SAND2020-8935PE

RAPID ANALYSIS OF MISSION SOFTWARE SYSTEMS

Todd Jones

September 10, 2020

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

Sandia National Laboratories
SAND Number: SAND2020-6931 PE

The Problem

AP Photo/Mark Zaleski

- **Question:**
  - How do I keep electricity flowing?

The Problem

Photo from http://carleyknight6.wordpress.com/

Backdoors Keep Appearing In Cisco's Routers

by Lucian Armasu July 19, 2018 at 10:00 AM - Source:

Over the past few months, not one, not two, but five different backdoors joined the list of security flaws in Cisco routers.

- **Question:**
  - How safe is my data on the internet?

The Problem

US Army Photo/Nate Allen

Secret Back Door in Some U.S. Phones Sent Data to China, Analysts Say

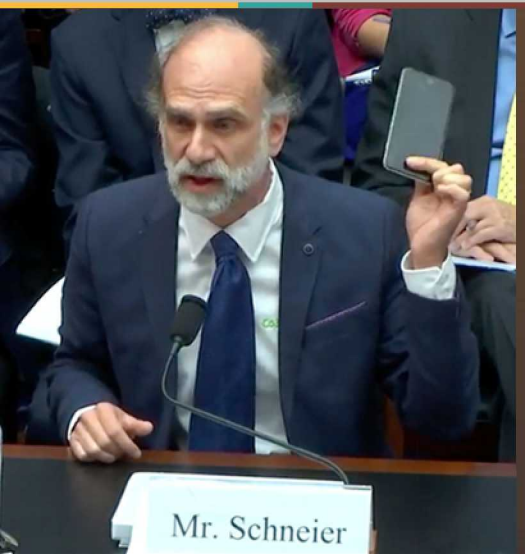
点击查看本文中文版

By MATT APUZZO and MICHAEL S. SCHMIDT NOV. 15, 2016

WASHINGTON — For about $50, you can get a smartphone with a high-definition display, fast data service and, according to security contractors, a secret feature: a backdoor that sends all your text messages to China every 72 hours.

- **Question:**
  - How do I keep sensitive information private?

A Common Thread



This is not a phone.

It is a computer that makes phone calls.

A refrigerator is a computer that keeps things cold. An ATM machine is a computer with money inside. Your car is not a mechanical device with a computer. It is a computer with four wheels and an engine. - Bruce Schneier

A Common Thread

What makes all these systems work is ***software***.

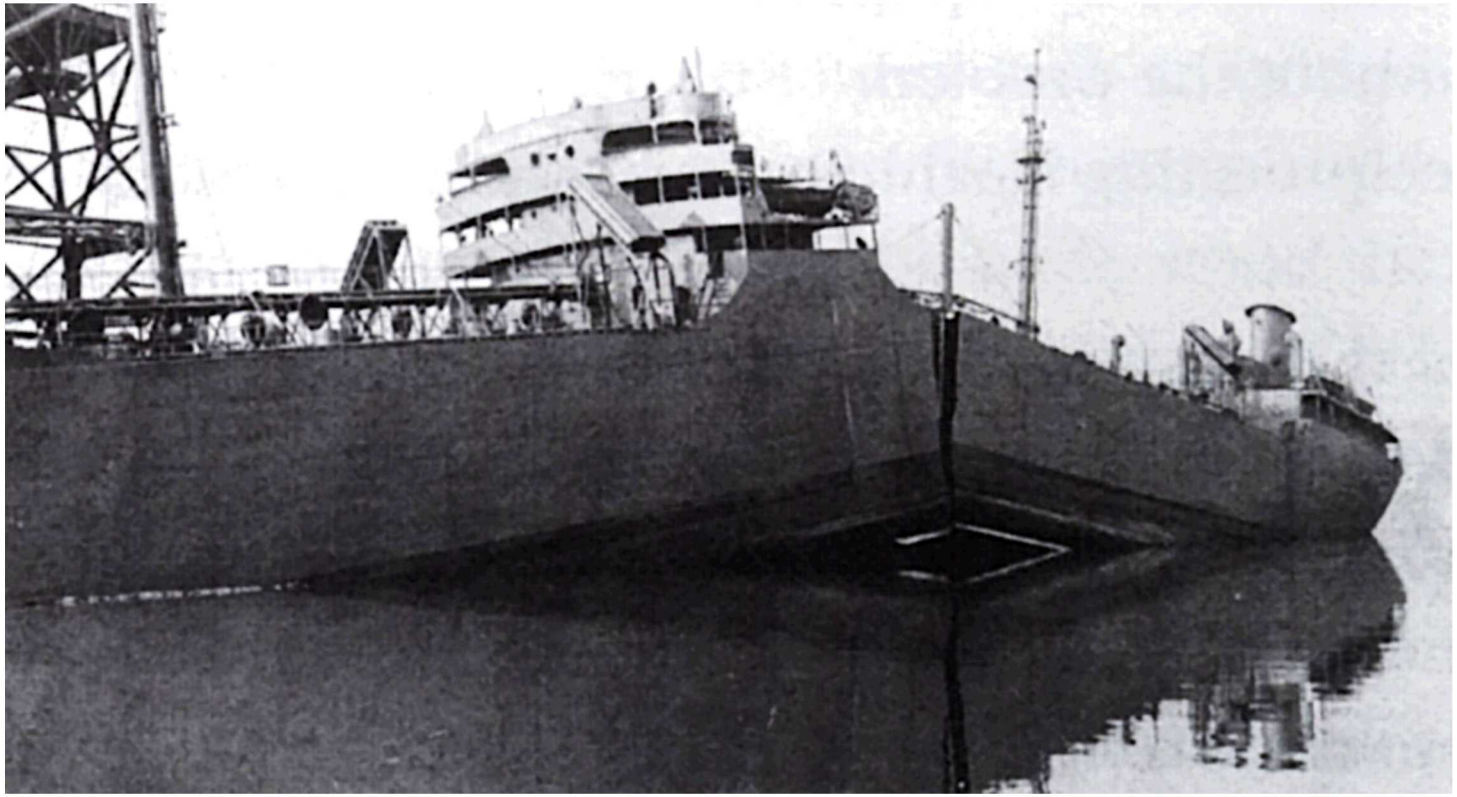
Software manages the
information and ***operation*** of these systems.

We want to answer questions about software.

The Challenge

Understanding

- Grade of steel
- Bulkheads
- Rivets
- Propeller
- Lubricants/cooling
- Effect of saltwater
- Effect of cold/ice
- Radar characteristics
- Operational range



Physical Systems

1940

The Challenge

Understanding

- Grade of steel
- Bulkheads
- Rivets
- Propeller
- Lubricants/cooling
- Effect of saltwater
- Effect of cold/ice
- Radar characteristics
- Operational range

Physical Systems

1940

A ship isn't a ship.
It's a computer that floats and moves.

TAKEAWAY:

Software is now the weak link.

We need to be able to pose and answer questions about software systems as well as we do about physical systems.

- Communications
- Radar/sonar
- Navigation
- Propulsion
- Power production
- Fire detection and suppression
- Engineering
- Fresh air/water

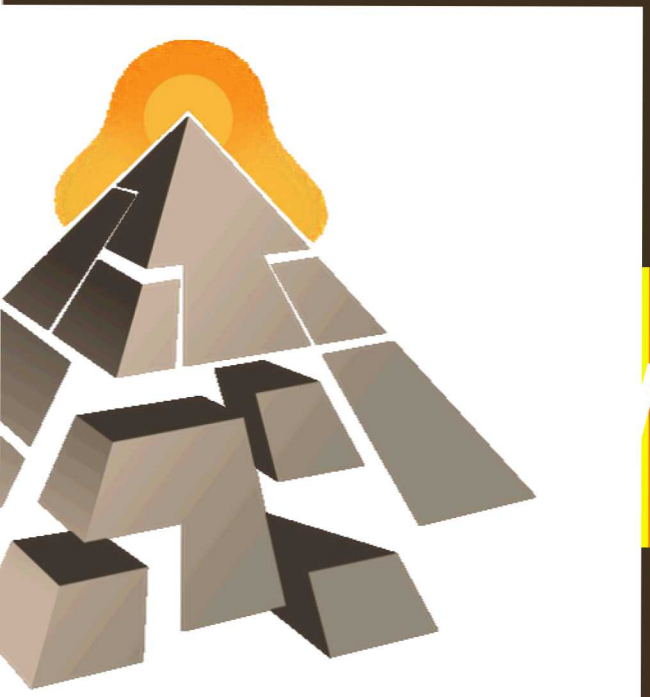
All the software:

…written under contract

…embedded in the commercial components

...used for design/evaluation/testing/support

2020

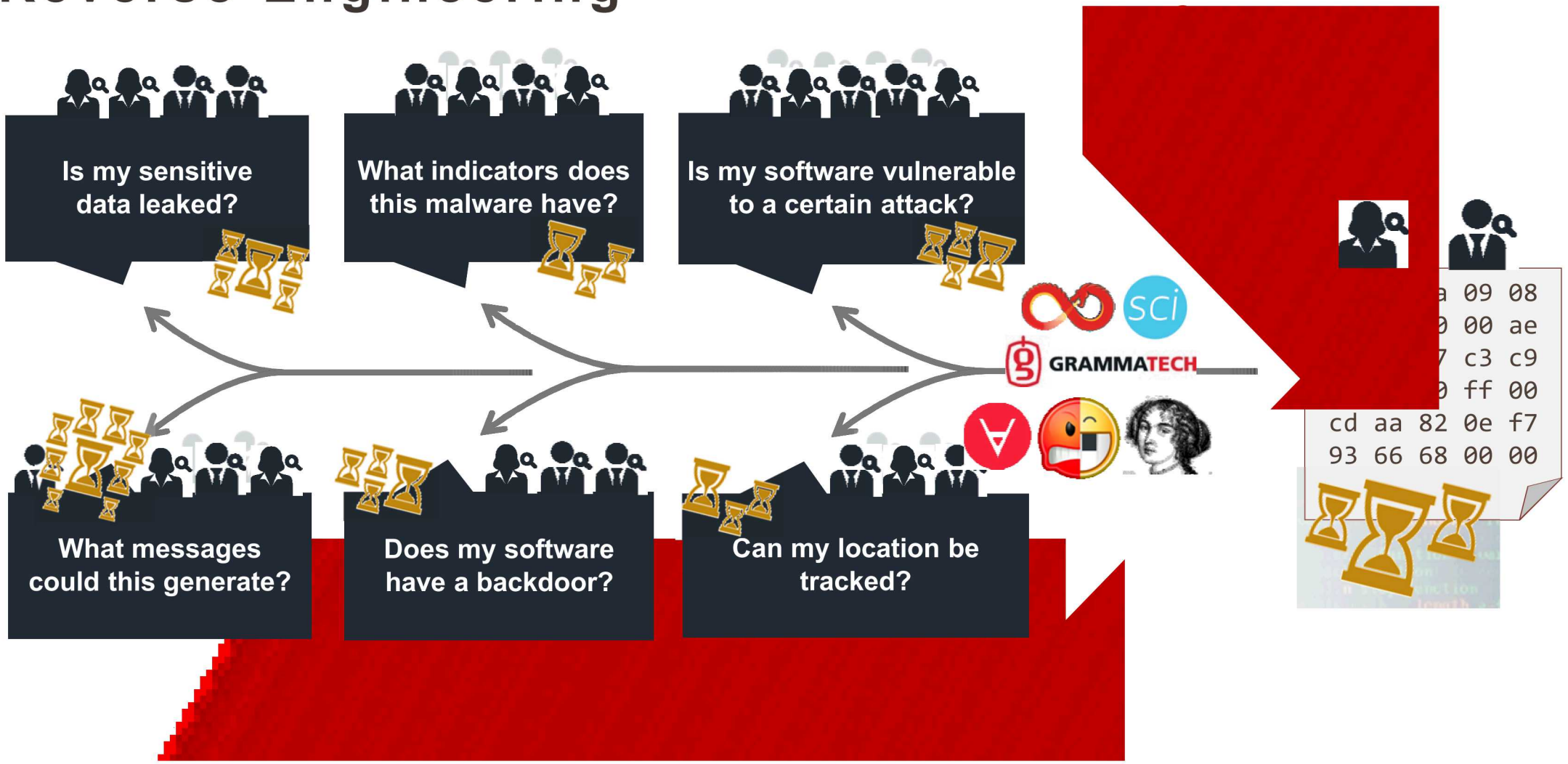


Why is there a problem?

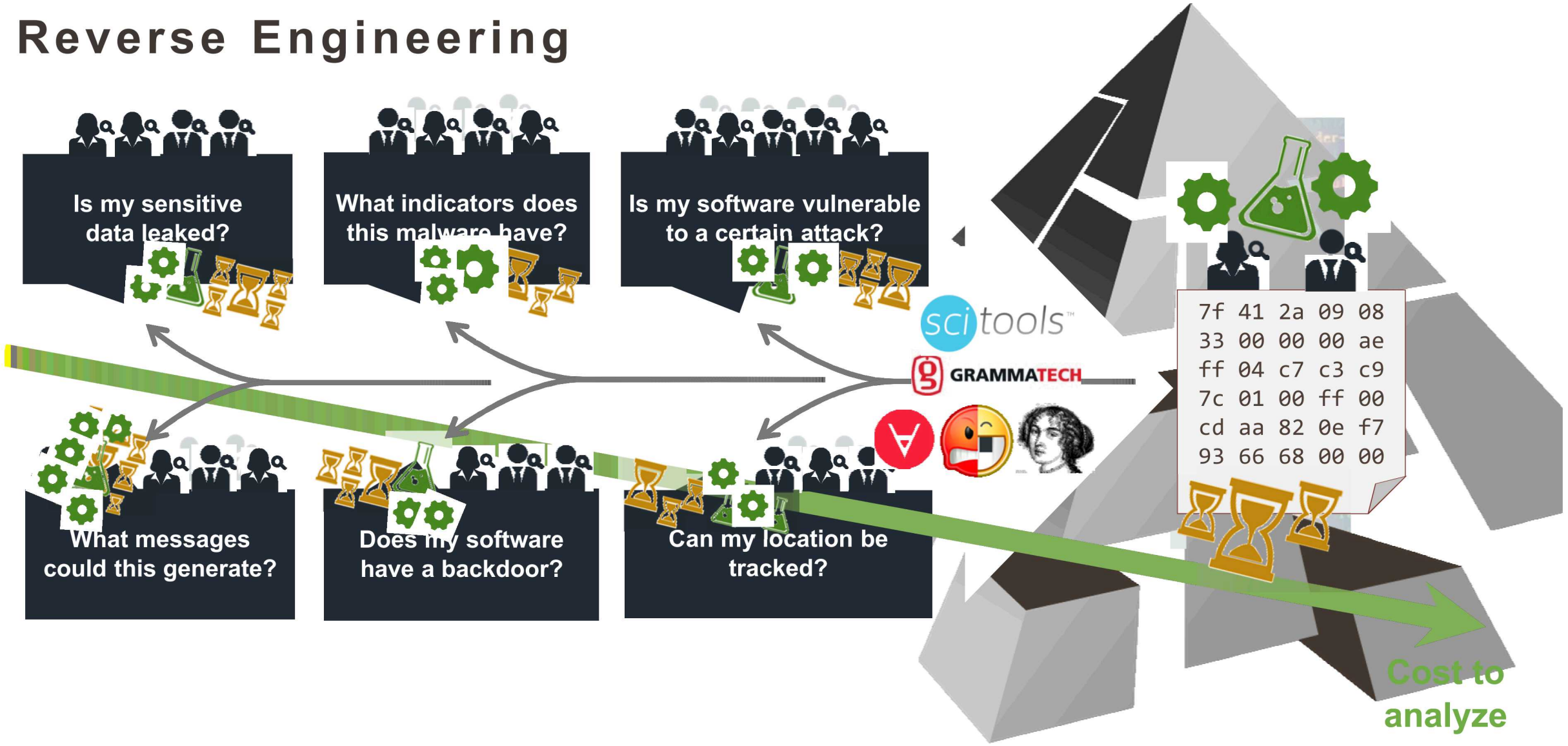
Software Development

```
7f 41 2a 09 08
33 00 00 00 ae
ff 04 c7 c3 c9
7c 01 00 ff 00
cd aa 82 0e f7
93 66 68 00 00
```

Reverse Engineering



Reverse Engineering



The Problem is Growing

MORE PLACES

Software driving more and more systems

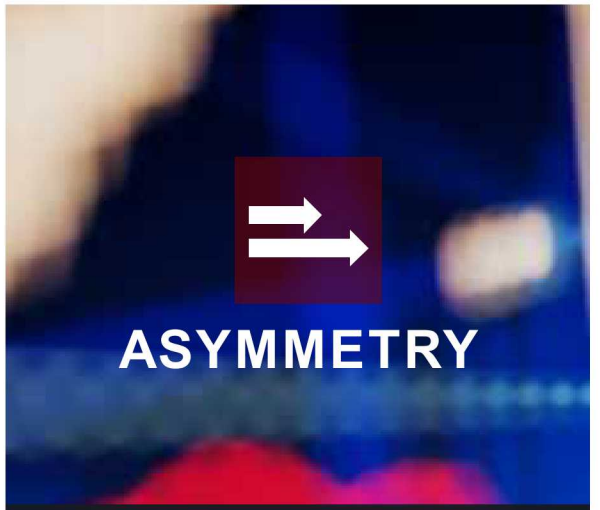
BIGGER

Increasing software size and complexity

Increasing number of versions

FASTER

Increasing rate of updates



ASYMMETRY

Developer tools improving faster than RE tools

Our Approach

Mission-driven, end-to-end focus - not a staffing company

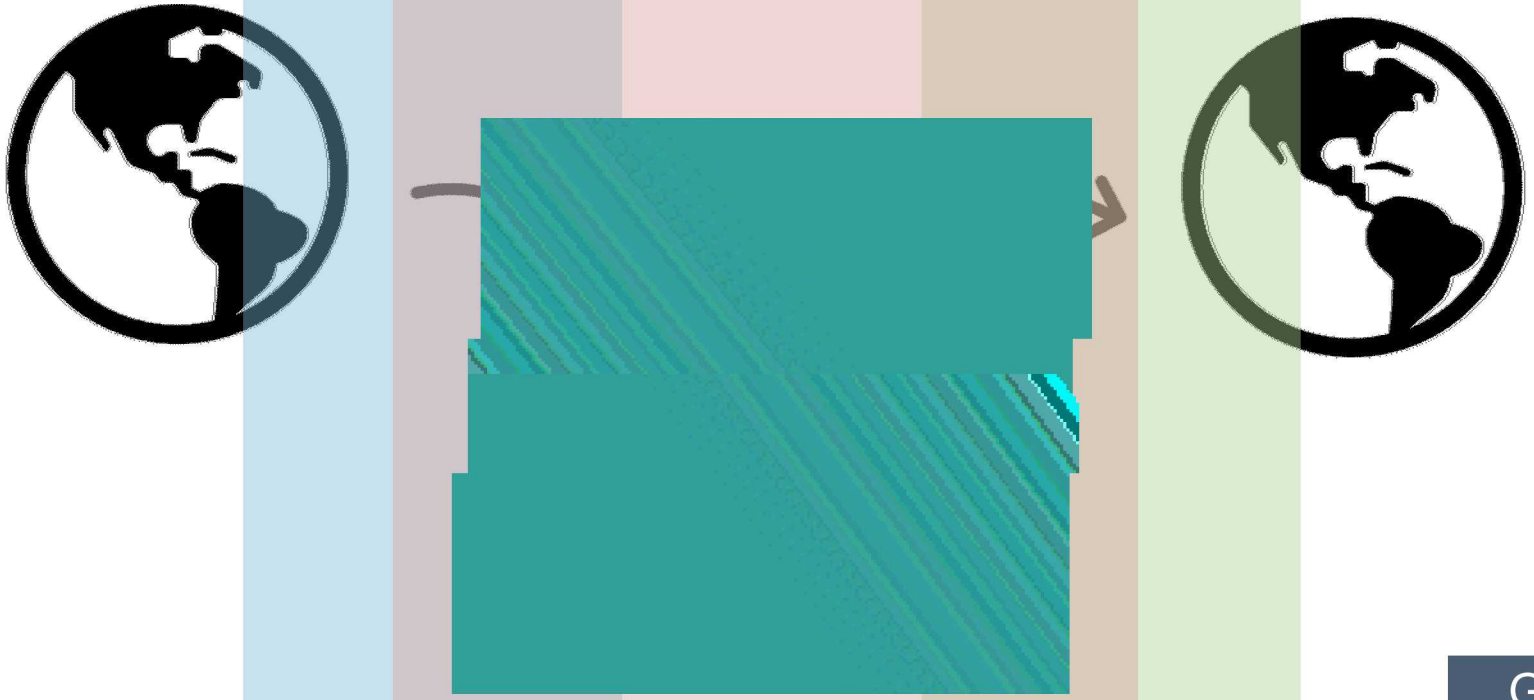
Learn by Doing: Case studies

HIDDEN TRIGGER DETECTION

NETWORK PROTOCOL EXTRACTION

INDICATORS OF COMPROMISE

End to End



Get the question right

Decompose and solve

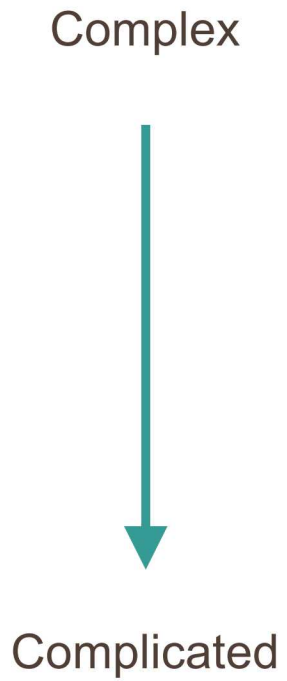
Make the answer count

Get the Question Right

- **A (snarky) Murray Gell-Mann description of the Richard Feynman "algorithm"**
  1. Write down the problem
  2. Think hard
  3. Write down the solution
- **Corollary:**
  - Be able to state the problem before you attempt to solve it
- **We partner with human factors experts to:**
  - Identify meaningful government mission problems that we have some hope of addressing

Getting the Question Right – Indicators of Compromise

- **Is my network safe?**
- **Are there adversaries operating on my network?**
- **Are any of these malicious tools being run on my network?**
  - Malicious tools identified out of band
- **What are the observable artifacts of this software running?**
  - Created files
  - Registry entries
  - Hostname lookups

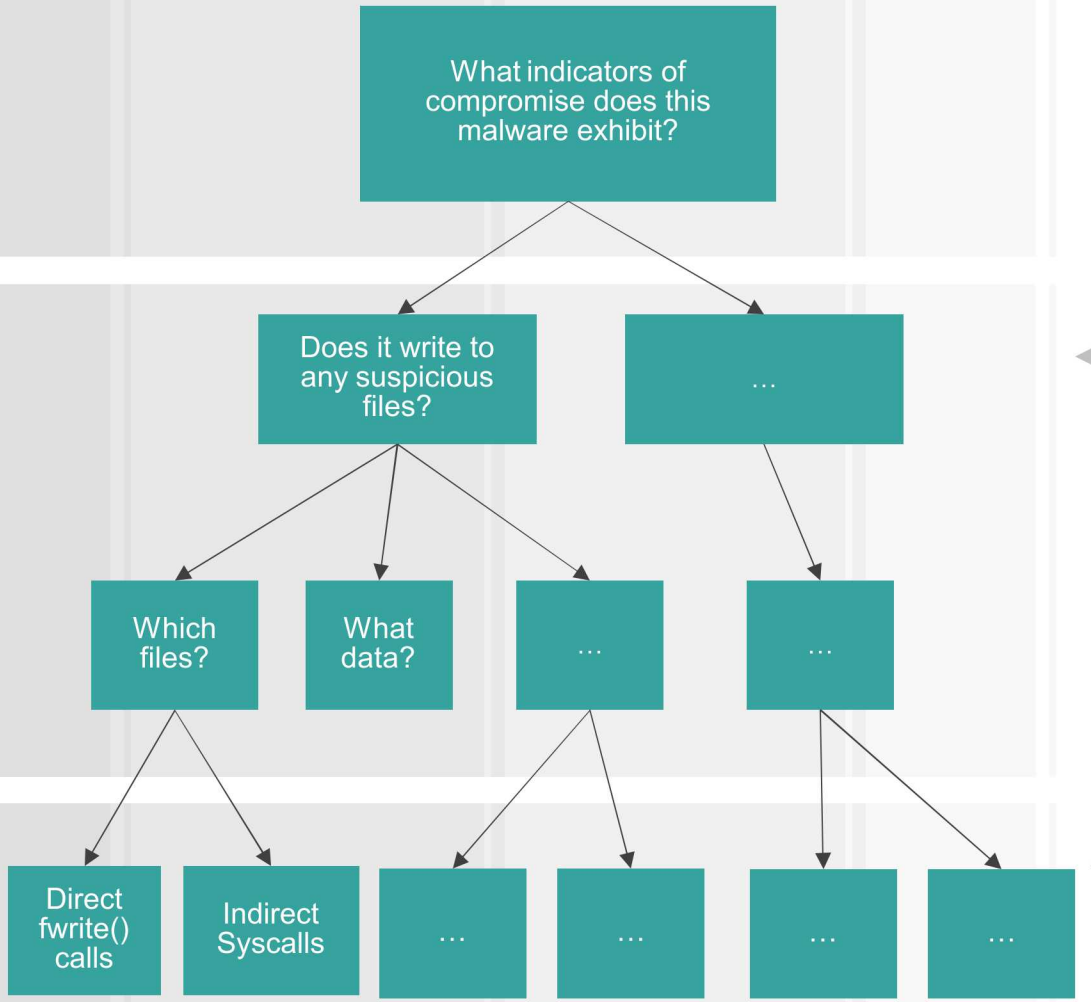


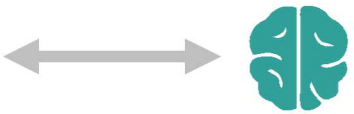
Decompose the Question

High level question

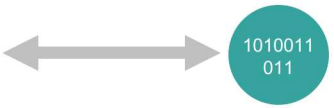
Reasoning about supporting evidence

Gather Properties



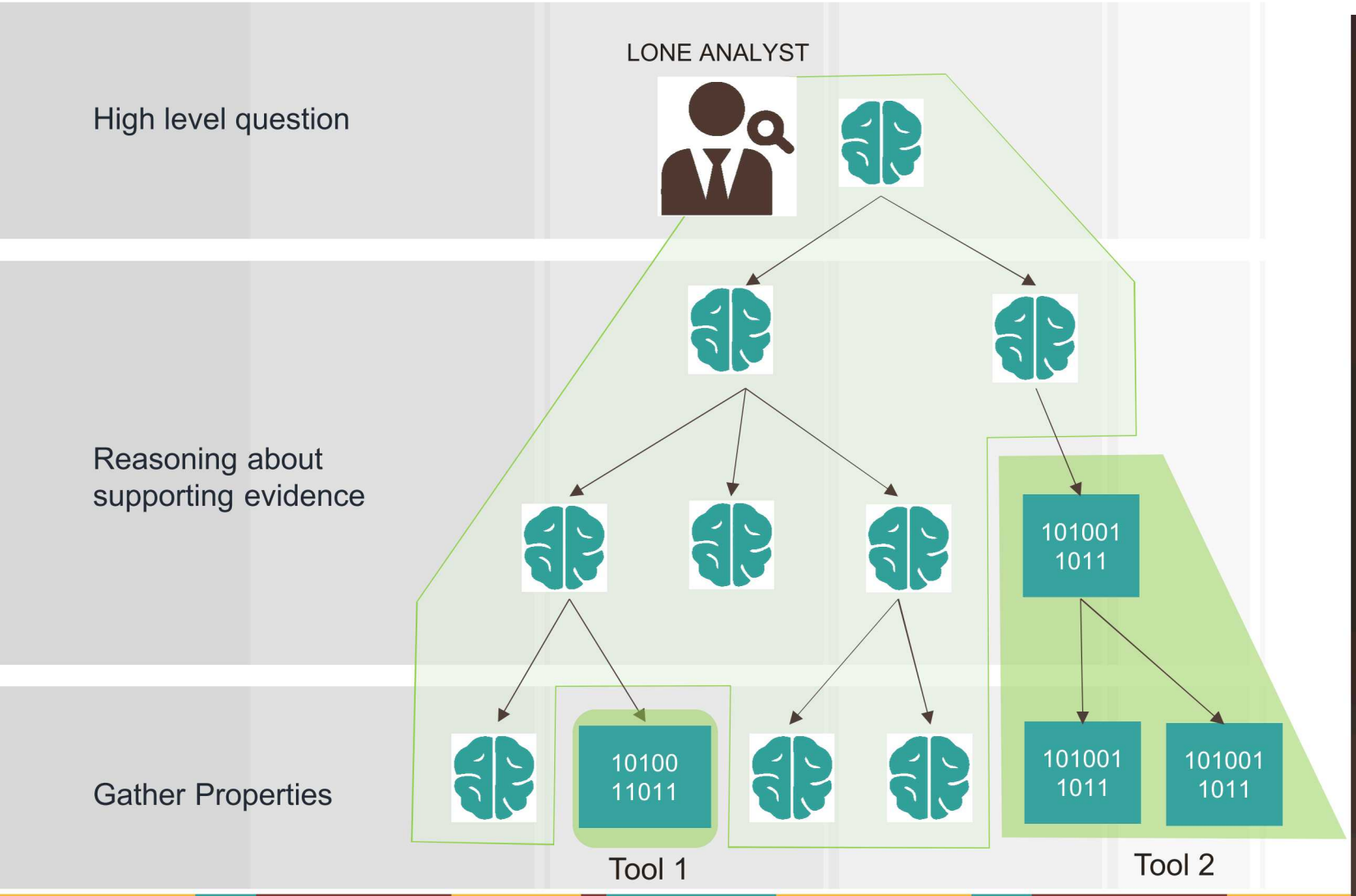


- Abstraction
- Inference
- Deduction



- Searching
- Pattern Matching
- Computation

Current Division of Labor



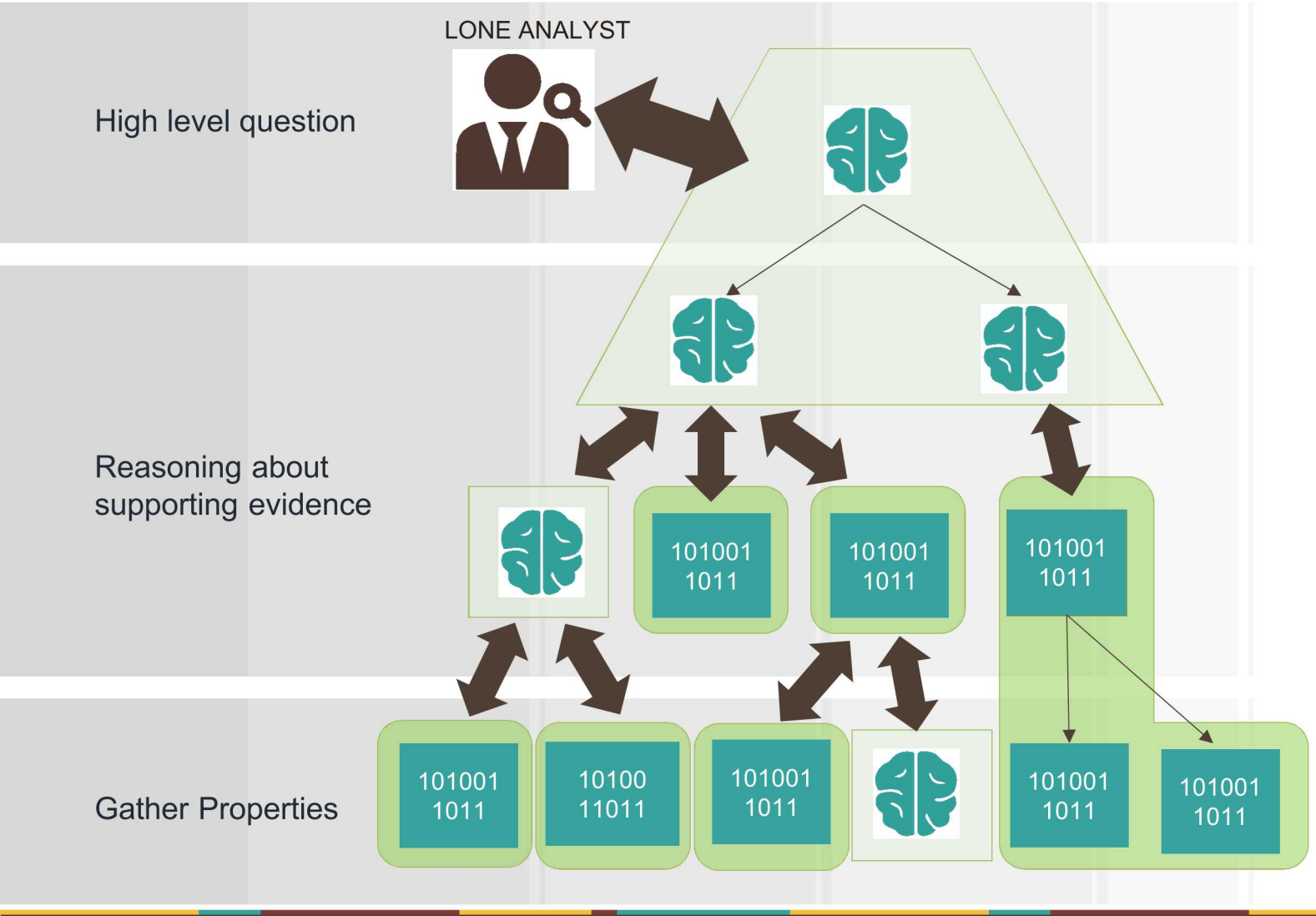
OBSERVATIONS:

- Analyst doing repetitive tasks
- Analyst solely responsible for mountains of data and reasoning
- Analyst experiences cognitive overload

PROBLEM:

Poor division of labor between the human and the tools

Incremental Improvement



Develop knowledge transfer interfaces:

- human⇔human
- tool⇔human
- tool⇔tool

Benefits:

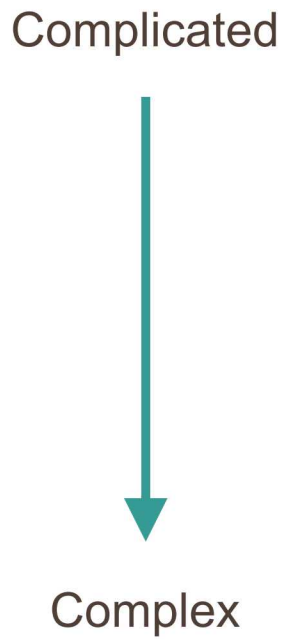
- Balanced division of labor
- Reusable analysis components
- Enable collaborative sensemaking

Decompose the Problem – Indicators of Compromise

- **Observable artifacts are the result of system calls**
- **What system calls are invoked by the program?**
  - Obfuscated dynamic library loading and function lookup
- **What are the arguments passed to identified system calls**
- **Often reduces to inferring the value of strings at system call sites**
- **Use an existing analysis – constant propagation**
  1. Lift binary using BAP
  2. Special purpose, flexible memory and pointer model
  3. Constant propagation abstract interpretation domain

Make the Answer Count

1. **Estimate the confidence we should have in answers we provide**
   - Verification and validation
   - Efforts to produce collections of ground truth binaries
     - Currently focusing on the lifting phase of the analysis
     - Correct disassembly, CFG recovery, and pointer analysis
2. **Get answers to the right place in the most useful form**
   - Sets of strings as signatures for NIDS and HIDS
   - Be explicit about uncertainty
   - Consider other places where constant propagation results are useful
     - Interactive reverse engineering and analysis tools



Final Thoughts

RAMSeS

- Static analysis
- Existing, third party binaries in support of program understanding
  - Focus on solving problems of practical, not just theoretical interest
  - Prefer to use or adapt what exists today

- What the plan *isn't*
  - Traditional program verification
  - Infrastructure for scaling analyses
  - One Tool to rule them all
  - A recipe for perfection (Progress not Perfection)
  - A quick fix

Questions?

ramses@sandia.gov
http://www.sandia.gov/ramses