



ACM BOOKS Collection II

This organizational history relates the role of the National Science Foundation (NSF) in the development of modern computing. Drawing upon new and existing oral histories, extensive use of NSF documents, and the experience of two of the authors as senior managers, this book describes how NSF's programmatic activities originated and evolved to become the primary source of funding for fundamental research in computing and information technologies.

The book traces how NSF's support has provided facilities and education for computing usage by all scientific disciplines, aided in institution and professional community building, supported fundamental research in computer science and allied disciplines, and led the efforts to broaden participation in computing by all segments of society.

Today, the research and infrastructure facilitated by NSF computing programs are significant economic drivers of American society and industry. The NSF has advanced the development of human capital and ideas for future advances in computing and its applications.

This account is the first comprehensive coverage of NSF's role in the extraordinary growth and expansion of modern computing and its use. It will appeal to historians of computing, policy makers and leaders in government and academia, and individuals interested in the history and development of computing and the NSF.

<http://books.acm.org>
<http://store.morganclaypool.com/acm>



Computing and the National Science Foundation 1950-2016 *Building a Foundation for Modern Computing*

Peter A. Freeman
W. Richards Adrion
William Aspray

ISBN: 978-1-4503-7271-8
DOI: 10.1145/3335772

Computing Reviews

Connect with our Community of Reviewers

“I like CR because it covers the full spectrum of computing research, beyond the comfort zone of one’s specialty. I always look forward to the next Editor’s Pick to get a new perspective.”

- Alessandro Berni



Association for
Computing Machinery

ThinkLoud

www.computingreviews.com

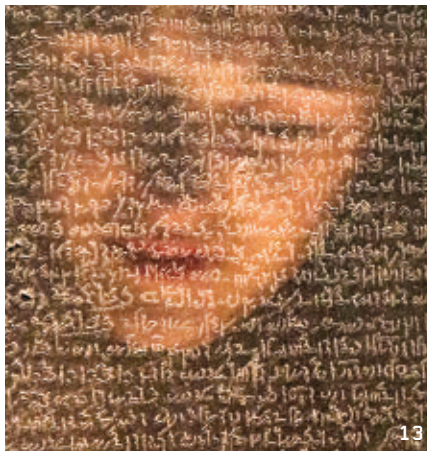
Departments

- 5 **From the President**
New Ways to Think About CS Education
By Cherri M. Pancake
-
- 7 **Cerf's Up**
Multisensory Adventures
By Vinton G. Cerf
-
- 9 **Letters to the Editor**
Adding More Color to Patch Picture
-
- 10 **BLOG@CACM**
Teaching Teachers to Offer STEM to Undergrads
Mark Guzdial considers how teaching computing to undergraduates requires better education of the teachers as well.
-
- 27 **Calendar**

Last Byte

- 128 **Q&A**
Reinventing Virtual Machines
The notion of scalable operating systems led Mendel Rosenblum to virtual machines, which have revolutionized datacenters and enabled modern cloud computing.
By Leah Hoffmann

News



- 13 **Dead Languages Come to Life**
Artificial intelligence automates the translation of extinct languages.
By Gary Anthes
-
- 16 **Machine Learning, Meet Whiskey**
Technologies are coming increasingly closer to approximating the human senses of taste and smell.
By Gregory Mone
-
- 18 **How Universities Deploy Student Data**
Personalizing efforts to drive greater student retention and success.
By Esther Shein

Viewpoints

- 22 **Computing Ethics**
The Temptation of Data-Enabled Surveillance
Are universities the next cautionary tale?
By Alan Rubel and Kyle M.L. Jones
-
- 25 **Technology Strategy and Management**
Artificial Intelligence and the Future of Professional Work
Considering the implications of the influence of artificial intelligence given previous industrial revolutions.
By Mari Sako
-
- 28 **Kode Vicious**
Master of Tickets
Valuing the quality, not the quantity, of work.
By George V. Neville-Neil
-
- 30 **Viewpoint**
Why Is Cybersecurity Not a Human-Scale Problem Anymore?
Examining the structure of the enterprise attack surface in view of the relative ease with which cyberdefenses can be subverted.
By Gaurav Banga
-
- 35 **Viewpoint**
Organizing Family Support Services at ACM Conferences
Seeking to improve access to conferences and provide support for attendees with children.
By Audrey Girouard, Jon E. Froehlich, Regan L. Mandryk, and Mark Hancock
-
- 39 **Viewpoint**
A Taxonomy of Automated Assistants
Rating your intelligent (human or automated) assistant.
By Jerrold M. Grochow

Special Section

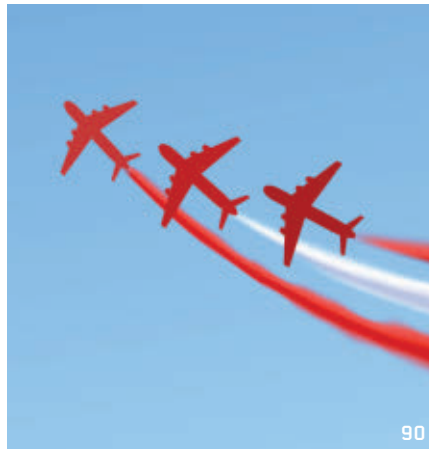


- 44 **East Asia and Oceania Region**
This technology-rich region—including the countries of Southeast Asia, Asia-Pacific, and Oceania—has many success stories to tell. Articles within this section explore an array of emerging technologies, research projects, and real-world experiences in such areas as smart cities, cybersecurity, digital health initiatives, fake news detection, digital heritage preservation, digital economy enhancements, seL4 and 5G deployment, and much more.



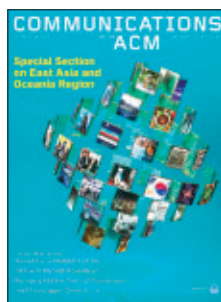
Watch the co-organizers discuss this section in the exclusive *Communications* video. <https://cacm.acm.org/videos/east-asia-and-oceania-region>

Practice



- 90 **Managing the Hidden Costs of Coordination**
Controlling coordination costs when multiple, distributed perspectives are essential.
By Laura M.D. Maguire
- 97 **Cognitive Work of Hypothesis Exploration During Anomaly Response**
A look at how we respond to the unexpected.
By Marisa R. Grayson

Articles' development led by **acmqueue** queue.acm.org



About the Cover: *Communications'* series of regional special sections turns the spotlight this month on the nations that make up East Asia and Oceania. The cover mosaic is a collection of images pulled from the articles within, providing a glimpse of the technologies and projects explored throughout the section. Cover illustration by Spooky Pooka at Debut Art.

IMAGES IN COVER COLLAGE: Shonan Village photo courtesy of Shonan Village/shonan-village.co.jp; Boe Declaration cover courtesy of Pacific Island Forum/forumsee.org; Handprint photo courtesy of Ralph Regenvanu/Twitter; SIOW conference photos courtesy of Singapore International Cyber Week/Twitter.com/SICWSG; Hammock photo by maloff/Shutterstock.com; Exhibit photo by Travel_Adventure/Shutterstock.com; Security scanner photo by Aizuddin Saad/Shutterstock.com; Robot waiter photo by Justin Adam Lee/Shutterstock.com; Singapore airport photo by Damian Lugowski/Shutterstock.com. Additional stock images from Shutterstock.com.

Contributed Articles



- 104 **Cyber Warranties: Market Fix or Marketing Trick?**
Risk transfer options offer hope, but little more.
By Daniel W. Woods and Tyler Moore



Watch the authors discuss their work in the exclusive *Communications* video. <https://cacm.acm.org/videos/cyber-warranties>

Review Articles

- 108 **The Antikythera Mechanism**
The discovery of this calculating machine is so significant that part of the history of ancient technology must be rewritten.
By Herbert Bruderer

Research Highlights

- 118 **Technical Perspective**
An Answer to Fair Division's Most Enigmatic Question
By Ariel D. Procaccia
- 119 **A Bounded and Envy-Free Cake Cutting Algorithm**
By Haris Aziz and Simon Mackenzie



ACM, the world's largest educational and scientific computing society, delivers resources that advance computing as a science and profession. ACM provides the computing field's premier Digital Library and serves its members and the computing profession with leading-edge publications, conferences, and career resources.

Executive Director and CEO

Vicki L. Hanson

Deputy Executive Director and COO

Patricia Ryan

Director, Office of Information Systems

Wayne Graves

Director, Office of Financial Services

Darren Ramdin

Director, Office of SIG Services

Donna Cappel

Director, Office of Publications

Scott E. Delman

ACM COUNCIL

President

Cherri M. Pancake

Vice-President

Elizabeth Churchill

Secretary/Treasurer

Yannis Ioannidis

Past President

Alexander L. Wolf

Chair, SGB Board

Jeff Jortner

Co-Chairs, Publications Board

Jack Davidson and Joseph Konstan

Members-at-Large

Gabriele Kotsis; Susan Dumais; Renée McCauley; Claudia Bauzer Medeiros; Elizabeth D. Mynatt; Pamela Samuelson; Theo Schlossnagle; Eugene H. Spafford
SGB Council Representatives
 Sarita Adve and Jeanna Neefe Matthews

BOARD CHAIRS

Education Board

Mehran Sahami and Jane Chu Prey

Practitioners Board

Terry Coatta

REGIONAL COUNCIL CHAIRS

ACM Europe Council

Chris Hankin

ACM India Council

Abhiram Ranade

ACM China Council

Wenguang Chen

PUBLICATIONS BOARD

Co-Chairs

Jack Davidson and Joseph Konstan

Board Members

Phoebe Ayers; Nicole Forsgren; Chris Hankin; Mike Heroux; Nenad Medvidovic; Tulika Mitra; Michael L. Nelson; Sharon Oviatt; Eugene H. Spafford; Stephen N. Spencer; Divesh Srivastava; Robert Walker; Julie R. Williamson

ACM U.S. Technology Policy Office

Adam Eisgrau

Director of Global Policy and Public Affairs
 1701 Pennsylvania Ave NW, Suite 200,
 Washington, DC 20006 USA
 T (202) 580-6555; acmpo@acm.org

Computer Science Teachers Association

Jake Baskin

Executive Director

COMMUNICATIONS OF THE ACM

Trusted insights for computing's leading professionals.

Communications of the ACM is the leading monthly print and online magazine for the computing and information technology fields. *Communications* is recognized as the most trusted and knowledgeable source of industry information for today's computing professional. *Communications* brings its readership in-depth coverage of emerging areas of computer science, new trends in information technology, and practical applications. Industry leaders use *Communications* as a platform to present and debate various technology implications, public policies, engineering challenges, and market trends. The prestige and unmatched reputation that *Communications of the ACM* enjoys today is built upon a 50-year commitment to high-quality editorial content and a steadfast dedication to advancing the arts, sciences, and applications of information technology.

STAFF

DIRECTOR OF PUBLICATIONS

Scott E. Delman
 cacm-publisher@cacm.acm.org

Executive Editor

Diane Crawford

Managing Editor

Thomas E. Lambert

Senior Editor

Andrew Rosenbloom

Senior Editor/News

Lawrence M. Fisher

Web Editor

David Roman

Editorial Assistant

Danbi Yu

Art Director

Andrij Borys

Associate Art Director

Margaret Gray

Assistant Art Director

Mia Angelica Balaquiot

Production Manager

Bernadette Shade

Intellectual Property Rights Coordinator

Barbara Ryan

Advertising Sales Account Manager

Ilia Rodriguez

Columnists

David Anderson; Michael Cusumano;
 Peter J. Denning; Mark Guzdial;
 Thomas Haigh; Leah Hoffmann; Mari Sako;
 Pamela Samuelson; Marshall Van Alstyne

CONTACT POINTS

Copyright permission

permissions@hq.acm.org

Calendar items

calendar@cacm.acm.org

Change of address

acmhelp@acm.org

Letters to the Editor

letters@cacm.acm.org

WEBSITE

http://cacm.acm.org

WEB BOARD

Chair

James Landay

Board Members

Marti Hearst; Jason I. Hong;
 Jeff Johnson; Wendy E. MacKay

AUTHOR GUIDELINES

http://cacm.acm.org/about-communications/author-center

ACM ADVERTISING DEPARTMENT

1601 Broadway, 10th Floor
 New York, NY 10019-7434 USA
 T (212) 626-0686
 F (212) 869-0481

Advertising Sales Account Manager

Ilia Rodriguez
 ilia.rodriguez@hq.acm.org

Media Kit acmm mediasales@acm.org

Association for Computing Machinery (ACM)

1601 Broadway, 10th Floor
 New York, NY 10019-7434 USA
 T (212) 869-7440; F (212) 869-0481

EDITORIAL BOARD

EDITOR-IN-CHIEF

Andrew A. Chien
 eic@cacm.acm.org

Deputy to the Editor-in-Chief

Morgan Denlow
 cacm.deputy.to.eic@gmail.com

SENIOR EDITOR

Moshe Y. Vardi

NEWS

Co-Chairs

Marc Snir and Alain Chesnais

Board Members

Tom Conte; Monica Divitini; Mei Kobayashi;
 Rajeev Rastogi; François Sillion

VIEWPOINTS

Co-Chairs

Tim Finin; Susanne E. Hambrusch;
 John Leslie King; Paul Rosenbloom

Board Members

Terry Benzel; Michael L. Best; Judith Bishop;
 Lorrie Cranor; Boi Falting; James Grimmelmann;
 Mark Guzdial; Haym B. Hirsch;
 Richard Ladner; Carl Landwehr; Beng Chin Ooi;
 Francesca Rossi; Len Shustek; Loren Terveen;
 Marshall Van Alstyne; Jeannette Wing;
 Susan J. Winter

PRACTICE

Co-Chairs

Stephen Bourne and Theo Schlossnagle

Board Members

Eric Allman; Samy Bahra; Peter Bailis;
 Betsy Beyer; Terry Coatta; Stuart Feldman;
 Nicole Forsgren; Camille Fournier;
 Jessie Frazelle; Benjamin Fried; Tom Killalea;
 Tom Limoncelli; Kate Matsudaira;
 Marshall Kirk McKusick; Erik Meijer;
 George Neville-Neil; Jim Waldo;
 Meredith Whittaker

CONTRIBUTED ARTICLES

Co-Chairs

James Larus and Gail Murphy

Board Members

Robert Austin; Kim Bruce; Alan Bundy;
 Peter Buneman; Jeff Chase;
 Yannis Ioannidis; Gal A. Kaminka;
 Ben C. Lee; Igor Markov;
 Lionel M. Ni; Doina Precup;
 Shankar Sastry; m.c. schraefel; Ron Shamir;
 Hannes Werthner; Reinhard Wilhelm

RESEARCH HIGHLIGHTS

Co-Chairs

Azer Bestavros, Shriram Krishnamurthi,
 and Orna Kupferman

Board Members

Martin Abadi; Amr El Abbadi;
 Animashree Anandkumar; Sanjeev Arora;
 Michael Backes; Maria-Florina Balcan;
 David Brooks; Stuart K. Card; Jon Crowcroft;
 Alexei Efros; Bryan Ford; Alon Halevy;
 Gernot Heiser; Takeo Igarashi;
 Srinivasan Keshav; Sven Koenig;
 Ran Libeskind-Hadas; Karen Liu; Greg Morrisett;
 Tim Roughgarden; Guy Steele, Jr.;
 Robert Williamson; Margaret H. Wright;
 Nikolai Zeldovich; Andreas Zeller

SPECIAL SECTIONS

Co-Chairs

Sriram Rajamani, Jakob Rehof,
 and Haibo Chen

Board Members

Tao Xie; Kenjiro Taura; David Padua

ACM Copyright Notice

Copyright © 2020 by Association for Computing Machinery, Inc. (ACM). Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or fee. Request permission to publish from permissions@hq.acm.org or fax (212) 869-0481.

For other copying of articles that carry a code at the bottom of the first or last page or screen display, copying is permitted provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center; www.copyright.com.

Subscriptions

An annual subscription cost is included in ACM member dues of \$99 (\$40 of which is allocated to a subscription to *Communications*); for students, cost is included in \$42 dues (\$20 of which is allocated to a *Communications* subscription). A nonmember annual subscription is \$269.

ACM Media Advertising Policy

Communications of the ACM and other ACM Media publications accept advertising in both print and electronic formats. All advertising in ACM Media publications is at the discretion of ACM and is intended to provide financial support for the various activities and services for ACM members. Current advertising rates can be found by visiting <http://www.acm-media.org> or by contacting ACM Media Sales at (212) 626-0686.

Single Copies

Single copies of *Communications of the ACM* are available for purchase. Please contact acmhelp@acm.org.

COMMUNICATIONS OF THE ACM

(ISSN 0001-0782) is published monthly by ACM Media, 1601 Broadway, 10th Floor New York, NY 10019-7434 USA. Periodicals postage paid at New York, NY 10001, and other mailing offices.

POSTMASTER

Please send address changes to *Communications of the ACM* 1601 Broadway, 10th Floor New York, NY 10019-7434 USA

Printed in the USA.



Association for Computing Machinery





Cherri M. Pancake

DOI:10.1145/3382126

New Ways to Think About CS Education

RECENT ARTICLES IN *Communications* have touched on our professional responsibilities to society, including the importance of attracting more and diverse people to computing, integrating ethics into the technology development cycle, and mitigating the effect of our conferences on the environment. Two new approaches could have impact on both the quality of CS education and those societal concerns.

Reinforcing What Students Learn about Ethics

For decades, accreditation agencies have required that CS students be exposed to the ethical responsibilities of a professional, typically through a course dedicated to societal aspects of computing. But as technology has taken over more aspects of daily life, concern has grown about whether CS graduates really learn enough about ethics and accountability. Harvard University has come up with a unique solution, called “Embedded EthICS.” It combines two tactics: interspersing ethical discussions throughout the curriculum, and engaging the help of philosophy faculty to co-teach relevant material. The evolution of the initiative was described in the August 2019 issue (p. 54).

Taking an integrative approach reinforces the importance of ethics in creating and applying technology. Teaching ethics at just one point in the curriculum doesn’t really convey the extent to which we, as technologists, need to be alert to the ethical aspects of everything we do. The notion of pairing ethical scenarios with different topics should have been obvious long ago; the accompanying table shows examples from common courses in the CS degree curriculum.

Partnering with philosophy experts adds

Courses typically included in the CS curriculum.

| Course | Types of Ethical Challenges Addressed |
|-------------------------|---|
| Networks | Social media, fake news, and the ethics of censorship |
| Programming Languages | Ethics in software verification and validation |
| Data Systems | Privacy in the design of data systems |
| Usability | Inclusive design and equality of opportunity |
| Artificial Intelligence | Machines and moral decision-making |

another dimension. In computing, we focus on finding the best algorithm to solve a problem, but do not always analyze the impact of solving it in this way. Learning more about the framework of philosophical thinking gives students tools to recognize when they are actually making ethical as well as technical choices—and that failure to consider those consequences may lead to unwise choices. (Learn more at embeddedethics.seas.harvard.edu)

Creating New Pathways into CS

The shortage of computer scientists is a worldwide problem, but it has proven difficult to create new pathways into computing. Oregon State University tackled the problem by targeting a largely untapped resource: people with degrees from other disciplines who want to move into CS. It allows individuals without any computing background to complete a second degree in CS, on an accelerated schedule (similar to medical schools offering accelerated MD degrees to researchers with a Ph.D. in science) and from their own location.

The goal was attracting new audiences to CS, so a key component is that although it is an accredited BS program, no prior knowledge is assumed. Students come from fields as diverse as music, forestry, and healthcare, with about 25% already holding advanced degrees (including Ph.D. and MD). Almost half report no prior experience with either program-

ming or online studies. Students can pace their coursework if they need to balance study with work or caregiving responsibilities.

In order to retain students, it was necessary to create a virtual environment that mirrors the on-campus experience. Software was developed to mimic the “active learning” format used in campus classrooms, including small-group instruction and interactive sessions with instructors and teaching assistants. Students have the same advising, tutoring, and career services support they would receive on campus, and a new near-peer mentoring program pairs entering students with experienced ones. An active online community has developed where past and current students interact with peers and potential students. (Learn more at eecs.oregonstate.edu/academic/online-cs-postbacc.)

If adopted widely, both approaches have ancillary benefits that could prove transformative as well. The Harvard initiative keeps ethical issues in the forefront of teachers’ minds, too. The Oregon State initiative helps increase the diversity of the computing workforce, and raises awareness of how remote participation can mimic “being there” without the environmental impact of travel or commuting. As recent *Communications* authors have pointed out, all three of those cultural changes are important to the future of our profession. **□**

ACM Transactions on Social Computing (TSC)

Open for
Submissions

Seeks to publish work that covers the full spectrum of social computing including theoretical, empirical, systems, and design research contributions



ACM Transactions on Social Computing (TSC) seeks to publish work that covers the full spectrum of social computing including theoretical, empirical, systems, and design research contributions. The editorial perspective is that social computing is fundamentally about computing systems and techniques in which users interact, directly or indirectly, with what they believe to be other users or other users' contributions. TSC welcomes research employing a wide range of methods to advance the tools, techniques, understanding, and practice of social computing, including: theoretical, algorithmic, empirical, experimental, qualitative, quantitative, ethnographic, design, and engineering research. Social computing will continue to be shaped by foundational algorithmic, econometric, psychological, sociological, and social science research and these broad-based perspectives will continue to have a profound influence on how social computing systems are designed, built and how they grow.

For more information and to submit your work, please visit:

TSC particularly solicits research that designs, implements or studies systems that mediate social interactions among users, or that develops or studies theory or techniques for application in those systems. Examples of such social computing systems include, but are not limited to: instant messaging, blogs, wikis, social networks, social tagging, social recommenders, collaborative editors and shared repositories.

tsc.acm.org



Association for
Computing Machinery



Vinton G. Cerf

DOI:10.1145/3383671

Multisensory Adventures

IN THIS COLUMN, I want to draw your attention to two books. One has been published to great acclaim and the other is still in process. They resonate with a visceral intensity for which I was honestly unprepared and surprised. The first, *Multisensory Experiences, Where the Senses Meet Technology*, by Carlos Velasco and Marianna Obrist, is to be published by Oxford University Press. The authors explore concepts we experience every day but don't necessarily understand fully. We are familiar with the five senses (sight, sound, touch, taste, smell). Our brains transduce these physical phenomena into neural pulses that flood along many pathways and interact in many ways. Interestingly, all these senses are translated into essentially similar neural signals but they are processed in a complex and interconnected neural web producing what we call *experience*.

Sensory memories are powerful. When I smell cigar smoke, I am transported back to my early childhood with my grandfather and mother during World War II. The olfactory memory recalls sights, sounds, scenes, and other sensory phenomena. Our senses are linked, not only in real time but also in recall.

We are learning more about our senses as our ability to measure phenomena improves. Taste is more than sweet, sour, bitter, salty, umami, and metallic since it is also heavily combined by our brains with smell. Most of the taste of wine is olfactory, for example. If reincarnation is real, I hope I am reincarnated as a wine-tasting dog! We are also learning there are neural circuits in the brain that literally function as spatial maps allowing us to navigate to places we have been before. It is remarkable how quickly these maps can be constructed.

This book posits that it is possible to deliberately create multisensory experiences in the form of art and to explore how experience can be altered with subtle changes in sensory ambience. A scene can go from cheerful and uplifting to scary and threatening depending on the choices made for background sound. For me, this illustrates viscerally, how experience is influenced by multimedia inputs.


Understanding and appreciating multisensory experience can come in many forms. The famed chef José Andrés invented new restaurants where especially talented chefs produce tiny, bite-sized tapas that produce unexpected multisensory effects. One dish is made with liquid nitrogen (I kid you not!) and after you pop this into your mouth, your outgoing nasal breath makes you look like a dragon. Another looks like a little palm frond beach hut, but it is actually a deconstructed Caesar salad. Your eyes tell you one thing but your mouth and nose tell you otherwise. You will find many examples of these kinds of artificially generated multisensory experiences accounted for in the book. I think you will find it an eye-opening experience to read what

We are learning there are neural circuits in the brain that literally function as spatial maps allowing us to navigate to places we have been before.

these authors have to say about the conflation of our senses produced by high-dimensional synthetic effects. There is nothing simple about our senses and this book explains why.

The second book I call to your attention is the 2015 Pulitzer-prize winning *All the Light We Cannot See*, by Anthony Doerr. By good fortune, I read this book after reading *Multisensory Experiences*. Marie-Laure is blind and her father is teaching her to navigate the town they live in. It is difficult, but eventually success comes. Now experience the multisensory ensemble she uses to navigate home:

"... one snowy Tuesday in March she squats on her heels on the sidewalk. The faintly metallic smell of the falling snow surrounds her. Listen. Cars splash along streets, and snowmelt drums through runnels; she can hear snowflakes tick and patter through the trees. She can smell the cedars in the Jardin des Plantes a quarter mile away. Here the Metro hurtles beneath the sidewalk: that's the Quai Saint-Bernard. Here the sky opens up, and she hears the clacking of branches: that's the narrow stripe of gardens behind the Gallery of Paleontology... They walk up their street now, she is sure of it. They are outside their building. Marie-Laure finds the trunk of the chestnut tree that grows past her fourth-floor window, its bark beneath her fingers. Old friend..."¹

I strongly recommend both books. May your senses combine in collaborative celebration! 

Reference

1. Doerr, A. *All the Light We Cannot See*. Scribner, 2014. Kindle Ed., 40–41.

Vinton G. Cerf is vice president and Chief Internet Evangelist at Google. He served as ACM president from 2012–2014.

Copyright held by author/owner.

ACM Welcomes the Colleges and Universities Participating in ACM's Academic Department Membership Program

ACM offers an Academic Department Membership option, which allows universities and colleges to provide ACM Professional Membership to their faculty at a greatly reduced collective cost.

The following institutions currently participate in ACM's Academic Department Membership program:

- Abilene Christian University
- Afrisol Technical College, Zimbabwe
- Alfred State College
- Amherst College
- Appalachian State University
- Augusta University, School of Computer and Cyber Sciences
- Ball State University
- Bellevue College
- Berea College
- Binghamton University
- Boise State University
- Bridgewater State University
- Bryant University
- California Baptist University
- Calvin College
- Clark University
- Colgate University
- Colorado School of Mines
- Columbus State University
- Cornell University
- Creighton University
- Cuyahoga Community College
- Denison University
- European University (Tbilisi, Georgia)
- Franklin University
- Gallaudet University
- Georgia Institute of Technology
- Georgia State University Perimeter College
- Governors State University
- Harding University
- Harvard University
- Harvey Mudd College
- Hochschule für Technik Stuttgart - University of Applied Sciences
- Hofstra University
- Hope College
- Howard Payne University
- Indiana University Bloomington
- Kent State University
- Klagenfurt University, Austria
- Madinah College of Technology, Saudi Arabia
- Massasoit Community College
- Messiah College
- Metropolitan State University
- Missouri State University
- Modesto Junior College
- Monash University, Australia
- Montclair State University
- Mount Holyoke College
- New Jersey Institute of Technology
- New Mexico State University
- Northeastern University
- Ohio State University
- Old Dominion University
- Pacific Lutheran University
- Pennsylvania State University
- Potomac State College of West Virginia University
- Purdue University Northwest
- Regis University
- Rhodes College
- Rochester Institute of Technology
- Rutgers University
- Saint Louis University
- San José State University
- Shippensburg University
- Simmons University
- Spelman College
- St. John's University
- Stanford University
- State University of New York at Fredonia
- State University of New York at Oswego
- Stetson University
- Trine University
- Trinity University
- Union College
- Union University
- Univ. do Porto, Faculdade de Eng. (FEUP)
- University at Albany, State University of New York
- University of Alabama
- University of Arizona
- University of California, Riverside
- University of California, San Diego
- University of Colorado Boulder
- University of Colorado Denver
- University of Connecticut
- University of Houston
- University of Illinois at Chicago
- University of Jamestown
- University of Liechtenstein
- University of Lynchburg
- University of Maribor, Slovenia
- University of Maryland, Baltimore County
- University of Memphis
- University of Namibia
- University of Nebraska at Kearney
- University of Nebraska Omaha
- University of New Mexico
- University of North Dakota
- University of Pittsburgh
- University of Puget Sound
- University of Southern California
- University of St. Thomas
- University of the Fraser Valley
- University of Victoria, BC Canada
- University of Wisconsin–Parkside
- University of Wyoming
- Virginia Commonwealth University
- Wake Forest University
- Wayne State University
- Wellesley College
- Western New England University
- William Jessup University

Through this program, each faculty member receives all the benefits of individual professional membership, including *Communications of the ACM*, member rates to attend ACM Special Interest Group conferences, member subscription rates to ACM journals, and much more.

DOI:10.1145/3383390

Adding More Color to Patch Picture

I READ “AUTOMATED PROGRAM REPAIR” with interest (Dec. 2019, p. 56–65). This is exciting technology that, if successful, holds out the promise of substantially improving software quality. While the article highlights systems developed by the first and third authors (GenProg, SemFix, Angelix), it omits quantitative data that can provide a more complete picture of the capabilities of extant program repair systems. My hope is this quantitative data can help researchers and practitioners better understand the capabilities and current limitations of this promising technology.

The most complete evaluation of the GenProg system was reported in Le Goues et al.,^{1,2} which examines results for a superset of the defects originally considered in Le Goues et al.³ Unfortunately, as reported in Qi et al.⁷ and communicated to the authors of Le Goues³ in fall of 2014, the experimental setup contains a variety of test harness and test script issues. When these issues are corrected, the results show that GenProg does not fix 55 of 105 bugs, as one might reasonably expect from reading the title of the article. Instead, GenProg fixes only two bugs, highlighting the remarkable ineffectiveness of GenProg as an automatic patch generation system. Moreover, only 69 of the reported 105 bugs are bugs—the remaining 36 are deliberate functionality changes.

I note this ineffectiveness may not be widely recognized—despite being informed of these results in fall of 2014, and despite the publication of Qi,⁷ at press time, websites maintained by the authors of GenProg still do not reflect the corrections required to accurately represent the capabilities of the GenProg system (for example, see <https://squareslab.github.io/genprog-code/>).

For comparison, the Prophet system,⁶ the current state of the art on this benchmark set, generates correct patches for 18 of the 69 defects. But for another 21 defects, Prophet generates incorrect patches that nevertheless validate. This situation requires developers to manually filter the vali-

dated patches, with developer evaluation effort and false positives an important concern.

These quantitative results can provide insight into why current commercial automatic patch generation systems such as those discussed in the article focus on specific defect classes such as null dereference defects. Focusing on these classes enables the development of more narrowly tailored techniques that can aspire to fix a larger proportion of the defects with fewer false positives.^{4,5}

In the near term, I think we can expect patch generation systems that focus on specific defect classes to play an increasingly prominent role in maintaining large software systems. Because of the substantial redundancy present in and across most large software systems, as well as the availability of multiple sources of information such as revision histories present in software repositories, I would expect efforts directed at broader classes of defects to pay off in the future. Of course, accurate reporting of relevant results can play an important role in helping the field progress.

Martin Rinard,
Cambridge, MA, USA

References

1. Le Goues, C. et al. The ManyBugs and IntroClass benchmarks for automated repair of C programs. *IEEE Trans Software Engineering* 41, 12 (Dec. 2015), 1236–1256.
2. Le Goues, C., Brun, Y., Forrest, S. and Weimer, W. Clarifications on the construction and use of the ManyBugs benchmark. *IEEE Trans. Software Engineering* 43, 11 (Nov. 2017), 1089–1090.
3. Le Goues, C., Dewey-Vogt, M., Forrest, S. and Weimer, W. A systematic study of automated program repair: Fixing 55 out of 105 bugs for \$8 each. In *Proceedings of the 34th Intern. Conf. Software Engineering* (Zurich, Switzerland, June 2–9, 2012), 3–13.
4. Long, F. Automatic patch generation via learning from successful human patches. Ph.D. thesis, MIT, Cambridge, USA, 2018.
5. Long, F., Amidon, P. and Rinard, M. Automatic inference of code transforms for patch generation. In *Proceedings of the 11th Joint Meeting on Foundations of Software Engineering* (Paderborn, Germany, Sept. 4–8, 2017), 727–739.
6. Long, F. and Rinard, M. Automatic patch generation by learning correct code. In *Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symp. Principles of Programming Languages* (St. Petersburg, FL, USA, Jan. 20–22, 2016), 298–312.
7. Qi, Z., Long, F., Achour, S. and Rinard, M.C. An analysis of patch plausibility and correctness for generate-and-validate patch generation systems. In *Proceedings of the Intern. Symp. Software Testing and Analysis* (Baltimore, MD, USA, July 12–17, 2015), 24–36.

CS + CS

I read “When Human-Computer Interaction Meets Community Citizen Science” (Feb. 2020, p. 31–34) with interest given my own, multidisciplinary exploration of similar territory. The authors do a nice job of describing the increasingly wide range of citizen science activities. Not only do many leading the expansion of citizen science refer to it as CS, a challenge for those of us who use that term for computer science, but that recent expansion has been occasioned by the launch and growth of online platforms, laying a foundation for the intersection of the two kinds of CS, as is implicit in the article.

I led a small team at RAND that has published two small reports on community citizen science. *The Promise of Community Citizen Science*² came out in 2017; *Community Citizen Science: From Promise to Action*¹ came out in 2019. So, while we would like to think we were the ones to introduce the concept, we applaud the work of Yen-Chia Hsu and Illah Nourbakhsh and hope that we can find a way to collaborate.

Marjory S. Blumenthal,
Washington, D.C., USA

References

1. Chari, R. et al. *Community Citizen Science: From Promise to Action* (2019); https://www.rand.org/pubs/research_reports/RR2763.html
2. Chari, R. et al. *The Promise of Community Citizen Science* (2017); <https://www.rand.org/pubs/perspectives/PE256.html>

Editor-in-Chief response

It's great to see excitement and energy in this important area!

Andrew A. Chien, Chicago, IL, USA

The *Communications* Web site, <http://cacm.acm.org>, features more than a dozen bloggers in the BLOG@CACM community. In each issue of *Communications*, we'll publish selected posts or excerpts.

twitter

Follow us on Twitter at <http://twitter.com/blogCACM>

DOI:10.1145/3381906

<http://cacm.acm.org/blogs/blog-cacm>

Teaching Teachers to Offer STEM to Undergrads

Mark Guzdial considers how teaching computing to undergraduates requires better education of the teachers as well.



**Mark Guzdial
Inventing Computing
Education to Meet All
(Especially Teachers)
Undergraduates' Needs:
CUE.NEXT Workshops**

<http://bit.ly/2RVYGk6>

January 3, 2020

In December, I had the honor of providing a keynote talk at the CUE.NEXT Workshop in Washington, D.C. The goal of the CUE.NEXT workshop series is to envision the future of computing in undergraduate education. At the website (<https://cue.northwestern.edu/>), organizers Larry Birnbaum, Susanne Hambrusch, and Clayton Lewis describe the purpose of the workshop:

Computing and computer science have become relevant to undergraduate education in all disciplines. Academic institutions are challenged to meet the demand of the growing and increasingly diverse student body seeking to learn more about computing, computer science, and the role of computation in their own disciplines. Courses and curricula aimed at teaching the fundamental technologies of computer

science to CS majors often do not meet the needs of this wider student audience ... The goal of the CUE.NEXT workshops is to initiate a national dialog on the role of computing in undergraduate education. Computing educators and CS departments, as well as colleagues and academic units representing other stakeholder disciplines, will work together to understand and address the challenges.

U.S. academic institutions participating sent a team of CS and other-than-CS faculty to develop approaches for providing CS education to other-than-CS undergraduates. I was there for the day, participating in two breakout sessions and giving my keynote. The first breakout session was amazing; I learned new perspectives about the challenges of growing CS education in K–12 in the U.S.

Educating K–12 Teachers

The first breakout sessions were organized by discipline. All three CUE.NEXT workshops included faculty interested in STEM, engineering, humanities, arts and media, economics, and social sciences. I went to a breakout session on

Education. Around the table were more than a dozen education faculty interested in providing new or developing teachers (pre-service) the computing education they would need.

Because of my work with the ECEP Alliance (<https://ecepalliance.org/>), I know education faculty pushing for computing in teacher development, and I'm reasonably familiar with teacher policies in a handful of states. The education faculty around the table at the CUE.NEXT workshop were from states I wasn't familiar with. They were faculty in math and science education dealing with new demands for computing education. This was a different sample of education faculty than I had worked with before.

I'll structure my report in terms of quotes I have in my notes.

"Lots of education faculty think CS is going to go away."

I was surprised by that. I think of computing education as a new literacy (as described on the Computing Education Research Blog, <http://bit.ly/2v1Ug1M>), and while we're currently pretty small (see data in my earlier blog, <http://bit.ly/2vHAGs1>), we're growing. CS education for all is inevitable, or so I thought. There have been lots of education fads over the years (as observed in <http://bit.ly/2Scy90N>); many education faculty think teaching programming is just the latest.

"CS isn't real or relevant, vs. hammer to nail."

There are several comparisons between computer science and shop classes. Shop classes include things like woodworking and automotive re-

pair. Those are real, and deal with concrete things. Computer science doesn't seem nearly as real or relevant to the teachers in their programs.

"I'm teaching elementary school, so I don't need programming."

The education faculty around my table saw that as a problem, that teachers focused on elementary school saw programming as scary and not really necessary or helpful. They talked about using cute robots as a way of drawing in elementary school teachers.

"Science education is so packed. We can't fit programming in."

Most requirements in teacher professional development programs in the states represented around the table are fixed by state law or regulation. The schools have few options. Particularly in programs in science and mathematics, there are *no* electives; all the classes are required. There is simply no place for another course to teach computer science. If we want science and mathematics teachers (among others) to learn about computing, we have to integrate it into existing classes. We heard about faculty fitting programming into science or mathematics methods ("how to teach X") courses, or into educational technology courses.

"Less trying to teach programming to teach programming, and more programming to serve the domain."

Nobody around the table with me was trying to teach future computer science teachers. They had no resources to do that. They were teaching future teachers to use programming to serve the domains they're teaching.

I asked the science ed faculty around the room what integration approaches they were using: Bootstrap Physics (<https://www.bootstrapworld.org/materials/physics/>), or Project GUTS for Middle School Science (<https://www.project-guts.org/>), or CT-STEM with NetLogo (<https://ct-stem.northwestern.edu/>), or something else? There was a long pause, then one professor spoke up:

"I don't know most of the words you just said."

Projects grown out of computing education are mostly unknown to math and science educators. The education faculty at the table with me were "growing their own." They were working with Scratch or App Inventor and trying to map from science standards in their state to computing activities.

There was a long discussion about barriers to entry for teachers. Math classes are one barrier; a lot of potential future teachers never make it past math classes. When math content is taught by education faculty, future teachers do better, but that creates a huge load on education faculty. Burnout is a significant problem. The education faculty are worried computer science might become the next barrier to teachers.

If we really want pre-service teachers to learn about computer science, these are the lessons that we have to hear. We have to learn about their problems and help to address them.

We Don't Need CS Faculty to Teach CS

My keynote slides are at <http://bit.ly/31G9EgR>.

I started with a review of the state of high school CS education today, to make the point few U.S. high school students ever take a CS course. If you want your undergraduates to learn CS, you have to put it in your undergraduate curriculum.

I told three stories: About how we designed the Media Computation course for liberal arts majors at Georgia Tech, about the need for Computational Literacy, and about how CS pedagogy needs to change to reach other-than-CS majors. (The last part is available as a video at <http://bit.ly/395y59T>.)

I ended with a claim that undergraduate teachers of computer science do not have to be computer science faculty. The most important knowledge a CS teacher needs is about the practices of computing within that discipline. CS faculty is likely know about software development practices; they don't necessarily know how computing is used in science, engineering, mathematics, liberal arts, or social sciences. The next most important knowledge is called *pedagogical content knowledge*: how to teach CS well. CS faculty tend not be well-informed about how to teach CS well (as described in the blog post at <http://bit.ly/2UjSbca>). It's not clear there's an advantage in having CS faculty teaching non-CS majors — and we don't have the capacity in U.S. universities. We should share the load.

Mark Guzdial is a professor in the Computer Science & Engineering Division, and in the Engineering Education Research program, of the University of Michigan.

© 2020 ACM 0001-0782/20/4 \$15.00

Coming Next Month in COMMUNICATIONS

A Snapshot of the Frontiers of Fairness in Machine Learning

The Bibliometric Approach for Detecting the Gender Gap in Computer Science

Indistinguishability

Reading in the Panopticon

Beyond the 'Fix-It' Treadmill

Revealing the Critical Role of Human Performance in Software

When Technology Goes Awry

Computers Do Not Make Art, People Do

Measuring and Mitigating OAuth Access

Token Abuse by Collusion Networks

Plus the latest news about solving for sensitivity, chipping with RISC-VS, and information gerrymandering.

The Pragmatic Wisdom of Michael Stonebraker

Making Databases Work

This book celebrates Michael Stonebraker's accomplishments that led to his 2014 ACM A.M. Turing Award "for fundamental contributions to the concepts and practices underlying modern database systems."

The book describes, for the broad computing community, the unique nature, significance, and impact of Mike's achievements in advancing modern database systems over more than forty years. Today, data is considered the world's most valuable resource, whether it is in the tens of millions of databases used to manage the world's businesses and governments, in the billions of databases in our smartphones and watches, or residing elsewhere, as yet unmanaged, awaiting the elusive next generation of database systems. Every one of the millions or billions of databases includes features that are celebrated by the 2014 Turing Award and are described in this book.

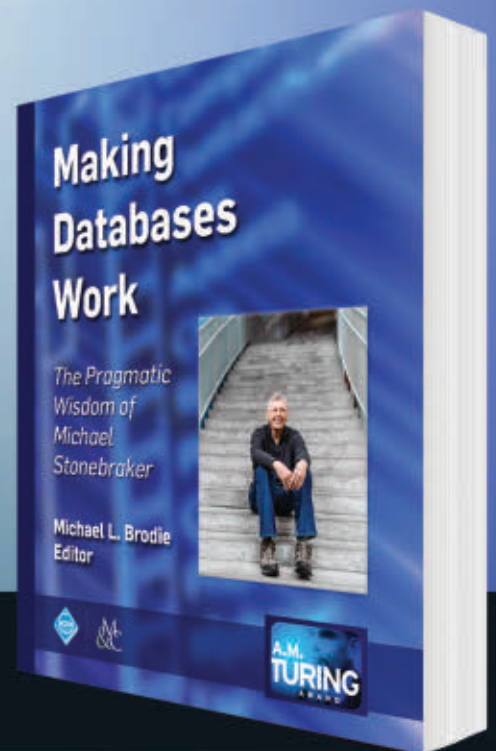
Edited by Michael L. Brodie

ISBN: 978-1-94748-719-2

DOI: 10.1145/3226595

<http://books.acm.org>

<http://www.morganclaypoolpublishers.com/acm>



ACM BOOKS

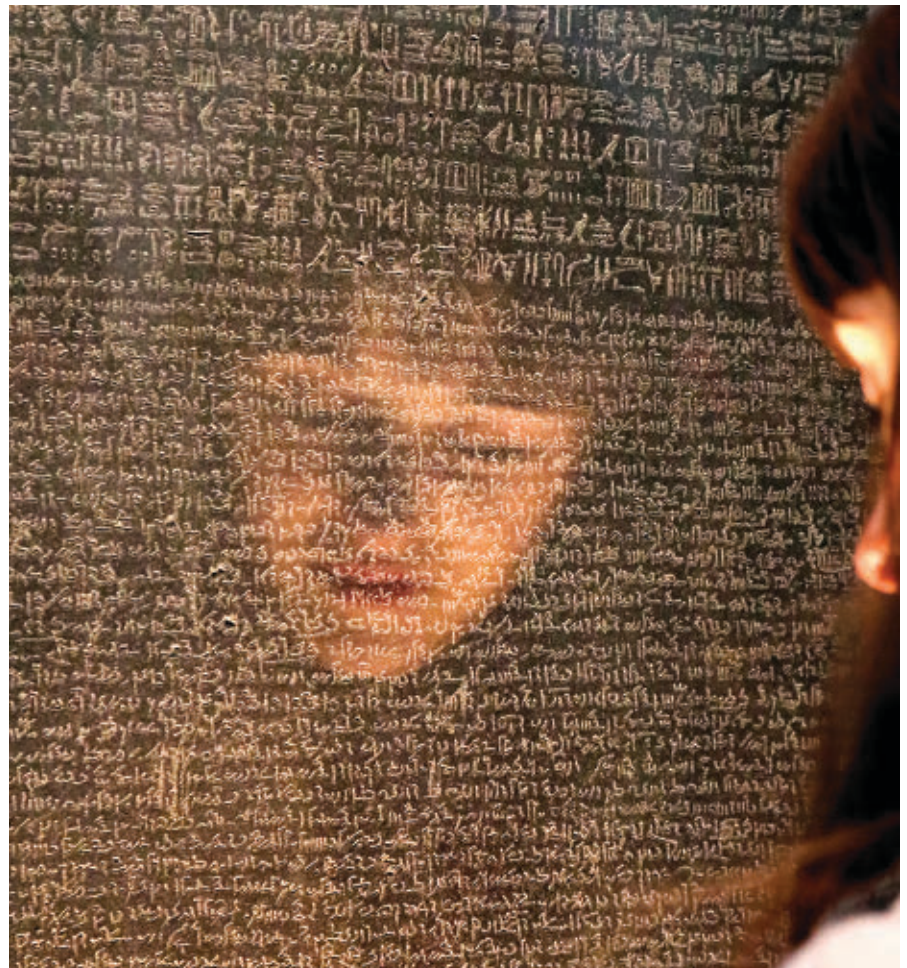
Dead Languages Come to Life

Artificial intelligence automates the translation of extinct languages.

DRIVEN BY ADVANCED techniques in machine learning, commercial systems for automated language translation now nearly match the performance of human linguists, and far more efficiently. Google Translate supports 105 languages, from Afrikaans to Zulu, and in addition to printed text it can translate speech, handwriting, and the text found on websites and in images.

The methods for doing those things are clever, but the key enabler lies in the huge annotated databases of writings in the various language pairs. A translation from French to English succeeds because the algorithms were trained on millions of actual translation examples. The expectation is that every word or phrase that comes into the system, with its associated rules and patterns of language structure, will have been seen and translated before.

Now researchers have developed a method that, in some cases, can automatically translate extinct languages, those for which these big parallel data sets do not exist. Jiaming Luo and Regina Barzilay at the Massachusetts Institute of Technology (MIT) and Yuan Cao at Google were able to automate the “decipherment” of Linear B—a Greek



By the Numbers

There are 7,000 known, distinct languages in the world

- ▶ 2,500 are endangered (children no longer learn to speak them)
- ▶ 230 have become extinct since 1950 (there are no speakers left)

Source: UNESCO

language predecessor dating to 1450 B.C.—into modern Greek. Previous translations of Linear B to Greek were only possible manually, at great effort, by language and subject-matter experts. The same automated methods were also able to translate Ugaritic, an extinct Semitic language, into Hebrew.

How It Works

In a recent paper, *Neural Decipherment via Minimum-Cost Flow*, the three computer scientist authors describe a two-step process. The first part, operating at the character level, uses a conventional neural network to predict the correct character in each word in the decipherment, based on prior knowledge of the patterns that tend to match across the two languages. The second step uses a linear program to minimize deviations of the derived vocabulary from the previously manually translated vocabulary. The two steps iterate back and forth and together attempt to translate words that are “cognates”—words with the same derivation.

“These two parts are complementary to each other,” says Luo, a Ph.D. student at MIT. “The linear program provides a global perspective that looks at the entire derived vocabulary, and it can utilize the information that is not readily available for a conventional local neural net. On the other hand, neural nets are very good at extracting local, character-level patterns that are harder to formulate using a linear program.”

The lack of parallel data for training, and the scarcity of ancient texts, make decipherment the “ultimate low-resource challenge for both humans and machines,” the researchers say in their paper. A typical manual decipherment “spans decades and requires encyclopedic domain knowledge, prohibitive manual effort, and sheer luck.”

Taylor Berg-Kirkpatrick, an assistant professor in the department of computer science and engineering at the University of California, San Diego, was not involved in this work, but has done similar re-

search in unsupervised translation by machine learning. As he explains, the three scientists writing the paper “explicitly pose the problem of cognate matching as a combinatorial optimization problem, and conceptually divide the problem into alphabet- and cognate-matching. What their paper does that’s quite new is make use of high-capacity neural nets within this framework. It works quite well.”

The three scientists employ “a smart way of biasing the model to be simple, while letting it still be flexible,” Berg-Kirkpatrick says. “It doesn’t reorder characters; it just looks at the input left-to-right and changes the characters to the new alphabet, possibly deleting or inserting a character here and there. It’s a sort of fuzzy alignment, not a concrete alignment.”

Automated translation without extensive training on parallel datasets generally requires developers to know in advance how two languages are related, with similar alphabets, structures, and patterns. These patterns are predictable and tend to repeat, and they can be matched across languages that are from the same family. For example, English verbs often assume the past tense with “-ed” added, while German verbs do that with the “ge-” prefix. “The stronger the prior knowledge—the inductive bias—that you put into the algorithm, the less

A lack of parallel training data and scarce ancient texts make decipherment the “ultimate low-resource challenge for humans and machines.”

data you will need,” says Cao, a research engineer at Google.

The researchers knew from earlier manual translations that Linear B and Greek have many cognates. While these cognates have the same origin, they have changed in slightly different ways over the years. The neural network-linear program system could use this knowledge to iterate through successively accurate decipherments, says Regina Barzilay, a professor of computer science and artificial intelligence at MIT. The method could translate between Linear B and Greek, but not directly from Linear B to French because those languages are too dissimilar, she says.

The algorithms were able to decipher Linear B with 67% success, meaning two-thirds of the cognate pairs were translated correctly. The other third was translated incorrectly based on the earlier, manually created dictionary, while non-cognates were not considered. “We need to find another way to solve that piece; words that don’t have a common origin or that are not in the manually created database,” says MIT’s Luo.

Luo says commercial systems like Google Translate work at the semantic level, converting entire sentences from one language to another in a way that tries to preserve their meaning. However, “decipherment” does only character-mapping and word-matching of cognates and does not get at the overall meaning of a block of text.

Application to Other Languages

In a sense, the translations of Linear B and Ugaritic by neural network were only proofs of concept—albeit important ones—as those languages were already translated. The next step—which computer scientists say will be much more difficult but not impossible—will be to try the ideas on so-far-undeciphered languages.

Barzilay says earlier manual attempts to decipher Linear B failed because researchers didn’t think Linear B was related to Greek. They struggled unsuccessfully for years to translate it into other languages, not succeeding until 1953, when the British architect and linguist Michael Ventris tried Greek. “So, finding the right related language is really crucial,” Barzilay says.

The three colleagues are now working to extend their methods to other extinct

Neural decipherment via minimum-cost flow may have application beyond even language translation, in areas like DNA sequence-alignment.

languages; for example, Iberian. However, Iberian, which was used by the indigenous people of the Iberian Peninsula (present-day home of Spain and Portugal) more than two millennia ago, has never been deciphered by any means, and presents several difficulties. One is that some of the existing texts lie in large monolithic blocks of characters, making it difficult to identify discrete words.

Even more problematic is that there is no agreement as to what other known language(s) might share a common origin with Iberian. There is debate, indeed controversy, among archaeological linguists as to whether the ancient Iberian language is related to Basque, to Aquitanian (a precursor of the Basque language) or to some other extinct language. “We believe that by automating the available data, we can shed light objectively on the subject to bring some understanding to what happened, and to European history,” Luo says.

In the meantime, there is more work that could be done with Linear B and Ugaritic, Cao says. The methods described in the recent paper only model the “surface form” (individual letters, not whole words at once) of the input text so as to find cognates. “What we didn’t do is consider the semantics of words, the context of words,” he says. “Similar words tend to occur in similar contexts, so that’s a big piece we have to add to the algorithm. And what about non-cognates, or phrases? These are much harder, but not impossible. We are working on things like that.”

Does society really need to know more about languages that have not been spoken for centuries? “You might as well ask what’s the point of doing research on archaeology,” Cao says. “This

is sort of archaeology for languages.” At a minimum, these advanced machine learning and artificial intelligence techniques will be a big help to scholars who previously lavished huge efforts on manual translations, he says.

“Their paper serves as a new demonstration that constraint-based methods and neural methods can work together to solve unsupervised problems,” Berg-Kirkpatrick says. “I hope this will serve to reinterest the NLP [natural language processing] community in historical decipherment problems.”

Neural decipherment via minimum-cost flow potentially may have application even beyond language translation. For example, the concepts might be applied to DNA sequence-alignment, in which biologists try to find small matching segments of DNA, called “motifs,” on the same or similar strands of DNA. That problem resides in a broad class of applications called “correspondence induction.”

Further Reading

Asgari, E. and Schütze, H. Past, Present, Future: A Computational Investigation of the Typology of Tense in 1000 Languages, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, April 2017 <https://www.aclweb.org/anthology/D17-1011/>

Berg-Kirkpatrick, T. and Klein, D. Simple Effective Decipherment via combinatorial Optimization, *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, July 2011 <https://www.aclweb.org/anthology/D11-1029/>

Duh, K. Bayesian Analysis in Natural Language Processing, *Computational Linguistics*, Vol. 44, Issue 1, March 2018, p.187-189 <http://bit.ly/2Vwn2Rn>

Luo, J. Cao, Y. and Barzilay, R. Neural Decipherment via Minimum-Cost Flow: from Ungaritic to Linear B, eprint [arXiv:1906.06718](https://arxiv.org/pdf/1906.06718), June 2019, <https://arxiv.org/pdf/1906.06718.pdf>

Robinson, A. Lost Languages: The Enigma of the World’s Undeciphered Scripts, Thames & Hudson, reprint edition, <https://amzn.to/2B1LAZc>

Snyder, B. and Barzilay, R. Unsupervised Multilingual Learning for Morphological Segmentation, *Proceedings of ACL-08: HLT*, Assoc. for Computational Linguistics, June 2008, <https://www.aclweb.org/anthology/P08-1084/>

Gary Anthes is a technology writer and editor based in Arlington, VA, USA.

© 2020 ACM 0001-0782/20/4 \$15.00

ACM Member News

AT THE INTERSECTION OF COMPUTATIONAL GEOMETRY AND ROBOTICS



“In elementary school, I loved geometry. In high school, I fell in love with computers,” recalls Dan

Halperin, a professor in the School of Computer Science of Israel’s Tel Aviv University. “In college, I discovered computational geometry, which brings the two together.”

Halperin earned undergraduate degrees in math and computer science, and his master’s and Ph.D. degrees in computer science, all from Tel Aviv University.

After receiving his doctorate, Halperin spent three years in the computer science robotics laboratory of Stanford University in California, before returning to Israel to join the faculty of Tel Aviv University.

One research emphasis during his career has been robust geometric computing, which concerns how geometric algorithms may be successfully implemented. Now, however, Halperin works at the intersection of robust computational geometry and robotics.

“I focus on motion and separability in tight settings, and this is manifested in multi-robot motion planning, assembly planning, and a variety of modern manufacturing techniques,” Halperin says. The emphasis is on tight settings, where the implementation of algorithms is difficult and subtle, calling for specialized tools like the Computational Geometry Algorithms Library (CGAL).

CGAL provides access to an open source library of geometric algorithms. Halperin sees CGAL as unique in this area, because it can help solve some challenging problems in implementing such algorithms. “CGAL is something special, in my view of robotics,” he says.

Halperin sees many opportunities, problems, and mysteries to solve in robotics, which he expects will provide him with ample material to work on in the coming years.

—John Delaney

Machine Learning, Meet Whiskey

Technologies are coming increasingly closer to approximating the human senses of taste and smell.

PICKING OUT THE differences between high-end whiskeys might be easy for a seasoned Scotch drinker, but until recently, this skill eluded artificial systems. Now researchers at the University of Glasgow have developed an artificial tongue capable of distinguishing between drams of Glenfiddich, Glen Marnoch, and Laphroaig whiskeys with 99% accuracy.

The “tongue” itself consists of tiny gold receptors that measure just 100 nanometers across, or approximately one-thousandth the width of a human hair. These bits of gold, which exhibit unusual optical properties at the nanoscale, function as artificial taste buds. When exposed to a liquid, the receptors change color. The researchers measure and track the changes across multiple receptors, then build up a statistical model of a given liquid’s attributes.

Overall, the system works as a human tongue does, according to University of Glasgow biomedical engineer Alasdair Clark, the lead researcher on the project. “If we humans are given a beverage, we can tell if it’s whiskey or water, but we couldn’t tell you the chemicals inside,” Clark notes. “The artificial tongue works in a similar way. It can’t tell you about the chemicals, but you can train it to recognize particular beverages.”

Scientists have been working to develop artificial tasting systems and electronic noses since the late 1980s, but those initial efforts were stalled by technological shortcomings and a lack of understanding of the physiology and information processing behind taste and smell.

Today, that situation has changed.

“We have a lot more detailed information about how the biological systems behind taste and smell work now,” says biochemical engineer Krishna Persaud of the U.K.’s Universi-



The smartphone displays the results of Hypertaste, an AI-assisted “tongue” app from IBM Research that uses electrochemical sensors to analyze complex liquids.

ty of Manchester. “The technology has also improved, in terms of computing, electronics, and pattern recognition. This has allowed much more powerful instruments to be developed.”

The Complexity of Taste and Smell

By the early 1990s, researchers had realized that smell and taste would not be as easy to replicate as vision or hearing. “It’s very easy to define what we’re measuring with vision or audio, but for smell that is highly, highly complex,” says physicist Jan Mitrovics of JLM Innovations, a German e-nose technology company. “We can build cameras that see much better than a human, but when it comes to smell, to odor, that’s not really possible.”

The human olfactory system relies on approximately 400 different types of smell receptors. A given chemical might activate roughly 10% of these receptors in the nose, and our brains then process this information to identify a particular smell. “If you think of a biological nose, what you have are lots of sensors which can detect chemicals, but they’re not very selective,” says Per-

saud. “When you sniff an odor, which very often contains hundreds of compounds, you are producing patterns of signals, and that pattern of information is then interpreted by your brain as the odor of an apple, coffee, or a flower.”

While mimicking this approach in artificial systems remains a challenge today, the ability to share training data and machine learning models has led to advances, according to materials scientist Patrick Ruch of IBM Research – Zurich. Ruch and his group recently developed Hypertaste, an artificial tongue that works like human taste and smell. Hypertaste consists of an array of polymers deposited on electrode pads. Different liquids will activate different polymers, generating unique voltages. By measuring and recording the voltage patterns, Hypertaste compiles a kind of digital fingerprint of a liquid.

“A sommelier would taste a lot of different wine samples to estimate different types of taste and sensory attributes,” Ruch explains. “In a similar way, we have to train an electronic tongue before it gives us useful information. We can train it to recognize

different liquids by associating digital fingerprints with these liquids.”

Hypertaste is sensitive enough to distinguish between mineral waters from different regions with an accuracy of up to 97%. The system, packaged into a small device that can be clipped to the edge of a glass, relays data to a mobile app, which transfers the data to the cloud, where it is fed into a machine learning model. This allows Hypertaste to be used on liquids it has not been exposed to before. “Within a second, you get a reply back from the model that calculates a confidence score,” Ruch says, “and based on that, it tells you the class of liquids in the training database most similar to the tested one.”

At JLM Innovations, Mitrovics and his team develop similar technologies focused on smell, including the SniffPhone, a European Union-funded project that involved multiple institutions and could yield a non-invasive medical diagnostics tool. When a patient breathes into the handheld SniffPhone, an array of sensors captures a pattern of the person’s breath. The information is transferred to a cloud server for processing, and the results are presented to a doctor.

Mitrovics has been involved in artificial smell since the early 1990s, and he says such capabilities were not possible until recently. “The sensor technology has evolved and become cheaper, more reliable, and much smaller,” he notes. “We can build small intelligent devices that use artificial intelligence and pattern recognition, and that has led to many developing applications.”

A World of Applications

The medical diagnostics potential of smell has been recognized for a while—dogs can alert diabetics when their blood sugar or insulin levels drop too low—but there are other commercial applications as well. The same holds for artificial taste. The tongue developed by Clark and his team could be used to identify counterfeit whiskey, for instance.

Similarly, an artificial taste and smell system developed by the Aromyx Corporation of Mountain View, CA, is being tested for quality-control applications. One Aromyx customer buys large volumes of fruit juice. Occasionally the juice spoils, and the company needs to know as quickly as possible.

Today, the company identifies spoilage using human testers, but Aromyx is testing a more quantifiable alternative. Aromyx develops biosensors modeled after the actual receptors used in human taste and smell. In this case, Aromyx identified an olfactory receptor that responds to spoiled juice, and reproduced the receptor in its artificial system to pick out the signal. “It takes all that human subjectivity out of it and they can measure directly how much of that bad flavor is present,” says Aromyx CEO Josh Silverman.

Artificial noses could be used to sense changes in environmental conditions indoors, and prompt people to open windows. They could more precisely measure the olfactory properties of coffee to ensure large companies maintain a consistent flavor profile in their beans. The potential applications are numerous, but at least one significant hurdle remains before these systems can approach the power of human taste and smell.

“We’re doing pattern recognition,” says Mitrovics from JLM, “and you need lots of data to do good pattern recognition.”

Building Better Datasets

In his company’s case, that means testing medical diagnostic applications with more patients. Clark’s artificial tongue would benefit from more data, too. If it were exposed to thousands of samples, then trained on a machine learning model, Clark predicts his artificial tongue could identify the characteristics of liquids it had never been exposed to previously.

Scientists involved with an IBM project focused on applying artificial intelligence to olfaction recently did exactly that. IBM Research team leader and geneticist Pablo Meyer was able to work with a dataset consisting of ratings from 49 people who ranked and categorized the olfactory characteristics of 500 different molecules using 21 simple words, such as flowery, sweet, and sour. Based on these results, his group then trained a machine learning algorithm using the physical characteristics of the molecules being smelled, and this model was able to accurately predict their smell ratings. Their system can look at a molecule and predict its olfactory characteristics. “The

breakthrough is that for the first time we could really predict words that would describe a molecule,” Meyer explains. “We can predict what something is going to smell like.”

This was all possible, Meyer says, because of the quality of the data. “What has advanced the field is that we were able to find a new dataset,” he notes. “A large single dataset can be that good.”

The next challenge will be compiling more of these datasets, which is no small hurdle given the subjectivity of taste and smell. “It’s not like you’re taking a picture with a smartphone camera or a professional camera; those pictures all look the same, so you can train from one to another,” notes Ricardo Gutierrez-Osuna, a professor in the department of computer science and engineering of Texas A&M University. “With cameras, you have the primary colors: red, green, and blue. With smells, there are no known primary odors.”

Still, researchers are encouraged by recent developments, and point to a bright future for these systems driven by continued improvements in sensors and computing, and increasing knowledge of how the biological analogues work. “The technology that is going to unlock the potential of these artificial tongues and noses is really just emerging,” says Ruch. ■

Further Reading

Macias, G., Sperling, J., Peveler, W., Burley, G., Neale, S., and Clark, A. Whisky tasting using a bimetallic nanoplasmonic tongue. *Nanoscale*, 2019, 11, 15216-15223.

E.D. Gutierrez, E.D., Dhurandhar, A., Keller, A., Meyer, P., and Cecchi, G.A.

Predicting natural language descriptions of mono-molecular odorants. *Nature Communications*, 9(1), 4979, 2018.

Persaud, K.

Towards Bionic Noses. *Sensor Review*, 2017; Vol. 37, No. 2. pp. 165-171.

Ruch, P., Hu, R., Capua, L., Temiz, Y., Paredes, S., Lopez, A., Barroso, J., Cox, A., Nakamura, E., and Matsumoto, K.

A portable potentiometric electronic tongue leveraging smartphone and cloud platforms. 2019 ISOCs/IEEE International Symposium on Olfaction and Electronic Nose (ISOEN).

Gregory Mone is a Boston-based science writer and the author, with Bill Nye, of the Jack and the Geniuses novels.

© 2020 ACM 0001-0782/20/4 \$15.00

How Universities Deploy Student Data

Personalizing efforts to drive greater student retention and success.

JON REID ACKNOWLEDGES he's a bit of a procrastinator, and his study habits can be defined as "more loose and less structured." So when the now-third-year senior learned during his freshman year about a personalized education tool being offered at his school, the University of Michigan (U-M), Reid was immediately on board.

The second semester of Reid's freshman year, he took a Statistics 250-level course, which he says is "notoriously difficult at Michigan." Although Reid is a history major, he was required to take a quantitative reasoning class. "The first day I went into a lecture hall of maybe 300 students and I felt completely overwhelmed," he recalls. "Math is not my strong suit."

Reid says the professor showed a quick tutorial on ECoach in the lecture hall, and students were even offered "a very small amount of extra credit to use it." The concept of receiving personalized support with resources and a checklist of what to do before an exam was appealing because it's "not my personality to go up to a professor."

ECoach helped Reid keep "everything fresh and centralized" and created a plan of attack for the class. "I felt I was staying ahead, and I was very confident going into the first exam." The tool even sent Reid personalized feedback on how he did (he received an A- on the first exam), including the median grade, where he fell and some tips to improve his score.

"I was very taken aback by that. I have never had a class where there was follow up on an exam with feedback and encouragement," he says.

ECoach is the brainchild of U-M professor Timothy McKay, who began looking at student data in his large Introduction to Physics classes in 2008, trying to understand who was succeeding and who was struggling. "Looking

at the data made me recognize the differences in backgrounds and goals, and the reasons for taking physics and the affect toward it," McKay says. Some students were enthusiastic and some were terrified, he says.

"I found myself wanting to be speaking differently to every one of my students," McKay says, "and do that in a way that was informed by who they are, where they're coming from and their trajectory; all things you'd like to know in order to coach them effectively."

ECoach is designed for first-year students taking science, technology, engineering, or mathematics (STEM) classes, who can access the ECoach website with a single sign-on. The personalization ECoach offers is based on data the university already has, such as what courses a student has taken, what their grades have been, and information from their admissions application, "so we can note what their high school background was and their standardized testing" scores, according to McKay.

Students are also asked questions, such as whether they are frightened about taking a particular class. If so, "it's important that we talk to them and let them know if they work with us and follow the advice [in ECoach], they can be successful," McKay says.

A Sense of Urgency

Today, it is becoming the norm rather than the exception for colleges and universities to utilize the data they have within student learning management systems (LMSS) and student information systems (SISs) for academic purposes. The reasons are not entirely altruistic; yes, higher education officials want to help students be successful while in school, but they also want to do whatever they can to keep them there, so the dollars continue to flow in both from students and from state and federal funding.

Data analytics is being done more frequently because "people are starting to really get a sense of how urgent it is" to keep students in school, says Glenda Morgan, a senior director and analyst in market research firm Gartner's Higher Education division.

"Tuition is a big income stream and the proportion of state funding has gone down dramatically and the portion students are paying has gone up," she says. There has also been a mind shift from not just getting students into college but "increasingly, it's 'Okay, we got them in the door; let's get them to succeed and get out on the other side,'" Morgan says.

Also, the analytics tools have gotten better, so campus officials are becoming more knowledgeable about what it takes to help students, especially those at risk of dropping out.

Although analytics in higher education is still in the early stages of implementation, 40% of CIOs say they will receive increased funding for it in 2019, according to Gartner research on the "Top 10 Business Trends Impacting Higher Education in 2019."

"More and more demand is being put on campuses to use analytics to improve student success ... because higher ed campuses are responsible or accountable to numbers of entities" such as accreditors, the federal government for financial aid, and their state for funding, notes Linda Baer, a senior consultant in higher education.

Some states even use performance-based funding for higher education, says Baer. Another compelling reason: in addition to stakeholder accountability, the traditional student population is declining, she says. "One way to counter the [decline in] incoming enrollment is to improve retention and graduation rates, so people are focused on 'how do we keep students staying here instead of dropping out?'"

Right now, one of the hottest uses of data is nudge technology, which incorporates behavioral economics principles to help enhance student outcomes, Morgan says.

Student success was listed as one of Educause's top 10 most pressing IT issues for 2019, according to Kathie Pelletier, director of Student Success Community Programs at the non-profit higher ed association.

"We interpret this as institutions taking responsibility for students' success, not just in the classroom, but as a whole person," Pelletier says. "Responsible use of trusted data is a key enabler here."

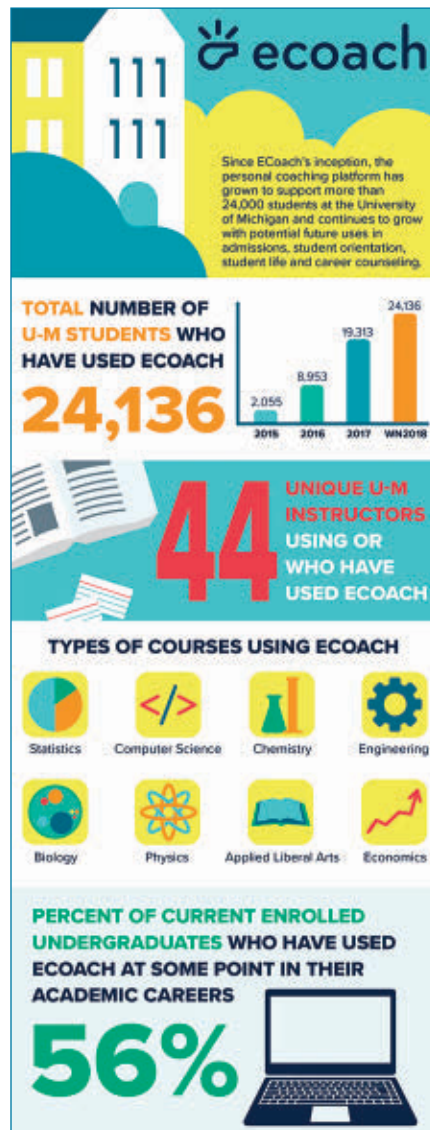
In terms of deployment of analytics technologies, only 7% of institutions have deployed predictive analytics for student success institutionwide and 58% are expanding or planning to do so, according to Educause's 2019 Strategic Technologies report, Pelletier says. Course-level learning analytics are slightly less widespread, with 3% institutionwide deployment and 35% planning or expanding, she adds.

"From these data we might expect student success analytics to become widespread at some point, but the additional complexity of measuring learning, potentially requiring faculty to approach assessment differently, may slow down the deployment of course-level learning analytics," Pelletier says.

Universities are certainly feeling the pressure to help students succeed. In a 2018 joint study by the National Association of Student Personnel Administrators, the Association for Institutional Research, and Educause, 88% of 970 senior-level student affairs professionals strongly agreed or agreed that to stay competitive, they must continue to invest in student success analytics. Among the respondent institutions' goals for conducting student success studies, 96% said they wanted to improve student outcomes from interventions, 71% said they sought more efficient delivery of programs or services, and 39% cited elimination/reduction of programs shown not to contribute significantly to student success.

What Schools Are Doing

At the University of San Francisco (USF), student data within the LMS



Data on the University of Michigan's ECoach digital platform.

such as attendance, grades, and assignments are being used to determine their progress, says vice president and Chief Information Officer Opinder Bawa. "For the first couple of weeks, we're very vigilant on [capturing] that information so we can do early intervention," he says. If a faculty member notices a student struggling, they can use technology developed by USF's IT department to alert school counselors to reach out to that student, Bawa says.

USF IT also built a mobile app called USFMobile, which pulls data from USF's Salesforce system used for advising and counseling, and can generate alerts, he says.

The only people who have access to USF's LMS data are faculty members,

and counselors have access to Salesforce. The university is working on putting in place rules for data governance, Bawa adds.

Eventually, student data also will be used to "really understand what [students] are doing, and what we can do to help them: Are they involved in clubs and sports?" he says. "If not, how can we get them involved and to be more successful at the university?"

However, Bawa adds, "We have to be very thoughtful in which data to use and how much data to use, so we don't cross the line into privacy."

Right now, USF has an 84% graduation rate in six years; Bawa says the goal is to "move the line" higher.

The University of Missouri is piloting a 12-question student success survey this year to help new students integrate to the campus. The purpose is to glean how students are adjusting in terms of academics, social, and financial concerns, and whether they are satisfied with the institution and plan to return, says director of Strategic Initiatives and Assessment for Student Affairs Ashli Grabau. The data is stored in the university's SIS.

If a student indicates any concerns in the survey, the information is used "to reach out to them to make a high-touch response," says Grabau. The impetus for the survey was "to really improve our student success, especially for new students, and for them to feel connected to the institution and get the resources they need to feel financially secure, and resources to improve their academic experience and improve their sense of belonging," she says.

The top three issues the survey revealed were focused around course struggles, class attendance, and financial concerns, according to Grabau; there were no surprise findings. "I think it reinforced that it's important to check in with students early on to see what they're struggling with and intervene, rather than halfway through the semester."

Although the university had a record 87.9% students return this year from last year's freshman class, "It's not as high as we want it to be," Grabau says. "Our goal is 93% and we're doing pretty good, but to move from good to great takes ... a high-touch approach."

Addressing Data Governance

Higher education observers say that as student data is more frequently being sliced and diced, it raises concerns about privacy, interpretation, and misuse.

“One reason students aren’t successful is because they take courses that aren’t useful or they change their mind about their major,” says Gartner’s Morgan. Some predictive analytics tools suggest courses of study and majors based on profiles of others who have been successful in those majors, she says. “They can be hugely successful, but I do worry that sometimes people aren’t questioning the assumptions behind algorithms.”

One of the biggest potential risks with student data is related to equity, says Educause’s Pelletier. “Disaggregating data allows institutions to identify and address potentially hidden opportunity gaps for historically underrepresented students. When institutions are not appropriately disaggregating data based on race, ethnicity, gender, and socioeconomic status, not only do they risk missing insights about the needs of various subpopulations, but they risk furthering the structural barriers and increasing equity gaps.”

Another equity concern related to data is in the way institution officials talk to students about what they “see” in their data, she adds. Consequently, adequate training for advisors, faculty, and other support staff in how to provide culturally responsive guidance is critical to promote a sense of belonging for students, which is a key driver of retention, Pelletier says.

Additionally, “Sharing the wrong data with students, or sharing the right data in the wrong way, can lead to alienating experiences,” Pelletier says.

While Baer believes in “know your students, know your data,” she says the next big principle is having data governance and policies in place. “That is how you secure the data, because a big tenet is to keep student data private,” she says. “And it’s the law.”

“Analytics systems are the new, bright shiny object of higher education, but it’s a new way of looking at the data,” and officials must maintain that tenet as the demand grows for student information in real time, agrees Colleen Carmean, founder and president of the Ethical Analytics Group, an organization created

“Disaggregating data allows institutions to identify and address potentially hidden opportunity gaps for historically underrepresented students.”

to help higher education create student and institutional success via data. Since universities often do not have staff with experience or thoughtfulness on how to use data ethically, a new trend is the rise of the data privacy officer, says Carmean, who is also former associate vice chancellor for academic innovation at the University of Washington-Tacoma.

Carmean recommends universities make certain students are aware of the intentions of any data initiatives being done on their behalf, and to be clear about motivations when sending out email messages asking for information.

“We may think we know what’s good for them, but it’s their data,” Carmean says.

By and large, universities are on top of protecting student data, Baer says. “Our trouble is that computer systems are as strong as you can make them, but there’s always security problems so campuses have to stay very vigilant on that.”

Many campuses also have ethics committees, so there are policies for how student data can be used, as well as who can use it, for what purpose, and how they protect it, Baer says.

Data management and governance was ranked #8 on the Educause top 10 issues list in 2019. Safeguarding student data has been a theme for a while, but the type of data available and the number of systems using and storing data have increased in volume and complexity, Pelletier says.

There are resources that can help. For institutions using cloud services, for example, the Higher Education Cloud Vendor Assessment Tool (HECVAT) can provide guidance in

managing risks to the confidentiality, integrity, and availability of sensitive institutional information and the personally identifiable information of constituents, according to Pelletier.

Ultimately, as schools migrate toward a student-focused experience and student success initiatives continue to gain momentum, they will continue seeking ways to engage students in defining what success looks like for them. Schools will also continue leveraging technology to send nudges, reminders, and resources based on each student’s own goals, Pelletier says, all in the name of helping them feel more connected.

Citing the quote “With great power comes great responsibility,” she says the promise of using trusted data to drive outcomes in student experience and student success must be tempered with employing an ethical approach to the collection, storage, and use of data.

“Privacy and security are important for all,” Pelletier says, “yet higher education has a particular responsibility to ensure our data practices are making institutions more student-ready, especially for historically underserved students.” **C**

Further Reading

Baer, L.L. and Carmean, C. *An Analytics Handbook. Moving From Evidence to Impact*, 2019, The Society for College and University Planning. <http://bit.ly/2Cfyl7R>

Responsible Use of Student Data in Higher Education, Stanford University CAROL & Ithaca S-R Project, 2018. <http://ru.stanford.edu/>

Ekowo, M. and Palmer, I. *The Promise and Peril of Predictive Analytics in Higher Education*, 2016, New America. <http://bit.ly/2pqwxpx>

Defining Student Success Data. Recommendations for Changing the Conversation, 2018. Higher Learning Commission. <http://bit.ly/34v7i4s>

Institutions’ Use of Data And Analytics for Student Success, 2018, National Association of Student Personnel Administrators, the Association for Institutional Research, Educause. <http://bit.ly/36DiWMM>

Esther Shein is a freelance technology and business writer based in the Boston area.

© 2020 ACM 0001-0782/20/4 \$15.00

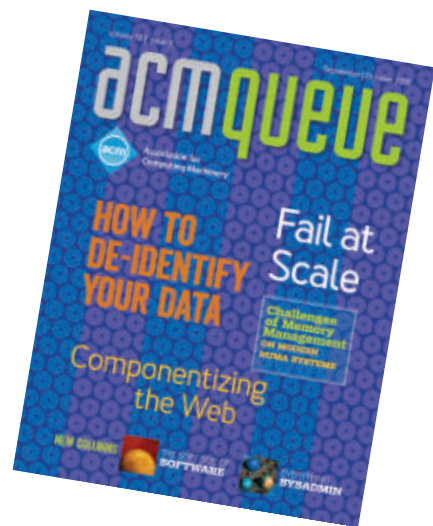
acmqueue

Check out the acmqueue app

FREE TO ACM MEMBERS

acmqueue is ACM's magazine by and for practitioners, bridging the gap between academics and practitioners of the art of computer science. For more than a decade *acmqueue* has provided unique perspectives on how current and emerging technologies are being applied in the field, and has evolved into an interactive, socially networked, electronic magazine.

Broaden your knowledge with technical articles focusing on today's problems affecting CS in practice, video interviews, roundtables, case studies, and lively columns.



Keep up with this fast-paced world on the go. Download the mobile app.



Association for
Computing Machinery

Desktop digital edition also available at queue.acm.org.
Bimonthly issues free to ACM Professional Members.
Annual subscription \$19.99 for nonmembers.

Computing Ethics

The Temptation of Data-Enabled Surveillance

Are universities the next cautionary tale?

THERE IS INCREASING CONCERN about “surveillance capitalism,” whereby for-profit companies generate value from data, while individuals are unable to resist.⁹ Non-profits using data-enabled surveillance receive less attention. Higher education institutions (HEIs) have embraced data analytics, but the wide latitude that private, profit-oriented enterprises have to collect data is inappropriate. HEIs have a fiduciary relationship to students, not a narrowly transactional one (see Jones et al.⁷). They are responsible for facets of student life beyond education. In addition to classrooms, learning management systems, and libraries, HEIs manage dormitories, gyms, dining halls, health facilities, career advising, police departments, and student employment.

HEIs collect and use student data in all of these domains, ostensibly to understand learner behaviors and contexts, improve learning outcomes, and increase institutional efficiency through “learning analytics” (LA). ID card swipes and Wi-Fi log-ins can track student location, class attendance, use

of campus facilities, eating habits, and friend groups. Course management systems capture how students interact with readings, video lectures, and discussion boards. Application materials provide demographic information. This data is used to identify students needing support, predict enrollment demands, and target recruiting efforts.

These are laudable aims. However, current LA practices may be inconsistent with HEIs’ fiduciary responsibilities. HEIs often justify LA as advancing student interests, but some projects ad-

vance primarily organizational welfare and institutional interests. Moreover, LA advances a narrow conception of student interests while discounting privacy and autonomy. Students are generally unaware of the information collected, do not provide meaningful consent, and express discomfort and resigned acceptance about HEI data practices, especially for non-academic data (see Jones et al.⁷).

The breadth and depth of student information available, combined with their fiduciary responsibility, create a duty that HEIs exercise substantial restraint and rigorous evaluation in data collection and use. Consider several recent examples.

Three Cases

Movement tracking. Based on student ID card swipes, a university researcher mapped student movements and social networks and built student retention models. It plans to use Wi-Fi router data to form even more detailed understandings and to share this information with advisors.¹ It is unclear whether students are aware their data is collected, have

Current learning analytics practices may be inconsistent with higher education institutions’ fiduciary responsibilities.



opted in, or provided informed consent. To act in students' interests, this HEI could have provided substantial information before the study started (including its rationale), the ability for students to easily opt-out, and a clear policy about collection and use of movement-tracking data. This university is one of many engaging in intensive student tracking, either by dedicated beacons, Wi-Fi check-ins, or phone apps.⁴

Third parties. In April 2018, researchers from Pearson publishing revealed they had conducted an experiment by incorporating encouraging "growth mindset" messages into a learning software interface and testing (without students' knowledge or consent) whether they affected students' performance.⁶ This demonstrates the value of student information to third parties for non-educational goals, including corporate profit. Similarly, Piazza, maker of a popular question-and-answer app required by many instructors, has sold student data based on students "opting-in" through a pre-checked box on the app sign-up page. Acting in students'—rather than vendors'—interests demands that HEIs develop stronger controls to

protect students and avoid learner data becoming part of surveillance capitalism. At the beginning of the relationship with vendors, HEIs should require that edTech companies make opting-in difficult. If the data collection involves interventions or data sales, HEIs should re-evaluate the relationship, and perhaps require the companies compensate students for their data.

Intensive advising. Like many HEIs, Georgia State University (GSU) has struggled to ensure students (in particular those from underrepresented backgrounds) complete their degrees. In 2011, GSU developed a system tracking academic and financial information that alerts advisors about risk factors (for example, an unsatisfactory grade in a key course). GSU's six-year graduation rate rose from 48% in 2011 to 55% in 2018.⁵ Moreover, students of color, Pell-eligible, and first-generation students now graduate at higher rates than the student body overall.³ The GSU case is often described as a LA success. However, GSU simultaneously hired dozens of new advisors and substantially increased student advising. Hence, it is not just an analytics program; it is an

advising program informed by data.

Certainly, the improved student outcomes are important, but it is not clear how much is due to LA, exactly what GSU's advising interventions actually are, and whether they provide the best outcomes for each student or just for GSU. Do they steer students away from challenging courses, reducing agency and potential for excellence, or provide tutoring services for at-risk students—increasing their agency and capabilities? It is unclear the increased student surveillance improved social good, or that GSU considered all of the relevant moral trade-offs. It is crucial to ensure we do not learn the wrong lesson and retain the analytics at the expense of the advising. Acting in students' best interests would require a robust, ongoing evaluation.

Doing It Differently

The cases mentioned in this column illustrate a range of issues in LA, and each demonstrates how HEIs can better fulfill their responsibilities to advance student interests. HEIs have responsibilities before developing LA programs, while student data is collected and analyzed (especially by third parties), and after analytics

have been incorporated. Here, we make explicit some of the specific responsibilities of HEIs have as information fiduciaries that can guide their actions.

Diverging interests. Fiduciaries have a responsibility to act in the best interests of their clients, although not necessarily to act *only* to advance client interest. Data analytics may provide insights for HEIs and help them fulfill their responsibilities to educate students and marshal resources effectively. However, it is an open question whether LA will live up to that promise or that primarily *student* interests motivate LA research. The tracking and third-party data use cases do not advance student interests.

Moreover, universities collect and use data somewhat indiscriminately because it is potentially relevant to their educational and custodial missions. Yet this “relevance condition” is insufficient to justify data collection, analysis, and use.⁸ Any student data is potentially relevant to educational objectives and it is impossible to tell a priori what data will actually be useful. Hence, a collection principle based solely on potential relevance is no limitation at all.

Fostering trust and trustworthiness. Students attend HEIs believing they are trustworthy, will respect students as individuals, and will not implement systems that subordinate student rights and interests for the sake of institutional or third-party goals. Yet systems currently being built and deployed create opportunities for greater privacy intrusions (the tracking case) and for institutional benefit (movement tracking, third-party use). Students have little knowledge of how they are surveilled, typically have no ability to opt out, are uneasy about data collection by HEIs, and cannot control use of their data. It is worth asking whether their trust is misplaced, or whether they must simply acquiesce because of the social and economic value of a college education.⁴

The right benefits. The benefits of LA may be quite limited. The exemplary LA system (GSU) rests substantially on advising resources, not data, with small increases in student success. It is unclear whether success stems from funneling students into easier courses or from collecting and analyzing troves of student data.

HEIs do have obligations to advance the educational interests of at-risk stu-

Institutional interests and student interests are not identical, and we should not assume they align.

dents. However, it does not follow that HEIs should subordinate students' privacy and autonomy interests for the sake of (speculative) retention and achievement rates. Student support (social, advising, tutoring, financial, mental health) should come first, and long before impinging on other interests.

Full account of student interests. We should also be wary of narrowly constraining student interests. The aforementioned cases focus on student academic achievement, but students have other interests as well, including privacy. They should not have to forego that interest for a marginal (and speculative) return in academic achievement. This should seem familiar to professionals who bristle at overweening surveillance by supervisors.

Full range of higher education's aims. Higher education has a number of aims, including developing communication, critical thinking, understanding and appreciation of diversity, and development of rewarding employment and careers.² These are fostered by helping students develop their autonomy. Yet close monitoring of student movement, social networks, and daily habits is an imposition on student privacy, a key element in developing and exercising individual autonomy. If we value student autonomy, we ought to curtail student surveillance.⁸

Third parties. Students' social networks, their travel around campus, their health, political, and religious activities are inferable from movements. If that information is valuable for HEIs, it will be of interest to others: potential employers, the FBI (pursuant to business records requests under Patriot Act section 215), software vendors, and more. Student privacy protections under FERPA are limited. Moreover, the more student data collected the greater the risk for data breaches. Hence, LA carries

risks that we should consider in determining whether it advances student interests overall.

Conclusion

To be clear, we are *not* opposed to LA tout court. Some student data collection will advance legitimate educational and custodial goals while respecting student privacy and fostering autonomy. However, LA and data-enabled surveillance can begin as tools for social good but slide into morally suspect territory, especially in immersive institutions with fiduciary responsibilities like HEIs. Institutional interests and student interests are not identical, and we should not assume they align. Governments, corporations, and HEIs should avoid data collection and analysis as a matter of convenience and for specious reasons. When HEIs aim to advance student educational interests, they should primarily do so in ways that are consistent with interests in privacy and autonomy. In short, HEIs should conduct LA in ways that justify the substantial trust that students place in them. That will require restraint and evaluation at the beginning, during, and at the end of LA projects. **□**

References

1. Blue, A. Researcher looks at 'digital traces' to help students. University of Arizona News. (Mar. 7, 2018); <http://bit.ly/31rCzVP>.
2. Bok, D.C. *Our Underachieving Colleges: A Candid Look at How Much Students Learn and Why They Should Be Learning More*. Princeton University Press, Princeton, NJ, 2006.
3. Ekowo, E. and Palmer, I. The promise and peril of predictive analytics in higher education: A landscape analysis. *New America* (Oct. 24, 2016).
4. Harwell, D. Colleges are turning students' phones into surveillance machines, tracking the locations of hundreds of thousands. *Washington Post*. (Dec. 24, 2019); <https://wapo.st/3bjnsSX>.
5. Hefling, K. The 'Moneyball' solution for higher education. *Politico*. (Jan. 16, 2019).
6. Herold, B. Pearson tested 'social-psychological' messages in learning software, with mixed results. *Education Week* (Apr. 17, 2018); <http://bit.ly/39bKuJw>.
7. Jones, K.M.L., Rubel, A., and LeClere, E. A matter of trust: Higher education institutions as information fiduciaries in an age of educational data mining and learning analytics. *Journal of the Association for Information Science & Technology* (2019), 1–15; doi: 10.1002/asi.24327.
8. Rubel, A. and Jones, K.M.L. Student privacy in learning analytics: An information ethics perspective. *The Information Society* 32, 2 (Feb. 2016), 143–159; doi: 10.1080/01972243.2016.1130502.
9. Zuboff, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs, New York, NY, 2019.

Alan Rubel (arubel@wisc.edu) is an associate professor in the Information School and director of the Center for Law, Society & Justice at the University of Wisconsin-Madison, Madison, WI, USA.

Kyle M.L. Jones, (kmlj@iupui.edu) is an assistant professor in the School of Informatics and Computing at Indiana University-Indianapolis (IUPUI), Indianapolis, IN, USA.

Copyright held by authors.

Technology Strategy and Management

Artificial Intelligence and the Future of Professional Work

Considering the implications of the influence of artificial intelligence given previous industrial revolutions.

IF YOU ARE a software engineer or a data scientist, your job did not exist a century ago. A century from now, your job will most likely look quite different. One driving force behind such work transformation is artificial intelligence (AI). Dwelling on the nearer term, the next decade or two, projections on the proportion of today's jobs that are susceptible to automation vary enormously—from 9% (in OEDC countries)² and 47% (of 702 occupations)⁷ to 96% (740 out of 769 occupations).^{6,8,9} Why is there such wide variation? What makes it difficult to predict with greater precision? And should we alter the way we think about jobs, given that better education is no longer a protection against risk of technological unemployment? This column addresses these questions, so that we might make better decisions about the future of work for our children and grandchildren.

History of Automation and Its Impact on Jobs Before AI

We are in the midst of the so-called fourth Industrial Revolution that fuses advances in AI, robotics, the Internet of Things, 3D printing, genetic engineering, quantum computing, and other technologies to bring about enormous improvements in efficiency and productivity. A brief historical review of how technologies in the earli-



er industrial revolutions affected work helps trace implications for the current industrial revolution.

The first Industrial Revolution was associated with the advent of the steam engine in the 18th century, enabling the mechanization of production. Productivity increased in textile and other factories as they switched their energy source from watermill to steam. The second Industrial Revolution in the 19th century was triggered by electricity and the application of scientific principles, which led to the proliferation of mass production. Im-

ages of the early 19th-century Luddites who destroyed textile machinery as a way of protesting against mechanization and of Charlie Chaplin hardly keeping pace on the assembly line in the film *Modern Times* remind us that a fundamental change in work was taking place. Industrial engineers designed production processes, and factory workers executed the pre-planned tasks typically on an assembly line. In effect, craftsmen's work was disaggregated into standardized tasks that could be carried out by semi-skilled operatives.

The third Industrial Revolution began with the emergence of computing machinery in the 1950s. This led to further automation of manufacturing using Computer-Aided Design (CAD), Computer-Aided Manufacturing (CAM), and the application of digital technology to communications (with the Internet), banking (with ATMs), and other service industries including logistics. While the first two industrial revolutions led to the substitution of mechanical power for human brawn (that is, muscle and handiwork), the third Industrial Revolution computerized human brainwork of the repetitive and routine sort (that is, financial calculation using spreadsheets). And computerization automated a broad range of white-collar workers including clerical, technical, and professional workers.

What's Different with AI?

AI automates tasks normally requiring human intelligence, a definition that distinguishes itself from automation of manual tasks. How does AI change what human workers do? Before the spread of machine learning (ML), enabled by massive processing power and data storage capabilities, the following propositions became compelling.

First, a job consists of interrelated tasks, and it is tasks, not jobs, that are automated by computers.³ 'Routine' tasks, which are easy to codify, tend to be automated; 'non-routine' tasks that require more tacit problem-solving capabilities, intuition, creativity, and persuasion are difficult to automate. For example, financial analysts build models to automate the prediction of stock prices, but they cannot automate the task of interpreting the prediction results and advising their clients. A key reason for the wide variation in the share of jobs 'at risk of computerization,' mentioned earlier, is due to different ways in which task-level analysis is aggregated into jobs. Second, some tasks are substituted for by computers, but others complement computers.⁴⁻⁵ For example, while automating calculation substitutes for human calculation, the rise in demand for automated calculation leads to a demand for complementary tasks in computer programming by humans.

With the advent of machine learning (ML), these propositions must be

modified. First, *the distinction between 'routine' and 'non-routine' tasks may cease to make sense in considering what human tasks are substituted for, and complemented by, AI.* We might think of AI as pushing the frontier of what are non-routine tasks, as machines become more capable of codifying what used to be tacit. However, an alternative, quite different, way of thinking about this is to consider machines performing tasks in such a way the 'routine' vs. 'non-routine' distinction no longer applies. This is because machines can follow rules that do not need to reflect the rules that human beings follow at all. For example, rules followed by a machine—be it a self-driving car or a facial-recognition algorithm—are not the same as the rules followed by humans engaged in the same activity. We therefore wish to avoid the 'AI fallacy'¹² of believing the only way to develop AI systems is to replicate the thinking process of human experts. This has become obvious with ML, letting machines infer rules from abundant data, in contrast to 'expert systems' for which a human domain specialist must articulate a set of rules for a machine to follow.

Second, *new tasks will be created by AI, but how these tasks will be bundled into jobs, new or existing, remains uncertain.* One study classifies new jobs (though they are really tasks) that AI will create into three types—trainers, sustainers, and explainers.¹³ *Trainers* teach AI systems—chatbots or digital assistants such as Siri and Alexa—how they should perform, especially in showing empathy or in detecting sarcasm. *Explainers* as algorithm forensic analysts know the inner workings of complex algorithms, and can explain to nontechnical human

AI automates tasks normally requiring human intelligence, a definition that distinguishes itself from automation of manual tasks.

resource professionals, for example, how a recruitment system identified the best candidate for a job. *Sustainers* help ensure AI systems are operating as designed, and that unintended consequences are addressed with appropriate urgency. Training, explaining, and sustaining appear to require different knowledge bases. Will they, then, be carried out by separate expert groups? Or will these roles be incorporated into existing occupations (for example, compliance professionals incorporating the sustaining task)? We do not yet know the answer to these questions.

Impact of AI on Professional Jobs

What is the implication of the preceding discussion on professional work of the future? Professional jobs are part of the 'professional, technical, and managerial occupations' in government statistics, and include scientists and engineers, as well as accountants, lawyers, and doctors. We focus on professionals, defined as possessing an expertise, a body of knowledge, and a service ethic,¹¹ because they are said to be most threatened by AI. This focus may provide a clue to the possible emergence of new professions.

First, just as craftwork was disaggregated into tasks in the second Industrial Revolution, professional work has been subjected to task disaggregation in the third and fourth Industrial Revolutions. Legal practice of giving advice to clients, for example, may be disaggregated into the tasks of defining the client's problem, reviewing documents around the problem, and explaining the result of that analysis to the client. With the application of AI and ML in particular, machines substitute for the task of reviewing documents. Machines also require complementary task inputs from data scientists and project management professionals working in multidisciplinary teams.¹ Thus, it helps to identify which professional tasks can be substituted by AI, and which tasks are complemented by AI.

Second, technological frontier is not the only factor determining what machines do and what humans end up doing. Humans may remain in the loop to ensure maintaining social and ethical norms. In particular, professional norms, sometimes buttressed by government regulation, enable a

professional group to claim an exclusive domain of expertise. For example, only doctors with a license can practice medicine, and only lawyers with a license can practice law. AI adoption in their specific field may lead to delegating what they used to do to machines, but the oversight function is likely to remain with those licensed professionals.

Third, the way tasks might be re-bundled into professional jobs depends on professional control. In many workplaces, from hospitals to business corporations, multidisciplinary teams of different professionals are becoming important. This multidisciplinary teamwork is likely to lead to more ‘hybrid professionals’ who develop a relational capability vis-à-vis expertise in other areas.¹⁰ At hospitals, doctors may extend their domain to take control of the delivery of good quality patient care, not just to treat patients. Moreover, a new medical practitioner might emerge to ensure the delivery of AI-assisted patient care. Similarly, computing professionals themselves may incorporate all of the training, explaining, and sustaining roles in AI adoption. But an equally plausible scenario is the rise of new AI professionals that focus on the new tasks of training, explaining, and sustaining. The revised ACM Code of Ethics and Professional Conduct obliges computing professionals to “monitor the level of integration of their systems into the infrastructure of society.”^a However, the more ubiquitous the AI technology becomes, the more challenging it would be to incorporate these new tasks into the existing computing profession.

Thus, the professional control perspective gives some hint at, but leaves open a variety of resolutions to, how AI may give rise to new occupations. Because social and ethical concerns are paramount, how tasks are bundled into professional work is likely to differ from occupation to occupation.

Conclusion

The impact of AI on the future of work should be framed in terms of tasks, not jobs, automated by AI. AI substitutes some tasks, complements others, and creates new tasks. How this complex interplay of substitution, complementarity, and creation rebundles tasks

The impact of AI on the future of work should be framed in terms of tasks, not jobs, automated by AI.

into existing or new jobs remains uncertain. For this, we must take account of social and professional norms over and above technological feasibility. The professions perspective is useful for considering the future of work, as professionals are increasingly expected to become ‘hybrid’ in capability relating to and sometimes incorporating expertise in other areas. □

a See <http://bit.ly/2GYtNFq>

References

1. Armour, J. and Sako, M. AI-enabled business models in legal services: From traditional law firms to next-generation law companies. *Journal of Professions and Organization* 7, 1 (2020).
2. Arntz, M., Gregory, T., and Zierahn, U. The risk of automation for jobs in OECD countries. OECD, Paris, France, 2016.
3. Autor, D.H. Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives* 29, 3 (Mar. 2015), 3–30.
4. Brynjolfsson, E. and McAfee, A. *The Second Machine Age*. W.W. Norton & Company, New York, 2014.
5. Brynjolfsson, E. and Mitchell, T. What can machine learning do? Workforce implications. *Science* 358, 6370, 2017, 1530–1534.
6. Delloit, B., Mason, R., and Wallace-Stephens, F. *The Four Futures of Work: Coping with Uncertainty in an Age of Radical Technologies*. RSA, London, 2020.
7. Frey, C.B. and Osborne, M.A. The future of employment: How susceptible are jobs to computerization? *Technological Forecasting and Social Change*. 114 (2017), 254–280.
8. Manyika, J. et al. *A Future that Works: AI, Automation, Employment, and Productivity*. McKinsey Global Institute, 2017.
9. Muro, M., Whiton, J., and Maxim, R. What jobs are affected by AI? Brookings Metropolitan Policy Program, Washington, D.C., 2019.
10. Noordegraaf, M. Hybrid professionalism and beyond: (New) Forms of public professionalism in changing organizational and societal contexts. *Journal of Professions and Organization* 2, 2 (Feb. 2015), 187–206.
11. Sako, M. The business of professionals. *Commun. ACM* 56, 7 (July 2013), 30–32.
12. Susskind, R.E. and Susskind, D. *The Future of the Professions: How Technology will Transform the Work of Human Experts*. Oxford University Press, USA, 2015.
13. Wilson, H.J., Daugherty, P., and Bianzino, N. The jobs that artificial intelligence will create. *MIT Sloan Management Review* 58, 4 (Apr. 2017), 14.

Mari Sako (mari.sako@sbs.ox.ac.uk) is Professor of Management Studies at Saïd Business School, University of Oxford, U.K.

Copyright held by author.

Calendar of Events

Mar. 30–Apr. 3

SAC '20: The 35th ACM/SIGAPP Symposium on Applied Computing, Brno, Czech Republic, Sponsored: ACM/SIG, Contact: Chih-Cheng Hung, Email: chung1@kennesaw.edu

Apr. 20–24

ICPE '20: ACM/SPEC International Conference on Performance Engineering, Edmonton, Canada, Co-sponsored: ACM/SIG, Contact: Jose Nelson Amaral, Email: amaral@cs.ualberta.ca

Apr. 21–24

CPS-IoT Week '20: Cyber-Physical Systems and Internet of Things Week 2020, Sydney, Australia, Sponsored: ACM/SIG, Contact: Wen Hu, Email: wen.hu@unsw.edu.au

Apr. 21–25

ICCPs '20: ACM/IEEE 11th International Conference on Cyber-Physical Systems, Sydney, Australia, Sponsored: ACM/SIG, Contact: Linda Bushnell, Email: lb2@uw.edu

Apr. 21–25

IPSN '20: The 19th International Conference on Information Processing in Sensor Networks, Sydney, Australia, Sponsored: ACM/SIG, Contact: Kusy Branislav, Email: brano.kusy@csiro.au

Apr. 22–24

HSCC '20: 23rd ACM International Conference on Hybrid Systems: Computation and Control, Sydney, Australia, Co-sponsored: ACM/SIG, Contact: Aaron Ames, Email: ames@caltech.edu

Apr. 25–30

CHI '20: CHI Conference on Human Factors in Computing Systems, Honolulu, HI, Co-sponsored: ACM/SIG, Contact: Philippe Palanque, Email: palanque@irit.fr



Article development led by **acmqueue**
queue.acm.org

Kode Vicious Master of Tickets

Valuing the quality, not the quantity, of work.

Dear KV,

I hope it is OK to write a letter to you as I am not a programmer, but I enjoy your columns (very funny and informative!) and I see you sometimes discuss topics other than code. I work in IT support, and the company I work for evaluates us based on the number of tickets we close. This seems to be a poor way to judge performance, as it is easy to game the system. For instance, you can take a lot of trivial tickets and close them quickly and look like a rock star. I guess my question is more about how to value work, which is maybe too big a topic for a KV column.

The Tickets That Exploded

Dear Tickets,

Questions about how work is valued can definitely go beyond the scope of a short column, given there are whole areas of politics, economics, and the social sciences devoted to these questions. It is also the case that wars—both cold and hot—have been fought over this topic, and still the question is unsettled.

Many silly metrics have been created to measure work, including the rate at which tickets are closed, the number of lines of code a programmer writes in a day, and the number of words an author can compose in an hour. All of these measures have one thing in common: They fail to take into account the quality of the output. If Alice writes 1,000 lines of impossible-to-read, buggy code in a day and Carol writes 100 lines of well-crafted,



easy-to-use code in the same time, then who should be rewarded?

Plenty of companies have chosen the easier metric—which is to reward Alice—because she appears to be more productive. This false measure of productivity stems from the nature of industrial work from which modern knowledge work (which includes IT and programming, as well as plenty of other endeavors) sprang. In an industrial system, the steps required to make something like a shirt are broken down (some might say “atomized”), such that any able-bodied worker could do the single task

assigned them for hours on end at a measurable rate.

In such a system, the people were used as machines and were eventually replaced by machines, because what was needed was not knowledge or skill but simply the ability to assemble—based on a predetermined plan—a set of objects, such as sleeves, collars, and buttons, into a larger object, a shirt. Of course, if you have ever shopped for a shirt you know they vary in quality. Cheap clothing is often made by machines and with minimal human intervention, because a machine can be told how to assemble

a shirt, just not always a very well-tailored one. The highest-quality, expensive clothing is still made in part by hand, because making something of quality requires not just rote movements, but also intellect and thought.

The very same tug of war has existed in knowledge work since its inception. What management wants is maximum output for minimum input, and input in this case is you, the human, who is doing the work. When management implements silly metrics such as those mentioned here, you have only a few choices.

The first choice is to game the system such that you can actually do minimum work to get the maximum benefit, usually by taking on trivial tasks that can easily be completed and will show at the end of each month that you accomplished more tasks than anyone else. After a while, that game becomes boring, but in large organizations you can do quite well with it, get promoted to management and then foist your own silly metrics on your new underlings, thereby perpetuating the pain and suffering you

once suffered and helping to bring about the downfall of civilization through a vicious cycle of stupidity. I find that this is the most common and unfortunate reaction.

A second choice is to take on difficult and challenging tasks, to learn from what you have taken on, and to hone your skills. You can then take those skills to a company that values good work and write a flaming goodbye letter to your former company. You might even mail it to everyone before you leave or just post it on one of the many anonymous company review sites that now litter the Internet.

A third and probably more difficult choice is to convince management to choose useful metrics, ones that take into account not just the rapidity but also the difficulty of the work being undertaken and the quality of the output. The problem is, I fear, that you are in an organization where this sort of discussion cannot take place. Once some companies have a metric, they will hold on to it like a drowning person, even if they find out their perceived float is actually an anchor.

KV favors option two. Learn what you can and then go find a place where your work is valued and where you are not simply being used as a tool.

KV

Q Related articles
on queue.acm.org

Gardening Tips

A good library is like a garden.

Kode Vicious

<https://queue.acm.org/detail.cfm?id=1870147>

GitOps: A Path to More Self-service IT IaC + PR = GitOps

Thomas A. Limoncelli

<https://queue.acm.org/detail.cfm?id=3237207>

System Administration Soft Skills

How can system administrators reduce stress and conflict in the workplace?

Christina Lear

<https://queue.acm.org/detail.cfm?id=1922541>

George V. Neville-Neil (kv@acm.org) is the proprietor of Neville-Neil Consulting and co-chair of the *ACM Queue* editorial board. He works on networking and operating systems code for fun and profit, teaches courses on various programming-related subjects, and encourages your comments, quips, and code snips pertaining to his *Communications* column.

Copyright held by author.



Advertise with ACM!

Reach the innovators and thought leaders working at the cutting edge of computing and information technology through ACM's magazines, websites and newsletters.

Request a media kit with specifications and pricing:



Ilia Rodriguez
+1 212-626-0686
acmm mediasales@acm.org

The National Academies of SCIENCES • ENGINEERING • MEDICINE

ARL Distinguished Postdoctoral Fellowships

The Army Research Laboratory (ARL) Distinguished Postdoctoral Fellowships provide opportunities to pursue independent research in ARL laboratories. Fellows benefit by working alongside some of the nation's best scientists and engineers, while enhancing the mission and capabilities of the U.S. Army and the warfighter in times of both peace and war.

ARL invites exceptional young researchers to apply. Fellows must display extraordinary abilities in scientific research and show clear promise of becoming future leaders. Candidates are expected to have already successfully tackled a major scientific or engineering problem or to have provided a new approach or insight, evidenced by a recognized impact in their field. ARL offers five named Fellowships; two of these positions are open for the 2020 competition.

Fellowships are one-year appointments, renewable for up to three based on performance. The award includes a \$100,000 annual stipend, health insurance, paid relocation, and a professional travel allowance. Applicants must have completed all requirements for a Ph.D. or Sc.D. degree by October 1, 2020, and may not be more than five years beyond their doctoral degree as of the application deadline. For more information and to apply, visit www.nas.edu/arldpf.

Online applications must be submitted by May 29, 2020 at 5 PM EST.



Viewpoint

Why Is Cybersecurity Not a Human-Scale Problem Anymore?

Examining the structure of the enterprise attack surface in view of the relative ease with which cyberdefenses can be subverted.

RARELY A DAY goes by that we don't see news about the poor state of affairs in cybersecurity. From data breaches at Target, the U.S. Office of Personnel Management, Sony, Disney, Yahoo!, Equifax and Marriot, the drumroll continues unabated. We are now in a world, where it's a matter of when, not if, an organization is compromised by a cyber-attack.

Most of us think of cybersecurity as a series of controls (tools and knobs) that an organization has to implement, and it seems perplexing why cyber-defenders in the situations mentioned here failed to take the necessary steps to protect themselves. Our focus on addressing cybersecurity challenges has been around inventing new controls (or enhancing existing ones) and implementing them correctly in the enterprise. This is an inadequate view.

In this Viewpoint, we show why cybersecurity is a very difficult problem. The enterprise attack surface is massive and growing rapidly. There are practically unlimited permutations and combinations of methods by which an adversary can attack and compromise our networks. There is a big gap between our current tools and methods, and what is needed to get ahead of cyber-adversaries.

The Enterprise Attack Surface

In order to better understand the nature and structure of the enterprise at-

tack surface, let's take a quick look at the abstract picture of the attack surface as shown in Figure 1.^a On the *x*-axis we have the different parts of the enterprise's extended network *where things can go wrong* from a cyber-security standpoint. On the *y*-axis we have the specific *ways in which these things will go wrong*—also known as *attack vectors* or *breach methods*.

The *x*-axis includes the organization's traditional infrastructure (servers, databases, switches, routers, and so forth), applications (standard and custom), endpoints (managed, unmanaged, mobile and fixed, IoTs, industrial controllers, and so forth), and cloud apps (sanctioned and unsanctioned).

At the right end of the *x*-axis, we have the organization's third-party vendors. The *x*-axis effectively repeats itself recursively in the organization's supply chain, where each third-party vendor is an entity with an *x*-axis and attack surface just like that of the organization, and this brings risk into the enterprise network because of certain trust relationships. The ellipses on the *x*-axis indicate that these categories of assets are large sets. It is quite difficult for most organizations to even enumerate their *x*-axis with accuracy.

^a Readers may want to zoom in on the images in the next few figures to see the axes properly.

On the *y*-axis, we have the different methods of attacks—starting from simple things like weak and default passwords, reused password, passwords stored incorrectly on disk, or transmitted in the clear, on to more complex things like phishing, social engineering, and unpatched software. Further down the *y*-axis, we have zero-day vulnerabilities—security bugs that are “unknown” until they are first used by an adversary. There are quite literally 100s of items on the *y*-axis in dozens of categories.

Each point in this *x*-*y* graph represents one way by which adversaries can compromise an enterprise asset. Note that each point is a vector not just a single number. To see this, consider the highlighted point of Figure 1, Line-of-business apps (*x*-axis) and shared passwords (*y*-axis). This is the idea that perhaps an enterprise employee's password for a personal account (for example, for Yahoo! or LinkedIn) is the same as their password for one of their enterprise app accounts. So, if Yahoo! or LinkedIn is breached, and the passwords were stolen (and were not properly hashed on disk) then the enterprise has a problem,¹¹ perhaps one million enterprise app accounts with reused passwords—easy ways for the adversary to get unauthorized access (Figure 2).

Most cyber-defenders have no idea what this “Password Sharing Risk Vector” looks like for their business. The

Figure 1. The enterprise attack surface.

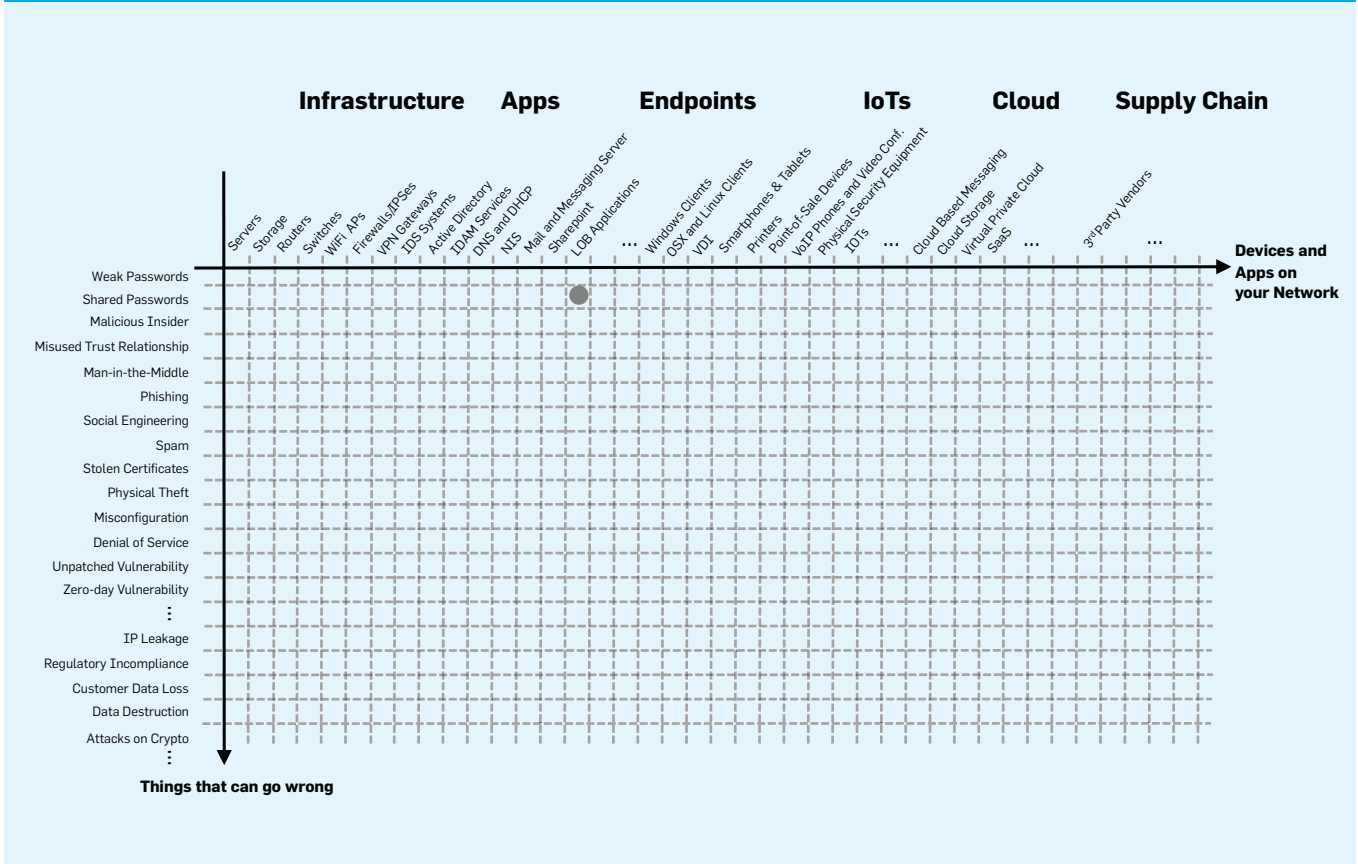


Figure 2. Password reuse in line-of-business apps.

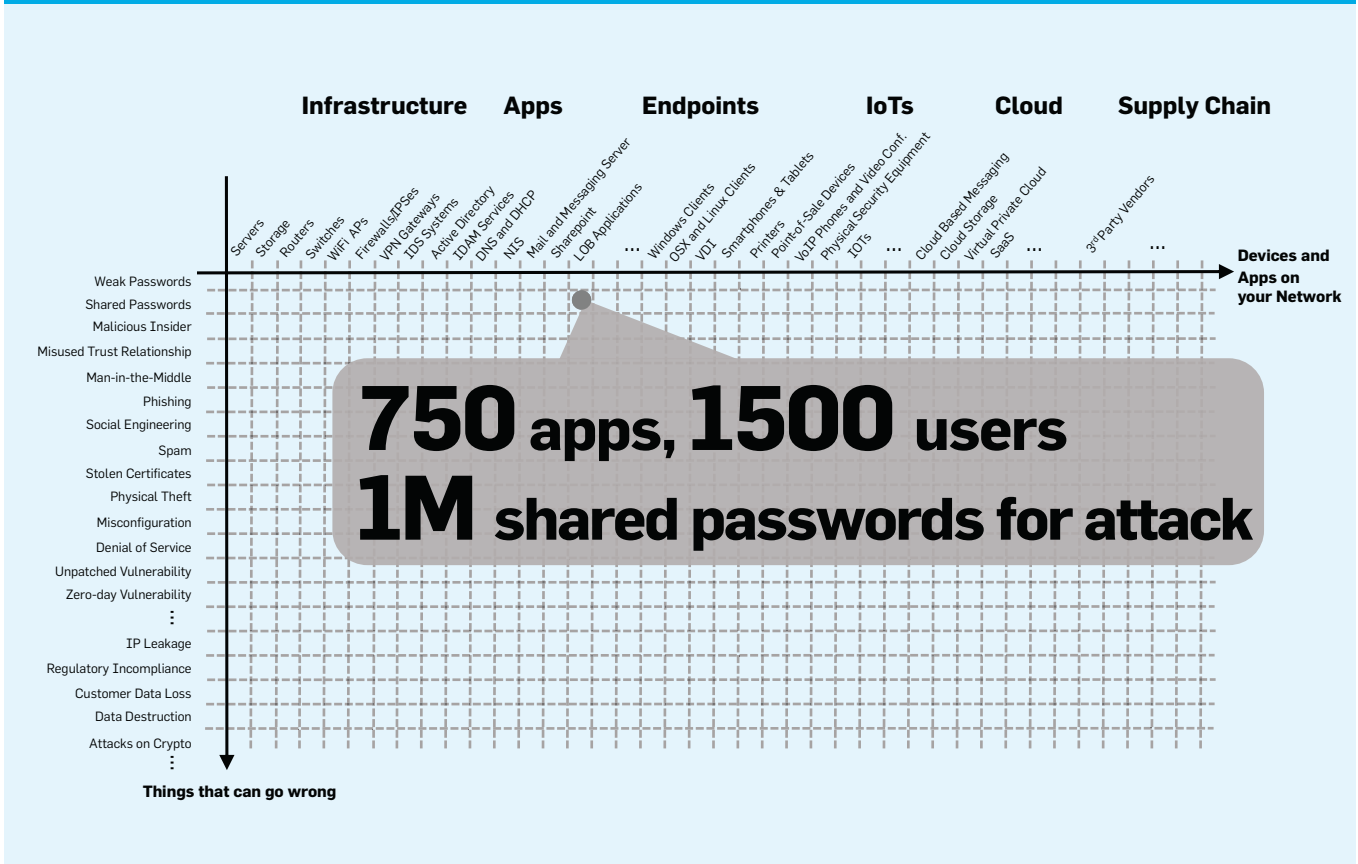


Figure 3. The Equifax breach.

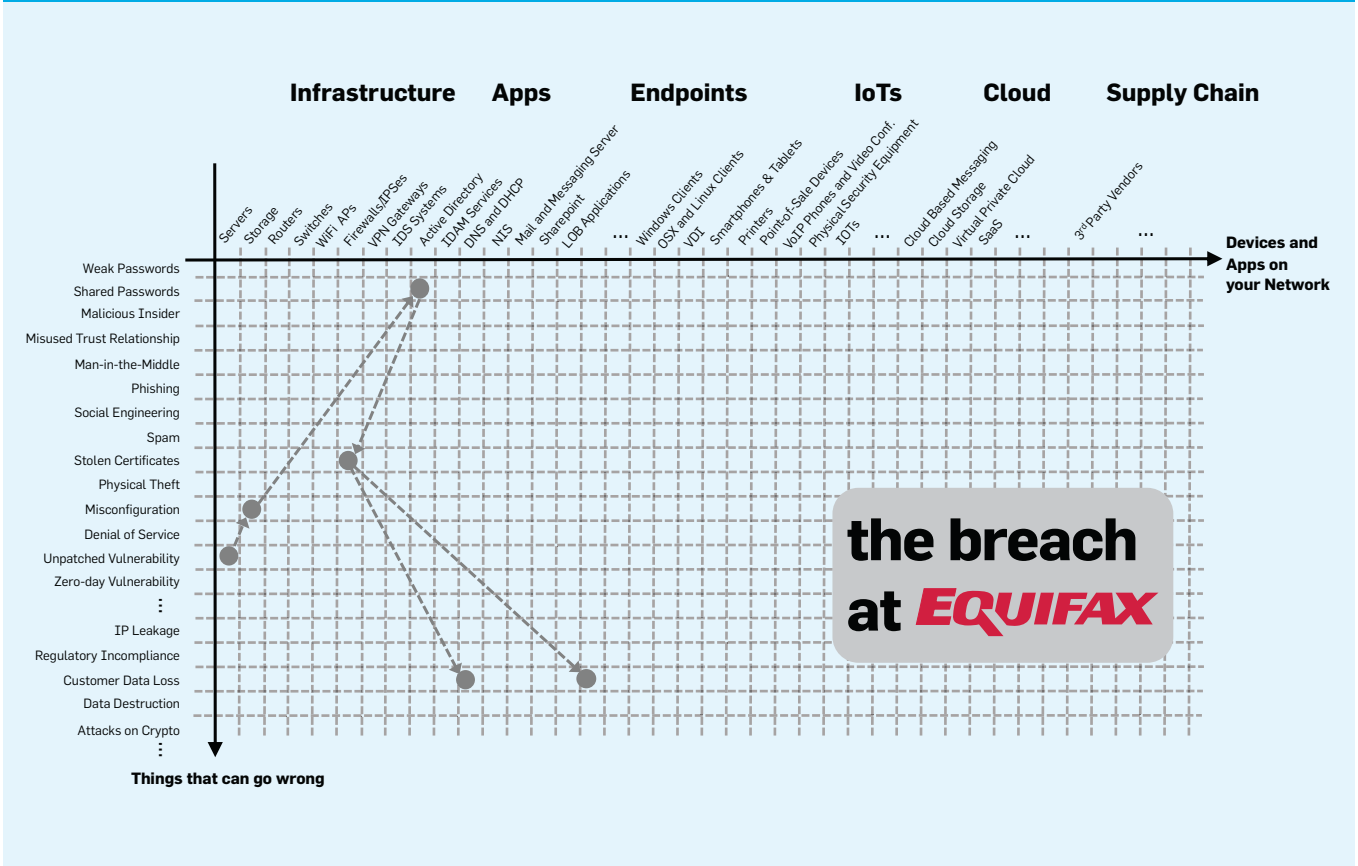
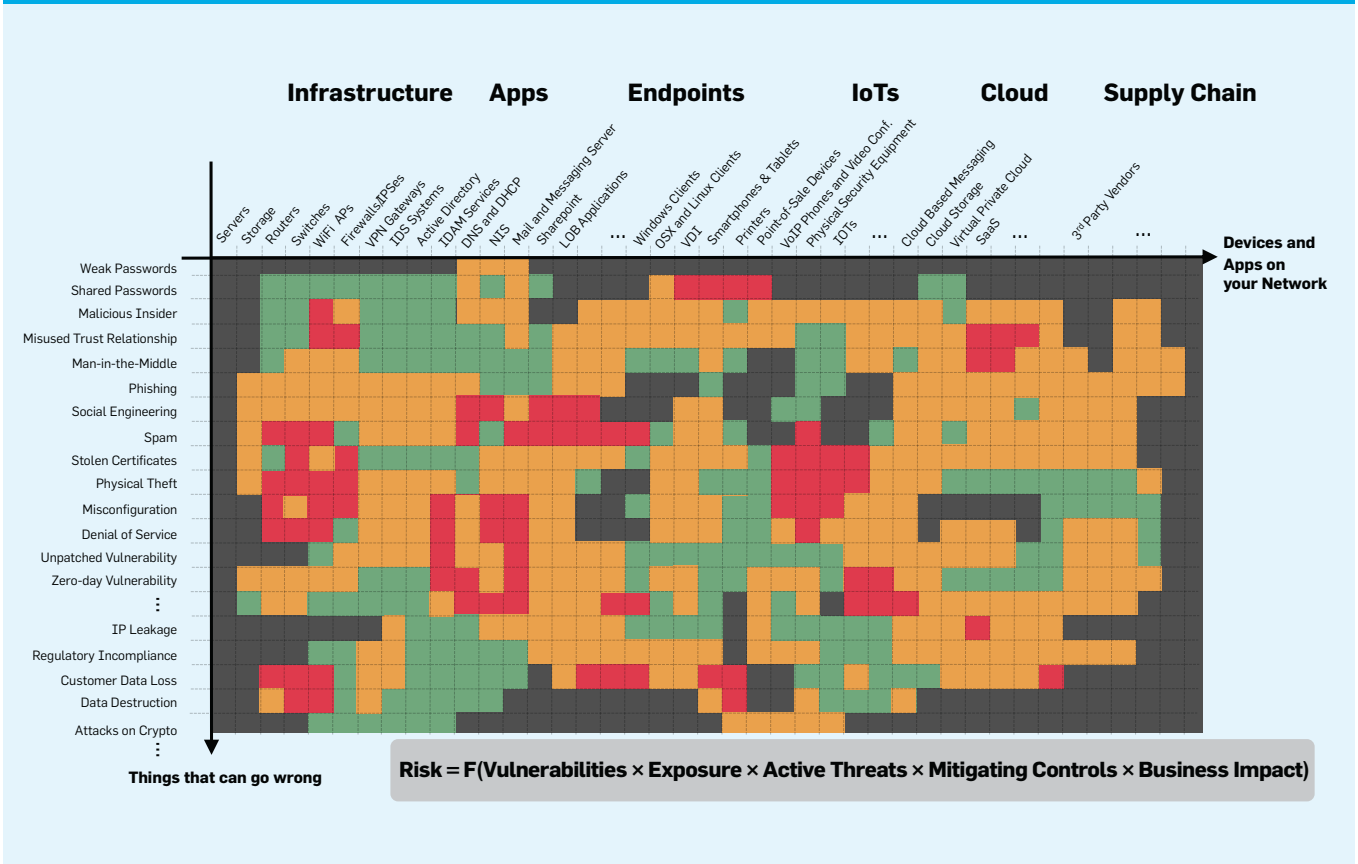


Figure 4. The risk calculation and heatmap.



Verizon Data Breach Investigations Report⁹ claims more than 80% of breaches involve password issues at some stage of the breach.

This gigantic x - y plot is the enterprise attack surface. In a typical breach, adversaries use some point on this attack surface to compromise an (Internet-facing) asset. Other points are then used to move laterally across the enterprise, compromise some valuable asset, and then to exfiltrate data or do some damage. Figure 3 shows how the Equifax breach⁸ unfolded. As you can imagine, the number of items on both of these axes grow as we adopt new technologies in the digital transformation of our businesses.

What Is Our Breach Risk?

For a CISO or CIO, the likely top-of-mind question is at what points on this attack surface is the enterprise at risk. What does the risk-heat map look like, that is, what is $Risk = Likelihood \times Impact$ at each point of the enterprise attack surface?

We must consider the points of the attack surface where we have *vulnerabilities*, for example, unpatched software. We also must factor in *exposure* due to usage—a device with unpatched Internet Explorer is not necessarily a critical risk if the default browser of the user is Chrome. We should prioritize *real threats*—considering what is currently fashionable with (or possible for) adversaries, and not waste time worrying about theoretical issues when we have numerous open security issues that present clear and present danger.

Furthermore, we must take *mitigating controls* into account—the enterprise's investments into security controls like firewalls, anti-phishing systems, EDR, and so forth. We want credit for the points on the attack surface where the enterprise has successfully mitigated risk, provided the control is working.

Finally, not every point on the x -axis is equally important. We have to distinguish between critical or important assets and those which are not so important and estimate business impact of a compromised asset.

The resulting computation and risk heat map factoring in *vulnerabilities*, *exposure*, *threats*, *mitigating controls* and *business criticality* might look

Cybersecurity checklists lull you into a false sense of security.

like Figure 4.^b This risk calculation is not simple. To accurately understand an enterprise's security posture and breach risk, we need to (repeatedly) solve a hyper-dimensional math problem over the tens (or hundreds) of thousands of assets and 100+ attack vectors. Attacks need to be modeled as a chain of probabilistic events, starting from the compromise of some asset exposed to the Internet, followed by attack propagation to compromise additional valuable assets.

In order to improve cybersecurity posture and decrease breach risk, we must reason about what actions will bring about the greatest reduction of breach risk for the enterprise. This also requires calculating cyber-resilience—the ability of an enterprise to limit the impact of cyber-attacks.² *Analyzing and improving enterprise cybersecurity posture is not a human-scale problem anymore.* Plugging in some numbers into Figure 4, for an organization with a thousand employees, there are over 10 million time-varying signals that must be analyzed to accurately predict breach risk. For an organization with 100,000 employees, we must incorporate several 100 billion time-varying signals in the risk calculation.

Cybersecurity Practice Today

Traditional methods such as vulnerability assessment (for example, with Qualys, Rapid7, or Tenable) and penetration testing are only able to analyze a small fraction of the attack surface of Figure 1. These legacy methods produce output that is voluminous, unprioritized, and often irrelevant (for example, asking to patch an IE CVE on a laptop where the default browser is Chrome). Most organizations are un-

able to keep their systems patched and free of known vulnerabilities.

Due to a lack of a viable proactive strategy, much effort and money goes into detecting and reacting to cybersecurity events. This is the investment to set up and operate Security Information and Event Management systems (SIEMs) and Security Operations Centers (SOCs). Logs and alerts from enterprise systems are collected and analyzed for indicators of compromise. False positive are a huge challenge, and attackers slip through routinely. Recent studies peg the average dwell time of undetected attackers in the enterprise at approximately 200 days.⁵

Some security teams run cybersecurity checklists aligned against frameworks such as NIST 1.1,⁶ CIS 100, or SOC 2. This is an important step in the right direction, but inadequate. Compliance checklists don't work in cybersecurity because of the scale of the underlying math, our constantly changing software, and a dynamic adversary.^{1,3} In hundreds of conversations with CIOs and CISOs of the Fortune 1000 over the course of the last few years, it is clear to this author that the vast majority of organizations do not have an accurate (much less real-time) inventory of their assets—they cannot enumerate the x -axis of Figure 1. Moreover, many important security attributes of assets (y -axis in Figure 1), such as reused passwords, are omitted from the cybersecurity checklist because they are deemed too difficult to measure.⁴ More generally, individual security practitioners routinely make decisions to accept critical risk factors on their checklists because IT has not figured out how to mitigate these factors and still keep the environment operational. All this leads to a systematic buildup of risk that the CIO or CISO is not aware of. Cybersecurity checklists lull you into a false sense of security.

Ultimately, a poor understanding of the massive attack surface results in waste, frustration, and anxiety. Most discussions on cybersecurity posture and risk between the board of directors and C-suite execs are based on gut and incomplete data. Organizations are unable to answer simple questions such as “What is the risk to our intellectual property from cyber-attacks?”

^b Figure 4 was generated using real data from a Fortune 1000 customer of Balbix.

Distinguished Speakers Program

A great speaker can make the difference between a good event and a WOW event!

Students and faculty can take advantage of ACM's Distinguished Speakers Program to invite renowned thought leaders in academia, industry and government to deliver compelling and insightful talks on the most important topics in computing and IT today. ACM covers the cost of transportation for the speaker to travel to your event.

speakers.acm.org



Association for
Computing Machinery

New products are routinely launched without much thought about cybersecurity.

We spend a lot of money on security tools, but don't know what we are getting in return in terms of reduction in breach risk. New products are routinely launched without much thought about cybersecurity. In spite of millions of dollars of annual security spending, most enterprises are just one bad click, one reused password, or a single unpatched system away from a major breach.

Cybersecurity Education

It is useful to note the huge gap between the requirements for cybersecurity professionals who can understand and address the challenges of a practically unlimited attack surface (Figure 4), and the education and training being offered in university computer science programs. A recent study⁷ noted none of the top 10 CS undergraduate programs require a cybersecurity course in order to graduate. While a small (but increasing) number of professional master's degree cybersecurity programs are now being offered by top 40 CS departments, these tend to focus on a handful of basic elements of computer security, particularly crypto—and how to secure a small number of points on the attack surface of Figure 1 using existing tools. Some programs teach incidence response and forensics.

Our cybersecurity training programs do not consider cyber-insecurity as a networkwide (probabilistic) risk optimization problem. We are not teaching our future technologists how to design and create cyber-resilient distributed systems, or how machine learning, automation, and data visualization can serve as powerful tools to understand and mitigate cyber-risk.

Call to Action

As a discipline, CS must start thinking of cybersecurity as a probabilistic risk optimization problem. This author's organization has done some work on being able to discover and quantify cybersecurity posture in the spirit of this Viewpoint. We have developed a system that makes continuous observations of the extended enterprise from multiple vantage points including network, endpoint, configuration, and logs. This data is analyzed by an ensemble of machine learning models to surface inventory, business impact, breach likelihood, and cyber risk. The system also provides a prioritized set of possible mitigating actions to reduce risk along with simulation tools to estimate the pro-forma ROI of contemplated mitigating actions. Early experience with this system at many Fortune 1000 organizations tells a bittersweet story of both despair (red heatmaps) and hope (we can now measure, so we will improve).

Much research needs to be done on understanding the principles of cyber-resilient distributed systems design and feasibility of bolting-on cyber-resilience enhancing controls on top of legacy systems. The work done on developing zero-trust frameworks like BeyondCorp¹⁰ is a good beginning.

CS educators should reevaluate course curriculum in order to better prepare students for the cybersecurity realities highlighted in this Viewpoint. **□**

References

1. Bailey, K. Why compliance does not equal security; <http://bit.ly/2H3LymD>
2. Banga, G. Balbix Blog: What is cyber-resilience? (2017); <http://bit.ly/2Sr0XCD>
3. Banga, G. Cybersecurity 101 for the C-suite and board members; <http://bit.ly/39hA0YS>
4. Das, A. et al. The tangled web of password reuse. NDSS Symposium 2014; <http://bit.ly/3biv2o>
5. IBM. *Cost of a Data Breach Study by Ponemon* (2018); <https://ibm.co/374AK1Z>
6. NIST. Framework for Improving Critical Infrastructure Cybersecurity (2018); <http://bit.ly/3biv9Zm>
7. Syed, S. CloudPassage blog: U.S. universities get "F" for cybersecurity education (2016); <http://bit.ly/20BIkLr>
8. U.S. Government Accountability Office. Report to Congressional Requesters: Actions Taken by Equifax and Federal Agencies in Response to the 2017 Breach; <http://bit.ly/2Sv2DLH>
9. Verizon. *Data Breach Investigations Report* (2017); <https://vz.to/2Stn2R4>
10. Ward, R. and Beyer, B. BeyondCorp: A new approach to enterprise security. *login*: 39, 6 (June 2014); <http://bit.ly/20xo3GE>
11. Wikipedia. 2012 LinkedIn hack (2012), <http://bit.ly/2SqnLmf>

Gaurav Banga (gaurav@balbix.com) is the Founder and CEO of Balbix, Inc., in San Jose, CA, USA.

Copyright held by author.

Viewpoint

Organizing Family Support Services at ACM Conferences

Seeking to improve access to conferences and provide support for attendees with children.

CONFERENCES MATTER. THEY offer a one-stop forum for academics and industry professionals alike to communicate recent findings, meet new people, foster collaborations, maintain connections, and nurture a sense of community. Conference attendees who are also parents, however, face a notable barrier: how to attend and gain the aforementioned career benefits while also balancing childcare responsibilities—a challenge Calisi et al.² refer to as the “childcare-conference conundrum.” While all primary caretakers of children are affected, women often experience greater disadvantages due to multiple factors (for example, biological, cultural) further affecting their community participation and, ultimately, their careers. Indeed, recent work studying family formation on academic careers found that a “baby penalty” negatively affected women’s but not men’s career mobility, with a larger negative impact for women of color.⁴

To help address these issues, academic and industry conferences are increasingly offering some sort of family-support services (for example, on-site nursing rooms and/or childcare).¹ Here, we describe planning, organizing, running, and assessing family-support services at ACM CHI2018—a large (3,500 participants) multidisciplinary conference of researchers and practitioners in human-computer interaction and design; CHI2018 was located in Montréal, Canada.



Smaller attendees appreciated the colorful CHI2018 giant hashtag.

Our goals with this Viewpoint are to provide insights into the logistics, policies, and complexities of offering family services at CHI and to reflect on and offer guidelines and best practices for future ACM conferences.

History of Services Offered at CHI

Unlike many academic conferences, CHI has a rich, though inconsistent, history of offering family support services. In 1996, Allison Druin started the CHIkids program—a fun, interactive “CHI daycamp” based on Druin’s research in children-focused participatory design.³ To help run

CHIkids, Druin contracted a professional childcare service and invited members of the CHI community to volunteer as “CHIkids Leaders.” To help cover childcare costs and to provide computer equipment to the children, Druin sought out and received external sponsorships.

Because detailed records are not available for these early initiatives, it is not possible to comprehensively analyze demand, successful outcomes, challenges, and cost; however, an archived CHIkids page from 1999 indicates 38 children enrollees (ages six months to 14 years old). An additional

archived page from CHI2001 discusses two childcare options: one for \$85/day intended for six-month to six-year-olds and the other called “CHICamp” for \$130/day for 7–14-year-olds (which amounts to \$121/day and \$185/day in today’s dollars, respectively). Though CHIkids and its later incarnation CHICamp required a significant grassroots effort to organize and run, the program continued for nine years.

While some community-based initiatives occurred in the interim (such as social media groups to coordinate babysitters), CHI did not offer on-site childcare again until 2016 when Jofish Kaye and Druin co-chaired the conference. This reintroduction of services was, in part, a reaction to certain previous flashpoints in which children were unable to accompany a parent into a conference social function (for example, the reception), but also seen by Kaye and Druin as an important initiative to increase the inclusiveness of the CHI conference. To formalize the effort, they created a childcare chair position within the organizing committee. While a symbolic success in highlighting family inclusivity to the CHI community, childcare enrollment did not meet expectations (approximately 15 children participated). Consequently, childcare services were not provided at CHI2017 and no one was appointed to the role of family services chair. It is with this context that we began planning for CHI2018.

Planning Family Support Services at CHI2018

To plan for CHI2018, we began with a background survey soliciting feedback

about potential family support services and childcare options. We summarize the main findings here; see <http://bit.ly/2tDntjo> for a full report. In all, we had 95 respondents, including 66 faculty, 17 students, and 10 people from industry. Of these, 69.5% indicated being likely or very likely to use childcare services at CHI2018. Most respondents preferred on-site childcare (92%) followed by on-site shared nannies (62%) and independent babysitters/nannies (38%). They estimated needing childcare services for 108 children at CHI2018.

When asked about the maximum pay rate for a full day (eight hours) of childcare for one child (a closed-form question ranging from \$60 USD/day to \$140/day), the most common response was \$80/day (27.8% of respondents). Over 69% preferred \$100/day or less; however, 14% of respondents were willing to pay \$140 or more. When asked about who would pay for childcare—for example, the parent or the employer—72.2% responded paying themselves “out-of-pocket” and only 12% thought that they could get some or all of the childcare costs covered by their institution (14% were not sure).

Informed by this background survey, our own research into family support services at other conferences, and available childcare options in Montréal, Canada (the conference location), we met with conference leadership and discussed options. There were three main considerations:

► **Financial.** The general chairs allocated \$15,000 to subsidize on-site childcare, which was approved by the SIGCHI Steering Committee and ACM.

This was justified with the survey data described here, combined with our shared desire to provide an important service, particularly beneficial to women and women of color.⁴ Further, the childcare provider needed to be contracted before knowing how many children would use the service, requiring a financial commitment from the general chairs.

► **Space.** With up to 23 parallel tracks at CHI, it can be complex to find space for non-program activities, such as on-site childcare and a nursing room.

► **Liability.** ACM is required to pay insurance to cover any liability for anything that occurs on the premises of the conference. The addition of on-site family services introduces additional concerns, including: children accessing the site need to be approved and tracked, just like adults; local laws governing the presence of children at events with alcohol; and the potential harm that could come to children must be included in the liability policy, such as a child harming themselves with a toy provided by the conference.

Running Family Support Services

Based on parents’ needs voiced during the survey and organizational constraints, the conference and family chairs decided to support attendees with a multipronged approach: a child pass, on-site childcare services, and a nursing room. We hoped this would enable broad attendance for CHI2018.

Child Pass: We included a \$10 USD pass for children 0–18 to accompany their parents, both to provide access to the conference center and to make children feel welcome. It included a conference badge. We used the child pass registrations to communicate information to parents. Children were welcome to all on-site events, except for technical paper presentations as we thought that the potential disruption to speakers and audience members was greater than the benefit, given that all papers were livestreamed.

Childcare: We offered on-site childcare via KiddieCorp, a professional on-site childcare service. Children had a large common room with toys, snacks, and activities such as crafts, toys, and books. The Kiddi-



The chalkboard column was quite popular with children and adults in the CHI2018 exhibit hall.

eCorp team members are uniformed, screened, and experienced employees who have completed the KiddieCorp training program. The services were in English, for children 6 months to 12 years old.

Childcare was available during the technical sessions of the main conference (not during the workshops due to space constraints) at the cost of \$10 USD/hour. Parents registered for childcare in advance on the KiddieCorp site or onsite. Based on pre-conference registrations, we contracted the daycare services to care for up to 18 children simultaneously with four daycare providers for each block of time.

Nursing room: We provided a quiet, semiprivate nursing room for feeding and changing. The nursing room had comfortable sofas, a changing table, nursing chairs, a play mat, a kettle, and a fridge to store milk. Signs in the nursing room directed caregivers to online live streams of paper sessions to watch on their own devices.

Assessing Family Support Services at CHI2018

In total, family support services at CHI2018 enabled 61 children to accompany their parents at the conference with the conference badges, and 24 children from 17 families used the on-site childcare services. The service was full on the first two days of the conference and used at over 80% capacity for the last two. Many attendees commented on the higher proportion of children around the conference site, typically quite positively. This is supported by the general post-CHI survey—sent to all 3,372 attendees—in which 63% agreed the conference was family-friendly, with only 8% disagreeing.

To gain understanding of how childcare services at CHI were used and perceived, we designed a custom family services follow-up survey, which was advertised on social media and emailed directly to attendees who registered for a child pass and/or indicated interest for childcare at the conference. Our 62 respondents included 33 who identified as female, 24 as male, and 2 as nonbinary or gender nonconforming; 47 reported being parents. Of respondents who brought their children to the conference, 77% stated the family support services



The childcare room with the theme of “Camp CHI” at CHI2018.

made a difference in being able to attend the conference. The top three reasons indicated for bringing a child included: not having a good caregiving/childcare option at home (32%), enjoying travel with children (18%), and providing an enriching educational experience (18%).

Of parents that did not bring children, half reported considering it before ultimately deciding no. Their top three reasons included: travel costs (45%), preferring not to travel professionally with children (32%), or that the services were considered inappropriate for the child’s age (32%).

In general, parents were satisfied with the services provided: all parents (14 of 14) were satisfied or very satisfied with the onsite childcare, 83% (14 of 17) with the child pass, and 75% (3 of 4) with the nursing room. Similar to the general survey, 79% of respondents felt that children were welcome at the conference, with 7% indicating “not really.”

The survey also solicited open-ended feedback, which was largely positive, emphasizing the impact these services had on the conference: *Extremely happy with the on-site childcare program this year! This service made a HUGE difference to our travel decision and experi-*

ences this time; and Thank you because of this type of effort, I plan to attend CHI in the future!!! Most of the negative comments were about the high cost; one was about safety concerns.

Four common themes emerged in the suggestions: reducing cost, publishing more information prior to the conference to help with planning (for example, about the space or local child-related information), extending childcare hours to cover lunch and pre-conference workshops, and more conference-related activities for older children (five years old or older).

Reflections and Recommendations

I’m not a parent, nor do I intend to be, but [childcare] is a critical service to our community, especially for students, junior scholars, single parents, and others who may otherwise not be able to attend (and if they can’t attend, they can’t publish in the proceedings).

—Post-Conference Survey Response

Here, we reflect on our findings and experiences running family support services at CHI2018, most of which should be generally applicable to other academic conferences.

Get support. Financial, logistical,

and even philosophical support from conference organizers and community leadership is critical. Financial support includes daycare costs and sponsorship, logistical support involves finding daycare providers, booking rooms, handling registration as well as covering insurance and liability issues, and philosophical support helps align the community toward investing in these services.

Have Family Support Chairs. To demonstrate a commitment to inclusivity, general chairs should consider dedicated co-chairs to manage family support services, to assess the needs of the community, source options, and implement the vision for and logistics of providing support for families. Family co-chairs can also ensure the services are running smoothly during the conference.

On-site care should ideally be subsidized by general registration fees or external sponsors. The selected childcare service should be insured, provide well-vetted professionals, and offer care throughout the day (including lunch break) with 15–20-minute buffers before the start and end of the day's conference activities. Advanced sign up is recommended with a required partial deposit so the conference can plan and adjust to demand.

Advertise early and broadly about care services so that people can plan submissions and travel. Information should be available on conference websites with regular social media reminders. Facilitate a social network of parents so they can get in touch with each other, ask questions, coordinate common resources.

Support children of all ages. Because different age groups require different types of activities and caretaker-to-child ratios, carefully consider the following groups when planning your services: infants (0–1), toddlers (1–5) and older children (6+).

Provide a nursing room: Parents need a quiet, semi-private room to feed their children, which can also be used for babies napping or general rest. We recommend including a mini-fridge to store milk, a bottle warmer or a kettle, and a changing pad. If possible, consider providing a live video feed of conference sessions.

Welcome children to events. Parents will appreciate the opportunity to bring their children to social events, demonstrations, and exhibit halls. Organizers should explicitly check policies of offsite venues to ensure that children are welcomed at official conference events. If not, we recommend announcing in advance when events can only be attended by adults, so parents can plan accordingly.

Think about diversity. While most of the CHI attendees that used childcare were faculty, this may have been due to costs. Communities should think about how to better support a more diverse set of parents, particularly student (for example, via subsidized childcare costs).

Offer registration discounts for children and caretakers (for example, spouses, grandparents), to allow families to attend certain conference events together. Badges can include insurance coverage (under the conference's insurance policy).

PC meeting support. Consider offering childcare support at an in-person program committee (PC) or other organizational meetings. As a survey respondent stated: *For me, the PC meeting is far more of a problem as a parent than CHI itself... I've declined PC committee invites for [multiple years] now. Several other parents amongst my colleagues are in the same situation.*

Give it time. Organizing family services is an ongoing process. Because of cultural norms, shifting expectations, and the need to plan conference submissions and travel far in advance, we estimate three to four years of continuous family support offerings are necessary before we can truly study and understand community impact.

Communicate to non-caregiving attendees. Announce to all attendees where children are welcome. If in paper sessions, communicate this to presenters in advance to ensure they are not taken by surprise. Further, exhibitors, demonstrators, and other presenters must be made aware that people under 18 years old may be present.

Donations. We recommend conferences donate all the family support services artifacts bought for the conference to a local women's shelter.

Toward the Future

The ACM SIGCHI community aims to be inclusive and diverse. SIGCHI conferences are taking steps to increase the participation and success of underrepresented groups in HCI. The advice page on organizing a SIGCHI sponsored conference includes a paragraph on the need to develop a policy regarding children at conferences.^a We encourage other special interest groups to adopt similar policy and encourage their members to consider how to improve access to conferences or include supporting attendees with families.

This conference attendee summarized clearly the need for these services: *THIS. It is so needed. I am terrified this will be a "one off" (even having a nursing room has varied from year to year). Programs like these make parenting in the community visible and send the right message about participation from primary caregivers who also happen to be HCI researchers. Oh, and I've also made professional connections I otherwise wouldn't have, because we brought our kids and wanted to connect as CHI parents to swap tips AND talk about research!* ■

a See <http://bit.ly/2Uy90oN>

References

- Bos, A.L., Sweet-Cushman, J., and Schneider, M.C. 2017. Family-friendly academic conferences: A missing link to fix the "leaky pipeline"? *Politics, Groups, and Identities*, 1–11; <http://bit.ly/2UAtSHK>
- Calisi, R.M. and a Working Group of Mothers in a Working Group of Mothers in Science. 2018. Opinion: How to tackle the childcare-conference conundrum. In *Proceedings of the National Academy of Sciences of the United States of America* 115, 12 (2018), 2845–2849; <http://bit.ly/2H6pvLY>
- Druin, A. Cooperative inquiry: Developing new technologies for children with children. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (1999), 592–599; <http://bit.ly/3bjMXmN>
- Mason, M.A. The baby penalty. *The Chronicle of Higher Education*. (2013); <http://bit.ly/2vanMTP>

Audrey Girouard (audrey.girouard@carleton.ca) the CHI 2018 Family Co-Chair, is an associate professor at Carleton University in Ottawa, Ontario, Canada.

Jon E. Froehlich (jonf@cs.washington.edu) the CHI 2018 Family Co-Chair, is an associate professor at the University of Washington, in Seattle, WA, USA.

Regan L. Mandryk (regan@cs.usask.ca) the CHI 2018 General Co-Chair, is a professor at the University of Saskatchewan, in Saskatoon, Saskatchewan, Canada.

Mark Hancock (mark.hancock@uwaterloo.ca) the CHI 2018 General Co-Chair, is an associate professor at the University of Waterloo, Waterloo, Ontario, Canada.

Viewpoint

A Taxonomy of Automated Assistants

Rating your intelligent (human or automated) assistant.

AUTOMATED CARS ARE in our future—and starting to be in our present. In 2014, the Society of Automotive Engineers (SAE) published the first version of a taxonomy for degree of automation in vehicles from Level 0 (not automated) to Level 5 (fully automated, no human intervention necessary).⁸ Since then, this taxonomy has gained wide acceptance—to the point where everyone from the U.S. government (used by the NHTSA⁵) to auto manufacturers to the popular press are talking in terms of “skipping level 3” or “everyone wants a level 5 car.”¹ As technology gets developed and improved, having an accepted taxonomy helps ensure people can talk to each other and know they are talking about the same thing. It is time for one of our computing organizations (perhaps ACM?) to develop an analogous taxonomy for automated assistants. With Siri, Alexa, Cortana, and cohorts selling in the “tens of millions”² and with more than 20 competitors on the market,⁷ having an easily understandable taxonomy will help practitioners and end users alike.

There is already a significant body of literature aimed at improving the design and use of automated assistants in both industry and academic arenas (with a variety of category names for these devices and systems, using some combination of “automated,” “digital,” “smart,” “intelligent,” “personal,” “agent,” and “assistant”), as the bibliographies of cited works



show. Some recent work focused on task content, use cases, and features. The task content of human activity has been widely studied over a long period of time, but Trippas et al.⁹ note that “how intelligent assistants are used in a workplace setting is less studied and not very well understood.” While not presenting a taxonomy of assistants, this type of task content analysis could be used as an aid in intelligent assistant design. Similarly, Mehrotra et al.⁴ studied interaction with a desktop-based digital assistant with an eye to “help guide develop-

ment of future user support systems and improve evaluations of current assistants.” Knote et al.³ evaluated 115 “smart personal assistants” by literature and website review to create a taxonomy based on cluster analysis of design characteristics such as communications mode, direction of interaction, adaptivity, and embodiment (virtual character, voice), and so forth—a technology and features-based taxonomy. A commercial study of 22 popular “intelligent ... or automated personal assistants”⁷ reported “Intelligent Agents can be classified

based on their degree of perceived intelligence and capability such as simple reflex agents, model-based reflex agents, goal-based agents, utility-based agents and learning agents.” While this is an arguably useful taxonomy, it also primarily addresses the technology used and not the actual use of the automated assistant. The website additionally presents editor and user ratings of ease of use, features, and performance that may be of value to end users.

The taxonomy suggested here focuses on the end-user view of “work output,” and while this approach uses a somewhat subjective measurement scale, further work might incorporate objective data such as is suggested in the references. The U.S. Department of Labor has commissioned work that provides detailed analysis of the job content, knowledge, and skills required of human assistants at various levels that could be used in further refining this taxonomy.⁶ Furthermore, the current taxonomy does not address the time an assistant might spend in performing a task, another factor that might be used in expanding a classification scheme.

Work-Output Capability Taxonomy

The work-output or “capability” taxonomy for assistants I have long used is based on observation of the skills and experience of human assistants over the past 40 years. Today’s administrative or personal assistants (the human kind) perform a wide range of functions, albeit with highly varying levels of accuracy, knowledge, skill, enthusiasm, and initiative. The best are professionals with superior skills who genuinely earn their titles—they provide highly valuable (and valued) assistance to the people they work with, leveraging their abilities in pursuit of the goals of the organization that employs them. These people should (and mostly do) enjoy all the kudos, benefits, and satisfaction that comes from being a professional recognized for excellent work.

Experience working with assistants of all ranks and skills has led me to want to expand the ranks of the best, whether in the future they will be human, automated, or human-augment-

ed-with-automation. Anyone who has worked with an assistant knows that if the assistant is not very good (for example, produces sloppy or inaccurate work, or takes longer than the expected or allotted time), the person or device is more often going to be a source of frustration and annoyance than assistance.

This assistant capability scale, while initially designed to rate human assistants, can readily form the basis for an intelligent automated assistant scale. As described in the examples below, it ranges from Level 1 (Entry Level Assistant) to Level 5 (Super Assistant). The key work-output characteristics of each level reflect an integration of skills and experience as follows:

Level 1: Work-output based on passively performing specifically assigned tasks;

Level 2: Work-output based on actively performing assigned tasks, developing related sub-tasks;

Level 3: Work-output based on using basic general knowledge and experience to understand specifically assigned duties and perform readily discerned tasks;

Level 4: Work-output based on using broad knowledge and experience, general and in the task area, to understand broadly assigned duties and perform implied tasks; and

Level 5: Work-output drawing on all available knowledge and experience from a variety of sources, general and in the task area, to infer useful duties, executed without supervision—just like you would have done them if you had the time, or even better!

Adapting this to automated assis-

The taxonomy suggested here focuses on the end-user view of “work output.”

ants, the levels could be linked to the human assistant scale: Level 1 “performs like an entry level assistant” to Level 5 “performs like a super assistant.” Or, they could follow the SAE approach and be more descriptive of the level of automation versus human intervention required: Level 0 “no automation,” Level 1 “requires significant human input, supervision and review,” Level 2 “requires some human input, supervision and review,” Level 3 “requires some human input and review,” Level 4 as “requires minimal human input and some review” and Level 5 as “requires minimal human input and review.” Admittedly these are subjective names, but examples help clarify, and they can still be related to the human assistant capability scale.

Work-Output Examples

It may be easiest to understand the taxonomy by example. Several are presented here representing typical tasks for an administrative assistant. As the taxonomy should encompass a broad range of uses of intelligent assistants and other kinds of tasks, the listed examples should be taken only as illustrative of the work-output at each level of capability.

Example 1: Preparing for a one-on-one meeting.

Level 1: An entry level assistant gets a call from your boss for a meeting and puts it on your calendar.

Level 2: An assistant gets the call, puts it on your calendar, and lets you know when your boss wants to see you.

Level 3: A good assistant finds out what the topic is and tells you that too.

Level 4: A really good assistant asks you what materials you need, assembles them, puts them in a folder, and gives them to you in time to review for the meeting.

Level 5: A super assistant finds most of the materials based on learning what the meeting is about, and calls your attention to what additional information you might also want to prepare.

Example 2: Preparing for a conference call.

Level 1: An entry level assistant sees that you put a conference call on your calendar.

Level 2: An assistant sees that you didn't note the phone number and password and asks you about them.

Level 3: A good assistant asks who the other attendees are and notes that on your calendar too.

Level 4: A really good assistant asks about attendees and topics, what information you might want distributed in advance, whether you want a reminder sent out, and so forth. [Overreach at this level, which should not happen at Level 5, would be calling the conference call organizer to request a change in agenda without first consulting with you.]

Level 5: A super assistant figures all this out based on your past behavior, on the title of the conference call, on the names of the other attendees, and just does it!

Example 3: Preparing for a candidate interview.

Level 1: An entry level assistant schedules a candidate interview and puts it on your calendar.

Level 2: An assistant puts the person's résumé in a folder and gives it to you before the interview.

Level 3: A good assistant gives you this information in enough time for you to review before the candidate comes in.

Level 4: A really good assistant looks up additional information about the candidate on the Web.

Level 5: A super assistant reads all this information and highlights interesting points for your attention.

Example 4: Arranging business travel.

Level 1: An entry level assistant puts an out-of-town meeting on your calendar.

Level 2: An assistant makes your travel arrangements as per your instructions.

Level 3: A good assistant looks up travel alternatives and brings them to you, and then makes your travel arrangements according to your instructions.

Level 4: A really good assistant validates your trip schedule, makes sure you can get from one place to another, arranges cars and pickups, and goes over it with you several days before your trip, in time to make changes if necessary. [Overreach at this level, which shouldn't happen at Level 5, would be booking dinner

The key work-output characteristics of each level reflect an integration of skills and experience.

at an expensive restaurant and box seats at a Broadway play.]

Level 5: A super assistant does all this based on your past trips, adds maps of the areas you are visiting, maps showing buildings you are going to (including any security arrangements), any other complex instructions, information on local sites that might be of interest, and so forth.

Example 5: Arranging a large meeting or event.

Level 1: An entry level assistant is told, step-by-step, what to do to plan a meeting you are hosting for colleagues from multiple locations, and requires that you follow up to ensure these things are done in the way you want.

Level 2: An assistant is told the general outlines of the event and the tasks to be done in preparation, and is able to follow through with most, reporting back to you.


Level 3: A good assistant discusses the general outlines of the event, comes up with the tasks, and reports back to you on any issues.

Level 4: A really good assistant does that and comes up with suggestions on how to deal with the issues.

Level 5: A super assistant suggests to you what needs to be done to have a really great meeting—and then does it all!

Conclusion

The human assistant scale presented in this Viewpoint can be readily (although subjectively) applied to intelligent automated assistants to help developers (and perhaps the systems themselves) improve their capabilities. I have used this capability scale to help human assistants understand the

kind of things they ought to be working on to improve their capabilities. It (or I) has not uniformly succeeded in that regard, although it should not be expected that assistants will necessarily perform at the same level for all types of tasks. The question is whether your assistant (human, automated, or human augmented with automation) is performing at Level 1 now, but can the assistant perform at Level 5 with some coaching. Perhaps Level 2 on some tasks and Level 4 on others? How do Alexa, Siri, and Google Assistant rate on various types of tasks—Level 1 on some, Level 5 on others? Will our future assistants be all digital, or will the super assistants of the future be the human ones who figure out how best to augment their skills with their own digital assistants? As for now, I am betting on the latter—at least until AI makes further advances into the realm of adding “the human touch.” Either way, fairly soon everyone will want to skip Level 3 and have a Level 5 intelligent assistant. 

References

1. Davies, A. Everyone wants a Level 5 self-driving car—Here's what that means. *Wired Magazine* (Aug. 26, 2016); <http://bit.ly/31xX5EJ>.
2. Kilgore, T. Amazon to 'double down' on Alexa as sales far exceeded expectations. *MarketWatch*; <https://on.mktw.net/2Stc0JA>.
3. Knote, R. et al. Classifying smart personal assistants: An empirical cluster analysis. In *Proceedings of the 52nd Hawaii International Conference on System Sciences 2019*, 2019; <http://bit.ly/31HvYqt>.
4. Mehrotra, R. et al. Hey Cortana! Exploring the use cases of a desktop based digital assistant. In *Proceedings of the First International Workshop on Conversational Approaches to Information Retrieval (CAIR'17)*, Tokyo, Japan; <http://bit.ly/2S78ZIs>.
5. National Highway Traffic Safety Administration. Automated vehicles for safety; <http://bit.ly/2ups21a>.
6. O*NET OnLine. Summary Report for Executive Secretaries and Executive Administrative Assistants; <http://bit.ly/2GZNt5v>.
7. Predictive Analytics Today. Top 22 intelligent personal assistants or automated personal assistants; <http://bit.ly/2UwiTi7>.
8. SAE International. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016_201609; <http://bit.ly/2vdMG4A>.
9. Trippas, J. et al. Learning about work tasks to inform intelligent assistant design. In *Proceedings of the Conference on Human Information Interaction and Retrieval (CHIIR '19)*, (Mar. 10–14, 2019, Glasgow, Scotland, U.K.); <http://bit.ly/385LKgX>.

Jerrod M. Grochow (jgrochow@mit.edu) is a research affiliate with the interdisciplinary cybersecurity research consortium at MIT Sloan School of Management (<https://cams.mit.edu>) in Cambridge, MA, USA. He is retired from a career in IT management in both industry and academia.

The author wishes to thank the reviewers and the Associate Editor for their thoughtful comments that led to my refining this Viewpoint.

Copyright held by author.

ACM ON A MISSION TO SOLVE TOMORROW.

Dear Colleague,

Without computing professionals like you, the world might not know the modern operating system, digital cryptography, or smartphone technology to name an obvious few.

For over 70 years, ACM has helped computing professionals be their most creative, connect to peers, and see what's next, and inspired them to advance the profession and make a positive impact.

We believe in constantly redefining what computing can and should do.

ACM offers the resources, access and tools to invent the future. No one has a larger global network of professional peers. No one has more exclusive content. No one presents more forward-looking events. Or confers more prestigious awards. Or provides a more comprehensive learning center.

Here are just some of the ways ACM Membership will support your professional growth and keep you informed of emerging trends and technologies:

- Subscription to ACM's flagship publication ***Communications of the ACM***
- Online books, courses, and videos through the **ACM Learning Center**
- Discounts on registration fees to ACM Special Interest Group conferences
- Subscription savings on specialty magazines and research journals
- The opportunity to subscribe to the **ACM Digital Library**, the world's largest and most respected computing resource

Joining ACM means you dare to be the best computing professional you can be. It means you believe in advancing the computing profession as a force for good. And it means joining your peers in your commitment to solving tomorrow's challenges.

Sincerely,



Cherri M. Pancake
President
Association for Computing Machinery



Association for
Computing Machinery

Advancing Computing as a Science & Profession

SHAPE THE FUTURE OF COMPUTING. JOIN ACM TODAY.

www.acm.org/join/CAPP

SELECT ONE MEMBERSHIP OPTION

ACM PROFESSIONAL MEMBERSHIP:

- Professional Membership: \$99 USD
- Professional Membership plus ACM Digital Library: \$198 USD (\$99 dues + \$99 DL)

ACM STUDENT MEMBERSHIP:

- Student Membership: \$19 USD
- Student Membership plus ACM Digital Library: \$42 USD
- Student Membership plus Print *CACM* Magazine: \$42 USD
- Student Membership with ACM Digital Library plus Print *CACM* Magazine: \$62 USD

- Join ACM-W:** ACM-W supports, celebrates, and advocates internationally for the full engagement of women in computing. Membership in ACM-W is open to all ACM members and is free of charge.

PAYMENT INFORMATION

Name _____

Mailing Address _____

City/State/Province _____

ZIP/Postal Code/Country _____

- Please do not release my postal address to third parties

Email Address _____

- Yes, please send me ACM Announcements via email
- No, please do not send me ACM Announcements via email

- AMEX VISA/MasterCard Check/money order

Credit Card # _____

Exp. Date _____

Signature _____

Purposes of ACM

ACM is dedicated to:

- 1) Advancing the art, science, engineering, and application of information technology
- 2) Fostering the open interchange of information to serve both professionals and the public
- 3) Promoting the highest professional and ethics standards

By joining ACM, I agree to abide by ACM's Code of Ethics (www.acm.org/code-of-ethics) and ACM's Policy Against Harassment (www.acm.org/about-acm/policy-against-harassment).

I acknowledge ACM's Policy Against Harassment and agree that behavior such as the following will constitute grounds for actions against me:

- Abusive action directed at an individual, such as threats, intimidation, or bullying
- Racism, homophobia, or other behavior that discriminates against a group or class of people
- Sexual harassment of any kind, such as unwelcome sexual advances or words/actions of a sexual nature

BE CREATIVE. STAY CONNECTED. KEEP INVENTING.



ACM General Post Office
P.O. Box 30777
New York, NY 10087-0777

1-800-342-6626 (US & Canada)
1-212-626-0500 (Global)
Hours: 8:30AM - 4:30PM (US EST)

Fax: 212-944-1318
acmhelp@acm.org
www.acm.org/join/CAPP

East Asia & Oceania Region Special Section



ILLUSTRATION BY SPOOKY POOKA AT DEBUT ART.
FOR CREDITS ON IMAGES IN COLLAGE, SEE P. 3.

Welcome

WELCOME TO THE special section covering East Asia and Oceania. Our region includes Southeast Asia, Oceania, and Asia-Pacific countries including Japan and Korea. The articles within this section—noted as Hot Topics and Big Trends—highlight the research and innovation emerging from the region as well as helps to strengthen the research collaboration and communication with various regions of the world.

The inventive minds of the researchers and practitioners in the region have put computing technology to great use as illustrated in diverse applications ranging from preserving cultural heritage to services designed to enhance the digital economy. These themes constitute some of the subjects explored in the Hot Topics section.

Big Trends in the region range from research trends to trends involving significant investments by local governments to support specific disciplines. These trends include technological advances such as advances in 5G, research advances in program analysis and trustworthy computing, and government initiatives in artificial intelligence and healthcare.

No section spotlighting this part of the world would be complete without a discussion of the collaboration and engagement efforts across the region. A workshop, held at the National University of Singapore in August 2019, strengthened our efforts in this direction. We discovered several existing conferences, such as AsiaCrypt, that build regional research networks. We thus present an article narrating experiences from the AsiaCrypt community. In addition, we present an article describing the Shonan meetings—a significant initiative from Japan’s National Institute of Informatics (NII) that provides a forum for research interaction across topics in computer science.

We hope this special section serves to encourage *Communications* readers to engage more with researchers and research efforts in our region.

—*Sue Moon, Ann Nicholson, and Abhik Roychoudhury*
Co-organizers of East Asia and Oceania Region Special Section

Sue Moon (sbmoon@kaist.edu) is Chair Professor at KAIST in the Republic of Korea.

Ann Nicholson (Ann.Nicholson@monash.edu) is a Professor and Deputy Dean Research in the Faculty of Information Technology at Monash University, Australia.

Abhik Roychoudhury (abhik@comp.nus.edu.sg) is Provost’s Chair Professor at the National University of Singapore.

Copyright held by owners/authors.



First row from left to right: Ping-Hai Hsu, Lihan Chen, Vivy Suhendra, Thomas Ho, Karen Teh, Abhik Roychoudhury, Ann Nicholson, Lam Kwok Yan, Sue Moon, Sally Cunningham. Second row from left to right: Jing Ma, Jakob Rehof, Sriram Rajamani, Lee Ching Yi, Fariz Darari, Ariel Liebman, Carsten Rudolph, Raphael Phan, Chris Bain, Ben Leong, Campbell Wilson, and Haibo Chen.

EDITORIAL BOARD

EDITOR-IN-CHIEF

Andrew A. Chien
eic@cacm.acm.org

DEPUTY TO THE EDITOR-IN-CHIEF

Lihan Chen
Morgan Denlow
cacm.deputy.to.eic@gmail.com

CO-CHAIRS, REGIONAL SPECIAL SECTIONS

Sriram Rajamani
Jakob Rehof
Haibo Chen

SPECIAL SECTION CO-ORGANIZERS

Sue Moon
KAIST
Ann Nicholson
Monash University
Abhik Roychoudhury
National University of Singapore



Watch the co-organizers discuss this section in the exclusive *Communications* video.
<https://cacm.acm.org/videos/east-asia-and-oceania-region>

Hot Topics



53

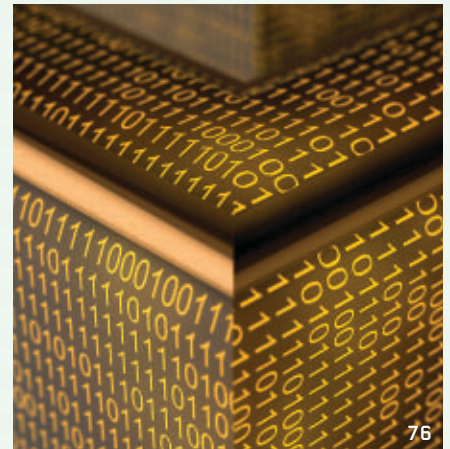
- 48 **The NII Shonan Meeting in Japan**
By Ken-ichi Kawarabayashi
-
- 50 **Capturing Cultural Heritage in East Asia and Oceania**
By Sally Jo Cunningham, Fariz Darari, Adila Krisnadhi, and Annika Hinze
-
- 53 **Cybersecurity in Pacific Island Nations**
By Carsten Rudolph, Sadie Creese, and Sameer Sharma
-
- 55 **Singapore's Cybersecurity Ecosystem**
By Karen Teh, Vivvy Suhendra, Soon Chia Lim, and Abhik Roychoudhury
-
- 58 **Innovating Services and Digital Economy in Singapore**
By Thomas Ho Chee Tat and George Loh Chee Ping

Big Trends



64

- 60 **AI Singapore: Empowering a Smart Nation**
By Sintia Teddy-Ang and Abigail Toh
-
- 64 **Digital Healthcare Across Oceania**
By Chris Bain and Abraham Oshni Alvandi
-
- 68 **Detecting Fake News in Social Media: An Asia-Pacific Perspective**
By Meeyoung Cha, Wei Gao, and Cheng-Te Li
-
- 72 **seL4 in Australia: From Research to Real-World Trustworthy Systems**
By Gernot Heiser, Gerwin Klein, and June Andronick



76

- 76 **Advances in Security Research in the Asiacrypt Region**
By Raphaël CW Phan, Masayuki Abe, Lynn Batten, Jung Hee Cheon, Ed Dawson, Steven Galbraith, Jian Guo, Lucas Hui, Kwangjo Kim, Xuejia Lai, Dong Hoon Lee, Mitsuru Matsui, Tsutomu Matsumoto, Shiho Moriai, Phong Nguyen, Dingyi Pei, Duong Hieu Phan, Josef Pieprzyk, Huaxiong Wang, Hank Wolfe, Duncan Wong, Tzong-Chen Wu, Bo-Yin Yang, Siu-Ming Yiu, Yu Yu, and Jianying Zhou
-
- 82 **5G Commercialization and Trials in Korea**
By Dong Ku Kim, HyeonWoo Lee, Seong-Choon Lee, and Sunwoo Lee
-
- 86 **Asia's Surging Interest in Binary Analysis**
By Sang Kil Cha and Zhenkai Liang

Profession | DOI:10.1145/3378546

The NII Shonan Meeting in Japan

BY KEN-ICHI KAWARABAYASHI

JAPAN'S NATIONAL INSTITUTE of Informatics (NII) launched its inaugural NII Shonan Meeting in February 2011. It was the first international conference of informatics in Asia, following in the style of the Dagstuhl seminars in Germany, designed to bring together the world's leading researchers and engineers to discuss open problems and challenges. More than 140 meetings have been held since then, and the number of participants totaled approxi-

mately 3,500 by November 2019. NII supports all the administrative arrangements for organizers and covers approximately half the fee for every academic participant (including room, board, and meeting fees). We currently hold 20–25 workshops/year. Sometimes, we also have summer/winter schools.

A recent trend in computer science conferences is to host “mega” conferences. For example, both the NeurIPS and CVPR events have attracted thousands of participants. There are many workshops

co-located with these mega conferences, but because they usually last a day or less it is very difficult for participants to interact with each other. This prompts the research community to have smaller workshops that last at least a few days. Dagstuhl is the most successful of this kind and an excellent model for the NII Shonan Meeting.

A typical NII Shonan event focuses on smaller but emerging areas of informatics with around 25 participants. Each meeting invites experienced and young promising researchers in academia and industry for interdisciplinary discussions. As of November 2019, participants have represented 60 countries and regions: 38% from Asia, 37% from Europe, 20% from North America and South America, 4% from Oceania, and less than 1% from Africa. In the last few years, about 13% of the participants are female researchers.

In such international surroundings, the NII Shonan Meeting encourages close communication among participants. The venue—Shonan Village Center—offers a conducive environment for such gatherings, located far from downtown on a hilltop with lush greenery. Each meeting lasts 4–5 days, including a half-day excursion. By the end of every workshop, each participant has learned the research interests and expertise of all fellow attendees through days of discussions and friendly interactions.

To hold a Shonan Meeting, organizers must submit a proposal to the NII's academic committee. Experts exam the documents and organizers are notified about the results. Although similar to the Dagstuhl's review process, we often request proposal revisions, taking reviewers' comments into account. Sometimes, organizers are

A typical NII Shonan event focuses on smaller but emerging areas of informatics with around 25 participants.



Every NII Shonan Meeting takes place at the scenic Shonan Village Center in Japan.

asked to make major changes to their proposals; in fact, on occasion they are asked for multiple revisions.

The computing areas covered by the NII Shonan Meet are very diverse. Indeed, in the last two years, software, theory, programming languages, and architecture (in a broad sense) have been the major focus, but they each occupy at most 20%. There are several workshops in the area of natural language processing, machine learning, databases, and security. This is a bit different from the Dagstuhl model in that approximately 30% of those workshops focus on theory (in a broad sense), followed by programming languages (with around 15%). Moreover, a Dagstuhl seminar typically hosts 40 researchers, while a NII Shonan Meeting hosts 25 researchers on average. It appears the NII Shonan Meeting focuses

on more specific areas while Dagstuhl's scope seems broader.

Remarkably, an increasing number of frequent participants have been creating research communities centered on their particular research topics. One example is a community on visualization formed after the No.046 seminar,^a that now organizes new workshops almost every year. Another group on engineering adaptive systems (EASy) has also been reuniting after No.004.^b

An important aspect of the NII Shonan Meeting is that collaborations have enabled researchers to publish books and papers and to obtain research grants. Researchers from No.120^c issued a paper on

a No.046; <https://shonan.nii.ac.jp/seminars/046/>

b No.004; <https://shonan.nii.ac.jp/seminars/004/>


c No.120; <https://shonan.nii.ac.jp/seminars/120/>

An important aspect of the NII Shonan Meeting is that collaborations have enabled researchers to publish books and papers and to obtain research grants.

visual analytics and another group from No.074^d published a book on immersive analytics. Springer has published three monographs as a book series on Shonan workshops written by participants, one of which was written by EASy researchers. In addition, some researchers received research funding from NSF, SNSF, and JSPS, submitting proposals discussed during the

d No.074; <https://shonan.nii.ac.jp/seminars/074/>

Shonan seminars.

Thus, the NII Shonan Meeting has contributed to research achievements by giving participants the opportunity to access broad ideas on specific topics. One of the missions of the NII Shonan Meeting is to welcome more participants from various backgrounds to this significant destination for informatics researchers. 

Ken-ichi Kawarabayashi is a professor and Deputy Director General at the National Institute of Informatics in Tokyo, Japan.

© 2020 ACM 0001-0782/20/4 \$15.00

Capturing Cultural Heritage in East Asia and Oceania

BY SALLY JO CUNNINGHAM, FARIZ DARARI, ADILA KRISNADHI, AND ANNIKA HINZE

TO CAPTURE CULTURAL heritage is to capture the experience of people who are directly involved in creating, witnessing, and maintaining cultural heritage objects. Ideally, the people accessing digital representations of cultural heritage objects are able to understand the significance underlying the objects. The question is how to capture (the experience of) cultural objects in digital form. Various modalities exist for representing cultural heritage: unstructured textual data, possibly including images or videos, as well as structured data.

To illustrate our approaches, we pick one cultural heritage representative: *rendang*, one of the five national dishes of Indonesia, believed to have existed as early as the 15th century. From textual sources, we may learn

about rendang and its history. According to Nurmufida et al.,⁸ rendang is a traditional cuisine originating from West Sumatra, with beef and coconut milk as its main ingredients. From image sources, we may see what rendang looks like. For example, Wikimedia Commons contains images of rendang that show that, despite being similar to curry, rendang is actually drier. Next, we may wonder how to cook rendang. A simple YouTube search provides a wide selection of videos showing how to cook rendang, with the chefs ranging from local Indonesians to international chefs. In fact, the word rendang originates from how it is cooked; that is, slowly (*merandang*, in the Minangnese language).

Textual Data

Much information about heritage is already avail-



able in traditional books. Recipe variations for rendang are published in many cookbooks, and its cultural significance can be traced through the centuries in novels, newspapers, and historic books. Many of these can now be accessed through their digitized versions, for example, via the HathiTrust Digital Library (www.hathitrust.org). A quick search for rendang reveals cookbooks, introductions to the Indonesian language, and to Javanese culture. However, not all search results refer to the dish; for example, Raffles¹⁰ refers to “Mangsa rendang” as the season of rain, and Rendang is also a district

in Bali. The same search will also miss references to *kalio*, a wetter version of rendang. Potential resources in other languages (using spelling variations) may also be lost. Thus, false positives (wrong hits) and false negatives (omitted hits) may occur in the search results.

Semantic text analysis is one approach to make such hidden information accessible and to help avoid irrelevant search results. The Capisco system⁶ avoids the need for complete semantic text markup by using an automatically generated Concept-in-Context (CiC) network. The network is

Capisco has been shown to provide quality semantic search results for English-language texts, with promising early results for other languages.

seeded by semantic concepts and their context of use as identified from Wikipedia texts. When doing a semantic search for rendang in Capisco, the user would specify they are interested in the dish; consequently, only those digital sources that contain words in the context of cooking and eating are

search results for English-language texts, with promising early results for other languages. The support for cross-language semantic search and multilingual texts (particularly relevant in bicultural New Zealand) is currently being investigated.

In addition to seeding the CiC network via

challenging for any individual nation. The lack of access to records and documents is severely limiting, and many scholars in Pacific nations must manually build their own collections to support their research. The preservation of heritage information in digital libraries is a recurring

cultural heritage information can also be consumed in a machine-friendly way, realized through structured data. One of the local initiatives to capture structured data about cultural heritage artifacts is by Putra et al.,⁹ who developed BudayaKB, a knowledge base storing Indonesian cultural heritage metadata. BudayaKB extracts entities of cultural heritage from Indonesian Wikipedia and presents the types and locations of the entities using machine-readable RDF triples.¹² Through BudayaKB, applications for cultural heritage can be developed easily, as the data is captured into a structured form ready to be queried. BudayaKB contains around 3,200 cultural heritage entities; a third of those are about food. Data in BudayaKB also is linked to Wikidata, a crowd-sourced knowledge-base hub. Information about rendang can be found in BudayaKB, as a type of traditional food, coming from West Sumatra. Figure 1 shows a SPARQL query,¹¹ asking for traditional foods of provinces in Indonesia.

The query results, filtered to focus on Sumatera Barat (West Sumatra), show not only rendang, but also other traditional foods from the same region like



selected. Texts about the Rendang district in Bali, as well as the rain season, would be excluded from search results (as they do not refer to the dish). In the Capisco CiC network, both rendang and kalio are semantic concepts that are flagged as potentially synonymous in the context of Indonesian cooking. Thus, the search results would not only contain all texts referring to rendang recipes, but also those containing kalio recipes. Capisco's semantic index can be exported to be used as data enrichment in existing digital libraries.⁵ Capisco has been shown to provide quality semantic

Wikipedia, Capisco also allows scholars to develop and refine their own set of relevant concepts. The system is unique in that the scholar is supported from initial exploration of digitized documents through to the creation of a publicly accessible collection.⁴ Particularly relevant for the Asia-Pacific region are Capisco's use cases of heritage collections and digital repatriation. While rich collections of historic documents exist for many South Pacific Island nations, the identification of these widely scattered documents and their compilation into coherent collections is

theme in the ICADL series of Asia-Pacific Conferences on Digital Libraries.

Structured Data

In addition to human consumption, it is of particular importance that

Through BudayaKB, applications for cultural heritage can be developed easily, as the data is captured into a structured form ready to be queried.

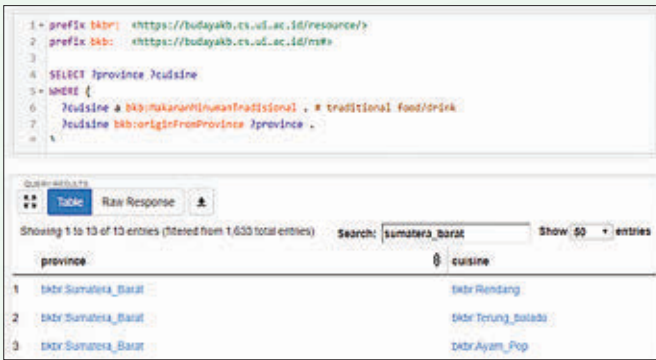


Figure 1. Cultural heritage in structured data enables querying.

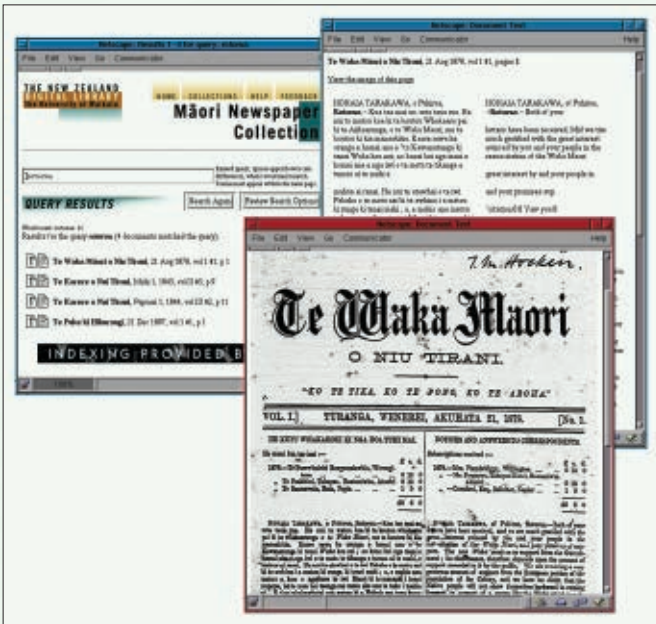


Figure 2. Niupepa collection of Māori newspapers presented in Greenstone.

terung balado (eggplants with chilies) and ayam pop (steamed chicken). From the links to Wikidata, one can learn that rendang is not only a beef dish, but also has variations with chicken, fish, and lamb. These examples demonstrate how structured data may facilitate knowledge discovery in cultural heritage.

Curation Support

The main challenge for effective and sustainable cultural heritage preservation is the development of tools to support domain experts in curating collections of cultural heritage information, rather than hand-


crafting code for creating and maintaining specific digital heritage objects. One such tool is the Greenstone digital library system that allows collection owners to index and present a searchable and browsable version of their documents online. The Niupepa collection of historic Māori newspapers (see Figure 2) uses the Greenstone system for presenting heritage information.²

Many existing computational techniques focus on storage and processing of digital data, but heritage information poses the additional challenge of preserving or returning the cultural context to

objects in a heritage collection. Because heritage documents and digital objects are scattered across a variety of resources, they are often stripped of their interpretation and the connection to the experiences of their originating people. Further development of any of these tools must be based on collaboration between indigenous domain experts and software engineers to ensure these connections are renewed and the cultural objects are treated with appropriate respect.

Current Challenges

Recent years have seen renewed debates about the return of heritage objects to indigenous communities.³ However, for intangible cultural heritage objects, such as information about the traditions of rendang, simple return is not a viable option and digital repatriation offers a possible alternative. Capisco is a suitable tool to semantically search vast existing collections to identify relevant documents, such as missionary reports, anthropological monographs, early geographic surveys, and Victorian tourist reports. These then can be compiled into portable collections, which may be returned to the indigenous peoples whose images, cultures, and histories were captured in the identified documents. In this challenging work, the rights of indigenous people relating to the collection and curation of data about their culture and heritage must be acknowledged.⁷ Appropriate processes and data representations need to

be developed by, with, and under guidance of the affected communities.¹ 

References

- Ahuriri-Driscoll, A. et al. Scientific collaborative research with Māori communities: Kaupapa or Kūpapa Māori? *AlterNative: An Intern. J. Indigenous Peoples* 3, 2 (2007), 60–81.
- Appertley, M., Cunningham, S.J., Witten, I.H. and Keegan, T.T. Niupepa: An historical newspaper collection. *Commun. ACM* 44, (May 2001), 86–87.
- Bell, J.A., Christen, K. and Turin, M. After the return: Digital repatriation and the circulation of Indigenous knowledge. *Museum Anthropology Review* 7, 1–2 (2013).
- Cunningham, S.J., Hinze, A.M., Bainbridge, D., Taube-Schock, C. and Ryan, T. Building heritage document collections for Pacific Island nations using semantic-enriched search. In *Proceedings of the Samoa Conference III*. National University of Samoa, 2015.
- Hinze, A., Bainbridge, D., Cunningham, S.J., Taube-Schock, C., Matamua, R., Downie, J.S. and Rasmussen, E. Capisco: Low-cost concept-based access to digital libraries. *Intern. J. Digital Libraries* 20 (2019), 307–334; <https://doi.org/10.1007/s00799-018-0232-3>.
- Hinze, A., Taube-Schock, C., Bainbridge, D., Matamua, R. and Downie, J.S. Improving access to large-scale digital libraries through semantic-enhanced search and disambiguation. In *Proceedings of the 15th ACM/IEEE-CS Joint Conf. Digital Libraries*. ACM, 2015.
- Kukutai, T. and Taylor, J., eds. *Indigenous Data Sovereignty: Toward an Agenda*. Anu Press, 2016.
- Nurmufida, M., Wangrinen, G.H., Reinalta, R. and Leonardi, K. Rendang: The treasure of Minangkabau. *J. Ethnic Foods* 4, 4 (2017), 232–235; <https://doi.org/10.1016/j.jef.2017.10.005>
- Putra, H.S., Mahendra, R. and Darari, F. BudayaKB: Extraction of cultural heritage entities from heterogeneous formats. In *Proceedings of the 9th ACM Intern. Conf. Web Intelligence, Mining and Semantics*, 2019, 6:1–6:9.
- Raffles, T.S. *The History of Java*. Printed for Black, Parbury, and Allen, London, 1817.
- W3C. SPARQL 1.1 Query Language, 2013; <https://www.w3.org/TR/sparql11-query/>
- W3C. RDF 1.1 Primer, 2014; <https://www.w3.org/TR/rdf11-primer/>

Sally Jo Cunningham is an associate professor in the Department of Computer Science at the University of Waikato, Hamilton, New Zealand.

Fariz Darari is an assistant professor in the Faculty of Computer Science, Universitas Indonesia and co-director of Tokopedia-UI AI Center, Indonesia.

Adila Krisnadi is an assistant professor in the Faculty of Computer Science, Universitas Indonesia and co-director of Tokopedia-UI AI Center, Indonesia.

Annika Hinze is an associate professor in the Department of Computer Science at the University of Waikato, Hamilton, New Zealand.

Cybersecurity in Pacific Island Nations

BY CARSTEN RUDOLPH, SADIE CREESE, AND SAMEER SHARMA

THE OCEANIA REGION is at a crossroads, with physical security challenges headlined by a changing climate providing an existential threat to many of the Pacific Island nations that call the region home. Furthermore, the region has become a geopolitical battleground with major actors including Australia, China, the European Union, New Zealand, and the U.S. all working to gain influence.

During the Pacific Islands Forum in late 2018, Pacific Island leaders outlined their security concerns through the Boe Declaration—a pronouncement that looked to establish an expanded concept of security.^a One such se-

a Pacific Island Forum Secretariat. Boe Declaration on Regional Security, 2018; <https://www.forumsec.org/boe-declaration-on-regional-security/>



curity issue that has begun to emerge and will only accelerate without individual and collective action is the threat of cyberattacks and cybercrime in the region. Most of the Pacific Island nations have either recently or are in the process of drastically increasing the access to international connectivity through subma-

rine cables that enable the potential transformation to digital economy and digital society of such nations. However, the rapid expansion of fast access to the globally connected Internet has also increased the risks of Pacific Island nations becoming victims of cyberattacks and cybercrime.

A good deal of ad hoc support does exist. Examples of such activities include international organizations offering hands on capacity building cybersecurity workshops for government representatives or specialized training for police forces. A truly global initiative that began in April 2018 is now supporting Pacific Island nations in developing and strengthening their cybersecurity capacity. The not-

for-profit Oceania Cyber Security Centre (OCSC) in Melbourne, Australia, has become the first regional partner of the Global Cyber Security Capacity Centre (GCSCC), at the University of Oxford in the U.K. and has teamed up with United Nations International Telecommunication Union (ITU) to complete cybersecurity capacity maturity assessments of countries in Polynesia, Micronesia, and Melanesia. This work targets 15 countries for cybersecurity capacity reviews based on the Cybersecurity Capacity Maturity Model for Nations (CMM).^b While five reviews have been completed, the

b Cybersecurity Capacity Maturity Model for Nations, 2017; <http://bit.ly/2QXAZpU>

The rapid expansion of fast access to the globally connected Internet has also increased the risks of Pacific Island nations becoming victims of cyberattacks and cybercrime.

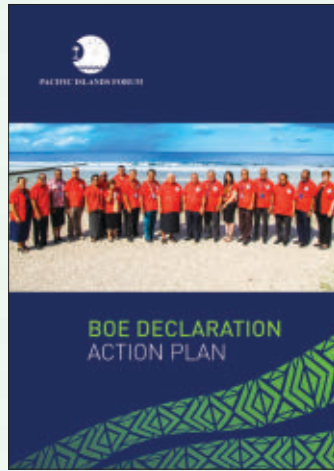
first comprehensive report has now been approved for released by the Samoan government and has been published by the Ministry of Communications and IT.^c

The process, the choice of stakeholders involved, the resulting report, and all subsequent policy decisions and projects are owned by the individual countries. However, researchers from the University of Oxford and the eight OCSC member universities in Victoria, Australia, accompany the project with research exploring the various themes and evolving cybersecurity context of the region.

While the detailed reports are confidential (as long as they are not published by the particular country), there are general learnings and themes emerging from the work completed in Samoa, Tonga, Vanuatu, Papua New Guinea, and Kiribati. It shows that cybersecurity is seen in a very wide sense. In addition to the core areas, particularly important topics are malicious use of social media, fake news, deception, and fake accounts. One main

^c Cyber-security Capacity Review Independent State of Samoa, 2018; <https://mcit.gov.ws/2019/04/03/cyber-security-capacity-review-2018/>

CMM reports contain many recommendations and constitute a first step toward building capacity and developing cybersecurity strategies and policies.



The Boe Declaration on Regional Security expounds on the need for and methods to obtain greater “human” security, including the handling of cyberthreats to safety. Pacific Island leaders signed the pronouncement by handprint.

issue is the lack of qualified people to support governments, companies, education, and the public in general. Further, there is a high probability for people with higher qualification to emigrate to other more developed countries. The accompanying research and future CMM reviews will reflect on what it means for

Boe Declaration Action Plan

Purpose

This Action Plan provides a framework for Forum Members to implement the Boe Declaration on Regional Security, in collaboration with key regional and international stakeholders.

The Action Plan sets out to positively and/or proactively shape our regional security environment by progressing specific, achievable and targeted activities under the relevant strategic focus areas prioritised under the Boe Declaration on Regional Security.

The Action Plan will be supplemented by a 12-month activity matrix which is a rolling (organic) document with activities to be reviewed by Forum Members and relevant stakeholders on an annual basis.

Strategic Context

The Action Plan is positioned in the context of the 2014 Framework for Pacific Regionalism, the Blue Pacific Narrative and the 2018 Boe Declaration on Regional Security and existing Forum Security Declarations. The Action Plan's component on engagement and advocacy is aligned to the 2018 Forum International Engagement and Advocacy Strategy.

Framework for Pacific Regionalism

In 2014, Forum Leaders endorsed the Framework for Pacific Regionalism (FPR) espousing their vision for the Pacific as “a region of peace, harmony, security, social inclusion, and prosperity so that all Pacific people can lead free, healthy and productive lives”.

Under the FPR, the regions security objective is defined as “security that ensures stable and safe human, environmental and political conditions for all”.

The Blue Pacific Narrative

In 2017, the Forum Leaders endorsed the Blue Pacific narrative as a call for Forum Members to work together as one Blue Continent recognising it as a catalyst for stronger and deeper Pacific regionalism.

Also in 2017, recognising the shared strategic value of our region through the Blue Pacific narrative, Forum Leaders noting the changing dynamics of our geopolitical environment and regional security landscape called for the development of a new, fit for purpose regional security declaration that recognises an expanded concept of security inclusive of human security, humanitarian assistance, promoting environmental security, and regional cooperation in building resilience to disasters and climate change.

The Boe Declaration on Regional Security

The Secretariat undertook rigorous and extensive consultations with Members, CHOP, regional law enforcement secretariats and relevant stakeholders to develop a new regional security declaration building off existing Forum security declarations but which accounted for an evolving regional security environment. In 2018, Leaders endorsed the Boe Declaration on Regional Security at their meeting in Niue (attached as [Annex 1](#)).


The Boe Declaration is a call to action for stronger and cohesive regional security cooperation and contribution through the assertion of our collective will and collective voice as peoples of the Pacific.

It recognises that climate change remains the single greatest threat to the livelihoods, security and wellbeing of the peoples of the Pacific. It further recognises that the Pacific is faced with a regional security environment confronted with complex security challenges framed by an expanded concept of security, within a dynamic geopolitical environment.

Pacific Island nations to expose themselves to the advantages and risks of the Internet community and what pathways can be developed toward higher maturity.

CMM reports contain many recommendations and constitute a first step toward building capacity and developing cybersecurity strategies and policies. They clearly show that particular characteristics of island countries in the Pacific require individual approaches to cybersecurity and digital transformation in general. Therefore, reviews are now followed by projects for capacity building, for example, to establish technical capabilities for incident response, to create awareness programs

satisfying individual requirements of the countries, or critical infrastructure protection.

The joint projects in the Pacific are part of a global initiative on cybersecurity capacity building for nations and represent a model for other regions, with a long-term vision of a network of regional centers, for which the OCSC is the reference model. 

Carsten Rudolph is an associate professor and the director of Oceania Cyber Security Centre at Monash University, Melbourne, Australia.

Sadie Creese is a professor and the director of the Global Cyber Security Capacity Centre in the Oxford Martin School at the University of Oxford, U.K.

Sameer Sharma is a Regional Director a.i. of the International Telecommunication Union, United Nations Specialized Agency for Information and Communications Technology, Regional Office for Asia-Pacific, Bangkok, Thailand.

Singapore's Cybersecurity Ecosystem

BY KAREN TEH, VIVY SUHENDRA, SOON CHIA LIM,
AND ABHIK ROYCHOUDHURY

A SUCCESSFUL DIGITAL ECONOMY requires cybersecurity to be a vital enabler, protecting the interests of individuals and businesses and enabling the resilience of businesses and services. Since 2013, Singapore's medium- to long-term directions for cybersecurity is to develop R&D expertise and capabilities to improve the trustworthiness of cyber infrastructures and systems with an emphasis on security, reliability, resilience, and usability among government agencies, academia, and industry. Various initiatives to support research, innovation, and enterprise have been implemented under the Whole-of-Government National Cybersecurity R&D (NCR) Programme.⁸ The program supports a synergistic range of initia-



tives to advance technological state-of-the-art in thematic National Satellites of Excellence in universities, grants for local research projects, international research collaborations, and joint technology developments with industry. Innovation is fostered through cross-sector R&D

discussions and partnerships and fast-tracked by national testbeds for safe and repeatable cybersecurity experiments.

Research Impact

Research entities in Singapore have adopted a multidisciplinary, mission-oriented approach in solving cybersecurity problems with notable outcomes. There are several such examples of research impact in cyber-security being achieved by Singapore's institutions including in software security, systems security, and Internet of Things (IoT) security.

A noticeable impact has been achieved in the field of vulnerability detection in programs, namely fuzz testing. AFLFast² is an extension of the widely

used AFL¹⁶ developed at Google, a greybox fuzzer, which uses lightweight program instrumentation to gain coverage information for guiding program path exploration. AFLFast achieved tenfold speed-up over AFL using strategies to gravitate path exploration toward low-frequency paths, which enabled it to expose several previously unreported CVEs that could not be exposed by AFL in 24 hours. It contributed to the runner-up team Codejitsu at DARPA Cyber Grand Challenge (2016) and has been integrated into mainstream AFL.

Scantist⁹ is a university spin-off with technologies for scalable vulnerability scanning and analysis at binary as well as source code levels, providing

Innovation is fostered through cross-sector R&D discussions and partnerships and fast-tracked by national testbeds for safe and repeatable cybersecurity experiments.



Scenes from Singapore International Cyber Week held in October 2019.

vulnerability management tools with low effort and expertise requirements. It combines static analysis in the form of signature-based matching and metrics to detect vulnerable functions, with dynamic analysis in the form of smart fuzzing to discover memory corruption vulnerabilities. The tools produce highly targeted remediation advice to allow quick and accurate fixes.

Anquan¹ is another spin-off providing distributed ledger and trusted computing platforms for financial markets. It was appointed as a technology partner, alongside Deloitte and Nasdaq, in a 2018 project by the Monetary Authority of Singapore (MAS) and Singapore Exchange (SGX) to develop delivery versus payment (DvP) capabilities for reduced risk settlement of tokenized assets across different blockchain platforms.⁶ Anquan's DvP solution design in this project is based on its permissioned blockchain with capabilities developed by the research group in Singapore,^{5,10} including scalability through a network sharding technique, security protection against malicious nodes, a smart contract language amenable to formal verification, and privacy with hardware-rooted trusted execution environment.

Research in cyber-physical system security has also generated sophisticated algorithms, software, and devices to detect physical, sensor, network, and information attacks.¹⁵ Among the practical outcomes is VVATER,¹¹ a mixed-reality visualization of process states and attacks in operational technologies such as a water treatment

and distribution plant, to help operators investigate and respond to attacks timely and comprehensively without advanced cybersecurity skills. A key novelty of VVATER is its ability for visualizing the interconnection of various infrastructures in historical plant operation and path of attacks in complex scenarios, as well as the resulting process anomalies and whether or not the anomaly is detected.

Support for Research, Innovation, Enterprise

Ecosystem support plays an important role in ensuring research endeavors are responsive to and impactful on cybersecurity needs of the industry and society. Building on the research successes, Singapore has set up three National Satellites of Excellence: on Trustworthy Software Systems at the National University of Singapore, on Mobile Systems Security at the Singapore Management University, and on Secure Critical Infrastructure at the Singapore University of Technology and Design. These satellites provide strategic thrusts in a focus area and help to develop the research and innovation ecosystem in Singapore, working closely with various national initiatives such as the Singapore Cybersecurity Consortium.

The Singapore Cybersecurity Consortium¹² is an organized construct to grow communities, foster partnerships across academia, industry, and agencies, and seed technology explorations around research to multiply and amplify its impact. Operating environment challenges and related research out-

comes are discussed in its thematic Special Interest Groups, leading to better appreciation of research capabilities, problems for research, and joint innovation development.

The National Cybersecurity R&D Laboratory⁷ and iTrust Laboratories¹⁴ are shared infrastructures facilitating enterprise-IT and OT security research experimentation, technology evaluation, and training. Research teams from academia and industry seeking to commercialize cybersecurity technologies are mentored on customer discovery and product positioning in the Lean LaunchPad Singapore: Cybersecurity Track,⁴ which integrates both business and technological perspectives. Complementing the effort in this space is Innovation Cybersecurity Ecosystem at Block 71 (ICE71),³ which provides entrepreneurship, accelerator, upscaling programs for start-ups, contributing to ecosystem growth in ASEAN.

Positioning Singapore as a Regional Cybersecurity Hub

Leveraging this comprehensive R&D foundation and its reputation as a trusted financial hub, Singapore is well-positioned to be a cybersecurity hub for the region. It attained the status of a Common Criteria Certificate Authorising Nation in January 2019. With this status, developers based in Singapore can enjoy lower costs and shorter time in attaining an internationally recognized certification mark. This facilitates the exportability of cybersecurity products produced in Singapore. The Singa-


pore International Cyber Week¹³ is the region's most established annual cybersecurity event, providing an ideal platform to discuss, strategize, and form partnerships across the nations.

All such efforts help to nurture the cybersecurity innovation ecosystem in Singapore and the region, which remains locally rooted and globally connected. This regional-global interplay is indeed a marked characteristic of all cybersecurity initiatives in the region featured in this section. The cybersecurity capacity maturity assessments of countries in the Pacific region (for more information, see the Rudolph et al. article in this section) is part of a global initiative on cybersecurity capacity building and is an application of the research on the Cybersecurity Capacity Maturity Model for Nations developed in the U.K.'s University of Oxford. The assessment project is accompanied with research on the evolving cybersecurity context of the region, with findings feeding back to the research on the model itself with possible benefits to other regions. Asiacypt, the regional flagship IACR conference for advances in security and cryptography research, gathers researchers in Asia and Oceania for closer collaboration while staying aligned to the international body of IACR and making borderless research contributions. (For more information on Asiacypt, see the Phan et al. article in this section). These initiatives and ours nurture the cybersecurity ecosystem in different but connected ways—a thriving innovation ecosystem

The Singapore International Cyber Week is the region's most established annual cybersecurity event, providing an ideal platform to discuss, strategize, and form partnerships across the nations.

would enhance the germination of ideas as well as accelerate the technology transfer and industry adoption of research results, which in turn supports the building and maturing of cybersecurity capabilities in the region.

Future Research Areas

Our R&D for the advancement of a secure smart nation does not end here. We will continue to focus R&D on security and the healing of software stacks in autonomous vehicles and the IoT, including curtailing attacks coming from nonfunctional domains. Future research areas will also focus on safe and dependable interactions between the physical worlds of sensors, motors, actuators, and robotics, and the cyber world of data processing, artificial intelligence, networking, and control systems to better protect interests and enabling the resilience of businesses and services in a digital economy of IoT and actions. 

References

1. Anquan Capital, 2019; <https://www.anquancapital.com/>.
2. Böhme, M., Pham, V-T and Roychoudhury, A. Coverage-based Greybox fuzzing as Markov Chain. In *Proceedings of the 2016 ACM SIGSAC Conf. Computer and Communications Security*, 1032-1043.

3. Innovation Cybersecurity Ecosystem at Block 71—ICE71; <https://ice71.sg/>.
4. Lean Launchpad Singapore. Past projects, 2019; <https://nus.edu/2T9k7zd>.
5. Luu, L. et al. A Secure sharding protocol for open blockchains. In *Proceedings of the 2016 ACM SIGSAC Conf. Computer and Communications Security*, 17–30.
6. Monetary Authority of Singapore, SGX, and Deloitte. Delivery versus Payment on Distributed Ledger Technologies, 2018; <http://bit.ly/2Qsw15a>.
7. National Cybersecurity R&D Laboratories, 2019; <https://ncl.sg/>.
8. National Research Foundation. National Cybersecurity R&D Program, 2019; <https://www.nrf.gov.sg/programmes/national-cybersecurity-r-d-programme/>.
9. Scantist. Vulnerability management. Simplified, 2019; <https://scantist.com/>.
10. Sergey, I., Kumar, A. and Hobor, A. Scilla: A Smart Contract Intermediate-Level Language, 2018; <https://arxiv.org/abs/1801.00687>.
11. Shrivastava, S. Virtual and mixed reality for security of critical city-scale cyber-physical systems. *iTrust Times 1*, (Apr–Jun 2019). Singapore University of Technology and Design.
12. Singapore Cybersecurity Consortium, 2019; <https://sgcsc.sg/>.
13. Singapore International Cyber Week, 2019; <https://www.sicw.sg/>.
14. Singapore University of Technology and Design. iTrust Labs Home; <https://itrust.sutd.edu.sg/itrust-labs-home/>.
15. Taormina, R. and Galelli, S. Deep-learning approach to the detection and localization of cyber-physical attacks on water distribution systems. *J. Water Resources Planning and Mgmt* 144, 10 (2018), 04018065.
16. Zalewski, M. American fuzzy lop (2.52b), 2017; <http://lcamtuf.coredump.cx/afl/>.

Karen Teh is Senior Deputy Director of Cybersecurity R&D at the National Research Foundation, Singapore.

Vivy Suhendra is Executive Director of the Singapore Cybersecurity Consortium.

Soon Chia Lim is the director of the Cybersecurity Engineering Centre for the Cyber Security Agency of Singapore.

Abhik Roychoudhury is Provost's Chair Professor at the National University of Singapore.

Copyright held by authors/owners.

Innovating Services and Digital Economy in Singapore

BY THOMAS HO CHEE TAT AND GEORGE LOH CHEE PING

SINGAPORE HAS BEEN investing in information and communication technologies (ICT) since the 1990s, and more recently, has been driving toward digitalization and automation for the transformation of the society, government, and economy. This article reports on Singapore's technology development in the services and digital economy (SDE) and concludes with insights for research over the next five years.

Singapore is ranked as the world's second most digitally competitive country by IMD.³ The nation's spending in technologies like artificial intelligence (AI), as shown in Figure 1 compared to international, is critical to the success of Singapore's Smart Nation and Digital Economy vision. Singapore's Re-

a <http://bit.ly/35xjmCm>

search Innovation & Enterprise (RIE) 2020 plan continues to “play a key role in driving Singapore's future economy, as well as address national and societal priorities.” According to the Deputy Prime Minister Heng Swee Keat, who also serves as chairman of the National Research Foundation (NRF), the RIE's \$19 billion investment, “will support the future economy council's efforts to encourage adoption of digitalization and automation across Singapore's economic sectors.”⁵

Nation Imperatives

With its genesis in interactive digital media in 2006, SDE was established in 2016.¹ It focused on the development of digital capacities and cross-cutting core competencies for the three domains shown in Figure 2. The SDE strategy consisted of developing human talents to learn, create, and innovate new digital technologies like AI, privacy and trust, cyber-



The Changi Airport in Singapore is considered one of the most technologically advanced airports in the world.

security, data analytics, and digital twinning.

The SDE competencies described here serve three national strategic imperatives that continue to keep Singapore competitive in trade and commerce. First, *to maintain Singapore's strategic position as a neutral, trusted node* in key spheres of global activity, which would be strengthened with advanced capabilities in cybersecurity, data privacy preservation, distributed digital ledger, and other technologies for the interconnected global digital economy. NRF established the National Cybersecurity R&D program to develop technologies for national strategic requirements. The National Cybersecurity Laboratories and Consortium ensure

forefront technologies are always available to guard Singapore's digital assets against cyber threats.

Second, *to achieve Singapore's Smart Nation vision—to make Singapore “an outstanding city in the world ... for people to live, work, and play in, where the human spirit flourishes.”* Prime Minister Lee Hsien Loong said at the 2019 Smart Nation Summit that the Smart Nation push “is applying technology to solve real problems that will make a difference to people's lives, across the whole of society” to improve the quality of life for all. The Jurong Lake District became Singapore's Smart Nation testbed.² It demonstrated technologies like smart lighting that minimizes energy consumption according to

NRF will invest in the development of sectoral applications using 5G communication technologies as well as beyond 5G.

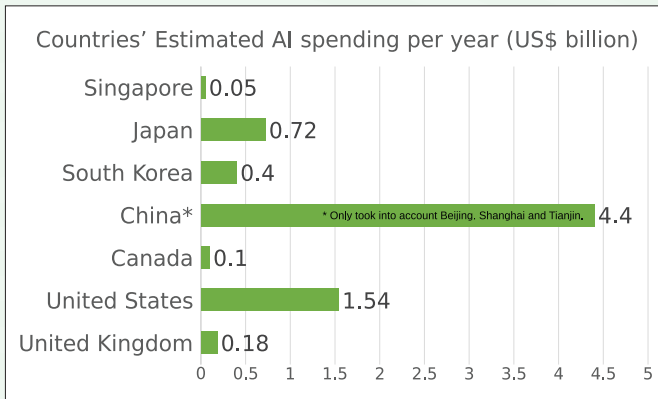


Figure 1. International benchmarking of Singapore's AI spending (2017).

ambient luminance as well as autonomous vehicles and sensor-augmented traffic light technologies³ for advanced mobility transportation and pedestrian safety.

Third, to overcome Singapore's manpower constraint and aging workforce and to increase productivity. NRF established The National AI Singapore (AISG) R&D program⁴ in 2017 to bring together universities, research institutes, and companies to innovate game-changing AI and automation capabilities. To date, it has completed nine industrial projects with 29 on-going and 363 more in the pipeline. The Singapore Data Science Consortium

was formed with universities and companies as members to do data analysis for Singapore's cosmopolitan society. It has completed nine industrial projects and hosted five industry-academic exchange events.

Going Forward

Singapore will continue to invest in SDE R&D in order to maintain and develop the capabilities mentioned here, such as AI, cybersecurity, and privacy-preserving, as well as invest in new R&D areas to support the national initiatives. Communication and connectivity are key technology components of digital systems and solutions. In addition,

NRF will invest in the development of sectoral applications using 5G communication technologies as well as beyond 5G. IoT devices and platforms will support and transform business models and solutions. A few of the emerging applications include connected cars, smart manufacturing, and smart/green building. IoT technologies will be an integral part of the Smart Nation vision and will be an area of R&D investment for SDE. Quantum technologies have already transformed some businesses. NRF intends to invest substantially to translate the quantum sciences developed in the Center of Quantum Technology (CQT) into

quantum technologies, which include quantum key distribution (QKD), post-quantum encryption, quantum algorithms, and computing.

The approach to achieve Singapore's desired outcomes of optimizing resources and shaping Singapore's economy involved identifying key technological focus areas and filtering them with required strategic criteria. Once the R&D areas are identified, action plans are formulated toward execution and achievement. 

References

- Hio, L. Big push for science and tech research, services and digital economy. *The Strait Times* (Jan. 9, 2016); <https://www.straitstimes.com/singapore/services-and-digital-economy>.
- Kelleher, J. In-depth look at the smart nation trials at Jurong. *Opengovasia*, Feb. 1, 2019; <https://www.opengovasia.com/in-depth-look-at-the-smart-nation-trials-at-jurong/>
- Institute for Infocomm Research. IHI Corporation (SAINT); <https://www.a-star.edu.sg/i2r/partnership/item/itemid/>
- National Research Foundation, AI Singapore, Nov. 7, 2018; <https://www.nrf.gov.sg/programmes/artificial-intelligence-r-d-programme>
- Teng, A. Innovation, transformation, and investment in workers key for a vibrant S'pore economy (Today Online, May 10, 2018); <http://bit.ly/2T4odIB>

Thomas Ho Chee Tat is Head of Services and Digital Economy at the National Research Foundation Singapore.

George Loh Chee Ping is Director of Services and Digital Economy at the National Research Foundation Singapore.

Copyright held by authors/owners.

RESEARCH INNOVATION & ENTERPRISE (RIE): CLOSER INTEGRATION OF STRATEGIES

- RIE planning to be oriented along four major technology domains, supported by three cross-cutting horizontals

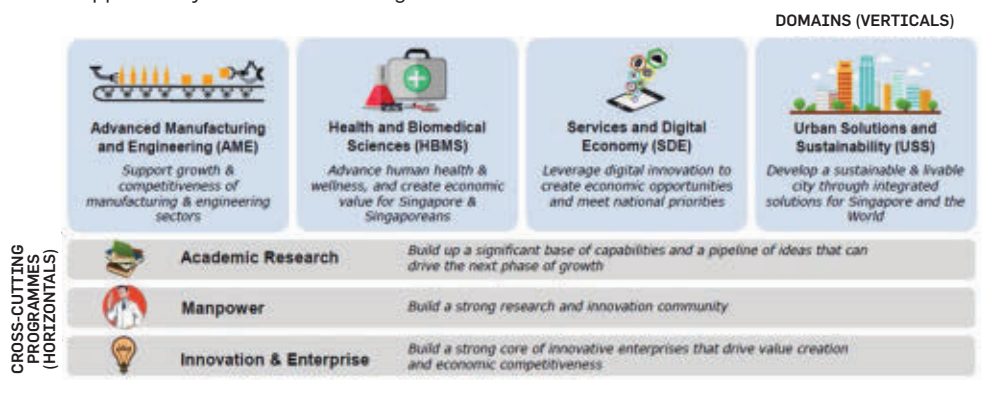


Figure 2. RIE framework for years 2015–2020.

BY SINTIA TEDDY-ANG /AI SINGAPORE
AND ABIGAIL TOH /AI SINGAPORE

AI Singapore Empowering a Smart Nation

AI SINGAPORE (AISG) was launched in June 2017 as an integrated, impact-driven, research and innovation program in artificial intelligence (AI) for the entire country. As a national initiative, AISG brings together the strength of Singaporean research bodies in Singapore's Autonomous Universities (AUs) and research institutes, together with the vibrant ecosystem of AI start-ups and companies developing AI products, to perform use-inspired research, create innovative AI solution, and develop the talent to power Singapore's AI efforts.

To achieve Singapore's national mission, AISG's activities are anchored around four key pillars:

► **AI research** is geared toward building deep AI research capabilities in Singapore through fundamental research. This pillar invests in research

for next-generation AI techniques/ algorithms beyond deep learning, and aims to strengthen Singapore's scientific leadership as a key player in the global AI race.

► **AI technology** is focused on creating significant economic and social impacts by tackling national or global challenges using AI. Under this pillar, the first AI in Health Grand Challenge was launched in June 2018, where the challenge statement was "How can AI help primary care teams stop or slow disease progression and complication development in three "Highs" (or 3H)—Hyperglycaemia (diabetes), Hyperlipidaemia (high cholesterol), and Hypertension (high blood pressure)—patients by 20% in five years?" This AI challenge is a novel approach to multidisciplinary collaboration and translation from research to practice. It has enabled local and international collaborators to come together to address Singapore's national health challenge.

► **AI industry innovation** accelerates the adoption of AI technology in the industry through proof-of-concept projects and talent development. Under the 100Experiments program, AISG has undertaken 50 AI projects for our local industry and has deployed more than 10 projects into production.^a Our award-winning AI Apprenticeship Programme (AIAP),^b which is a full-time, nine-month program to train and groom local Singaporean AI talent, has seen about 60 engineers graduate from the program as of October 2019. The plan is to train up to 500 Singaporean AI apprentices over the next five years.

► **AI Makerspace** is a first for Singapore. Leveraging on the intellectual properties (IPs) and experiences from the other three pillars, AI Makerspace hosts an AI knowledge base (that is, open source AI libraries, API, and

a As of Dec 2019.

b AISG's AI Apprenticeship Programme (AIAP) was awarded the 2019 IDC Singapore Digital Transformation Award for Talent Accelerator in Oct 2019.



100Experiments is AISG's flagship program to solve industry business problems through the design and development of AI solutions, translate AI IPs from academia to industry, and help companies build their own AI teams.

datasets) and will be the first national platform for accessing cutting-edge AI tools and solutions for research and commerce. Targeted at small and medium enterprises (SMEs) and start-ups, AI Makerspace will help industries jump-start their AI journey by providing access to resources for experimentation, such as curated datasets from industry and government, cutting-edge AI tools, and supercomputing resources specialized for AI workloads.

100Experiments

100Experiments (100E) is AISG's flagship program to solve industry business problems through the design and development of AI solutions, translate AI IPs from academia to industry, and help companies build their own AI teams. An organization can propose a problem statement where no commercial-off-the-shelf AI solution exists, but can potentially be solved through AISG's ecosystem of researchers and research IPs within nine to 18

months. AISG will assemble a team of AI researchers and engineers from Singapore's research and development ecosystem to work on an organization's problem statement. Through a collaborative process, a company's existing technical manpower will work alongside a team of AI researchers and engineering assembled by AISG to develop AI solutions while helping the company build up its internal AI capabilities.

Two examples of successful 100E projects include:

- ▶ **Automatic assessment of chronic wounds in diabetic and elderly patients.** The current practice of assessing wounds is time-consuming as it takes an average of about 30 minutes per wound. The quality of information gathered is also very subjective and relies on the experience of the attending nurse(s). Only about 2% of the nurse population in Singapore are qualified wound nurses, thus it is not possible to offer

DOI: 10.1145/3378418

Developing AI for Law Enforcement in Singapore and Australia

BY LAM KWOK YAN AND CAMPBELL WILSON

NTU SPIRIT Smart Nation Research Centre, together with the Singapore Judiciary, has successfully developed an Intelligent Case Retrieval System (ICRS) using AI capabilities. ICRS enables efficient retrieval of relevant precedent cases through the use of continuously adaptive AI/data analytics approaches. The use of such tools can help the legal profession to understand case details and perform legal research by trawling through the case repositories at a faster and more accurate rate to obtain the most relevant case precedents and identify possible outcomes in different areas of law.

The value of ICRS is to better enable all parties to evaluate the strengths or weaknesses of their cases. With better quality legal submissions, judges too are assisted in their decision-making processes, thus elevating the quality of judgments delivered. The ultimate aim is to fortify the domestic confidence in our courts and boost Singapore's international reputation as a leading Judiciary in the region and the world.

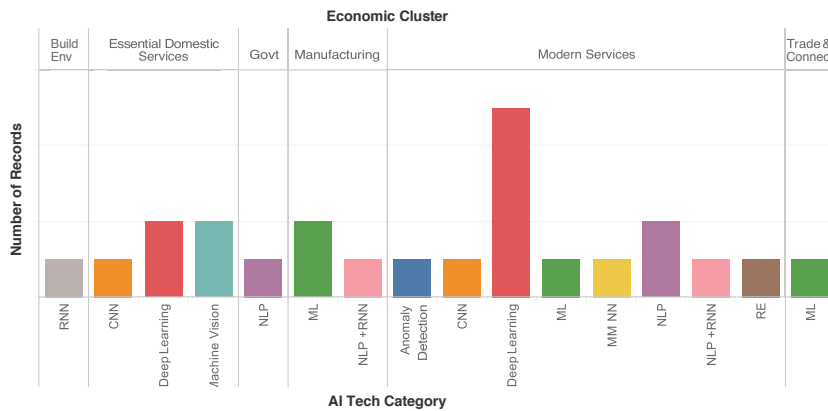
In Australia, Monash University and the Australian Federal Police have joined forces, exploring the use of AI for fighting crime and protecting the vulnerable. They have launched a joint research lab known as AI for Law Enforcement and Community Safety (AiLECS). The lab grew out of previous work into the automated classification of online child exploitation material. The abhorrent material encountered by investigating police reflects the terrible harm endured by victims, while viewing it inflicts psychological damage on investigators. The lab is improving the techniques to classify this and other distressing material using AI, while also undertaking research into the ethics and explanations of such technologies in law enforcement.

The Singapore and Australian research teams have hosted a series of joint dialogues exploring opportunities for collaboration in further developing AI technologies for law enforcement agencies and the judiciary.

Lam Kwok Yan is a professor and the director of Smart Platform Infrastructure Research on Integrative Technology, and director of Strategic Centre for Research in Privacy-Preserving Technologies and Systems, Nanyang Technological University, Singapore.

Campbell Wilson is Associate Dean (International) and co-director of the Monash University/Australian Federal Police AiLECS Lab, Monash University, Australia.

© 2020 ACM 0001-0782/20/4

Figure 1. Spread of AI technologies across 100E projects.

AI Makerspace will help industries jump-start their AI journey by providing access to resources for experimentation, such as curated datasets from industry and government, cutting-edge AI tools, and supercomputing resources specialized for AI workloads.

adequate care to chronic wound patients with existing practices.


AI Singapore worked together with a local start-up, KroniKare, to develop a Wound Scanner that uses computer vision, image processing, and semantic segmentation on an AI-driven handheld device that mimics the wound analysis by specialists. The KroniKare Wound Scanner is the first AI-based diagnostic tool to be registered in Singapore under the Health Sciences Authority (HAS) Class B medical device. The scanner uses multi-spectral images to automatically analyze and report chronic wound conditions, and therefore allow healthcare institutions to better document wound conditions, triage patients, and allocate resources for wound management. This capability has resulted in improved patient outcomes in terms of early detection and faster interventions for major wound complications and abnormalities. The digitized documentation also produced more accurate records and wound assessment time has been reduced to only 30 seconds, thus significantly reducing nurses' workloads. The tool also helped train and enhance the skills of junior nurses.

► **AI solution to transform the online search experience for Asian travelers.** With English as the dominant language being used online by 25% of all Internet users, today's search engines are extremely efficient in understanding travel search queries and providing query resolutions in the English language. However, when dealing with travel search queries conducted

in Asian languages such as Japanese, Korean, simplified Chinese, and traditional Chinese, the performance of the search engines decline significantly and the accuracy of query resolution dips.

The Expedia Group and AI Singapore's project team are working to leverage natural language processing and machine learning to develop an AI-based model to enhance search query understanding and resolution in Japanese before extending the model to other Asian languages to enhance online search efficiency. When completed, the AI solution will enable Expedia Group to deepen its understanding of travel search query patterns and nuances in Asian languages, and equip the travel platform with the ability to better serve the needs of Asian travelers by improving the accuracy and efficiency of search query resolution.

The AI technologies typically used in the 100E projects are deep learning, computer vision, and natural language processing. The accompanying figure summarizes the spread of AI technology across 100E projects in the various economic sectors.

Over the next few years, AISG will focus on encouraging national and international research collaborations, accelerate the adoption of AI, and grow the AI talent pipeline for Singapore. 

Sintia Teddy-Ang is director of Planning and Operations at AI Singapore.

Abigail Toh is head of Marketing and Communications at AI Singapore.

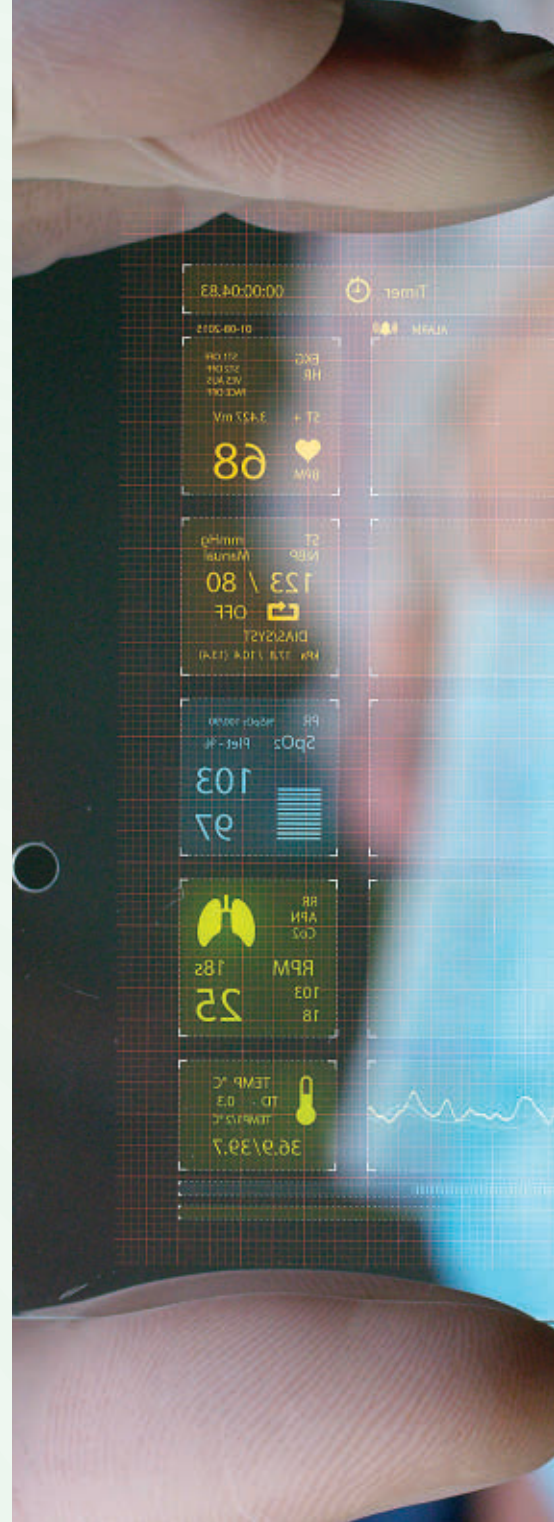
BY CHRIS BAIN /MONASH UNIVERSITY
AND ABRAHAM OSHNI ALVANDI /MONASH UNIVERSITY

Digital Healthcare Across Oceania

MANY NATIONS HAVE begun a serious exploration of digital health (DH)^{16,17} to address the pressures on their healthcare systems.^{1,3,9,12} Citizens and technologists are also driving related changes that facilitate growth in DH—for example, through the greater use of mobile technologies, sensors, extended reality, and artificial intelligence (AI).

These technologies are part of DH as defined by the World Health Organization (WHO),¹⁴ enabling new care paradigms that are distinguishable from older technologies like electronic medical records (EMRs). The impact of these DH technologies in healthcare in each country is captured through the concept of Digitally Enabled Healthcare Ecosystems (DEHEs).

In this article, we address three questions in several national settings in Oceania (Malaysia,



Singapore, and Australia). These key questions are:

1. What is the current state of progress of DH?
2. What is the apparent social license for DH? (for example, given concerns about privacy).
3. In the next 10 years, is DH likely to fully take hold? (for example, as evidenced by major disruption of traditional healthcare).

Here, we explore these questions based on local expert opinion in the case of Malaysia and Singapore cross-refer-



encing with publicly available information. Let's start by considering Malaysia.

Malaysia

The country's early attempt in transforming healthcare was the of launching its telemedicine program in 1997.⁴ According to Shah Yasin, head of medicine at Monash University Malaysia, "the ... government has been talking about telehealth for more than two decades."

Another early contribution in this space was the Malaysian Health Data Warehouse (MyHDW); a key initiative

commenced in 2013. MyHDW has been promoted as a sound national platform for the integration of health data (from a variety of public and private health sources) intended to improve decisions on patient care.⁵

In addition, nearly 20% of public hospitals in the country have implemented EMRs, but challenges to full implementation remain due to software incompatibility and inadequate ongoing support.⁷ As in some other countries, private hospitals in Malaysia use systems that are not necessarily compatible with those

of other (including public) facilities, according to Yasin. Despite some of these strides, there remain problems with a lack of integration between systems, a lack of private healthcare involvement, and inadequate long-term investment.

In 2019, the Malaysian Director General (DG) of Health flagged a shift to a more modern DH approach, away from the historic focus on telemedicine. He stated: "We have agreed for the Telemedicine Development Group (TDG) to be rebranded as DHM (that is, Digital Health Malaysia) ... to support develop-

In Singapore, some DH systems have been designed to move care beyond the hospital and provide valuable services to the community.

ment, research, and innovation of DH initiatives in Malaysia.”¹¹

The DG further stated, “These (DH tools and interventions) are what we need to embrace digital healthcare ... to transform the current healthcare system to one of high impact, reasonable cost, with great outcomes. ... The Ministry of Health shall provide its network of hospitals and health clinics as testbeds for ... digital innovations ...”

This plan is in alignment with greater recent Internet penetration in the country, allowing wearables and mobile devices to be more widely used to track behavioral aspects of healthcare.¹³ While there have been some public concerns expressed about privacy in DH, in general, it does not seem to be a major concern, according to Yasin.

Singapore

In Singapore, the social license for DH is quite clear, according to leading expert Adam Chee, a local representative of WHO’s international DH experts.¹⁵ Chee contends that Singapore’s citizens are generally savvy when it comes to digitalization and do understand the benefits of DH. This relative maturity is reflected in the actions of the Singapore government, which is transforming the country into a hotbed for research, innovation, and enterprise in smart healthcare systems. The practice is in line with the country’s Healthcare 2020 Masterplan¹⁰ by which it aims to change its healthcare landscape through the mantra of “better health and better future for all.” The key goals of the plan are to deliver accessibility, quality, and affordability in healthcare.

In turn, the local IT industry and health sector have driven various e-health or DH plans.^a Amid the recent growth, some DH systems have been designed to move care beyond the hospital and provide valuable services to the community (for example, the Multi-Dose Medication Management (MMM) system). Some other projects have also been introduced to take the concept of “health” beyond traditional mass healthcare and redefine it as the “healthiness of people” (for example,

a IHIS Yearbook 2016–2017; https://www.ihis.com.sg/About_IHiS/Pages/year_book.aspx

Smart Health TeleRehab).^b

Despite a recent analysis of Singapore’s “e-health” by a major European collaborative^c that described the nation state as the “undisputed continental leader” in the area, Chee argues that Singapore is still somewhat at the “upper-middle level” in terms of achievement. He says that “although we observe a heavy adoption of DH in various isolated areas, there still exists a need to better integrate (not just interface) the solutions at a patient (and clinical user) level. There is also a deep concern regarding data governance, security, and privacy given recent incidents in the healthcare sector.”

In order to further strengthen the DH environment, Chee thinks a decade is a rather short time to effect any major structural changes, but he notes that certain aspects of traditional healthcare are being changed in Singapore, including management of health data and rehabilitation. More significantly, he expects that creative start-ups will play more of a role, for example, through utilizing cutting-edge AI.

Australia

Having considered two very different DEHEs in Malaysia and Singapore, let’s now consider the situation in Australia, where there has been a long history of achievements in health IT and DH, but arguably in relative silos.⁶ In 2017, the Australian Digital Health Agency (ADHA) prioritized several areas in its national strategy to be achieved by 2020.²

While DH in Australia is at an exciting stage—with an ever-increasing number of healthcare providers moving onto electronic platforms in support of clinical care, and the presence of a vibrant start-up ecosystem (for example, see DoseMe^d and Life Whisperer^e)—it’s still something of a “tale of two nations.” For example, the general practice sector has enjoyed approximately 90% clinical computerization for many years, yet many government and privately funded hospitals still do not have EMRs.

The My Health Record (MyHR)^f roll out, using an opt-out consent model,

b HealthHub; <https://www.healthhub.sg/about-us>

c <http://bit.ly/2ZJp2ry>

d <https://doseme-rx.com/>

e <https://www.lifewhisperer.co/>

f <https://www.myhealthrecord.gov.au/>

recently generated fierce debate, with various interest groups seeking to influence the choices faced by Australian citizens. Many of the concerns that were highlighted in public conversation were in the areas of data custodianship and privacy. Ultimately, about 90% of Australian citizens have not opted out having a MyHR.

This debate highlights the social licence for moving to a fully expressed DH future in Australia is not clearly established, especially when one considers MyHR is just one part of that ultimate DEHE. Such a future, with various DH systems (for example, Alive Cor by Kardia^g) and DH interventions (for example, Propeller Health^h) routinely used to provide safe and efficient care would still seem some way off given they require trust in third parties around personal data custodianship in order to have widespread uptake.


We do anticipate, however, that the situation in Australia will be substantially different in 10 years due to a spirit of entrepreneurship and innovation among local experts working in DH—especially among graduates moving into healthcare. These individuals are showing an interest in DH and what it may offer them and their patients. They will, however, be constrained to some extent by the influence of the status quo in terms of poorly designed remuneration models for clinicians and other critical inhibitors.

Summary

The development of DH is distinctive in Australia, Singapore, and Malaysia due to variations in their IT infrastructure, demographics, consumer expectations, economic, and cultural settings. When extrapolated more broadly, this suggests a diversity in the evolution of DEHEs in Oceania.

While Singapore compares favorably by some measures, the Australian and Malaysian DEHEs are not as advanced as in the U.S.—the long-standing home of the Personal Connected Healthcare Alliance,ⁱ for example; or in Israel, having exported a number of DH products to the U.S.^{j,k}

In the three nations examined, challenges remain in the integration of DH

systems that, if organized well, could facilitate improved wellness healthcare for all citizens. Issues of data custodianship and the maintenance of privacy have also arisen in at least two of the nations. The ultimate challenge for all three, however, is in turning promising levels of technical achievement into sustained improvements in healthcare. 

i <https://www.pchalliance.org/>

j <http://bit.ly/2SXH2gP>

k <https://prn.to/2QM6jb4>

References

1. Alami, H., Gagnon, M.P., Wootton, R., Fortin, J.P. and Zanaboni, P. Exploring factors associated with the uneven utilization of telemedicine in Norway: A mixed methods study. *BMC Medical Informatics and Decision Making* 17, 1 (2017), 180.
2. Australian Digital Health Agency. Australia's National Digital Health Strategy; https://conversation.digitalhealth.gov.au/sites/default/files/adha-strategy-doc-2ndaug_0_1.pdf.
3. Department of Health and Social Care. The future of healthcare: Our vision for digital, data and technology in health and care, 2018.
4. Government of Malaysia Ministry of Health. Malaysia's Telemedicine Blueprint: Leading health care into the information age. Kuala Lumpur, 1997; <http://www.moh.gov.my/moh/resources/auto%20download%20images/5ca1b20928065.pdf>
5. Government of Malaysia Ministry of Health. Malaysian Health Data Warehouse (MyHDW). 2011–2013; <http://www.moh.gov.my/moh/images/gallery/publications/myhdw%202011-2013.pdf>.
6. Hambleton, S.J., and Aloizos, J. Australia's digital health journey. *General Practice* 185 (2006), 84–87.
7. Kamal, J.I.A. Implementation of electronic medical records in developing countries: Challenges and barriers. *Development* 7, 3 (2018).
8. Lau, F. From siloed applications to national digital health ecosystems: A strategic perspective for African countries. *Improving Usability, Safety and Patient Outcomes with Health Information Technology: From Research to Practice*, 2019, 404.
9. Scottish Government. Scotland's digital health and care strategy—enabling, connecting and empowering, 2018; <http://bit.ly/2FfiB6j>
10. Singapore Healthcare Masterplan 2020; <http://bit.ly/2ZKONaW>
11. Speech by Malaysia's Director General of Health; <http://healthcaredtoday.com.my/telemedicinedg.html>
12. Swindells, M. The NHS IT Strategy. NHS England, 2017; <https://www.england.nhs.uk/blog/the-nhs-it-strategy/#draft-placemat/>
13. World Health Organization. Meeting on Strengthening Health Information Systems for Sustainable Development Goals and Universal Health Care Monitoring in the Western Pacific Region, Manila, Philippines (22–24 Jan. 2019).
14. World Health Organization. WHO Definition of Digital Health; <http://bit.ly/36l8J60>
15. World Health Organization. WHO DH Roster of Experts; <http://bit.ly/2u6gqQ2>
16. World Health Organization. WHO Global Observatory for eHealth, 2016. Atlas of E-Health Country Profiles: The Use of E-Health in Support of Universal Health Coverage: Based on the Findings of the Third Global Survey on E-Health, 2015.
17. World Health Organization. WHO Guideline: Recommendations on digital interventions for health system strengthening: Web supplement 2: Summary of findings and GRADE tables (2019)

Chris Bain is a professor of Digital Health in the Faculty of IT at Monash University, Melbourne, Australia.

Abraham Oshni Alvandi is a research assistant in Digital Health in the Faculty of IT at Monash University, Melbourne, Australia.

g Kardia by Alivecor; <https://www.alivecor.com/>

h <https://www.propellerhealth.com/>



DH development is distinctive in Australia, Singapore, and Malaysia due to variations in their IT infrastructure, demographics, consumer expectations, economic, and cultural settings.



BY MEEYOUNG CHA /KAIST, WEI GAO /SINGAPORE MANAGEMENT UNIV., AND CHENG-TE LI /NATIONAL CHENG KUNG UNIV.

Detecting Fake News in Social Media: An Asia-Pacific Perspective

IN MARCH 2011, the catastrophic accident known as “The Fukushima Daiichi nuclear disaster” took place, initiated by the Tohoku earthquake and tsunami in Japan. The only nuclear accident to receive a Level-7 classification on the International Nuclear Event Scale since the Chernobyl nuclear power plant disaster in 1986, the Fukushima event triggered global concerns and rumors regarding radiation leaks. Among the

false rumors was an image, which had been described as a map of radioactive discharge emanating into the Pacific Ocean, as illustrated in the accompanying figure. In fact, this figure, depicting the wave height of the tsunami that followed, still to this date circulates on social media with the inaccurate description.

Social media is ideal for spreading rumors, because it lacks censorship. Confirmation bias and filter-bubble effects further amplify the spread of unconfirmed information. Upon public outcry, independent fact-checking organizations have emerged globally, and many platforms are making efforts to fight against fake news. For example, the state-run *Factually* website in Singapore has been known to clarify falsehoods since its inception in May 2012, which was followed recently by the implementation of the Protection from Online Falsehoods and Manipulation Act (POFMA) in October 2019. In Taiwan, the government officially created a feature on the website of the Executive Yuan (the executive branch of Taiwan’s government) to identify erroneous reporting and combat the spread of fake news. Taiwan’s Open Culture Foundation has also developed and introduced the well-known anti-fake fact-checking chatbot *Cofacts* in May 2018. The Indonesia government since 2018 has held weekly briefings on hoax news; that same year, the country revised its Criminal Code to permit the imprisonment for up to six years of anyone spreading fake news. Governments in the Asia and Oceania region, including South Korea, Singapore, Japan, Taiwan, Philippines, Cambodia, Malaysia, have enacted relevant laws to prevent fake news from spreading.

Nonetheless, fact-checking of fake news remains daunting, and requires tremendous time and effort in terms of human investigation. Moreover, it is prone to low efficiency and inadequate coverage due to the complexity of the topics being checked, and is incapable of keeping up with the fast



The recent presidential election in Taiwan was fraught with fake news and disinformation, inflaming supporters on both sides.

production and diffusion of falsehoods online. This article will review some of the latest techniques to automatically debunk fake news, many of which were initiated in the Asia and Oceania region.

Research on understanding and debunking false information spans multiple disciplines, including social psychology, information management, and computer science. Computational approaches to automate fact-checking have attracted interest, especially in data mining, natural language processing, and artificial intelligence. Existing approaches primarily rely on training classifiers, for which past events or claims are gathered and labeled as real or fake, and significant

features are extracted to generate appropriate data representation. Recent techniques on deep learning further improve performance and enable the interpretability of fact-checking by representation learning and attention-based models. We categorize these methods into four streams.

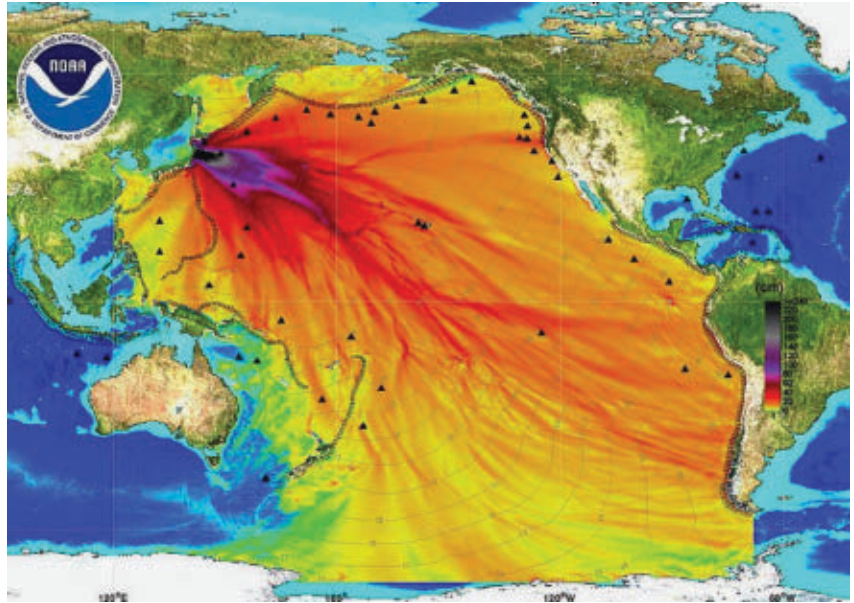
► *Feature-engineering method*: One of the first data-driven research projects on fake news was initiated in East Asia,³ which tried to build a classifier to debunk fake news using features crafted from temporal (for example, frequency of spikes over time), structural (density, clustering), and linguistic (for example, usage of negation and persuasion words) factors. Among them, linguistic features were found to be consistently

effective over short (that is, several days) and long (that is, two months) time.² Modeling time-evolving evidence as features based on social interactions can further boost performance.⁵ However, this stream requires ad-hoc processing of raw data for constructing features, which are painstakingly detailed, biased, and labor-intensive.

► *Matching-based method*: Assuming that false claims will spark responses from skeptical users who question their veracity, false rumors can be detected by searching and clustering tweets.¹¹ Text patterns of skepticism about factual claims are particularly useful, such as expressions like “Really?” While rule-based matching requires manual specification, semantic matching can

Governments in the Asia and Oceania region, including South Korea, Singapore, Japan, Taiwan, Philippines, Cambodia, Malaysia, have enacted relevant laws to prevent fake news from spreading.

A map falsely introduced as showing the spread of radioactive seepage from the Fukushima region (<https://www.snopes.com/fact-check/fukushima-emergency/>). The original image was produced by the National Oceanic and Atmospheric Administration, @NOAA.



be a viable alternative. False claims may be debunked by retrieving evidence from relatively reliable sources, such as checked news and Wikipedia articles, by an evidence-aware neural attention model that learns to highlight words related to the claim.⁷

► *Representation-learning method:* Deep neural networks, such as Recurrent Neural Networks (RNN), are exploited to learn latent representations of text content based on low-level features such as term frequency or word embeddings.⁴ Some modeled rumors as information campaigns through Generative Adversarial Networks (GAN).⁶ Recently, a co-attention network-based model was developed not only to find the correlation between posts and their comments for more accurate detection, but also to automatically identify which users, sentences, comments, and words contain fake signals.⁸

► *Multimodal method:* News spread in social media exhibits multiple modalities, such as text, image, and social context. Automatic fact-checking is trending to evidence collection and consolidation from multimodal information. Example methods include the RNN-based multimodal feature extraction and fusion model for rumor detection.¹ Adversarial neural networks have been developed for multimodal

fake news detection by learning an event's invariant representation,¹⁰ which removes tight dependencies of features on the specific events in the training data for a better generalization. Tri-relationship embedding models publisher-news relations and user-news interactions simultaneously to recognize fake news in the early stage of news dissemination.⁹

Technically, the approaches were developed from non-learning-based methods, traditional supervised learning methods, neural representation learning methods toward semi-supervised, and more recently, unsupervised methods. Content-wise, information was adopted from a single modality, such as pure text or images, toward the combination of multiple modalities. Semantically, prior methods began with shallow patterns, and hand-crafted features have been advanced by utilizing automatic feature learning, which is now tending to cross fact-checking using more sophisticated learning techniques across heterogeneous content and structures. For practical applications, the learning algorithm is shifting from employing a massive collection of user responses to relying on a limited set of observations, pursuing higher detection accuracy in the early stage of news propagation. Moreover, to reduce fake

news, stakeholders require models to provide explainable outcomes that highlight which users and publishers are creating fake news, on which topics, through what types of textual and social manners.

Research on fact-checking is still in its infancy. Existing approaches bear several noticeable limitations because social media content comprises different modalities that reflect the dynamics and diversity of current events. Supervised models trained entirely on historical events that consider different types of features as a single representation can cause overfitting when the training set is multifaceted. Furthermore, while most algorithms learn to assess and discriminate inquiries, viewpoints, and stances of users on newsworthy claims, the outcome can suffer from low recall, since not all falsehoods may spark responses on social media. Also, many fake responses are intentionally created by certain groups of users to fool or attack people from other groups. Since the responses could come from everyday users or just be pertinent, the information they convey is generally too noisy and subjective to deduce a reliable veracity assessment.

To encourage advanced research in fighting against fake news, Taiwan's Ministry of Science and Technology (MOST) has created a special call for relevant research projects for which it will provide funding support. South Korea's National Research Foundation (NRF) and the Japan Science and Technology Agency (JST) also financially support research projects on detecting and tracking disinformation. We believe more and more countries in Asia and Oceania region will devote resources to the war on fake news.

The fake news phenomenon is taking new turns. YouTube and instant messaging (IM) services (for example, Whatsapp, Kakaotalk) are emerging as hotbeds of fake news. According to a survey conducted by the Korea Press Foundation, 34% of Korean YouTube viewers report having watched or received videos containing fake news. Taiwan's Open Government Foundation g0v reported that in 2017 only 46% of chatbot responses on that nation's most popular IM app LINE is correct. Fake news on streaming platforms and

IM services is particularly concerning because it contains visual content, which is more persuasive than mere text posts. Also, IM may reinforce the credibility of fake claims because people are more likely to follow trusted social contacts blindly. Data synthesis techniques like GANs can produce high-quality fabricated videos of celebrities and politicians that may appear in fake news. Intelligent fact-checking chatbots that incorporate a GAN to retrieve and generate natural-language evidence and explanations, such as the Cofacts program in Taiwan, are being progressively developed and implemented on IM services to catch false information and prevent it from spreading instantly. Consequently, the task of detecting fake news in the era of big data, social media, and artificial intelligence calls for greater attention from the research community. ■


References

1. Jin, Z., Cao, J., Guo, H. and Luo, J. Multimodal fusion with recurrent neural networks for rumor detection on microblogs. In *Proceedings of ACM Multimedia 2017*, 795–816.
2. Kwon, S., Cha, M., Jung, K. Rumor detection over varying time windows. *PLOS ONE* 12, 1 (2017), e0168344.
3. Kwon, S., Cha, M., Jung, K., Chen, W. and Wang, Y. Prominent features of rumor propagation in online social media. In *Proceedings of ICDM 2013*, 1103–1108.
4. Ma, J. et al. Detecting rumors from microblogs with recurrent neural networks. In *Proceedings of IJCAI*, 2016, 3818–3824.
5. Ma, J., Gao, W., Wei, Z., Lu, Y. and Wong, K.-F. Detect rumors using time series of social context information on microblogging websites. In *Proceedings of CIKM*, 2015, 1751–1754.
6. Ma, J., Gao, W. and Wong, K.-F. Detect rumors on Twitter by promoting information campaigns with generative adversarial learning. In *Proceedings of WWW*, 2019, 3049–3055.
7. Papat, K., Mukherjee, S., Yates, A. and Weikum, G. DeclarE: Debunking fake news and false claims using evidence-aware deep learning. In *Proceedings of EMNLP*, 2018, 22–32.
8. Shu, K., Cui, L., Wang, S., Lee, D. and Liu, H. dEFEND: Explainable Fake News Detection. In *Proceedings of KDD*, 2019, 395–405.
9. Shu, K., Wang, S. and Liu, H. Beyond News Contents: The Role of Social Context for Fake News Detection. In *Proceedings of WSDM*, 2019, 312–320.
10. Wang, Y. et al. EANN: Event Adversarial Neural Networks for multi-modal fake news detection. In *Proceedings of KDD*, 2018, 849–857.
11. Zhao, Z., Resnick, P. and Mei, Q. Enquiring minds: Early detection of rumors in social media from enquiry posts. In *Proceedings of WWW*, 2015, 1395–1405.


Meeyoung Cha is an associate professor in the School of Computing, Korea Advanced Institute of Science and Technology and the Institute for Basic Science, South Korea.

Wei Gao is an assistant professor in the School of Information Systems at Singapore Management University.

Cheng-Te Li is an associate professor in the Institute of Data Science and Department of Statistics at National Cheng Kung University, Tainan, Taiwan.



Fact-checking of fake news remains daunting, requiring tremendous time and effort.



BY GERNOT HEISER /CSIRO DATA61/UNSW SYDNEY,
GERWIN KLEIN /CSIRO DATA61/UNSW SYDNEY,
AND JUNE ANDRONICK /CSIRO DATA61/UNSW SYDNEY

seL4 in Australia: From Research to Real-World Trustworthy Systems

TEN YEARS AGO, the functional correctness proof of the seL4 microkernel marked the first time a complete operating system (OS) kernel had been verified to the source-code level.⁴ This means there was a machine-checked proof that the implementation in the C language satisfied the kernel's specification, expressed in mathematical logic.

Much has happened since then: We have extended the verification to show the kernel enforces desired security and safety properties, we have removed

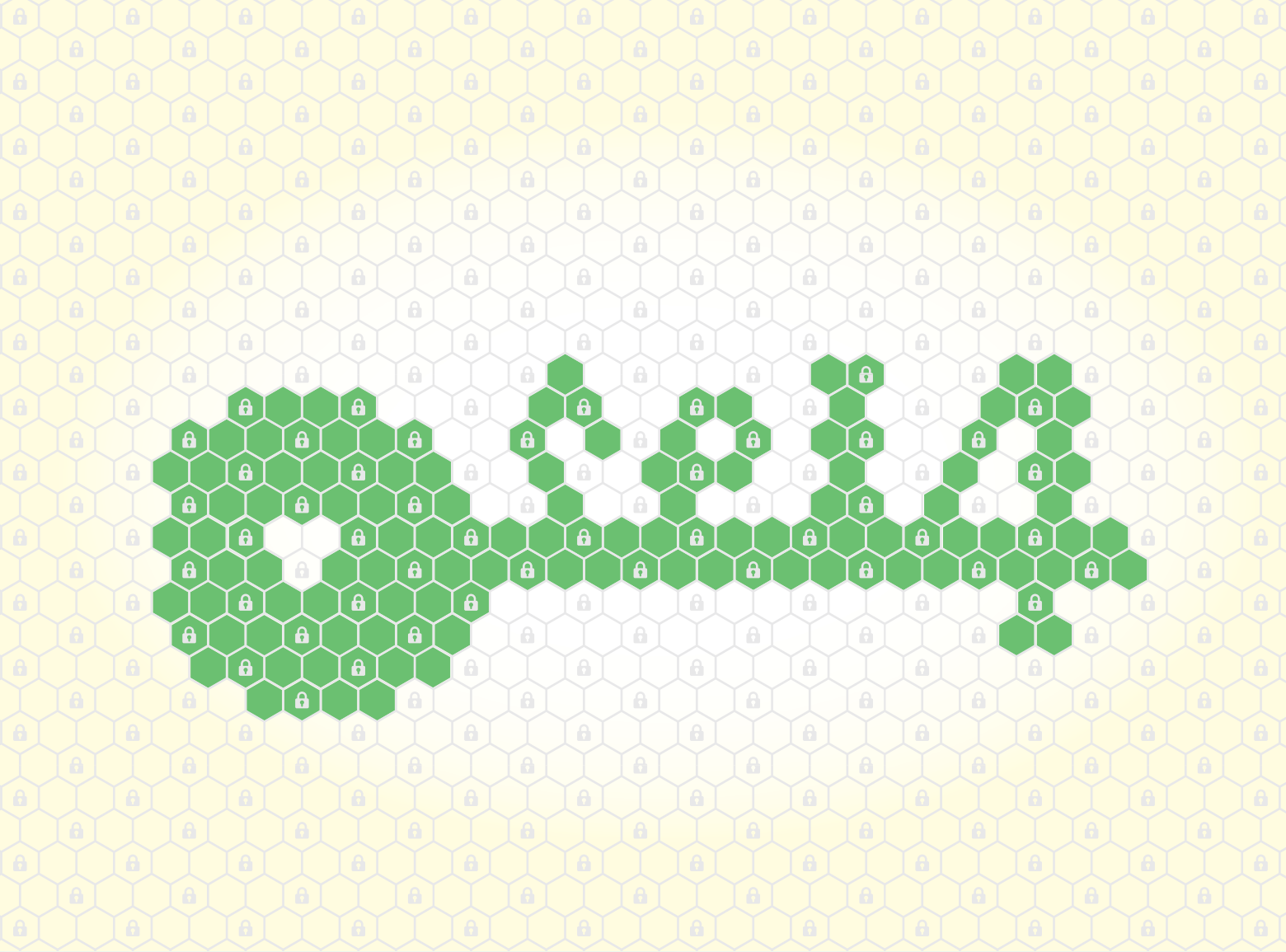
the need to trust the compiler, and we verified implementations for processor architectures other than the original Arm v6. We used experience from deploying seL4 in a number of real-world systems to evolve the kernel and its proofs to support a broader class of use cases, and we have made significant progress toward extending the assurance to systemwide properties. We will provide a brief overview of these developments, as well as ongoing research.

What is seL4?

An operating system consists of a kernel, the software that runs in the privileged mode of the hardware, and system services that are programs running in the unprivileged hardware mode (user mode). A microkernel is the irreducible part of the kernel: It only contains code that must run privileged. The kernel is the most dangerous part of the system: if anything goes wrong here, there is nothing to protect the rest of the system. Microkernels minimize the kernel to maximize the chance of getting it right.

To make this work, the microkernel must be very general in what it provides—it must allow constructing arbitrary system functionality on top. As such, microkernels strive to be *policy free*, and provide only general mechanisms that can be used to implement arbitrary policies. For example, the microkernel does not have a notion of a security policy, and it does not provide a file system (with its implied policies of access control and management of persistent storage), it does not provide a process model. Instead it provides primitive notions of address spaces, threads of execution, and low-level communication primitives (referred to as IPC).

Minimality and policy freedom combined with high performance have been the defining design principle of the L4 family of microkernels, which go back to the mid 1990s. seL4, developed 2004–2009, takes policy freedom to a new level: It does



not even manage physical memory; the kernel has no heap, and user-level managers must provide the kernel with memory to store meta-data, such as page tables and thread control blocks. Kernel memory thus becomes subject to memory-management policies defined at the user level. An implication is that user-level memory partitioning extends into the kernel, making it easier to reason about isolation in an seL4-based system.

Besides that, seL4 continues the trademark performance focus of the L4 microkernel family, outperforming any other OS in terms of message-passing cost, despite its strong assurance story. This is summarized in our principle “security is no excuse for poor performance.”

Verification: More Properties, More Platforms

The original functional-correctness proof was between an operational

model of the kernel and the C source of its implementation. It applied to 32-bit Arm v6 processors. We later reverified it for the Arm v7 architecture (a minor change). More importantly, we developed a translation validation tool chain, which used a formalization of the Arm instruction-set architecture (ISA) developed at Cambridge to prove the executable binary produced by the GCC compiler and linker have the same semantics as the C code, that is, the compilation did not introduce bugs. This took the compiler, as well as our assumptions on C semantics, out of the trusted computing base (TCB).

We also proved the kernel model had desirable isolation properties, specifically that in a suitably configured system, seL4 can enforce confidentiality, integrity, and availability using capability-based access control, although the confidentiality notion does not cover timing channels. This

implies that seL4 enables the construction of systems that are secure in a well-defined sense.

Furthermore, we performed a sound and complete worst-case execution time (WCET) analysis of the kernel for specific Arm v6 processors, the first published case of such an analysis for a protected-mode OS. This means it is possible to reason about the timeliness of hard real-time systems built on seL4. Klein et al.⁵ offers a detailed description of these verification achievements.

We have since ported and verified the kernel on other architectures: The functional correctness proof for the 64-bit x86 architecture was completed in 2018, while proof of functional correctness, as well as translation correctness for 64-bit RISC-V processors, is scheduled to complete in early 2020. We verified kernel extensions for hardware-supported virtualization on Arm v7 in 2017, making seL4 a fully verified hypervisor.

Evolving the Kernel

While based on more than a decade of experience with earlier L4 kernels,³ seL4's system model differs from these in a number of ways. The most obvious is its radical separation of policy and mechanism, by delegating all spatial resource management to user level. Experience with this model led to a number of incremental changes, which resulted in re-verifying the evolved kernel.

However, the original kernel's most significant shortcoming was its handling of time, which it had inherited from its L4 ancestry—a highly simplistic and somewhat unprincipled scheduling model. We had earlier flagged this as the last outstanding significant kernel-design issue.³

We solved this problem by introducing *scheduling contexts*, which extend capability-based access control to time.⁷ Other than concurrent work at Washington University, this represents the first capability model of time that is suitable for hard real-time systems and mixed-criticality systems, and it achieves this without compromising seL4's trademark superior performance. The implementation of this revised resource-management model is presently undergoing verification, which we expect to complete in early 2020.

Security by Architecture: Building Trustworthy Systems

The seL4 mechanisms are powerful, but also low-level; this is the cost of taking policy freedom further than ever before. A simple (if not trivial) system, consisting of one process invoking another, already consists of about 50 seL4 objects that user-level code must manage. Such a system is depicted in the gray-background box on the LHS of the accompanying figure. It shows two processes, each represented by a thread object that contains scheduling parameters and capabilities to other objects: a Cspace made up of Cnodes that store capabilities to objects the process has the right to access, and a Vspace, which describes the process's memory map (a thin abstraction over page tables). The two processes each have a capability to an Endpoint (communication port) object, with Send right for the

one process, and Receive right for the other, which allows the processes to communicate (in one direction).

Building practical systems calls for a higher abstraction level (which will inevitably introduce policy and thus might be domain-dependent). Developers think in terms of architectures represented by boxes, representing data/active entities, and lines connecting them, representing communication. Such an abstraction is presented by our CAMkES component framework, whose boxes are processes with well-defined and seL4-enforced interfaces, an example is shown at the top of the figure. CAMkES allows presenting security policies architecturally: The security policy is observed if we can guarantee that communication can only happen where it is explicitly allowed by the architecture, that is, along lines connecting boxes.

We support this by a number of certified, automatic transformations. The developer specifies the architecture in the CAMkES architecture-description language (ADL). The first transformation (box arrow on the right) generates from the ADL the communication “glue” code, that is, the seL4 system call invocations that provide the communication indicated by the lines in the architecture specification. The generated C code is compiled and linked together with the code implementing the functionality of the boxes

to create the system image.

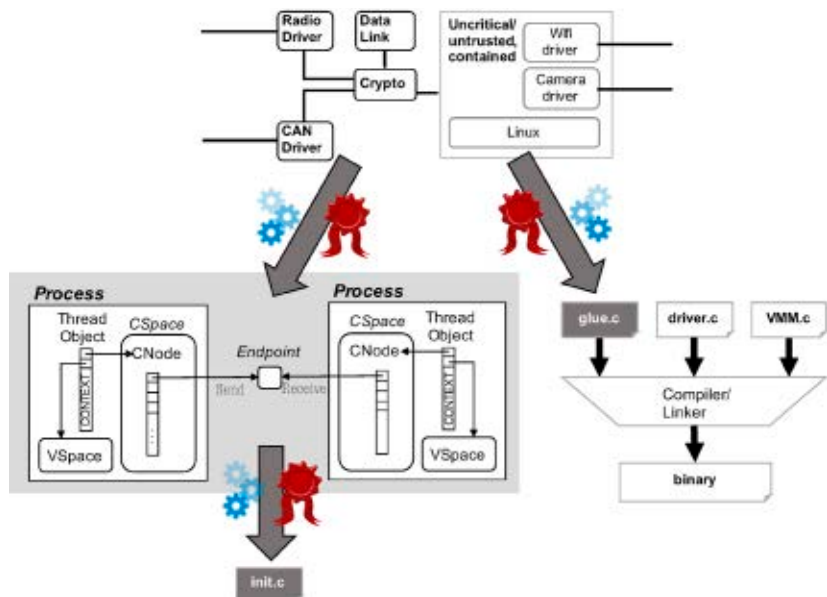
A second language—CapDL—describes the capability distribution in a system. We generate the low-level representation of the system by compiling the CAMkES ADL spec into CapDL (top left box arrow). The CapDL representation is the low-level representation (gray box) discussed earlier. For clarity, we only show a small part of the CapDL representation of the architecture at the top (corresponding to just two of the boxes and one connecting line).

The architecture compilation comes with a proof procedure, which ensures the CapDL spec only allows the interactions specified at the ADL level.

Finally, our CapDL tool chain generates from that representation the system initialization code that produces the respective seL4 objects and assigns capabilities (bottom left box arrow). This CapDL initializer is currently verified at the model level, with code verification under way. Together these tools guarantee the user-level initialization code boots the system into a state that is correctly represented by the CAMkES specification, and thus allows reasoning about security properties at that architectural level.⁶

CAMkES components can in principle be implemented in any programming language. Security/safety-critical, that is, *trusted*, components should be verified to ensure their trustworthi-

Architecture enforcement. Parts in gray are generated with assurance.



ness. This is possible for C programs, as demonstrated by the verification of seL4, but easier and more cost effective in higher-level languages, especially functional languages that guarantee type and memory safety. We employ two such languages—CakeML and Cogent.

CakeML⁸ is an ML dialect which comes with a certifying compiler as well as a runtime system that has been proved functionally correct. It is a fully fledged functional language suitable for implementing a wide class of applications, and its semantics supports verification of programs in the HOL4 and Isabelle proof assistants. We are using CakeML to build verified components of application code, as well as in the CapDL initializer.

Cogent is a simpler language, aimed at implementing systems code.¹ For performance predictability it is non-managed (has no garbage collector), yet the type system ensures memory safety, similar to Rust. The language is intentionally Turing-incomplete, which means it is supplemented by a library of (manually verified) abstract data types. The Cogent compiler produces C code, together with an Isabelle embedding of the program semantics and a proof certificate that the C code has the same semantics. We have used Cogent to implement file systems, simple device drivers, as well as protocol layers.

Real-World Deployments

We open sourced seL4 in 2014 under the GPL v2 license. This led to the kernel being designed into a number of security- and safety-critical systems. Among them are autonomous aerial and land vehicles under the DARPA HACMS program, where together with project partners we performed an *incremental cyber retrofit*, re-architecting easily compromised real-world systems into something that resisted attacks by professional penetration testers.⁶

Other uses are a secure communication device that has been security-evaluated and approved for defense use up to *secret*, and a military cross-domain device able to securely connect to networks of different classifications (which is undergoing security evaluation). A number of safety-critical

systems are under development, including medical devices, autonomous passenger cars, and industrial control systems.

Present Activities

In order to scale up community engagement we are setting up an *seL4 Foundation*, similar to those for the Linux and RISC-V communities. We expect the setting-up process to be completed some time in 2020. The aims of the Foundation include a broad-based membership, commercial sponsorship, and an acceleration of the provision of components and tools for seL4 developers. This complements our engineering efforts aimed at providing generic frameworks and infrastructure, to lower the bar of adopting seL4 for building trustworthy systems in a wide class of application domains.

Ongoing research activities include incorporating and verifying time protection mechanisms for principled prevention of timing channels.² We are also working on verifying the multicore version of seL4, a challenging task due to the fact that for performance reasons the kernel is rich in data races. Finally, we are working on our ultimate goal: Proving security or safety of a complete, real-world system. This will require a combination of all the building blocks discussed here, as well as significant further innovation in formal analysis and proof techniques.


Systems Verification Around the World

seL4 was the first general-purpose operating system (or any significant system component) whose implementation was proved correct. It triggered a wealth of projects on verifying systems code but is still unique in not compromising on performance and being designed for real-world use. This has just been recognized by the ACM SIGOPS Hall of Fame award.

Some other high-profile verification projects include work at MIT, on proving integrity properties of file systems in the presence of crashes, and most recently, tackling some of the concurrency issues in file systems—verifying concurrent systems is extremely difficult and will take time to fully solve.

The CertiKOS kernel at Yale showed that formal OS kernel verification is possible for multicore processors in principle, at least in a simpler, more restricted hypervisor/separation kernel setting that sacrificed performance for easier verification. The main verification technique used there does not directly extend to the more complex seL4 kernel, but some of the underlying ideas can be potentially applied.

A group at the University of Washington is specializing on fully automated verification of systems code, such as file systems and simple OS kernels, and is exploring the limits of what automated verification can achieve.

Researchers at Microsoft demonstrated in the IronClad and IronFleet systems that verified OS-level properties can be lifted up to an entire distributed system and its protocols, albeit still assuming correctness of network drivers and communication. Nevertheless, this work shows our goal of systemwide properties on a verified kernel can be achieved in principle. 

References

1. Amani, S. et al. Cogent: Verifying high-assurance file system implementations. In *Proceedings of the 2016 Intern. Conf. Architectural Support for Programming Languages and Operating Systems*. Atlanta, GA, USA, 175–88.
2. Ge, Q., Yarom, Y., Chothia, T. and Heiser, G. Time protection: The missing OS abstraction. In *Proceedings of the 2019 ACM EuroSys Conf.*, Dresden, Germany.
3. Heiser, G. and Elphinstone, K. L4 microkernels: The lessons from 20 years of research and deployment. *ACM Trans. Computer Systems* 34, 1 (2016), 1:1–1:29.
4. Klein, G. et al. seL4: Formal verification of an operating-system kernel. *Commun. ACM* 53, 6 (2010), 107–15.
5. Klein, G., Andronick, J., Elphinstone, K., Murray, T., Sewell, T., Kolanski, R. and Heiser, G. Comprehensive formal verification of an OS microkernel. *ACM Trans. Computer Systems* 32, 1 (2014) 2:1–2:70.
6. Klein, G., Andronick, J., Kuz, I., Murray, T., Heiser, G. and Fernandez, M. Formally verified software in the real world. *Commun. ACM* 61, 10 (Oct. 2018) 68–77.
7. Lyons, A., McLeod, K., Almatary, H. and Heiser, G. Scheduling-context capabilities: A principled, light-weight OS mechanism for managing time. In *Proceedings of the 2018 ACM EuroSys Conf.* Porto, Portugal.
8. Tan, Y.K., Myreen, M., Kumar, R., Fox, A., Owens, S. and Norrish, M. The verified CakeML compiler backend. *J. Functional Programming* (Feb. 29, 2019).

Gernot Heiser is chief research scientist at CSIRO Data61 and Scientia Professor and John Lions Chair at UNSW Sydney, Australia.

Gerwin Klein is chief research scientist at CSIRO Data61 and a conjoint professor at UNSW Sydney, Australia.

June Andronick is chief research scientist at CSIRO Data61 and a conjoint associate professor at UNSW Sydney, Australia.

Copyright held by authors/owners.
Publication rights licensed to ACM.

RAPHAËL CW PHAN
Monash University, Malaysia

MASAYUKI ABE
NTT, Japan

LYNN BATTEN
Deakin University, Australia

JUNG HEE CHEON
SNU, Korea

ED DAWSON
QUT, Australia

STEVEN GALBRAITH
University of Auckland, New Zealand

JIAN GUO
NTU, Singapore

LUCAS HUI
ASTRI, Hong Kong

KWANGJO KIM
KAIST, Korea

XUEJIA LAI
SJTU, China

DONG HOON LEE
Korea University, Korea

MITSURU MATSUI
Mitsubishi Electric, Japan

TSUTOMU MATSUMOTO
YNU, Japan

SHIHO MORIAI
NICT, Japan

PHONG NGUYEN
University of Tokyo, Japan

DINGYI PEI
Guangzhou University, China

DUONG HIEU PHAN
University of Limoges, France

JOSEF PIEPRZYK
CSIRO Data61, Australia

HUAXIONG WANG
NTU, Singapore

HANK WOLFE
University of Otago, New Zealand

DUNCAN WONG
CryptoBLK, Hong Kong

TZONG-CHEN WU
NTUST, Taiwan

BO-YIN YANG
Academia Sinica, Taiwan

SIU-MING YIU
HKU, Hong Kong

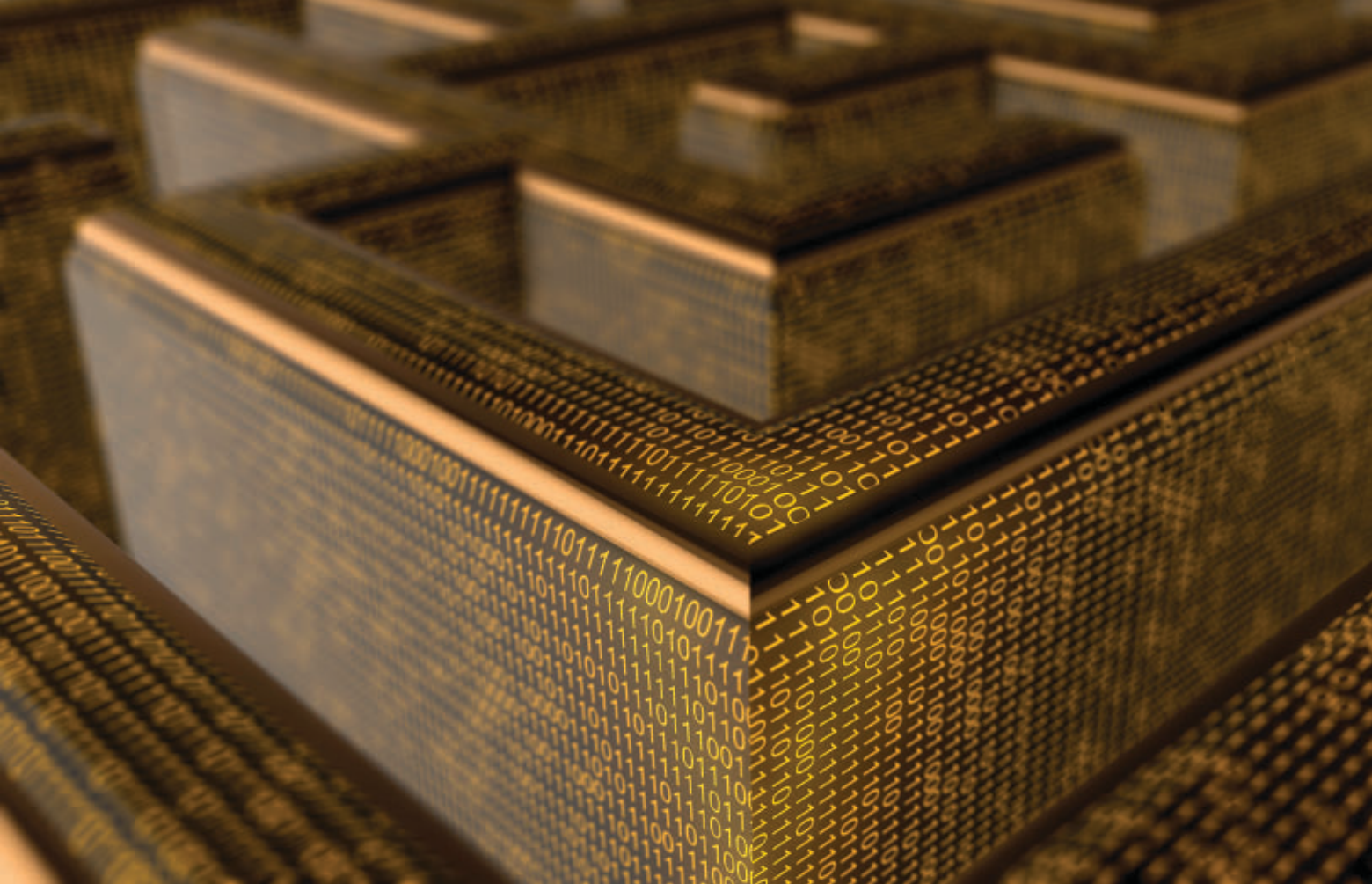
YU YU
SJTU, China

JIANYING ZHOU
SUTD, Singapore

Members of the International Association for Cryptologic Research explore regional work and collaboration activities.

Advances in Security Research in the Asiacrypt Region

THE CRYPTOGRAPHIC AND security research community is very closely knit in the sense that irrespective of which country we are in, or which region we come from, we are aligned to a single formal association—the International Association for Cryptologic Research (IACR). Regional flagship IACR conferences allow fellow researchers to gather regularly. The three main



regions hosting security research conferences are the U.S. (Crypto), Europe (Eurocrypt), and Asia + Oceania (Asiacrypt). These events provide a sense of belonging and are one main reason researchers continually involve themselves in such regional activities. In this article, we focus on regional collaborative activities as well as highlight each country's security research.

Asiacrypt. In 1990, the first Auscrypt was held in Sydney, spearheaded by Jennifer Seberry and Josef Pieprzyk, both of UNSW. The aim was to have a regional collaborative venue similar in style to Crypto and Eurocrypt, both of which started in the 1980s. In 1991, Japanese cryptographers initiated Asiacrypt in Fujiyoshida as the first crypto conference in Asia. It was organized by Hideki Imai (YNU), Ron Rivest, and Tsutomu Matsumoto. After Auscrypt 1992 at Gold Coast, Australian cryptographers combined Auscrypt into Asiacrypt by hosting Asiacrypt 1994 in Wollongong. As a bi-annual conference, Asiacrypt 1996 and Asiacrypt 1998 were hosted in Korea and China, respectively; followed by Asiacrypt 1999

in Singapore. Subsequently a strong suggestion was made to IACR by the Asiacrypt Steering Committee for the conference to be held annually. As a result, Asiacrypt has been the annual IACR flagship conference for the Asian region since 2000 and holds the same status within the industry as the Crypto and Eurocrypt conferences. Each year, more than 200 papers are submitted to Asiacrypt of which about one-third are selected by multiple peer review to be published in the proceedings. Attendance on average has been over 250 participants.

Australia and New Zealand. In 1996, the first Australasian Conference on Information Security and Privacy (ACISP) was held in Wollongong and organized by Josef Pieprzyk and Jennifer Seberry. The event is designed to complement, not to compete with, Asiacrypt. The concept of an Australia-New Zealand security conference was to encourage researchers to communicate and provide feedback to develop strong papers suitable for Asiacrypt and journals in various areas of information security and privacy. The venues move around universities and cities

in Australia and New Zealand in order to encourage inclusiveness.

Japan and Korea. There are workshops jointly organized with countries in Asia. The Japan-Korea Joint Workshop on Information Security and Cryptology (JW-ISC) was initiated by major research institutes in Japan and Korea in 1993. It was later extended to the Asia Joint Conference on Information Security (AsiaJCS) in 2005. These events provide opportunities for young researchers in Asia to extend their activities by giving talks to international audiences and exchanging their ideas with invited experts.

Country-Specific Initiatives for Security and Crypto

China. In 2005, the State Key Laboratory of Information Security (SKLOIS) of the Chinese Academy of Sciences organized the first international conference on Information Security and Cryptography (Inscrypt, formerly CISC) in Beijing. ChinaCrypt is another major crypto event organized by the Chinese Association for Cryptologic Research (CACR) since 2007.

ChinaCrypt is the biggest regional

Regional collaboration map based on IACR flagship conferences (CRYPTO, Eurocrypt, Asiacypt) for the last six years 2019–2014.

| | 2019 | 2018 | 2017 | 2016 | 2015 | 2014 |
|------------------|------|------|------|------|------|------|
| Australia | C | | C | C | | S |
| China | AHJS | JKS | AJS | AJS | J | JKS |
| Hong Kong | C | | T | | | |
| Japan | C | CS | CS | CS | C | CS |
| Korea | | C | | | | CS |
| Singapore | CT | CJ | CJ | CJ | | ACJK |
| Taiwan | S | | H | | | |

forum for Chinese researchers and practitioners in cryptography where talks are given in Chinese except for non-Chinese invited speakers. At each conference, organizations on behalf of their cities bid to host the next year’s ChinaCrypt, voted by the CACR committee. In 2018, CACR initiated the National Cryptographic Algorithm Design Competition to solicit and evaluate innovative cryptographic algorithms nationwide. In September 2019, 24 proposals advanced to the second round; and winners will be announced in 2020. The Chinese National Standards for cryptographic algorithms currently in use include the SM2, SM3, SM4, and SM9 that correspond to algorithms for public-key cryptography, cryptographic hash functions, block ciphers, and identity-based cryptography, respectively. SM2, SM3, and SM9 have acquired international recognition under ISO/IEC14888-3/AMD1 and ISO/IEC10118-3:2018.

Hong Kong. The Centre for Information Security and Cryptography (CISC) was formally established in HKU in 1998. Since then, cryptography became a major research area in

Hong Kong. Besides contributions to many fundamental cryptographic primitives, researchers in Hong Kong were among the first to apply cryptographic techniques in various application areas such as education, digital forensics, privacy-preserving computations in VANET (vehicular ad hoc network), smart grids, bioinformatic computations, and recently on blockchains and cryptocurrencies. In 2017, Asiacypt was held in Hong Kong for the first time.

Japan. The three-year Cryptography Research and Evaluation Committees (CRYPTREC) was organized in 2000 to evaluate cryptographic techniques publicly applied and widely used in industries and select those that excel in security and implementation. In 2003, the Ministry of Internal Affairs and Communication and the Ministry of Economy, Trade, and Industry publicized the list of ciphers recommended for the procurement of e-government. Annual symposiums, known as the Symposium on Cryptography and Information Security (SCIS) since 1984 and the Computer Security Symposium (CSS) since 1998, are organized

by local research communities, the Institute of Electronics, Information and Communication Engineers (IEICE) and Information Processing Society (IPSJ) in Japan, respectively. These symposiums have jointly attracted over 1,200 participants in recent years.

Korea. The Korea Institute of Information Security and Cryptology (KIISC) was established in 1990 by the government to promote the academic advancement of information security and cryptology. KIISC organizes two domestic annual conferences every summer and winter, and since 1998 timed the annual International Conference on Information Security and Cryptology (ICISC) to occur a week before Asiacypt. In 2000, KIISC began hosting the Workshop (now renamed World Conference) on Information Security Application (WISA) to bridge academia and industries, focusing on the practices and applications of information security and cryptology. WISA has been consistently co-sponsored by the Ministry of Science and ICT (MSIT), the Korea Internet and Security Agency (KISA), the National Security Research Institute (NRI), the Electronics and Telecommunications Research Institute (ETRI) and the leading domestic security industries. Korea has its own block-cipher standards—SEED, ARIA, and HIGHT—for domestic applications listed in the international standard. KIISC has hosted Asiacypt (1996, 2004, 2011), FSE 2010, and CHES2014, and will be hosting Asiacypt2020 and PQCrypto2021.

Malaysia. The inaugural Mycrypt



Asiacypt 2019, Kobe, Japan.



Asiacypt 2020, Daejeon, South Korea.

was pioneered by Raphaël Phan in 2005 to expose the local security community to the culture of international crypto conferences. Held in Kuala Lumpur, the program chairs were Serge Vaudenay and Ed Dawson (QUT) and featured an unconventional cryptography session spotlighting a paper on the new notion of questionable encryption by Adam Young and Moti Yung. Subsequently, Phan led the bidding team to host Asiacrypt in Malaysia for the first time in 2007. This was held in Kuching in the Borneo state of Sarawak. Mycrypt was rejuvenated in 2016, focusing on paradigm-shifting unconventional crypto, with Phan and Yung as the program chairs. In terms of government initiatives, the Malaysian government agency CyberSecurity Malaysia initiated in 2016 an effort similar to CRYPTREC (Japan) to propose a list of trusted cryptographic algorithms for Malaysia. MySEAL produced a recommended list of existing algorithms in 2017.

Singapore. In 2018, two multidisciplinary research centers each funded by a \$10 million grant from the National Research Foundation of Singapore were set up under Nanyang Technological University (NTU) and National University of Singapore (NUS), respectively. The centers are to develop capabilities, technologies, and skilled manpower toward scalable and customized privacy preserving technologies that are aligned with national priorities of Singapore in the services and digital economy of the Research, Innovation, and Enterprise (RIE2020) plan.

Taiwan. Taiwan's burgeoning cryptographic community hosted Asiacrypt 2003 and 2014. Taiwan is known for specialties in post-quantum cryptography and cryptographic implementations. For example, Bo-Yin Yang (Academia Sinica) was the co-inventor of Ed25519, a widely used elliptic curve digital signature scheme, which is a de facto standard on the Internet, and of which the U.S. National Institute of Standards and Technology (NIST) has said will be part of the U.S. standard FIPS 186-5. Taiwanese scholars contributed to the multivariate digital signatures MQDSS and Rainbow in the second

round of the NIST post-quantum standardization process.

Vietnam. The Vietnam Cryptographic Branch was formed in 1945 soon after independence, and the development of cryptography was exclusively realized by the government's secret agencies such as the Ban Co Yeu (Cryptographic Bureau). The first official collaboration between a secret agency and a public research institute occurred in 1987 in a project called M-87, led by Phan Dinh Dieu (VNU)—the founder of the Vietnamese IT Society. Since then, research in cryptography has increased in public institutions as well as public universities where cryptography/security courses were created. The first international cryptology conference in Vietnam (Vietcrypt) was held in 2006, led by Vietnamese researchers abroad: Khanh Nguyen (Singapore), Phong Nguyen and Duong Hieu Phan (France) and Duy Lan Nguyen (Australia) and supported by Phan Dinh Dieu (as general chair). Vietcrypt attracted numerous prominent cryptographers including Jacques Stern, Tatsuaki Okamoto, Phil Rogaway, and inspired new young students to follow cryptography research. Subsequently, Vietnam hosted its first Asiacrypt in 2016, led by Duong Hieu Phan as one of the general chairs, along with Ngo Bao Chau (VIASM).

Security Research Highlights of the Asiacrypt Region

Australia. The pioneers of Australian crypto include Jennifer Seberry (Wollongong) and Ed Dawson (QUT) who have made research contributions to the field of symmetric cryptography. Yuliang Zheng (Monash) introduced the signcryption notion in 1997, which to date has been cited over 1,500 times. The aim of signcryption is to provide both confidentiality and authentication of messages more efficiently than by independent encryption and signing, by means of intertwining both cryptographic operations. Josef Pieprzyk (Macquarie) in joint work with Nicolas Courtois that has been cited over 1000 times, proposed a method to break ciphers whose S-boxes can be expressed by an over-defined system of algebraic equations. Desmedt,



Asiacrypt has been the annual flagship conference for the Asian region since 2000 and holds the same status within the industry as the Crypto and Eurocrypt conferences.





Research collaborations in the Asiacrypt region actively occur among security researchers and cryptographers.



Pieprzyk, Steinfeld (Macquarie), and Wang at Crypto 2007 studied the secure n -party computation in the passive, computationally unbounded attack model of the n -product function $f_G(x_1, \dots, x_n) = x_1 \cdot x_2 \cdots x_n$ in an arbitrary finite group (G, \cdot) , where the input of party P_i is $x_i \in G$ for $i=1, \dots, n$. The problem of constructing such protocols was reduced to a combinatorial coloring problem in planar graphs. Ron Steinfeld (Monash) is one of Australia's leading researchers in lattice cryptography and homomorphic encryption. His papers with Stehlé at Asiacrypt 2010 and Eurocrypt 2011 have been influential to lattice cryptography using cyclotomic rings and made major inroads to the efficiency of homomorphic encryption. Lynn Batten (Deakin) and Xun Yi (MIT) have made considerable contributions to privacy-preserving digital cash, notably the first detailed analysis of e-commerce-based personal information distribution from the purchaser viewpoint and solving the change-giving problem; as well as to privately querying aggregated medical data such that the privacy of the data, the query, the querier, and data owner are guaranteed simultaneously.

China. Xiaoyun Wang (Shandong) at Crypto 2005 broke the established hash functions SHA-0, SHA-1 (cited over 1,800 times), following her earlier work at Eurocrypt 2005, which broke the hash function MD5 (cited over 1,700 times) used in many commercial systems.

Japan. Within symmetric cryptology, Mitsuru Matsui (Mitsubishi) introduced the technique of linear cryptanalysis, which was the first attack that effectively broke the Data Encryption Standard (DES); this work has been cited in excess of 3,300 times and remains one of the top cryptanalysis techniques to guard against any new cipher designs. More recently, Yosuke Todo (NTT) invented the cryptanalytic technique of division property-based integral attack, which has broken the full rounds of the MISTY-1 cipher that had been proven secure against two top attack techniques, that is, differential and linear cryptanalysis. Tetsu Iwata (Nagoya) and Kaoru Kurasawa (Ibaraki)

designed the one-key CBC MAC (OMAC1 aka CMAC) block cipher-based message authentication code (MAC), which has made the CRYPTREC list and specified as NIST SP 800-38B.

For the context of public-key crypto, the Fujisaki-Okamoto (NTT) at Crypto 1999 is a well-known technique to upgrade the security of public-key encryption schemes from chosen plaintext resistance to adaptively chosen ciphertext security. It has been cited close to 1,000 times to date. The Okamoto identification scheme at Crypto 1992 is the first identity scheme that withstands adaptive attacks in the random oracle model, by proposing a new way of embedding an instance of the discrete-logarithm problem in the security proof. The notion of structure-preserving signatures (SPS) was introduced by Abe (NTT), Fuchsbauer, Groth, Haralambiev, and Ohkubo (NICT) at Crypto 2010. SPS is a proof-friendly signature scheme over pairing groups allowing to efficiently prove signatures without explicitly showing the signature, the message nor the verification key. It is used as a building block in privacy-preserving cryptographic protocols such as anonymous credential systems. Sakai, Ohgishi, and Kasahara developed the first identity-based key exchange scheme in 2000, which is the first constructive application of pairings over elliptic curves that preceded the first identity-based encryption (IBE) scheme by Boneh and Franklin in 2001. This opened the door to pairing-based cryptography that is widely deployed in real-world applications now. Tatsuaki Okamoto (NTT) and Katsuyuki Takashima (Mitsubishi) at Asiacrypt 2009 introduced a technique called dual pairing vector spaces that extends pairing groups to a vector space. It was a breakthrough to many applications such as attribute encryption and functional encryption scheme. In the context of real-world security, Tsutomu Matsumoto et al. (YNU) in their work that has been cited over 1,000 times, showed the surprising result that fingerprints artificially copied on silicone rubber fingers could effectively cheat fingerprint identification

systems, leading to a new research direction of liveness detection for fingerprint systems.

Korea. Kwangjo Kim et al. (KAIST) suggested a set of DES S-boxes that can be resistant against differential cryptanalysis and linear cryptanalysis (Asiacrypt 1991). Jung Hee Cheon et al. (SNU) developed cryptanalytic techniques for the CLT13 (Eurocrypt 2015) and CLT15 (Eurocrypt 2016) multilinear maps, as well as against the obfuscations based on the GGH13 (CRYPTO 2018) and GGH15 (Crypto 2019) multilinear maps. Cheon et al. also contributed substantially to solving variants of the discrete log problem (DLP), which directly affects the hardness of many security systems, as well as designed the first-known homomorphic encryption operating on real numbers that was subsequently used for privacy-preserving machine learning (Asiacrypt 2017). Korean researchers have also contributed substantially to research on broadcast encryption with improved efficiency (computation and transmission), identity-based encryption, and functional encryption with extra features such as tighter security reduction, revocability, and reduced key size, as well as on authenticated key exchange protocols including features such as dynamic grouping and forward secrecy.

Malaysia. Cryptographic research in Malaysia dates back to the early 2000s, and throughout the last two decades, the cryptographic community of actively publishing researchers in Malaysia remains largely the same. Main novel contributions from the community include new types of block cipher cryptanalysis techniques (double slides, realigning slides, overlapping multiset attacks, and bitslice tuple attacks) by Raphaël Phan, cryptanalysis of provable security (Phan, Bok-Min Goi and Wei-Chuen Yau), and variants of identification schemes by Swee-Huay Heng et al.

New Zealand. Early pioneers of information and computer security research in New Zealand include Hank Wolfe (Otago), Lech Janczewski (Auckland), Ray Hunt (Canterbury), Peter Gutmann (Auckland), and Ian

Welch (Victoria). Theoretical research in cryptography began with the return to New Zealand of Steven Galbraith (Auckland). Quantum computers, if they can be built to run Shor's algorithm at large scale, will break most public key cryptosystems in widespread use today. Two major research directions in post-quantum crypto are lattice crypto and isogeny crypto, based on mathematical problems that cannot be solved by Shor's algorithm. Galbraith is a major contributor to lattice cryptanalysis and introduced a significant optimization to lattice signatures, which has been used in submissions to the NIST PQCrypto standardization process. In isogeny crypto, Galbraith is a leading researcher—he illustrated an adaptive attack on isogeny crypto that has widely influenced the field (Asiacrypt 2016), and has made major advances (Asiacrypt 2017, Eurocrypt 2019) in isogeny signatures.

Singapore. Jian Guo (I²R), Thomas Peyrin (NTU), and Axel Poschmann (NTU) designed the PHOTON lightweight hash function family (Crypto 2011), which was adopted as an ISO standard in 2016. Hongjun Wu's (NTU) hash function design JH was selected as one of the five finalists of the SHA-3 competition in 2011; Hongjun Wu's authenticated encryption schemes ACORN and AEGIS, as well as Deoxys-II co-designed by Ivica Nikolić (NUS) and Thomas Peyrin (NTU) were selected as the final winners of the CAESAR competition in 2019. NTU's Jian Guo, Thomas Peyrin, and Hongjun Wu also contributed significantly on the cryptanalysis of symmetric-key primitives. Joseph Liu and Jianying Zhou et al. designed a lightweight identity-based signature scheme, which was adopted as an ISO standard in ISO/IEC 29192-4 in 2013. Yanjiang Yang and Jianying Zhou et al. designed password-based anonymous entity authentication scheme, which was adopted as an ISO standard in ISO/IEC 20009-4 in 2017.


Cross-Country Research Collaborations

Research collaborations in the Asia-crypt region actively occur among security researchers and cryptographers. To showcase this, consider the

papers published for the IACR flagship conferences—Crypto, Eurocrypt and Asiacrypt—in recent years. The accompanying table gives the summary of joint papers published in the past six years at these three main venues that involved researchers across multiple Asiacrypt countries, and if so, which countries they collaborated with. China actively collaborates, mainly with Australia, Japan, and Singapore. Moreover, Japan and Singapore also collaborate closely.

Ideas for even closer collaborations. Besides the practice of East Asia and Oceania countries taking turns hosting the flagship Asiacrypt conference, and cross-country research collaborations and joint publications, the following ideas could further strengthen collaborations and are worth exploring:

- ▶ Joint Ph.D. studentships among Asiacrypt universities: Universities in Asiacrypt countries could devise jointly supervised and jointly awarded Ph.D. degrees wherein the Ph.D. student spends a year in a host university where the co-supervisor is based.

- ▶ Research visits co-located with Asiacrypt: Arrange co-located research visits before and after Asiacrypt to be hosted by universities in the Asiacrypt-hosting country. This way, researchers aiming to attend Asiacrypt could leverage the trip as a research visit to a university in that country before or after the Asiacrypt period. 

References

1. Biham, E., Shamir, A. Differential cryptanalysis of DES-like cryptosystem. In *Proceedings of 1990 Crypto*, 2–21.
2. Courtois, N. and Pieprzyk, J. Cryptanalysis of block ciphers with overdefined systems of equations. In *Proceedings of 2002 Asiacrypt*, 267–287.
3. Fujisaki, E. and Okamoto, T. Secure integration of asymmetric and symmetric encryption schemes. In *Proceedings of 1999 Crypto*, 537–554.
4. Matsui, M. Linear cryptanalysis method for DES cipher. In *Proceedings of 1993 Eurocrypt*, 386–397.
5. Matsumoto, T. Gummy and conductive silicone rubber fingers. In *Proceedings of 2002 Asiacrypt*, 574–576.
6. Wang, X., Yin, YL, and Yu, H. Finding collisions in the full SHA-1. In *Proceedings of 2005 Crypto*, 17–36.
7. Wang, X. and Yu, H. How to break MD5 and other hash functions. In *Proceedings of 2005 Eurocrypt*, 19–35.
8. Zheng, Y. Digital signature or how to Achieve cost (signature and encryption) << cost(signature) + cost(encryption). In *Proceedings of 1997 Crypto*, 165–179.

For information regarding this article, contact Raphaël CW Phan, raphael.phan@monash.edu.

© 2020 ACM 0001-0782/20/4

DOI:10.1145/3378430

BY DONG KU KIM /YONSEI UNIV., HYEONWOO LEE /
DANKOOK UNIV., SEONG-CHOON LEE /GIGA KOREA
FOUNDATION, AND SUNWOO LEE /INSTITUTE
OF CONVERGENCE TECHNOLOGY

5G Commercialization and Trials in Korea

SINCE KOREA HAS a limited ICT R&D fund compared to other IT global countries, its strategy was essential to achieve its global competence in each generation of mobile communication. Just after the rollout of the world's first 5G service, the government took the next step by announcing the 5G + strategy to promote the 5G application to a wide-ranging industry and create a sustainable 5G ecosystem leading to new growth engines. In this article, we focus on the government-industry 5G collaborations, including the R&D roadmap and promotion to the 5G commercialization, the global collaboration, the first 5G experience, and 5G vertical trials to make the 5G-enabled industrial transformation take place in Korea.

The development of an electronic digital switching system called TDX in the 1980s, the world's first



CDMA mobile service in the 1990s, and the nationwide wired and mobile broad Internet networks in the 2000s are the key advances that made it possible for Korean consumers to easily adopt new technologies such as LTE and 5G. In 2018, the handset penetration rate of South Korea was similar to western Europe, where LTE adaption was 84% with 99.95% coverage and 65Mbps downlink capacity.⁴ Data consumption has been stagnant since 2018, but it has now increased to around 25GB/month after 5G commercialization, which is more than 2.5 times over LTE.

In 2013, the 5G Forum was established as the public-private 5G pro-



A researcher stands in front of an aerial view of Seongnam, South Korea, in 2018, exploring the transformation 5G will bring to urban life in the area.

motion think tank by distributing a national vision and 5G mobile communication strategy. Having recalled the first commercialization of LTE in Korea in July 2011, and the world-first nationwide network of LTE in 2013, the government-industry alliance for 5G began much earlier compared to previous generations. The 5G Forum published 16 white papers, including vertical-specific ones.¹ Figure 1 summarizes the key public-private efforts to 5G commercialization in services, R&D, ecosystem creation, and spectrum. Its strategy had three phases: The first phase focused on R&D to secure core technologies for equipment, and the second phase was aimed at 5G demon-

stration at PyeongChang Winter Olympics using Pre-5G technology. The goal of Phase 3 was to move forward with 5G commercialization. The PyeongChang Winter Olympics in early 2018 became the momentum gaining opportunity to secure 5G equipment and terminal development early, making sure that 5G service rolled out in April 2019, eight months before the initial plan.

Global Collaboration

As of Oct. 2019, 62 operators from 34 countries started 5G commercial services with a limited footprint. Indeed, Korea is rapidly ramping up the infrastructure aimed at nationwide coverage. Ericsson’s 2019 mobility report

expected 13 million 5G subscribers by end of 2019 (where Korea counts for more than one third) and one billion 5G subscribers by 2023. In order to establish a global 5G ecosystem and promote the timely introduction of 5G, collaboration has played a key role in two tracks—joint promotion and joint research.

5G Forum established bilateral and multilateral collaborations with Chinese IMT-2020 Promotion Group, European 5G-IA (Infrastructure Association), Japanese 5G-MF (Mobile Promotion), 5G Americas, 5G Brazil, UK5G, 5GTurkey, Indonesia i5G, Taiwan 5G-Alliance, WWRF, and FuTURE Forum. They focused on identifying,

PHOTO BY JEAN CHUNG/BLOOMBERG VIA GETTY IMAGES

Figure 1. Journey of government-industry collaboration to 5G commercialization.

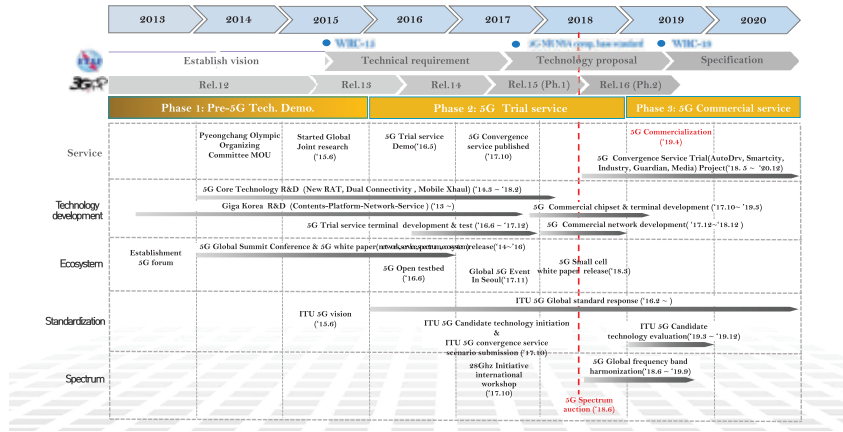


Figure 2. KT 5G coverage; SKT 5G coverage; and LGU+ 5G coverage as of 2019.10.07.

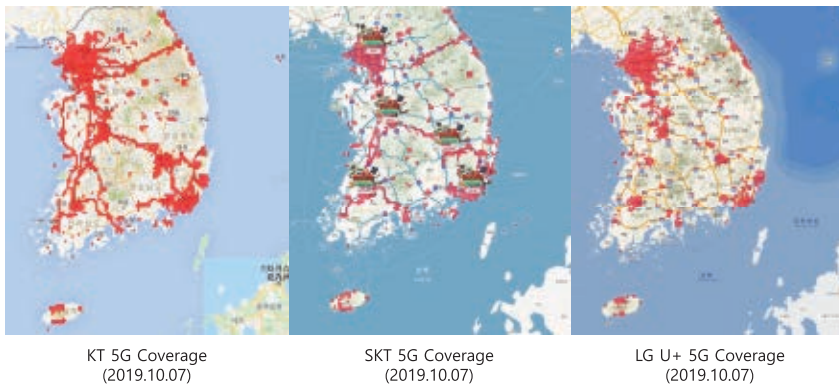


Figure 3. Examples of LGU+ Baseball Live; SKT AR Zoo; and KT Neckband 360.

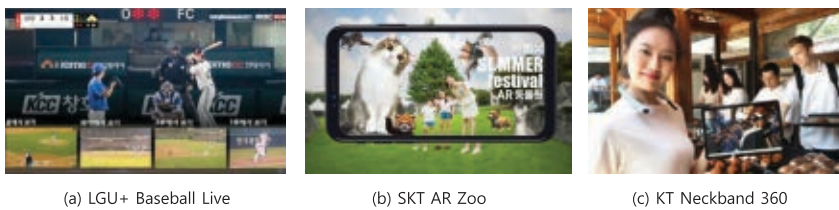


Figure 4. Trial cities.



sharing, and promoting the harmonized 5G visions, spectrum, and roadmap. Joint research with global partners currently include three R&D projects: 5G phase-2 R&D on advanced new radio millimeter-wave access technology, network inter-operability with the European Union, and high-speed vehicle application R&D with the U.K. Vertical specific university projects are also ongoing, such as 28GHz 5G V2X communication² and the tactile Internet.³

World's First 5G Experience

After 5G spectrum allocation in June 2018, each mobile carrier finally launched the world-first smartphone-based 5G commercial service in April 2019. It took no more than a year from spectrum allocation to commercial service.

Its subscriber base showed rapid growth, surpassing one million subscribers less than 70 days after commercial release. As of Nov. 2019, 5G subscribers have surpassed 4.3 million, which is mainly due to their unprecedented subsidy, tariff plans offering more data than 4G, and attractive VR / AR content services. Figure 2 shows the 5G coverage of three operators, which keep ramping up in 85 cities and major highways by the end of 2019.

The carriers competitively launched VR, AR, and high-definition live streaming services applying 5G bandwidth improvements; handset manufacturers also introduced innovative technologies such as foldable phones and wearable terminals. Figure 3 shows some of the 5G applications of mobile operators. Figure 3(a) is the LGU+ baseball live service, where you can watch baseball games of KBO (Korean baseball league) in the view angle as you want with low latency, improving the sense of realism. Figure 3(b) is a location-based mixed reality (MR) zoo service that SKT provides at some of the public parks in Seoul, bringing a particular fictitious animal of high-quality video to the real-life environment over 5G networks of ultra-low latency graphic processing at the MEC. Figure 3(c) is KT Neckband-360 that enables the live stream of the surrounding with 360-videos to give their friends virtual experiences. It can also be easily applied to private

surveillance or public disaster relief. Recently, telecom operators either already started or plan to begin some streaming game service with global cloud giants. Though it takes time to evaluate how much users appreciate early 5G services compared to LTE, more real-time applications of immersive media continue to come into the market.

It is also interesting that heavy data users in LTE are consuming only around 65% more after having switched to 5G service. It pointed out that the B2C market has already saturated in the LTE. Many collaborations are happening now in Korea for developing B2B applications enabled by the Rel-16 infrastructure along with cloud infrastructure and the proliferation of edge services, such as autonomous vehicle V2X, C-ITS service, smart factory, and smart healthcare.

5G Industrial Trials

5G promises to open the service big-bang era, along with the combined technologies of AI inference capability, edge cloud computing, and industrial IoT. For Korea to brace for the age of service big bang, 5G Forum along with MSIT and Giga KOREA Foundation set up the six task forces on Dec. 2016 to develop 5G convergence service scenarios that will likely to be commercialized in Korea by 2030. Many experts from three operators, manufacturers, institutes, and academia were brought together and published the representative service scenarios in the fields of AI, robots, smart city, autonomous industry, media, and public safety sector in 2017. Among those service scenarios, the Korean government chose five of them, which eventually led to 5G industrial trials starting in August 2018.

Figure 4 shows five ongoing trials, their consortium leaders, and the city location of testbeds, which are autonomous driving, smart factory, 5G media, public safety, and smart city. Another new trial in the fashion industry is beginning this October.

The 5G-enabled autonomous driving trial focuses on 5G V2X enabled bus, cloud-assisted remote-driving service, and URLLC-assisted traffic-control service, which includes the access of both 3.5GHz and 28GHz

spectrum. The crucial use case of a 5G factory trial is 5G-enabled autonomous multifunction robots, which is being tested in the real factory. 5G smart city trials focus on wireless intelligent CCTV and drone service. 5G public safety trials focus on critical evacuation service in the tunnel accident, in-building fire, and collapse of infrastructures. 5G media focuses on remote medical care, live remote entertainment, and VR experience. It is particularly promising that the trial consortium made sure to develop 5G industrial prototype modules to apply 5G real industrial connectivity in their use-cases by the end of this year.

Last year, the government announced the 5G+ strategy to further promote 5G business trials, which invests \$160 billion dollars in 5G industrial R&D by 2026. As the national SOC infrastructure ages after six years, the government plans to develop and expand the 5G friendly B2G business first in order to encourage the private sector to step in investment.

Conclusion

The 5G Forum keeps promoting the collaboration of government, ICT, and vertical industries, institutes, and academia to integrate 5G into the vertical industries. Finally, 5G Forum has already started laying out 6G vision. In fact, the 6G national project is slated to launch in 2021 and last for eight years. 

References


1. 5G Forum: <https://www.5gforum.org>
2. Cho, Y.J., Suk, G.-Y., Kim, B., Kim, D.K. and Chae, C.B. RF lens-embedded antenna array for mmWave MIMO: Design and performance. *IEEE Commun. Mag.* 56, 7 (July 2018), 42–48.
3. Kim, K.S. et al. Ultrareliable and low-latency communication techniques for tactile Internet services. In *Proceedings of the IEEE 107*, 2 (Feb. 2019), 376–393.
4. OpenSignals. State of LTE: <http://bit.ly/370dtCw>

Dong Ku Kim (dkkim@yonsei.ac.kr) is a professor at the School of Electrical & Electronic Engineering at Yonsei University, as well as chair of the 5G Forum executive committee, Seoul, Korea.


HyeonWoo Lee (woojaa@dankook.ac.kr) is a professor at the college of SW Convergence at Dankook University as well as a vice-chair of the 5G Forum executive committee, Seoul, Korea.

Seong-Choon Lee (sclee0328@gkf.kr) is the president of the Giga KOREA Foundation and a member of the 5G Forum executive committee, Seoul, Korea.

Sunwoo Lee (sunwoo.lee@kt.com) is the head of Infra Laboratory at the Institute of Convergence Technology and a member of the 5G Forum executive committee, Seoul, Korea.



Ericsson's 2019 mobility report expected 13 million 5G subscribers by end of 2019 (where Korea counts for more than one third) and one billion 5G subscribers by 2023.



BY SANG KIL CHA /KAIST AND ZHENKAI LIANG /
NATIONAL UNIVERSITY OF SINGAPORE

Asia's Surging Interest in Binary Analysis

BINARY CODE ANALYSIS (binary analysis, for short) is a vital security approach for protecting commercial off-the-shelf (COTS) software and understanding malware, where there is no source code available. From the perspective of computer security, it is imperative to analyze binary code, as source-level scrutiny does not always reveal lurking software bugs due to compiler or interpreter misbehavior.

Since the late 1990s, there has been significant research interest worldwide on binary analysis.

The BitBlaze project (by Carnegie Mellon University and University of California, Berkeley)¹⁴ is one of the few pioneering research prototypes that incorporates a variety of tools for binary analysis, such as VINE for static analysis, TEMU for dynamic analysis, and Rudder for symbolic execution. Following up on BitBlaze, BAP (by Carnegie Mellon University)³ provides a wealth of APIs that can be used to build a custom binary analyzer, while DECAF (by Syracuse University)⁷ provides efficient, platform-neutral support for dynamic binary analysis. Angr (by University of California, Santa Barbara)¹³ offers a user-friendly platform for common binary analysis tasks, such as disassembly, instrumentation, and symbolic execution that are utilized by an active user community.

In recent years, the expanding cyber infrastructure supporting governments and industry sectors in Asia has ignited strong interests in software security, leading to a surge of research activities in binary analysis in the region. Thanks to the broad international support, researchers have begun to drive new ideas and delve into various research topics in binary analysis, such as automatic exploit generation, vulnerability discovery, and automated reverse-engineering. This article outlines the key research developments and trends in binary analysis led by researchers in the East Asia and Oceania region.

Growing Research Interest in Vulnerability Discovery and Exploit Generation

Fuzzing is a popular technique to discover software vulnerabilities by randomly or semi-randomly generating inputs to tested software programs. It has been an area of focus for researchers from China, Korea, and Singapore. AFLFast (from the National University of Singapore)² is one noteworthy project used by many security practitioners today to find security vulnerabilities in software. It served as the catalyst for grey-box fuzzing research: many



research papers on grey-box fuzzing came out after its release.¹² CollAFL (by Tsinghua University, China)⁶ and Eclipser (by the Korea Advanced Institute of Science and Technology, or KAIST)⁴ are examples of regional efforts in improving fuzzing efficiency.

Another research focus is on new types of memory vulnerabilities and exploits. Data-oriented exploitation (by National University of Singapore)^{8,9} is a new type of memory error exploit that works by manipulating non-control data of the program without hijacking the control flow of a target program, which brings the expressiveness of data-oriented exploits to a new level.

Automatic Exploit Generation (AEG, by Carnegie Mellon University)¹ is a pioneering work of automatically finding and exploiting the vulnerabilities of a program, which incorporates analysis techniques such as symbolic

execution and fuzzing. AEG is important for software security, as one can apply the technique to discover vulnerabilities and quickly fix security-relevant vulnerabilities prior to the release of software products. AEG typically requires binary code, as it is not possible to figure out the exact memory layout by simply looking at the source code. The importance of AEG was realized by the Cyber Grand Challenge (CGC), the first hacking competition between machines hosted by DARPA in 2016. To expedite the AEG process, one needs to search for exploitable states in program paths diverging from crashing inputs found by fuzzing. Revery (by the University of Chinese Academy of Sciences and Tsinghua University, China)¹⁵ tackles this challenge by adopting a control-flow stitching technique.

Since the CGC, several Asian countries have started to organize similar

competitions. Korea ran the AI-based Automated Vulnerability Discovery Challenge in 2018, and Japan also hosted the Automatic Cyber Hacking Challenge at the Code Blue Conference in 2018. Both competitions were focused on attacks, that is, binary-level exploitation, rather than defenses, unlike CGC, where binary patching played a crucial role. However, the Korean Ministry of Science, ICT, and Future Planning (MSIT) have announced this year's competition would include both attacks and defenses.

Efforts in Building Scalable Binary Analysis Frameworks

Binary analysis faces a constant challenge due to the ever-increasing demands of researchers and cybersecurity responders. The dramatic increase of binary analysis tasks calls for frameworks that reduce human effort and boost productivity, which



The expanding cyberinfrastructure supporting governments and industry sectors in Asia has ignited strong interests in software security, leading to a surge of research activities in binary analysis in the region.




can be scaled to meet real-world demands. Based on the common abstractions used in binary analysis, such as intermediate representation (IR), instruction semantics, data flow, control flow, and others, research in the region aims to develop automatic methods to generate abstractions and optimize analysis processes, as well as enabling analysis by various tools to inter-operate.

Many of the efforts carried out in the Asia region aim to produce solid building blocks for scalable binary analysis frameworks. For example, writing the specification of instructions from scratch is largely an error-prone task; instruction-set manuals typically comprise thousands of pages of descriptions written in natural language. To write a binary analysis front-end, which translates binary code into an IR, one should carefully read the manuals and implement the logic. This process requires tremendous engineering effort and thus, many researchers consider it too costly to investigate. TaintInduce (by the National University of Singapore and the Chinese Academy of Sciences)⁵ is a project to automatically generate taint rules (or data-flow properties) without manual specifications; it infers taint rules based on observations of instruction executions. In addition, TaintInduce proposes a common definition and API for taint rules, enabling follow-up work to be built on a common knowledge base.

In 2019, researchers from KAIST made public their binary analysis framework, called B2R2.¹⁰ This was the first attempt to build a binary analysis framework in this region. It focused on optimizing the performance and accuracy of a binary analysis front end, which was often neglected by binary analysis researchers, as the front end had been regarded as a simple translation module. However, they presented various optimization techniques, including parallel lifting and big-integer splitting, that achieve an order-of-magnitude improvement in the performance of the front end. The same research team published a novel technique for finding semantic bugs that appeared in IRs for binary analysis, which can

benefit any binary analysis framework available today.¹¹

Concluding Remarks

Binary analysis has been gaining popularity in Asia. Built on the momentum of vulnerability discovery and exploit generation, as well as the foundational work to build scalable platforms for binary analysis, we hope to see more active regional research collaboration in the field. With extended collaborative effort, we believe researchers in this region will trigger major technological breakthroughs in the future. 

References

1. Avgerinos, T., Cha, S.K., Hao, B.L.T. and Brumley, D. AEG: Automatic exploit generation. In *Proceedings of the Network and Distributed System Security Symp.*, 2011.
2. Bohme, M., Pham, V.-T. and Roychoudhury, A. Coverage-based grey-box fuzzing as Markov chain. In *Proceedings of the ACM Conf. Computer and Communications Security*, 2016.
3. Brumley, D., Jager, I., Avgerinos, T. and Schwartz, E.J. BAP: A binary analysis platform. In *Proceedings of the Intern. Conf. Computer-Aided Verification*, 2011.
4. Choi, J., Jang, J., Han, C. and Cha, S.K. Grey-box concolic testing on binary code. In *Proceedings of the Intern. Conf. Software Engineering*, 2019.
5. Chua, Z., Wang, Y., Băluță, T., Saxena, P., Liang, Z. and Su, P. One engine to serve'em all: Inferring taint rules without architectural semantics. In *Proceedings of the Network and Distributed System Security Symp.*, 2019.
6. Gan, S. et al. CollAFL: Coverage sensitive fuzzing. In *Proceedings of the IEEE Symp. Security and Privacy*, 2018.
7. Henderson, A., Prakash, A., Yan, L.K., Hu, X., Wang, X., Zhou, R. and Yin, H. Make it work, make it right, make it fast: building a platform-neutral whole-system dynamic binary analysis platform. In *Proceedings of the Intern. Symp. Software Testing and Analysis*, 2014.
8. Hu, H., Chua, Z., Adrian, S., Saxena, P. and Liang, Z. Automatic generation of data-oriented exploits. In *Proceedings of the USENIX Security Symp.*, 2015.
9. Hu, H., Shinde, S., Adrian, S., Chua, Z., Saxena, P. and Liang, Z. Data-oriented programming: On the expressiveness of non-control data attacks. In *Proceedings of the IEEE Symp. Security and Privacy*, 2016.
10. Jung, M., Kim, S., Han, H., Choi, J. and Cha, S.K. B2R2: Building an efficient front-end for binary analysis. In *Proceedings of the NDSS Workshop on Binary Analysis Research*, 2019.
11. Kim, S., Faerevaag, M., Jung, M., Oh, S.J.D., Lee, J. and Cha, S.K. Testing intermediate representations for binary analysis. In *Proceedings of the Intern. Conf. Automated Software Engineering*, 2017.
12. Manès, V.J. et al. The art, science, and engineering of fuzzing: A survey. *IEEE Trans. Software Engineering*, 2019.
13. Shoshitaishvili, Y. et al. (State of) the art of war: Offensive techniques in binary analysis. In *Proceedings of the IEEE Symp. Security and Privacy*, 2016.
14. Song, D. et al. BitBlaze: A new approach to computer security via binary analysis. In *Proceedings of the Intern. Conf. Information Systems Security*, 2008.
15. Wang, Y., Zhang, C., Xiang, X., Zhao, Z., Li, W., Gong, X., Liu, B., Chen, K., Zou, W. Revery: From proof-of-concept to exploitable (one step towards automatic exploit generation). In *Proceedings of the ACM Conf. Computer and Communications Security*, 2018.

Sang Kil Cha is an assistant professor at KAIST, South Korea.

Zhenkai Liang is an associate professor at the National University of Singapore.

© 2020 ACM 0001-0782/20/4

ACM Transactions on Quantum Computing (TQC)

Open for
Submissions

Publishes high-impact, original research papers and select surveys on topics in quantum computing and quantum information science



Recent advances in quantum computing have moved this new field of study closer toward realization and provided new opportunities to apply the principles of computer science. A worldwide effort is leveraging prior art as well as new insights to address the critical science and engineering challenges that face the design, development, and demonstration of quantum computing. Alongside studies in physics and engineering, the field of quantum computer science now provides a focal point for discussing the theory and practice of quantum computing.

ACM Transactions on Quantum Computing (TQC) publishes high-impact, original research papers and select surveys on topics in quantum computing and quantum information science. The journal targets the quantum computer science community with a focus on the theory and practice of quantum computing including but not limited to: quantum algorithms and complexity, models of quantum computing, quantum computing architecture, principles and methods of fault-tolerant quantum computation, design automation for quantum computing, issues surrounding compilers for quantum hardware and NISQ implementation, quantum programming languages and systems, distributed quantum computing, quantum networking, quantum security and privacy, and applications (e.g. in machine learning and AI) of quantum computing.

For more
information
and to submit
your work,
please visit:

tqc.acm.org



Association for
Computing Machinery

Article development led by [acmqueue](https://queue.acm.org)
queue.acm.org

Controlling coordination costs when multiple, distributed perspectives are essential.

BY LAURA M.D. MAGUIRE

Managing the Hidden Costs of Coordination

IT STARTED WITH 502 errors. Almost immediately a flood of user reports swamped the service's community Slack channel.

A user posted "Getting 502s?" at 9:22 A.M., and within minutes 40 other users responded with the Yes and MeToo emojis.

Also at 9:22 A.M., in an ops channel, an incident had been opened by an on-call engineer, and the site reliability engineers responsible for the service had been paged out. By 9:23 A.M. five responders were checking logs and dashboards.

At 9:25 A.M.—less than two minutes after an initial tentative question indicated there may be an issue—the first notification was pushed out to users. This was aimed at slowing the influx of user reports from the 77,000-plus user community.

In less than seven minutes, eight hypotheses about the nature of the problems had been proposed by the responders. In that same period, five of those had been investigated and discarded.

Within the first 10 minutes of the incident, the responders had been directly in touch with the 4,700 users in their community channel, opened tickets with three dependent services' support teams, and coordinated among a response squad of 10.

Diverse players are engaged when IT systems run at speed and scale. This becomes immediately apparent when the service is disrupted. Those whose work depends on the system functioning, both directly and indirectly, are compelled to get involved either to help with resolution or to seek more information so they can adjust their goals and priorities to account for the degraded (or absent) service.

Often, because of the business-critical nature of the service or four nines service-level agreements, a service outage triggers an all-hands-on-deck page for multiple responders. This core group represents a small fraction of the roles involved, however. Even with a brief look at an incident response, it becomes apparent that performance in resolving service outages in these systems is about rapid, smooth coordination of these multiple, diverse players, as expressed in the accompanying figure.

Joint activity distributed among this collective takes place across scripted and unscripted efforts such as recognizing the disruption, taking actions that safeguard the system from further decline, diagnosing the source(s) of the problems, determining potential solutions, cross-checking a fix before the code gets pushed, as well as a whole suite of after-action activities.

Even in relatively smaller scale systems, the incident response can become less about diagnosis and repair of service outages and more about managing the needed capabilities of



multiple responders, the potential benefits that could be realized by having more participants available to assist, and the needs of the stakeholder groups. This coordination incurs additional demands. For example, for their skills and experience to be useful to the current flow of events, incoming responders need to be briefed and understand tasks delegation according to a required sequencing of tasks. Doing this requires a substantial amount of effort—particularly as the severity of the outage or number of responders increases or the uncertainty grows.

In the high-consequence world of managing service delivery for critical digital infrastructure, the time pressure to diagnose and repair an outage is enormous.¹ While resources may be

readily available, it can be extraordinarily challenging to use them as the tempo of the incident escalates and the efforts to stop a cascade of failures occupy all the attention of the response team.

Herein lies the crux of the issue: The collaborative interplay and synchronization of roles is critical,^{12,13,15} but prior research has shown poor coordination design incurs cognitive costs for practitioners, specifically, the additional mental effort and load required to participate in joint activities.^{5,6} This can be particularly exacerbated in the digital services domain where it plays out across geographically distributed groups. Using examples from critical digital services, this article explores the nature of coordination costs and how software engineers experience them

during a service outage. These findings provide new directions for design to control costs of coordination in incident response.

Hidden Costs of Coordination

The choreography needed for smooth operation is effortful,⁷ particularly when the system is under stress. But these efforts are difficult to discern and typically not separated from expected “professional practice” within a field. This choreography arises as “an escalating anomaly can outstrip the resources of a single responder quickly. There is much to do and significant pressure to act quickly and decisively. *“To marshal resources and deploy them effectively requires a collection of skills that are related to but different from*

those associated with direct problem solving. But to be effective, these resources must be directed, tracked, and redirected. These activities are themselves demanding.”¹⁸

That this collection of skills goes largely unnoticed is not surprising. The fluency with which expert practitioners manage these coordination demands minimizes the visibility of the efforts involved.¹⁹ It is only when the coordination breaks down that it comes to the forefront. Difficulties in synchronizing activities, disruptions to the smooth flow of task sequences, or conversation explicitly aimed at trying to organize multiple parties are examples of evidence that coordination breakdowns have occurred.

It is worth separating out the choreography needed for coordination from the costs that those activities incur. An example of this occurs when recruiting new resources to an incident response—just one function in joint activity. The associated overhead costs include:

- ▶ Monitoring current capacity relative to changing demands
- ▶ Identifying the skills required
- ▶ Identifying who is available
- ▶ Determining how to contact them
- ▶ Contacting them
- ▶ Waiting for a response
- ▶ Adapting current work to accommodate new engagement (waiting, slowly completing tasks to aid coordination)

- ▶ Preparing for engagement
 - ▷ Anticipating needs
 - ▷ Developing a ‘critsit’ or status update
 - ▷ Giving access/permissions to tools and coordination channels
 - ▷ Generating shared artifacts (dashboards, screenshots)
- ▶ Dealing with access issues (inability to join web conferences or trouble establishing audio)

These overheads seem relatively benign—they are implicit features of any joint activity. And that is precisely the point: They can be a minimal burden in normal operations and therefore disregarded as worthy of support in explicit design. In high-tempo, time-pressured, and cognitively demanding scenarios, however, these burdens increase to the point of overloading already burdened responders. Think of a loss of engine power during the first few minutes of flight or an unexpected event during a spacewalk—seconds count here and any additional friction in cognitive work matters. Now think of the speed at which critical digital services operate—*microseconds* count and the hidden coordination costs can matter in previously unconsidered ways. The cognitive costs of coordination matter in incident-response processes. Now let’s consider how poor coordination design impacts engineering teams responsible for system reliability.

The Need for Coordination Design

Highly technical system operation is increasingly non-located. Demands for near-perfect reliability and the burnout this can generate for on-call engineers has given rise to different models of 24/7 systems management to distribute calls across time zones. Even when a team may be geographically collocated, outages happen in off-hours or when members of the team may be traveling, in meetings, or otherwise unavailable for face-to-face interactions. This means incident response should be designed to accommodate entirely remote joint activity.

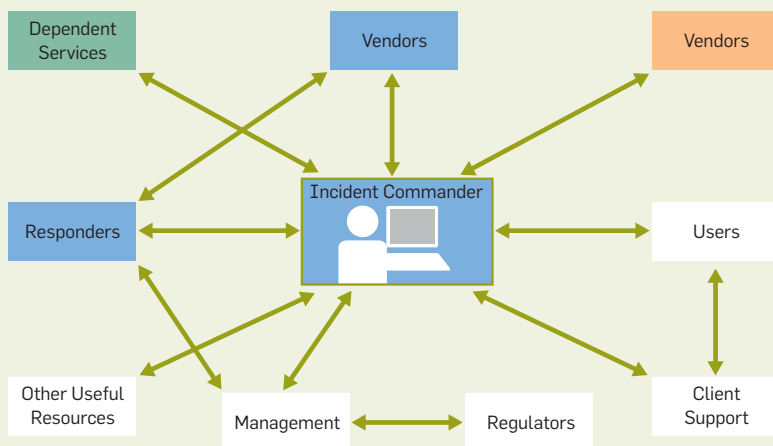
The need for good coordination design transcends the software community: Increasingly, other industries that were not typically geographically dispersed in the past are taking advantage of technological capabilities to distribute their workforces to optimize cost or available talent (providing just-in-time expertise).

Current coordination design focuses on the structure of handling support, including triage methods whereby run-books or troubleshooting algorithms are used by less experienced support engineers before escalating to experts or through geographically dispersed support networks that “follow the sun.” These formats can decrease the need to wake up expert resources when the system goes down, but these configurations do not eliminate the need for coordination design. The requirements are shifted in ways that can escalate situations, compounding the coordination demands of the event as other stakeholders get engaged.

Let’s follow this through with an example. When anomalies generate the need to page the on-call staff, these direct responders begin gathering. Simultaneously, other stakeholders with an interest in the problem are also drawn in. Users may begin flooding support channels and ticketing systems trying to determine if the service is degraded or if their system is wonky, or dependent services may experience problems and begin asking for information. This coordination “noise” makes it challenging to determine if these are all the same problems, related, or unrelated.

With diagnostic and safeguarding activities commanding substantial at-

Resolving service outages.



tion, additional resources are then needed to triage this influx of reports and sort through the incoming data to minimize data overload.¹⁶ As the incident progresses and the concern over impact grows, escalations to management bring in even more participants as senior executives begin pressing for more details or demanding the service be restored. Customer support roles facing urgent requests from clients will seek information to pass along.

Despite the substantial number of parties involved, systems are rarely designed with explicit attention to the coordination requirements. When they are, typically it is to: centralize response coordination through an incident commander; design an overly prescriptive process management perspective that fails to account for the hidden cognitive work of coordination; or depend on tooling that fails to fully support the dynamic, nonlinear manner in which incident response happens. These methods do not necessarily support the cognitive work of coordination the way they are intended.

Attempts at Supporting Coordination

Some would argue that coordination design is fundamental for developing and deploying technology in distributed systems such as critical digital infrastructure (CDI). But process-driven coordination design—emphasizing distributed tasks instead of joint activity—will not address the needs described earlier. One example of process-driven industry best practice surrounding coordination during service outages—borrowed from disaster and emergency response domains—is an incident command system (ICS). Central to this model is assigning an incident commander (IC) and ensuring disciplined adherence to the shared ICS across the roles and groups involved. Let's look at how these two tenets can actually limit resilient incident-response practices.

Attempt 1: Assigning an IC. The intent of the IC role is to manage the coordination requirements of the involved parties by directing the activities of others and holding the responsibility for taking timely decisions. Under certain conditions (for example, in low-tempo scenarios with few involved parties or reasonably known



The need for good coordination design transcends the software community: Increasingly, other industries that were not typically geographically dispersed in the past are taking advantage of technological capabilities to distribute their workforces to optimize cost and available talent.



and predictable event outcomes) this may be an appropriate configuration. In these contexts (or these phases of an incident), the cognitive and coordinative demands are manageable without design for coordination.^{7,12,13} Routine events can be handled without undue stress.

Escalations that move a situation to nonroutine or exceptional, however, dramatically increase the cognitive activities needed to cope and generally do not follow a predictable course. As demands grow, an incident-command structure tends to become a workload and activity bottleneck that slows response relative to the tempo of cascading problems.²⁰ Working both *in* and *on* the incident forces attention to be divided across the “inherent” roles of the position. For example, the IC needs to be tracking the details of the incident to be prepared to anticipate and adapt to rapidly changing conditions, but too much effort spent on forming an accurate assessment of the situation takes away from managing the coordination across roles. In reverse, trying to centrally manage who does what when tends to fall behind the pace of events and challenges, making the trouble harder to resolve and the joint activity harder to synchronize.

This is not an inconsequential point. Being an effective choreographer of the joint activity demands current, accurate knowledge *and* the ability to redirect attention to the orchestration of the players coming in and out of the event alongside their changing needs. In addition, what is seen as the IC maintaining organizational discipline during a response can actually be undermining the sources of resilient practice that help incident responders cope with poorly matched coordination strategies and the cognitive demands of the incident.

Attempt 2: Enforcing operational discipline to follow the ICS. Previous studies in software have shown different strategies for coping with workload demands such as dropping tasks (known as *shedding load*), deferring work to later, or reducing the quality of the work performed.² Other attempts to balance the workload sink with the value of the coordination call for adding more resources, but this comes with costs as well. In

poorly designed systems, resources needed to help handle the demands are unable to be brought into play smoothly without disrupting the work under way to control the adverse effects of the event.

Herein lies a paradox: You have resources available but are unable to make them useful. Concurrently, their attempts to become useful are counterproductive—new responders coming into an audio bridge or ChatOps channel need to ask for a briefing, and the updating disrupts the flow of activity. This can drive the formation of side channels among select responders where diagnostic work can take place uninterrupted. Creating this peripheral space is necessary to accomplish cognitively demanding work but leaves the other participants disconnected from the progress going on in the side channel.

Unless you have been “on the fireline” of an event of this sort, it can be easy to minimize the tension inherent in these situations. It’s worth restating: the systems studied in coordination research are often life-critical or otherwise high-consequence. Despite the importance of coordination, timely actions must be taken to cope with anomalies as they threaten to produce failures. When high costs of coordination could undermine the ability to keep pace with the evolving demands of the anomalous situation, people responsible for the outcomes will, of necessity, adapt. Incident response in critical digital infrastructure systems is not exempt. In fact, the speed and scale at which CDI operates, coupled with the challenges of a distributed team connected through technology, make the domain particularly susceptible to interference from excessive costs of coordination.

In observations of critical events and post-mortems, adaptations to create subgroups in different channels that are separate from the “official” incident response occur repeatedly.⁹ Often, postmortems misinterpret these forms of adaptation to high costs of coordination. Retrospective discussions portray these adaptations as contrary to the ICS protocols and therefore lead to efforts to block people from forming these channels. The behavior is actually an adap-



Being an effective choreographer of the joint activity demands current, accurate knowledge and the ability to redirect attention to the orchestration of the players coming in and out of the event alongside their changing needs.



tive strategy to cope as coordination becomes too expensive. Rather than forcing responders to bear significant attentional and workload costs, it is advisable to facilitate shifting various lines of work to subgroups while supporting connecting the progress or difficulties into the larger flow of the response.

The emergency services community has begun to recognize the limitations of the ICS,⁴ as have other domains where command and control or hierarchical methods are giving way to more flexible teaming structures.^{10,11} When practices such as ICS are adopted across domains, it is important to pay close attention to the critiques and findings from other large-scale, multi-agent coordination contexts. In doing so, it is possible to limit the unintended adverse impacts when real-world demands of one setting challenge the practices imported from another.

These findings about how people in an incident response adapt when high costs of coordination threaten the critical cognitive work are an important source of design seeds to guide innovations.

Attempt 3: Using technology to facilitate coordination. The term *computer-supported cooperative work* (CSCW) was coined by Irene Greif and Paul Cashman in the early 1980s to describe the emerging field of computers mediating the coordination of activity across people and roles.³ Since then, advances in technological capabilities, the omnipresence of the computer in the workplace, and the proliferation of automated processes have solidified the importance of CSCW, while rendering it redundant since almost all forms of joint activity have become computer-mediated.

Still, this field has three main themes that are of particular interest in CDI: the use of collaboration software platforms; the coordination of joint activity between humans and bots; and the nature of reciprocity in human-automation teaming.

Collaboration Software Platforms

Not surprisingly, because of the changing needs of the work environment and the technical capabilities of the workforce, software engineering has driven innovation and the development of tooling and practices for collabora-

tive work. Online software platforms take traditional offline activities such as project management planning, issue tracking, group discussion, and negotiation of shared work and enable real-time collaboration of participants across a distributed network.

The platforms have shifted from expensive, proprietary forms of file sharing to broadly accessible, cloud-based tools that can be quickly adopted across both formal and ad hoc groupings. Lowering the barrier to collaboration in this way eases the coordination costs of transient, single-issue demands and of early exploratory efforts. This means collaborative work can be facilitated more rapidly with less overhead. Flexible coordination structures also provide the ability to adapt their use to the problem demands.

The resilience demonstrated in the earlier example of forming side channels to manage high costs of coordination was facilitated by the ease with which direct messages could be sent or new channels could be spun up. Supporting rapid reconfiguration into smaller, ad hoc teams enables smooth transitions as activity is distributed across continuously changing groups of participants. This collection of attributes—adapting to changing problem demands, dynamic reconfiguration of resources, and smooth coordination—is critically important in high-consequence work and a prominent feature of groups that are skilled at distributed joint activity in many domains.

Designing technology that can aid these capabilities is a means to control the costs of coordination. While many of these platforms optimize coordination costs on one criterion (rapid reconfiguration), ChatOps platforms exact penalties in coordinating *with* the tools themselves. For example, while the practice of ChatOps allows traceability that *could* support bringing new responders up to speed, the packed message-list format of the tooling is poorly designed to do so.¹⁴ Responders coming into an event that is under way must scroll through the list of text, searching for the relevant lines of inquiry still in consideration, key decision points, and other important contextual information to gain a current understanding of the situation.

These seemingly trivial aspects of design matter greatly. Think back to the tension inherent in high-tempo operations when seconds matter and expert resources are in high demand. Those who are likely to be drawn in to join in the response efforts on a service outage frequently possess specialized skills that are often scarce. As such, they may not be brought into the event until later stages, at which time the tempo or propagation of failure drives a need for taking urgent action. Poor design renders ChatOps nearly useless as a tool for sensemaking as people come into an evolving and increasingly pressured situation.

Coordinating Joint Activity Across Humans and Machines

The last subsection shifted the framing of controlling the costs of coordination. Initially, cost of coordination referred to the additional efforts to accommodate the tasks and interactions inherent in joint activity. In human-human coordination the costs of the interaction are borne by both parties, and “investments” may be made by relaxing individual or short-term goals in the service of accommodating shared or longer-term goals. Working jointly distributes the costs across the participants. The preceding subsection introduced an important distinction: Interacting with tools and automation skews the costs. There are many coordination costs in human-machine teaming that go unnoticed or are exacerbated by tool design.

For example, the initial expenditure of effort to set up tooling designed to aid in various functions of anomaly response, such as monitoring or alerting, can be substantial. Engineers responsible for assembling their own stacks spend considerable effort in: assessing the appropriateness of a tool for a given purpose; evaluating it relative to their team’s needs; considering the technical capabilities needed to understand how it functions; learning how it works; maintaining an accurate mental model as new features are added; determining appropriate configurations; performing maintenance to ensure that old configurations are removed or updated as demands change; tolerating the lack of context sensitivity that can result in unnecessary alerting; pro-

viding access and permissions to the users on the team; constructing security measures to prevent inadvertent changes; and making changes and adjustments as new tools are integrated. (The list could continue.) These are all examples of how coordinating with machines have costs for their human counterparts. If the tool were a human colleague, the amount of effort you would need to expend to ensure it remained a relevant team member might give you pause; however, this fundamental asymmetry that unduly burdens the human team members with additional costs to compensate for the limitations of automation is characteristic of current-day human-machine teams.^{6,7}

A key (and often overlooked) aspect of the dynamics of teamwork across human-human and human-machine networks is the degree to which the participants in the joint activity *consider* the goals, workload, and needs of others and adapt their actions accordingly.

Recognizing the Dynamics of Reciprocity

Choreographing technologically mediated joint activity can enable greater opportunities for reciprocity when the technology is designed to combat excessive costs of coordination.¹⁷ For example, studies of NASA’s space-shuttle mission control during critical events reveals many patterns of effective joint activity. Of particular interest, many people join in beyond those who are titularly responsible. The technology that mediates communication in the control room and back-rooms facilitates bringing people up to speed as they join in from being off duty, with low burdens on the people currently handling the anomalies.¹³ The additional personnel provide diverse perspectives, especially as each flight controller increasingly focuses on his or her scope of responsibility as the anomalous situation unfolds. The ability to “look in and listen in” has been widely documented as a benefit to smooth coordination.^{8,12}

It’s not difficult to see the parallel between mission control and CDI in the rapid escalation in the number of stakeholders (other responders, users, customer support, management) during a service outage. Technologies that

enable this and other abilities for joint activity in a fully distributed network without adding extra burdens provide a means for people whose skill, experience, and knowledge *could* be useful to the event but who have not been explicitly drawn in can ready themselves to assist should the need arise. Being current on the event progression, yet untethered to specific responsibilities, offers an opportunity for reframing through fresh perspectives (see Grayson article in this issue).

In outlining these three attempts at supporting coordination, it's clear that technology both affords lower-cost coordination by supporting adaptive capacity and exacerbates high-cost coordination through asymmetrical burdens on the human side. In CDI environments, where technology can be rapidly developed and deployed, designs can easily add unintended costs for joint activity unless the tools are explicitly designed to support coordination.

Conclusion

Coordination remains an integral part of large-scale, distributed work systems, but the lack of coordination design for joint activity continues to add hidden cognitive costs for practitioners. These efforts and load are related to the additional work of enabling smooth synchronization across multi-party groupings as the cognitive work of anomaly response is completed in high tempo, evolving incident scenarios. Recall the opening case, in which the escalating incident brought in multiple, diverse, and distributed perspectives, each with a vested interest in the event progression.

Each participant was necessary to managing the outage both directly and indirectly, and the ChatOps forum enabled their participation. Closer examination across a number of cases, however, reveals a paradox: The platforms themselves both facilitate and hinder coordination. The easy formation of side channels enables engineers to adapt through flexible reconfiguration outside of the main response efforts, but bringing new responders up to speed is made difficult by the structure of a packed message-list design.

Some of the common tactics thought to control the costs of coordination include adopting incident command

structures, specifically the IC role. Using collaborative software platforms and adopting technologies to aid in coordination have been shown in actual cases to reveal limits and unrecognized implications for cognitive work. Nevertheless, all of these areas provide opportunities to choreograph smoothly in high-tempo, multi-agent events, especially by supporting the ability to adapt when the costs of coordination climb too high.

Some initial considerations to control cognitive costs for incident responders include: assessing coordination strategies relative to the cognitive demands of the incident; recognizing when adaptations represent a tension between multiple competing demands (coordination and cognitive work) and seeking to understand them better rather than unilaterally eliminating them; widening the lens to study the joint cognition system (integration of human-machine capabilities) as the unit of analysis; and viewing joint activity as an opportunity for enabling reciprocity across inter- and intra-organizational boundaries.

Controlling the costs of coordination will continue to be an important issue as systems scale, speeds increase, and the complexity rises in the problems faced during anomalies that disrupt reliable service delivery. ■

Related articles on queue.acm.org

The Calculus of Service Availability

Ben Treynor, Mike Dahlin, Vivek Rau, Betsy Beyer
<https://queue.acm.org/detail.cfm?id=3096459>

Collaboration in System Administration

Eben M. Haber, Eser Kandogan, Paul Maglio
<https://queue.acm.org/detail.cfm?id=1898149>

Distributed Development Lessons Learned

Michael Turnlund
<https://queue.acm.org/detail.cfm?id=966801>

References

- Allspaw, J. Trade-offs under pressure: heuristics and observations of teams resolving Internet service outages. M.S. thesis, 2015. Lund University; <https://lup.lub.lu.se/student-papers/search/publication/8084520>.
- Grayson, M.R. Approaching overload: Diagnosis and response to anomalies in complex and automated production software systems. M.S. thesis, 2018. The Ohio State University; https://etd.ohiolink.edu/pg_10?::NO:10:P10_ETD_SUBID:174511.
- Grudin, J. Computer-supported cooperative work: History and focus. *Computer* 27, 5 (1994), 19-26; <https://www.microsoft.com/en-us/research/wp-content/uploads/2017/01/IEEEComputer1994.pdf>.
- Jensen, J., Waugh Jr., W.L. The United States' experience with the Incident Command System: What

we think we know and what we need to know more about. *J. Contingencies and Crisis Management* 22, 1 (2014), 5-17; <http://bit.ly/2ROsJXm>.

- Klein, G. The strengths and limitations of teams for detecting problems. *Cognition, Technology & Work* 8, 4 (2006), 227-236; <https://link.springer.com/content/pdf/10.1007%2Fs10111-005-0024-6.pdf>.
- Klein, G., Feltovich, P.J., Bradshaw, J.M., Woods, D.D. Common ground and coordination in joint activity. *Organizational Simulation*. W. Rouse and K. Boff, eds. Wiley, New York, NY, 2004, 139-184; http://jeffreymbradshaw.net/publications/Common_Ground_Single.pdf.
- Klein, G., Woods, D.D., Bradshaw, J., Hoffman, R.R., Feltovich, P.J. Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems* 19, 6 (2004), 91-95; <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1363742>.
- Luff, P., Heath, C., Greatbatch, D. Tasks-in-interaction: Paper and screen-based documentation in collaborative activity. In *Proceedings of the ACM Conf. Computer-supported Cooperative Work*, 1992, 163-170; <https://dl.acm.org/citation.cfm?id=143475>.
- Maguire, L.M. Controlling the cognitive costs of coordination in large-scale, distributed systems: In search of a model of choreography to support joint activity. Dissertation, 2020, forthcoming; The Ohio State University; https://www.researchgate.net/profile/Laura_Maguire5.
- Nemeth, C.P. Groups at work: Lessons from research into large-scale coordination. *Cognition, Technology & Work* 9, 1 (2007), 1-4; <https://link.springer.com/article/10.1007/s10111-006-0049-5>.
- O'Leary, R., Bingham, L.B., eds. *The Collaborative Public Manager: New Ideas for the 21st Century*. Georgetown University Press, 2009; <https://muse.jhu.edu/book/13036>.
- Patterson, E.S., Watts-Perotti, J., Woods, D.D. Voice loops as coordination aids in space shuttle mission control. *Computer Supported Cooperative Work* 8, 4 (1999), 353-371; <http://bit.ly/2NyppgB>.
- Patterson, E. S., Woods, D. D. 2001. Shift changes, updates, and the on-call architecture in space shuttle mission control. *Computer Supported Cooperative Work* 10(3-4), 317-346; <https://link.springer.com/article/10.1023/A:1012705926828>.
- Potter, S. S., Woods, D.D. 1991. Event-driven timeline displays: Beyond message lists in human-intelligent system interaction. In *Proceedings of the IEEE Intern. Conf. Systems, Man, and Cybernetics*; <https://ieeexplore.ieee.org/abstract/document/169864>.
- Watts-Perotti, J., Woods, D. D. 2009. Cooperative advocacy: an approach for integrating diverse perspectives in anomaly response. *Computer Supported Cooperative Work (CSCW)* 18(2-3), 175-198; <https://link.springer.com/article/10.1007/s10606-008-9085-4>.
- Woods, D.D. Cognitive demands and activities in dynamic fault management: abduction and disturbance management. *Human Factors of Alarm Design*. N. Stanton, ed. Taylor & Francis, London, U.K., 1994, 63-92.
- Woods, D.D. Essentials of resilience, revisited. *Handbook on Resilience of Socio-Technical Systems*. M. Ruth and S.G. Reisemann, eds. Edward Elgar Publishing, 2019, 52-65; <http://bit.ly/38edu2F>.
- Woods, D.D., ed. STELLA Report from the SNAFUcatchers Workshop on Coping with Complexity. SNAFU Catchers Consortium, 2017; <http://stella.report/>.
- Woods, D.D., Hollnagel, E. *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*. CRC Press, Boca Raton, FL, 2006.
- Woods, D.D., Patterson, E.S. How unexpected events produce an escalation of cognitive and coordinative demands. *Stress, Workload, and Fatigue*. P.A. Hancock and P.A. Desmond, eds Mahwah, N.J. L. Erlbaum, Mahwah, NJ, 2001; http://cse.leng.ohio-state.edu/productions/laws/laws_mediapaper/2_4_escalation.pdf.

Laura M.D. Maguire is a researcher in the Cognitive Systems Engineering Lab at The Ohio State University. She has been a researcher with the SNAFU Catchers Consortium since 2017 and works closely with digital service companies on incident response practices, tool development, design, and contextual research. She tweets on cognitive systems engineering at @LauraMDMaguire.

Copyright held by owner/author.
 Publication rights licensed to ACM.

A look at how we respond to the unexpected.

BY MARISA R. GRAYSON

Cognitive Work of Hypothesis Exploration During Anomaly Response

WEB PRODUCTION SOFTWARE systems currently operate at an unprecedented scale, requiring extensive automation to develop and maintain services. The systems are designed to adapt regularly to dynamic load to avoid the consequences of overloading portions

of the network. As the software systems scale and complexity grows, it becomes more difficult to observe, model, and track how the systems function and malfunction. Anomalies inevitably arise, challenging incident responders to recognize and understand unusual behaviors as they plan and execute interventions to mitigate or resolve the threat of service outage. This is *anomaly response*.¹

The cognitive work of anomaly response has been studied in energy systems, space systems, and anesthetic management during surgery.^{10,11} Recently, it has been recognized as an essential part

of managing Web production software systems. Web operations also provide the potential for new insights because all data about an incident response in a purely digital system is available, in principle, to support detailed analysis. More importantly, the scale, autonomous capabilities, and complexity of Web operations go well beyond the settings previously studied.^{8,9}

Four incidents from Web-based software companies reveal important aspects of anomaly response processes when incidents arise in Web operations, two of which are discussed in this article. One particular cognitive func-

tion examined in detail is hypothesis generation and exploration, given the impact of obscure automation on engineers' development of coherent models of the systems they manage. Each case was analyzed using the techniques and concepts of cognitive systems engineering.^{10,11} The set of cases provides a window into the cognitive work "above the line"³ in incident management of complex Web operation systems.

It seems easy to look back at an incident and determine what went wrong. The difficulty is understanding what actually happened and how to learn from it. Hindsight bias narrows the ability to learn as it leads an after-the-fact review to oversimplify the situation that people faced and miss the real difficulties. When Web requests are failing and customers cannot access content, however, people ask questions about what is malfunctioning, what is the underlying problem driving the disturbances observed, or what interventions will mitigate or resolve the problems being experienced?

Software engineering consists of sense-making in a highly dynamic environment with extensive and sometimes puzzling interdependencies in a network of systems mostly hidden "below the line." Problems produce effects and disturbances that propagate and appear distant from the source driving behavior—effects at a distance in highly interdependent processes.⁷ Observing and tracing the behaviors of multiple automated processes and subsystems is difficult and ambiguous.

The people engaged in resolving the incident bring mental models about how the different components, functions, and subsystems are interconnected and update these models as they explore possible explanations. Understanding and resolving anomalies can require connecting experiences and knowledge gained from handling multiple past incidents. But no models of how the system works are identical or complete, so understanding the event requires work to integrate information and knowledge across the diverse perspectives. Hypothesis exploration and planning interventions are collaborative processes across distributed parties that could be all around the world (see Maguire's article in this issue).

What Is Anomaly Response and Hypothesis Exploration?

Anomalies come in many forms, though the cognitive work of responding to them involves a basic set of key functions. An anomaly has two qualities: it is abnormal and unexpected, such as strangely slow response times for loading a home page or high network demand during a typically low-traffic period in the middle of the night. Recognizing an anomaly as a discrepancy between what is observed and what is expected of system behaviors depends on the observer's model of what the system is doing in that context. Anomalies are events that require explanation, since the current model of the system does not fit what is observed.

In other words, anomalies are triggers to generate and explore hypotheses about what is going on that would, if true, account for the anomalous behaviors.¹⁰ Multiple anomalies can build up over time as problems propagate through highly interdependent networks and as actions are taken to counter abnormal behaviors. Anomalies become an unfolding set of unexpected findings matched by generating an unfolding set of candidate hypotheses to test as potential explanations.

The cognitive work of anomaly response involves three interdependent lines of activity: anomaly recognition, in which practitioners collect and update the set of findings to be explained; hypothesis exploration, in which practitioners generate, revise, and test potential explanations that would account for the findings; and response management or replanning, in which practitioners modify plans in progress to maintain system integrity, mitigate effects on critical goals, and determine what interventions can resolve the situation. Each of these is time-dependent and requires revision as evidence about anomalies and their driving sources comes in over time, remedial actions are taken that produce additional surprising effects, and pressure to resolve the situation builds even when uncertainty persists.

Given a set of findings to be explained, hypothesis exploration generates and tests candidates. Interestingly, research shows that the difficulty of hypothesis exploration increases as the scale of interdependencies increas-

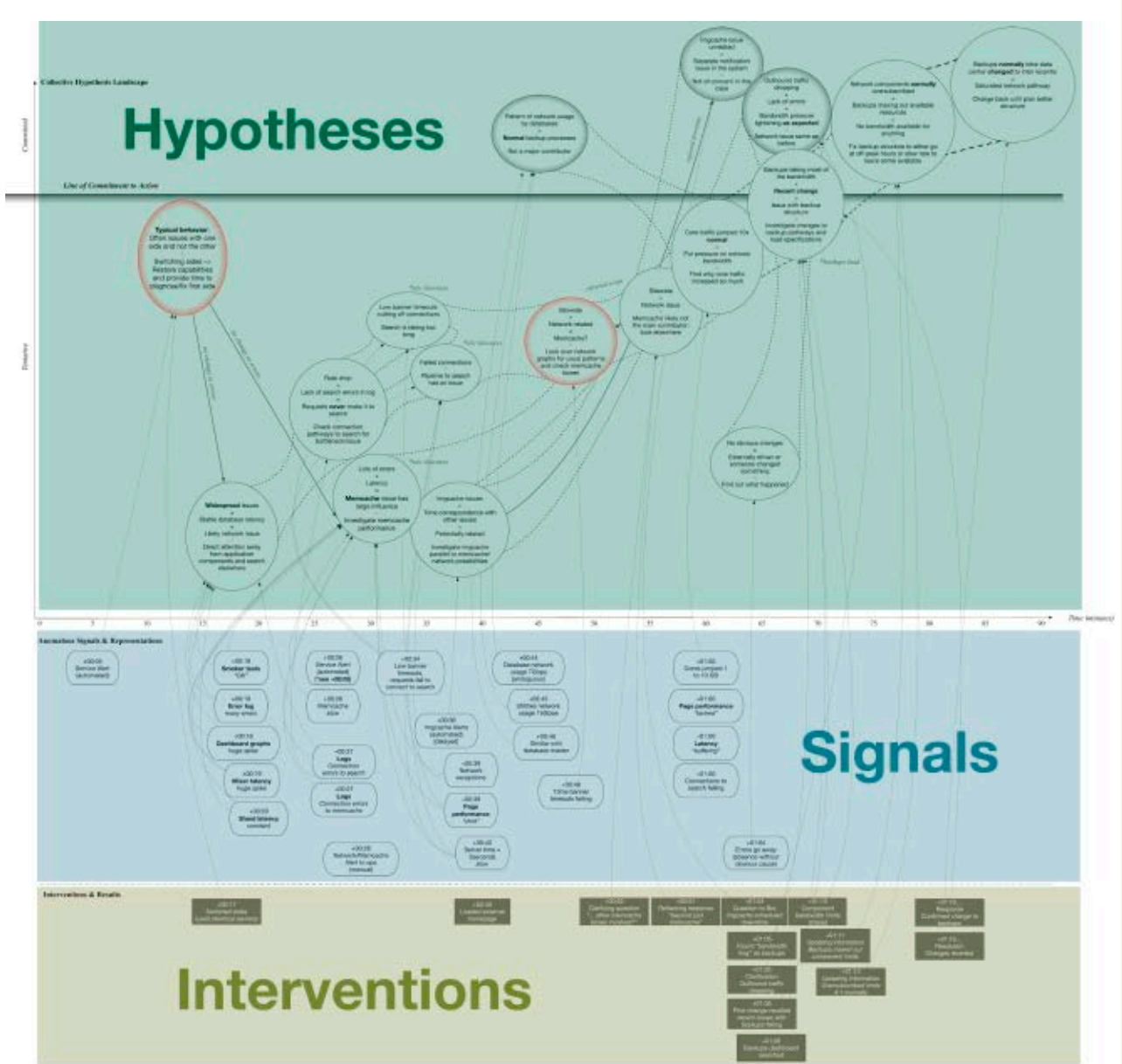
es. In hypothesis generation, the goal is to broaden the set of hypotheses under consideration as candidate explanations for the pattern of findings and avoid premature narrowing. Research strongly suggests that diverse perspectives—in the right collaborative interplay—help generate a broader set of candidates. New events will occur while hypothesis exploration is under way, and these may reinforce current working hypotheses, elaborate on the current set of possible hypotheses, or overturn the current working hypothesis. The incredible speed of automation can quickly change the information landscape, further complicating hypothesis exploration.

Anomaly Response Cases in Web Operations

The study under discussion here is based on four cases sampled from the larger corpus of available cases in the database developed by the SNA-FUcatchers Consortium (<https://www.snafucatchers.com/about-us>), a group of industry leaders and researchers focused on understanding how people cope with complexity and produce resilient performance in the operation of critical digital services. The definition of *incident* varies among organizations, though most capture circumstances around service losses or degradations (for example, Huang et al.⁴). The relevant parties and company-specific details were de-identified for the analysis.

Chat-log files were gathered from postmortem records as the primary data source for each of the cases. The chat logs were either from IRC (Internet Relay Chat) or Slack, depending on the main communication technology used at the time. The chat logs do not directly show the actions of the engineers on the system; they do record intention and plans that individuals expressed above the line in the process of responding to anomalies and the noted signals crossing the line.¹² Additionally, the chat operations demonstrate the emergence of anomalies in the observer's implied stance, given the written updates in the main channels. The data records were supplemented with knowledge-elicitation sessions with individuals who had direct knowledge of the incidents.

Figure 1. Collective hypothesis landscape.



The analysis used process-tracing methods.⁶ Over several iterations, the communication logs were analyzed by applying a lightweight coding scheme based on the cognitive work of anomaly response and macrocognitive functions.^{5,7} The focus was on several key processes, including events; hypothesis generation; model revisions; interventions; and stance.² These five aspects captured the expectations and communication flow of the engineers responding to cascading disturbances.

This article focuses on the results of hypothesis generation and explo-

ration. The engineers communicate active theories to provide direction for diagnostic search, as well as broadening the hypothesis-exploration space with contributions from multiple perspectives. The evolution of the hypothesis space is marked and laid out diagrammatically (figures 1–3), featuring activities such as adding or ruling out hypotheses; findings that support active hypotheses; hypothesis modifications; revisiting past hypotheses; mental model updates; and points of confusion and uncertainty.

Visualizing the Hypothesis Exploration Space

Over the course of the cases, software engineers offered many hypotheses. The anomalous signs and signals prompted new ideas to emerge, as well as supporting the evolution or dismissal of other explanations. The chat channels enabled open communication about these hypotheses in a collective landscape that was theoretically available to all participants at any time during the incident. The parallel cognitive paths were laid out for each case to show the diverse patterns of action and insight brought to bear.

The top portions of the diagrams, exemplified in Figure 1, portray the different hypotheses in the shared landscape. Each bubble has a condensed version of evidence and proffered conclusion. The hypothesis-exploration space (top portion) is marked with various hypothesis bubbles, which are supported by the anomalous signs (middle) and the shared interventions and results (bottom). The line of commitment separates the hypotheses that were acted upon, though may have been proven false or irrelevant depending on the case. The hypotheses are connected, showing both divergence and convergence over time. Notably, some hypotheses were ultimately dismissed (red outline at minutes 15 and 50) or noted as irrelevant to the matter at hand (black outline at minutes 50, 60, and 70). The line of commitment marks a point where action was taken, often in spite of uncertainty.

The middle portion of each figure supports the upper section with specific moments of anomalous signs and signals. Each marker denotes the time since the incident's start and anomalous state. The bottom portion shows interventions and clarifying questions the engineers made during the inci-

dent. These actions could be diagnostic or therapeutic depending on the case.

Both the signals and the interventions have arrows driving toward single or multiple hypotheses in a similar alignment to that of representations toward above-the-line practitioners, as shown in Cook.³ The hypotheses are generated above the line and are motivated from the anomalies and interactions arising from the line of representation.

Next let's examine two of the case studies, in which the narrative of investigation and mitigation is shown in a graphical timeline representation.

A Case of Widespread Latency

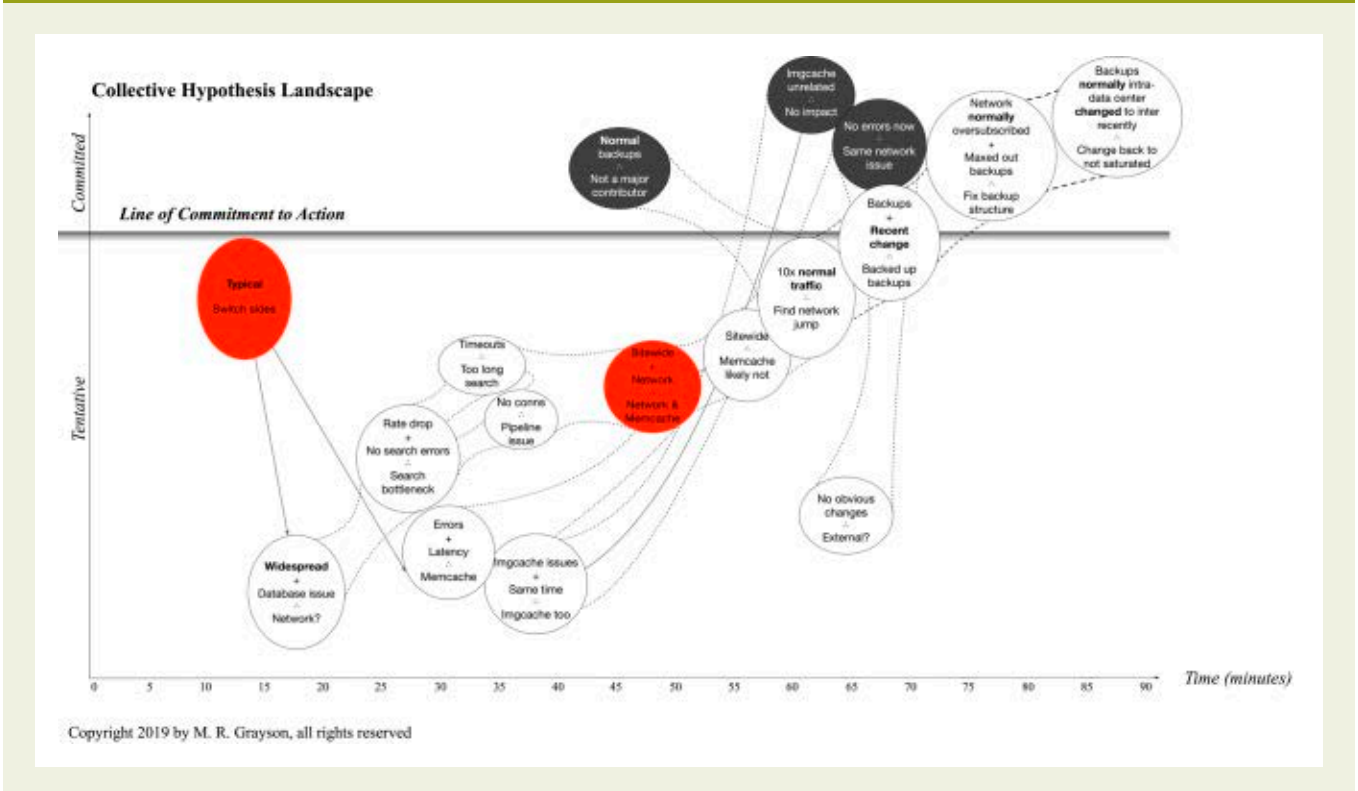
Background: two datacenters house the databases and servers needed to run a website. Backups are periodically stored to protect the data if anything goes wrong and are kept separately within each datacenter. Network pipelines connect the two, as well as connecting them to the terminals that the software engineers use to sustain the site. Little do they know that what appears to be a reasonable minor change will have widespread consequences.

A search engineer is on call when an automated alert is triggered. He and another engineer gather several peo-

ple online in the predawn hours to diagnose the anomalous behavior in multiple systems. Increased latency, lag, and connection issues are rampant across search, memcache, monitoring charts, and the production website. The network connection issues are the driving source of the overload, though it is not immediately obvious to the search and operations engineers. They see the symptoms of the network degradation but have limited access to trace the underlying issues. Their initial hypotheses are tentative as they gather more information.

As seen in the top section in Figure 2, early action is taken to switch the server groups. The hypothesis-exploration space (top portion) is supported by the anomalous signs (middle) noted by the engineers in the chat logs, as well as the shared interventions and results (bottom). The hypotheses eventually converge in this case to surpass the relative line to commit to a plan of action. Often a problem with one half would not affect the nearly identical other half. That hypothesis is quickly abandoned, however, to follow other ideas about what is driving the widespread disturbances.

Figure 2. Anomalous signals and representations.




The first warning signs come from a custom email alert by one product engineer that there are connection errors to the API servers. The errors are intermittent and hard to reproduce consistently among multiple individuals. The normal alerting thresholds are not triggered, so the product engineer alerts other knowledgeable engineers in operations and infrastructure to assist. No errors are appearing in the typical logging monitors, which seems very strange to the responders. They determine the connections are not reaching the servers but cannot support their hypotheses about where the problem resides. One person recalls that a recent teeing junction was put in place to test certain boxes before adding them to production by diverting a small amount of traffic to them.


As shown in Figure 3, the software engineers decide to remove the testing boxes early on, which does not resolve the issue. The shared hypothesis space (top) in this case features several divergent hypotheses, some of which are acted upon and eventually combined. The anomalies (middle) are fairly consistent throughout the incident, while the interventions (bottom) deal with questioning and following different lines of inquiry.

Many hypotheses are discussed and available to the participants in the chat, though it takes revisiting old ideas and integrating new ones to form any satisfactory conclusions. Although the participants later realize that some hypotheses were incorrect, the information they discover by following several paths of inquiry evolve into other hypotheses.

The initial responders have limited access to the load-balancer logs and are mainly focused on the application-level components, without much success. A network engineer with load-balancer access is called in to check on the connections and finds an old rule on the same cluster of servers in question. After removing the rule and the most recently added teeing rule, the engineer finds that the errors disappear. Together with the infrastructural engineer who implemented the teeing rule, they determine that the rules had unexpectedly interacted and reactivated the old one, which directed requests to a box that no longer existed.



Hypothesis exploration is complicated by the interacting effects hidden below the observable monitors.



The basic assumption for the load-balancer structure is that it will not send traffic to a server that is not able to handle it. It accomplishes this purpose with a health check, a short request response verifying that the server has available capacity. The teeing rules, however, do not automatically have this health check and can send requests to a box that might not exist. Furthermore, the interaction of multiple rules could have some influence on this check if one set has a valid target while the other does not. It is difficult for the engineers to estimate the downstream impact of the dropped requests for their system's functions and the end users' experience since the fraction of diverted requests is so small and invisible to monitoring.

Prior testing rules were not actively curated and are left on the machines without overt influence over the network traffic flow. A normal testing structure is added that inadvertently reactivates an older junction rule that instead sends traffic to a box that has been decommissioned. After this discovery is made, there is still confusion as to why the new rule interacted with the old one when they should have been independent. Hypotheses are acted upon without great certainty as to the effects, such as removing the teeing rules completely. The network engineer with the most access and directability on the load balancer also struggles to understand the entanglements that triggered the sinkhole of dropped requests. Eventually, a few theories emerge as plausible explanations for the apparent zombie rule, although no definitive consensus is reached without further testing.

Exploring New Hypothesis Landscapes

The groups of engineers in each of these cases explored their hypothesis spaces in different ways, though both had common challenges in reaching their incidents' "conclusions." The true nature of incidents is in the continual flow of day-to-day operations rather than the short duration captured in a postmortem. Nevertheless, the captured cases do show relevant patterns such as exploring both narrowly and more broadly.

The first case saw the ideas converging toward a fairly confident plan of action. In contrast, the second case had hypotheses that the engineers committed to early on and many divergent paths without a clear resolution. Both had initial responses that were unsuccessful in accomplishing the desired result, but provided additional information to direct subsequent hypotheses. Each probe into the system and search spurred new ideas to be added to the collective hypothesis space.

The engineers in both cases demonstrated a vital skill of interpreting data and providing context to the ambiguous signals. The underlying automation is opaque, especially when performing highly autonomous functions such as distributing network traffic in a load balancer. Effects emerging at a distance from their sources, however, presented in highly interpretable ways. Each case demonstrated a different set of signals observable to the engineers over the course of the incident. Another side effect of the interdependent, opaque network is masking, which obscures the automation-driven functions that might be relevant. The diversity of pathways through which overload can occur and surface is a symptom of the complexity of the network, which requires deeper and more informative measures for investigation.

Hypothesis exploration is complicated by the interacting effects hidden below the observable monitors. Limited measurable signals, masking, and strange loops restrict the human responders' abilities to understand the systems and take appropriate corrective actions. Time also affected the scope of investigation. Recent changes were given priority as likely contributors to the current issue, even when evidence may have supported other explanations.

It is much more difficult to trace changes disjointed in time, such as a week prior or long-term choices that left latent effects waiting to be activated by specific circumstances. One major difficulty in tracing anomalies in complex software systems is the system's constant state of change. Hundreds of updates occur each day, varying in length, though their impacts might not be felt until much later as a


cumulative effect. Current alerting platforms often provide localized information, which can help support focused hypotheses and also narrow the scope of investigation. The responders in these cases were hindered by the lack of observability into the underlying dynamics of the automation.

The tracings of the hypothesis-exploration space in this article specifically reveal unexpected patterns for incident management: there are multiple committed hypotheses and interventions above the line; many hypotheses are generated in a short time; more actions and hypotheses are continually made after one was committed; and the opaqueness and complexity of the highly dynamic systems produce effects at a distance that complicated hypothesis exploration.

The complexity of the system and autonomous functions drove investigators to collaborate and explore multiple hypotheses in responding to the anomalies. Diverse perspectives expanded the hypotheses considered and beneficially broadened the scope of investigation. Explicit comments by engineers updating each other's mental models were frequent in the chat logs; this finding supports Woods' theorem on the importance of finding and filling gaps in understanding.⁸

The chartings of hypothesis evolution also demonstrate the influence of a collective idea space via the communication channel. Early divergence of multiple hypotheses led to some tentative commitments to action, as well as ruling out irrelevant contributions. The discarded ideas often helped other paths gain momentum toward a general convergence sufficiently explaining the anomaly. The unique experiences, skill sets, and roles of the individual responders contributed to resolving the complex challenges.

Ultimately, sharing ideas and investigating several hypotheses broadened the engineers' views of the problem enough to find reasonable solutions. Whether it was a cumulative progression of evidence or eureka moments after finding the right monitoring source, the incident responders were able to intervene and protect the functionality of the software systems. High-reliability continuous development and deployment pres-

ures engineers to keep pace with change and adapt to constant challenges. Their hypothesis exploration should be supported by the tools they use every day because they are already solving problems that end users never even know about. 

Related articles on queue.acm.org

The Debugging Mindset

Devon H. O'Dell

<https://queue.acm.org/detail.cfm?id=3068754>

Searching Vs. Finding

William A. Woods

<https://queue.acm.org/detail.cfm?id=988405>

User Interface Designers, Slaves of Fashion

Jef Raskin

<https://queue.acm.org/detail.cfm?id=945161>

References

- Allspaw, J. Trade-offs under pressure: Heuristics and observations of teams resolving Internet service outages. Master's thesis. Lund University, Lund, Sweden, 2015.
- Chow, R., Christoffersen, K., Woods, D.D. A model of communication in support of distributed anomaly response and replanning. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 44, 1* (2000). Sage Publications, 34–37.
- Cook, R.I. Above the line, below the line. *Commun. ACM* 63, 3 (Mar. 2020), 43–46.
- Huang, P. et al. Gray failure: The Achilles' heel of cloud-scale systems. In *Proceedings of the 16th Workshop on Hot Topics in Operating Systems*. ACM, 2017, 150–155.
- Klein, G. et al. Macrocognition. *IEEE Intelligent Systems* 18, (2003), 81–85.
- Woods, D.D. Process tracing methods for the study of cognition outside of the experimental psychology laboratory. *Decision making in Action: Models and Methods*. G.A. Klein, J. Orasanu, R. Calderwood, C.E. Zsombok, eds. CT: Ablex Publishing, Westport, CT, 1993, 228–251.
- Woods, D.D. Cognitive demands and activities in dynamic fault management: Abductive reasoning and disturbance management. *Human Factors in Alarm Design*. N.A. Stanton, ed. Taylor & Francis, Bristol, PA, 1994, 63–92.
- Woods, D.D., ed. STELLA: Report from the SNAFUcatchers Workshop on Coping with Complexity, 2017; <https://snafucatchers.github.io>.
- Woods, D.D. The strategic agility gap: how organizations are slow and stale to adapt in a turbulent world. *Human and Organizational Factors in High-Risk Companies*. F. Daniellou and R. Amalberti, eds. Foundation for Industrial Safety Culture, Toulouse, France, 2018.
- Woods, D.D. and Hollnagel, E. Anomaly response. *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*. CRC/Taylor & Francis, Boca Raton, 2006, 69–95.
- Woods, D.D. and Hollnagel, E. Automation surprise. *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*. CRC/Taylor & Francis, Boca Raton, 2006, 113–142.
- Woods, D.D., Patterson, E.S. and Roth, E.M. Can we ever escape from data overload? A cognitive systems diagnosis. *Cognition, Technology & Work* 4, 1 (2002), 22–36.

Marisa Grayson is an award-winning cognitive systems engineer at Mile Two, LLC in Dayton, OH. She is a member of the SNAFU Catchers Consortium and the Cognitive Systems Engineering Lab at Ohio State University.

Copyright held by author/owner.
Publication rights licensed to ACM.

DOI:10.1145/3360310

Risk transfer options offer hope, but little more.

BY DANIEL W. WOODS AND TYLER MOORE

Cyber Warranties: Market Fix or Marketing Trick?

WHEN BUYING A second-hand car you are at the mercy of the dealer. The dealer knows which cars were treated well by past owners and which are likely to break down within a few months. When buying an information security product, the vendor has a better idea of how effective the product truly is. In both cases, the seller has information the buyer lacks.

Economists refer to this phenomenon as a market with asymmetric information. Akerlof¹ suggested this leads to a “market for lemons” dominated by lower quality goods (aka lemons in the case of used cars). Consumers cannot differentiate between lemons and quality used cars. Akerlof’s model suggests only lemons would be sold in such a market.

Car dealers offer warranties to overcome this problem. If the used car breaks down within, say,

six months, the dealer must pay for its repair. This discourages dealers from selling lemons with lengthy warranties. Consequently, the length of the warranty provides information about how likely the vehicle is to break down.

Returning to information security, vendors have started attaching cyber warranties to information security products with no additional fee. Will cyber warranties better align incentives in the market for information security products? Or are they marketing tricks riddled with coverage exclusions hidden in the fine print of the terms and conditions?

Might Cyber Warranties Remedy the Market for Lemons?

A natural first question to ask is why warranties might succeed in addressing the market for lemons where other mechanisms have failed. Akerlof¹ identified possible solutions including brand reputation, certification, liability laws, and warranties.

Linking brand reputation to the effectiveness of products is difficult because they appear to be working until an attack succeeds, which happens infrequently. Reputation systems are further limited by commercial sensitivity preventing information from being pooled across organizations. Vendors instead signal quality by speaking at

» key insights

- **Because security is often unobservable, products are not rewarded in the marketplace for being more secure. Warranties could correct this information asymmetry by signaling product quality.**
- **Cyber-product warranties provide limited coverage for replacing or repairing products, whereas cyber-incident warranties cover the consequences of an attack.**
- **Terms and conditions vary widely across vendors. Prescriptive security requirements are more prevalent in warranties than in cyber insurance policies.**
- **Exclusions convey useful information about product limitations to those who read the fine print.**



conferences, publishing security research, and through marketing activities. The latter can lead to (arguably deceptive) claims about product functionality that may not reflect reality.

External experts could certify the effectiveness of the product. Past history shows certification firms face incentives to skip on assessment. A framework for certifying computer systems as secure “motivated the vendor to shop around for the evaluation contractor who would give his product the easiest ride.”² Even if such incentives were overcome, there are difficulties in using laboratory experiments to establish real world security.

Liability laws could shift the costs of an ineffective product back onto the vendor. This might incentivize vendors to create more effective products and even force firms selling ineffective products out of the market. However, the resistance to software liability is well documented.^{3,4} To prove vendors liable for creating a defective product, the product in question must be shown to have caused the injury. Establishing such proximate cause is fiendishly difficult, given the constellation of security controls employed by firms.

So why might cyber warranties succeed where other approaches have failed? Certification incurs large up-

front costs regardless of effectiveness, whereas warranties only incur a cost when the product fails to mitigate an attack. Consequently, vendors with more effective products incur less cost in offering warranties. The barriers to adoption can be overcome by individual firms unilaterally offering warranties—courts need not assign liability nor governments pass legislation.

This article evaluates three viewpoints on the role of warranties. The theoretical view argues cyber warranties can align incentives and fix a dysfunctional market, as put forward in Woods and Simpson.⁵ A skeptical view characterizes cyber warranties as mar-

keting tricks offering little meaningful coverage to the buyer. The conciliatory view holds that while warranties do not significantly change the incentive to invest, they do prevent vendors from overexaggerating the functionality of products. Which viewpoint best describes reality can be answered empirically by inspecting the terms of the warranties, which we undertake next.

What Do Cyber Warranties Cover?

We searched for combinations of the terms “warranty,” “indemnity,” “information,” “security,” and “cyber” using a popular search engine. We stopped when further results revealed no new warranties attached to information security products. Some vendors provide a description of the warranty without the actual contract, we included these descriptions in our corpus if they were detailed enough for our purposes. This resulted in a corpus of 15 warranties attached to information security products.


Inductive analysis identified *coverage*, *obligations*, and *exclusions* as the main components of the warranties. Coverage describes which costs the vendor will indemnify and the total indemnification limit. Obligations describe what the buyer must do for the warranty to be valid. Exclusions describe which circumstances invalidate coverage.

Consumers should first ask whether the product comes with a product or incident warranty. Of the 15 warranties, two-thirds were only triggered by defective hardware or software. We will call these *cyber-product warranties* from now on, denoted by *P* in the table. Cyber-product warranties offer to repair or replace the product, denying coverage for first- or third-party costs resulting from an attack.

Cyber-incident warranties (denoted by *I* in the accompanying table) cover the consequences of an attack. The firms offering cyber-incident warranties sell intangible products and services like source code review, network monitoring, or back-up services. Four of the five cyber-incident warranties in our sample covered first-party costs like notifying customers and hiring consultants for forensic investigation, public relations or legal review. One vendor explicitly covered ransomware payments and nothing else



The conciliatory view holds that while warranties do not significantly change the incentive to invest, they do prevent vendors from overexaggerating the functionality of products.



(denoted I^{RWP}). None of the warranties (*I* or *P*) cover regulatory fines or third-party liability. The amount of coverage ranged from \$10,000 to \$5,000,000 depending on the size of the buyer.

Obligations on the buyer can be classified into install-time, ongoing and post-incident. Ongoing and install-time obligations are most common. The majority (denoted *V* in the table) use vague terms like proper maintenance and operation without a concrete definition for what this entails. However, some warranties are exceptions in providing prescriptive obligations (denoted *P*). These vendors tend to offer higher limits. For example, one vendor requires a “differential security analysis” whenever the buyer modifies software covered by the warranty. Another vendor requires the client to relinquish write access to the product and allow the vendor to configure security functions like the whitelist. Post-incident obligations concern when and how the client must notify the vendor after discovering the incident.

There is significant diversity in terms of what cyber-incident warranties exclude. For example, a back-up provider excludes “any breach due to weak or stolen credentials” or denial of service. A monitoring product excludes breaches that are not a result of Advanced Persistent Threat (APT) activity. A firm offering source code review excludes coverage if the attack results from unknown vulnerabilities, defined elaborately using the Common Vulnerabilities and Exposures (CVE) database and a list of 122 known vulnerabilities.

Consumers might worry about the vendor’s ability to fund the indemnity payment. Some cyber-incident warranties were backed by insurance. For example, one vendor claimed to be “underwritten by an A-rated, internationally known insurance carrier.” A different vendor suggested their relationship with insurers meant purchasing the product “could result in better terms on cyber insurance.”

Our corpus represents close to the population of cyber-incident warranties while only comprising a sample of cyber-product warranties. The latter are predominantly offered by firms selling physical devices to be deployed in the buyer’s network. The corresponding warranties are less diverse and less likely to be announced

Examining the details within cyber warranties.

Columns refer to: a description of the vendor, whether we have the actual contract (Yes (Y) or No (N)), the type and amount of coverage offered and whether there are obligations (Vague (V) or prescriptive (P)) for install time, ongoing, or post-incident.

| Description | Contract | Coverage Type | Coverage Amount | Install time | Ongoing | Post-incident |
|----------------------|----------|------------------|--------------------------------------|--------------|---------|---------------|
| Routers | Y | P | Repair or Replace | | V | |
| Access control | Y | P | Repair or Replace | V | Y | |
| Security tokens | Y | P | Repair or Replace | | | P |
| Network management | Y | P | Repair or Replace | | | |
| End-point protection | Y | P | Repair or Replace | V | Y | |
| Network architecture | Y | P | Repair or Replace | V | V | |
| Various products | N | P | Repair or Replace | | V | |
| Firewalls | Y | P | Repair or Replace | | V | |
| Routers | Y | P | Repair or Replace | V | V | |
| Firewalls | Y | P | Repair or Replace | V | V | |
| Source code review | Y | I | \$5,000,000 | | V | 30 days |
| Back-up services | Y | I | \$10,000/\$50,000 | V | P | 30 days |
| End-point | N | I | \$100,000/\$500,000/\$1,000,000 | | | P |
| Monitoring | Y | I | \$1,000,000 or 2x License fee | V | V | |
| End-point protection | N | I ^{RWP} | \$1,000,000 or \$1,000 per end-point | | | |

publicly. This can be contrasted with cyber-incident warranties, which are announced publicly to generate coverage from security reporters.

Looking Forward

Warranties must transfer non-negligible amounts of liability to vendors in order to meaningfully overcome the market for lemons. Our preliminary analysis suggests the majority of cyber warranties cover the cost of repairing the device alone. Only cyber-incident warranties cover first-party costs from cyber-attacks—why all such warranties were offered by firms selling intangible products is an open question. Consumers should question whether warranties can function as a costly signal when narrow coverage means vendors accept little risk.

Worse still, buyers cannot compare across cyber-incident warranty contracts due to the diversity of obligations and exclusions. Ambiguous definitions of the buyer’s obligations and excluded events create uncertainty over what is covered. Moving toward standardized terms and conditions may help consumers, as has been pursued in cyber insurance, but this is in tension with innovation and product diversity.

The scope of the product drives warranty terms and conditions. The source code review firms are reasonable in only indemnifying losses resulting

from known vulnerabilities with a corresponding CVE number, which protects the vendor from incurring costs from zero-day attacks. Exclusions like the monitoring firm only indemnifying losses resulting from AdvaPT activity are less reasonable given non-APT attacks are presumably easier to detect and much more common.

Warranties with many obligations and exclusions at least communicate the attached product’s limitations. Prescriptive ongoing obligations from end-point protection firms demonstrate how security is about more than just buying the right product. In fact, the expertise of security professionals is so important that one firm invalidates coverage unless the buyer relinquishes write access to the platform.

Theoretical work⁵ suggests both the breadth of the warranty and the price of a product determine whether the warranty functions as a quality signal. Our analysis has not touched upon the price of these products. It could be that firms with ineffective products pass the cost of the warranty on to buyers via higher prices. Future studies could analyze warranties and price together to probe this issue.

In conclusion, cyber warranties—particularly cyber-product warranties—do not transfer enough risk to be a market fix as imagined in Woods.⁵ But this does not mean they are pure marketing

tricks either. The most valuable feature of warranties is in preventing vendors from exaggerating what their products can do. Consumers who read the fine print can place greater trust in marketing claims so long as the functionality is covered by a cyber-incident warranty. **□**

References

1. Akerlof, G.A. The market for 'lemons': Quality uncertainty and the market mechanism. *Uncertainty in Economics*. P.Diamond and A. Rothschild, eds. Elsevier, 1978, 235–251.
2. Anderson, R. Why information security is hard—An economic perspective. In *Proceedings of the 17th Annual Computer Security Applications Conf. IEEE*, 2001, 358–365. IEEE, 2001.
3. Linden, D. and Rashid, A. The effect of software warranties on cybersecurity. *ACM SIGSOFT Software Engineering Notes* 43, 4 (2019), 31–35.
4. Rustad, M.L. and Koenig, T.H. The tort of negligent enablement of cybercrime. *Berkeley Tech. LJ* 20 (2005), 1553.
5. Woods, D.W. and Simpson, A.C. Cyber-warranties as a quality signal for information security products. In *Proceedings of the 9th Conf. Decision and Game Theory for Security*. Springer, 2018, 22–27.

Daniel W. Woods (daniel.woods@uibk.ac.at) is a post-doctoral researcher in the Department of Computer Science at the University of Innsbruck, Austria. He completed this work as a Fulbright Cyber Security Scholar.

Tyler Moore (tyler-moore@utulsa.edu) is the Tandy Associate Professor of Cyber Security and Information Assurance in the Tandy School of Computer Science at the University of Tulsa, OK, USA.

Copyright held by authors/owners. Publication rights licensed to ACM.



Watch the authors discuss this work in the exclusive *Communications* video. <https://caacm.acm.org/videos/cyber-warranties>

The discovery of this calculating machine is so significant that part of the history of ancient technology must be rewritten.

BY HERBERT BRUDERER

The Antikythera Mechanism

UNTIL THE DISCOVERY of the Antikythera Mechanism, astrolabes were often considered the earliest analog mathematical devices. Such complex gearwork as in this astronomical calculator, however, only appeared (again) much later, especially in medieval clockworks. Leonardo da Vinci (1452–1519) knew gears, as his drawings show. Heron of Alexandria (1st century) used cogwheels for his pantograph. The construction of analog measuring and drawing instruments (for example, sectors, proportional dividers, compasses) and logarithmic circular and cylindrical slide rules was comparatively simple. Planimeters and (mechanical) differential analyzers were sophisticated. The first mechanical calculating machines were invented in the 17th century (Wilhelm Schickard, Blaise Pascal, Gottfried Leibniz). These digital devices required stepped drums, pinwheels, and accumulators. In the second half of the 20th century there was a competition between electronic analog computers and electronic digital computers.

This article is not about new groundbreaking insights. Rather, it presents an overview of decades of effort and different views. The review is not aimed at experts, but at computer scientists who are interested in the history of technology.

Some consider the Antikythera Mechanism (see Figures 1–14)—an astronomical calculator—as the world’s first analog calculator. This article is based on an international survey among the leading specialists for the Antikythera Mechanism and an adaptation of a chapter of the book on the history of computing by the author and its English translation.^{3,4,5} For more explanation, see the sidebar “Structure of the Antikythera Mechanism.”

It is not easy for most laypeople to understand the movements of the celestial bodies. The structure of the Metonic dial “was very unusual, having two distinct centers. We will call one of them the axial center because this was the location of the axle or arbor bearing the dial’s pointer ... On the right side of the plate, a series of five concentric semicircular slots ... were cut through the plate. On the left side was another series of five concentric semicircular slots, centered on the secondary center.¹⁷ (To learn more about the structure, the astronomical functions, and the details of the gearing, see Seiradakis,¹⁹ Edmunds,⁹ and Freeth.¹⁴)

Astrolabe, Planetarium, or Calendar Calculator?

For a long time, the purpose of the Antikythera Mechanism—a bronze

» key insights

- We do still not yet know who constructed the Antikythera Mechanism and where it was made.
- It seems unlikely the great Greek scientist and engineer Archimedes invented the Antikythera Mechanism, but he may well have contributed to the development of such mechanisms.
- There must have been a long tradition in building such complex astronomical gear works.

gearwork in a wooden case—was unknown. Was the approximately 32cm–33cm high, 17cm–18cm wide, and at least 8cm deep, shoe-box sized device an astrolabe, a planetarium, or a calculator?

This question has been resolved today, unlike many others. It is an *astronomical calculating machine*. The device determines the approximate position of the sun, the moon and—as can be inferred from the texts on the device—possibly the (five then known) planets and serves as a calendar. It predicted or described solar and lunar eclipse possibilities based on the Saros cycle and calculated the phases of the moon. The machine also showed the data for the four Panhellenic games (the Isthmia, Olympia, Nemea, and Pythia) as well as the minor Naa of Dodona and Halieia of Rhodes. The scales are concentric on the front. The major cycles of the back (Metonic and Saros dials) were spiral-shaped. Greek (astronomical and technical) texts were found on the covers of both sides of the device. Almost certainly the machine contains over 40 individual gears (including any plausible reconstruction of the lost planetary gearwork). The fixed programmed calculator was presumably operated by a lateral knob or a crank. The development of such astronomical instruments apparently began in the 3rd century B.C. The mechanism is so mature that it can hardly be a unique device. In the history of culture, technology, and science, gearworks are of outstanding importance. Such complex constructions only reappear in Europe with the astronomical tower clocks in the 14th century, more than 1,000 years later. The Antikythera Mechanism is probably the world's the first analog calculator. Alexander Jones of New York University believes the device was designed primarily for educational and philosophical purposes. It certainly was good for demonstrations. For astronomers it was probably too imprecise, and unsuitable for navigation.

Figure 1. The complex astronomical calculator, over 2,000 years old, was discovered in 1901 in the sea off the Greek island of Antikythera. The discovery of the mysterious technological marvel was a big surprise. It is still unknown where the device was manufactured and who invented it. The opinions about its age vary by about 120 years. There are numerous physical and virtual replicas. Research groups from Greece, the U.K. and the U.S. are trying to elicit the last secrets from the device. The Antikythera mechanism and the astrolabes are considered to be the first analogue calculators (courtesy of National Archaeological Museum, Athens/Costas Xenikakis).



Figure 2. The front of the Antikythera Mechanism shows seven hands (sun, moon, five planets) and a double ring scale (outside: Egyptian calendar, inside: zodiac). At the top and bottom of the digital model is the Parapegma inscription (courtesy of Hublot, with data from the Antikythera Mechanism Research Project).



Figure 3. The rear side of the mechanism (digital reconstruction) has two spiral scales. Above: Metonic cycle with display of the Callippic cycle (restored) and the Panhellenic games; below: Saros cycle for eclipses and Exeligmos cycle (courtesy of Hublot, with data from the Antikythera Mechanism Research Project).



Structure of the Antikythera Mechanism

Front middle

- ▶ two concentric dials
- ▶ sun pointer for the movements of the sun
- ▶ moon pointer for the movements of the moon and for the moon phases
- ▶ probably five planetary hands

outer dial

- ▶ Egyptian calendar
- ▶ 365 days
- ▶ 12 (Egyptian) months with 30 days + 5 additional days

inner dial

- ▶ zodiac

Rear above

- ▶ dial with the Metonic cycle (19 years) (Metonic spiral)
- ▶ 235 lunar months: 125 months with 30 days and 110 months with 29 days

inside this dial:

- ▶ probable dial with the Callippic cycle (76 years)
- ▶ dial with Panhellenic, Naa, and Halieia games (four quadrants)

below

- ▶ dial with the Saros cycle
- ▶ 223 lunar months (around 18 years)
- ▶ glyphs indicating solar and lunar eclipses

inside this dial:

- ▶ Exeligmos cycle (54 years)
- ▶ 669 lunar months

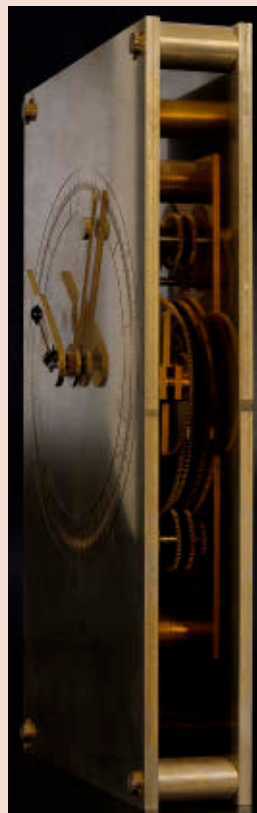
Annotations

The concentric dial on the front side had several pointers, the four or five dials on the back side had one pointer on each. The Metonic and the Saros cycles were spirals, the other cycles of the back face were circular dials. The Antikythera Mechanism skillfully determines which months have 29 or 30 days. The star calendar (parapegma) is written on the front of the case. The dial with the Callippic cycle (76 years) is not actually preserved. The version of the Egyptian calendar represented on this dial assumed a constant calendar year of 365 days. There were no leap days. For this reason, the Egyptian calendar scale was manually adjustable to allow different alignments with the Sun's apparent motion through the zodiac.

Figure 4. Computer-aided simulation of the front side by Tony Freeth (courtesy of Tony Freeth, Images First Ltd., London).



Figure 5. The Swiss clockmaker Ludwig Oechslin built a replica with sun and moon pointers, but without the presumed planetary hands (courtesy of Ochs und junior AG, Lucerne).



Why astronomical calculators have rarely survived? The bronze gears of the Antikythera Mechanism were only 2mm thin. This fine, fragile construction may explain why hardly any devices have survived. Metal was precious and was therefore recycled. These devices were not gold-plated artifacts with precious stones.

The *Antikythera Mechanism Research Project*. In 2005, an international research community called Antikythera Mechanism Research Project was founded. It is mainly made up of experts from the fields of astronomy, physics, astrophysics, mathematics, engineering, history of technology and science, archaeology, and classicists. For further information, see <http://antikythera-mechanism.gr>.

When was the astronomical calculator found? The Ionian island of Antikythera, originally called Aigila, is located between the Peloponnese peninsula and Crete, opposite Kythera (hence Antikythera). The shipwreck was found by sponge divers in 1900. The Antikythera Mechanism came to light in the summer of 1901 (probably July). Dives were also carried out in 1953, 1972, and several since 2012 up to the present year. The mechanism is only partially preserved and consists of 82 damaged fragments. Fundamental investigations have only been carried since the 1950s. Tomographic methods were also used for this purpose.

When did the ship go down? As can be seen from the finds of coins and amphorae, the ship sank between 70 B.C. and 50 B.C. This date period is generally accepted. "Around 60 B.C., a ship was wrecked of the northeast coast of a small island called Aigila in the straits between Crete and the Peloponnese ... The exact character of the ship is not known, but it was probably a large merchant vessel, perhaps about 40 meters long."¹⁷ The vessel may have been on its way from Asia Minor to the western Mediterranean. Perhaps the sail freighter was about 10m wide and could load 250 tons. The cargo included silver and bronze coins dating between 85 B.C. and 60 B.C.

When was the ship built? According to the 2010 radiocarbon analyses by Andrew Wilson (University of Oxford), the wood used for shipbuilding

is estimated (with a probability of 84.8%) to have been cut between 211 B.C. and 40 B.C. This can be seen from the new calibration curves of radiocarbon dating (C-14 method, 14C method).⁷ Since wooden ships do not last indefinitely long and are not permanently seaworthy, the boat is likely to have been manufactured at the earliest a few decades before the shipwreck.

When was the astronomical calculator built? Opinions on the year of manufacture of the astronomical calculator vary widely. The estimates range from 205 B.C. to 50 B.C. Christián Carman, James Evans, and Tony Freeth assume a production approximately 205 B.C. Michael Edmunds, Paul Iversen, Alexander Jones, and Michael Wright however believe in a much later production at a time when the ship was much closer to sinking.

Michael Edmunds from the University of Cardiff writes: “The present

best estimate of its construction date is around the middle of the range 150 B.C.—60 B.C.—although a date as early as early as 220 B.C. is not completely ruled out.”⁷ The astrophysicist adds: “My preferred period is 140 B.C.—70 B.C. But there must have been earlier, probably simpler versions. So, one would guess that similar mechanical devices might date from 200 B.C. or maybe 250 B.C.”⁷ According to Edmunds, there are references in the literature mentioning that astronomical devices were made or at least known from 250 B.C. to at least 500 A.D.

The historian of science and classicist Alexander Jones writes in this context: “We are obviously a long way from being able to put together a coherent story of the evolution and eventual degeneration of the ancient tradition of astronomical mechanisms, but there is enough evidence to suggest that complex and scientifically ambi-

tious mechanisms were being made at least through the three centuries from about 100 B.C. to A.D. 200, and that the people who were most likely to encounter them were mechanics, philosophers, and scientists.”¹⁷

Paul Iversen from Case Western University, Cleveland, believes in a late production of the Mechanism: “I would say the Mechanism was manufactured soon before the shipwreck of about 70 BCE–50 BCE, but in any case, probably not more than one generation, or about 100 BCE at the earliest.”¹⁶

Jones shares a similar opinion: “A far simpler hypothesis, however, is that the Mechanism was made somewhere around the Aegean not long before the shipwreck and was on its way to its intended home by a route that would next have proceeded up the Adriatic toward, say, Brundisium, stopping somewhere along the way

Figure 6. Digital reconstruction by Tony Freeth: on the left the front side with the two concentric scales and seven pointers, on the right the back side with the spiral scales, among others for the display of solar and lunar eclipses (courtesy of Tony Freeth, Images First Ltd., London).



to deliver part of the cargo. Occam's razor thus makes it probable that the Mechanism was commissioned by someone who lived in or near Epirus in the first half of the first century B.C."¹⁷

Jones adds: "I argued that the archaeological context favors the hypothesis that the Mechanism was new when it was lost in the wreck, because otherwise it becomes difficult to account for the presence of an antique object that was manifestly made for a locality west of the Aegean in a cargo

originating in the Aegean and destined for points west."¹⁷

Michael Wright (formerly of the Science Museum, London): "There is, however, no good argument for suggesting that the instrument was designed that early [205 B.C.] and there is a counterargument that the several displays were adjusted to mutual agreement in a way that could not have been done before the latter half of the second century B.C. The most likely explanation is that the designer of these

displays drew on old information."²⁰

The physicist Wright, however, rejects the assumption the instrument was new when the ship sank. Part of the device was mechanically confused. He writes: "I think it very unlikely that the instrument was very old at the time because I think it simply would not have lasted very long without being destroyed by use and handling. I suggest that it was probably built with a generation or so of its loss; that is, within a few decades of 100 B.C."²⁰

Figure 7. The replica from Thessaloniki is equipped with a cover on both sides in contrast to other real and digital replicas (courtesy of 3D Solidforms, Thessaloniki).



Figure 8. A view into the complex gearing of the replica of Thessaloniki (courtesy of 3D Solidforms, Thessaloniki).



Figure 9. This illustration shows the front side of the reconstruction by Markos Skoulatos and is fully functional (courtesy of Markos Skoulatos).

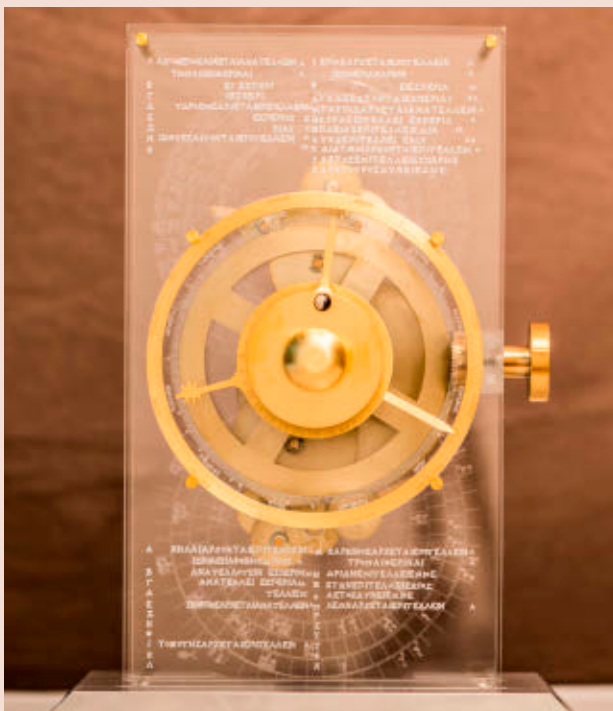
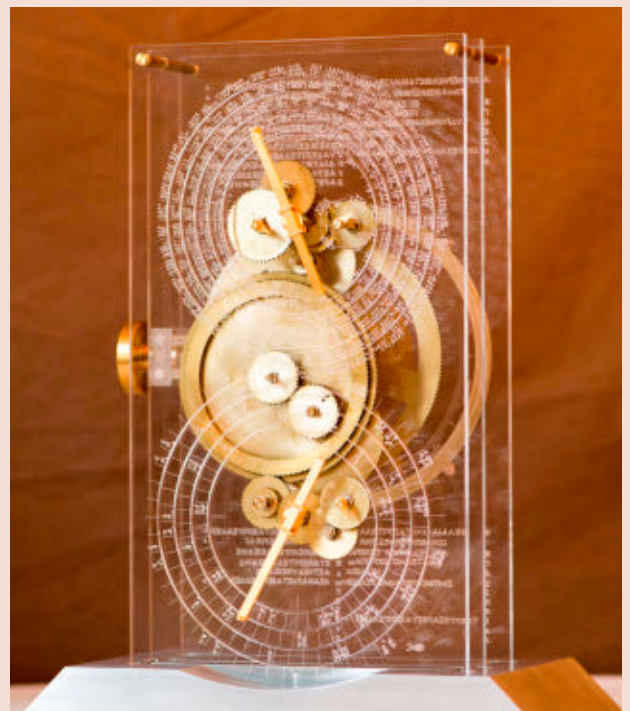


Figure 10. The illustration shows the rear side of Skoulatos' model. The transparent model allows one to follow the functioning of the mechanism (courtesy of Markos Skoulatos).



The physicist James Evans from the University of Puget Sound, Tacoma, WA, tends to assume an early production: “The eclipse predictor best fits an 18-year Saros cycle that started in 205 B.C. One or two Saros cycles later would also work, though with somewhat larger errors. Of course, we cannot rule out the possibility that it was built considerably later but using an out-of-date eclipse cycle.”¹¹

London-based Tony Freeth, on the other hand, assumes a construction around 205 B.C.: The prediction of the solar and lunar eclipses is based on the Saros cycle. The display on the back of the mechanism is intended to allow the determination of its age.¹³

According to Christián Carman and James Evans, the eclipse dial works best when the full moon of the first month of the Saros cycle reaches May 12, 205 B.C.⁶

If the astronomical calculator had been manufactured by Archimedes during his lifetime, it would have been approximately 150 years old when the ship sank. Such an early production does not seem very plausible.

Where was the mechanism made? The origin of the mechanism is unknown. Sicily was once thought to be the place of production, current thinking puts Rhodes as the likely location.

Possibilities include Alexandria, Pergamon, Syracuse and Rhodes. Syracuse had the advantage of any heritage left by Archimedes, but the problem that it was sacked in at the time of his death in 211 B.C., although something may have remained. The best candidate must be Rhodes, a port at which the Antikythera ship had called (judged by some of its cargo) not long before its wreck. Rhodes was a highly technological naval center around 100 B.C. with a fine bronze industry and an astronomical tradition. It is also one place where we know that a similar contemporary device was reputedly made and seen.”¹⁶

Edmunds adds that the star calendar (Parapegma) on the wheels corresponds to the geographical latitude of Rhodes and that Cicero had seen a comparable device on Rhodes in the first century B.C.⁸

The classicist and epigrapher Iversen contends the computer was most

Figure 11. Markos Skoulatos and Georg Brandl also created a digital model, operated via a portable computer or a smartphone (courtesy of Markos Skoulatos and Georg Brandl).



Figure 12. This picture shows the help menu for the front side (courtesy of Markos Skoulatos and Georg Brandl).

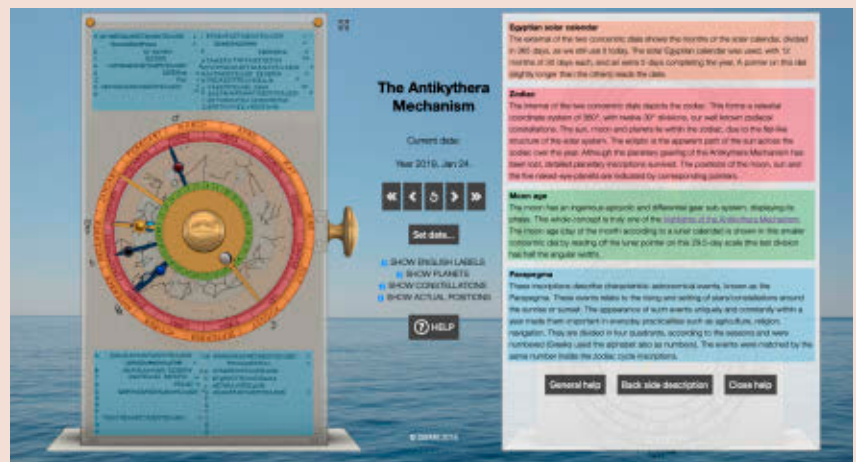
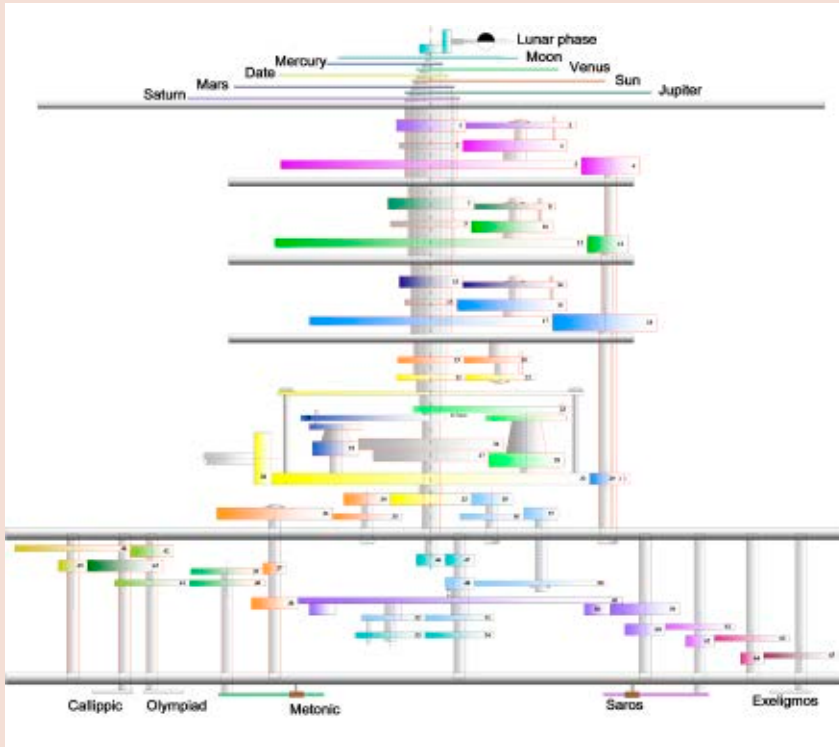


Figure 13. This picture shows the help menu for the rear side (courtesy of Markos Skoulatos and Georg Brandl).



Figure 14. One of the 48 models designed by two Chinese researchers for the design of the Antikythera Mechanism. Above are shown the hands for sun, moon and five planets, below four astronomical cycles and the course of the Panhellenic games (courtesy of Jian-Liang Lin and Hong-Sen Yan).



likely to be manufactured in Rhodes for several reasons:¹⁶

- ▶ The games of Halieia, played in honor of the sun god, appear beside the Panhellenic games on the back of the device. The Halieia are mostly attested only on Rhodes or the territory it controlled on the mainland opposite the island known as the Rhodian Peraia. They are mentioned in the Doric not Attic-Ionic dialect, the latter of which is customary on the rest of the Mechanism. Dorian halios means sun.

- ▶ We know from writings that such devices were manufactured in Rhodes at the time of the shipwreck. This is evidenced by writings written during this period. Cicero describes such a machine built by his teacher, Poseidonios. This scholar lived on Rhodes. At this time, Geminus, who in his treatise explains the astronomical theory underlying the gearwork, probably also worked on this island.

- ▶ In the 1st century B.C., around the time of its demise, the island of Rhodes was a center for astronomy.

- ▶ The astronomers of Rhodes used the Attic-Ionic dialect in the 1st centu-

ry. On the island, a single inscription written in this scientific language was found.

- ▶ The alphabetical lists of the annual astronomical events (for example, solstice, equinox) for the sun and the fixed stars (Parapegma) contained on the front of the Mechanism are best for the northern latitudes 33.3 to 37.0 and thus for Rhodes. However, they do not apply to Epirus (Western Greece) or Alexandria.

Iversen considers it unlikely that Archimedes built the device, because it makes use of certain findings on the movement of the sun and moon, which are attributed to Hipparchos (around 150 B.C.). For example, Poseidonios or an employee of his school could be considered as the creator. The customer is likely to be from Epirus due to the epirotic calendar (Metonic spiral).¹⁶

Cicero mentions in *De Natura Deorum*, book 2, (45 B.C.), the “sphaera” of Posidonius. The word “sphaera” (Greek sphaera) has several meanings: “We need to be careful to distinguish mechanized planetaria from certain other concepts and categories of ob-

jects that may be described in similar language. The Greek word sphaera (or Latin sphaera) may refer to an astronomical mechanism but was also appropriate for a simple globe.”¹⁷

The star calendar is also associated with the Greek author Geminus, who presumably lived in the 1st century.

The mathematician Freeth, on the other hand, assumes the original form of the astronomical computer originated from Archimedes. This famous scholar, who died in 211 B.C., lived in the Corinthian colony in Syracuse. According to Cicero, the great Greek scholar is said to have made such an instrument. The sophisticated state of the mechanism was a surprise. Until the discovery of Price, nothing comparable was known in ancient Greek technology. Freeth writes: “I personally think it is likely that the original design came from Archimedes and he started the tradition of making these devices. The Antikythera Mechanism is simply a later version of the Archimedes design. But there is little hard evidence. ... The sophistication of the mechanism, when uncovered by Price, was astonishing, given what had previously been known about ancient Greek technology.”¹³

According to Wright, however, the mechanism of Archimedes was a completely different device, namely a mechanical celestial globe. The English man, who also works as a mechanic, has reconstructed it. Cicero, who reported in *De Re Publica* (book 1, 54 B.C.–51 B.C.) and in *Tusculanae Disputations* (book 1, 45 B.C.) on the “sphaera” of Archimedes, has lived, however, over 100 years later than the outstanding Sicilian researcher.

Kyriakos Efstathiou (University of Thessaloniki) suggests that at this time one of the most important Greek astronomers, Hipparchos, lived in Rhodes. Many researchers believed he, his disciple Poseidonios, or someone from the astronomy school there could be considered the creator of the mechanism.¹⁰

The most convincing assumption is the Antikythera Mechanism comes from the surroundings of Poseidonios. The stoic philosopher had no astronomical or craft skills himself. There are obviously close relations between the Mechanism and Hipparchos as well as Geminus.

Rebuilds

There are numerous real replicas (reconstructions) of the Antikythera Mechanism and some virtual models (simulations). Among the best known real (that is, physical) replicas were the devices of Ioannis Theofanidis (Greece, 1934), Derek de Solla Price and Robert Deroski (U.S.), Allan Bromley and Frank Percival (Australia), John Gleave (U.K.), Michael Wright (U.K.) as well as John Seiradakis and Kyriakos Efstathiou (Greece). Further models have been produced by Dionysios Kriaris (Greece), Massimo Vicentini (Italy), and Tatjana van Vark (Netherlands). However, some replicas are not operational and differ from the original design. 3D Solidforms sells the devices developed in cooperation with the Aristotle University of Thessaloniki. Markos Skoulatos and Georg Brandl (Germany) have built new real and virtual replicas.

Digital replicas are available, for example, from Tony Freeth (U.K.). In Switzerland, the Mechanism was also reconstructed, for example by Ludwig Oechslin (formerly International Watch Museum in La Chaux-de-Fonds). Matthias Buttet created for Hublot SA, Nyon VD a watch that incorporates the functions of the Antikythera Mechanism.

The most important (real) replicas are the reconstructions of Michael Wright and the Antikythera Mechanism Research Project (2006). Wright, the leading model maker of Antikythera Mechanism, justly points out that a computer-generated 3D image does not have the same persuasive power as a physical model. In the artificial world there is neither mass, inertia, force, friction, nor elastic or inelastic deflection. Questions of material strength and wear properties are excluded.

The physicist Markos Skoulatos of the University of Technology in Munich designed a real reconstruction and then followed with a digital model, developed together with the physicist Georg Brandl. The mechanical reconstruction exhibits less friction and the virtual model high accuracy.

Is the Antikythera Mechanism an Analog Device?

The (non-programmable) mechanism

of Antikythera is sometimes called the world's oldest analog "computer." The gear trains refer to the orbits of the planets. Due to this similarity, the mechanical calendar computer appears as an analog device. In addition to the input, the output of this machine is also analog: the dials (scales) and the continuously rotating hands.

However, the calendars are calculated digitally. The number of teeth on the gears is always an integer. The ratio between two gears is always a rational number. These relations reflect the celestial movements, for example, in the Metonic cycle (x cycles in y years, where x and y are integers). Rational numbers are numbers that can be represented as a quotient of two integers. In addition to the number zero, this includes all (positive and negative) whole and fractional numbers.

The display is analog, but the gear-work operates digitally. Most researchers regard the astronomical marvel as an analog device. However, it can also be understood as a mixed calculator. Astronomical clocks are also hybrid, as are digital clocks with analog display. A numerical display is more precise than analog pointers but less comfortable to read.

For further explanations see *Computation and its Limits* by P. Cockshott, L.M. Mackenzie, and G. Michaelson (Oxford Press 2012).

Conclusion

The current state of research can be summarized as follows: The Antikythera Mechanism, discovered in 1901, was lost about 60 B.C. when a Roman merchant ship sank in the Mediterranean Sea. The complex astronomical calculator was probably built on the island of Rhodes near the Greek philosopher Poseidonios. The client for the teaching material seems to be a person in northwestern Greece.

Acknowledgment

I am very grateful to Michael G. Edmunds, Kyriakos Efstathiou, James C. Evans, Tony Freeth, Paul A. Iversen, Alexander R. Jones, and Michael T. Wright for their helpful answers in connection with my survey and for the valuable comments of the reviewers. Markos Skoulatos and Georg Brandl provided exciting information on their

new project. In addition, I would like to thank all for the permission to reproduce the fascinating images. □

References

- Bignasca, A., Lagogianni-Georgakarakos, M., Kaltsas, N., Vlachogianni, E. Eds. *Der versunkene Schatz. Das Schiffswrack von Antikythera*, Antikenmuseum Basel und Sammlung Ludwig, Basel 2015
- Bitsakis Y. Ein antiker mechanischer Kosmos. *Antike Welt* 5 (2015), 27–32
- Bruderer, H. *Meilensteine der Rechentechnik I. Mechanische Rechenmaschinen-Rechenschieber – historische Automaten – wissenschaftliche Instrumente*. de Gruyter Oldenbourg, Berlin/Boston, 2018; <https://www.degruyter.com/view/product/480555>.
- Bruderer, H. *Meilensteine der Rechentechnik II. Erfindung des Computers – Elektronenrechner – Entwicklungen in Deutschland, England und der Schweiz*. de Gruyter Oldenbourg, Berlin/Boston, 2018; <https://www.degruyter.com/view/product/503373>.
- Bruderer, H. *Milestones in Analog and Digital Computing*. Springer Nature Switzerland AG, Cham, 3rd edition, 2020; <https://www.springer.com/de/book/9783030409739>.
- Carman, C.C., Evans, J.E. On the epoch of the Antikythera Mechanism and its eclipse predictor. *Archive for History of Exact Sciences* 68, 6 (Nov. 2014), 693–774
- Edmunds, M.G. The Antikythera mechanism and the mechanical universe. *Contemporary physics* 55, 4 (Dec. 2014), 263–285
- Edmunds, M.G. Personal communication, (U.K., Nov. 22, 2017).
- Edmunds, M.G. and Freeth, T. Using computation to decode the first known computer. *Computer* 44 (July 2011), 32–39.
- Efstathiou K. Personal communication, (Greece, Nov. 23, 2017).
- Evans J.C. Personal communication, (U.S., Nov. 21, 2017).
- Freeth, T. Eclipse prediction on the ancient Greek astronomical calculating machine known as the Antikythera Mechanism. *Plos One* 9, 7 (July 30, 2014) e103275; <https://doi.org/10.1371/journal.pone.0103275>
- Freeth, T. Personal communication, (U.K., Nov. 6, 2017).
- Freeth, T. Revisiting the eclipse prediction scheme in the Antikythera mechanism. *Palgrave Commun.* (2019); doi.org/10.1057/s41599-018-0210-9
- Iversen, P.A. The calendar on the Antikythera mechanism and the Corinthian family of calendars. *Hesperia* 86, (2017), 129–203
- Iversen, P.A. Personal communication, (U.K., Nov. 20, 2017).
- Jones, A.R. *A Portable Cosmos. Revealing the Antikythera Mechanism, Scientific Wonder of the Ancient World*. Oxford University Press, NY 2017.
- Jones, A.R. The Antikythera mechanism and the public face of Greek science. In *Proceedings of From Antikythera to the Square Kilometre Array: Lessons from the Ancients Conference* (Kerastari, Greece, June 12–15, 2012).
- Seiradakis, J.H. and Edmunds, M.G. Our current knowledge of the Antikythera mechanism. *Nature Astronomy* 2 (Jan. 2018), 35–42
- Wright, M.T. Personal communication (England, U.K., Nov. 22, 2017).

Useful websites

- ▶ <http://www.antikythera-mechanism.gr/>
- ▶ <http://www.antikythera-mechanism.gr/data/models/computer-models>
- ▶ <http://www.antikythera-mechanism.gr/data/models/solid-models>
- ▶ <http://www.antikythera-mechanism.gr/history/people/michael-t-wright>
- ▶ <http://www.antikythera-mechanism.gr/museum>
- ▶ <https://www.hublot.com/antikythera/>
- ▶ <https://www.theantikytheramechanism.com/>

Herbert Bruderer (bruderer@retired.ethz.ch; herbert.bruderer@bluewin.ch) is a retired lecturer in computer science at ETH Zürich; more recently, he has been a historian of technology and was co-organizer of the International Turing Conference at ETH Zürich in 2012.

Copyright held by author/owner.
Publication rights licensed to ACM.

Attention: Undergraduate and Graduate Computing Students

There's an **ACM Student Research Competition (SRC)**
at a SIG Conference of interest to you!



Association for Computing Machinery
Advancing Computing as a Science & Profession

SPONSORED BY Microsoft

It's hard to put the **ACM Student Research Competition** experience into words, but we'll try...



"Attending ACM SRC was a transformative experience for me. It was an opportunity to take my research to a new level, beyond the network of my home university. Most important, it was a chance to make new connections and encounter new ideas that had a lasting impact on my academic life. I can't recommend ACM SRC enough to any student who is looking to expand the horizons of their research endeavors."

David Mueller
North Carolina State University | SIGDOC 2018



"Participating in the ACM SRC was a unique opportunity for practicing my presentation skills, getting feedback on my work, and networking with both leading researchers and fellow SRC participants. Winning the competition was a great honor, a motivation to continue working in research, and a useful boost for my career. I highly recommend any aspiring student researcher to participate in the SRC."

Manuel Rigger
Johannes Kepler University Linz, Austria | Programming 2018



"The SRC was a great chance to present early results of my work to an international audience. Especially the feedback during the poster session helped me to steer my work in the right direction and gave me a huge motivation boost. Together with the connections and friendships I made, I found the SRC to be a positive experience."

Matthias Springer
Tokyo Institute of Technology | SPLASH 2018



"I have been a part of many conferences before both as an author and as a volunteer but I found SRC to be an incredible conference experience. It gave me the opportunity to have the most immersive experience, improving my skills as a presenter, researcher, and scientist. Over the several phases of ACM SRC, I had the opportunity to present my work both formally (as a research talk and research paper) and informally (in poster or demonstration session). Having talked to a diverse range of researchers, I believe my work has much broader visibility now and I was able to get deep insights and feedback on my future projects. ACM SRC played a critical role in facilitating my research, giving me the most productive conference experience."

Muhammad Ali Gulzar
University of California, Los Angeles | ICSE 2018



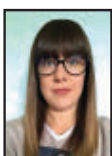
"At the ACM SRC, I got to learn about the work done in a variety of different research areas and experience the energy and enthusiasm of everyone involved. I was extremely inspired by my fellow competitors and was happy to discover better ways of explaining my own work to others. I would like to specifically encourage undergraduate students to not hesitate and apply! Thank you to all those who make this competition possible for students like me."

Elizaveta Tremsina
UC Berkeley | TAPIA 2018



"The ACM SRC was an incredible opportunity for me to present my research to a wide audience of experts. I received invaluable, supportive feedback about my research and presentation style, and I am sure that the lessons I learned from the experience will stay with me for the rest of my career as a researcher. Participating in the SRC has also made me feel much more comfortable speaking to other researchers in my field, both about my work as well as projects I am not involved in. I would strongly recommend students interested in research to apply to an ACM SRC—there's really no reason not to!"

Justin Lubin
University of Chicago | SPLASH 2018



"Joining the Student Research Competition of ACM gave me the opportunity to measure my skills as a researcher and to carry out a preliminary study by myself. Moreover, I believe that "healthy competition" is always challenging in order to improve yourself. I suggest that every Ph.D. student try this experience."

Gemma Catolino
University of Salerno | MobileSoft 2018

Check the SRC Submission Dates: <https://src.acm.org/submissions>

- ◆ Participants receive: \$500 (USD) travel expenses
- ◆ All Winners receive a medal and monetary award. First place winners advance to the SRC Grand Finals
- ◆ Grand Finals Winners receive a handsome certificate and monetary award at the ACM Awards Banquet

Questions? Contact Nanette Hernandez, ACM's SRC Coordinator: hernandez@hq.acm.org



research highlights

P. 118

**Technical
Perspective**
**An Answer to Fair
Division's Most
Enigmatic Question**

By Ariel D. Procaccia

P. 119

**A Bounded and Envy-Free
Cake Cutting Algorithm**

By Haris Aziz and Simon Mackenzie

Technical Perspective

An Answer to Fair Division's Most Enigmatic Question

By Ariel D. Procaccia

THE CAKE-CUTTING PROBLEM is the brainchild of the noted mathematician Hugo Steinhaus, who formulated and studied it in the early 1940s, even as he was hiding from the Nazis who occupied his native Poland. The question Steinhaus asked is one that must have occurred to many others too (albeit probably under circumstances that afford greater accessibility to cake): How does one fairly divide a cake between multiple people? The difficulty is that the cake is heterogeneous, and the participants have different preferences, so simply giving them pieces of equal size will not do.

On a conceptual level, Steinhaus' main insight was that fairness—an ostensibly abstract idea—can be specified mathematically. One particular notion has emerged as the epitome of fairness: envy-freeness, which means that each participant prefers her piece of cake to the piece given to any other participant.

So, what is an algorithm that would produce an envy-free division of the cake for two players? Easy: I cut; you choose. I am not envious because I am indifferent between the two pieces, and you get the piece that you prefer. And for three players? That is already tricky. The general case was open for decades and considered a major open problem, until it was solved in 1995 by Steven Brams and Alan Taylor.

The envy-free protocol of Brams and Taylor is guaranteed to terminate with an envy-free allocation of the cake. However, the running time of this algorithm is unbounded: by carefully tuning the preferences of the participants, it is possible to make the protocol perform an arbitrarily large number of steps. Consequently, as soon as Brams and Taylor solved the envy-free cake-cutting problem, they immediately launched a new problem to the top of fair division's most wanted list: the existence of a *bounded* envy-free cake-cutting protocol.

The allocation of indivisible goods gives rise to questions that are both mathematically challenging and deeply practical.


This problem stood its ground for two decades, until it was cracked by Aziz and Mackenzie in 2016. Their solution is presented in the following paper; it is an intricate protocol, which builds on previous ideas while adding several ingenious ingredients into the mix.

So, should aspiring cake cutters hang up their spurs? Not quite yet. The number of steps required by the Aziz-Mackenzie protocol is bounded by a function of the number of participants—but that function grows comically fast. In fact, for just two participants, the bound is a number whose number of digits is so large that it itself has almost 20,000 digits! This begs the question of whether there exists a cake-cutting protocol that is both envy free and computationally efficient.

Looking beyond cake cutting, fair division algorithms have been transitioning from theory to practice. For example, envy-free solutions to the rent division problem—assigning rooms to housemates and dividing the rent between them—are used widely. Other applications include allocating computational resources, assigning seats in college courses to students, and even eliminating gerrymandering

through provably fair political redistricting procedures.

The allocation of indivisible goods, in particular, gives rise to questions that are both mathematically challenging and deeply practical. A paradigmatic use case is dividing a jewelry or art collection between several heirs. Since the goods cannot be split between participants, envy-freeness cannot be guaranteed. Instead, let us allow one participant to prefer the bundle of goods allocated to another, but it must be the case that removing *any* good from the bundle of the latter participant will always eliminate the former's envy; this property is known as *envy-freeness up to any good*, or EFX for short. Furthermore, suppose the value each participant has for a bundle of goods is simply the sum of values of individual goods. Is an EFX allocation guaranteed to exist? This fundamental and deceptively accessible question is open. In my view, it is the successor of envy-free cake cutting as fair division's biggest problem.

Taking a broader perspective, in recent years the term “fairness” is being used by machine learning researchers to refer to lack of bias or discrimination. This viewpoint is rooted in Rawlsian ethics and might seem to be at odds with the preference-based notions of fairness favored by eight decades of research in fair division. Nevertheless, there are strong synergies between the two fields. In particular, well-established fairness notions like envy-freeness can—and should—help guide the design of ethical AI. 

Ariel D. Procaccia is Gordon McKay Professor of Computer Science at Harvard University, Cambridge, MA, USA.

Copyright held by author.

A Bounded and Envy-Free Cake Cutting Algorithm

By Haris Aziz and Simon Mackenzie

Abstract

We consider the well-studied cake cutting problem in which the goal is to find an envy-free allocation of a divisible resource based on queries from agents. The problem has received attention in mathematics, economics, and computer science. It has been a major open problem whether there exists a discrete and bounded envy-free protocol. We report on our algorithm that resolved the open problem.

1. INTRODUCTION

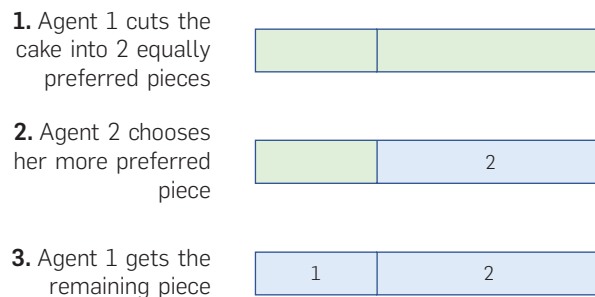
The cake cutting problem is a fundamental mathematical problem in which the ‘cake’ is a metaphor for a heterogeneous divisible resource represented by the unit interval $[0, 1]$.⁴ The resource could represent time, land, or some computational resources. The goal is to allocate the cake among n entities that are referred to as ‘agents.’ Each agent’s allocation consists of a collection of subintervals. Each of the agents is assumed to have additive and nonnegative valuations over segments of the interval. A cake-cutting algorithm/protocol interacts with the agents in order to identify a fair allocation.

One of the most important criteria for fairness is *envy-freeness*. An agent envies another if she would have preferred to receive the other’s piece rather than hers. A cake cutting protocol/algorithm is called *envy-free* if each agent is guaranteed to be nonenvious if she reports her real valuations. If a protocol is envy-free, then an honest agent will not be envious even if other agents misreport their valuations.

The interaction of the protocol with the agent uses two types of queries EVALUATE and CUT. EVALUATE asks an agent i to report her value for the subinterval between two points x and y . CUT asks an agent i to choose a point y such that the interval between a given x and y is worth a given value t . This natural query model was popularized by Robertson and Webb.⁸

How does an envy-free protocol look like? This is perhaps best illustrated with the most famous envy-free cake cutting protocol. It is the *Cut and Choose Protocol* for two agents. One agent is asked to divide the cake into two equally preferred pieces. The other agent is then asked to pick her most preferred piece whereas the cutter gets the remaining piece. The protocol is explained pictorially in Figure 1. It is formally specified as Algorithm 1. For a piece of cake D (which is just a subset of the cake), we write $V_i(D)$ to denote the agent i ’s value for the piece D . The proof that the Cut and Choose Protocol is envy-free is as follows. Agent 1 gets one of the equally preferred pieces so she is not envious. Agent 2 gets the piece that she prefers at least as much as the other piece so she is also not envious.

Figure 1. Cut and choose protocol.



Algorithm 1. Cut and choose protocol.

Input Cake $R = [0, 1]$ and two agents 1 and 2.

Output An envy-free allocation of R .

- 1: Ask agent 1 for her value $V_1(R)$. Then ask agent 1 to place a mark x on the cake so that $V_1(0, x) = V_1(x, 1)$. Divide the cake into two pieces $[0, x]$ and $[x, 1]$.
- 2: Ask agent 2 for her value $V_2(0, x)$ and $V_2(x, 1)$. If $V_2(0, x) \geq V_2(x, 1)$, give agent 2 piece $[0, x]$. Otherwise, give piece $[x, 1]$ to agent 2.
- 3: Give agent 1 the remaining piece.

Is there a cake cutting algorithm that is envy-free for three, four, or more number of agents? The question has been the topic of intense study in the past decades. It dates back to the work of mathematician Hugo Steinhaus who presented the cake cutting problem in the 1940s.^{8,9} For an enjoyable overview of the history of the cake cutting problem, we refer to the Communications of the ACM paper by Procaccia⁷ the popular book by Brams and Taylor.⁴ For the case of three agents, an elegant protocol was independently discovered by John L. Selfridge and John H. Conway around 1960. Before our work, a general envy-free cake cutting algorithm using a finite number of steps and cuts was proposed by Brams and Taylor.³ However, it can require an arbitrarily large number of steps, even for four agents. This led to the question of whether there exists a bounded envy-free algorithm. In other words, does there exist an envy-free algorithm that has a provable bound on the number of steps which is only dependent on a function of n (the number of agents)?

In this paper, we report on the first bounded and envy-free cake cutting algorithms.^{1,2} Next, we present the ideas behind the general algorithm.

The original version of this paper, entitled ‘‘A Discrete and Bounded Envy-Free Cake Cutting Protocol for Any Number of Agents,’’ was published in the *Proceedings of the 57th Symposium on Foundations of Computer Science*, (2016), 416–427.

2. THE PROTOCOL: AN OVERVIEW

At a high level, our protocol (which is referred to as the Main Protocol) allocates a large enough portion of the cake in an envy-free manner. After that, it tries to add some small portions of the unallocated cake to the allocated part in a structured and envy-free manner with the goal to reduce the problem to envy-free allocation for a smaller number of agents. Throughout the protocol, there is a partial allocation of the cake that is maintained to be envy-free. By partial we mean that the whole cake may not be allocated.

The Main Protocol makes calls to other protocols (in particular the Core Protocol, Discrepancy Protocol, and the GoLeft Protocol) in order to find an envy-free allocation. The Core Protocol is used to obtain an envy-free partial allocation. The Main Protocol applies it many times on the unallocated cake to make the unallocated cake smaller and smaller.

After finding a large enough envy-free partial allocation, the Main Protocol uses two possible ways to decompose our problem into one involving a smaller number of agents. The first case is when we find a situation where some agents are mainly interested in one part of the unallocated cake and other agents are mainly interested in the remaining part. This discrepancy in valuations of the agents is exploited by the Discrepancy Protocol. If the first case does not arise, we use the GoLeft Protocol to exchange suballocations of agents to enable one set of agents to “dominate” the other agents. The dominating agents think they will not be envious of the dominated agents even if one of the dominated agents gets all the unallocated cake. In that case, we reduce our problem to that involving the remaining cake and the dominated agents. Domination is a key idea on which our protocol is based and which helps us reduce our problem to a smaller problem. See Figure 2 for an overview of the Main Protocol.

Figure 3 presents a realizable sequence of steps that capture some of the key ideas of our protocol.

3. THE PROTOCOL: MORE DETAILS

In this section, we give more details of each of the components of the Main Protocol.

3.1. Core protocol

A crucial building block of our protocol is the Core Protocol which finds a partial allocation that is envy-free.

Figure 2. A bird’s-eye view of our protocol.

Main Protocol

Goal: Find an envy-free allocation of the whole cake.

1. Call the **Core Protocol** (that finds an envy-free partial allocation) several times to get a larger and larger envy-free allocation.
2. Decompose the problem into one with a smaller number of agents via two possible ways:
 - a) Call the **Discrepancy Protocol** (that exploits how agents value different parts of the unallocated cake): we get two smaller subproblems.
 - b) If there is no discrepancy, call the **GoLeft Protocol** (implements exchanges of some pieces to enable one set of agents to dominate the other agents). We get one smaller subproblem (with less number of agents).

The Core Protocol asks one of the n agents—the “cutter”—to divide the cake into n equally preferred pieces. Recall that this step is similar to the first step of the Cut and Choose Protocol. It then finds a possibly partial allocation in which each agent’s allocation is a contiguous piece of the cake. Each agent receives one of the pieces defined by the “cutter”. The agents may get the pieces in trimmed form. We guarantee the cutter as well as at least one other agent to get a full piece, and that no agent envies another agent. Another feature of the allocation is that for each piece that is partially allocated, the exact point at which it has been cut off corresponds to the mark by another agent to ensure she is not envious of that piece. When we first designed the Core Protocol, it was designed to establish the existence of an allocation that satisfies the properties discussed above. Once the existence of such an allocation is established, there is a simpler way to define a protocol which achieves such an allocation. The general idea for the simplified version was made explicit in an interesting and detailed follow-up paper, which solved the sister problem for the case of chores or burnt cake (agents have nonpositive valuations).⁵ Here we present a simplified version of the Core Protocol (Algorithm 2) for cake cutting. The protocol requires at most $(n!)^2n$ queries.

Algorithm 2. (Simplified) core protocol.

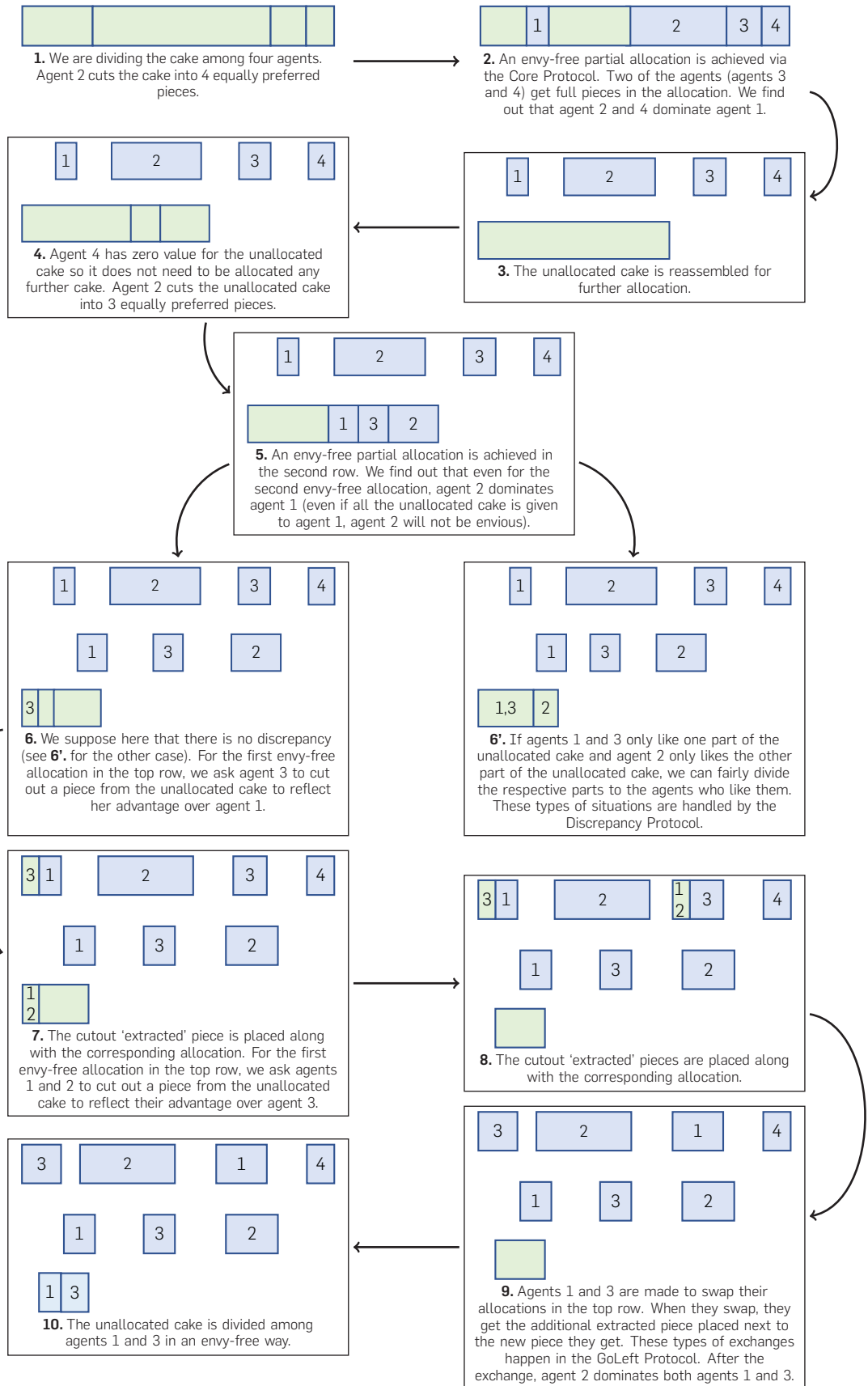
Input Agent set N , a cutter $i \in N$ and cake R .
Output An envy-free allocation of cake $R' \subseteq R$ for agents in N and an updated unallocated cake $R \setminus R'$.

- 1: Ask agent i to cut the cake R into n equally preferred pieces (p_1, \dots, p_n) .
- 2: **for** each permutation (a'_2, \dots, a'_n) of $N \setminus \{i\}$ **do**
- 3: **for** each permutation (p'_1, \dots, p'_n) of the n pieces (p_1, \dots, p_n) **do**
- 4: Give p'_1 to i
- 5: **for** $j = 2$ to n **do**
- 6: Give p'_j to a'_j . Ask a'_j to trim any of the pieces p'_{j+1}, \dots, p'_n if needed so the value of the pieces does not exceed the value a'_j has for her allocation p'_j .
- 7: **if** the allocation p corresponding to the permutation of agents and pieces is envy-free **then**
- 8: **return** the allocation p (which is called a Core allocation)
- 9: **else**
- 10: Reattach the trimmed parts to regain the original pieces.

In the Core Protocol, the cutter agent gets a full piece. Another agent also gets a full piece. So from the cutter’s perspective, at least $2/n$ of the cake is allocated by one call of the Core Protocol. Equivalently, the cutter thinks that her value of the remaining cake is at most $(n - 2)/n$ of her value of the full cake.

If we call the Core Protocol with a different cutter each time to further allocate the unallocated cake, we just need n calls of the Core Protocol to obtain an envy-free partial allocation which also satisfies proportionality (gives each agent value at least $1/n$ of the whole cake). Algorithm 3 does exactly that and in $n!^2n^2$ queries finds a partial allocation that

Figure 3. Illustration of some of the ideas of the protocol. The terminal states are 6' and 10.



satisfies envy-freeness and proportionality—two of the most important fairness concepts.^a The remainder of the paper describes what to do when we do want to allocate the whole cake.

Algorithm 3. An envy-free and proportional protocol.

Input Agent set N and cake R .

Output An envy-free and proportional allocation of the cake that may not allocate the whole cake.

- 1: **for** $i = 1$ to n **do**
- 2: **if** there is unallocated cake, **then** run the Core Protocol on the unallocated cake with i as the specified cutter.
- 3: **return** the allocation.

3.2. Domination and significant advantage

As the Core Protocol by itself is not powerful enough to allocate all the cake in bounded time, we rely on the idea of *domination* with the goal to decompose our problem into one involving a fewer number of agents. In this section, we denote an agent i 's allocation by X_i .

Recall that in an envy-free allocation, each agent i thinks she has an advantage (even if it is zero advantage) over each other agent j :

$$V_i(X_i) - V_i(X_j) \geq 0.$$

Domination is an extreme form of advantage. An agent i *dominates* another agent j if she is not envious of j even if the unallocated cake R is given to j :

$$V_i(X_i) - V_i(X_j) \geq V_i(R).$$

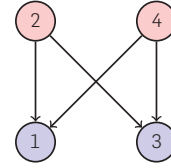
The other protocols are used with the following objective in mind: find a set of agents $A \subset N$ such that each agent in $N \setminus A$ dominates each agent in A . In order to ensure that each agent in some set $N \setminus A$ dominates each agent in A , it requires changing the current allocations of the agents as well as the unallocated cake. While doing these changes, we ensure that the current partial allocation remains envy-free. By identifying such a set $N \setminus A$, we reduce the problem to envy-free allocation for a smaller number of agents. The agents in $N \setminus A$ are not envious whatever the unallocated cake is allocated among agents in A . This crucial idea is illustrated in Figure 4.

Dominance of an agent i over another agent j has a close relation with agent i considering herself as having a ‘significant advantage’ over j . In order to define significance, we consider a suitable large constant bounded by some function over n . For a partially allocated cake, and piece a , an agent i finds value $V_i(a)$ significant if the value is at least $V_i(R) \left(\frac{n-2}{n}\right)^B$ where R is the unallocated cake

Significance of a piece is with respect to the residue, so if the residue becomes smaller, a significant value remains significant. The rationale for defining a significant value is that if an agent i thinks she has a significantly higher value for her allocation than she has for agent j 's allocation, then this significant advantage can be changed into domination

^a Note that finding an envy-free allocation that may be partial is a trivial problem: allocate nothing!

Figure 4. In the figure, an agent points to another agent if the former dominates the latter. Suppose we find an envy-free partial allocation among four agents such that each agent in $\{2, 4\}$ dominates each agent in $\{1, 3\}$. Then we can simply allocate the remaining cake among agents in $\{1, 3\}$ in an envy-free way.



by calling the Core Protocol a bounded number of times with i as the cutter. We explain this idea below.

Suppose we partially allocate the cake and agent i gets allocation X_i whereas agent j gets allocation X_j . Suppose that agent i thinks she has a significant advantage over agent j :

$$V_i(X_i) - V_i(X_j) \geq V_i(R) \left(\frac{n-2}{n}\right)^B.$$

Consider the situation where we run the Core Protocol over the unallocated cake R with agent i as the specified cutter and we do it B times so that the eventual unallocated cake is R^* . Then

$$V_i(R^*) \leq \left(\frac{n-2}{n}\right)^B V_i(R).$$

Thus, after B calls of the Core Protocol, agent i who previously had a significant advantage over agent j now dominates her:

$$V_i(X_i) - V_i(X_j) \geq V_i(R) \left(\frac{n-2}{n}\right)^B \geq V_i(R^*).$$

When we get a Core Protocol outcome, the cutter already has a significant advantage over the agent who got the least cake in the cutter's estimation. This significant advantage can easily be converted into domination by calling the Core Protocol. The main challenge is to obtain domination relations between more pairs of agents. Throughout the main protocol, the tentative partial allocation remains envy-free. Secondly, if an agent dominates another agent, the domination is maintained despite updates to the allocation.

3.3. Extraction

After we have called the Core Protocol on the updated unallocated cake, a sufficient but bounded number of times, we are in a position to *extract* from the residue. In each of the calls of the Core Protocol, there was a corresponding envy-free allocation. By envy-freeness, in each such allocation, each agent j has a nonnegative advantage over another agent i . For each of the Core allocations and for each $i, j \in N$, agent i is asked to extract a piece from the unallocated cake of value of the advantage over j in that Core allocation.

The extracted piece e will be in consideration to be attached to i 's corresponding allocated piece so that j is indifferent between her allocation and i 's allocation. If j 's intended extraction has a significant value, we do not extract because we only want to extract pieces from the remainder which are not significant for all the agents. If the intended extraction is not significant, we put it on a side for

consideration for attachment. If it cannot be made unambiguously insignificant, then we say that the piece is discrepant and we call the Discrepancy protocol which either exploits or ‘eliminates’ this discrepancy.

Figure 5 shows how agents extract pieces from the unallocated cake R . In the figure, we consider extractions by agents 2, 3, and 4 based on their advantage over agent 1. Agent 2 thinks that her advantage over agent 1 is of the same value as her value for the leftmost extracted piece. Agent 4 thinks that her advantage over agent 1 is of the same value as the sum of her values for two leftmost extracted pieces.

The extracted pieces will be attempted to be attached to agent 1’s piece as indicated in Figure 6.

Suppose we have a set of Core Protocol allocations and the corresponding extracted pieces placed in the appropriate order. We call a set of Core Protocol allocations *isomorphic* to each other if for each piece c_i in agent i ’s allocation, the agents who extracted cake from the residue and associated to c_i are the same and did so in the same order. Later, we will identify a subset of Core Protocol allocations that are isomorphic to each other. Isomorphic allocations will be considered later by the GoLeft Protocol.

3.4. Discrepancy protocol

When pieces are being extracted from the residue, it may be the case that one of the pieces e in consideration for extraction is significant for some agent. In that case, the piece is not extracted and the Discrepancy Protocol is called that either eliminates or exploits this discrepancy. The discrepant piece e is kept aside from the residue. If the piece is “almost significant,” we can make it significant by reducing the residue by calling the Core Protocol a bounded number of times.

By doing this, either the discrepant piece becomes unambiguously significant or we still have the case that some agents consider e significant and others do not. The first case is helpful because there is no discrepancy in terms of significance and our protocol makes use of this consistency. In the second case, if there exists some $i \in N$ such that $V_i(R)/n < V_i(e) < V_i(R)n$, we continue running the Core Protocol with agent i as the cutter. By doing so, we achieve in a bounded number

Figure 5. Agents extracting pieces from the remaining cake up to their advantage over agent 1.

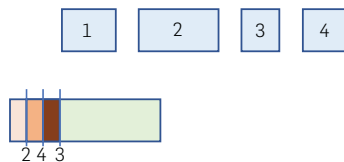
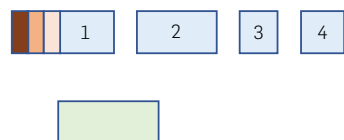


Figure 6. Extracted pieces placed next to agent 1’s allocation for the purpose of attachment.



of calls of the Core Protocol the situation where for each agent i , either $V_i(e) \geq V_i(R)n$ or $V_i(e) \leq V_i(R)/n$. This situation is explained in Figure 7.

Algorithm 4. Main protocol—high-level sketch.

Input: Cake R and a set of agents N .

Output: An envy-free allocation.

- 1: **Core Allocations:** generate core allocations by repeatedly dividing the unallocated cake via the Core Protocol a bounded number of times.
- 2: **Extraction:** extract pieces from the residue corresponding to the core allocation pieces as long as the pieces are not significant for any agent as explained in Section 3.3. While extracting pieces, if some piece a is significant for at least one agent, call the Discrepancy Protocol as explained in Section 3.4. It ensures that now either all agents consider the piece significant (in which case it is not attached) or we decompose the main problem into two subproblems for the Main Protocol where some agents are to be given a and the others are to be given the remaining unallocated cake.
- 3: **GoLeft:** Call the GoLeft Protocol to attach extracted pieces to the corresponding Core allocations. The GoLeft protocol returns a subset of agents $A \subset N$ such that each agent in $N \setminus A$ dominates each agent in A . The central idea of GoLeft is to facilitate exchanges of suballocations of agents.
- 4: Call the Main Protocol to allocate the remaining cake to agents in A .
- 5: **return** allocation of the cake to the agents.

If $V_i(e) \geq V_i(R)n$, then i is predominantly interested in e rather than the residue. If $V_i(e) \leq V_i(R)/n$, then i is predominantly interested in R . Because the piece that agents are predominantly interested in has n times more value than the other piece, any agent who gets an envy-free (and hence proportional) allocation of the preferred piece also gets at last $1/n$ value of the preferred piece. The value is at least as much as the value of the piece that is less preferred.

3.5. Main protocol

Continuing to call the Core Protocol on the updated remaining cake gives no guarantee that the cake will be allocated fully even in finite time. Hence, we need to use other protocols which are called by the Main Protocol. We gave an intuitive idea of the Main Protocol in Figure 2. We give a more detailed high-level sketch of the protocol in the form of Algorithm 4.

The first two stages of the Main Protocol are making calls to the Core Protocol to further allocate the cake and then to

Figure 7. Discrepancy. Agents in A think that the left part has n times more value than the right part. Agents in $N \setminus A$ think that the right part has n times more value than the left part. In that case, if we allocate the left part to A in an envy-free way and the right part to $N \setminus A$ in an envy-free way, we obtain an overall envy-free allocation for N .



implement the extraction as explained in the previous sections. While pieces are being extracted, we may have to call the Discrepancy Protocol. Throughout the steps of the Main Protocol, we maintain an envy-free allocation as well as keep track of the updated unallocated cake. After that, the Main Protocol calls the GoLeft Protocol. In the subsequent section, we give further details of the GoLeft Protocol.

3.6. GoLeft protocol

In this section, we give an overview of the GoLeft Protocol (Algorithm 5). When the GoLeft Protocol is called, we already have a bounded number of envy-free allocations due to the calls to the Core Protocol. We also have extracted pieces from the residue that will be considered for attachments to the corresponding Core allocations of the agents.

The purpose of extracting pieces from the residue is that we can attach them to the corresponding Core allocation piece of i so that j is indifferent between her allocated piece and i 's piece. This makes it easier for j to switch one of her pieces if she gets the additional insignificant extraction. Making agents exchange their allocations while additionally giving them additional extracted pieces is useful to diversify the relations of agents having a significant advantage over others.

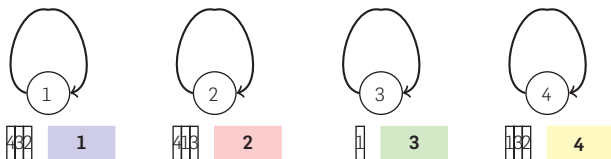
We elaborate on why attachment is helpful to obtain additional significant advantages. Let us say that in a number of Core allocations, agent k has a significant advantage over agent i 's allocation and agent j has an insignificant advantage over i 's allocation. In order for k to have a significant advantage over j rather than i , we want to make some local envy-free preserving operations so that j gets i 's allocated piece along with j 's insignificant extraction corresponding to j 's advantage over that piece of i 's.

Permutation graph. When the GoLeft Protocol starts, it first identifies a working set S of C Core allocations from out of the C' Core allocations that we focus on. As C' is chosen to be large enough, we can find C Core allocations that are isomorphic. The protocol then constructs a *permutation graph* corresponding to the working set of isomorphic allocations.

In the permutation graph, each node i corresponds to an agent i who holds a set of isomorphic pieces along with her attached extracted pieces in the working set of isomorphic allocations S . Agent i points to agent j if j holds isomorphic pieces in S that have had all attachments up till i 's extracted pieces. Each node has an indegree of one. Initially, the permutation graph consists of all nodes having self-loops (see Figure 8).

We divide the nodes of the permutation graph into sets T and T' . Set T is the set of nodes/agents such that the isomorphic

Figure 8. Initial state of the permutation graph along with the corresponding state of an allocation representative of the working set of isomorphic allocations.



pieces held by them in S have not had $n - 1$ attachments). T' is the set of nodes/agents such that the isomorphic pieces held by them in S have had $n - 1$ attachments.

The protocol identifies a cycle in the permutation graph that includes at least one node i from T . Such a cycle always exists. In each of the working set S of isomorphic allocations, we implement an exchange of pieces held by agents in the cycle: each agent in the cycle is given the piece corresponding to the node that the agent points to in the cycle. After implementing the exchange, the permutation graph is updated to reflect the exchange. In the exchange, if an agent gets an inferior piece, she always gets the additional extracted pieces associated with it up till the agent's extractions. Hence, each agent's value from her allocation is preserved in each allocation in S even if she gets a different piece than in the original Core allocation. For any agent i , as long as no agent gets extracted pieces beyond i 's extraction, i will not be envious. In the GoLeft protocol, it can be the case that some agent j gets extracted pieces beyond i 's extracted pieces but before any such attachments in the last part of the GoLeft protocol, we ensure that no envy arises.

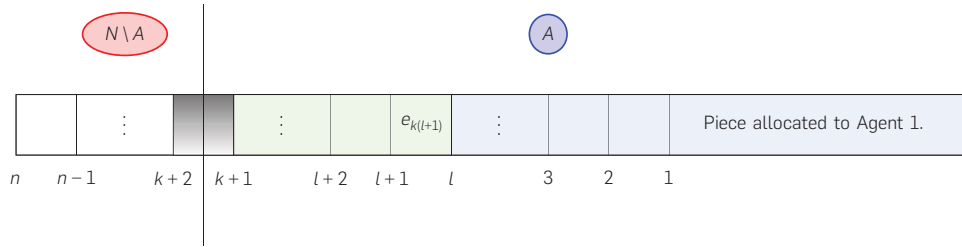
After implementing the cycle, we focus on a node $i \in T$ that was in the cycle. For agent/node i , we know that for all allocations in the working set S , agent i has been allocated the original isomorphic pieces c_k as well as all associated pieces up till i 's extracted piece. If the piece of cake agent i is currently allocated in the allocations S has no more extracted pieces left to attach to it, but it has not had $n - 1$ attachments, this means that all agents who have not had their corresponding piece extracted/attached have a significant advantage over agents who have had an extracted piece attached. In this case, the GoLeft Protocol returns the set of dominated agents to the Main Protocol and we are left with a smaller envy-free allocation problem because it involves a fewer number of agents.

In case node i does not lead to an exit from the GoLeft Protocol, we know that there are associated pieces that can still be attached to the isomorphic pieces held by i in the working set of Core allocations S . We focus on the next set of associated pieces $e_{k(l+1)}$ that we are interested to attach to the pieces c_k that have already had associated pieces $e_{k2}, e_{k3}, \dots, e_{kl}$ attached in their corresponding main pieces c_k (see Figure 9). Additionally attaching pieces $e_{k(l+1)}$ to pieces c_k is useful in making the agent who extracted them interested in the pieces c_k because of the additional $e_{k(l+1)}$ as well as the previous attachments.

Avoiding envy when attaching extracted pieces. Naively attaching the pieces can be problematic and spoil the envy-freeness of the allocation that we maintain. We deal with the issue as follows.

- The agents who did not extract pieces associated with the c_k pieces as well as agents who extracted pieces that have not been attached are asked to 'reserve' a big enough subset $S' \subset S$ of allocations in which they value the difference between their bonus value for c_k and the extracted pieces currently attached to c_k the most. These allocations S' are removed from S and their remaining unattached associated pieces are sent back to the residue. By maintaining the advantages in the Core alloca-

Figure 9. Illustration of the GoLeft protocol on a particular piece of cake that is originally allocated to agent 1. Agents $k + 2$ to n will not go left and are the prospective dominators because they find the shaded space between the trims of $k + 2$ and $k + 1$ significant. Agents 2 to $k + 1$ are the agents that go left.



tions S' , such agents will not be envious even if some agent in $\{1, \dots, l\}$ additionally gets all other extracted pieces $e_{k(l+1)}$ in the remaining Core allocations in S .

Algorithm 5. GoLeft protocol—high-level sketch.

Input: Set of C' allocations, extracted pieces corresponding to the C' Core allocations, and residue R .

Output: A set of agents $A \subset N$ such that agents in $N \setminus A$ dominate agents in A .

- 1: Select C isomorphic allocations (set S); Build the permutation graph.
 - T , the set of nodes with pieces for which $n - 1$ extracted pieces have not been attached.
 - T' , the set of nodes with pieces for which $n - 1$ extracted pieces have been attached.
- 2: **while** there is a node in T **do**
- 3: Find a cycle that includes a node that is from T (such a cycle always exists).
- 4: In the cycle identified, let each agent in the cycle get the allocation she points to up till her extractions.
- 5: **if** there is a piece p corresponding that is not from T' but has no more associated pieces to be attached **then**
- 6: Consider the set of agents A who either owned the original piece p or whose extracted pieces have already been attached to p . Return the dominated set of agents A .
- 7: **Attachment:** consider the set of isomorphic Core allocation pieces $\{c_k\}$ that have already had associated pieces $\{e_{k2}\}, \{e_{k3}\}, \dots, \{e_{kl}\}$ attached to them but some extracted pieces have not been attached. Attach in a subset of the allocations in C' the set of extracted pieces $\{e_{k(l+1)}\}$ to the set of pieces $\{c_k\}$, thus making $\{c_k\}$ desirable to the agent who extracted $\{e_{k(l+1)}\}$. In order to attach the pieces without creating envy, a subset $S', S'' \subset S$ of Core allocations is removed from the working S of Core allocations. The Core allocations in $S' \cup S''$ do not undergo attachments or further changes. Update the permutation graph to reflect the attachment. If the piece has had all $n - 1$ extracted pieces attached, add the corresponding node to T' and make every node point to it.

- The agents indexed from 1 to l who have all already had their extracted pieces attached to c_k are asked to choose a high enough fraction of the Core allocations in S in which

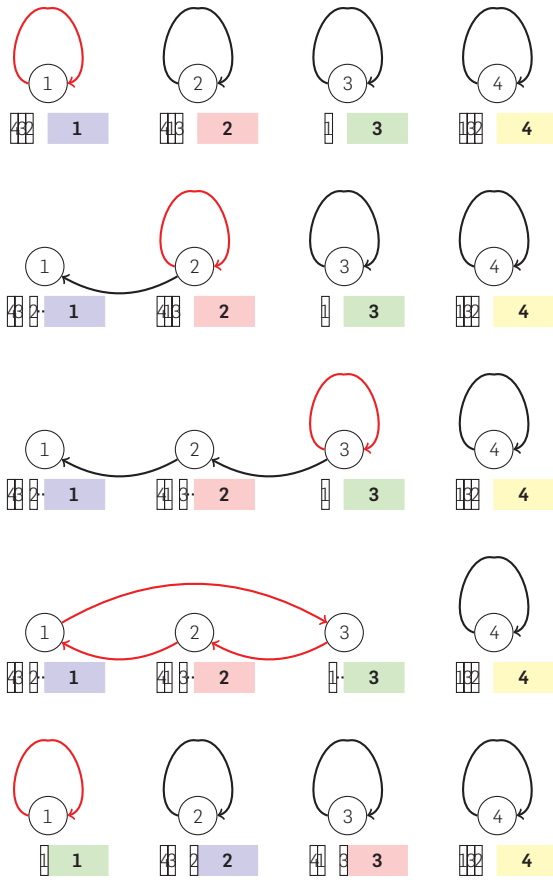
they value the $e_{k(l+1)}$ pieces. We call these allocations S'' . The $e_{k(l+1)}$ pieces from S'' are bunched together and the Main Protocol is called to divide this cake in an envy-free way among the agents indexed from 1 to l where l is strictly less than n . As envy-freeness implies proportionality, they derive enough value that they will not be envious if the agent indexed $l + 1$ gets all other pieces in set $e_{k(l+1)}$. The corresponding set of allocations S'' is then removed from consideration for updates.

Hence, each time we attach isomorphic extracted pieces $e_{k(l+1)}$ to isomorphic pieces c_k , we ‘freeze’ allocations $S' \cup S''$ from the working set S and still maintain an envy-free allocation. Note that in the allocations that remain in S , agents may currently hold a different isomorphic piece than they previously did, but as they also hold the corresponding attachments associated with the isomorphic piece, each agent’s total value in each isomorphic allocation in S stays the same. In Figure 10, we show the states of the permutation graph and the corresponding representative Core allocation as well as the corresponding extracted pieces.

When the protocol attaches extracted pieces $e_{k(l+1)}$ to allocated pieces c_k currently held by agent l , it deletes the incoming edge of node/agent l and replaces it by an edge coming from agent $l + 1$ who extracted pieces in $e_{k(l+1)}$. Intuitively, $l + 1$ is now willing to be allocated to c and its attached pieces instead of her current pieces in S . We delete previous edges to ensure that until termination, nodes in T have an in-degree of exactly 1, which guarantees that no matter the cycle involving a node in T found by the protocol, we will make progress towards termination. The following example shows how progress is made in attaching extracted pieces to the working set of isomorphic Core allocations.

EXAMPLE 1. In Figure 10, we demonstrate how the permutation graph along with the working set of isomorphic allocations changes in the GoLeft Protocol. Note that even when the representative allocation changes, there still exist allocations isomorphic to the previous representative allocations but these allocations have been removed from consideration from the working set of allocations. The colored/shaded pieces represent the pieces given by the Core Protocol to each agent. The small pieces on the left of the colored pieces are extracted pieces, each labeled by the agent who extracted it. At first, the extracted pieces are associated with a specific allocated piece. Then they are attached to it

Figure 10. Permutation graph along with the corresponding state of an allocation representative of the working set of isomorphic allocations.



(represented by the dotted lines). Finally, when a colored/shaded piece is real-located to a new agent, the extracted pieces attached to it are also allocated to the new agent (in the diagram we now aggregate the extracted piece to the main piece). In the second state of the isomorphic allocation, agent 2 points to agent 1 because the piece extracted by agent 2 has been attached to 1's held piece. In the third state of the isomorphic allocation, agent 3 points to agent 2 because the piece extracted by agent 3 has been attached to 2's held piece. In the fourth state of the isomorphic allocation, agent 1 points to agent 3 because the piece extracted by agent 1 has been attached to 3's held piece. In the fifth state, the agents 1, 2, and 3 exchange their currently held piece and are allocated cake up to their extracted piece. In the fifth (last) state of the isomorphic allocation, agent 1 holds a piece up till her extraction but neither agent 2 or 4 extracted pieces for the piece that agent 1 holds. This means that agents 2 and 4 have a significant advantage over agent 1. Initially, the piece was held by 3 and still is in discarded isomorphic allocations. This implies a significant advantage of 2 and 4 over 3. Therefore, agent 2 and 4 can be made to dominate 1 and 3.

By attaching enough extracted pieces in the appropriate order, the GoLeft Protocol finally arrives at a point where there is some isomorphic set of pieces c_k in the set S for which all possible associated pieces have been attached but there is some set of agents $N \setminus A$ who do not have associated pieces. The reason

agents in $N \setminus A$ could not extract such pieces is because they had a unanimous significant advantage over the agent indexed 1 who got the pieces c_k . By gradually attaching (unanimously insignificant) associated piece to pieces c_k and ensuring that all agents who did extract the corresponding pieces do get some isomorphic piece in c_k (along with the associated insignificant attachments), we make sure that agents in $N \setminus A$ now dominate agents in A . At this point, we can return from the GoLeft Protocol. We have successfully reduced our envy-free allocation problem to that involving less number of agents. By recursively calling the Main Protocol to allocate the remaining cake to agents in the smaller set A , we eventually allocate the whole cake.

4. CONCLUSION

We presented a high-level overview of our bounded envy-free protocol. The protocol has an upper bound that is a power tower of six n 's. In the other direction, any envy-free protocol requires at least $\Omega(n^2)$ queries.⁶

We additionally show that even if we do not run our protocol to completion, it can find in at most n calls of the Core Protocol a partial allocation of the cake that achieves proportionality (each agent gets at least $1/n$ of the value of the whole cake) and envy-freeness. If we allow for partial allocations, an interesting open problem is the following one: can envy-freeness and proportionality be achieved in a polynomial number of steps?

Acknowledgments

Haris Aziz is supported by a UNSW Scientia Fellowship. He thanks Xin Huang, Sven Koenig, Omer Lev, Bo Li, and Simon Rey for helpful feedback. He also thanks Simon Rey for his help in making some of the figures. □

References

1. Aziz, H., Mackenzie, S. A discrete and bounded envy-free cake cutting protocol for four agents. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)* (ACM Press, 2016), 454–464
2. Aziz, H., Mackenzie, S. A discrete and bounded envy-free cake cutting protocol for any number of agents. In *Proceedings of the 57th Symposium on Foundations of Computer Science (FOCS)* (2016), 416–427
3. Brams, S.J., Taylor, A.D. An envy-free cake division protocol. *Am. Math. Month J.* 102 (1995), 9–18.
4. Brams, S.J., Taylor, A.D. *Fair Division: From Cake-Cutting to Dispute Resolution*. Cambridge University Press, 1996.
5. Dehghani, S., Farhadi, A., Taghi Hajiaghayi, M., Yami, H. Envy-free
6. Procaccia, A.D. Thou shalt covet thy neighbor's cake. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)* (AAAI Press, 2009), 239–244.
7. Procaccia, A.D. Cake cutting: Not just child's play. *Commun. ACM* 7, 56 (2013), 78–87.
8. Robertson, J.M., Webb, W.A. *Cake Cutting Algorithms: Be Fair If You Can*. A. K. Peters, 1998.
9. Steinhaus, H. The problem of fair division. *Econometrica*, 16 (1948), 101–104.

Haris Aziz (haris.aziz@unsw.edu.au), UNSW Sydney, Australia; CSIRO Data61, Sydney, Australia.

Simon Mackenzie (simon.mackenzie@data61.csiro.au), CSIRO Data61, Sydney, Australia.

[CONTINUED FROM P. 128] **When you worked through all the problems, your prototype took eight hours to boot.**

But it still had all the debugging stuff in it, along with everything else we needed to figure out when something went wrong. At the time, my wife was running the company, and she said, “eight hours, that’s not going to work.”

And I told her, “this is incredibly good news. We know everything we have to do to run Windows; we just need to figure out how to make it run fast enough.”

VMware was founded in 1998. Have your views changed at all since then about the relationship between universities and entrepreneurship?

Researchers, especially in applied fields like systems, are always trying to make an impact. And one of the biggest ways you can make an impact is to take your ideas and change the way that industry does something—that’s sort of a best-case outcome. I’m very glad I was able to launch a company and move the industry to do things in a way that creates better outcomes for everyone.

The other aspect of it is probably not as noble—you can make a lot of money. You look around, and there’s incredible wealth being created. If you’re not part of that, you’re going to end up being left behind or maybe not feel as successful.

VMware obviously had enormous impact on the industry. Are there things you would have done differently, knowing how the cloud computing industry has evolved since then?

VMware was fabulously successful in terms of a business venture. It had a unique product, and it was able to charge money for it. But one of the reasons it didn’t become a standard virtualization platform is that major cloud computing vendors opted for an open-source solution rather than paying a high price for a proprietary piece of software.

I struggled with that at VMware. I really wanted to figure out how to make a product that was both successful and universally used. Would I do things differently now? It’s a challenge. When you do something new, it takes a lot of resources to figure out how to do it, and you want to get rewarded for that.

I know that people are launching more open-source companies now, but I just could not figure out how to do it.


How else has the field changed over the course of your career?

It used to be, in computer science, you would come up with an idea, and you basically looked at all the good things it could do and all the positive scenarios it supported. Now, we’re seeing some very scary unintended consequences. Technology has become such a prominent part of the world, and it’s done a lot of good, but it’s also enabled some not-so-good things. People expect computer scientists to anticipate those scenarios, and some of them imagine that all we technologists need to do is take more ethics courses. I’m a little skeptical of that.

You’re back at Stanford after a leave of absence at BeBop, a development platform acquired by Google in 2016. What are you working on?

Stanford has a rule that in any seven-year window, you can take two years off for a sabbatical. In 1998, I took a leave to found VMware, and then more recently for BeBop, but I like it here at Stanford, with new students and new ideas coming in all the time.

One of the things I’ve been looking at recently is a very old problem. If you have a bunch of computers—in a datacenter, for example—they all have clocks. We still use a pretty old protocol called network time protocol, or NTP, to synchronize those clocks. Essentially, it involves stage messages saying, “I think my time is this. What do you think your time is?” You end up with clocks that are synchronized pretty well from a human point of view, because everyone’s device looks like it’s showing about the same time.

But for a computer, in a program, it’s kind of worthless. I could easily send a message from one computer and find out that it arrived earlier than I sent it. So a colleague, Balaji Prabhakar, and I have been working on a new clock sync algorithm that can synchronize clocks down into the single-digit nanoseconds. 

Leah Hoffmann is a technology writer based in Piermont, NY, USA.

© 2020 ACM 0001-0782/20/4 \$15.00



2018 JOURNAL IMPACT
FACTOR: 6.131

ACM Computing Surveys (CSUR)

ACM Computing Surveys (CSUR) publishes comprehensive, readable tutorials and survey papers that give guided tours through the literature and explain topics to those who seek to learn the basics of areas outside their specialties. These carefully planned and presented introductions are also an excellent way for professionals to develop perspectives on, and identify trends in, complex technologies.



For further information
and to submit your
manuscript,
visit csur.acm.org

Q&A

Reinventing Virtual Machines

The notion of scalable operating systems led Mendel Rosenblum to virtual machines, which have revolutionized datacenters and enabled modern cloud computing.

STANFORD UNIVERSITY PROFESSOR Mendel Rosenblum, recipient of the inaugural ACM Charles P. “Chuck” Thacker Breakthrough in Computing Award, developed his groundbreaking virtual machines in the late 1990s as a way of enabling disparate software environments to share computing resources. Over the next two decades, these ideas would transform modern datacenters and power cloud computing services like Amazon Web Services, Microsoft Azure, and Google Cloud. Here, Rosenblum talks about scalability, systems design, and how the field has changed.

Virtual machines were pioneered by IBM in the 1960s. What prompted you to revisit the concept back in the 1990s?

When I got to Stanford, I joined up with John Hennessy, who was building a very large supercomputer with shared memory that scaled up to 4,000 processors. His group needed an operating system, because existing systems couldn’t run on a machine of that size. That prompted me to start thinking about scalable operating systems. At the same time, I was working on another project about building operating systems on modern hardware. And I began trying to build simulation environments that you could run operating systems on.

This is the so-called SimOS machine simulator.

It was a piece of software that would run and look enough like the



hardware machine that you could boot an operating system and all its applications on it. It was much, much slower than a real machine, but it let us model how the hardware was doing under realistic workloads.

Sequent Computer Systems, which was developing an operating system for multiprocessor machines, was interested in your work on scalable operating systems, but told you their team was too small to implement such a major change.

They also said they needed the machine to be able to run Microsoft Windows. So when I was on the plane back from Portland, it occurred to me that maybe if we just brought back the idea of virtual machines, and use that to sort of carve up these big machines we were building, we

“Technology has become such a prominent part of the world, and it’s done a lot of good, but it’s also enabled some not-so-good things.”

would be able to run existing operating systems on them.

Soon afterward, you and your students built a virtual machine monitor that could run multiple copies of commodity operating systems on a multiprocessor—starting with a MIPS processor and then, when you decided to commercialize your work, the Intel X86.

The Intel X86 was the dominant processor at the time, though we didn’t really understand how complex it was. It was technically known as not virtualizable because it didn’t have adequate hardware support. So we had to figure out some techniques for doing that. Linux was pretty simple, because it didn’t use very much of the X86. But Windows was taking forever, and we kept discovering new things about the X86 architecture to figure out how to do it. [CONTINUED ON P. 127]

Today's Research Driving Tomorrow's Technology

The ACM Digital Library (DL) is the most comprehensive research platform available for computing and information technology and includes the ongoing contributions of the field's most renowned researchers and practitioners.

Each year, roughly 20,000 newly published articles from ACM journals, magazines, technical newsletters and annual conference volumes are added to the DL's complete full text contents of more than 550,000 articles.

The DL also features the fully integrated and comprehensive bibliographic index, *The Guide to Computing Literature*—a continually updated index featuring millions of publication records from over 5,000 publishers worldwide.

For more information, please visit

<https://libraries.acm.org/>

or contact ACM at

dl-info@hq.acm.org

ACM

DL

DIGITAL
LIBRARY



OCTOBER 25 - 29, 2020
TALLINN, ESTONIA

nordichi2020.org

SHAPING EXPERIENCES, SHAPING SOCIETY

The biannual conference, NordiCHI, is the main Nordic forum for Human-Computer Interaction (HCI) research. NordiCHI addresses the field broadly and is a meeting place for researchers from academia and industry, designers, practitioners, educators and others working within HCI and related disciplines.

The theme of NordiCHI 2020 is “shaping experiences, shaping society” and our goal is to raise awareness about the effect we have on society through the experiences we enable, facilitate or craft.

Key Dates

April 15th '20

Submission deadline for research paper abstracts

April 20th '20

Submission deadline for research papers

June 23rd '20

Notifications sent

