schober
2024 IEEE Division V Director-Elect:	Cecilia Metra

BOARD OF GOVERNORS

Term Expiring 2024:

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

Term Expiring 2025:

İlkay Altıntaş, Mike Hinchey, Joaquim Jorge, Rick Kazman, Carolyn McGregor, Andrew Seely

Term Expiring 2026:

Megha Ben, Terry Benzel, Mrinal Karvir, Andreas Reinhardt, Deborah Silver, Yoshiko Yasuda

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 Tubbs
Director, Membership Development:	Eric Berkowitz
Director, Periodicals & Special Projects:	Robin Baldwin

IEEE BOARD OF DIRECTORS

President & CEO:	Thomas M. Coughlin
President-Elect:	Kathleen A. Kramer
Director & Secretary:	Forrest (Don) Wright
Director & Treasurer:	Gerardo Barbosa
Past President:	Saifur Rahman
Director & VP, Educational Activities:	Rabab Ward
Director & VP, Publication Services & Products:	Sergio Benedetto
Director & VP, Member & Geographic Activities:	Antonio Luque
Director & President, Standards Association:	James E. Matthews III
Director & VP, Technical Activities:	John Verboncoeur
Director & President, IEEE-USA:	Timothy T. Lee



Interview With Ronnie Chatterji, Coordinator for the Creating Helpful Incentives to Produce Semiconductors and Science Act

Shane Greenstein , Harvard Business School, Boston, MA, 02163, USA

President Joe Biden signed the Creating Helpful Incentives to Produce Semiconductors and Science Act (CHIPS) and Science Act on 9 August 2022. The act provides billions of dollars in subsidies and tax credits for manufacturing semiconductors on U.S. soil, research on semiconductors, workforce training, and investment in equipment. On 20 September 2022, the White House announced that Ronnie Chatterji would serve as the White House Coordinator for CHIPS Implementation at the National Economic Council. Before that, Aaron “Ronnie” Chatterji served as the chief economist for the Department of Commerce since April 2021.

Chatterji is currently on leave from his position as the Mark Burgess & Lisa Benson-Burgess Distinguished Professor of Business and Public Policy at Duke University’s Fuqua School of Business. He also holds a secondary appointment at Duke’s Sanford School of Public Policy. Ronnie received his Ph.D. degree from the Haas School of Business at the University of California, Berkeley and his B.A. degree in economics from Cornell University.

The CHIPS and Science Act affects the work experience of many IEEE members. To keep members informed, Shane Greenstein, “Micro Economics” columnist for *IEEE Micro*, recently interviewed Dr. Chatterji.

Shane Greenstein: The Chips and Science Act had many provisions, however, let’s focus on those most relevant to the economics of integrated circuits. For decades, the U.S. government has had a hand in integrated circuits by procuring military equipment and funding some R&D activities. This bill significantly expands budgetary and tax commitments beyond those areas. Which are the most significant commitments? Why?

Ronnie Chatterji: The CHIPS and Science Act is a historic commitment to semiconductor manufacturing and R&D in the United States. It includes a \$39 billion program to support manufacturing, including the construction of new fabs (fabrication plants) and strengthening key parts of the supply chain. There is an \$11 billion program to advance R&D at the Department of Commerce, including the establishment of a National Semiconductor Technology Center (NSTC). And there are significant investments to stand up new programs at the U.S. Departments of Defense and State as well as the National Science Foundation. Another key component of CHIPS and Science is a 25% investment tax credit that the U.S. Department of the Treasury is leading on. So it is a true whole-of-government approach, as it must be, given the scale of the challenge. My job at the National Economic Council is to ensure that the Act is implemented across government and achieves our objectives.

Greenstein: The bill had many motivations. Let’s break them down and focus on one economic aspect at a time. There are national interests in supporting a supply of frontier processors for domestic users, for example, developing artificial intelligence for large language models. Yet, other industries desire an ample supply of inexpensive chips and do not need the frontier, say, in automobiles. The former involves the production of new facilities, while the latter could include retrofitting established facilities and accommodating nondomestic pools of suppliers. So what do we think about supporting those different types of goals?

Chatterji: Chips go into all kinds of products, from our most advanced defense systems, to our cars, electric grid, routers, and refrigerators. During the pandemic, we saw how disruptions in the supply chain for auto chips drove large price increases, accounting for a significant share of inflation in 2021. So we have both economic and national security motivations for establishing a strong manufacturing base for chips in the

United States. The CHIPS program, particularly the 9902 program out of the U.S. Department of Commerce for manufacturing incentives, is specifically designed to support both of these interests and ensure we have adequate supply of leading-edge and mature chips.

Greenstein: Let's explore the economics of supply-chain resilience. There is collective industry interest in supporting resilience, but, potentially, no firm is incentivized to coordinate and lead the effort. How does your activity invest in realizing that goal? Another topic is workforce requirements. Companies want a trained workforce, but perhaps none individually invest in it. What do you think about that?

Chatterji: The pandemic and geopolitical events like Russia's war in Ukraine have changed the ways companies think about their supply chains. Compared to two years ago, companies are making significant investments in supply-chain resilience and acquiring insights they never had before about where their inputs come from. It is a big change, and the Biden administration is undertaking several initiatives to support supply-chain resilience. One of the most notable is the recent Notice of Funding we released from the CHIPS program to invest in critical areas of the semiconductor supply chain. It is not enough to build fabs here. We also have to make sure that the myriad of suppliers that support a successful fab are here too. And so in every program we design, we are including incentives to increase supply-chain resilience. I also see evidence that the private sector is changing the way they do business when it comes to supply chains. Consider the recent announcements by auto companies to launch partnerships with key suppliers of chips or batteries to try to increase supply-chain resilience. My view is that the collective recognition in the public and private sector that supply-chain resilience is a top priority will drive the creation of a new paradigm, leading to large investments of the kinds we would not have seen 10 years ago.

We are also very active on the workforce front. In every aspect of the CHIPS program, we are asking applicants to provide specific information about how they will create and sustain high-quality jobs, and how they will partner with local governments, community groups, and labor organizations to create a skilled pipeline of workers from every corner of America. More than 45 colleges across 17 states have launched semiconductor training programs, and a handful of innovative corporate partnerships have been launched. Our National Science Foundation is investing \$200 million in workforce, and a key component is making sure we have the right curriculum to prepare workers for the

jobs that are actually available. We need to get this right. Like the supply chain, we cannot be successful unless we get the workforce piece right.

Greenstein: This type of program raises the economic question about whether government policy pays firms to do what they would have done anyway, or whether it crowds out other investment and R&D activities. What do you think about those concerns?

Chatterji: This is an important concern that I think about every day. The good news is we have actually crowded-in over \$200 billion in private-sector investment in the American chip industry since the legislation was introduced. As we continue to implement this program, we will keep pushing to use taxpayer dollars as efficiently as possible and incentivize new investments to achieve our economic and national security goals.

I ALSO SEE EVIDENCE THAT THE PRIVATE SECTOR IS CHANGING THE WAY THEY DO BUSINESS WHEN IT COMES TO SUPPLY CHAINS.

Greenstein: Yet another question is, "Does the industry invest enough in R&D with a long-term payoff for the country?" But some long-term trends are hard to forecast. Fabless production has taken over much of the worldwide capacity, but not all of it. In addition, Moore's law has slowed and is estimated to end sometime this decade. What do you think about addressing long-term goals with such moving targets?

Chatterji: This is the hardest part of the job. First, you need a great team with financial, technical, public and private-sector experience. They have to be incentivized to look around corners and challenge existing assumptions about how the industry will evolve. I see evidence of this every day in my interactions with my colleagues at the government agencies implementing this program. Second, you have to design a program that can evolve as the industry does and make investments that can be evaluated at key milestones in case the strategy needs to change. Third, you have to set clear goals and focus on achieving what you can with available funds on the time frame that has been set out from the Act.

Greenstein: What have you liked the most about your job?

Chatterji: I feel fortunate to be at the center of one of the most innovative and ambitious economic and national security initiatives of our time, and to work with a team that shares a common vision for what we

can and must do. I try to remind myself of that each day, and that feeling is my favorite part of the job.

Greenstein: Which part of your academic work informed your efforts in government, and which part of your government work do you expect to take back to academics?

I WANTED TO UNDERSTAND HOW GOVERNMENT REGULATIONS IMPACTED BUSINESS STRATEGY AND HOW BUSINESSES INTERACTED WITH GOVERNMENT LEADERS.

Chatterji: I did my Ph.D. at the Haas School of Business in a program called *Business and Public Policy*. I went to this program because I had an interest in the intersection between the public and private sectors. I wanted to understand how government regulations impacted business strategy and how businesses interacted with government leaders. Like a lot of graduate students, I was surprised that, despite my lofty interests, the first two years of coursework was mostly

technical and theoretical. I eventually got to the phenomenon I was interested in, but not before doing a lot of tool building.

Looking back, it was the foundation at Berkeley that set me on a course to be in this position today, implementing a program at the intersection of business and public policy and leveraging my economics background, along with an understanding of technology policy and innovation. My academic career, including my time as faculty member at Duke University's Fuqua School of Business, has allowed me to develop the kinds of skills to break down complicated problems like the ones we try to solve every day in the CHIPS program and use the best available data and insight to answer them.

Returning to Duke, I will spend at least some of my time returning to my business and public policy roots. I want to better understand how these industrial policies around the world are changing the incentives for where and how businesses invest and what the impact of these investments is going to be. 🌍

SHANE GREENSTEIN is a professor with Harvard Business School, Boston, MA, 02163, USA. Contact him at sgreenstein@hbs.edu.

ADVERTISER INFORMATION

Advertising Coordinator

Debbie Sims
Email: dsims@computer.org
Phone: +1 714-816-2138 | Fax: +1 714-821-4010

Advertising Sales Contacts

Mid-Atlantic US, Northeast, Europe, the Middle East and Africa:
Dawn Scoda
Email: dscoda@computer.org
Phone: +1 732-772-0160
Cell: +1 732-685-6068 | Fax: +1 732-772-0164

Southwest US, California:
Mike Hughes
Email: mikehughes@computer.org
Cell: +1 805-208-5882

Central US, Northwest US, Southeast US, Asia/Pacific:
Eric Kincaid
Email: e.kincaid@computer.org
Phone: +1 214-553-8513 | Fax: +1 888-886-8599
Cell: +1 214-673-3742

Midwest US:
Dave Jones
Email: djones@computer.org
Phone: +1 708-442-5633 | Fax: +1 888-886-8599
Cell: +1 708-624-9901

Jobs Board (West Coast and Asia), Classified Line Ads

Heather Buonadies
Email: hbuonadies@computer.org
Phone: +1 623-233-6575

Jobs Board (East Coast and Europe), SE Radio Podcast

Marie Thompson
Email: marie.thompson@computer.org
Phone: +1 714-813-5094

Get Published in the New *IEEE Transactions on Privacy*

**This fully open access journal is
now soliciting papers for review.**

IEEE Transactions on Privacy serves as a rapid publication forum for groundbreaking articles in the realm of privacy and data protection. Be one of the first to submit a paper and benefit from publishing with the IEEE Computer Society! With over 5 million unique monthly visitors to the IEEE Xplore® and Computer Society digital libraries, your research can benefit from broad distribution to readers in your field.

Submit a Paper Today!


Visit computer.org/tp to learn more.



DEPARTMENT: ART ON GRAPHICS

Changing Aesthetics in Biomolecular Graphics

Laura A. Garrison , Mohn Medical Imaging and Visualization Centre, Haukeland University Hospital, 5021, Bergen, Norway and also Bouvet ASA, 5058, Bergen, Norway

David S. Goodsell , Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, 92037, USA and also Research Collaboratory for Structural Bioinformatics Protein Data Bank, Institute for Quantitative Biomedicine, Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, New Brunswick, NJ, 08903, USA

Stefan Bruckner , Institute for Visual and Analytic Computing, University of Rostock, 18059, Rostock, Germany

Aesthetics for the visualization of biomolecular structures have evolved over the years according to technological advances, user needs, and modes of dissemination. In this article, we explore the goals, challenges, and solutions that have shaped the current landscape of biomolecular imagery from the overlapping perspectives of computer science, structural biology, and biomedical illustration. We discuss changing approaches to rendering, color, human–computer interface, and narrative in the development and presentation of biomolecular graphics. With this historical perspective on the evolving styles and trends in each of these areas, we identify opportunities and challenges for future aesthetics in biomolecular graphics that encourage continued collaboration from multiple intersecting fields.

The structural biology community was an early adopter of computer graphics, driven by the need to visualize and explore the complex 3-D shapes of biological molecules such as proteins and DNA. This is no less true today, and structural biologists rely on an advanced suite of molecular graphics tools as a central part of their research pipeline, as well as for dissemination and outreach. Over the 50 or so years from the first biomolecular visualization to the rich graphics environment today, the aesthetics of biomolecular graphics have changed and matured, driven by multiple orthogonal demands. In the 1960s, the hardware was often difficult to access and software was limited, so the field of biomolecular visualization was limited to a small community of experts who developed software and provided access to methods for the

production of visual materials. Imagery was strongly influenced by the hardware: lines and points in the interactive Evans and Sutherland MultiPicture System, monochrome raster screens, and pen plotters. As consumer hardware continued to improve and the structural biology community became increasingly convinced of the utility of biomolecular graphics, an explosion of method development ensued, and the best techniques were adopted and made more user friendly. Throughout this period of development, the applications and modes of dissemination strongly shaped the aesthetics of the methods and the visualizations that were produced (Figure 1). In this article, we explore several aspects of visualization aesthetics and how the changing computer graphics environment has shaped the imagery that we see today.

RENDERING APPROACHES LEVERAGE THE STRENGTHS OF GRAPHICS HARDWARE

Rendering approaches for biomolecular data have evolved with advancements and expansions in

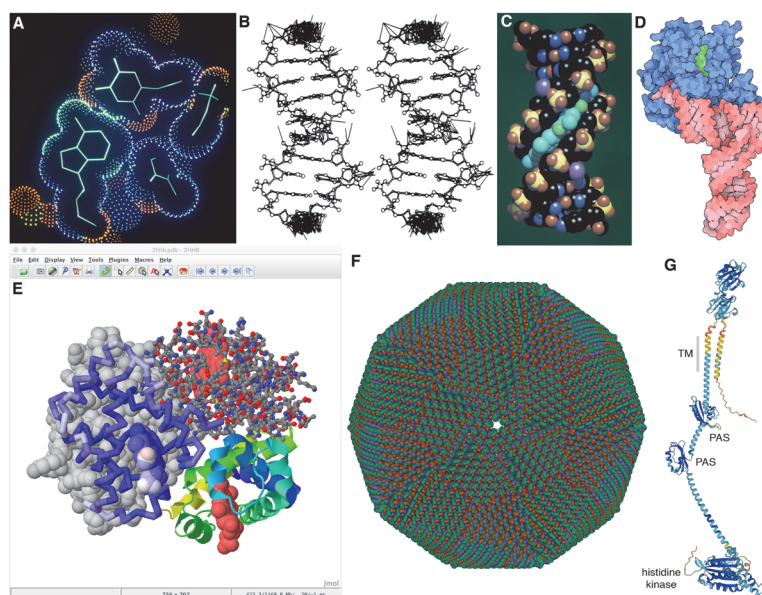


FIGURE 1. Biomolecular Graphics Development. Images created over a span of three decades. (A) Dot surface and wireframe for a complex of DNA with an inhibitor, displayed interactively on an Evans and Sutherland MultiPicture System with custom software (PDB ID 6bna). (B) Wireframe with extra lines to depict lattice contacts, created with a modified version of ORTEP. Printed using a pen plotter and presented in publication as a stereo pair for 3-D viewing (PDB ID 126d). (C) Raster space-filling image of DNA (atomic colors) and inhibitor (cyan) with Phong shading, created with several hours of computation with custom software (PDB ID 6bna). (D) Nonphotorealistic rendering of tRNA and elongation factor created with custom software (PDB ID 1ttt). (E) Natural language scripting of Jmol allows nimble, interactive access to multiple rendering options, as seen in this fanciful image of hemoglobin in four styles created with ten minutes of effort (PDB ID 2hhb). (F) Mol* is leading a new generation of web-based biomolecular viewers, here providing interactive exploration of faustovirus (PDB ID 5j7v), currently the largest structure in the Protein Data Bank. (G) Predicted structure of PleC, a protein involved in formation of a bacterial microdomain, with most confident regions in dark blue and least confident regions in orange (AlphaFold2 ID AF-P37894-F1).

capabilities of graphics hardware and software. Oak Ridge Thermal-Ellipsoid Plot Program^a (ORTEP), developed in 1965 out of Oak Ridge National Laboratory, dominated protein crystallography for many years due to widespread use of pen plotters to create publication-quality images. These images have a beautiful economy of line [Figure 1(b)].

With higher powered graphics hardware, more photorealistic rendering styles were increasingly used for biomolecular graphics. Photorealistic rendering techniques aim to mimic the look of real-life objects as closely as possible. Interestingly, biomolecular structures measure below the wavelength of visible light, so in some sense the term “photorealism” is, strictly speaking, a misnomer in this context. However, more advanced simulation of illumination serves to make biomolecular

structures possibly more relatable by embedding them in a larger space with light and shadow.

During the 1990s, amidst a large body of work in graphics on exploring nonphotorealistic rendering techniques which mimic the aesthetics of hand-drawn artwork, these approaches also became popular for depicting biomolecular structures. Toon shading, also known as cel shading, for instance, refers to a class of nonphotorealistic rendering approaches inspired by cartoons and comics and their use of shades and tints. This is a relatively inexpensive rendering style from a computational perspective, and is visually simplistic while conveying the information necessary to understand the structural arrangement of the molecules in question. Goodsell’s illustrative style is an attempt to capture some of the visual effectiveness of ORTEP images [see Figure 1(d)]. Today, toon shading is readily accessible in interactive graphics, and provided by, for example, Mol* and Protein Explorer.

^a[Online]. Available: <http://www.chem.ucla.edu/~harding/IGOC/O/ortep.html>

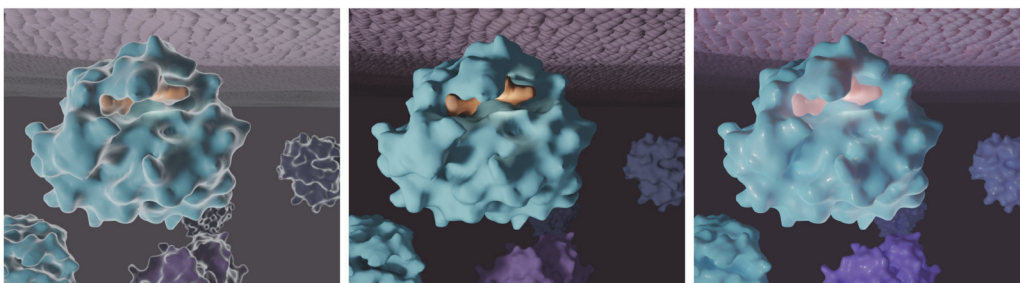


FIGURE 2. Demonstration of three common historical and contemporary biomolecular graphics rendering styles on a basic biomolecular scene with a ligand bound to a receptor on the interior of a membrane: (Left) “Scanning electron microscope (SEM)-look” material, (Middle) ambient occlusion (AO) layered with a matte material, and (Right) subsurface-scatter and clearcoat layers applied to a physics-based rendering material.

Other rendering approaches have also been developed to mirror the appearance of structures under a given acquisition method [i.e., scanning electron microscopy (SEM) (Figure 2, Left)]. This method applies a fresnel effect to a basic Phong shader, where structures appear to have a white outline that becomes thicker as the view angle becomes more oblique to the eye. This “SEM” shading technique has largely fallen out of fashion. However, when in use, the fresnel effect is usually more subtle (e.g., the CDC’s visualization of the COVID-19 virion, crafted by Alissa Eckert and Dan Higgins^b). Ambient occlusion (AO) in rendering output provides helpful depth cues to biomolecular structures and has become more prevalent with better and more efficient hardware and software for computer graphics (Figure 2, Middle). With ambient occlusion, a separate rendering pass calculates the exposure of each point on an object to ambient light. For example, the cavity of a molecule appears darker because it is more occluded from light. With improved hardware and software solutions enabling physically based rendering in most 3-D applications, computationally “expensive” features like global illumination, subsurface scattering, and clearcoat are more accessible and frequently used (Figure 2, Right). Molecules with this treatment can have a “gummy bear” appearance that is more editorial than strictly educational, and has been especially popular in editorial graphics, for example, in stylized illustrations for journal covers or in marketing materials for pharmaceutical products.

Today, advancements in science and technology pose an entirely new set of rendering challenges,

providing new capabilities and requiring new aesthetic decisions. Virtual reality and augmented reality are finally becoming a useful tool in research and education, driven by the increased affordability of consumer-level hardware. Currently, the need for seamless response during navigation imposes limitations on the complexity of VR scenes, requiring artists and developers to pare down the rendering options for objects being depicted. 3-D printing has also benefited from affordable options, now available to hobbyists and classrooms. Most commercial machines are limited to rigid, monochrome builds, but clever designs with snap-together parts and magnets are used to expand the types of stories that can be captured in the models. 3-D models also impose strict limitations on the types of representations that may be used, to ensure that the model is buildable and strong enough to handle.

RESEARCHERS ARE UTILIZING AN INCREASINGLY NUANCED APPROACH TO COLOR

Choice of color is arguably one of the most impactful decisions made during the creation of a biomolecular visualization. Coloring strategies can draw viewer attention to key features or molecules of interest, or encode physical or functional features of the molecule to aid in exploration and analysis, such as a ligand in a pathway or binding site of a receptor molecule. These perceptual tricks may include coloring key molecules in light or highly saturated colors that contrast with the surrounding molecules and environment. Here, the use of color is more often about drawing attention to the intended structures rather than encoding particular structural or functional properties of the molecules, for example, hydrophobicity.

^b[Online]. Available: <https://phil.cdc.gov/Details.aspx?pid=23311>

Early approaches to color were largely driven by technology and nascent traditions. Color printing was expensive and predominant technologies were pen plotters and monochrome or 8-bit raster screens. Much of early computer-generated imagery was produced and published in black-and-white. As color became increasingly feasible, published imagery was often colored using default settings, leading to a predominance of saturated colors. The CPK coloring scheme popularized by Linus Pauling and his physical molecular models [carbon black, oxygen red, nitrogen blue, hydrogen white, Figure 1(c)], was the de facto standard for most research graphics. Interestingly, developers immediately encountered a problem as color interactive hardware became available: how to deal with black carbons on a black screen. The most common approach at the time, which was still quite wedded to saturated color, was to use green.

As color screens became more common in laboratories, an explosion of experimentation followed, leading to multiple color palettes for specific needs. Some of these are now only rarely used, such as amino-acid-specific colors based on the “Shapely” physical models, whereas others showed widespread utility and are provided as hard-wired options in most current packages. These include coloration of entire biomolecular subunits [see Figure 1(d)], coloring of properties such as electrostatics and hydrophobicity, and coloration that highlights local structural features such as protein secondary structure or position in a polymer chain [see Figure 1(e)].

Today, we enjoy an environment that is filled with options for coloring, and a great freedom to customize coloring based on our personal ideas and preferences. Current molecular graphics software typically provide a menu of traditional options with more-or-less standardized color choices, paired with flexible methods for selecting atom sets and assigning custom colors to them. Given this ability, color is frequently used, especially by biomolecular animators and designers, to evoke a variety of moods or feelings when viewing a molecule or scene, and the choice of palette may be tuned to appeal to the intended audience and use case. Current research in coloring aesthetics is layering advanced capabilities on these basic coloring choices. For example, the Viola laboratory is exploring methods to transition smoothly from molecule-based coloring to atomic coloration in multiscale systems as viewers transition from whole-cell views to individual molecules (Figure 3).

We have observed a growing interest in the biomolecular visualization community to standardize coloring schemes. This has great advantages. For example,

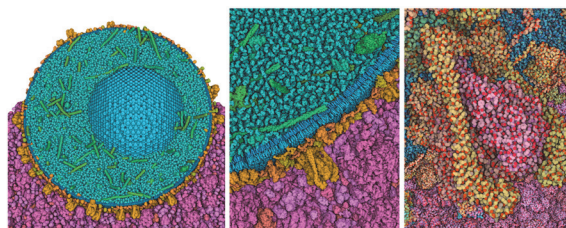


FIGURE 3. Adaptive multiscale representation and coloring.

Model of an insulin secretory granule (blue, green, and orange) and cytoplasm (magenta) is displayed with subunit colors and coarse surfaces at left. As the user zooms in, the view progressively changes to a full atomic representation with atomic color (right). Images by Ludovic Autin.

it would greatly aid with issues of accessibility, by promoting the availability and use of color-blind-friendly palettes. Standards also unify a field. For example, the CPK coloring scheme, given its excellent provenance and widespread use, is instantly recognizable by most viewers, allowing facile comparison when viewing figures from multiple labs. The confidence coloring scheme currently used in AlphaFold2 structures [see Figure 1(g)] is rapidly becoming a similar de facto standard. Standards, however, can only codify current knowledge and thus may potentially inhibit creative exploration as the field of structural biology continues to grow. Current molecular graphics tools typically express both of these views, providing turnkey methods to apply the (currently) most useful standards while also streamlining the ability to develop customized palettes.

BIOMOLECULAR IMAGERY HAS GROWN TO ENCOMPASS LARGER NARRATIVES AND PERSONAL STYLES

Biological stories are growing larger to span new experimental results from atoms to cells, and visualization options provide myriad opportunities for building new and effective visual explorations. When designing and executing these stories, we are always faced with three orthogonal challenges: 1) technical capabilities of turning data into images, 2) the needs of the intended audience, and 3) our own personal artistry. Much of the early history of biomolecular visualization was centered around the creation of figures for research publications. These were effectively “molecular portraits” that presented the structure and, hopefully, some aspects of their function. These portraits were created with a handful of programs, and most

often did not stray far from the default coloring and rendering options.

Today, the audiences for biological stories have expanded. Education and more general outreach, e.g., public exhibitions, have pushed content authors to incorporate narrative devices to make molecules and their environments more engaging and comprehensible to broader audiences. These tend to require artistic expertise and knowledge of advanced 3-D software, such as Blender, Autodesk Maya, or Maxon Cinema 4-D. Pharmaceutical company growth and marketing initiatives for new drug developments have helped drive cinematic storytelling approaches to biomolecular reactions and pathways. New software solutions, like Moleculumentary,¹ allow content authors to semiautomatically create tailored narratives to disseminate scientific content. Tools like this enable users without deep expertise in 3-D animation and design to create narratives for education and outreach.

The narratives themselves have also grown larger, with an explosion of new Big Data and methods for accessing and visualizing these data. Continued advances in structure determination methods, such as the current resolution revolution in cryoelectron microscopy, and recent advances in protein structure prediction, such as AlphaFold2 and RoseTTAFold, are radically increasing the number and complexity of atomic structures that are available for detailed depiction of biomolecular structures. Visualization methods are faced with depiction of larger and larger datasets, interactively, on the web. We no longer can limit ourselves to presenting only a basic portrait of a molecule. We now need to explore uncertainty in predicted structure models [see Figure 1(g)] and the dynamic aspects of single molecules, assemblies, and ensembles. Biomolecular dynamics simulations capture a vast amount of information, where only a few time slices may be of interest to the viewer. Choosing how to display this information in a digestible way requires multiscale thinking in both time and space. Connections to sequence and functional annotations need to be at our fingertips to explore bioinformatics data related to our subjects.

The current software and hardware environment for biomolecular visualization is robust and provides nimble opportunities for building a personal aesthetic within this data-rich environment. Figure 4 includes four pioneering artists who have helped shape current trends in biomolecular graphics design and animation. Drew Berry pioneered a cinematic style combining dynamic scene design, a unique approach to biomolecular motion, and immersive sound in his animations for general audiences. Gaël McGill has perfected an

editorial style with design decisions that produce arresting images for textbook and commercial applications. For example, the image in Figure 4 has circulated for a decade on social media as “the most detailed depiction of a cell.” Janet Iwasa creates data-heavy imagery in collaboration with researchers for use in publication and presentations, and has developed a direct style that is true to the science. Veronica Falconieri Hays creates dynamic portraits of molecules using contemporary methods for scientific illustration, often framed within larger stories that show their cellular context.

OUTLOOK AND CHALLENGES

Molecular graphics is mature, but still offers ample opportunities for research and software development in graphics and visualization, to address current limitations and challenges. Amazingly, most widely used tools with a significant visualization component still focus on fairly basic rendering/aesthetic approaches, and the vast majority of users will employ the default rendering and coloring when using these tools. As the field of biomolecular graphics becomes more mature with tools for creating content more widely available, our community could consider careful attention to defaults and presets provided by common methods. Additional guidance could support users with limited artistic skills to navigate the vast design space when creating biomolecular graphics. This is particularly important, as methods and tools are increasingly used by diverse user communities outside of structural biology, which also necessitates more extensive conventions that facilitate our understanding of molecules and their environments. To address these challenges, programs like Chimera and Mol* currently provide one-click preset views for different applications. The VIS and graphics community have begun building dedicated applications (such as MegaMol, CellPAINT, and Marion) that are more oriented to the needs of artists and are more specialized than generic graphics applications like Blender and Maya.

Accessibility of images is also becoming a stronger design specification when creating imagery, particularly in settings for nontechnical audiences. Colorblind friendly color palettes have emerged as a low-hanging fruit, and visualization tools are increasingly adopting default options that are distinguishable to those with limited vision. While this is a good first step, we need to do more to make biomolecular graphics accessible to segments of the population. Sound engineering can help engage

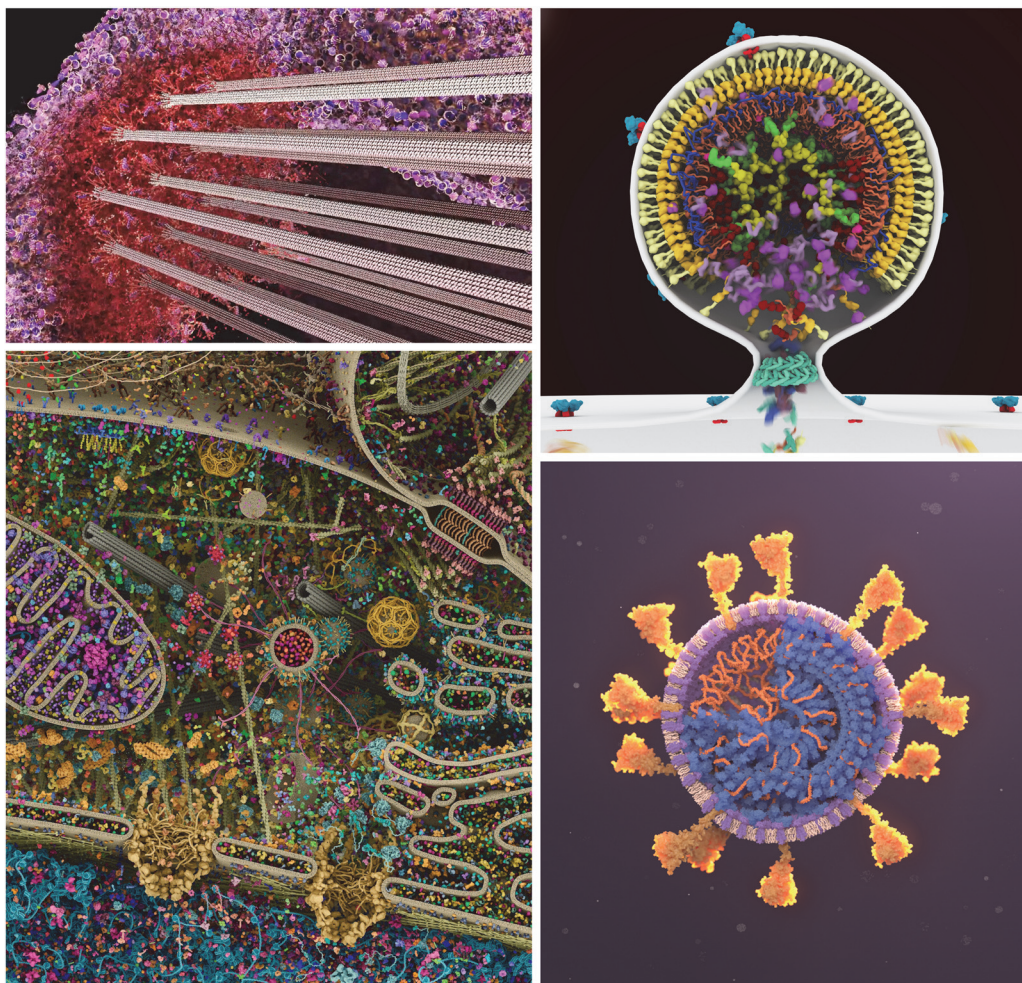


FIGURE 4. Large-scale narratives and personal aesthetics. Work from four contemporary artists shows state-of-the-art design decisions in complex biomolecular systems. Notice in each case how the approach to color and lighting is tuned for the audience, and representations are chosen to depict complexity at a level appropriate for the scene. (Upper left) Drew Berry creates groundbreaking video animations for general audiences with station WEHI.TV, such as this still image from a video animation of the kinetochore. (Lower left) Gaël McGill of Digizyme and Harvard Medical School creates dynamic editorial images and animations for textbooks and commercial clients, such as this complex image of the interior of a living cell created with Digizyme team member Evan Ingersoll. (Upper right) Janet Iwasa is a pioneer in the creation of data-rich animations for dissemination of information in research settings, such as this animation of budding of HIV-1 from an infected cell. (Lower right) Veronica Falconieri Hays is a Certified Medical Illustrator who creates captivating editorial and educational imagery and animations with beautiful interplay of color and light, as in this SARS-CoV-2 illustration.

and immerse viewers in an environment, but this can go further to encode real meaning to molecules for those with low or limited vision. Thinking about how to integrate audio and other sensory modalities to enhance the accessibility of biomolecular graphics—sound, haptics, physical models—can be critical to advance public understanding of molecules and their relevance in society.

There are still abundant areas that would benefit from creative development. For example, on the data-in side, cryo-EM is currently providing structural views of biomolecular assemblies of unprecedented complexity. Often these structures have a hierarchical organization, with multiple biologically relevant entities in one assembly. Currently, this hierarchical structure is difficult to define within existing graphics

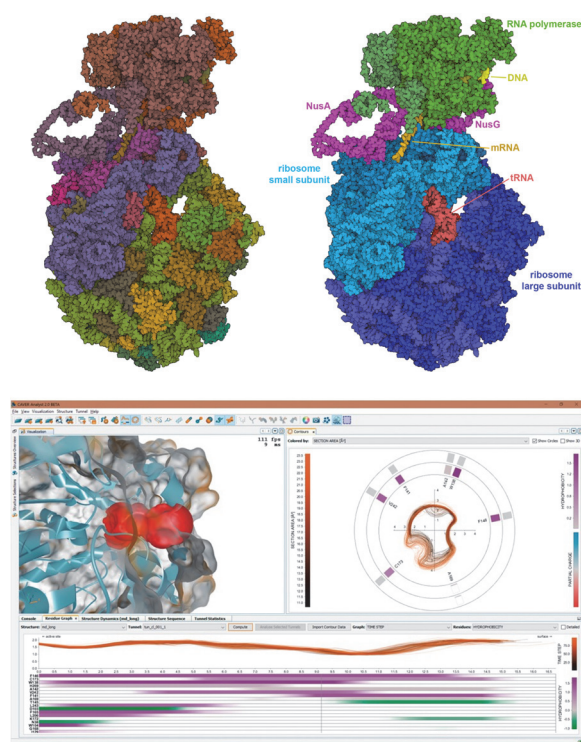


FIGURE 5. (Top) Hierarchical nature of large biomolecular assemblies. The expressome (PDB ID 6x9q) includes RNA polymerase and a ribosome tethered together by two processivity factors, NusA and NusG. This cryo-EM structure also includes a small fragment of DNA, a messenger RNA, and two tRNA molecules. Both figures are rendered interactively with Mol* at the RCSB PDB website, with user-tunable options for ambient occlusion, outlines, and flat shading. The left figure is colored using a default option based on chain instances in the coordinate file. The right figure is colored manually by selecting individual chains and choosing colors to highlight the structural hierarchy of biologically relevant subassemblies. *(Bottom) CAVER Analyst 2.0 software application.*² This system enables real-time visualization of tunnels and channels in biomolecular structures. The upper left panel depicts a molecular structure with a highlighted region of interest, in this case, a tunnel, with the accompanying right panel showing a cross-cut contour of the highlighted tunnel. Each contour indicates a time step from the underlying long MD trajectory. The bottom panel depicts the tunnel radius in profile along its length, with variation over time (each line is one time step) and the amino acids that form the boundary of the tunnel, ranked according to their influence on the tunnel's boundary. A consistent color design is used throughout to help researchers understand connections between data in the separate panels.

methods, requiring laborious manual selection and coloring of individual chains (Figure 5, Top). A more facile connection to functional annotations will be needed to address this problem. Another challenge lies in visual depiction of the underlying uncertainty of the model(s) responsible for the biomolecular graphic, and displaying the provenance of the data. While color-coding areas of confidence is a common approach to the former [see Figure 1(g)], aesthetic and easily identifiable display of such information is by no means a solved problem. Data provenance for dynamic data is another area of great need and current creative effort. Successful current approaches often use a dashboard-style visualization, where a surface model of the molecule is paired with other views that show key aspects of the structure and underlying data at each time step of the simulation (Figure 5, Bottom). On the data-out side, the fields of virtual reality and 3-D printing are still very much the Wild West, and creative approaches to aesthetics and design can lead quickly to effective results.

Molecular graphics, now and throughout its history, has been a multidisciplinary effort bringing together the talents of molecular biologists, computer scientists, and artists to build methods and imagery. This collaboration is continuing to expand to encompass growing fields of knowledge, for example, leveraging expert annotations from bioinformatics, best practices from perceptual science and science historians, and direct user feedback from educational evaluation experts. 🧑🏫

ACKNOWLEDGMENTS

This work was supported in part by the National Institutes of Health under Grant GM120604, in part by RCSB Protein Data Bank National Science Foundation under Grant DBI-1832184, in part by the National Institutes of Health under Grant GM133198, and in part by the U.S. Department of Energy under Grant DE-SC0019749.

REFERENCES

1. D. Kouřil et al., "Moleculumentary: Adaptable narrated documentaries using molecular visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 29, no. 3, pp. 1733–1747, Mar. 2023.
2. A. Jurčík et al., "Caver analyst 2.0: Analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories," *Bioinformatics*, vol. 34, no. 20, pp. 3586–3588, 2018.

3. A. J. Olson, "Perspectives on structural molecular biology visualization: From past to present," *J. Mol. Biol.*, vol. 430, no. 21, pp. 3997–4012, 2018.
4. X. Martinez et al., "Molecular graphics: Bridging structural biologists and computer scientists," *Structure*, vol. 27, no. 11, pp. 1617–1623, 2019.
5. B. Kozlíková et al., "Visualization of Biomolecular structures: State of the Art revisited," *Comput. Graphics Forum*, vol. 36, no. 8, 2017, pp. 178–204.
6. G. T. Johnson and S. Hertig, "A guide to the visual analysis and communication of biomolecular structural data," *Nature Rev. Mol. Cell Biol.*, vol. 15, no. 10, pp. 690–698, 2014.
7. S. I. O'Donoghue et al., "Visualization of macromolecular structures," *Nature Methods*, vol. 7, no. Suppl3, pp. S42–S55, 2010.
8. T. D. Goddard and T. E. Ferrin, "Visualization software for molecular assemblies," *Curr. Opin. Struct. Biol.*, vol. 17, no. 5, pp. 587–595, 2007.
9. D. S. Goodsell, "Visual methods from atoms to cells," *Structure*, vol. 13, no. 3, pp. 347–354, 2005.
10. J. H. Iwasa, "Bringing macromolecular machinery to life using 3D animation," *Curr. Opin. Struct. Biol.*, vol. 31, pp. 84–88, 2015.

LAURA A. GARRISON is a visualization researcher and designer affiliated with Bouvet ASA, 5058, Bergen, Norway, and the Mohn Medical Imaging and Visualization Centre, Haukeland University Hospital, 5021, Bergen, Norway. Garrison received

her Ph.D. degree in visualization from the University of Bergen, Bergen, Norway, for her work on multiscale visualization of human physiology. Her other research interests include human–computer interaction, accessibility in visualization, and visual data analysis. Contact her at laura.garrison@uib.no.

DAVID S. GOODSSELL divides his time between computational biology research and science outreach. His work in biomolecular visualization at Scripps Research includes modeling and artistic depiction of the cellular mesoscale and development of non-photorealistic rendering methods for molecular and cellular subjects. He creates outreach materials for the RCSB Protein Data Bank, including a popular monthly column that presents molecular structure and function for general audiences. Contact him at goodsell@scripps.edu.

STEFAN BRUCKNER is a professor with the University of Rostock, 18059, Rostock, Germany, where he heads the Chair of Visual Analytics. His research interests include all aspects of data visualization, with a particular focus on interactive techniques for the exploration and analysis of complex heterogeneous data spaces. He is the corresponding author of this article. Contact him at stefan.bruckner@uni-rostock.de.


Contact department editor Bruce Campbell at bcampbel01@risd.edu or department editor Francesca Samsel at fsamsel@tacc.utexas.edu.



www.computer.org/cga

IEEE Computer Graphics and Applications bridges the theory and practice of computer graphics. Subscribe to CG&A and

- stay current on the latest tools and applications and gain invaluable practical and research knowledge,
- discover cutting-edge applications and learn more about the latest techniques, and
- benefit from CG&A's active and connected editorial board.

 **IEEE COMPUTER SOCIETY**

 **IEEE**

DEPARTMENT: MEMORY AND STORAGE

The DNA Data Storage Model

Dave Landsman , Western Digital CorporationKarin Strauss , Microsoft Corporation

Reliably storing digital data in synthetic DNA fits naturally into the layered Open Systems Interconnect model and shares many parallels with reliably storing data using existing storage technologies and interfaces.

DNA data storage, or using synthetic DNA as a data storage medium, is being seriously considered as an archival storage solution due to its volumetric data density potential, data retention characteristics, sustainability, and potential for dramatically lower total cost of ownership versus existing storage technologies.

INTRODUCTION

The biotechnology industry has made DNA data storage possible today due to decades of investment in molecular-level technologies for medical and life sciences applications that now enable us to construct and read synthetic DNA, base by base. These fundamental capabilities make it possible to encode digital data into a sequence of bases (adenine, guanine, cytosine, and thymine, or AGCT), write that sequence as a set of corresponding DNA molecules (synthesis), store the molecules, prepare them for reading (retrieval), read them back as a sequence of bases (sequencing), and finally, decode the original digital data (Figure 1). To learn more about this process, see *Preserving Our Digital Legacy: An Introduction to DNA Data Storage*.¹

Even though synthetic DNA as a data storage medium is similar to traditional storage media in many ways, it is worth highlighting some key differences.

First, in traditional storage, the media is premanufactured (e.g., SSDs use NAND cells, HDD, and tape use magnetic domains on a platter or strip, respectively) and written by modifying the state of the media. For

such devices, capacity and throughput scaling require modifications to both media and write/read heads. In contrast, the most common DNA data storage method does not employ a premanufactured media substrate. (Note: Methods to attach DNA to a planar substrate, to serve as a “memory/storage cell”, are being considered; this article does not cover these methods.) Instead, the storage media—DNA molecules—are manufactured during write operations. In this case, DNA’s universal, fixed, and reader/writer independent physical structure enables throughput improvements without media changes, or data migration when adopting new writer/reader generations.

Second, because DNA media is detached from any array-based substrate, protocol information such as object identifiers and segment indices must be embedded within each DNA molecule in the DNA archive, for locating objects, object segmentation, etc. Despite this overhead, DNA can achieve much higher volumetric density as detached media than as array-based media.

Third, DNA has unique error characteristics as a medium for end-to-end storage. For example, in addition to substitutions (akin to bit flips), insertions and deletions may also occur. Encoders try to avoid certain sequences^{2,3,4} to reduce interference with the writing and recovery process.

Despite the uniqueness of synthetic DNA as a storage medium, there are many parallels with traditional data storage and storage interface mechanisms. In this article, we draw parallels between the DNA data storage model and the Open System Interconnection (OSI) model (Figure 2). Before we begin describing the DNA data storage layer model, we describe a few things about DNA.

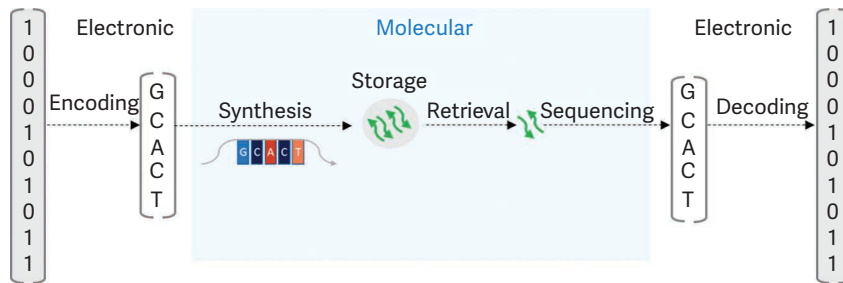
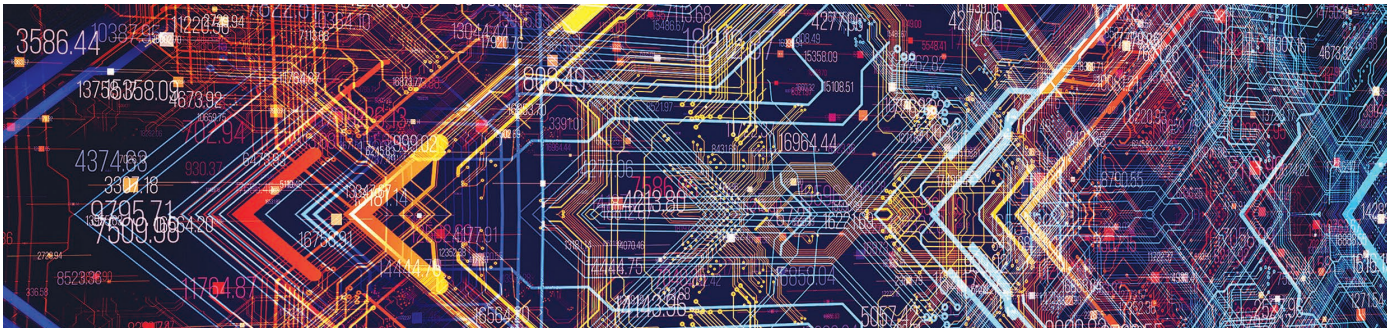


FIGURE 1. The DNA data storage system. Encoding and decoding are performed in the electronic domain and translate bits to bases and vice versa. Synthesis and sequencing are the interfaces to and from the molecular domain, creating and reading DNA sequences.

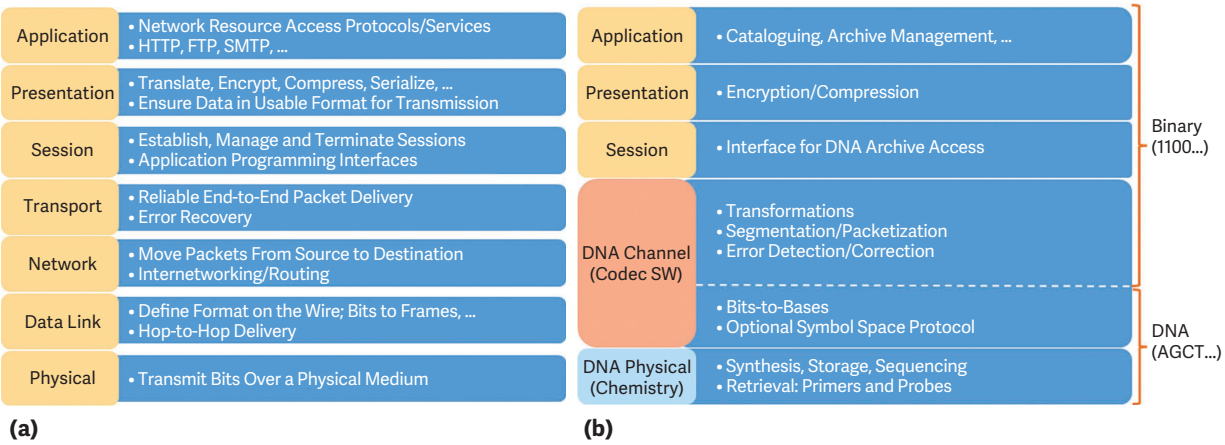


FIGURE 2. The layers of (a) the OSI model and (b) the DNA data storage model.

DNA MECHANICS IN BRIEF

We are most familiar with DNA as a “double-helix” (Figure 3), or dual-stranded DNA (dsDNA), where the base adenosine on one strand has a chemical binding affinity (complementarity) with the base thymine on the other strand, and the base cytosine has an affinity with the base guanine. Complementarity is used in nearly all of the techniques employed in DNA data storage, in the process called hybridization (Figure 4).

In organisms, cell division naturally replicates the genetic code. Cellular mechanisms separate the two

strands of the original dsDNA into two single-stranded DNA strands (ssDNA) and create two new dsDNA molecules from them, effectively copying the cell’s genetic information. Both ssDNA and dsDNA are used in different parts and applications of the DNA data storage pipeline.

Note that no organisms or cells are used for DNA data storage: synthetic DNA for data storage is constructed and manipulated through well-controlled chemical processes, covered in the section “The DNA Data Storage Physical Layer.”

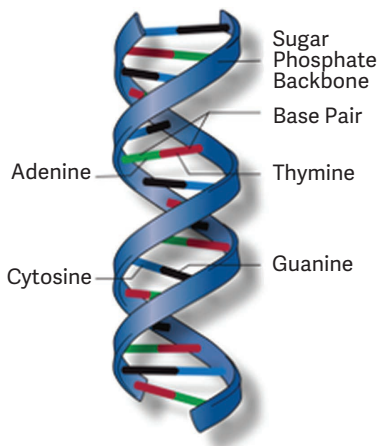


FIGURE 3. The DNA double helix. (Source: National Human Genome Research Institute; used with permission.)

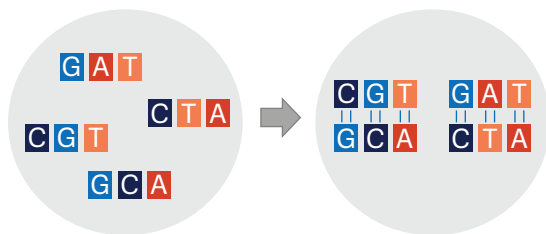


FIGURE 4. Hybridization: A process in which complementary bases bind to each other.

APPLICATION, PRESENTATION, AND SESSION LAYERS

The upper three layers of the OSI model map to the DNA data storage model effectively unchanged in function; at this level of abstraction, DNA is simply another storage medium.

The application layer is the interface closest to users. For example, since DNA data storage is best matched to long-term storage, the application layer is likely to define how data are logically organized for archival purposes, including metadata describing the data in the archive.

The presentation layer maps naturally to the preparation of bitstreams as input to the lower layers. It may include transformations such as encryption and compression. Like with other media, such functions may be accelerated by special-purpose (silicon-based) hardware support.

The session layer provides access to the DNA data storage interface. Various implementations are possible, but the most commonly discussed is object based. In this implementation, the interface provided by the session layer is basically an object store with a key-value schema. It offers basic primitives such as read/write of individual objects or all objects, as well as primitives with more complex semantics, such as indexed search. These commands translate to logical and physical storage operations at lower layers.

THE DNA DATA STORAGE CHANNEL LAYER

The DNA channel layer takes as input a bitstream or other digital object from the session layer and processes those bits to ensure that the DNA sequences written and read by the physical layer can be successfully decoded, enabling recovery of the original digital source data. The DNA channel layer is implemented as a software codec.^{2,3,4,5,6,7,11}

The DNA channel [Figure 5(b)] shares conceptual characteristics with a more “traditional” network/electrical channel [Figure 5(a)]. The DNA channel layer roughly incorporates the functionality of the transport through the data link layer in the OSI model. It receives a bitstream from the session layer, preparing it for handoff to the lower layers for the conceptual equivalent of “transmission,” which, in the DNA case, means writing, storing, recovering and reading DNA molecules. As there is no physical “wire” in DNA data storage, the data link layer functionality is embedded in the DNA codec, instead of in the transmitter/receiver pair (Figure 5, dashed lines).

When the bitstream is presented to the DNA channel layer, the following types of operations are performed, not necessarily in this order.

Packetization: In a network or fabric, the key function of the transport and network layers is packetization, routing, and flow control across the link. Neither routing nor flow control are relevant here, but packetization is. The longest strand of synthetic DNA that can currently be constructed base by base with standard chemistry, while maintaining sufficient accuracy, is around 300 bases. Even if every base in a strand could be used for payload data, a typical strand would encode tens of bytes. Thus, like in traditional networking,

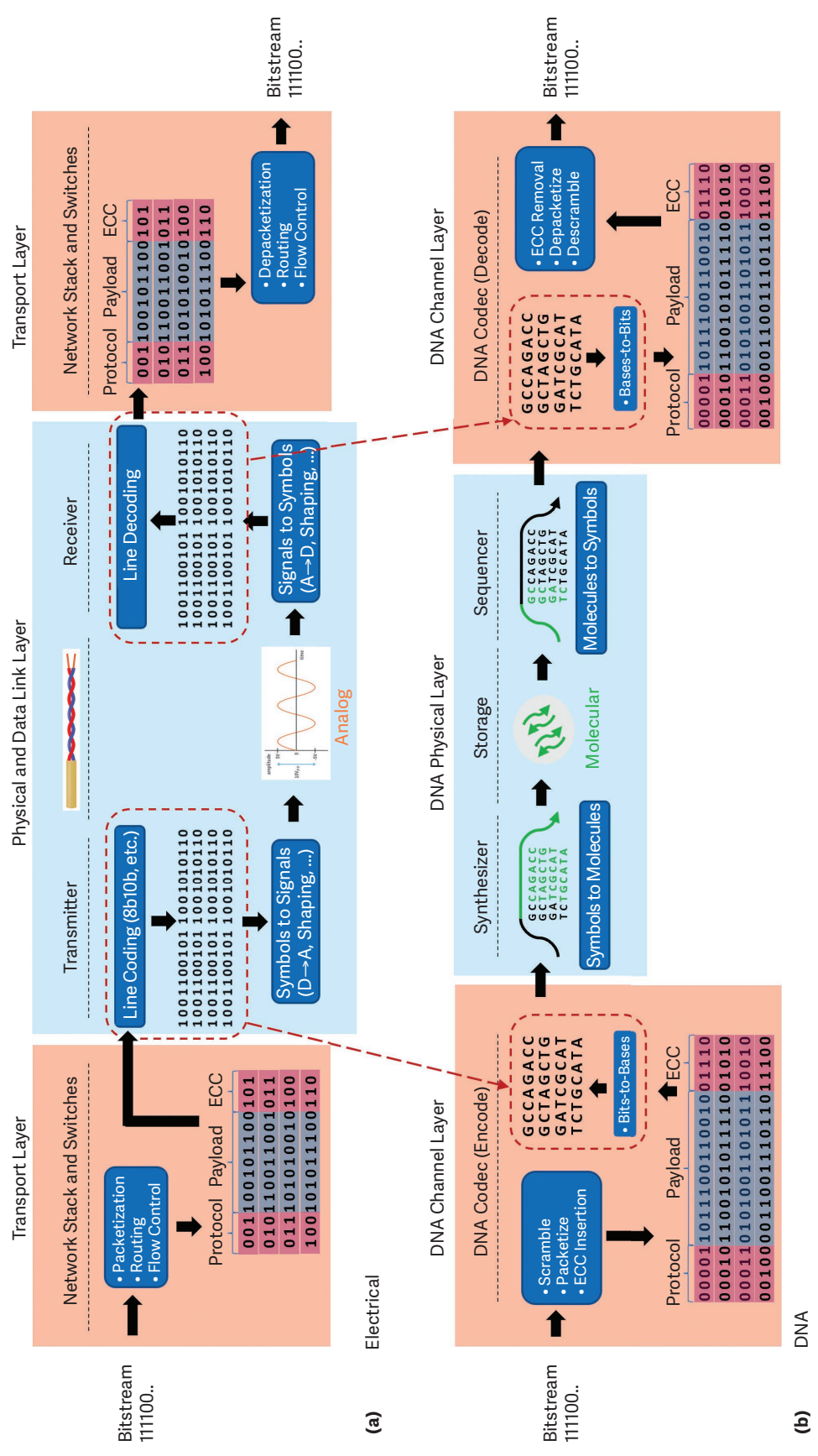


FIGURE 5. Comparing (a) the electrical channel model with (b) the DNA channel model.

where bitstreams must be broken into smaller pieces to fit limited-size packets, DNA data storage requires breaking data objects into many segments during encoding to fit limited-size strands. Reassembling these segments in the right order when sequencing requires adding indices to each segment before synthesis. The need for segment indices (and other protocol fields) introduces a tradeoff between the number of segments used to represent an object in a DNA pool (that is, maximum object size) and the number of bases used for the segment index (that is, index overhead).

IN SUMMARY, THE DNA CHANNEL LAYER ENCODES AN INPUT BITSTREAM SO THAT THE DNA SEQUENCES THAT ARE SENT TO THE PHYSICAL LAYER CAN BE EFFECTIVELY AND SUCCESSFULLY DECODED AFTER "TRANSMISSION."

Error correction: In constructing segments for storage, the DNA codec must add redundant information for error correction because random errors (base insertions, deletions, and substitutions) and erasures (missing sequences) may affect such segments through the DNA physical layer. Error correction can augment segments with additional data (inner code) or add segments that contain additional data (outer code). After the DNA is sequenced, the DNA codec uses this redundant information to recover lost sequences, correct remaining errors, and reliably deliver the original bitstream back to the upper layers.

Translation: To store digital data in DNA, the data must be translated from digital (0 or 1) to bases (A, T, C, or G). Although it is possible to simply map every two bits in a bitstream into one of the four possible bases, there are advantages to other mappings. For example, mappings of longer bitstreams to longer base sequences may better accommodate certain error correction methods and be more space efficient.

Transformations: With DNA, patterns of repeated bases (homopolymers), a high proportion of Gs and Cs (high GC content), and some other specific patterns

can cause errors. For example, some sequencing techniques exhibit errors with long homopolymers, and high GC content may affect sequencing preparation protocols. The transformations in the DNA channel layer avoid transforming digital bitstreams into such problematic patterns of bases, thereby reducing downstream errors and associated error correction overhead, and improving overall data reliability. This is conceptually similar to the actions in a network/electrical channel to mitigate analog effects. For example, a serial link maintains close to an equal number of ones and zeroes on the wire.

Optional DNA space protocol: In addition to the actions above, the codec may insert additional base sequences into the already transformed digital data (process not shown in Figure 3). The purposes are use-case specific but generally involve random access to or search within the DNA archive. An example is the addition of an object ID, assigned by the session layer to sequences belonging to the same object. The section "Retrieval" discusses this further.

After these steps, the resulting DNA sequences are handed over to the DNA physical layer.

In summary, the DNA channel layer encodes an input bitstream so that the DNA sequences that are sent to the physical layer can be efficiently and successfully decoded after "transmission" (synthesis, storage, retrieval, and sequencing). In doing this, the DNA channel uses transformations and processing steps that are well structured and similar in nature to those used in existing storage systems. In the next section, we discuss the DNA physical layer.

THE DNA DATA STORAGE PHYSICAL LAYER

The DNA physical layer [Figure 2(b)] involves the basic chemistry of DNA for writing (synthesis), physically preserving (storage), and reading (sequencing) DNA molecules. It also involves the basic chemistry of DNA for retrieval, which is both a general preparation step for sequencing and also an integral step in how systems implement DNA storage operation requests from the upper layers. We will first briefly discuss synthesis, storage, and sequencing and then cover retrieval as it is central to the functioning of logical storage operations for DNA data storage.

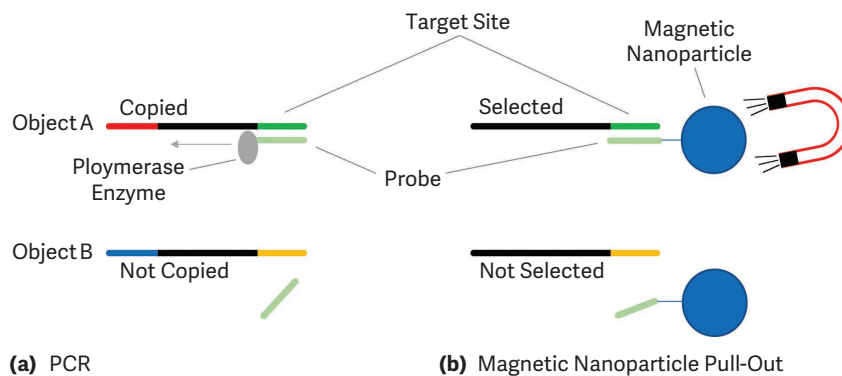


FIGURE 9. Two random access methods: (a) Probe attaches to the target site so that only the DNA sequences with the Object ID are copied during PCR. (b) Probe attaches to the target site of sequences with the Object ID, and magnetic nanoparticles used for molecular pull-out with a magnet.

Storage

After being written and before being read, the DNA molecules need to be stored in an environment that prevents them from degrading.^{9,10,11} While a wide variety of preservation methods are common in biotechnology broadly, DNA used for data storage is typically stored dry and in a chemically inert environment to increase storage density and stability. Significantly, with such methods, DNA for data storage shows potential for extremely long endurance at room temperature, supporting reliable long-term storage that is low cost and sustainable. We expect that large pools of DNA molecules (say, terabytes of storage) will be organized into even larger libraries where pools will be addressed using a physical coordinate system,¹² like tape libraries today.

Retrieval: Probes for storage operations

Retrieval defines how DNA molecules are extracted from a DNA archive and prepared for sequencing. The most basic operation is to read the entire archive, which is equivalent to reading all data in a tape from beginning to end. However, when a pool of DNA contains multiple objects, random access operations (for example, “seeking” the location of an object or searching for sets of objects) need to be performed on the pool.

There are a variety of methods to implement such operations, but two popular methods are PCR^{3,13,14} and DNA pull-out with magnetic nanoparticles¹⁵ (Figure 9).

PCR requires the session layer to assign a set of object IDs, which, at encoding time, the channel layer

appends as a set of predefined DNA base sequences to both ends of every DNA sequence belonging to an object in the archive; these base sequences are referred to as target sites. At retrieval time, probes, that is, short DNA sequences that perfectly complement the target site of one or more objects, attach to the target molecules (recall that A pairs with T, and C with G) and, along with special enzymes called polymerases, kickstart the PCR process. Over multiple PCR cycles, the attached probes (called primers when used in PCR), enable the polymerases to copy only the target DNA molecules, making them more abundant than molecules representing other objects. The result is a pool in which most of the DNA molecules represent the object(s) to be read.

DNA pull-out with magnetic nanoparticles uses the same principle of DNA attachment but a different method to make the molecules representing the object of interest more numerous in a solution. With this process, the channel layer only needs to append the object IDs as target sites at one end of every DNA sequence belonging to an object. At retrieval time, the probes are attached to magnetic nanoparticles. As the probes bind to target DNA molecules, the target DNA molecules can then be separated from the rest of the DNA molecules in the pool with a magnet.

For either random access method, additional preparation steps such as further PCR steps and DNA cleanup may be required, depending on the sequencer used for reading the information out.

Tradeoffs between different chemical implementations of object random access along dimensions

such as reliability, preparation, and reading overheads are an active area of research for DNA data storage,¹⁶ as well further work on implementing even more advanced operations such as search,¹⁷ content similarity search,¹⁸ file preview,¹⁹ etc.

This discussion has shown how, for the same read-object request from the session layer, different mechanisms can be used at the lower channel and physical layers, similar to how two NAND flash storage devices are implemented differently despite using the same command interface.

While DNA as a storage medium has fundamental differences from traditional storage, many of the data transformations and error processing considerations for DNA data storage have analogies to transmitting data through “traditional” network/storage electrical channels. It is our hope that framing DNA data storage implementation methods in the OSI layered storage model will help the nascent DNA data storage ecosystem evolve. 🌱

ACKNOWLEDGMENT

We thank Kyle Tomek, John Hoffman, Damien Le Moal, and Luis Ceze for their valuable contributions and feedback on this manuscript. We also thank our respective research teams and collaborators for their substantial contributions to this field. Dave Landsman is the corresponding author.

REFERENCES

1. “Preserving our digital legacy: An introduction to DNA data storage,” DNA Data Storage Alliance. [Online]. Available: <https://dnastoragealliance.org/dev/wp-content/uploads/2021/06/DNA-Data-Storage-Alliance-An-Introduction-to-DNA-Data-Storage.pdf>
2. N. Goldman et al., “Towards practical, high-capacity, low-maintenance information storage in synthesized DNA,” *Nature*, vol. 494, no. 7435, pp. 77–80, Feb. 2013, doi: 10.1038/nature11875.
3. L. Organick et al., “Random access in large-scale DNA data storage,” *Nature Biotechnol.* vol. 36, no. 3, pp. 242–248, Mar. 2018, doi: 10.1038/nbt.4079.
4. S. R. Srinivasavaradhan, S. Gopi, H. D. Pfister and S. Yekhanin, “Trellis BMA: Coded trace reconstruction on IDS channels for DNA storage,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2021, pp. 2453–2458, doi: 10.1109/ISIT45174.2021.9517821.
5. G. M. Church, Y. Gao, and S. Kosuri, “Next-generation digital information storage in DNA,” *Science*, vol. 337, no. 6102, p. 1628, Sep. 2012, doi: 10.1126/science.1226355.
6. W. H. Press, J. A. Hawkins, S. K. Jones Jr., J. M. Schaub, and I. J. Finkelstein, “HEDGES error-correcting code for DNA storage corrects indels and allows sequence constraints,” *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 31, pp. 18,489–18,496, Aug. 2020, doi: 10.1073/pnas.2004821117.
7. Y. Erlich and D. Zielinski, “DNA Fountain enables a robust and efficient storage architecture,” *Science*, vol. 355, no. 6328, pp. 950–954, Mar. 2017, doi: 10.1126/science.aaj2038.
8. B. H. Nguyen et al., “Scaling DNA data storage with nanoscale electrode wells,” *Sci. Adv.*, vol. 7, no. 48, Nov. 2021, Art. no. eabi6714, doi: 10.1126/sciadv.abi6714.
9. D. Coudy, M. Colotte, A. Luis, S. Tuffet, and J. Bonnet, “Long term conservation of DNA at ambient temperature. Implications for DNA data storage,” *PLoS One*, vol. 16, no. 11, Nov. 2021, Art. no. e0259868, doi: 10.1371/journal.pone.0259868.
10. L. Organick et al., “An empirical comparison of preservation methods for synthetic DNA data storage,” *Small Methods*, vol. 5, no. 5, May 2021, Art. no. 2001094, doi: 10.1002/smt.202001094.
11. R. N. Grass, R. Heckel, M. Puddu, D. Paunescu, and W. J. Stark, “Robust chemical preservation of digital information on DNA in silica with error-correcting codes,” *Angewandte Chemie*, vol. 54, no. 8, pp. 2552–2555, Feb. 2015, doi: 10.1002/anie.201411378.
12. S. Newman et al., “High density DNA data storage library via dehydration with digital microfluidic retrieval,” *Nature Commun.*, vol. 10, no. 1, Apr. 2019, Art. no. 1706, doi: 10.1038/s41467-019-09517-y.
13. S. M. Tabatabaei Yazdi, Y. Yuan, J. Ma, H. Zhao, and O. Milenkovic, “A rewritable, random-access DNA-based storage system,” *Scientific Rep.*, vol. 5, Sep. 2015, Art. no. 14138, doi: 10.1038/srep14138.
14. C. Winston, L. Organick, D. Ward, L. Ceze, K. Strauss, and Y. J. Chen, “Combinatorial PCR method for efficient, selective oligo retrieval from complex oligo pools,” *ACS Synthetic Biol.*, vol. 11, no. 5, pp. 1727–1734, Feb. 2022, doi: 10.1021/acssynbio.1c00482.
15. K. N. Lin, K. Volkel, J. M. Tuck, and A. J. Keung, “Dynamic and scalable DNA-based information storage,” *Nature Commun.*, vol. 11, no. 1, Jun. 2020, Art. no. 2981, doi: 10.1038/s41467-020-16797-2.
16. K. J. Tomek et al., “Driving the scalability of DNA-based information storage systems,” *ACS Synthetic Biol.*, vol.

- 8, no. 6, pp. 1241–1248, May 2019, doi: 10.1021/acssynbio.9b00100.
17. J. L. Banal et al., "Random access DNA memory using Boolean search in an archival file storage system," *Nature Mater.*, vol. 20, no. 9, pp. 1272–1280, Sep. 2021, doi: 10.1038/s41563-021-01021-3.
 18. C. Bee et al., "Molecular-level similarity search brings computing to DNA data storage," *Nature Commun.*, vol. 12, no. 1, Aug. 2021, Art. no. 4764, doi: 10.1038/s41467-021-24991-z.
 19. K. J. Tomek, K. Volkel, E. W. Indermaur, J. M. Tuck, and A. J. Keung, "Promiscuous molecules for smarter file operations in DNA-based data storage," *Nature Commun.*, vol. 12, no. 1, Jun. 2021, Art. no. 3518, doi: 10.1038/s41467-021-23669-w.
 20. C. Fuller et al., "The challenges of sequencing by synthesis," *Nature Biotechnol.*, vol. 27, pp. 1013–1023, Nov. 2009, doi: 10.1038/nbt.1585.
 21. S. Goodwin, J. McPherson, and W. McCombie, "Coming of age: Ten years of next-generation sequencing technologies," *Nature Rev. Genetics*, vol. 17, pp. 333–351, Jun. 2016, doi: 10.1038/nrg.2016.49.
 22. M. MacKenzie and C. Argyropoulos, "An introduction to nanopore sequencing: Past, present, and future considerations," *Micromachines*, vol. 14, no. 2, 2023, Art. no. 459, doi: 10.3390/mi14020459.
 23. D. Branton et al., "The potential and challenges of nanopore sequencing," *Nature Biotechnol.*, vol. 26, pp. 1146–1153, Oct. 2008, doi: 10.1038/nbt.1495.

DAVE LANDSMAN is a distinguished engineer at Western Digital Research, Milpitas, CA 95035 USA. Contact him at dave.landsman@wdc.com.

KARIN STRAUSS is a senior principal research manager at Microsoft Research, Redmond, WA 98052 USA and an affiliate full professor at the University of Washington, Seattle, WA 98195 USA. Contact her at kstrauss@microsoft.com.

IEEE Computer Society Has You Covered!

WORLD-CLASS CONFERENCES — Over 195 globally recognized conferences.

DIGITAL LIBRARY — Over 900k articles covering world-class peer-reviewed content.

CALLS FOR PAPERS — Write and present your ground-breaking accomplishments.

EDUCATION — Strengthen your resume with the IEEE Computer Society Course Catalog.

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

NETWORK — Make connections in local Region, Section, and Chapter activities.



Explore all member benefits
www.computer.org today!



Pervasive Healthcare: Privacy and Security in Data Annotation

Emma L. Tonkin , University of Bristol, BS8 1TH, Bristol, U.K.

Kristina Yordanova , University of Greifswald, 17489, Greifswald, Germany

Novel pervasive healthcare solutions often require extensive labeled data, both to train activity, behavior, and symptom detection systems, and to monitor the reliability of the system. The annotation process can be intrusive for participants and raise significant privacy and security issues. We discuss these challenges and identify mitigations.

The increasing use of pervasive healthcare monitoring and assistance systems has raised awareness both of the potential of these systems and of their potential shortcomings. The recent pandemic, alongside the challenges of supporting an ageing population in many countries, have placed increasing strain on healthcare systems worldwide. Pervasive monitoring offers the potential to inform clinicians' decisions, helping to ensure that patients receive the care they need. To know whether this promise can be realised, it is key to understand and monitor how well these systems work in real-world contexts as part of an ongoing validation and review process, to ensure that they are not used in environments or contexts in which they perform poorly. Validation may also look beyond model performance in order to evaluate the clinical relevance of a system, ensuring that it can provide the required functionality in the target user population to an adequate level to meet the needs of healthcare professionals. To be useful, a system must also be accessible to individuals making use of the data, such as clinicians or caregivers, and therefore it is also advantageous to monitor the system's usability to understand how effectively the information held is transmitted to these stakeholders. Consider, for example, a system designed to monitor the progression of dementia: The system should be able to make clinically relevant judgements about the extent to which everyday behaviors reflect cognitive decline, expressing these in such a way that their clinical relevance is clear.

Pervasive health monitoring systems are typically trained to recognize facets of a person's daily activities, health status, or behavior. To build models capable of detecting these on a wearable, mobile, or pervasive platform, a ground truth containing appropriately labeled data is necessary for training and testing purposes. The potential for bias in sensor development is significant, and awareness of the fragilities of the validation and training processes has increased in recent years. As a simple example, the shortcomings of medical technologies, such as pulse oximeters illustrate the lack of inclusivity in the validation process. These tools rely on near-infrared light transmission to noninvasively monitor arterial hemoglobin oxygen saturation and are sensitive to skin pigmentation. The consequences of such bias can be significant to healthcare. A 2022 study found that in the COVID-19 pandemic, Asian, Black, and Hispanic patients received less supplemental oxygen than White patients, associated with differences in pulse oximeter performance.

Activity and behavior recognition systems are typically much more complex, monitoring complex behaviors in many clinical contexts and potentially relying on a combination (fusion) of several sensors, so are likely to be sensitive to variation in many factors, including demographic, social, cultural, and lifestyle. Consider the deployment of a dementia-progression sensor in domestic homes. It must be robust to differences in individual activities resulting from differences in who we live with, how we approach everyday tasks and how we use our homes: for example, everyday activities such as cooking vary a great deal depending on what we choose to cook. These issues are likely to be further aggravated by temporal drift, as any initial training data provided by the users themselves becomes less relevant over time.³ This does not necessarily reflect intrinsic change, and could result from

causes as simple as the purchase of a new piece of kitchen equipment that changes the way we approach a cooking task. There are mitigating steps that can be taken to minimize the impact of these factors. However, to reduce the likelihood of poor diagnosis and outcomes as a result of tools that operate poorly in a particular context of use, and to reduce the risks of bias and unfairness, broader and more extensive validation and ground truthing is recommended—if deployed in the real world, systems should be robust enough to cope with the broad diversity seen in real-world conditions.

Collection and labeling of data can be invasive and time-consuming, and raise security and privacy risks for participants. Where data collected can be directly annotated, a subset of data collected during the ordinary course of a product's use may be manually annotated—as with voice agents like Alexa—to evaluate and improve the tool's performance, raising the risk of unauthorized disclosure, etc. In many applications that ordinarily make use of “opaque” data, that is, data which are difficult to label directly, supplemental modalities of data are required, such as video or audio, which support offline annotation. The community is therefore faced with the task of responsible data annotation: finding methods to collect the necessary ground-truth data in such a way as to maximize the quality of the result, while minimizing the impact and risk on the participant. This is further complicated by the observation that pervasive healthcare systems are often designed and deployed in contexts where legal frameworks and participant expectation coincide in a strong expectation of privacy, such as in private homes and healthcare settings.

CHALLENGE: ANNOTATING THE EVERYDAY LIVES OF PEOPLE WITH DEMENTIA

In what follows, we will discuss different challenges associated with ensuring privacy and security in data annotation. To illustrate these challenges, we will use the insideDEM study as a running example.⁷ The aim of the insideDEM project was to develop and analyse a dataset in order to understand whether a pervasive system could support clinical decision-making in dementia care. Nursing staff have limited time to observe and work with each patient, so potentially systems that could provide a good understanding of patient symptoms could help decision-making around changes in behaviour, as well as motivating decision-making around patient safety and wellbeing. To this end, we collected and analysed sensory data describing the everyday life of people suffering from dementia. The data were collected at two locations, one senior home with people in the late stages of

dementia and another senior home with people who were in the middle stages of dementia. The goal of the data analysis was to recognize the exhibition of challenging behavior associated with the progression of the disease and the collected data were annotated correspondingly online (i.e., at the time of data collection) using dementia care mapping (DCM), which is an online method where a trained mapper annotates the behavior of up to eight participants in a five minutes interval. DCM is a standard tool for measuring the quality of care of people with dementia and it has been shown to be suitable for assessing behavioral problems in activities of daily living. Later, a more accurate offline annotation was produced by looking at the video logs.

PRIVACY AND SECURITY CONCERNS IN ENHANCED DATA COLLECTION AND LABELLING

Security of healthcare data is a concern to users and hence an obstacle to adoption.⁵ High-profile existing controversies and past data breaches may be a factor, as may the novelty of the platforms used, which may result in variation in sensitivity to privacy and security. Users put high importance on the protection of health data and may have concerns about access by others (e.g., family members), how it is processed, and for what purpose.⁵ The perception that a platform is secure can facilitate adoption.

The collection of additional data for annotation may be viewed as a form of “enhanced surveillance,” data collection beyond that required to achieve a platform's primary aims. Steps that can be taken to support acceptance of such activity include: thorough, informative, and accessible information made available at all stages about annotation activities, including any privacy-enhancing steps and security provisions taken to limit the risks; meaningful engagement with participants and stakeholders to support co-design and to understand participant perspectives; developing effective explanation strategies and supports, rather than relying solely on documentation or terms and conditions; an ongoing consent model, rather than a one-off model; building and sharing an effective model of the risks of particular data types or modalities; employing technical measures (such as privacy-enhancing technologies) and organizational controls to protect participant data against privacy and security risks.

In the insideDEM example, we observed privacy concerns both by the family members of the participants and by the nurses in the senior homes. One family member refused for his relative with dementia to be video recorded, which resulted in a lack of video recordings in the first location, meaning that online annotation became an essential component of the dataset. Nurses also

shared that they felt monitored and controlled by the use of video technology and it was a challenge to convince them to participate in this study. As we note later in this article with regards to the risk of a data breach, the concerns of clinical staff and family members are well reasoned and often grounded in well publicized existing cases. For example, the participants' concerns in insideDEM echo the broader international debate around the use of cameras to monitor treatment of residents in nursing homes.² The potential for these data to be adapted to monitor staff more broadly or to be used in ways which compromise resident privacy is widely discussed.

When exploring the possibility of participation in a study, it is useful to bear in mind the ethical boundaries of persuasion and to strive for *rational* persuasion, in which the potential participant reaches their decision based on reasoned consideration of the reasons advanced by the researchers.⁸ It is important also to contend with the fact that as researchers in healthcare, many nonrational means of persuasion are also present, ranging from peer pressure to appeal to authority, which—especially given the hierarchies of power involved in a work context—may exceed what is appropriate. Coercion (force, leading to deprivation of choice) and manipulation (imposition of covert influence upon decision-making) must be avoided, and even persuasive methods are problematic in contexts in which it is not clear that the intended outcome is in the best interests of those being influenced.

In this note, we identify several steps that can be taken when designing an annotation task to understand and document the risks to those involved in collecting a dataset; to reduce the risks and reassure participants; and to communicate effectively about technologies, risks, benefits, and outcomes.

CENTERING PARTICIPANT AND STAKEHOLDER PERSPECTIVES

Participant and stakeholder engagement provides insight into the environment and its actors, and begins by discovering who is involved, who is affected by data collection, and their roles. Participatory design approaches, increasingly used in healthcare, involve participants directly in the design process in ways that empower them to influence the outcome. Deployment processes are often viewed as negotiations.⁴ This was also true in insideDEM, during which researchers liaised with nursing home staff, residents, and their family members to understand their views on the study and the technology used, with the ultimate aim of negotiating informed consent. We made use of their feedback to modify our approach in line with the issues and preferences raised. In particular, as mentioned

previously, we did not use video in one of the insideDEM locations.

While such compromises are vital, it is important to recognize the risks. In the case of insideDEM, following analysis, the online annotation used was shown to be very inaccurate. In the second location, in which video recordings were taken, we were able to reannotate the dataset based on the video logs. For the first location, as video was unavailable, we were unfortunately unable to reconstruct an accurate ground truth. This limitation in the ability to repair a damaged dataset is a risk that emerges from light-touch annotation approaches. A strategy that may help to mitigate these risks is to make use of additional pilot studies in minimally invasive environments, helping to inform researchers' knowledge of the options, identify effective mitigations, and provide accessible examples of what is planned. Pilot studies can also be used to support good, clear explanations of the planned research. Furthermore, participatory design methods such as co-design help to build in participant perspectives "from the ground up."

It is useful to recall that participant acceptance does not imply legality, just as what is legal is not necessarily acceptable to participants. Perceived data sensitivity from technology users does not necessarily align with legal and regulatory definitions and guidance.¹ Nonetheless, participant engagement is a key aspect of "privacy by design and by default," the broader aim of which is to identify ways of achieving our goals, while minimizing risks, foregrounding consent, and encouraging participants to remain in control of decisions taken around their data.

UNDERSTANDING THE RISKS OF "RICH" DATA

Data are commonly understood by reference to the ways those who collected the data thought it would be used. That is, the intent behind collecting a particular data type is often conflated with the capabilities of the data. However, for privacy concerns to be taken into account, it is important to realize that data may contain a great deal more information than the information we are interested in. Custodians of data do not always know what potential a dataset may have, or how it may be used by others.

Data suitable for annotation are by definition data that contain adequate information to support the labeling process. There is a high likelihood that such data contains more information than the labeling process requires. For example, to validate a speech-to-text system, it is important to know that the system has transcribed the text in a way that agrees with a human annotator. However, the content of the text may contain

a great deal of information about the data subject, their activities, and their acquaintances. So the ways in which data could be processed are a better signifier of their sensitivity than our intent in collecting them.

In the insideDEM project, the video data collected for annotation are an example of data that contain much more information than the features we are interested in. We were interested in the annotation and identification of challenging behaviors, such as pacing, apathy, and agitation, but, of course, the video log provides us also with a lot of information about other activities, interactions in the senior home, and persons that are not part of the study (guests or nurses), and if the data are not handled properly, it could provide a lot of private or sensitive information. To handle this problem, technical and organizational measures were employed. Only a few people who were cleared through the ethics proposal were allowed to see and annotate the data. The data were on a secure server, to which only certain people had access, and encrypted at rest. The ML researchers never saw or used the video logs.

PRIVACY-ENHANCING TOOLS (PETS) IN ANNOTATION

In the insideDEM project, online annotators were trained to annotate a certain set of challenging behaviors and ignore any other features identified. The drawback of the online approach is that contextual information is lost; the annotation is not very accurate, as compared to offline annotation based on video log. However, the use of selective attention has an analogue in offline annotation, specifically in the use of technologies designed to reduce the amount of unnecessary data collected, and to maximize the relevance of the representation shown to annotators.

A helpful guideline in personal data management is the principle of data minimization—collect only the data that are relevant and necessary for the purpose. Appropriate technologies can make this easier. For example, video streams can be preprocessed using PETs. These might for example employ simplified representations of the individual using an abstraction, such as a bounding box, silhouette, skeleton, or “cartoon depiction.” In this manner, identifiable features are substituted with alternative representations that are less identifiable but still carry relevant information.

The selection of appropriate preprocessing methods is a research task in itself. To be used effectively, it is important to have an understanding of what information is retained in the processed representation. This informs the use of these data for annotation, and is also helpful in mapping residual risk.

Whilst protective to participant privacy, such preprocessing has drawbacks. For example, the quality of annotations may suffer as a result of the reduced expressivity of the data. If it becomes desirable to reanalyze data with different labels or at a more detailed granularity, a representation deemed adequate for the initial study may be inadequate for the new approach. However, such technologies help to minimize the risk of data breaches. If unauthorized access is gained to data of this kind, a participant is less likely to be reidentified or harmed. It is also worth considering that appropriate preprocessing has the potential to facilitate the annotation task, for example, by highlighting relevant features or events.

MEANINGFUL CONSENT AND EXPLAINABILITY

Consent is a vital principle in data collection. Communicating clearly about data and how it is used and processed is an important part of ensuring that people understand what data are collected, what the risks and benefits are, and how these are managed, how the data will be used, and what precautions are taken to safeguard the data. Effective explanation in the context of pervasive healthcare is an active research area⁴ which, while it links to explainable AI, includes a broad range of concerns ranging from the establishment of common ground and shared mental models to negotiation of vocabularies, development of effective strategies for demonstrating and discussing technical functionality, user experience, and so on.

In our case study, there is a very clear tension between the potential of systems derived from our data to improve the everyday lives of people with dementia and of healthcare professionals and family members, and the potential for the intrusive process of data collection to cause harm. A shared understanding of the problems, the risks to participants, the mitigating factors that limit these risks, and the potential benefits of the research, are essential elements in supporting rational decision-making—not only in terms of the decision to participate, but in negotiating which elements of the research a potential participant wishes to take part in, and which they do not (e.g., *granular consent*).

While the process of negotiating consent is well understood, the long-term nature, unobtrusive profile, and “tool/agent” purpose of many pervasive systems complicates consent in several ways. First, consent once earned is not automatically valid forever. Rather, it may require regular review, taking into account several circumstances: for example, a person may: simply change their mind; no longer find participation useful or relevant; be unable to consent due to cognitive decline or a mental health condition. Consent decisions may also require review due to changes in the way data are used, which may be driven by

external forces or result from appropriation, the adaptation of existing technologies for new purposes.⁶

It is possible to justify some forms of data collection on a lawful basis other than consent, for example, EU law offers “Legitimate interest,” which applies when personal data are processed in a way that the data subject would expect and which offers a clear benefit (e.g., improvement in the performance of a system, improving outcomes for its users). Some personal data, however, are sensitive—relating to health, biometric information, or some other private information, and in this case, it is vital that those whose data are collected understand and agree with what is happening, or when they are unable to give consent, that an appropriate process is followed.

Such assurances were sought by several stakeholders during the insideDEM project; for example, one family member was very concerned that the data would be collected in such a way that data breaches could easily occur. This is by no means an unreasonable concern, as individuals who follow technology news will be aware of several high-profile cases in which consumer surveillance equipment resulted in severe data breaches. A keystone of meaningful consent is good, effective, and accurate explanation.

CONCLUSION

Pervasive healthcare applications are likely to become increasingly significant parts of healthcare monitoring. High-quality labeled data are key to building tools that work well in the many real-world contexts where they would be most useful. Yet data labeling receives little coverage in papers and articles, and the lessons learned by researchers are not always shared. This makes the task of ensuring that annotation is treated securely and in a privacy ensuring manner even more challenging than it might otherwise be. In this note, we described some of the associated challenges and attempted to identify potential solutions. 🤖

ACKNOWLEDGMENTS

The authors would like to thank the researchers from the Mobile Multimedia Information Systems Group at the University of Rostock who were involved in the insideDEM project for their input on the project’s development and challenges. This work was supported under the SPHERE Next Steps Project funded in part by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/R005273/1. It was also supported in part by the Mare Balticum Fellowship Program of the University of Rostock and in part by the Federal Ministry of Education and Research of Germany under Grant FKZ 16SV7349.

REFERENCES

1. R. Belen-Saglam, J. R. C. Nurse, and D. Hodges, “An investigation into the sensitivity of personal information and implications for disclosure: A UK perspective,” *Front. Comput. Sci.*, vol. 4, 2022, Art. no. 908245, doi: 10.3389/fcomp.2022.908245.
2. C. Berridge, J. Halpern, and K. Levy, “Cameras on beds: The ethics of surveillance in nursing home rooms,” *AJOB Empirical Bioeth.*, vol. 10, no. 1, pp. 55–62, 2019, doi: 10.1080/23294515.2019.1568320.
3. D. Bouchabou, S. M. Nguyen, C. Lohr, B. LeDuc, and I. Kanellos, “A survey of human activity recognition in smart homes based on IoT sensors algorithms: Taxonomies, challenges, and opportunities with deep learning,” *Sensors*, vol. 21, 2021, Art. no. 18, doi: 10.3390/s21186037.
4. R. Eardley et al., “A case study investigating a usercentred and expert informed ‘companion guide’ for a complex sensor-based platform,” *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 6, no. 2, Jul. 2022, Art. no. 93, doi: 10.1145/3534625.
5. C. Jacob, E. Sezgin, A. Sanchez-Vazquez, and C. Ivory, “Sociotechnical factors affecting patients’ adoption of mobile health tools: Systematic literature review and narrative synthesis,” *JMIR Mhealth Uhealth*, vol. 10, no. 5, May 2022, Art. no. e36284, doi: 10.2196/36284.
6. T. Jakobi, C. Ogonowski, N. Castelli, G. Stevens, and V. Wulf, “The catch(es) with smart home: Experiences of a living lab field study,” in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2017, pp. 1620–1633, doi: 10.1145/3025453.3025799.
7. F. Kruger, C. Heine, S. Bader, A. Hein, S. Teipel, and T. Kirste, “On the applicability of clinical observation tools for human activity annotation,” in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, 2017, pp. 129–134, doi: 10.1109/PERCOMW.2017.7917545.
8. N. Jacobs, “Two ethical concerns about the use of persuasive technology for vulnerable people,” *Bioethics*, vol. 34, pp. 519–526, 2020, doi: 10.1111/bioe.12683.

EMMA L. TONKIN is a research fellow with the Department of Electrical and Electronic Engineering, University of Bristol, Bristol, BS8 1TH, U.K., and an organizer of the Annotation of User Data for Ubiquitous Systems Workshop Series. She is the corresponding author of this article. Contact her at e.l.tonkin@bristol.ac.uk.

KRISTINA YORDANOVA is a full professor of data science with the University of Greifswald, 17489, Greifswald, Germany, and an organizer of the Annotation of User Data for Ubiquitous Systems Workshop Series. At the time at which this work was carried out, Kristina led the Cognitive Methods for Situation-Aware Assistive Systems (CoMSA²t) junior research group at the University of Rostock, DE. Contact her at Kristina.yordanova@uni-greifswald.de.

Privacy in the Era of 5G, IoT, Big Data, and Machine Learning

Elisa Bertino , *Purdue University*

We have today a number of technologies that, when combined, can support unprecedented applications and significantly enhance existing applications. These technologies include 5G cellular networks, big data, machine learning, and the Internet of Things (IoT). The combination of those technologies allows us to: a) increase our capacity for pervasive, fine-grained, and continuous data gathering and for the effective and efficient processing of these data (even with real-time guarantees); b) generate knowledge from data and

*THE INCREASED ADOPTION
OF WEARABLE DEVICES AND
CONTINUOUS DATA STREAMING
FROM THESE DEVICES ALLOWS A
PARTY TO COLLECT FINE-GRAINED
GEO-TEMPORAL DATA ABOUT
INDIVIDUALS.*

continuously evolve this knowledge, thus supporting recommendation systems and decision processes—also accompanied by suitable explanations; and c) make devices, control systems, and cyberphysical systems intelligent and autonomous.

However, many such technologies collect and/or use data, which often contain privacy-sensitive data. Collected data, even if anonymized by removing identifiers such as names or Social Security numbers, when linked with other data may lead to reidentifying

the individuals to which specific data items are related. Also, as organizations, such as governmental agencies, often need to collaborate, they exchange datasets, resulting in these datasets being available to many different parties. Privacy breaches also occur at many different layers (for example, networks, hosts, and applications) and components in our interconnected systems. An example of a privacy attack in the context of a cellular network is the ToRPEDO side-channel attack,¹ which exploits the paging protocol to track users.

On the other hand, security techniques implemented by applications, especially the mobile ones, often have vulnerabilities, which undermine privacy. Notable examples are vulnerabilities in authentication protocols, such as in conventional login-password-based authentication and in SMS-based one-time passwords, or the use of covert channels and side channels to bypass the permission systems of the underlying operating systems.² It is important to emphasize that security and privacy are two different requirements. However, security is a prerequisite for privacy. The use of machine learning techniques further threatens privacy because of attacks such as the inversion ones by which a party can infer the sensitive contents of the data samples used for training. Finally, the increased adoption of wearable devices and continuous data streaming from these devices allows a party to collect fine-grained geo-temporal data about individuals.

Given the many ways by which data privacy can be breached, one may wonder whether the battle for privacy is lost or whether something can be done. In this respect, it is important to notice that research and industry have proposed many privacy-preserving techniques over the last 20 years, ranging from cryptographic techniques, such as oblivious data

Digital Object Identifier 10.1109/MSEC.2022.3221171

Date of current version: 19 January 2023

structures, which hide data access patterns, and homomorphic encryption to data anonymization techniques, which transform data to make it more difficult to link specific data records to specific individuals or perturb the data. The problem of location privacy has also been the focus of extensive research both in the past and recently. Research efforts have also been devoted to investigating privacy-preserving techniques for data in the cloud, on smartphones, and in social networks. Finally, trusted environments have been developed that represent an important building block for privacy-preserving techniques.

However, despite the availability of many privacy-preserving techniques, we are still far from satisfactory solutions to privacy. The first reason is that privacy depends very much on user personal preferences, contexts, and culture. Therefore, we need privacy-preserving techniques that can be personalized. The other reason is that several privacy regulations have been defined over the years, such as the Health Insurance Portability and Accountability Act (HIPAA) and General Data Protection Regulation (GDPR). Technical solutions need to comply with these regulations—and in some cases, coming up with approaches is challenging, and also, these approaches need to be complemented by proper organizational processes.

Finally, it is important to notice that different privacy techniques must be used depending on the tasks to be executed on the data, such as record linkage, data analytic, and operational tasks, and on the specific transactions a user is executing, such as browsing the web, getting recommendations on movies, and buying products online. In the latter context, techniques proposed for privacy-preserving digital and for privacy-preserving e-commerce transactions are critical. Those techniques combined with network anonymizers and application-level privacy techniques are critical building blocks for holistic online privacy.

Given that we have all those privacy-preserving technologies, what more do we need for privacy? What we need are holistic privacy-preserving environments able to combine privacy-preserving approaches with adaptation to different user contexts and tasks to achieve “privacy protection in depth.” Users consider privacy important, but they often feel that privacy is complex to manage.

However, there is also the key question of “personal privacy versus collective safety,” as ultimately, the choice of making available (some of) our personal data, and thus renouncing some privacy, to benefit society is a personal choice. Users must be able to make informed decisions. Therefore, two challenging questions need to be addressed: 1) How can we make it possible for people to make decisions about “personal privacy versus collective safety”? 2) How can we make it possible to reconcile those two seemingly

*HOW CAN WE MAKE IT POSSIBLE
FOR PEOPLE TO MAKE DECISIONS
ABOUT “PERSONAL PRIVACY VERSUS
COLLECTIVE SAFETY”?*

opposing goals? We believe that data transparency and policy-based use of data are two key elements relevant to answering these questions. 🤖

REFERENCES

1. S. R. Hussain, “Privacy attacks to the 4G and 5G cellular paging protocols using side channel information,” in *Proc. Netw. Distrib. Syst. Secur. Symp.*, 2019, pp. 1–15, doi: 10.14722/ndss.2019.23442.
2. J. Reardon et al., “50 ways to leak your data: An exploration of apps’ circumvention of the android permissions system,” in *Proc. 28th USENIX Secur. Symp.*, 2019, pp. 1–10.

ELISA BERTINO is a professor with Purdue University, West Lafayette, IN 47907 USA. Contact her at: bertino@purdue.edu.



WWW.COMPUTER.ORG/COMPUTINGEDGE

DEPARTMENT: COMPUTING'S ECONOMICS

Scams, Frauds, and Crimes in the Nonfungible Token Market

Nir Kshetri, University of North Carolina at Greensboro

This article delves into scams, frauds, and deceptions in the nonfungible token (NFT) market. It also proposes a typology of cyberattacks and other malicious behaviors in the NFT space.

The nonfungible token (NFT) market is growing rapidly. According to Lithuania-based data acquisition and analysis company DappRadar, which tracks decentralized applications across multiple blockchains, the NFT market exceeded US\$23 billion in 2021 compared to less than US\$100 million in 2020.¹ The U.S. multinational investment bank and financial services company Morgan Stanley estimates that the NFT market could reach US\$240 billion in 2030.² This rapid growth in the NFT market has offered a wide variety of opportunities for scammers, fraudsters, and cybercriminals. The targets of such acts are creators and owners as well as consumers and buyers of NFTs. Investors in NFT projects have also been defrauded. Some perpetrators use techniques such as hacking and malware that are intended to gain unlawful access to victims' digital wallets that store NFTs and other cryptoassets. Others rely on novel but simple social engineering scams to convince victims to invest in fake schemes involving NFTs and to divulge sensitive information that can be used to breach cryptoaccounts.

Instances of other abusive practices such as insider trading have also been reported. In September 2021, the world's largest NFT marketplace, OpenSea, admitted that its product head was engaged in an insider trading scam. The scheme involved buying an NFT before it was advertised. When buyers' interest in the NFT increased, the asset would be sold at a higher price.³ In one trade, a digital artwork was bought for

US\$822 and sold for US\$4,000.⁴ In this article, I discuss how NFTs are vulnerable to breaches, bugs, and attacks as well as other types of scams, frauds, and deceptions. The article delves into cyberattacks and other malicious behaviors in the NFT space.

CYBERATTACKS AND OTHER MALICIOUS BEHAVIORS IN THE NFT SPACE

Cryptoassets, such as cryptocurrencies and NFTs, are vulnerable to cyberattacks at various levels. First, NFTs face risks related to the platform on which smart contracts run.⁵ That is, the blockchain behind the NFTs themselves could be vulnerable to hacking. For instance, Ethereum was hacked in 2016 by exploiting vulnerabilities in the code of the decentralized autonomous organization (DAO). Note that the DAO was launched by a group of Ethereum developers, who are run through smart contracts and do not need centralized management and the direct control of self-interested institutions. At the next level, exchanges that facilitate the trading of NFTs (for example, OpenSea) have their own vulnerabilities. Finally, cybercriminals can hack wallets that are used to store NFTs.

Cyberattacks targeting NFTs mainly include actions against NFT exchanges and wallets. While NFTs are based on blockchain, exchanges and marketplaces such as OpenSea and Rarible, function in a centralized manner.⁶ Thus, they cannot seize the benefits of decentralized technologies, such as peer review systems to identify and fix bugs. Consequently, they are vulnerable to breaches, bugs, and attacks. In September 2021, a bug in the OpenSea token market



led to the disappearance of 42 NFTs that were valued at more than US\$100,000.⁷

There are two types of wallets: hot ones (for example, accounts in an exchange/website-based wallets) and cold ones (for instance, those based on hardware or paper). NFTs that are stored in hot wallets are under the control of the wallet provider. For instance, custom protocols are used for accounts in cryptoexchanges, which are often based on a nonblockchain system.⁸ The majority of attacks involving NFTs have been carried out against hot wallets.

Social engineering, which involves emotional appeals, such as fear, pity, and excitement, to victimize targets, has been a major modus operandi of most NFT fraudsters. Those parties establish interpersonal relationships and create a feeling of trust and commitment to achieve their goals. Social engineering tricks are used to gain access to victims' private keys to accounts associated with NFTs. In other cases, victims may be lured to click malicious links and download files containing malware.

In addition to cybercrimes, many other unlawful and malicious behaviors occur in the NFT space. Perpetrators are taking advantage of the relative newness of the NFT market and potential victims' lack of understanding of such assets. Other key challenges include underdeveloped regulations around cryptoasset intellectual property rights, copyright theft, unauthorized replication of NFT artwork, and the creation of phony NFT artwork.⁵ For instance, scammers are creating and selling NFTs without the knowledge and consent of the owners of the assets that the NFTs represent.

Some NFT platforms have facilitated fraudulent practices by allowing transactions without proper due process and verification. Twinci, which describes itself as the first NFT social marketplace, permits anyone to open an account and start creating and collecting NFTs. A user can connect cryptocurrency wallets

such as Metamask and imToken, and automatically set up a profile for them. Note, too, that wallets such as imToken do not require email addresses or any other personal information to set up an account. Once a user connects on Twinci, he or she can upload an image of an artwork. Twinci mints a token of the image, and the NFT is ready to go to the marketplace. Twinci accountholders can name their price in a chosen cryptocurrency.

SCAMS, FRAUDS, AND CRIMES INVOLVING NFTS: A TYPOLOGY

Table 1 presents a typology of cyberattacks and other malicious behaviors in the NFT space. The vertical axis represents fraudsters' *modi operandi*. The horizontal axis shows the targets of the schemes. In this section, we discuss the nature of the crimes in each cell.

Cell 1

As mentioned, cybercriminals increasingly target digital wallets of NFT owners. In June 2021, an NFT artist, Fvckrender, reported that he was tricked into opening a file containing a virus delivered to his social media account,⁹ which enabled a criminal to access his digital wallets. He reported that the hacker stole 40,000 Axie Infinity tokens valued at US\$4 million within minutes.¹⁰ In a similar incident, in December 2021, an art curator and NFT collector reported the theft of 16 NFT tokens in a phishing attack. NFTs worth about US\$2.2 million were stolen from the collector's hot wallet.¹¹

Cell 2

As noted, NFT platforms face protocol risks such as hacking. Israeli cybersecurity company Check Point reported that it found vulnerabilities in OpenSea that could have enabled cybercriminals to sell malicious NFTs or trojanized digital art. Check Point researchers said that a security flaw in OpenSea made it possible for hackers to offer a malware-infected image file

TABLE 1. A typology of NFT scams.

Victim/target → Main element of the victimization strategy	Creators/owners of NFTs or the actual assets that NFTs represent	Consumers/buyers of NFTs or investors in NFT projects
Technology attacks, such as malware and hacking (mostly in combination with social engineering)	Cell 1 <ul style="list-style-type: none"> Attacks targeting digital wallets of NFT creators/owners 	Cell 2 <ul style="list-style-type: none"> Exploiting security flaws in NFT platforms Giveaway scams
Purely social engineering and other nontechnological attacks	Cell 3 <ul style="list-style-type: none"> Creating fake NFT customer service pages to lure NFT creators/owners Creating and selling NFTs without the knowledge and consent of the owner of the actual assets that NFTs represent Tricking artists into paying to mint NFTs of their assets 	Cell 4 <ul style="list-style-type: none"> Investment scams Tricking consumers into buying fake NFTs

The U.S. multinational investment bank and financial services company Morgan Stanley estimates that the NFT market could reach US\$240 billion in 2030.

as an NFT. For instance, a user could be lured with a free NFT. When he or she opened the NFT file, a series of malicious pop-ups pretending to be from OpenSea would deploy. One of them would request the user to connect his or her digital wallet. When the user did so, the hackers would steal funds. OpenSea patched the flaw when it was brought to the company's attention.¹²

Another category of scams involves giveaways and airdrops, in which fraudsters lure victims by offering free NFTs. In such a scheme, a fake account sends a message to users on social media, such as Twitter, telling them that they have won an NFT. Users are given a link to a fake website, which asks them to connect their digital wallet and enter their seed phrase.¹³ The criminals then steal NFTs and digital currencies and tokens in the wallet.¹⁴

Cell 3

Some scammers use fake customer service pages to trick NFT owners into divulging sensitive information. When creative producer and director Jeff Nicholas was trying to get help for a royalty issue from OpenSea in August 2021, a group of criminals masquerading as company employees scammed him. They invited Nicholas into a channel of the voice over Internet Protocol instant messaging and digital distribution platform Discord, called *OpenSea Support Server*. After hours of interaction, they convinced him to share his screen. When he did, they took a picture of the QR code synced to his private key, or seed

phrase, which enabled them to gain full access to his cryptoassets. They stole 150 ether (ETH) valued at about US\$480,000.¹⁵

Fraudsters also take advantage of the lack of clear regulations regarding the ownership of an NFT versus the ownership of the physical or digital object represented by the NFT.⁵ A distinct category of NFT scam involves creating and selling NFTs of works by high-profile artists without their knowledge and permission. Serbian artist Milos Rajkovic, who created video loops in which human faces and landscapes transform in strange ways (<http://sholim.com/biography.html>), was not involved with NFTs. In July 2021, he found that 122 of his works were for sale on OpenSea. While the first fakes were removed, another account posted the same works. Fraudsters exploit NFTs because many artists and collectors do not know about crypto. This makes the market an attractive target.¹⁶ To cite another example, a scammer listed the Chinese artist Qing Han's (known as Qinni) popular artwork *Bird Cage* on Twinci. The platform deleted the NFT and banned the account when the fraud was reported. However, other Twinci accounts had five listings connected to NFTs of Qing's work. Some were listed for as much 500 TWIN (Twinci's cryptocurrency) (1 TWIN = US\$0.54 on 25 November 2021).¹⁷

Scammers are also reported to be creating and selling NFTs in the metaverse that falsely appear to be created by luxury brands. This has raised questions around ownership and legality. For instance, there is no

clear answer to whether sales of branded digital items are legal if the brand did not participate in creating the products. In the metaverse and gaming platform Roblox, brands such as Gucci, Stella McCartney, and Nike have sold digital items. Users can also buy items that appear to be related to Burberry, Chanel, Prada, Dior, and Louis Vuitton despite the fact that these brands may not have been involved.¹⁸ Finally, scammers are said to approach artists to deceive them into paying money and cryptocurrencies, such as ETH, to have NFTs made from their work. The fraudsters then run off with the money.¹⁹

Cell 4

NFT investment scams have also proliferated. One example is the popular NFT project Evolved Apes, which is described on OpenSea as “a collection of 10,000 unique NFTs trapped inside a lawless land.”²⁷ Scammers took 798 ETH from the project’s funds in multiple transfers. The funds were derived from the initial public sale of NFTs and commissions on the secondary market and meant for project-related expenses.²⁰ More than 4,000 NFTs in the Evolved Apes offering were sold in a week.²¹ The artist who created the images was not paid. A social media competition was launched to create buzz. The winners did not receive the promised NFT prizes.²⁰ Cash giveaways were not delivered, and expenses for activities such as marketing and developing game and rarity tools, which are used by brands and creators to list NFT projects for a fee,²² were not paid. Scammers also trick consumers into buying fake NFTs. They copy social media accounts of reputable companies and create fake pages that closely resemble the originals. Using the accounts, they sell bogus NFTs.¹³

PROTECTING AGAINST NFT SCAMS

Creators and owners of NFTs, owners of assets that are potentially attractive for creating high-value NFTs, and consumers, buyers, and investors need to be aware of a wide variety of crimes and scams taking place in NFT marketplaces and exercise security precautions. Owners of NFTs and assets that can be minted into NFTs must be vigilant and take measures to ensure that their assets are not misused. Consumers should understand that buying an NFT is different from buying things on e-commerce websites. There

is little recourse for victims of NFT scams. There are often no refunds and few protections.

In addition, for Ethereum-based NFTs, volatility and gas fees could increase the costs to execute smart contracts.²³ There are also gas fees to transfer NFTs from marketplaces into personal cryptocurrency wallets. For example, in September 2021, *Time* magazine announced the sale of NFTs that consisted of 4,676 tokens tied to digital artwork. Each token was priced at 0.1 ETH (around US\$310 based on the price then of ETH). All tokens were immediately sold, which clogged the Ethereum blockchain network. The fees also increased drastically. Buyers spent about four times as much on transaction fees as they did on the NFTs.²⁴

It is also critical to understand NFT functions such as the storage of content and metadata. In most cases, an NFT is only a smart contract. Content and metadata are stored separately mainly because their files could be too large to hold on the Ethereum blockchain. Thus, while a contract may exist, the data can disappear. NFT markets such as OpenSea, Rarible, Foundation, and Nifty Gateway do not store images. They display only a media file linked with a code on the blockchain. If the media file is deleted from the actual source or the uniform resource locator to that source gets changed or breaks, a buyer may not be able to access an NFT from his or her digital wallet. For instance, online digital art NFT auction platform Nifty Gateway stores data with Cloudinary, a software-as-a-service company providing cloud-based image and video management. If Cloudinary shuts down, NFTs sold by Nifty Gateway may disappear.

Some argue that storing an asset as an interplanetary file system hash is better. The hash acts as an immutable fingerprint. Even in this case, a file can become unavailable if the only node storing it is disconnected from the network.²⁵ An Australian artist and programmer found that most of the images associated with NFTs were hosted in web 2.0 storage, which may lead to the “404: File not found” which means that a page does not exist.²⁶ An NFT can also be removed at the source if a platform’s terms of service, such as those related to copyrights, are violated.²⁸

In light of the proliferation of NFT-related investment scams, it is important to undertake due diligence of investment schemes. For instance, investors can use the Discord platform to understand the community

behind an NFT and get a feel for the project. They should interact with other members and follow topics of conversation. It is important to ask the creators questions about the project's technical aspects. A lack of substance in the discussion can raise a red flag. If the creators have a presence on Discord and respond with details, the project is more likely to be genuine. People associated with fake projects may try to create distractions. It is also important to check if a project creator has an inflated social media following with a high number of fake Twitter followers. For instance, Followeraudit.com (<https://www.followeraudit.com/?ref=alternativeassets.club>) can be used to track the number of active, inactive, and fake followers of a project.

NFTs have provided a number of avenues for criminals, and thus there is a wide range of fraudulent acts in the NFT market. While some scams require technical skills, such as malware and hacking, only social engineering is sufficient to victimize targets in other schemes. The potential problem of storage failure is also an important issue that needs NFT buyers' attention. Likewise, transaction fees may increase the amount buyers need to pay to get their NFTs. Because of the lack of a clear regulatory framework around the ownership of an NFT versus the physical or digital object being represented, some scammers are also creating and selling NFTs without the knowledge or permission of the owners. 🤖

REFERENCE

1. P. Herrera, "Dapp industry report," DappRadar, Dec. 17, 2021. <https://dappradar.com/blog/2021-dapp-industry-report>
2. I. Lee, "Budweiser is getting in on the NFT craze with its 'Key to the Budverse' line of ethereum-based collectibles," Markets Insider, Nov. 29, 2021. <https://markets.businessinsider.com/news/currencies/budweiser-budverse-nft-1936-gold-rare-core-token-collection-beer-2021-11>
3. A. Gupta, "OpenSea bans insider trading after employee defrauds buyers," Jumpstart, Sep. 20, 2021. <https://www.jumpstartmag.com/opensea-bans-insider-trading-after-employee-defrauds-buyers/>
4. A. Herena, "NFT trader OpenSea bans insider trading after employee rakes in profit," *The Guardian*, Sep. 16, 2021. [Online]. Available: <https://www.theguardian.com/technology/2021/sep/16/nft-trader-opensea-bans-insider-trading-after-employee-rakes-in-profit>
5. M. Fox, "The NFT market is now worth more than \$7 billion, but legal issues facing the nascent sector could hinder its growth, JPMorgan says," Markets Insider, Nov. 19, 2021. <https://markets.businessinsider.com/news/currencies/nft-market-worth-7-billion-legal-issues-could-hinder-growth-2021-11>
6. L. Keller, "Does content moderation on platforms like OpenSea amount to censorship?" Forkast, Dec. 17, 2021. <https://forkast.news/does-opensea-censor-nft-content/>
7. "100,000 worth of NFTs disappear forever, thanks to OpenSea bug," Investing, Sep. 09, 2021. <https://www.investing.com/news/cryptocurrency-news/100000-worth-of-nfts-disappear-forever-thanks-to-opensea-bug-2611477>
8. I. Novikov, "The three layers of cryptocurrency security," *Forbes*, May 3, 2018. [Online]. Available: <https://www.forbes.com/sites/forbestechcouncil/2018/05/03/the-three-layers-of-cryptocurrency-security/?sh=12e0ec3e29aa>
9. K. Crow, "NFT art the latest target for online fraudsters," *Financial News*, Aug. 26, 2021. [Online]. Available: <https://www.fnlondon.com/articles/nft-art-the-latest-target-for-fraudsters-20210826>
10. S. Millare, "Four tips for NFT artists to protect themselves from hacking and online theft," BitPinas, Jul. 2, 2021. <https://bitpinas.com/feature/four-tips-for-nft-artists-to-protect-themselves-from-hacking-and-online-theft/>
11. V. Chawla, "Bored Ape NFT collector loses \$2.2M in phishing scam," Crypto Briefing, Dec. 31, 2021. <https://cryptobriefing.com/bored-ape-nft-collector-loses-2-2m-in-phishing-scam/>
12. L. Ropek, "Gullible OpenSea users were vulnerable to 'malicious NFT' attacks, researchers say," Gizmodo, Mar. 28, 2021. <https://gizmodo.com/gullible-opensea-users-were-vulnerable-to-malicious-nft-1847850437>
13. K. Rees, "The 5 biggest NFT scams and how to avoid them," MakeUseOf, Oct. 21, 2021. <https://www.makeuseof.com/biggest-nft-scams-how-to-avoid/>
14. L. Alex, "Evaluating NFTs: How to know whether an NFT project is legit," Cryptonews, Oct. 9, 2021. <https://cryptonews.com/exclusives/evaluating-nfts-how-to-know-whether-an-nft-project-is-legit.htm>

15. A. Wang, "The NFT scammers are here," *The Verge*, Sep. 21, 2021. <https://www.theverge.com/22683766/nft-scams-theft-social-engineering-opensea-community-recovery>
16. "Scammers turn their attention to NFTs as the crypto subsector sees multimillion dollar mania," *Coin News*, Aug. 27, 2021. <https://thecoin.news/post/35827>
17. J. Kwan, "An artist died. Then thieves made NFTs of her work," *Wired U.K.*, Jul. 28, 2021. [Online]. Available: <https://www.wired.co.uk/article/nft-fraud-qinni-art>
18. M. McDowell, "The 'Baby Birkin' NFT and the legal scrutiny on digital fashion," *Vogue Business*, Jun. 15, 2021. [Online]. Available: <https://www.voguebusiness.com/technology/the-baby-birkin-nft-and-the-legal-scrutiny-on-digital-fashion>
19. "Scammers target Sacramento artists through crypto currency: A first-hand account of going down the rabbit hole," *Sacramento News & Review*, Nov. 10, 2021. [Online]. Available: <https://sacramento.newsreview.com/2021/08/20/scammers-target-sacramento-artists-through-crypto-currency-a-first-hand-account-of-going-down-the-rabbit-hole/>
20. E. Gen, "Investors spent millions on 'evolved apes' NFTs. Then they got scammed," *Vice*, Oct. 5, 2021. [Online]. Available: <https://www.vice.com/en/article/y3dyem/investors-spent-millions-on-evolved-apes-nfts-then-they-got-scammed>
21. "NFT buyers scammed as 'creator' bails, who could possibly have seen this coming?" *Kotaku*, Oct. 5, 2021. <https://kotaku.com/nft-buyers-scammed-as-creator-bails-who-could-possibly-1847806528>
22. "Top 7 NFT tools to find the best NFTs," *BeInCrypto*, Nov. 1, 2021. <https://beincrypto.com/learn/nft-tools/>
23. L. Daryanani, "Everything you need to know about the 5 categories of risk associated with DeFi," *AMBCrypto*, May 13, 2021. <https://ambcrypto.com/everything-you-need-to-know-about-the-5-categories-of-risk-associated-with-defi/>
24. W. Gottsegen, "Time's NFT launch sends gas fees spiraling: Keith Grossman, the magazine's president, admits the rollout was 'not ideal,'" *CoinDesk*, Sep. 24, 2021. <https://www.coindesk.com/business/2021/09/23/chaotic-time-magazine-nft-launch-sends-gas-fees-spiraling/>
25. J. Benson, "Yes, Your NFTs Can Go Missing—Here's What You Can Do About It. Most NFTs don't really permanently live on a blockchain. That's potentially a huge problem when it comes to storing them," *Decrypt*, Mar. 19, 2021. <https://decrypt.co/62037/missing-or-stolen-nfts-how-to-protect>
26. I. Walker, "Someone right-clicked every NFT in the heist of the century," *Kotaku*, Nov. 18, 2021. <https://kotaku.com/someone-right-clicked-every-nft-in-the-heist-of-the-cen-1848084379>
27. Unnamed Creator, "Evolved Apes Inc." *OpenSea*. <https://opensea.io/collection/evolved-apes-inc> (Accessed: Jan. 25, 2022).
28. R. Brahmbhatt, "NFTs are mysterious disappearing, here's how." *Interesting Engineering*. <https://interestingengineering.com/nfts-are-mysteriously-disappearing-heres-how> (Accessed: Jan. 25, 2022).

NIR KSHETRI is the "Computing's Economics" column editor and a professor in the Bryan School of Business and Economics, University of North Carolina at Greensboro, Greensboro, North Carolina, 27412, USA. Contact him at nbkshetr@uncg.edu.



IEEE MultiMedia serves the community of scholars, developers, practitioners, and students who are interested in multiple media types and work in fields such as image and video processing, audio analysis, text retrieval, and data fusion.

Read It Today!

www.computer.org/multimedia



Cryptographic–Biometric Self-Sovereign Personal Identities

Doron Drusinsky, Naval Postgraduate School

This article describes a hybrid of self-sovereign identities (SSIs), cryptographic authentication, and biometric authentication, which allows remote yet secure proof of identity, all while addressing privacy concerns. The proposed technique addresses cooperative attacks, which are unique to the SSI embodiment of blockchain technology.

The emergence of online services provides increasing opportunities for personal identities to be utilized online for routine activities, such as online shopping, social networking, online banking, and online political activities. The two prevailing identity management models are centralized and federated.

Centralized identities usually rely on a chain of trust of other centralized identities. For example, a corporate or government common access card (CAC) is issued after the person being identified (Pbi) presents a state-certified (centralized) identity document, such as a driver's license or passport; airport Transportation Security Administration screening is an application that requires a state-issued identity, while entering the premises of a large corporation requires a corporate ID. At the present time, proof of identity using a state-certified identity requires the proof of possession of a hard-to-forge identity document ("something I have"); supporting documentation is sometimes required as well (for example, for REAL ID), such as phone, utility, or property tax bills.

Clearly, whenever such a centralized identity is used, it is also exposed. Also clearly, the state or corporate issuing entity has full control over the issuance and renewal of the identities it certifies; in fact, the Pbi has no access to the records that contain his or her identity.

The upside of using centralized identities is that they are widely trusted. CAC cards enable remote authentication using a cryptographic scheme called *challenge–response* (CR), also known as the *Sigma protocol*¹: the CAC card contains a chip that stores a unique secret key for the Pbi, which is used to sign a random number the server generates; the server then verifies the signature using the Pbi's public key included in the Pbi's public identity certificate. That certificate, signed by a trusted certificate authority (whose own certificate is further signed in a chain of trust),² contains the Pbi's identity. Consequently, the CR procedure binds the three artifacts: the private key, the public key, and Pbi's identity.

A federated identity is based on a digital identity issued by some well-known online "prime" entity, such as Facebook or Google, that is subsequently used as an attestation of trust to further sign on to some other online service. The (limited) trust provided by a federated identity is based on the prime entity limiting the creation of identities so that only humans (or corporations) can create identities, with each creating, at most, one such identity.

Clearly, the prime entity has control over the identity in that it can remove the identity or block access to it at any time. In addition, prime entities have not been effective in limiting identities to being human Pbis; for example, bots have been known to have used fake Facebook accounts.³ It is also quite easy for an individual to create a plurality of Facebook accounts using a plurality of email accounts. Hence, this category of identities induces only limited trust. For example,



most readers could trust that a Facebook entity posting images of a known solar eclipse would be using a genuine identity. On the other hand, a Facebook entity posting images of stolen property has a reason to hide his or her true identity.

Question-and-answer (Q&A) sessions are a technique for establishing the trustworthiness of a claimed identity using randomly generated questions pertaining to the Pbl's history, such as former residences, the amount of debt owed, property ownership, or similar information about relatives. While a dishonest Pbl can possibly find the correct answers when allowed sufficient time, a Q&A session is typically time limited.

This article is concerned with a third kind of identity called a *self-sovereign identity* (SSI). SSIs are not controlled by or stored with a centralized entity or prime entity behind a federated identity. Rather, as described in the next section, SSIs are stored in a public, decentralized blockchain.

Allende López⁴ presents a high-level overview of SSIs using blockchain. In this article, I present a more detailed explanation of the blockchain implementation of SSIs with a special emphasis on remote protocols for remote biometric authentication.

The Sovrin Foundation⁵ offers the Sovrin Identity Network, a public permissioned ledger for SSIs, and describes itself as a global nonprofit organization "whose sole purpose is governance of this ledger and its surrounding ecosystem." Hence, the Sovrin Foundation is a nonprofit organization dedicated to the realization of an identity network that cannot be owned or controlled by any single company, organization, or government.

Centralized identities are always accompanied by some form of biometric authentication, such as a photograph or fingerprints. When using a centralized identity for onsite verification, such as during a border crossing, automated and manual techniques are used to correlate the assumed Pbl with the identity document. In contrast, this article is concerned with trusted remote Pbl verification.

A blockchain is a distributed data structure with no governing authority. A large, distributed collection of computers, owned by mostly unrelated entities, maintains a copy of the blockchain and repeatedly updates it using a distributed network formed by random neighbor links. Blockchains do not have a central gatekeeper, like a bank, to verify transactions. Rather, in its digital currency form, described in my recent *Computer* article,⁶ a blockchain is a decentralized ledger of digital transactions in which trust is obtained by achieving consensus. Trust in the blockchain is a result of the distributed system arriving at a consensus that transactions are approved by the payer, the payer has the funds to pay, the payee (one or more) receives the precise amount being paid, and a transaction is not duplicated or emitted from the ledger. Consensus is achieved by a form of crowd-sourced verification, performed by a large plurality of computers called *miners*. Both Bitcoin and Ethereum, the two largest cryptocurrencies, rely on a consensus mechanism using either the energy-wasteful "proof-of-work" technique or a "proof-of-stake" technique to maintain a trustable ledger of transactions.

Blockchain technology is emerging as a technology capable of implementing SSIs. The first part of the article describes a blockchain-based SSI implementation approach and some related concerns. The section "Biometric Proof of SSI Ownership" describes methods of establishing the unique link between the Pbl (the human entity), referred to as *Bob*, and his SSI using an integration of biometric authentication with blockchain-based SSIs. Such a link will prevent the use or abuse of Bob's SSI by someone else who either obtains Bob's secret key or, perhaps, even uses Bob's device (with or without permission) to masquerade as Bob.

In addition to its security-related qualities, the suggested method also ensures that Bob is alive when performing the transaction. This is especially important for applications such as voting or attestation for Social Security benefits.

TABLE 1. Blockchain use cases and associated properties.

Property use case	Split or combine	Transfer/sale as a whole	Update attributes	Onboarding initiated by user	Onboarding verified by blockchain
Cryptocurrency	Can split and combine coins	Yes	No	No	No
NFT	No	Yes	No	Yes	No
SSI	No	No	Yes	Yes	Yes

BLOCKCHAIN-BASED SSIs

In its SSI use case, a blockchain stores digital identities rather than transactions. Having identities stored in a decentralized ledger means that there is no governing entity that controls the use and verification of stored identities. Unlike their cryptocurrency counterparts, identities are nonfungible, that is, they cannot be traded, exchanged, or split. Hence, an SSI is a form of nonfungible token (NFT) that cannot be further exchanged. Table 1 lists the fungibility, onboarding, and modification properties of blockchain records for the three abovementioned blockchain use cases.

Like its cryptocurrency and NFT blockchain counterparts, an SSI identity needs to be verified to be added to the blockchain. More specifically, it needs to be verified as being genuinely true, having no duplication, and not being omitted from the ledger. Of these properties, it is only the first that is unique to the SSI blockchain use case.

We identify the following SSI-based blockchain operations:

1. onboarding an SSI into the blockchain
2. proof of identity using SSI
3. change of attributes (the change of the Pbl's personal identification attributes within the blockchain, such as his or her address, marital status, children, and so on, or the change of biometric models associated with the Pbl as a result of retraining the Pbl's biometric authenticator)
4. SSI removal from the blockchain
5. the recovery of private keys.

ONBOARDING AN SSI

An SSI identity is verified to be genuinely true during onboarding, that is, when a new identity is added to

the blockchain. When Bob wants to create his SSI, he initiates the process, adds attributes (such as his address and marital status), and then signs the transaction. At this point, Bob's identity is but a request; Bob still needs to provide some proof that he owns this SSI, or else his self-proclaimed SSI will not be trusted as being genuine.

Suppose M , Mindy's machine, is a subsequent blockchain verifier (a miner) that is tasked with verifying this and other SSI requests contained in the current block. M can rely on a centralized identity (for example, perform a CR session with the CAC card), use a federated identity verification, or use a Q&A session. Clearly, the level of trust induced by the verification (for instance, a federated identity is less trustworthy than a centralized one) should become an SSI attribute as well.

Suppose M uses a centralized identity as a proof of Bob's identity. Clearly, such an onboarding process is not self-sovereign.

When verification depends on Bob's interaction with N , as is the case when Bob participates in a Q&A verification session, there is an obvious potential for a dishonest verifier M_d to create fake identities by creating a fake interaction, such as a fake Q&A session. The blockchain's distributed consensus mechanism protects against such an attack using a slight tweak, as follows: when a new identity is suggested for verification, rather than it being part of a single block B of identities being verified, as in the Bitcoin case, it is now added to k blocks B_1, \dots, B_k ; this means that k successive miners M_1, M_2, \dots, M_k will each perform their own verification, each with its own Q&A session.

Suppose a fake identity I_{Fake} is distributed for verification by a dishonest miner M_d . I_{Fake} will be added to a sequence of k blocks B_1, \dots, B_k where verifiers M_d, M_2, \dots, M_k each perform their own Q&A sessions on the SSIs in their corresponding block, including I_{Fake} . As depicted in Figure 1, an honest verifier M_h ($2 \leq h \leq k$) will extend the blockchain with block B_h that does not contain I_{Fake} but will not extend a chain that ends with a block that does contain I_{Fake} . The proof-of-work and proof-of-stake penalty associated with being dishonest ensures that most of the M_i s online are honest; because the blockchain favors longer chains over

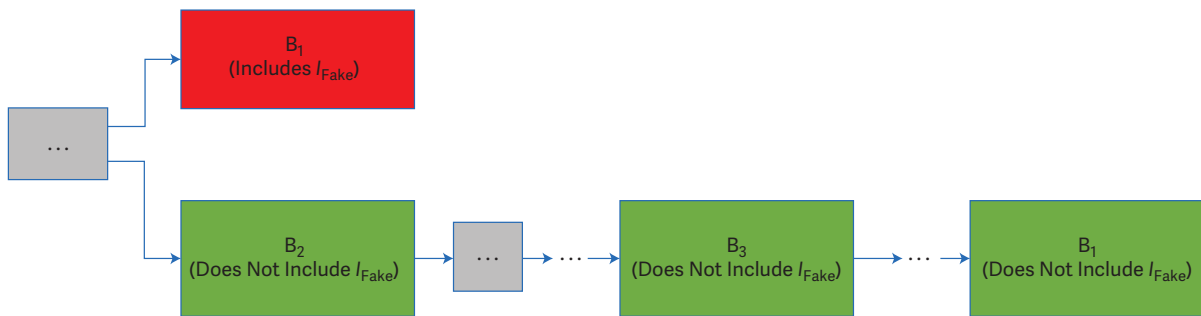


FIGURE 1. An honest verifier M_h ($2 \leq h \leq k$) will extend the blockchain with block for B_h that does not contain I_{Fake} but will not extend a chain that ends with a block that does contain I_{Fake} .

shorter ones, then I_{Fake} is effectively eliminated.

Note that, because Bob (the Pbl) needs time to answer questions in k Q&A sessions, we cannot expect blocks B_1, \dots, B_k to be successive. Therefore, the blockchain must provide a mechanism for verifier M_i to quickly locate preceding blocks in the blockchain (from those created within the past day, for example) that contain a suggested identity I ; a good search parameter is the public key stored in I .

An SSI record, therefore, contains the following:

- › the name and, possibly, other public attributes, such as the address
- › (potential) hidden centralized identity attributes, such as the hashed social security number
- › the Pbl's public key
- › the Pbl's public-key signature
- › verifier signatures and related public keys.

PROOF OF IDENTITY USING SSIs

As with cryptocurrency and NFTs, the owner of the SSI has a secret key in his or her possession, while the corresponding public key, also called the *SSI address*, is stored as part of the immutable SSI. Suppose Bob applies for a loan from the bank, and Alice at the bank is trying to prove that the person applying (in person

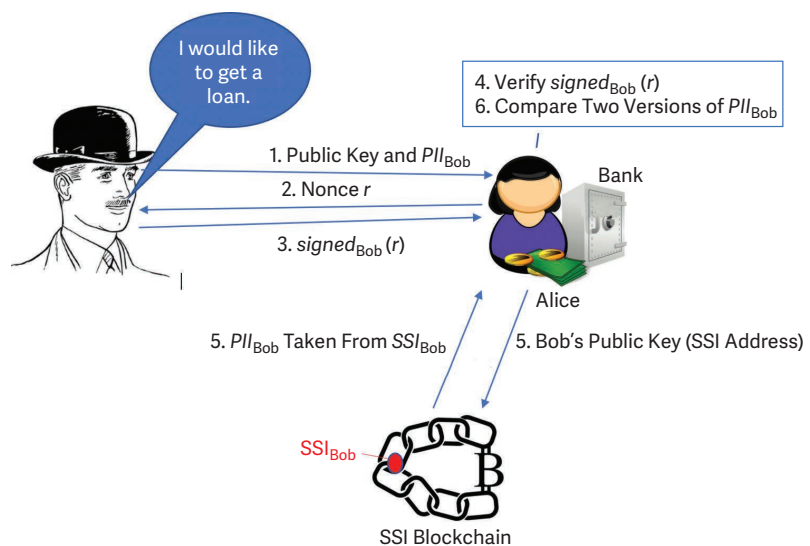


FIGURE 2. The process for *prove(Bob)*.

or remotely) is Bob, i.e., that his SSI identity is his personally identifiable information (PII), PII_{Bob} , that he supplied to Alice. Figure 2 depicts the proof using a CR session called *prove(Bob)*, based on the common Sigma cryptographic CR protocol^{1,7}:

1. Bob provides Alice with his public key, which is the blockchain address for his SSI (SSI_{Bob}), denoted $addr(SSI_{Bob})$. He also provides Alice with his PII, denoted PII_{Bob} .
2. Alice sends Bob a random number r .
3. Bob signs r with his secret key.
4. Alice verifies r and the signature using Bob's public key.

TABLE 2. The reasoning for the components of $prove(Bob)$.

Component	Reasoning
CR (steps 1–4 of Figure 2)	Prove that Bob's machine is Alice's counterpart in $prove(Bob)$ and not some adversary's machine that holds PII_{Bob} .
SSI_{Bob} comparison (steps 5 and 6 of Figure 2)	Prove that the claimed PII_{Bob} supplied by Bob's machine corresponds with the PII_{Bob} stored on SSI_{Bob} per the SSI address given by Bob's public key (which corresponds to Bob's private key stored on Bob's machine).

TABLE 3. The reasoning for the components of $prove_{ACERT}(Bob)$.

Component	Reasoning
Steps 1–4 of Figure 2	Prove that Bob's machine is Alice's counterpart in $prove_{ACERT}(Bob)$ and not some adversary's machine that holds PII_{Bob} .
Steps 5 and 6 of Figure 2	Prove that the claimed PII_{Bob} supplied by Bob's machine corresponds with the PII_{Bob} stored on SSI_{Bob} per the SSI address given by Bob's public key (which corresponds to Bob's private key stored on Bob's machine).
Abovementioned steps 1', 5', and 6'	Prove that Bob is holding BM_{Bob} when performing steps 1–6 of Figure 2.

5. Alice obtains SSI_{Bob} from the blockchain using $addr(SSI_{Bob})$.
6. Alice compares the two versions of PII_{Bob} : the one obtained from Bob and the one stored in the blockchain. Alice obtains a level of trust in SSI_{Bob} based on its trust attribute.

Table 2 contains the reasoning for applying each component of $prove(Bob)$.

BIOMETRIC PROOF OF SSI OWNERSHIP

The proof of identity $prove(Bob)$, described in the previous section, proves that the secret key stored in Bob's machine (M_{Bob}) corresponds to the public key stored in Bob's blockchain SSI; it does not address the question of whether it is actually Bob who is operating or holding M_{Bob} . This article suggests augmenting the SSI with biometric authentication. We propose two such augmentations schemes: *ACert* and *BodyCERT*.

Biometric proof of SSI ownership using *ACert*

ACert is the following rudimentary method for binding an individual's biometrics with his or her SSI: Bob's SSI blockchain record is expanded to contain a hash of his biometric authentication models,

denoted as $hash(BM_{Bob})$. As with other aspects of a blockchain record, or transaction, such a record is practically immutable; that is, an attacker cannot change it to replace $hash(BM_{Bob})$ with $hash(BM_{Attacker})$.

For verification using *ACert*, $prove(Bob)$ is modified to become $prove_{ACERT}(Bob)$ as follows:

- › **Step 1':** In step 1 of Figure 2, Bob includes a signed hash of the biometric authentication models used to authenticate himself [$hash(BM_{Bob})$] for $prove_{ACERT}$.
- › **Step 5':** In step 5 of Figure 2, Alice (M_{Alice}) also uses $hash(BM_{Bob})$ taken from SSI_{Bob} .
- › **Step 6':** In step 6 of Figure 2, Alice (M_{Alice}) adds a comparison: she compares the two versions of $hash(BM_{Bob})$.

Table 3 contains the reasoning for applying each component of $prove_{ACERT}(Bob)$. Clearly, when applying the abovementioned $prove_{ACERT}(Bob)$ process, one needs to trust the biometric authentication application on M_{Bob} to send BM_{Bob} to M_{Alice} if and only if Bob authenticates correctly.

Biometric proof of SSI ownership using *BodyCERT*

The $prove(Bob)$ and $prove_{ACERT}(Bob)$ processes suffer from an inherent weakness: it is not Bob the Pbl who performs the proof but, rather, M_{Bob} . While this weakness exists for cryptocurrency and NFT blockchains as well, there is one attack vector that is unique to an SSI—a cooperative attack, described as follows: Bob allows Carrie to use his device, and Carrie can thereafter claim to be Bob, using his SSI for the purposes of medical insurance fraud or some other illegal identity-based activity. A cooperative attack is unique to SSIs because, if Bob cooperates with Carrie to claim his NFT, for example, then this cannot really be considered as an attack: the resulting transaction was effectively done with Bob's blessing.

TABLE 4. The reasoning for the components of $prove_{BodyCERT}(Bob)$.

Component	Reasoning
Steps 1–4 of Figure 2	<ul style="list-style-type: none"> • Prove that Bob’s machine is Alice’s counterpart in $prove_{BodyCERT}(Bob)$ and not some adversary’s machine that holds PII_{Bob}. • Prove that Bob is holding BM_{Bob} when performing steps 1–6 of Figure 2.
Steps 5 and 6 of Figure 2	Prove that the claimed PII_{Bob} supplied by Bob’s machine corresponds with the PII_{Bob} stored on SSI_{Bob} per the SSI address given by Bob’s public key.

BodyCERT, the proposed solution to the cooperative attack problem, integrates biometrics with the key pairs, in contrast to *ACert*’s approach, where the biometrics are stored in SSI_{Bob} as adjacent attributes. This technique was detailed in my *Computer* article⁸; a short summary of its application process on Bob’s (mobile) device follows.

During biometric authentication training,

1. Generate a nonce, a random number for Bob’s device. This number is referred to as the *golden bitmask*, or the *Golden*.
2. Replace Bob’s secret key S_{Bob} with $S_{Bob}' = S_{Bob} \oplus Golden_{Bob}$, where \oplus is the exclusive-or operator.
3. Perform biometric authentication training on Bob’s device. Such training uses an array of m -bit classifiers, such as trees of a random decision forest⁹ or an ensemble of random forests; for simplicity, let’s assume it uses the first approach.
4. For each index $i = 0, \dots, m$, if $Golden_{Bob}[i] = 0$, then flip bit classifier $[i]$; that is, have the leaves return a flipped value (one instead of zero, and vice versa).
5. Apply error-correction encoding to $Golden_{Bob}$, resulting in a vector of error-correction bits denoted as $ECC_{Golden-bob}$.

Note how the secret key S_{Bob}' stored on Bob’s machine no longer matches the public key counterpart of S_{Bob} and cannot be used to generate a valid signature. Hence, when applying a cooperative attack, Carrie cannot use Bob’s device to impersonate Bob. Bob, however, can nonetheless generate a valid signature as follows.

For verification using *BodyCERT*, $prove(Bob)$ is modified to become $prove_{BodyCERT}(Bob)$:

1. Bob performs biometric authentication, thereby generating an m -bit vector V_{Bob} .
2. Bob applies error correction to V_{Bob} using $ECC_{Golden-bob}$, resulting in a so called actual

bitmask, abbreviated as $Actual_{Bob}$. The expectation, as I explained in a previous article,⁸ is that $S_{Bob} = S_{Bob}' \oplus Actual_{Bob}$.

3. Bob executes all steps of $prove(Bob)$, depicted in Figure 2, using the recovered S_{Bob} . The recovered S_{Bob} exists only in the runtime memory of the $prove()$ method; it is never serialized to long-term storage. In other words, the recovered S_{Bob} is volatile.

Carrie’s cooperative attack can be done in two ways: either Carrie performs a complete takeover of M_{Bob} , including the training of classifiers, or she simply uses M_{Bob} as is. In both cases, Carrie needs to recreate S_{Bob} from S_{Bob}' to identify as Bob. The difference between the two approaches, however, is that $ECC_{Golden-bob}$ is replaced with $ECC_{Golden-carrie}$ in the first case, while it remains intact in the second case. Hence, Carrie has a higher likelihood of success if she simply uses M_{Bob} as is, assuming Bob lets Carrie authenticate into it. In this case, as I explained in an earlier article,⁸ if the Hamming distance¹⁰ between V_{Carrie} and V_{Bob} is sufficiently small, then error correction using $ECC_{Golden-bob}$ could, potentially, regenerate S_{Bob} from S_{Bob}' using $Actual_{Carrie}$.

One form of armor-plating *BodyCERT* against such an attack is to use a rainbow of bit classifiers taken from a plurality of classification technologies—for example, having some bit classifiers implemented as random forest classifiers, others as orthogonal random features¹¹ classifiers, and yet others as support vector machine (SVM) classifiers.¹² While the Hamming distance between V_{Carrie} and V_{Bob} is obviously Pbl dependent (for example, V_{Carrie} might be close to V_{Bob} , while V_{David} is not), it is also technology dependent, such as random forests having a harder time distinguishing between Carrie and Bob than

SVM. Hence, the use of a plurality of classification technologies is expected to decrease the Hamming distance between V_{Carrie} and V_{Bob} . Table 4 contains the reasoning for applying each component of $\text{prove}_{\text{BodyCERT}}(\text{Bob})$.

CHANGE OF SSI ATTRIBUTES

There are two types of attribute changes: one uses biometric verification [for example, $\text{prove}_{\text{BodyCERT}}(\text{Bob})$], and the other involves a change of the biometric verification attributes themselves. To change PII attributes like the address, marital status, or issue date, Bob—the Pbl—would initiate a blockchain transaction by

THERE ARE TWO TYPES OF ATTRIBUTE CHANGES: ONE USES BIOMETRIC VERIFICATION [FOR EXAMPLE, $\text{prove}_{\text{BodyCERT}}(\text{Bob})$], AND THE OTHER INVOLVES A CHANGE OF THE BIOMETRIC VERIFICATION ATTRIBUTES THEMSELVES.

which his PII is changed to the new PII. The blockchain verifier/miner would verify the request by requesting proof that this is, indeed, Bob who is requesting the change using $\text{prove}_{\text{ACert}}(\text{Bob})$ or $\text{prove}_{\text{BodyCERT}}(\text{Bob})$. Note that the protocol can be extended so that Bob also adds some necessary supporting documents when he initiates the transaction.

There are two distinct scenarios in which Bob needs to change his biometric authentication: either Bob still has the device on which he can perform $\text{prove}_{\text{ACert}}(\text{Bob})$ or $\text{prove}_{\text{BodyCERT}}(\text{Bob})$, or he does not. The first case is but a special case of the PII attribute change discussed earlier. The second case requires a “reset,” that is, the removal of SSI_{Bob} followed by the fresh onboarding of a new SSI for Bob.

OTHER SSI-RELATED ACTIONS

Due to brevity constraints, we do not address other SSI-related actions, such as the removal of an SSI record and loss of private keys. These actions are described in the SSI literature (for example, by Allende López⁴).

PRIVACY CONCERNS

Being distributed and public, a blockchain SSI repository introduces obvious privacy issues. For example, Bob’s name, address, date of birth, and other PII (PII_{Bob}) would be available to any computer worldwide that downloads a copy of the blockchain. A simple way to mitigate this is to store a hash of Bob’s PII, denoted $\text{hash}(\text{PII}_{\text{Bob}})$, in the blockchain rather than the raw data. Suppose that Bob needs to identify himself using PII_{Bob} , such as in the loan example for Figure 2; he would then need to provide PII_{Bob} to the bank. A concern related to this approach is, therefore, that any future data breach at the bank would, effectively, allow an adversary to decrypt $\text{hash}(\text{PII}_{\text{Bob}})$. One solution to this concern is to add a field like “issue-date” PII_{Bob} ; clearly, any change to the “issue-date” attribute alone will induce a very different $\text{hash}(\text{PII}_{\text{Bob}})$. Hence, when Bob is notified of a data breach concerning PII_{Bob} , he can issue an attribute change transaction to SSI_{Bob} , as described in the “Change of SSI Attributes” section.

ACert also introduces an additional concern, as follows. ACert includes $\text{hash}(\text{BM}_{\text{Bob}})$ as part of the highly visible SSI for Bob. Such exposed $\text{hash}(\text{BM}_{\text{Bob}})$ allows David, an adversary, to onboard a new SSI record with $\text{PII}_{\text{David}}$ accompanied with $\text{hash}(\text{BM}_{\text{Bob}})$; this would be somewhat like the Division of Motor Vehicles allowing David to obtain a driver’s license with Bob’s photo on it.

Hence, we suggest two changes, one to the SSI blockchain record content and the other to the $\text{prove}_{\text{private}}(\text{Bob})$ protocol, as follows:

- Instead of storing $\text{hash}(\text{BM}_{\text{Bob}})$ in the SSI_{Bob} record, ACert stores the elliptic-curve point $Q_{\text{Bob}} = G \times \text{hash}(\text{BM}_{\text{Bob}})$, where G is the elliptic-curve generator, and \times is the product of the point G with the scalar $\text{hash}(\text{BM}_{\text{Bob}})$. Recall that the SSI_{Bob} record also contains Bob’s public key, which is mathematically equivalent to $G \times S_{\text{Bob}}$. Hence, Q_{Bob} can be thought of as being a public key whose private key counterpart is $\text{hash}(\text{BM}_{\text{Bob}})$.
- $\text{prove}_{\text{private}}(\text{Bob})$ does not send $\text{hash}(\text{BM}_{\text{Bob}})$ to Alice, signed or otherwise. Instead, $\text{prove}_{\text{private}}(\text{Bob})$ incorporates a second CR session in which the random number r is signed by

M_{Bob} using $\text{hash}(BM_{\text{Bob}})$ as the private key. Alice verifies the signature using Q_{Bob} as the public key, where Q_{Bob} is taken from SSI_{Bob} .

This article described a viable, self-sustainable identity framework implemented using blockchain technology integrated with remote cryptographic signatures and biometric authentication from a live biometric signal. The proposed approach protects privacy and against the misuse of identities in the form of cooperative attacks. 🌐

ACKNOWLEDGMENT

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. government. The U.S. government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright annotations thereon.

REFERENCES

1. "Proof of knowledge." Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Proof_of_knowledge#Sigma_protocols (Accessed: Mar. 28, 2022).
2. "Chain of trust." Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Chain_of_trust (Accessed: Mar. 28, 2022).
3. K. Krombholz, D. Merkl, and E. Weippl, "Fake identities in social media: A case study on the sustainability of the Facebook business model," *J. Service Sci. Res.*, vol. 4, no. 2, pp. 175–212, 2012, doi: 10.1007/s12927-012-0008-z
4. M. A. López, "Self-sovereign identity: The future of identity: Self-sovereignty, digital wallets, and blockchain," Inter-American Development Bank, Washington, DC, USA, 2020. Accessed: Mar. 28, 2022. [Online]. Available: <https://publications.iadb.org/en/self-sovereign-identity-future-identity-self-sovereignty-digital-wallets-and-blockchain>
5. "The inevitable rise of self-sovereign identity," The Sovrin Foundation, Provo, UT, USA, White Paper, 2017. [Online]. Available: <https://sovrin.org/wp-content/uploads/2018/03/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf> (Accessed: Mar. 28, 2022).
6. D. Drusinsky, "On the high-energy consumption of bitcoin mining," *Computer*, vol. 55, no. 1, pp. 88–93, 2022, doi: 10.1109/MC.2021.3123781.
7. R. Canetti and H. Krawczyk, "Security analysis of IKE's signature-based key-exchange protocol," in *Advances Cryptology — CRYPTO 2002*, vol. 2442. M. Yung, Ed. Berlin, Germany: Springer-Verlag, 2002, pp. 143–161.
8. D. Drusinsky, "Who is authenticating my e-commerce logins?" *Computer*, vol. 54, no. 4, pp. 15–24, 2021, doi: 10.1109/MC.2021.3055684.
9. G. Pierre and L. Gilles, "Learning to rank with extremely randomized trees," in *Proc. JMLR: Workshop Conf.*, 2011, vol. 14, pp. 49–61.
10. "Hamming distance." Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Hamming_distance (Accessed: Mar. 28, 2022).
11. F. X. X. Yu, A. T. Suresh, K. M. Choromanski, D. N. Holtmann-Rice, and S. Kumar, "Orthogonal random features," in *Proc. 30th Conf. Neural Inf. Process. Syst. (NIPS 2016)*, Barcelona, Spain, 2016, pp. 1983–1991.
12. M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul./Aug. 1998, doi: 10.1109/5254.708428.
13. "Man-in-the-middle attack." Wikipedia. Accessed: May 1, 2022. [Online]. Available: https://en.wikipedia.org/wiki/Man-in-the-middle_attack
14. "Real ID." DMV.CA.GOV. [Online]. Available: https://www.dmv.ca.gov/portal/driver-licenses-identification-cards/real-id/?gclid=Cj0KCQjw5-WRBhCKARIsAId9FIFbBmJx3xIN7YaOE5X1GGbKfKk750bjnTXQec4fjLHPmQL_9ZUs4aAsuoEALw_wcB (Accessed: Mar. 2022).
15. F. Boudot, B. Schoenmakers, and J. Traoré, "A fair and efficient solution to the socialist millionaires' problem," *Discrete Appl. Math.*, vol. 111, nos. 1–2, pp. 23–36, Jul. 15, 2001, doi: 10.1016/S0166-218X(00)00342-5.
16. "Socialist millionaire problem." Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Socialist_millionaire_problem (Accessed: Mar. 29, 2022).

DORON DRUSINSKY is a professor in the Naval Postgraduate School's Department of Computer Science, Monterey, California, 93943, USA. Contact him at ddrusins@nps.edu.

Career Accelerating Opportunities

Explore new options—upload your resume today

careers.computer.org



Changes in the marketplace shift demands for vital skills and talent. The **IEEE Computer Society Career Center** 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 Career Center keeps you connected to workplace trends and exciting career prospects.



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

Submit a paper to the new IEEE Open Journal of the Computer Society covering computing and informational 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 TRANSACTIONS ON BIG DATA

IEEE Transactions on Big Data is a quarterly journal that publishes peer-reviewed articles with big data as the main focus.

The articles provide cross-disciplinary, innovative research ideas and applications results for big data including novel theory, algorithms, and applications. Research areas include:

- Big data
 - Analytics
 - Curation and management
 - Infrastructure
 - Performance analyses
 - Semantics
 - Standards
 - Visualization
- Intelligence and scientific discovery from big data
- Security, privacy, and legal issues specific to big data

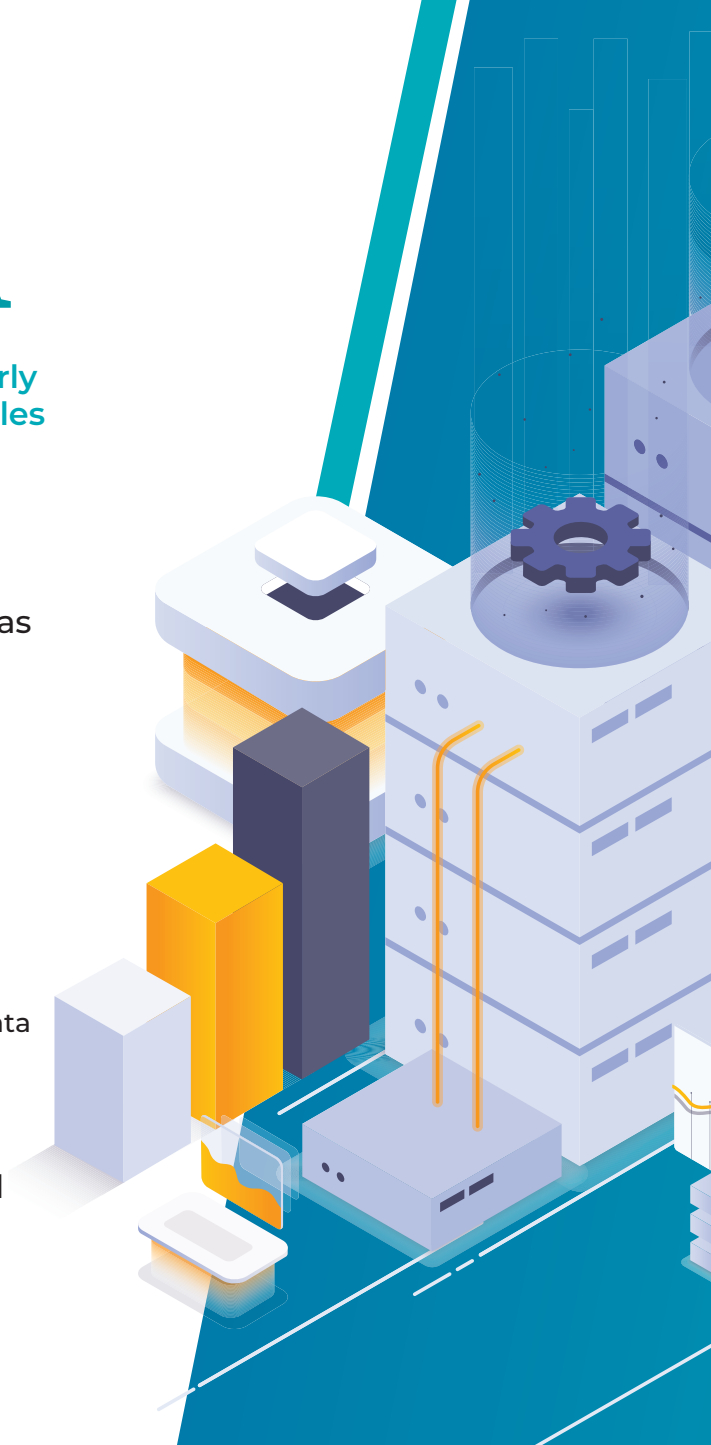
Applications of big data in the fields of endeavor where massive data is generated are of particular interest.

www.computer.org/tbd



Join the IEEE Computer Society
for subscription discounts today!

www.computer.org/product/journals/tbd



**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

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



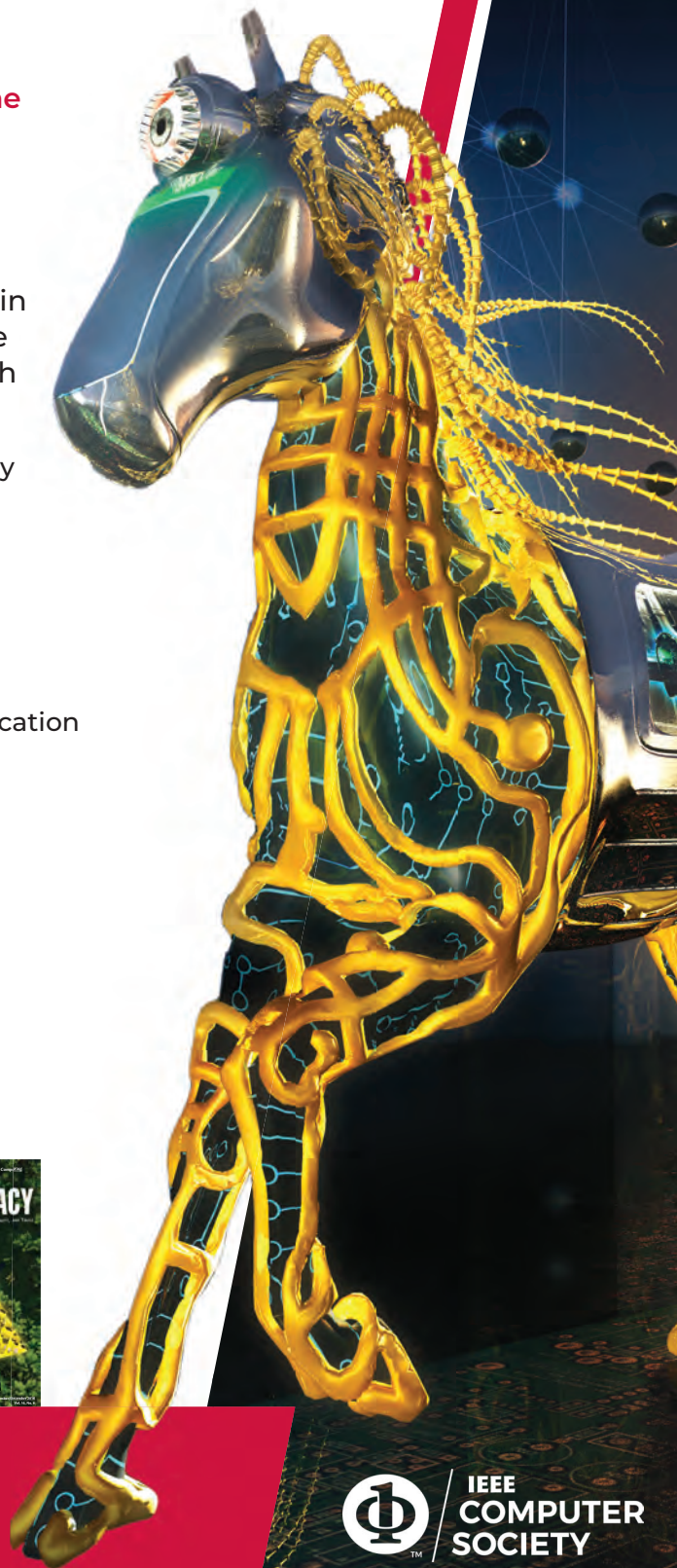
Join the IEEE Computer Society
for subscription discounts today!

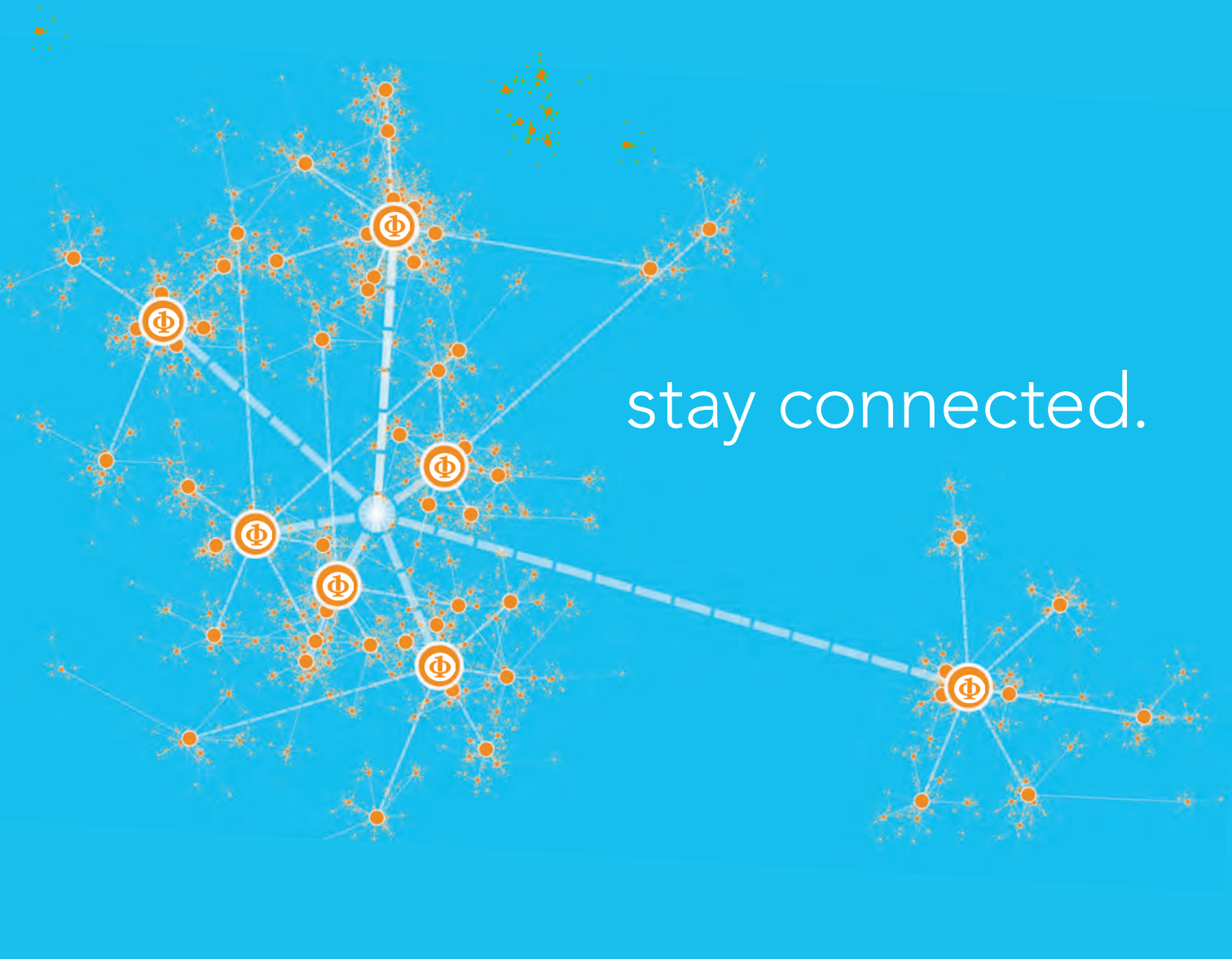
www.computer.org/product/magazines/security-and-privacy



IEEE
COMPUTER
SOCIETY

IEEE





stay connected.

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 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

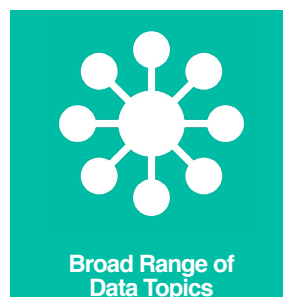
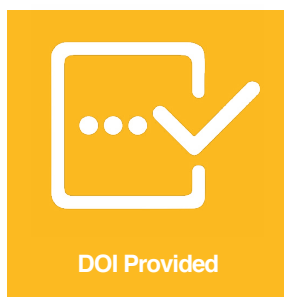
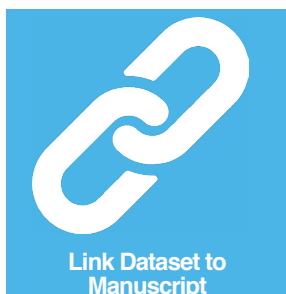


IEEE
COMPUTER
SOCIETY



SHARE AND MANAGE YOUR RESEARCH DATA

IEEE DataPort is an accessible online platform that enables researchers to easily share, access, and manage datasets in one trusted location. The platform accepts all types of datasets, up to 2TB, and dataset uploads are currently free of charge.



IEEE*DataPort*[™]

UPLOAD DATASETS AT [IEEE-DATAPORT.ORG](https://www.ieee-dataport.org)

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





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.

JUNE

3 June

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

4 June

- ICSA (IEEE Int'l Conf. on Software Architecture), Hyderabad, India
- WoWMoM (IEEE Int'l Symposium on a World of Wireless, Mobile and Multimedia Networks), Perth, Australia

10 June

- ARITH (IEEE Symposium on Computer Arithmetic), Malaga, Spain

16 June

- CVPR (IEEE/CVF Conf. on Computer Vision and Pattern Recognition), Seattle, USA

17 June

- SVCC (Silicon Valley Cybersecurity Conf.), Seoul, South Korea

19 June

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

24 June

- DSN (IEEE/IFIP Int'l Conf. on Dependable Systems and Networks), Brisbane, Australia
- MDM (IEEE Int'l Conf. on Mobile Data Management), Brussels, Belgium

- RE (IEEE Int'l Requirements Eng. Conf.), Reykjavik, Iceland

25 June

- CAI (IEEE Conf. on Artificial Intelligence), Singapore

26 June

- CBMS (IEEE Int'l Symposium on Computer-Based Medical Systems), Guadalajara, Mexico

27 June

- CS (IEEE Cloud Summit), Washington, DC, USA (Hybrid)

29 June

- ISCA (ACM/IEEE Annual Int'l Symposium on Computer Architecture), Buenos Aires, Argentina

JULY

1 July

- ICALT (IEEE Int'l Conf. on Advanced Learning Technologies), Nicosia, Cyprus

2 July

- COMPSAC (IEEE Annual Computers, Software, and Applications Conf.), Osaka, Japan

3 July

- IOLTS (IEEE Int'l Symposium on On-Line Testing and Robust System Design), Rennes, France

7 July

- SERVICES (IEEE World Congress on Services), Shenzhen, China

8 July

- CSF (IEEE Computer Security Foundations Symposium), Enschede, Netherlands
- EuroS&P (IEEE European Symposium on Security and Privacy), Vienna, Austria

15 July

- CISOSE (IEEE Int'l Congress On Intelligent and Service-Oriented Systems Eng.), Shanghai, China
- ICME (IEEE Int'l Conf. on Multimedia and Expo), Niagara Falls, Canada
- SCC (IEEE Space Computing Conf.), Mountain View, USA
- SMC-IT (IEEE Int'l Conf. on Space Mission Challenges for Information Technology), Mountain View, USA

22 July

- ICCP (IEEE Int'l Conf. on Computational Photography), Lausanne, Switzerland

23 July

- ICDCS (IEEE Int'l Conf. on Distributed Computing Systems), Jersey City, USA

24 July

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



AUGUST

7 August

- IRI (IEEE Int'l Conf. on Information Reuse and Integration for Data Science), San Jose, USA
- MIPR (IEEE Int'l Conf. on Multimedia Information Processing and Retrieval), San Jose, USA

19 August

- Blockchain (IEEE Int'l Conf. on Blockchain), Copenhagen, Denmark
- Cybermatics (IEEE Congress on Cybermatics), Copenhagen, Denmark

21 August

- HOTI (IEEE Symposium on High-Performance Interconnects), virtual
- RTCSA (IEEE Int'l Conf. on Embedded and Real-Time Computing Systems and Applications), Sokcho, South Korea

25 Aug

- HCS (IEEE Hot Chips Symposium), Stanford, USA

27 Aug

- SustainTech (IEEE SustainTech Expo: Technology Solutions for a Sustainable Future), San Diego, USA

SEPTEMBER

2 September

- VL/HCC (IEEE Symposium on Visual Languages and Human-Centric Computing), Liverpool, UK

15 September

- IISWC (IEEE Int'l Symposium

on Workload Characterization), Vancouver, Canada

- QCE (IEEE Int'l Conf. on Quantum Computing and Eng.), Montreal, Canada

16 September

- ACSOS (IEEE Int'l Conf. on Autonomic Computing and Self-Organizing Systems), Aarhus, Denmark
- e-Science (IEEE Int'l Conf. on e-Science), Osaka, Japan

23 September

- MASS (IEEE Int'l Conf. on Mobile Ad-Hoc and Smart Systems), Seoul, South Korea

24 September

- CLUSTER (IEEE Int'l Conf. on Cluster Computing), Kobe, Japan
- IC2E (2024 IEEE Int'l Conf. on Cloud Eng.), Paphos, Cyprus

OCTOBER

6 October

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

7 October

- SecDev (IEEE Secure Development Conf.), Pittsburgh

8 October

- DFT (IEEE Int'l Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems), Didcot, United Kingdom
- LCN (IEEE Conf. on Local Computer Networks), Normandy, France

11 October

- ICEBE (IEEE Int'l Conf. on e-Business Eng.), Didcot, United Kingdom

13 October

- VIS (IEEE Visualization and Visual Analytics), Tampa Bay, USA

20 October

- FOCS (IEEE Annual Symposium on Foundations of Computer Science), Chicago, USA

21 October

- ISMAR (IEEE Int'l Symposium on Mixed and Augmented Reality), Bellevue, Washington, USA

28 October

- CIC (IEEE Int'l Conf. on Collaboration and Internet Computing), Washington, DC, USA
- CogMI (IEEE Int'l Conf. on Cognitive Machine Intelligence), Washington, DC, USA
- ICNP (IEEE Int'l Conf. on Network Protocols), Charleroi, Belgium
- TPS-ISA (IEEE Int'l Conf. on Trust, Privacy and Security in Intelligent Systems, and Applications), Washington, DC, USA

Learn more about
IEEE Computer Society
conferences
computer.org/conferences



YOU HAVE ENORMOUS RESPONSIBILITY.



Protect yourself from risk.

1-800-493-IEEE (4333)

To learn more*, visit **IEEEinsurance.com/IEEEPL**



IEEE

**PROFESSIONAL LIABILITY
INSURANCE**

Program Administered by AMBA Administrators, Inc.
In CA d/b/a Association Member Benefits & Insurance Agency
CA License #0196562 | AR License #100114462
103213 (5/24) Copyright 2024 AMBA. All rights reserved