

## Maturing Cyber Security Using BioThreat Experiences and Resources

Norman Lee Johnson  
Tim Williams  
15 Jun 2009

njohnson@referentia.com  
twilliams@referentia.com



Referentia Systems - Hawaii-based high Tech company, focused on network security and operations. <http://referentia.com>

I'm not an expert in CS, Tim is.

Background: 25 years at Los Alamos Ntl Lab, simulations of all kinds (fluids to epidemics) and risk assessment in bio for prioritization of threats and gap determination.

Current research interests is simulation of "infectious idea" spread on social networks using epidemiological sim models and characterizing cultural-social behavior of users/attackers/defenders from network activity.

## Goal: Provide a new viewpoint for maturing cybersecurity

### What was it like to live in London 200 years ago?

- How common was disease?
- Life expectancy? What changed?

### Background

- Related work: Adaptive Immunity

### Maturity of Cyber and Bio

### Similarities

- Function-Process
- System

### Maturing Cyber with Bio

### Specific Guidelines

### Specific Examples



The Goal - Because we are so much in the trenches in responding to cyberthreats, we need to have a higher perspective of where we are and where we are going.

The outline is to provide a quick view of where we are.

Then look at the relative maturity of bio and cyber from a couple of perspectives.

Then look at why bio has something to give cyber based on looking at the similarities between the two systems.

Then based on an analysis of gaps and current cyber resources and bio resources, provide guidelines and specific examples of how bio can help cyber, both for technologies and a roadmap of development.

the details will be familiar, but the viewpoint may be new.

Background and outline

**From Pres. Obama's introduction of the report:**

- "...cyberthreat is one of the most serious economic and national security challenges we face as a nation."
- "...not as prepared as we should be, as a government, or as a country."
- "... from a few keystrokes on a computer -- a weapon of mass disruption."

**Lead by Melissa Hathaway, Senior Advisor to the Director of National Intelligence (DNI) and Cyber Coordination Executive**

- Reviewed more than 250 executive orders, policies and advisory reports
- Held 40 meetings with stakeholders
- Reviewed more than 100 papers submitted to it
- "Dealing with security piecemeal by different sectors and stakeholders, and dealing with security as a stand-alone issue, has not provided a secure infrastructure."

**A commentary made the observation:**

- "...It's like we're playing football and our adversaries are playing soccer"

This event and document captures well the current state of cyber:

-we are vulnerable - event to a cyber Katrina

-We aren't doing what we already know we need to do

-The problem is both structural (infrastructure that was never designed for security) and broad in scope (all aspects of cyber and life)

-=> Remarkable what wasn't said: no looking to the horizon. This is generally true in all policy and analysis documents.



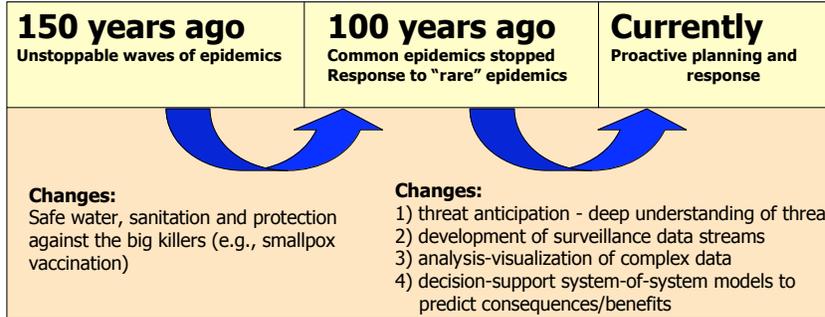
**Frequency and types of events**

**Depth and breadth of response to events**

Two ways to look at relative maturation

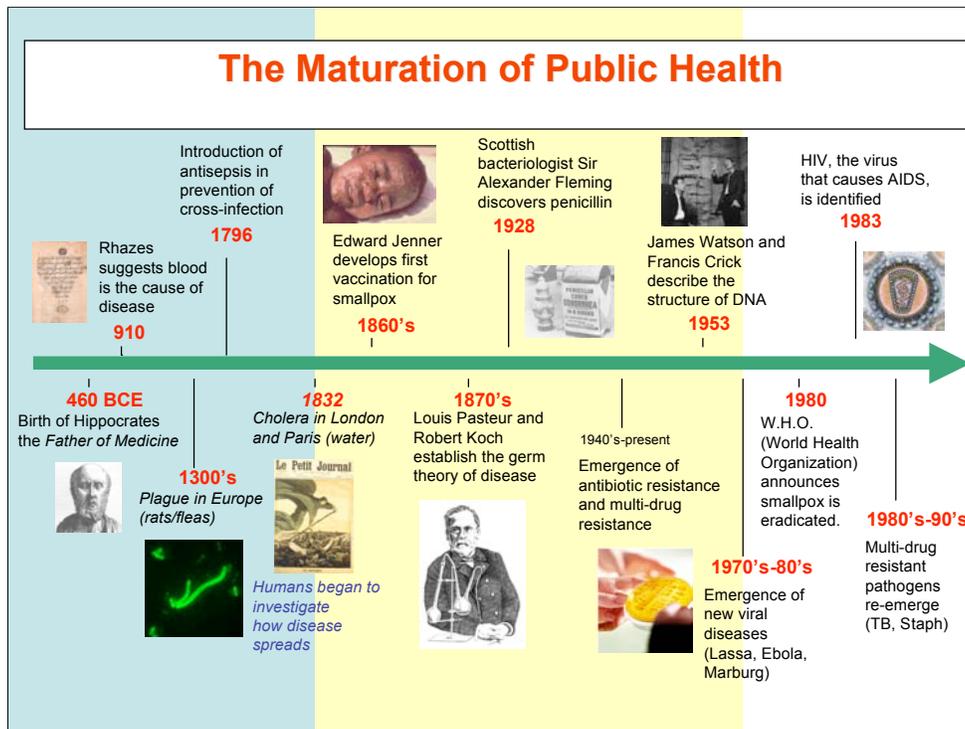


How Public Health was changed over 150 years....



What made the difference: managing public health at all levels

Cyber is in a transition from stage 1 to 2 (as for bio in 150 years to 100 years ago), but the long term solutions are in stage 3



Here is a more detailed view of a time line for bio.

What made the greatest change from stage 1 to 2 was understanding the threat and then creating a public health infrastructure that supported greater health.

The final stage started when we went from data poor to data rich - discovery of DNA and sequencing - enabled use to connect the sources to the consequences in full complexity.

We note that bio still has major problems as with HIV and multi-resistant pathogens from excessive use of antibiotics and now antivirals.

**Cyber protection: Policy scale**

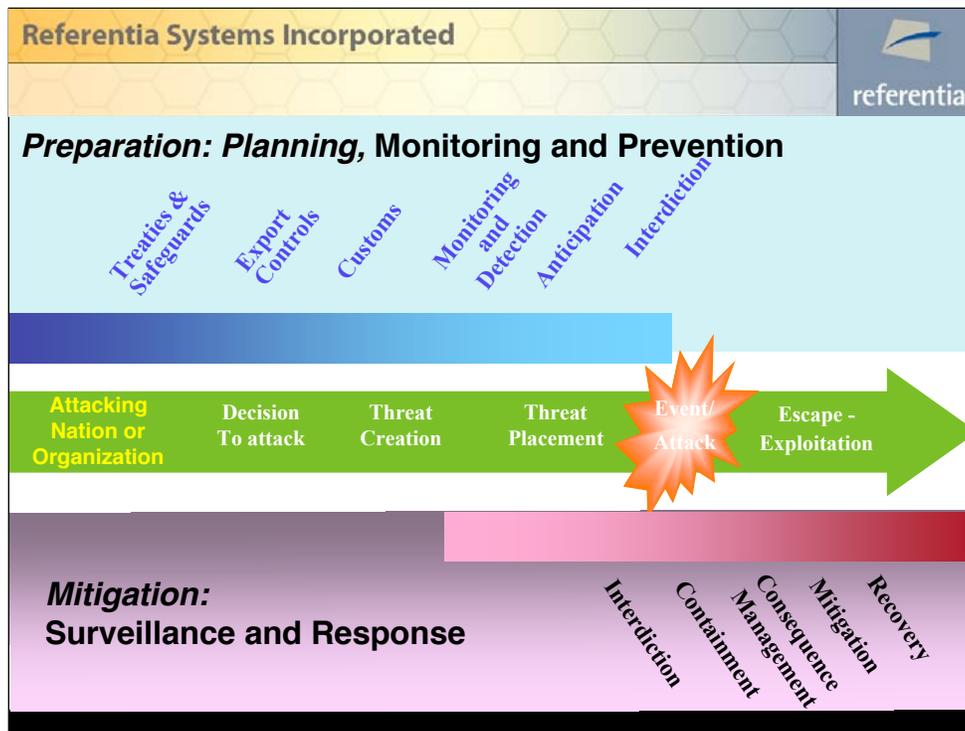
*This is what attackers do:*

**Attacking  
Nation/  
Organization/  
Individual**



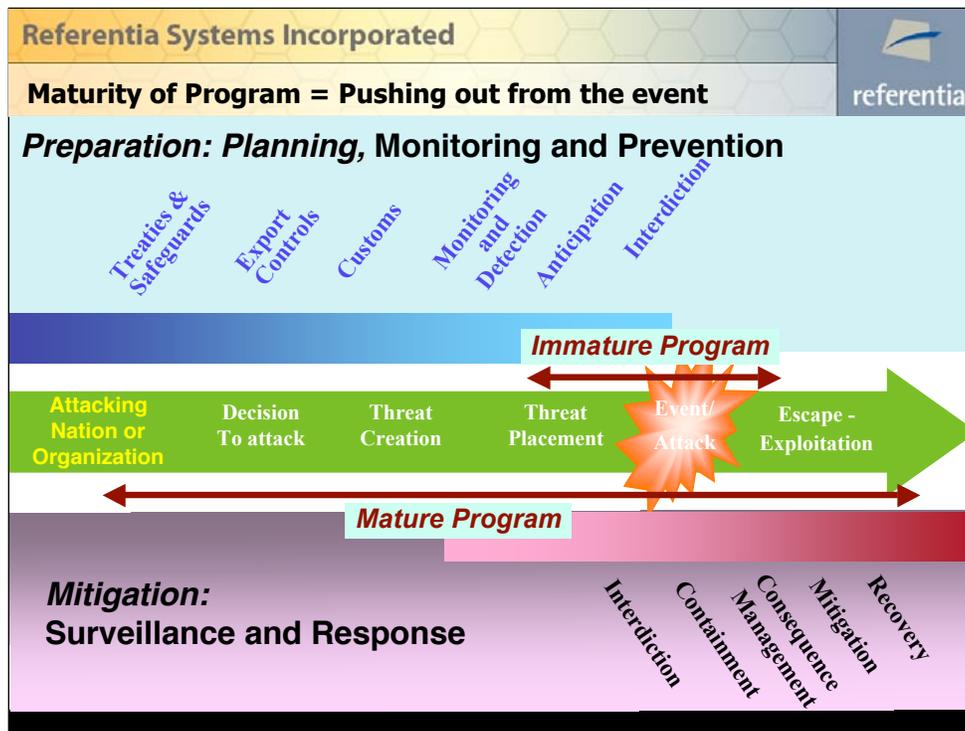
*How do we operationally respond?*

This is the same for all premeditated attackers. (excluding the emerging ones from nature)



Major points:

- The more mature your response to the threats, the more you push capabilities away from the event.
- sponsors are organized by each of these areas. Few span all areas.
- There are common technologies that support all aspects of this figure.
- In bio, Cost of programs that cover the spectrum of preparation and response is \$5-10 billion a year (excluding normal public health care).  
for cyber it probably is less than a 10th of this even including cyber “health care”



Major points:

- Most other threat areas (CBRNE and personnel) are mature. Cyber is the new one on the block and still figuring out the world.
- Why can't cyber learn from the other threat areas?



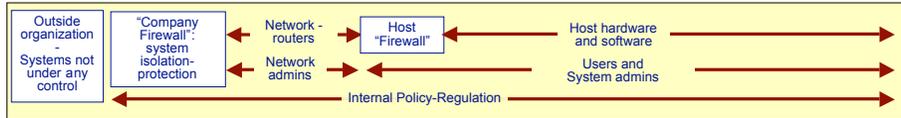
**Function-Process Similarities**

- The threat-host lifecycle (the infection process)

Now we established that cyber is less mature than bio,  
Now determine how bio can help cyber by examining similarities.  
Start with Function-Process similarities.

## The Lifecycle of a Threat in a Host System

Threats require a host or host systems - within which they attack, enter, exist, manipulate, steal resources, and evade. The life of a threat is a "threat lifecycle"



Threat Life-Cycle	Enter network	Evade detection	Move to host	Attack or Collect data	Replicate	Spread to other hosts	Exit or communicate outside	Repeat Cycle
Defender Actions	Protect from entry	Detect entry	Detect - Stop move	Detect - stop attack	Detect - stop replication	Detect - stop spread	Detect and/or deter communication	Assess damage, locate source, etc ...

### Examples of threat lifecycles:

**Viral threat:**

**Denial of service:**

**DNS/BGP spoofing:**

System of systems: a simple diagram that captures the main components of a typical system of networked hosts, the humans managing/using those systems and the procedures/etc. that influence the use of those systems.

See the paper for a longer description of the threat life-cycle: all the likely steps that occur when a threat attacks a host system. The life cycle is different for each treat (as in the examples below) and different for each type of host and is even different for different members of hosts within a type of host (all Macs or All Dell PCs).

Note that this figure can be evolved more by looking at the different signals occur at each step in the life-cycle. Etc.

Examples of lifecycles for typical threats:

Denial of service: this threat never initially enters the host, but could generate a state where entry is enabled - so this threat doesn't have much of a lifecycle [in biology, a chemical that stimulates a challenging host response would be the equivalent threat: it could weaken the immune systems such that entry and viability of a threat is then possible.]

Viral threat: This type of threats captures the full spectrum of the lifecycle above: from entry into the network - possibly in an email, entry into a host, activity of the virus, prorogation of the virus using the resources of the host - storage, emails, coms, etc.

Defender actions: these are the activities of the host or host system to protect from or in response to the threat.



## Function-Process Similarities

- The host system immune response options
  - Host immune state determines susceptibility
  - Host defense options are very similar - Layered defense systems :
    - Cell wall - firewall, with preferential transport
    - Innate immune response - always active
    - Adaptive immune response - takes time to work the first time
    - System isolation
    - Death of host

Immune state: set by prior infection, immunization, current infections, general health - and it determines susceptibility of the host.

The host defense options are essentially identical between cyber and bio.



### **System Similarities**

- Direct Consequences
- Secondary and indirect consequences

Direct: how that degraded performance of the host affects systems - directly

Indirect: how the absence of many hosts affect the rest of the system - via interdependence

See the paper for a detailed discussion of this similarity, and references on available bio resources to help here.



## **Develop programs that extend out from the event**

### **Similar challenges require similar solutions**

- Inherent chaotic nature of systems require a data-driven approach

### **From a Analysis of Cyber Gaps and Bio Opportunities**

- Data stream development
- Surveillance and situational awareness
- Analysis and visualization
- Decision support resources
  - Predictive/forecasting simulations
  - Consequence-benefit analysis resources
  - Resources to integrate all of the above

Programs: as before in the time line of attacker and defender programs - the bio experience provide a clear roadmap of the programs which much be implemented and funded.

Similar challenges lead to similar solutions: Common theme is that both cyber and bio require a data-driven approach.

Chaos comes from inherent nature of early dynamics - random hosts cause a different spread pattern that is not predictable (if you "played the tape again", it would look very different at early times)

Chaos from dynamics of attacker-defender innovation: because the attacker is generally after a goal, many possible attack scenarios are not use - only the ones perceived to achieve the goal. This often leads to probing the "don't know what I don't know" area of threats - which is very unpredictable and therefore leads to chaotic dynamics.

Approach taken: From how developing a mature planning/response prog. (as above) => currently available, gaps, bio opportunities.

\*\*\* the ideal way to determine gap is a risk assessment with a cost-benefit analysis of options. Cyber is no where near achieving this is a transparent and objective way.

In the paper we look in specific mission areas: cyber status, gaps, and bio resources that fill the gaps.

Here we take a different presentation approach that leads to the same solution, but is easier to present.

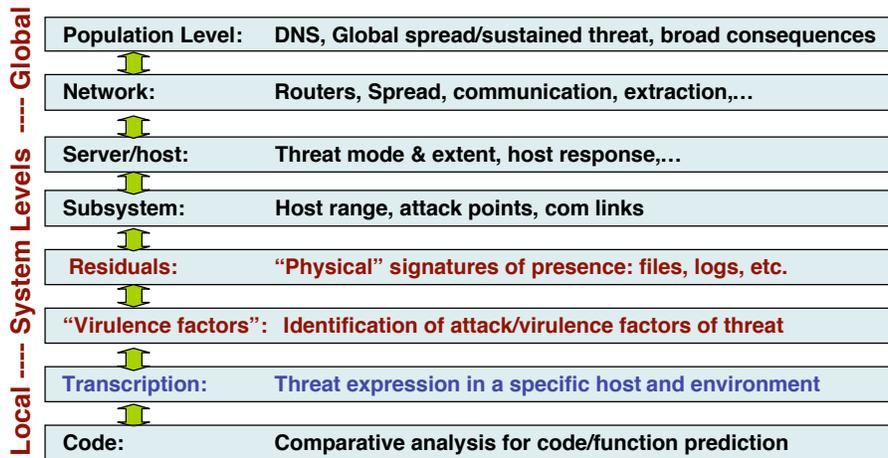
Referentia Systems Incorporated			
Analysis of Requirements, Gaps and Resources			referentia
Cyber Resources Required	Existing Cyber Resources	Cyber Gaps: Needed Resources	Enabling Bio-Resources
<b>Diverse cyber data:</b> providing historical and real-time data of current network topology and traffic; enclave, component and user activity, access, status	<b>Rich and more in development</b> - Network flow traffic types/volume; component types & programs used	Status of components: susceptibility, symptoms of attack, readiness, activity, threat level	Genome" threat data bases, "virulence" databases, current threats, current news
<b>Analysis and visualization of complex data streams:</b> past and situational health, attacks, losses; global-to-local drill down, weak-signal precursors, threat ID and attribution, intuitive analysis of large data sets	<b>In development</b> - Large data set analysis identifying trends and precursors, anomalous behavior, ideally automated	Health of network and components, direct and inferred attack status, syndromic precursors to attack ID, forensics, threat attribution, ...	Threat phylogeny, syndromic surveillance, health metrics, virulence change ID, forensic tools, responsiveness status, visualization resources
<b>Predictive models of future state/losses from an attack</b> given historical and current state, with transparency of outcome-to-cause and uncertainty quantification	<b>Scarce</b> - mostly academic simulations of network activity for limited threats; no exhaustive studies of tipping points	Databases of threats, standard threat models, emerging threat theory, effectiveness of response options	Epidemiological simulation resources, studies of mitigation options, coupled infrastructure sims, cost estimates,
<b>Consequence - benefit resources</b> including risk assessment, management and communication, expert-stakeholder conflict resolution, mission continuity	<b>Very limited for real-time response; limited for planning; limited fundamental understanding</b>	Metrics for mission readiness, threat-vulnerability mapping, integration of simulations	Standard threat scenarios for uniform preparedness, advanced risk assessment, adversary models,
<b>Decision-support integration of above for planning and response:</b> quantitative and transparent assessment of options, local-to-global cost-readiness tradeoffs, acquisition guidance, etc.	<b>Very limited</b> - currently wet-ware (human) based, no policy-level guidance on infrastructure acquisition, no operations support tools	Cost-benefit analysis of "what if" scenarios and response options; Risk management and communication	Threat anticipation-prediction, risk-based training, multi-stakeholder net-assessment studies, acquisition tools

Discussed in detail in the paper, but here's a different presentation that leads to the same conclusions

Note that possibly unfamiliar bio words are defined in the paper.

## A Multi-Level Threat View of Cyber Security/Defense

View the system as **signatures/activities/processes at different levels** - from small & localized to large & system-wide.

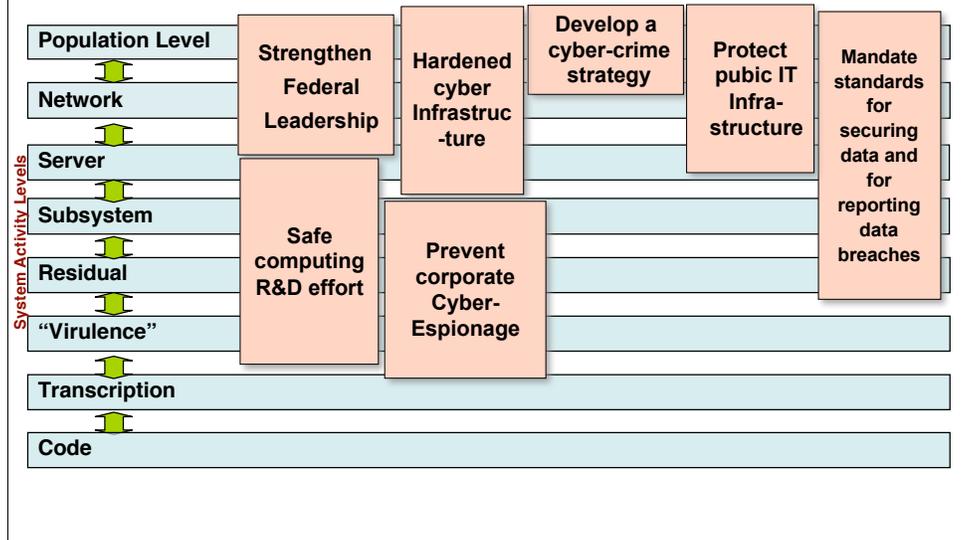


This is directly from a systems biology approach to threat-host systems - as viewed at different levels (in black) and signatures (in red) and processes (in blue). Below defines the possibly unfamiliar phrases.

- Code (analysis): Examining the smallest functional units of code in a threat and analyzing how they can behave and function - (qualified because code may not be used, e.g., be dormant, or the expression may depend on the environment, e.g., the server or host). [Genomics/proteomics: how information is coded at the smallest level - knowing these units are present provides a way of categorizing threats - e.g., phylogenetic trees - and captures the possible threat and function of the threat package]
- Transcription: Repeat as above for Code Analysis but done within the context of a host, thereby including how the environment affects the code expression
- “Virulence factors”: examines how certain codes are more destructive than others and looks at specific functional units that enable the heightened destructiveness. [virulence factors] In principle, one can estimate the “potential” destructiveness of a threat by the presence of certain coding. It is only the potential because the coding may not be called or may not have the right environment to work [ same as a genotype-phenotype in biology]
- Residual (analysis): examines the small signals/signatures of presence of a threat - based on: [Trace analysis used in Metabolomics is the “systematic study of the unique chemical fingerprints that specific cellular processes leave behind” - specifically, the study of their small-molecule metabolite profiles.]
- Subsystem analysis: This is familiar signatures/activities/processes within components of a self-contained host system - a server [comparable to a cell within a multicellular organism]. Because threats often use components of the host to hide and propagate [e.g., HIV]
- Server/host (analysis): This is the familiar host-level signatures/activities/processes. Many of the aspects are similar to the subsystem analysis, with the main difference being that a host/server is a more independent entity and therefore shows broader functions and more complexity, where the sub-system analysis might be limited to a very specialized function - say mail attachments or video chips/programs.
- Network (analysis): This is a familiar network level - the signatures, activities and processes that happen at a router level - all about communication between hosts. But because routers are computers themselves, there are aspects here that resemble some sub-system host analysis.
- The population level (analysis): this captures a collection of hosts within an enclave to the entire network and components. There is a great diversity of host and com networks that get introduced at a population level that results in places where threats can hide, steal resources, attack at different channels (ports), etc. There is also the spread of a threat at a population level that determines the ultimate impact of a threat as well as the time response required to protect from a new threat.

Example using this Landscape to understand Programs:

### White House program in cyber security Policy Initiatives tend to populate the top levels



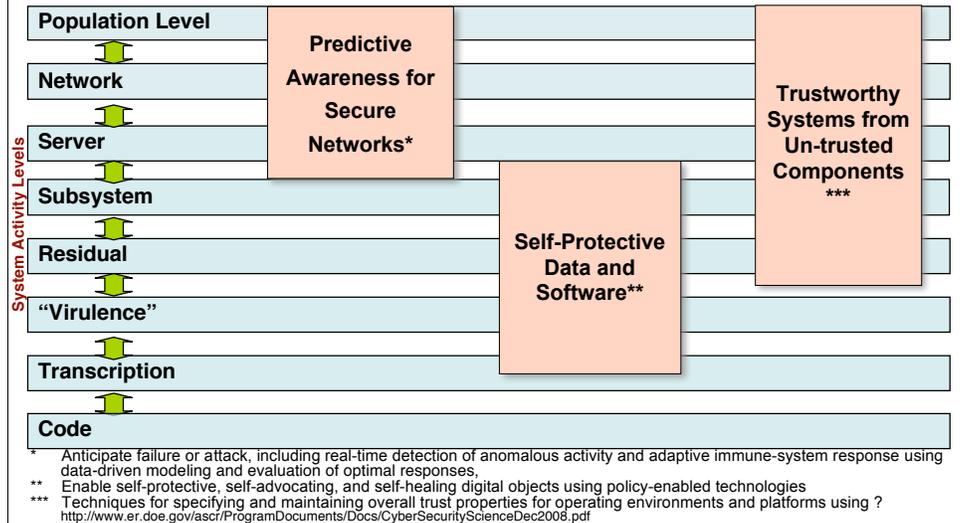
To show how this system landscape can be used, we apply it to different studies that have been released.

Here is a policy example: These areas were posted on the White House Web site under homeland security the Wed after Obama took office:

- **Strengthen Federal Leadership on Cyber Security:** Declare the cyber infrastructure a strategic asset and establish the position of national cyber advisor who will report directly to the president and will be responsible for coordinating federal agency efforts and development of national cyber policy.
- **Initiate a Safe Computing R&D Effort and Harden our Nation's Cyber Infrastructure:** Support an initiative to develop next-generation secure computers and networking for national security applications. Work with industry and academia to develop and deploy a new generation of secure hardware and software for our critical cyber infrastructure.
- **Protect the IT Infrastructure That Keeps America's Economy Safe:** Work with the private sector to establish tough new standards for cyber security and physical resilience.
- **Prevent Corporate Cyber-Espionage:** Work with industry to develop the systems necessary to protect our nation's trade secrets and our research and development. Innovations in software, engineering, pharmaceuticals and other fields are being stolen online from U.S. businesses at an alarming rate.
- **Develop a Cyber Crime Strategy to Minimize the Opportunities for Criminal Profit:** Shut down the mechanisms used to transmit criminal profits by shutting down untraceable Internet payment schemes. Initiate a grant and training program to provide federal, state, and local law enforcement agencies the tools they need to detect and prosecute cyber crime.
- **Mandate Standards for Securing Personal Data and Require Companies to Disclose Personal Information Data Breaches:** Partner with industry and our citizens to secure personal data stored on government and private systems. Institute a common standard for securing such data across industries and protect the rights of individuals in the information age.

Example using this Landscape to understand Programs:

**DOE's Report on Scientific R&D for CyberSecurity Dec 2008**



Here is an example from a technical perspective. We note that this was one of a few studies that recommended the development of predictive or forecasting resources, a common tool in the bio area but noticeable missing in cyber. There is also an emphasis on bottom up solutions, reflecting the distributed and decentralized nature of the system being addressed.

Three focus areas from the report:

\* Mathematics: Predictive Awareness for Secure Systems.

Goal: Provide capabilities to examine system or network behavior to anticipate failure or attack, including real-time detection of anomalous activity and adaptive immune-system response.

Research: Develop mathematical modeling techniques for complex information applications and systems, enabling data-driven modeling, analysis, and simulation of architectures, techniques, and optimal response to threats, failures, and attacks.

\*\* Information: Self-Protective Data and Software.

Goal: Create active data systems and protocols to enable self-protective, self-advocating, and self-healing digital objects.

Research: Develop techniques and protocols to provide data provenance; information integrity; awareness of attributes such as source, modification, trace back, and actors; and mechanisms to enforce policy concerning data confidentiality and access.

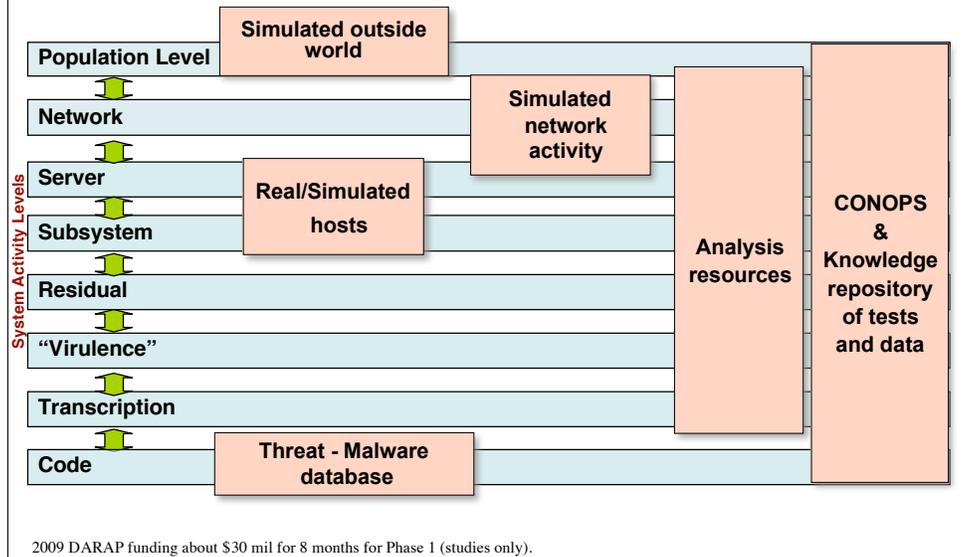
\*\*\*Platforms: Trustworthy Systems from Un-trusted Components.

Goal: Develop techniques for specifying and maintaining overall trust properties for operating environments and platforms.

Research: Develop approaches for quantifying and bounding security and protection, integrity, confidentiality, and access in the context of a system comprising individual components for which there are varying degrees of trust.

Example using this Landscape to understand Programs:

### DARPA's program in *National Cyber Range (NCR) Testbed*



Here is an example of a operational and military program (not recommendations).

We note that while some of the boxes span the range, there are large holes in the resources to really understand the threat-host dynamics, either at a component or system level. Hence, this is a resource strictly for testing resources, not predicting how they will function in other testbeds or in real systems.

Description for NCR from DARPA BAA: To achieve these goals the NCR will provide, as a minimum, the following objectives:

All necessary resources including but not limited to test facilities, utilities (power, water, etc), physical security, and heating, ventilation and air conditioning (HVAC).

All personnel necessary to design, operate, and maintain range, to include but not limited to management, administration, system administration and engineering personnel.

All necessary administration to include necessary certification/accreditation,

Concept of Operation (CONOP) development, security management, test scheduling, and processes.

The ability to replicate large-scale military and government network enclaves.

The ability to replicate commercial and tactical wireless and control systems.

The ability to connect to distributed, custom facilities and/or capabilities as necessary to incorporate specialized capabilities, effects, or infrastructures. Interactive test suites to design, configure, monitor, analyze, and release tests.

A robust range management suite.

A large pool of heterogeneous systems (nodes) as well as the ability to rapidly integrate new nodes.

The ability to rapidly generate and integrate replications of new machines.

The ability to integrate new research protocols. A test toolkit/repository for reuse of recipes and architectures.

Forensic quality data collection, analysis, and presentation. Realistically replicate human behavior and frailties.

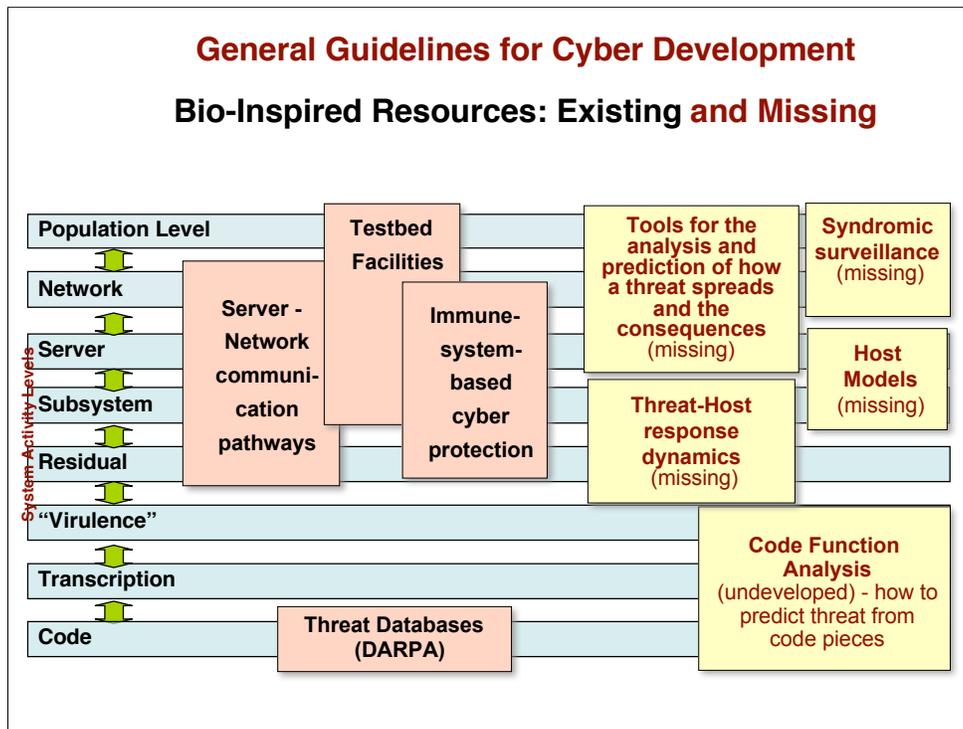
Realistic, sophisticated, nation-state quality offensive and defensive opposition forces. Dedicated on-site, support for installation, troubleshooting, and testing.

The ability to accelerate and decelerate relative test time.

The ability to encapsulate and isolate tests, data storage, and networks.

A knowledge management repository for test case samples and past experiences that can be used for future endeavors.

A malware repository.



Here is a summary of the previous slides and other programs not presented: What is observed is that there is a strong effort in characterization of the system components and developing resources for exploring the dynamics of the system at a high level. But there are few resources or plans that try to connect the lower levels with the upper levels - as being done in bio to develop a predictive capability. For example, if I observe a new threat (or a variation of an old one) for cyber, do I have any resources that will help me prediction how it will spread, what damage it will do and what my mitigation options are? And do I have the resources to discover new threats that I don't know yet (as in world wide coordinated bio surveillance)? The answer is no, no, no. And no. The boxes on the right in yellow are the resources that could be developed now using bio technologies which would begin to address these issues.

**Tools for analysis and prediction:** these are resources that connect the lower levels with the higher levels, coupled with consequences on other systems (e.g, economy, power generation). In the bio world these are either simulations, such as simulating an epidemic, or risk assessment resources.

**Syndromic surveillance:** These are used in the bio work to capture a threat based on the symptoms rather than the direct presence (code) of the threat. Because doctors (and cyber security) tend to look for what they already know, this is the only way in many scenarios to capture novel threats: by observing repeated patterns of symptoms in hosts.

**Host Models** - are "models" or useable descriptions or working systems that represent a general type of host that could be attacked. For example, an older PC with Windows XP of some version that connects to the internet thru irregular ethernet connection, surfing, network gaming, Outlook email and uses Microsoft Office. Some of these do exist.

**Threat-Host response dynamics:** these resources, when combined to host models, predict the host behavior from the lowest levels of the system (code or transcription). The equivalent in the bio world are the immune systems models for a variety of hosts and threats.

**Code function analysis:** how to predict the population level dynamics of a threat from its components



**Similar dynamic challenges require similar solutions**

- Inherent chaotic nature of systems require a data-driven approach

**Develop programs that extend out from the event**

**From a Cyber Gap Analysis**

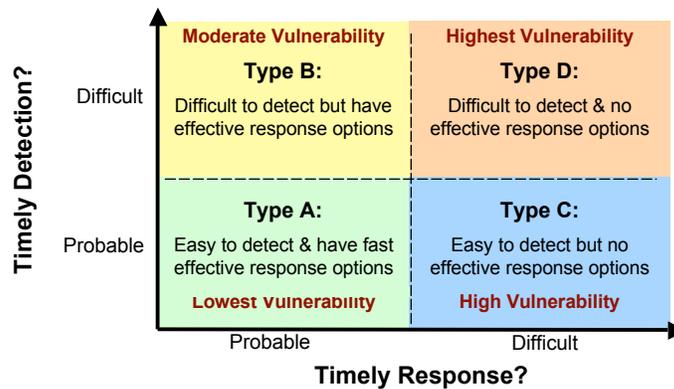
- Threat anticipation
- Surveillance and situational awareness
- Analysis and visualization
- Decision support systems-of-systems resources

**Two Specific Examples**

- Addressing the complexity of threat categorization
- Graded response to limit “regret” or degrade system performance

## Cyber Threat Types Are Complex

This Threat Chart is a way to simplify the complex landscape of threats



This approach to threat categorization is taken directly from a study of building protection from Chem-Bio (CB) threats for the National Academies and was originally developed by Norman Johnson to address the complex CB threat landscape. See: <http://books.nap.edu/openbook.php?isbn=0309109558>

Very powerful threat categorization because:

- Puts complex variety of threats in a comparable and understandable basis
- Links measurable attributes (timely detection and response) to outcome: vulnerability and consequences
- Points to where the biggest challenges occur: difficult detection and response

We need a similar threat landscape for cyber, if for nothing else, to simplify the communication of the cyber threat to less experienced stakeholders

In the paper we discuss how this figure can be extended in a third dimension to account for difference in response.

This landscape of threats could be extended to a third dimension to include consequences of response options (high/low) - including levels of regret - because threats that have timely detection and response options could differ greatly by the consequences of the response (e.g., continue normal operations or suspend all operations).

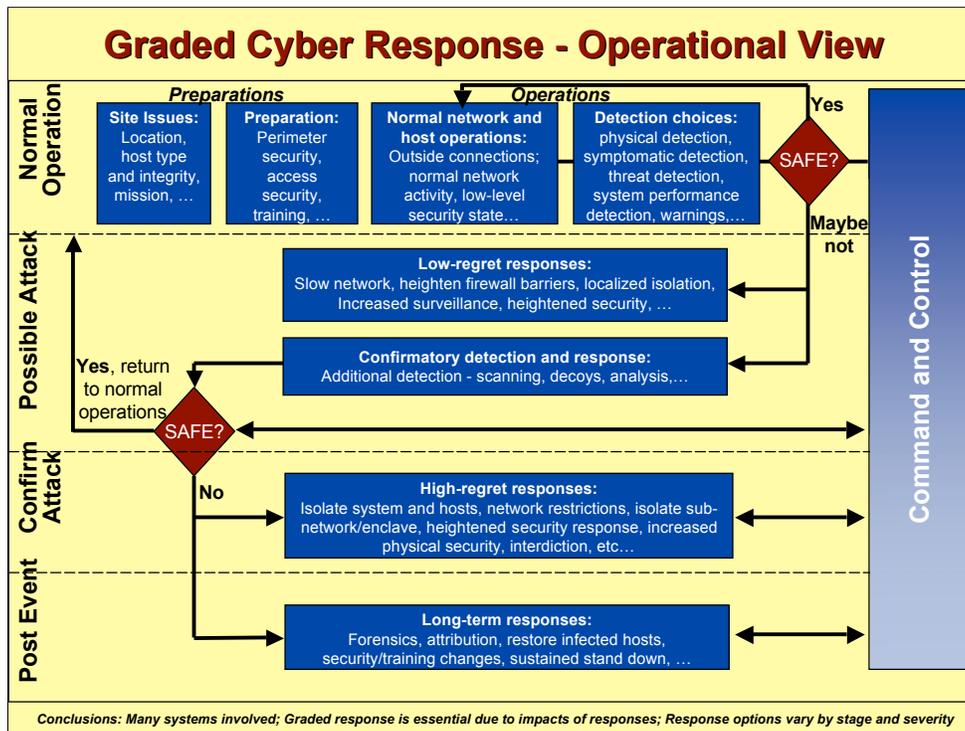
Full Reference: The Report was prepared by the National Research Council by the Committee on "Protecting Occupants of DOD Buildings

from Chemical and Biological Release". The report was sponsored by Defense Threat Reduction Agency. For more information,

contact the Board on Chemical Sciences and Technology at (202) 334-2156 or visit <http://nationalacademies.org/bcst>. Copies of

Protecting Building Occupants and Operations from Biological and Chemical Airborne Threats: A Framework for Decision Making

are available from the National Academies Press, 500 Fifth Street, NW, Washington, D.C. 20001; (800) 624-6242; [www.nap.edu](http://www.nap.edu).



Focus on the response-mitigation part of the previous figure.

Because computers and networks are now used for “normal” operations, this sets the context for operational response: a tiered response is required to achieve an appropriate level of “regret” in the response (regret being the negative impact to normal operations or “health” of the system). Because high regret responses cause disruption of normal possibly essential activities, a tiered response is essential to not cause more damage than the threat.

Tiered systems also match resources to the threat level. For example, the continual use of human analysis during normal operations is prohibitive. But in a tiered response they can be engaged as needed, In the bio world, certain surveillance options are very expensive even for high value assets.



## Summary of Using Bio to Mature Cyber

**Current policy and resource development are aligned with immediate needs, but policy lacks over-the-horizon thinking**

**Use the bio-threat programs as template and justification for the growth of federal programs and international engagement**

**Use the analysis herein to transfer specific technologies from bio domain**

**Define research areas from bio-domain lessons**

**What is a common unmet challenge to both?**

Characterization and prediction of the response of users/attacker/defenders accounting for behavioral, social and cultural differences.



Are we planning too much?

mattbuck.com



GeNeRAL-SHALL We ATtACK WHILe  
THEY'RE DOING THEIR ONGOING  
VULNERABILITY ASSESSMENTS—  
OR, WAIT UNTIL LATER?

Are we too little - too late?

