CERIAS Tech Report 2024-01

Grand Challenges in Trustworthy Computing at 20: A Retrospective Look at the Second CRA Grand Challenges Conference

by Richard DeMillo and Eugene H. Spafford
Center for Education and Research
Information Assurance and Security
Purdue University, West Lafayette, IN 47907-2086

Grand Challenges in Trustworthy Computing at 20:

A Retrospective Look at the Second CRA Grand Challenges Conference

Richard DeMillo (Georgia Tech)
Eugene H. Spafford (Purdue University CERIAS)

The 24th Annual Security Symposium of Purdue University's Center for Education and Research in Information Assurance and Security (CERIAS)[1] was held on Purdue University's West Lafayette campus on March 28-29, 2023. The meeting coincided with the 20th anniversary of the 2003 Gordon-Style Conference[2] that we co-chaired, sponsored by the Computing Research Association (CRA), to define Grand Challenges in Trustworthy Computing. By 2003, information and communication security (the term "cybersecurity" was not yet in widespread use) had already emerged as a critical enabler for the global digital revolution that was underway. The resulting Grand Challenges[3] were announced when the field was experiencing a historical inflection point and became a landmark for a generation of researchers and funding agencies.

The importance of cybersecurity is now well-established in public consciousness. It occupies policy discussions ranging from the integrity of elections and IT infrastructure protection to rules and regulations governing privacy and global information sharing. As many of the participants in the 2003 conference are still active in the field, we concluded that the CERIAS symposium would be an excellent opportunity to reflect on the Grand Challenge process for a new generation of cybersecurity professionals who will confront many of the same threats faced by their predecessors. The 2003 participants were invited to meet online over six weeks to review and comment on the successes and failures of the Grand Challenge recommendations. We summarized those discussions in a public panel session at the CERIAS Symposium. This article distills lessons learned from the 2003 workshop. It may ignite interest in new problems and agendas directed at securing computing infrastructure in a rapidly growing and evolving threat landscape.

[1] https://ceri.as
[2] In an effort to promote out-of-the-box (sometimes risky) thinking and open discussions, comments and communications at Gordon-style conferences are not attributed to individuals
[3] https://archive.cra.org/reports/trustworthy.computing.pdf

From the Perspective of 2003

When CRA asked us to organize the second in what would become a series of conferences directed to challenging the research community to address important problems in computing, the Internet's structure was, in many ways, still vague and indistinct. Less than 10% of the world's population was connected to the Internet in 2003 (today, the global Internet user population is estimated to reach nearly 70% of the human race).[4] It would be four years before the iPhone would be introduced to the world, Facebook did not yet exist (it was established in 2004), and it was still early in Google's evolution from one of many search engine companies to a dominant provider of Internet services (Gmail was not launched until 2004). Sun Microsystems and HP existed as two of the dominant computer companies in the world. Although the global market for online advertising had stalled at slightly more than $7 billion, it would grow 30-fold to nearly $210 billion over the next twenty years. Military and Intelligence applications had embraced computerization in the preceding decade. However, most critical infrastructure still relied on human operators and mechanical controls.

By 2003, Peter Neumann's "Risks to the Public" forum had been a regular feature of ACM SIGSOFT—and, eventually, Communications of the ACM (CACM)—for nearly twenty years. Inspired by ACM President Adele Goldberg's 1984 letter to ACM Membership citing how "Increasingly, human lives depend upon the reliable operation of systems,"[5] the Risks Forum was remarkable for the scant citations of what we would today call cybersecurity incidents. Nevertheless, epidemic-style attacks on network-connected devices were on the rise and accelerating at an alarming rate. When the conferees met, damages caused by cyber-attacks were on a path to tripling every year[6]. In January 2003, it took the SQL Slammer worm only 10 minutes to propagate, disabling half of the DNS root servers in the world and forcing critical services such as banking, 911 calls, and air traffic control offline.[7] Within six months, two more destructive malware attacks (the SoBig virus and Slammer worm) degraded network services worldwide. Preventing such attacks from crippling information and communications technology seemed beyond the reach of computer scientists and engineers. Additionally, there was widespread apprehension about the growing gap in numbers and training of professionals to confront the threats posed by nation-states, organized crime, and a generation of anarchists and criminals who all had access to the same technology that was used to defend vulnerable systems.

[4] https://Internetworldstats.com

[5] Communications of the ACM, February 1985 (pp. 131-133)

[6] That growth rate was not sustainable, but damages from cyber attacks did grow over the next twenty years by a factor of nearly 600 https://www.statista.com/statistics/267132/total-damage-caused-by-by-cyber-crime-in-the-us/

[7] https://www.zakon.org/robert/internet/timeline/

The national information and communications security strategy was heavily influenced by the events of September 11, 2001,[8] but was short on definite recommendations for understanding and countering known threats. The board of CRA[9], which represented the principal academic, governmental, and industrial computing research organizations, overwhelmingly favored a national meeting designed to produce a research agenda tuned to these trends.

A diverse group of experts was chosen based on vision statements crafted in response to a published request for participation[10]. The group assembled at the Airlie House in Northern Virginia in November 2003 to draft a list of challenges to drive research and development for the next generation. We structured the meeting to consider "out of the box" approaches to make infrastructure immune from attacks by various threat actors and thus more trustworthy to all users. Trial balloons, candidate challenges, and the venue's secluded atmosphere presented opportunities to argue priorities and tradeoffs. The steering committee sought a small set of high-level goals instead of an arbitrary "top 10" list. The result was a document summarizing four grand challenges.

Workshop results were announced in a briefing at the National Press Club and published as a report[11]. Although we cannot say precisely how much influence the meeting had, the four items appear to have affected the national information security and privacy research agenda[12]. The workshop influenced the report "Cyber Security: A Crisis of Prioritization"[13] from the PITAC (President's Information Technology Advisory Committee) to President Bush in 2005, the "Hard Problems List"[14] of the Infosec Research Council in 2005, and the report "Toward a Safer and More Secure Cyberspace"[15] by the National Academies in 2007; attendees of the Grand Challenges workshop were involved in the production of all three of these documents.

[8] https://www.us-cert.gov/sites/default/files/publications/cyberspace_strategy.pdf

[9] Both authors were CRA Board of Directors members in 2002-3.

[10] Out of 220 applicants who submitted statements to the steering committee, 50 attendees were selected to attend. The committee explicitly sought a wide range of expertise, constituencies, and level of seniority. The full list of participants is in the final workshop report.

[11] https://archive.cra.org/reports/trustworthy.computing.pdf

[12] https://www.nsf.gov/bfa/dias/policy/docs/cise_cleve.pdf

[13] https://www.nitrd.gov/historical/Pitac/Reports/20050301_cybersecurity/cybersecurity.pdf

[14] https://www.nitrd.gov/documents/cybersecurity/documents/IRC_Hard_Problem_List.pdf

[15] https://nap.nationalacademies.org/catalog/11925/toward-a-safer-and-more-secure-cyberspace

The Grand Challenges

Assumptions

In 2003, information technology was becoming increasingly pervasive in our daily lives. Computing infrastructure was becoming more complex and interconnected on a worldwide scale. The historical intertwining of Moore's Law and increasingly capable software continued to push the boundaries of computing power and complexity,[16] making it more difficult for developers to assure users that systems were trustworthy. Furthermore, Internet-driven digital transformation sweeping across various industries made it clear that computing security and trustworthiness were becoming critical to global infrastructure.

Conferees had to make assumptions about the landscape supporting such an overarching vision. First, the Internet and an array of public and private specialized networks were already playing critical roles in computing, and it was only possible to imagine progress by assuming a reliable network infrastructure. Furthermore, networks and underlying system infrastructure would have to be sound. We assumed a reliable end-to-end infrastructure to avoid building on shifting sands. Second, the scale of the assurance problems demanded a new generation of effective methods and tools for designing and building trustworthy systems. Finally, it seemed unreasonable that a single "silver bullet" solution would appear. Each system and application domain would require understanding the human and societal factors contributing to trustworthy systems.

These were ambitious assumptions, but the trajectory of technology was moving in the right direction. Increasingly powerful computers were already becoming smaller, cheaper, and more easily embedded in other systems. Computers were also more mobile. Networking and mobile computing were becoming pervasive, reaching a global community and inviting more of the human race to participate in the digital revolution. E-commerce, e-government, on-demand services, telecommuting, telemedicine, and entertainment were seen as among the industries that would ultimately be affected. With all this development, it was assumed that vast amounts of engineering data would be at system designers' fingertips, relieving enterprises of trial-and-error approaches to technology and policy[17].

[16] https://www.defensemedianetwork.com/stories/the-intertwined-history-of-darpa-and-moores-law/

[17] Conferees did not have uniform views about the possible negative side effects of amassing large amounts of data, e.g., that could compromise personal privacy or aid authoritarian governments.

In the years following the conference, the cybersecurity community made significant progress in strengthening underlying infrastructure, validating many of the group's assumptions. For example, the widespread adoption of encryption protocols such as the Secure Sockets Layer (SSL) and, later, Transport Layer Security (TLS) improved the overall trustworthiness of the network infrastructure. Secure networking protocols such as IPv6 and the Domain Name System Security Extensions (DNSSEC) added to network robustness, although the pace of widespread adoption seemed glacial to many of us. Similarly, the development of security tools and virtualization technologies made it possible to isolate different components of a computing system and reduce attack surfaces, enabling a more reliable end-to-end infrastructure. Advances in foundational areas, including cryptology, led to the widespread adoption of strong encryption and authentication mechanisms, making it more difficult for attackers to compromise systems and steal sensitive data.

The Challenges

The Grand Challenges were of a different order than the results of previous, similar efforts. To begin with, they were not incremental. Problems for which there was a clear pathway from existing knowledge to a solution never made it to the Grand list. Furthermore, a Grand Challenge should define its success criteria. The challenge statement should include a time frame and a deliverable solution to a well-posed technical problem. The conferees agreed that it would be impossible otherwise to say when a general statement of intent–no matter how ambitious–had been satisfied. Furthermore, Grand Challenges should be relevant to the direction of the field and accompanied by an explanation of why existing approaches were insufficient and required different methods. Finally, Grand Challenges should be "Grand" in the literal sense: They should excite the imagination. They should require a level of innovation or creative invention that commands the attention of the most capable and fearless scholars. They should be worthy of investment by the research community because of their potential for broad impact.

While whittling down scores of candidate ideas, conferees embraced an overarching vision of the general direction of computing technology. From a traditional engineering standpoint, computers became more reliable and supported varied policies and personal choices. New ways of approaching security would have to anticipate a future in which computers would be more intuitive and predictable. Such a future would require assuring end-user control over the flow of information code execution, resulting in systems that are easier to control and less brittle, adapting more readily to unanticipated physical conditions and use cases. Most importantly, researchers would find ways to ensure that security is an integral system property so that systems are secure by design. There would need to be understandable tools and methods for expressing trust.

Knowing that systems evolve and change over time, future-proofing system security (perhaps by reusing complex modules) would be of great practical importance. Such a future would require assuring end-user control over information flows and code execution, and addressing the impact of Moore's Law on the capabilities of bad actors. Attendees estimated that addressing these problems would cost $400-600 million over ten years. Failure would contribute to social disruption, political chaos, and significant lawlessness.

> GC 1: *Within the decade, eradicate widespread viral, spam, and Denial of Service attacks.*
>
> GC 2: *Create the scientific principles, tools, and development methods for building large-scale systems for operating critical infrastructure, supporting democratic institutions, and furthering significant societal goals, ensuring their trustworthiness even though they are appealing targets.*
>
> GC 3: *For the coming dynamic, ubiquitous computing systems and applications, create an overall framework to provide end users with comprehensible security and privacy that they can manage.*
>
> GC 4: *In the next ten years, aim to create and implement quantitative models, methods, and tools for managing information systems risks that are on par with quantitative financial risk management techniques.*

Hits and Misses

We asked the 2023 online panel to grade community performance in addressing the four challenges. The initial responses were discouraging. Many participants said, "Not a single challenge was met." Others pointed out (correctly) that we had been only dimly aware of the scale and trajectory of the problems we were addressing. GC1, for example, was assumed to be a five-year, $600 million problem – in retrospect, a woeful underestimate.

Some failures can be traced to assumptions that did not anticipate the pace and scale of technological change. One central area where the conference fell short was agreeing on the impact of inexpensive mobile technology on the overall security landscape[18]. The explosion of connected devices introduced new attack vectors and made it more

[18] A different but slightly overlapping group met in 2000 and more accurately predicted the impact of mobile and “always on” connectivity. Their conclusions were made available to the Grand Challenge workshop attendees, but they were explicitly disavowed by a majority of them. Cf. https://www.cerias.purdue.edu/site/blog/post/who_says_you_cant_predict_the_future/

challenging to secure systems. Consumer-grade IoT devices have now become ubiquitous, and their users need more training to appreciate the effect of connectedness on security and privacy. In a parallel enterprise trajectory, gigabit networks, powerful distributed computing capabilities, and the rise of cloud computing created new challenges for security and privacy. Challenges such as these were not adequately anticipated in 2003, and the cybersecurity community has had to play catch-up as new threats and risks evolved.

Other assumptions underestimated the capabilities of threat actors and their ability to influence the global distribution of technology to penetrate vulnerable systems. Ransomware and botnets are examples of attacks that were dimly (if at all) considered in 2003. The enabling e-commerce models for packaging, weaponizing, and selling malware did not yet exist in 2003, and sophisticated malware-based attacks were uncommon. Cryptocurrency, a core enabler of current online extortion and crime, was not imagined by workshop attendees. Supply chain attacks were understood as a potential issue in 2003, but the current magnitude and complexity of modern development were beyond our ken at the time.

It is apparent from transcripts of conference breakout sessions that eliminating epidemic-style attacks was thought to be susceptible to a complete solution rather than a continuing problem that would require the complete elimination of threats. The goal of eliminating threats must be addressed by more than the research enterprise and is certainly well beyond the resources that the estimated $600 million buys.

Others pointed out the lack of understanding of application subject matter areas. Here, for instance, is one version of GC2: "By November 2008, design, build and deploy an electronic system to safely and securely with 100% accuracy tabulate the votes in a national election." The problem statement does not apply to elections in the United States. A U.S. federal election comprises 10,000 or more independent contests, each using mutually incompatible legislation and rules determined by states and localities. When combined with human errors and misunderstandings, a fragmented electoral system makes 100% accuracy impossible. Even if a federal mandate were possible, the system envisioned by GC2 would ignore other aspects of conducting a complete election (guaranteeing ballot secrecy, for example) that have nothing to do with tabulation. We also now know[19] of entirely new risk vectors (disinformation, for example) enabled by social media that undermine trust in systems that support governmental and societal functions.

[19] Report On The Investigation Into Russian Interference In The 2016 Presidential Election (https://www.justice.gov/archives/sco/file/1373816/download)

Critiques of the original grand challenges have some merit. However, as became apparent in the online forum, they should be paired with progress achieved over the past twenty years. Foremost among these is the idea that cybersecurity is an "enabler" for designers. Similar to brakes that enable cars to go faster but with greater confidence, the purpose of security is to enable computing technology to be applied in high-stakes applications. Similarly, while eliminating attacks may not be achievable, cybersecurity research has reduced overall susceptibility and allowed technical and business solutions to reduce the incidence of DDOS attacks. Advances in cognitive security and theories of design have created interdisciplinary approaches to replace usable security with human-centered security.

Even the basic notion of trust has been reexamined in light of knowledge developed in pursuit of the grand challenges. Resilient and Zero Trust Architectures enable developers to build systems that detect, contain, and recover from compromised states, creating a more secure operating environment with unsecure system components.

New Challenges

Today's computing environment is vastly different than the one anticipated in 2003. To illustrate the nature and speed of change in cyber security, consider the developments in the few weeks leading up to CERIAS 2023. By early January 2023, it had already become clear that Large Language Models (LLM) might be an unpredictable, disruptive force shaping information technology: A hundred million users shared information in prompts, and individuals used these new engines to suggest threat models and began probing vulnerabilities in IT systems, breaking already fragile assumptions about scale and capabilities. The public launch of ChatGPT was followed immediately by a new National Cybersecurity Strategy[20] that promised to rebalance and realign existing approaches to take into account changed threats and economics of the cybersecurity marketplace.

The new strategy also promised to rebalance market forces. Cybersecurity measures had been added to the already long list of tasks heaped upon users who, in essence, assumed responsibility for applying security patches, tracking threats and vulnerabilities, and understanding how to detect and contain rogue software delivered to their computers by manufacturers who put it there, to begin with. Despite decades of research in so-called usable security, the unfairness of this approach had become apparent to many, including the 2003 conferees. The National Cybersecurity Strategy shifted the burden (and risk) from users to hardware and software vendors, drawing high praise from the 2003 GC committee.

[20] https://www.hsdl.org/c/abstract/?docid=875831

In addition to the scale and unanticipated capabilities of attackers, the growth of the cybersecurity workforce skills gap has become a dominant concern and stretched thin educational resources in ways not imagined in 2003[21]. In a similar vein, an increasingly divisive and contentious social and political scene illustrates the role that insiders, nation-states, political actors, and domestic terrorists might play in defining the threat landscape.

Almost all of these developments spawn problems that seem to require interdisciplinary thinking. Technology alone cannot sustain cybersecurity research and development. Sophisticated policy solutions, tools for law enforcement, and empirical methods not discussed by the 2003 conferees will undoubtedly play critical roles in defining cybersecurity challenges over the coming years. The field must expand to include problems in economics, psychology, law, social equity, international affairs, cyber-physical systems, and the basic philosophy of social media. Interestingly, this was the core concept when CERIAS was established in 1998 — to take an interdisciplinary approach to cybersecurity involving more than computing technology; only recently has the field at large begun to embrace that idea, and not all entities involved have yet done so.

Besides resurrecting and revising the original GC problems, new Grand Challenges to counter side channel, supply chain, insider, and domestic abuse threats deserve attention. The relationship of safety to security needs exploration. A well-grounded theory of privacy is still elusive, and an understanding of combined hardware and software security for emerging applications and potentially significant technologies such as blockchain and decentralized finance, quantum computing, and ML over extensive data is still in its infancy.

Summary and Recommendations

As the Grand Challenge committee recognized in 2003, today's speed of change and reliance on information technology is increasing. Now, as then, we confront on an unprecedented scale the risk of significant disruptions, including failures in power, transportation, and communication systems, privacy breaches, data tampering, and novel types of theft and fraud. To the 2003 threats from criminals, anarchists, extremists, cyber terrorists, and indiscriminate attackers, we add escalating attempts by nation-states, terrorist networks, and insiders attempting to hijack the tools of democratic governance. These attacks compromise security and, in the end, trust. A computing infrastructure must be resilient against such attacks to be considered trustworthy. Still, the challenges to achieving resilience are not obvious.

[21] https://www.nist.gov/document/workforcedemandonepager

The time is ripe for a new Grand Challenge Symposium in Cybersecurity. A new panel to define cybersecurity's 2023 Grand Challenges carries the same possibilities and drawbacks that existed twenty years ago. Still, we know from prior experience that such a convening has a lasting influence on the research community. It gives structure to debates and proposals that would otherwise occur in fragments. Over half of the principals of the 2003 GC Symposium became research leaders in cybersecurity in the last two decades, and almost all went on to prominence. Their experience debating the research challenges undoubtedly informed a generation of colleagues, students, and constituents, and rather than stifling debate–by prematurely declaring some problems important and others not–became the seed for more robust discussions.

Acknowledgments

The authors thank the following 2003 GC participants for their contributions and support in 2023 for this report: Virgilio Almeida, Annie Antón, Terry Benzel, Luis Bettencourt, David Clark, Steve Crocker, Jeremy Epstein, David Evans, Simson Garfinkel, Dan Geer, Anup Ghosh, Virgil Gligor, Seymour Goodman, Doug Jacobson, Jay Lala, Carl Landwehr, Peter Lee, Ruby Lee, Clifford Neuman, Cristina Nita-Rotaru, John Richardson, Angela Sasse, R. Sekar, Daniel Simon, Sean Smith, Anil Somayaji, David Stork, Ravi Sundaram, Bhavani Thuraisingham, Peter Wayner, and Jeanette Wing.

The 2003 workshop and report were supported, in part, by NSF grant CCR-0335324, and by the Computing Research Association.