



TRUSTED CI

THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

Science *and* Security: Sound Odd?

Trusted CI Fellows Report

January 27, 2021

For public release

Author, Laura Christopherson
Trusted CI Fellow
Renaissance Computing Institute

I served as a Trusted CI Fellow during 2020 while also working on the Cyberinfrastructure Center of Excellence (CI CoE) Pilot project.¹ In fact, it was through the CI CoE that I learned about Trusted CI and became interested in the fellowship. Over the past year, my work with CI CoE and the fellowship exposed me to the importance of information security in science. When I mentioned this intersection of science and security to friends or others outside of technology and academia, I often got puzzled looks. I think part of the confusion was because I was mainly talking about earth sciences (which is the type of research largely conducted by the research facilities that CI CoE supports) and I suppose people initially assumed I must be talking about health sciences. They could, of course, understand why security would be important in medicine. We all want our personal information (e.g., medical records) protected. And since March of 2020, COVID has been the leading story in all news sources, and those stories have included discussions about the importance of maintaining the integrity of COVID research data so that we can develop a vaccine as quickly as possible. It's a life-or-death issue.

When it was suggested that other kinds of sciences, other kinds of research, might need some protection however, then they seemed a little dumbfounded. For instance, I received comments to the effect of, "Well why would anyone want to steal images of a black hole? It's not private, confidential information. It's up there for all to see." And after all, don't we want to share this information? That's why scientists shared the first image of a black hole in April of 2019.² That gave me pause, I admit: Well yeah, that's not private, personal information. No person would be compromised in any way or suffer any harm if the read-outs from a particle accelerator were disclosed by WikiLeaks, right? Earth science is not a life-or-death situation, after all. There's no money in stealing data from Laser Interferometer Gravitational-Wave Observatory (LIGO)³ or IceCube.⁴ Furthermore, don't people want this information shared? What about all that "open science" jazz anyway?

While it may be true that cyber thieves would be less inclined to attempt to steal information or disrupt the activities of scientists when sexier alternatives are available (e.g., the bank accounts of millions of Wells Fargo users, presidential election tabulations, juicy emails between a senator and his mistress, design schematics of a nuclear warhead, personal health information of patients participating in a highly controversial drug trial), it is possible that cyber criminals—in targeting those sexier-alternatives—may unknowingly hit humble research organizations because they also happen to use the same systems that businesses and governments use. The SolarWinds hack⁵ is a good example of this, as described by Kim Milford, the executive director of the Research and Education Networks Information Sharing and Analysis Center at Indiana University. "While it does not seem at this time that higher education institutions or sensitive

¹ <https://cicoe-pilot.org>

² <https://www.jpl.nasa.gov/edu/news/2019/4/19/how-scientists-captured-the-first-image-of-a-black-hole/>

³ <https://www.ligo.caltech.edu/>

⁴ <https://icecube.wisc.edu/>

⁵ <https://www.insidehighered.com/news/2021/01/06/unraveling-solarwinds-hacks-fallout-higher-ed>

research secrets were the target of this attack, it is possible that hackers may have scooped up so much information they do not yet realize what they have," Milford said in an Inside Higher Ed article. In other words, cyberattackers may unknowingly steal a scientific Easter egg that they could crack later for what may turn out to be a goldmine of competitive research secrets.

When I think about the comments I received, I'm just not sure if the everyday Joe or Jane even thinks about science when considering the importance of security. I suspect they largely think about themselves and their personal information instead. It's only natural. But I wanted to understand how to articulate the importance of information security in science. So, I set out to find information on various questions around the intersection of science and security. Putting aside healthcare-related research (for the remainder of this paper), I wanted to know:

- How does the average American conceive of security? What are average concerns about information security? Do scientific research organizations even enter the picture for the average American? Does the average American think it is important to protect scientific research?
- Do scientific research facilities get attacked? What risks do they face?

I scoured the library's databases of research articles, traipsed through the Internet using a variety of search terms, perused various polling/social science research organization websites (e.g., Pew, the National Academies Press), trolled through popular online tech magazines and blogs, and was ultimately unable to find fully satisfactory answers to my questions. So, I asked Von Welch, director of Trusted CI, the NSF Cybersecurity Center of Excellence, if he knew of any reports of attacks on scientific research organizations. He was able to locate only two publications related to this subject: one from the Australian National University,⁶ reporting a breach to their administrative systems, and an FBI case study⁷ reporting attacks on military sites, federal research labs, universities, and other sites, discovered in 2004 and resulting in the arrest of a 19-year-old man in 2005.

It appears that there is dearth of information on:

- The public's awareness of or views on whether security matters in science
- Threats faced by scientific research organizations
- Consequences and impacts if scientific research organizations experience loss or damage to precious research findings.

In my review of information about attacks/security in the non-academic/research world, I uncovered two themes. One was about the nature of the attack and the second was about who is usually attacked. The nature of an attack is often described as a theft of some kind. There always appears to be some discussion of what was lost and its value, what the hacker sought as

⁶ <https://www.insidehighered.com/news/2021/01/06/unraveling-solarwinds-hacks-fallout-higher-ed>

⁷ Ricker, Kathleen & Barlow, James & Adams, Craig. (2008). FBI Major Case 216: A Case Study. 10.13140/2.1.2775.2644.

his reward. The most discussed prizes seem to be money (Wells Fargo bank accounts), power (presidential election tabulations), reputation (juicy emails), strength (nuclear warhead), or access to some deep secret (personal health information). Because earth science data won't really give you money, power, reputation, and strength in the way we usually think about those things, and because it won't give you access to deep, dark, personal secrets to leverage against your enemies, why would a cyber thief bother?

Frequently discussed targets of attacks were financial institutions (money), governmental institutions (power), nations (strength), and individuals (reputation and secrets). (Research bodies are also mentioned but they tend to be those that conduct biomedical research which I would still classify as reputation and secrets, because the data at risk is often personal information of specific individuals, and it is often the risk to these individuals that the discussion centers on.)

Had I surveyed the news over the past decade, I imagine I would have found very similar results... that most news stories primarily report on that which was stolen from individuals, profit seeking businesses, or governmental/national/political organizations. Off the top of my head, when I think of recent, big news stories about security, I think Russia and the 2016 presidential election, Facebook and Cambridge Analytica, Independence Blue Cross, Wannacry (ransomware), Target, Hillary Clinton's email server, Cal Cunningham (NC senatorial candidate), and Equifax, to name a few. I can't think of a single instance of any news story discussing an attack on an earth science research facility. Although my personal recollections don't confirm the absence of attacks (i.e., it just confirms that I haven't heard of any), I still ask you, my reader, did you hear of any? If you did, how many compared to the other kinds of attacks you also heard about? I suspect it's just not a hot topic for most news outlets.

The point of all this is to say that in spite of not finding any information that said, "Hell yeah, security is really important in science, for good reason," I still conclude that Hell yeah, security is really important in science for good reason... in fact the same reasons, but perhaps with a different way of thinking about them. First, I think the more mainstream definition of security and what it means to secure data might require expansion when discussing research. For instance, many of the research facilities we work with in CI CoE have to protect their data from harsh environmental conditions. IceCube is located at the South Pole. Its equipment could freeze, and data could be lost. So, the data must be protected... from the ice (less so probably than from some hacker).

Additionally, it may be worthwhile to rethink those more commonly discussed prizes (money, power, reputation, strength, and juicy secrets). If we concede that a cyberattacker is less likely to find these prizes from stealing scientific data, do they (money, power, reputation, strength, and juice secrets) enter the discussion at all? I would say yes, but in a different way. Instead of

being the reward at the end of the maze, I would argue that they are qualities inherent to science and so can't be stolen from it. They are not the hoped-for results of some activity (e.g., theft), but that which is intrinsic to science, and consequently, make it vital to protect science.

Money = Valuable

The NSF spends millions of dollars funding earth science research. If research activities are disrupted, if data is corrupted or lost, then that money has been wasted. So, although you may not get rich off of studying earthquakes or by stealing images of the moon, science is a priority in our society, and we've invested decades of money into it. The American public's tax dollars support scientific research, and we all want a good return on our investment. This affects us all.

Power = Powerful

It is said that "knowledge is power." Science seeks to uncover new knowledge, and it has empowered us in numerous ways. Consider a very simple and practical example of how science has improved our everyday lives. Because we sought to understand electricity and harness its power, we are able to enjoy the comfort of heating and cooling, have light to see by, and can enjoy hot meals cooked on a stove from ingredients preserved in a refrigerator. Science also tackles issues vital to our survival as a species on this planet. It explores questions about natural energy (which can be used to power medical devices that keep us alive), our carbon footprint (which impacts the resilience of Earth's ability to sustain life), and weather and climate change (which affects the habitability of Earth, important when considering future generations). So, in a sense, earth science is a life-or-death issue after all, but perhaps on a broader scale, because it concerns mankind.

Reputation = Noteworthy

Because we depend on science for so many things, it is important that the outcomes of scientific studies are accurate. If scientific data is put at risk, it calls into question the findings of scientific researchers. Years of work can be invalidated, reputations destroyed, and trust eroded. Each year in the history of our existence, we have continued to build upon this knowledge. We have a very sizeable bank account of knowledge from which to draw on and help us advance. Just as our personal or business bank accounts containing money ought to be protected, so should this wealth of knowledge the scientific community has socked away.

Strength/Bold

To use another cliché, it is said that "there is strength in numbers." Most of the science research facilities that we work with in the CI CoE serve thousands of scientists (students and professionals) from around the world. For example, the partnership of the Seismological Facilities for the Advancement of Geoscience (SAGE)⁸ and the Geodetic Facility for the Advancement of Geoscience⁹ estimate they serve, roughly, 10,000 scientists worldwide. NOIRLab¹⁰ (a collection of five observatories) estimates a user base of 3,000-5,000 per quarter. The Natural Hazards Engineering Research Infrastructure (NHERI) is composed of multiple units. One of those, DesignSafe,¹¹ which provides computation services for analyzing hazards data, estimates a user base of 5,000, with roughly 1,000 using their services each month. If each of these facilities serve approximately 1,000 people each month, then they collectively serve several hundreds of thousands of scientists (students and professionals) from a variety of earth-science disciplines throughout each year.

These facilities also manage very large datasets. NHERI-DesignSafe manages roughly 200 TBs of data. The Oceans Observatory Initiative¹² pulls in around 15,000 rows of data every 30 seconds, roughly 10 TBs of data every three months. The Cornell High Energy Synchrotron Source (CHESS)¹³ collects around 120 TBs every few months. SAGE ingests around 10 TBs of data per year and has a total archive of roughly 650 TBs that has been collected over 40 years. The image archive for NOIRLab manages almost five PBs of data.

Within these PBs of data is the possibility to uncover tremendous new insights about our world. Scientists from all over the globe depend on these facilities to support their work. Even though the present-day Galileo may not come to mind when thinking of information security, he and many others exist, and they rely on these PBs and PBs of data to uncover new knowledge about our world. Across the various facilities we serve in CI CoE, there exists a very strong userbase that uses extremely large, multifaceted datasets that may very well exceed the bytes needed to store the emails on Hillary Clinton's server, the Target credit card accounts that were breached, and the 2016 election tabulations that may have been tampered with.

⁸ https://www.iris.edu/hq/news/story/nsf_makes_5_year_93m_award_to_iris_to_manage_the_sage_facility

⁹ <https://www.unavco.org/about/about.html>

¹⁰ <https://noirlab.edu/public/>

¹¹ <https://www.designsafe-ci.org/>

¹² <https://oceanobservatories.org/>

¹³ <https://www.chess.cornell.edu/>

Secrets = Discoveries

Science data probably doesn't contain any personal information that might embarrass someone, put them in a negative light, or compromise their credit rating. However, it probably does contain an entire host of secrets that, unlike personal secrets, we want and need to uncover. For instance, LIGO collected data for more than a decade, waiting to discover new knowledge, before they finally detected gravitational waves that allowed us to look back 1.3 billion light years at two colliding black holes.

This discovery comes at the culmination of decades of instrument research and development, through a world-wide effort of thousands of researchers, and made possible by dedicated support for LIGO from the National Science Foundation. **It also proves a prediction made 100 years ago by Einstein that gravitational waves exist. More excitingly, it marks the beginning of a new era of gravitational wave astronomy** – the possibilities for discovery are as rich and boundless as they have been with light-based astronomy.¹⁴

So, to end on that auspicious note, I hope I have made a good case for the importance of security for science, in spite of the lack of research I was able to find in this area. It is because of this lack of work, that I will now try to convince you of one last thing: There needs to be more research in this area. In the Inside Higher Ed SolarWinds article, Kim Milford encourages "cybersecurity leaders to provide thought leadership and guidance" on this subject.

For my part, I suggest there be work around the following questions/themes:

What does security look like in science? What are the threats?

I have suggested that it may be unlikely that cyber thieves will target science when it doesn't really afford them the prizes they may typically seek. So, security may be less about guarding against malicious actors and more about making sure the data is well protected from other kinds of threats faced by so many research facilities. Perhaps this calls for redefining security when applied to science.

Are earth-science facilities the targets of malicious attacks? If so, how and why does this happen? How does that compare with the other threats they face?

Although I found little evidence of malicious attacks, Von Welch was able to locate information on the subject, so malicious attacks do happen. Why do they happen if they don't yield the same prizes that are stolen from other types of targets? Are there things to be gained—other

¹⁴ <https://www.ligo.caltech.edu/detection>

prizes—I did not imagine? If so, this may be very helpful information to technology professionals working in science. It could also expose other dimensions to the motivations of black hat hackers, which could be explored by social scientists as well as computer scientists.

Why is it important to protect science? How does science benefit us all?

I have suggested that when people think of security, they tend to think of themselves, their valuables, their secrets, their associations; and that perhaps this is why science may fail to come to mind when thinking about security. I have also attempted to point out that science has implications beyond the individual, group, or organization, that it concerns and benefits mankind. If this is so, then it is particularly important to raise awareness about the importance of protecting scientific data. I believe this will also help validate the work of technology professionals who stand guard at the gates of science. We hear about the latest scientific discovery and the scientists involved, but the contributions these guardians make to science may not be considered newsworthy. I get the sense that, as a result, they are often overlooked and perhaps not valued in the way they deserve. So, I urge both the science and technology community to work on changing this.

Finally, I think future Trusted CI Fellows are the perfect candidates to explore these questions and to publish on these subjects. I hope that I have inspired future Fellows to pursue these questions. May they achieve success no matter what they pursue in the future, and I wish them well.