Cyber Case Study Program

Starfruit Lab: A Cybersecurity Dilemma Case Study

Case Number ACI-03-2025

LTC Jason C. Brown, PhD

Editor in Chief: Karen Guttierri, PhD
Lead Editor: Volker Franke, PhD
Managing Editor: Anne Chance, PhD

About Case Studies

The ACI-TRENDS Global Cybersecurity Case Program is administered by the Army Cyber Institute at the United States Military Academy at West Point and publishes cybersecurity-related teaching cases, simulations and interactive exercises for use in academic and professional classrooms.

What are case studies?

Case studies are high-impact interactive learning tools structured around real or realistically simulated events placing learners in the role of decisionmakers confronting complex problems, tradeoffs, and uncertainty. Case studies immerse participants in the problem and its dilemmas.

Why use case studies?

Immersing participants cognitively and emotionally in this way requires them to grapple with ambiguity, incomplete information, and competing priorities while developing plausible courses of action.

Case studies are particularly important because cyber incidents unfold across technical, organizational, legal, and human domains simultaneously.

Case studies make invisible systems and cascading consequences visible, demonstrate how theory and policy apply under pressure, and build the analytical judgment required to anticipate, assess, and respond to real-world cyber threats.

Where to find ACI case study series.

More information on the case method can be found:

1. Link to case study page on TRENDS
2. Reference to VF article with link
3. **If there is a link to an ACI or WPP publication**

Acknowledgement

An earlier version of this case study was developed and published by the cyber college at Air University, Air Force Cyber College in Montgomery.

About the Author

Dr. Jason C. Brown is a research scientist and assistant professor at the Army Cyber Institute at West Point. He teaches risk management, organizational security, and systems-based decision making. As a futurist, he studies emerging threats, technological and social trends, and responses to those threats. A currently serving officer with the U.S. Army, Dr. Brown has worked within the intelligence, information operations, and cyber career fields. He has authored technical reports on the futures of extremism, information warfare, cyber enabled financial crimes, microtargeting, and Chinese soft power.

Starfruit Lab: A Cybersecurity Dilemma Case Study

Situation

Starfruit Lab is a fictional artificial intelligence company, developed by two Israeli entrepreneurs. Starfruit's proprietary machine learning algorithms were created to take advantage of vast quantities of anonymized personal and public health data, consumer genomics, and biomedical research to identify personalized dietary and legal supplement recommendations for elite athletes.

Starfruit Lab's technologies will be at the center of a ten-year research project monitoring cadet physical and mental performance through personalized dietary and supplement intake.

The Army Application Lab, an integration team at Army Futures Command (AFC), received approval for testing a business model that allows the Department of Defense to purchase startup and small businesses with high research potential to be managed by select research and educational organizations. AAL successfully purchased Starfruit Lab for $250 million three years after the startup was created. AAL temporarily assigned responsibility to manage and oversee Starfruit Labs to AFC's Artificial Intelligence Integration Center (AIIC) team under a new federal sandbox regulation program.

The same year, the U.S. Military Academy at West Point (USMA) received approval to establish a research center for Soldier Performance and Applied Health (SPAH). The SPAH Center is scheduled to take control of Starfruit's assets from AIIC in the next six months. Starfruit Lab's technologies will be at the center of a ten-year research project monitoring cadet physical and mental performance through personalized dietary and supplement intake. This health research project has been reviewed for human subject research implications, ethics, and safety considerations.

Mission

It's a simple mission: Your team is tasked to conduct asset, vulnerability, and threat assessments for Starfruit Lab interfacing with Army networks.

It is a decade in the future. You are part of the USMA Chief Information Officer's (CIO) team that will review Starfruit Lab's security policies, data storage, network interfaces, and computing systems. To do this, you are tasked to conduct asset, vulnerability, and threat assessments. You will need to develop recommendations on how Starfruit Lab's assets will comply with current and projected security policies to interface with USMA and Army enterprise networks. Army Cyber Command can provide one Cyber Protection Team (CPT) to assist with some of the technical and security procedures after your assessment is complete.

Dilemma #1

As part of the CIO assessment team for Starfruit's on-boarding process, you have spent long days and many weekends reviewing system configurations, policy & compliance documentation from Starfruit's original owners, and an in-depth threat assessment. Your team has already remediated dozens of potential vulnerabilities, installed patch after patch, and has tested interoperability with the West Point Research and Education Network version 2.0 (WREN 2.0) in a virtualized environment. You are finalizing the plan of action and milestone briefing to the CIO to recommend that Starfruit will be ready to interface with the WREN 2.0 in approximately seven days. This lines up

exactly when the SPAH's research project hits Phase II, which requires the AI engine to begin churning out individualized meal, fitness, and dietary supplement plans to 4,000 cadets.

The cyber protection team (CPT) tasked to you from Army Cyber Command to evaluate and prepare Starfruit for live deployment has three more days of evaluations and checks. The CPT team leader approaches you and reports several anomalies in the logs between a Jupyter server and the AI engine. It appears that the AI engine can seek a secondary data source and has pinged an external IP address traced back to Israel at least six times in the logs. There is also evidence of the "Explosive" remote access tool (RAT) in the web front-end, which could indicate infiltration by Lebanese hacktivists nicknamed Volatile Cedar. The CPT team leader suspects that the machine learning code of the AI engine may have been training on a data set that you don't entirely control.

The threat begins to reveal itself: You find evidence of the "Explosive" remote access tool (RAT) and log files showing unauthorized contact with an Israeli ISP.

Of course, the simple fix is to block access to the unknown IP traced back to Israel and remove the Explosive RAT, but that does not help you solve what other problems may be under the hood! Further consultation with Starfruit's software engineers about this problem suggest the fix could take anywhere from 2-4 months to debug and trace. However, simulation tests have been 99.94% accurate on modeling desired outcomes from the AI engine as it identifies the health and dietary plans for the test population, so the machine learning engine does not appear to be corrupted.

Delaying a live deployment of Starfruit onto the WREN 2.0 likely will put the SPAH research plan behind by several months. In a progress report meeting with SPAH leadership, you find out that the Phase II timeline is irrevocably linked to additional grant funding from the National Institutes of Health (NIH). If SPAH does not meet certain data analysis milestones in the next fiscal quarter (three months), the NIH can legally revoke the next iteration of the grant. This also means that three of the four Starfruit software engineers and six SPAH health and wellness specialists who are funded with the NIH grant would lose their jobs.

The first dilemma: completely verifying the threat puts the entire research program at risk; however, simulations suggest the threat vectors did not achieve their objective! What to do?!

The CIO and SPAH director ask you these questions: Has the training data for the AI been manipulated? Who was on the other end of the IP address? Are there other back doors within any other part of the software or any part of the system? What is your risk assessment recommendations given this new information?

What is your risk assessment recommendations given this new information?

Dilemma #2

In the threatcasting report entitled, "Digital Weapons of Mass Destabilization,"[1] the Army Cyber Institute, the Defense Threat Reduction Agency, and Arizona State University teamed up to investigate how the cyber landscape and weapons of mass destruction might generate new threats or become more disruptive in the next decade. During a data collection workshop,

1 Brian David Johnson, Jason C. Brown, and Josh Massad, Digital Weapons of Mass Destabilization: The Future of Cyber and Weapons of Mass Destruction, (Tempe, AZ: Arizona State University, 2020), https://threatcasting.asu.edu/publication/digital-weapons-mass-destabilization.

several teams imagined how DNA manipulation and at-home DNA kits are becoming cheaper and easier to access.

Over the next decade hackers working on behalf of the Chinese government successfully (and surreptitiously) infiltrate the SPAH genetic database to target random cadets. Using genetic data stolen from 23andme, Ancestry, and other commercial vendors,[2] the hackers subtly modify the records of these individuals, and the program distributes much stronger doses of performance enhancing supplements that are coded directly to the DNA of the recipient. They are also directed to eat foods with higher amounts of protein in their diets. Their exercise programs are also increased to take advantage of the raised levels of protein. This leads to the cohort of cadets gaining significantly more muscle mass and performance increases over those in the control group.

A foreign adversary has proven it can surreptitiously access and manipulate U. S. Soldiers at a genetic level.

The SPAH and CIO leadership are alarmed, not by the performance increase, but by the implications. A foreign adversary has proven it can surreptitiously access and manipulate U.S. soldiers at a genetic level. This could be a 'benign' test for a far more sinister capability: to degrade soldier performance, test genetically-specific bioweapons, or create deep ethnic divisions within the ranks. The fact that the hackers chose to improve performance only makes the ethical dilemma of how to respond more complex.

The health advisory team overseeing the cadets' fitness and wellness program have noticed the anomaly of cadets exceeding expectations and asked the entire SPAH and Starfruit teams what could be causing it. They initially ask about data manipulation or other reasons why the numbers are so high, so the CIO requests another Cyber Protection Team mission from Army Cyber Command to do a deep hunt. Over the course of several months, your cybersecurity hunt teams have discovered some level of access and manipulation of the DNA registers.

Dilemma 2: Do you take the entire project offline or are the results innocent enough ('just making people stronger') to leave it running? What are your recommendations to leadership?

The CIO and SPAH director ask for your assessment:

- What is the likelihood the training data for the AI has been manipulated, and what is the worst-case impact if it has?
- Who do you suspect was on the other end of the IP address, and what might they have gained?
- How likely are other back doors, and can we trust the AI engine's 99.94% accuracy simulation if a RAT was present?
- Given these unknowns, what is your final risk assessment and recommendation for the deployment?

2 Atlamazoglou, Stavros. "China Is Scooping up DNA Data to Target Foreign Spies — and You, the US Government Says." Insider, March 12, 2021. https://www.businessinsider.com/chinas-dna-data-could-allow-targeting-spies-dissidents-citizens-2021-3.