INTERNET INFRASTRUCTURE HEALTH METRICS FRAMEWORK (IIHMF)

A set of models and metrics to measure the public health of Internet infrastructure.

Prepared by the CyberGreen Institute
May 2021

CyberGreen

TABLE OF CONTENTS

1. INTRODUCTION

The Internet Infrastructure Health Metrics Framework (IIHMF) is a set of models and metrics to measure the "public health" of Internet infrastructure. This report explains what that means, how we do it, and our progress towards the goal.

As digital societies continue to evolve, digital economies must increasingly depend on resilient, trustworthy, and safe Internet infrastructure.

Cybersecurity concerns have, for many years, not only been discussed in technical and private sector circles but have also become a top priority in nation state intergovernmental agencies and even broader diplomatic discussions. Various governments have created cybersecurity agencies at strategic levels and forged strategic private-public partnerships.

The increasing frequency and scale of DDoS attacks[1] has translated to greater economic loss and has exposed vulnerabilities in critical infrastructure. For example, in 2012, six U.S. banks were the targets of sustained, complex, multi-pronged DDoS attacks. Some estimates put economic loss resulting from IT services downtime in the range of $300,000-$1,000,000 per hour.[2] [3]

Many governments have mandates to provide for and to protect their citizens. Ensuring economic stability and the availability of critical services is part of this. As healthcare, transportation, financial services, utilities, educational institutions, emergency services and most of our societal needs embrace digital technologies, governments must develop appropriate policy frameworks to ensure that digital services are available, reliable and trustworthy. States need to maintain security and influence by navigating a geopolitical environment in which power is earned and exercised through digital capabilities in infrastructure development.

One of the more fundamental challenges of this work has been deciding what constitutes "Internet infrastructure" and deciding what should be measured (see recommendation E4) to assess both its health and public health elements related to it. We have held a workshop, have sought advice and opinions from experts in multiple fields, and have explored existing literature. This work is both nascent and in a relatively new area, and this first phase serves as the foundation for a framework that aims to be adaptable as cyber threats evolve.

OVERVIEW OF REPORT

For clarity, this report focuses on the current results and a set of recommendations more than a project history. It is organized in three main sections:

- Summary of recommendations and contributions

1 Nicholson 2020
2 Ibid
3 "Calculating the Cost of Downtime" n.d.

- The supporting analysis
- Technical analysis and details

The recommendations and contributions, listed at the start, are explained within sections 2 and 3. The supporting analysis starts with a discussion of Internet infrastructure (section 4). We use a specific definition, tuned to the needs of this project. From there, in section 5, we discuss both health models, and how they differ from enterprise security models. ("cyber public health complements enterprise risk management like public health complements medicine.")

Having discussed cyber public health models, we turn to classical risk management, and compare and contrast those models to a new set of "impact models" created for this work in section 6.

This sets us up to be able to talk, in section 7, about the indicators we need to measure in order to discuss the public health of Internet infrastructure, and outline a scorecard we can build.

A scorecard is a tool for driving improvements and, sometimes, simply shining a light on a problem is sufficient to drive action. There are, however, complex reasons that people do not take action. For example, rates of vaccination are dropping in the United States because of many factors, including grounded and fantastic concerns about a new vaccine, fear of needles, language barriers, and allergies. When there are analogous reasons for low scores, it will be important to understand them as part of improvement.

The last major section of the report (section 8) is a technical deep dive into the indicators. A set of conclusions rounds out the report, and is followed by references, two appendixes (1 and 2) with some work products we generated along the way that no longer fit the overall report, and Appendix 3 which focuses on mitigation recommendations. Finally, Appendix 4 shows our thought process in the selection of indicators to measure and some initial measurement characteristics which might be used to create a scorecard.

SCOPE AND GOALS

We focus on understanding risks to a nation's Internet infrastructure as a subset of the cybersecurity risk a nation state is subjecting itself to. This differs from common enterprise risk management that have been the focus of most cybersecurity efforts. Will emergency services be able to communicate during natural disasters? Will utilities be available if there is a cyber attack on critical water or waste management facilities or an electric plant? Will healthcare services have reliable information that has not been modified through data manipulation or deletion? How capable are the banks and lending institutions to ensure citizen and business data is kept confidential so only authorized individuals have access to financial data? Public safety and national security are now reliant on a healthy and trustworthy Internet infrastructure.

The IIHMF will allow states to measure their overall risk, understand how it changes over time, and compare to other states. It also enables us to measure the health of Internet infrastructure using metrics and a model based on public health.

Being able to measure Internet infrastructure with a health model is new, and the work has come with a series of challenges. We have made some very sweeping brush strokes in this project, knowing that some of them may require adjustment as we learn more. We do this intentionally.

2. SUMMARY OF RECOMMENDATIONS

We break our recommendations into two major groups: recommendations for policymakers, writ broadly, and recommendations for further research by ERIA. The recommendations are organized so that many build on previous recommendations. This results in a different order from their appearance in the report.

The following list summarizes actionable recommendations for policymakers:

- P1. Conduct a census of critical Internet infrastructure in your country.
  - a. Craft criteria for inclusion, perhaps in groupings such as core IP routing, essential DNS, communications services used by Internet operators.
  - b. Evaluate how much downtime is tolerable per service.
  - c. Craft accounting guidance for assessing the cost of a problem so that reported numbers have consistency.
  - d. Maintain the census by repeating it now and then.
  - e. Share the census.
- P2. Evaluate national standards for security advice and the consistency and character of that advice. (By character, we mean, is the advice goal-centered, such as "be resilient" or activity-centric such as "run disaster recovery exercises quarterly.")
- P3. Mandate that companies reporting privacy breaches include information about the controls in place and their efficacy relative to the breach.
- P4. Create standards for incident and near miss reporting and investigation which show which controls were in place, and which functioned as intended. Our ability to ensure that our measures tie to effective controls, and thus our ability to define measures which should tie to effective improvement, is limited by our lack of outcome data.

The following list summarizes actionable recommendations for ERIA to further research:

- E1. Invest in models and datasets that illuminate risk and connect it to indicators which can be studied from outside the system (that is, measured at internet scale, rather than enterprise or SMB scale).
- E2. Investigate reasons that organizations are not acting on security advice. Some advice is like "exercise and eat well," and other advice is like "do not pollute." Eating well means forgoing hamburgers for a distant payoff. Not polluting requires spending money to make the world better. Both behaviors are important to public health, and each requires different forms of encouragement.

E3. Develop a fuller model of the mapping between cybersecurity issues and public health issues. This might take the form of a taxonomic flowchart, which uses the characteristics of a computer security problem, a harm to the infrastructure, or a technology-enabled harm for categorization decisions.
E4. Refine definitions of critical Internet infrastructure in ways that enable public health measures.
E5. Create a formula for an internet infrastructure health scorecard, and engage with local and international civil society on its content and uses.
E6. Run a pilot to measure internet infrastructure health and engage with the questions raised by preliminary data collection, analysis and comparison.
E7. Create a set of evaluation criteria that allow this project and its successors to assess observed measures.

3. SUMMARY OF CONTRIBUTIONS

This section summarizes the outputs of this project, which has been large and complex, and has produced an unusual variety of sub-deliverables; work products as we aim to craft a new type of scorecard.

1. Sets of Internet health indicators with analysis of the indicators and characterization of the measurements (Appendix 4).
2. A promising analysis of "Roles and Lifestyle Choices in the Internet Infrastructure Ecosystem" (Appendix 1).
3. A set of measures that require further investment and research but that could inform future work and iterations of the Framework (Appendix 4, "Tier 2" section).
4. Lessons learned about how measurement can be applied in this space.
5. Lists of challenges and obstacles to measurement (many, but not all, reflected in the recommendations above)
6. Three new models that tie technical Internet measures to public health ("Impact Models")
7. Mitigation recommendations for the issues we are concerned about (Appendix 3).

4. INTERNET INFRASTRUCTURE

To measure the health of something, we need to be able to state what we are assessing. For this report, we considered a minimalist definition of internet infrastructure. We took this approach because this work has revealed the complexity of measuring broadly scoped infrastructure. We have chosen a smaller scope for Internet infrastructure than some other projects for several reasons, including a broader scope of indicators, data availability and to manage the complexity of interconnectivity.

Our work shows that some of the assets listed above are not easily measurable. There is no consensus on what constitutes "Internet infrastructure", and it can vary from country to country. If there were country-by-country surveys or censuses of what counts as critical Internet infrastructure then this project or similar ones could start from there and measure what they enumerate.

For example, according to the Internet Infrastructure Coalition, Internet Infrastructure is a collective term for all hardware and software systems that are "responsible for hosting, storing, processing, and serving the information that makes up websites, applications, and content."[4]

Our selection is based on a more specific list, taken from a 2015 ENISA report.[5] That report considers asset types as follows:

- Protocols (e.g. routing, security, application, essential addressing)
- Hardware (e.g. network devices and servers)
- Software (e.g. operating systems, device drivers, software)
- Information (credentials, operational information, system configurations)
- Services (routing, applications, security, essential addressing)
- Interconnection (Internet exchange points, generic Internet provider)
- Infrastructure (cablings, building, power supply, cooling systems, etc)
- Human resources (operators, developers, managers, etc)

COMPONENTS OF INTERNET INFRASTRUCTURE FOR THE IIHMF

In the context of being able to diagnose the health of Internet Infrastructure, we have classified six components based on a combination of underlying, fundamental technologies and services, some of which will evolve over time, as well as some critical, dependent components. These are:

Open Services: These are network services that can be used for Distributed Denial of Service (DDoS) amplification. These are attacks where a large quantity of traffic is created which causes disruption of service or renders a service unavailable.

Routing: Devices on the Internet must be able to determine the path to take from a sender of information to the recipient. This is achieved through a means called 'routing'. Routing information must be reliable, available and trusted.

Domain Name System: The domain name system is a globally distributed, loosely coherent dynamic database of information. It maps names to IP addresses and is also used for other types of information dissemination. It is a fundamental service that must be reliable, available and trusted.

4 "What Is the Internet's Infrastructure? [Video]" 2019
5 Lévy-Bencheton et al. 2015

Email: All organizations use email to conduct day-to-day business and they rely on its availability and reliability. Email is used to coordinate responses to problems, to distribute information, and may be a critical part of responding to problems. We consider email to be internet critical infrastructure.

Certificates: A source of trust in online communications is asserting one's identity through some use of credentials. When we scoped this broadly to credentials, it included digital certificates along with passwords, hardware or software tokens and other means for asserting one's identity, but those were complex to measure. For example, trying to assess if leaked passwords can be used to login to a service may appear to be a hacking attempt.

Security protocols & services: An important part of trusted Internet infrastructure is fundamental security services and protocols that are utilized. This includes commonly utilized Virtual Private Networking (VPN) protocols such as SSL/TLS and IPsec.

We performed component classifications despite the definitional challenges with Internet infrastructure. There are other components and indicators which we studied in the preliminary stages of this work, but ultimately chose to focus on higher-impact and more accessible parts of the overall picture.

Each of the chosen components have various indicators associated with them that can be measured and from which derived metrics can characterize a specific cybersecurity risk factor (see Appendix 4).

Many of these are, on the surface, easy to count and measure. However, there are complexities in measuring each of these. For example:

Infrastructure Component	Sample Challenges to Measurement
DNS	• Need for a list of domains

Certificates	• Different standards for acceptable certificate ciphers
	• Decisions about time - if a certificate is gathered on day 1, expires on day 2, and is evaluated on day 3, do we need to check for an update?
	• If we look to credentials more broadly:
	    ○ Checking if leaked creds are blocked may require logging in
	    ○ Probing may appear to be hostile scanning

Table 1: Challenges to measurement

While this definition and ENISA's asset categorization may be broad and inclusive enough, there are underlying complexities to this work that necessitate a deeper dive and more consideration for this Framework and its ultimate output. and this brings us to recommendation E4:

Recommendation E4: Refine definitions of critical Internet infrastructure.

In doing this work, we have grappled with what counts as Internet infrastructure, and what is measurable from an Internet scanning perspective. For example Internet infrastructure requires electrical power, which may be provided for by a national grid, a microgrid, or local backup power (batteries and generators). Whatever choices have been made, it is hard, using only Internet tools, to assess the overall quality of Internet infrastructure. It is unclear how far we should go in terms of trying to measure electricity or telephone service that might be used to help debug and coordinate in the event of an outage. It is also unclear how "deep" to go into a country's infrastructure.

For example, if we consider a model such as shown in Figure 1, we might all agree that the country-level exchange on the left counts as infrastructure, and perhaps that the block by block network on the right does not, but where in the middle do we make that distinction? For our purposes, more importantly, how do we ensure that the things we are measuring are consistent across groups? If country 1 includes the town's Internet exchange point in its infrastructure, and country 2 does not, that makes comparisons more complex.

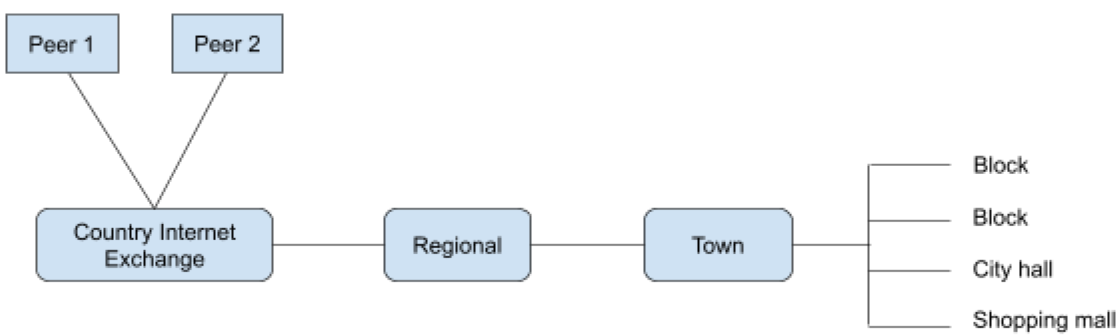

Figure 1: Example model of Internet infrastructure details

Recommendation P1: Conduct a census of critical Internet infrastructure in your country.

That is, produce a specific list of what is critical internet infrastructure. For example:

1. Network interconnections
    a. 10.1.2.3 is the network connection to the Company1 IX point in city.
    b. 192.168.2.3 is the network connection to the Company2 IX point in a different city
    c. cert.gov is a critical resource for distributing security information.

This recommendation may create the appearance of a conundrum, which is that the census will be attractive to attackers, and may be perceived as a "roadmap to attack." However, it is likely that at least nation state attackers can already generate such a list of a country's critical Internet infrastructure, and it may be less accurate, exposing other systems to attack. Our minimalist definition of internet infrastructure may be less sensitive than lists with more broad criteria. For example, our short list of facilities (IP routing) is of systems easily discovered by an attacker, and DNS is literally advertised globally.

We can, of course, make a list of the national DNS resolvers for a country but, in doing so, we discovered that the first several we were investigating did not use DNSec which was planned as a major component of our analysis.

Scoping down and analyzing only what we would measure in this subset has revealed important challenges in defining the infrastructure, defining public health measures, gathering data on compliance with those measures, and assessing the impact of issues with compliance.

For example, is www.cert.jp critical infrastructure? If www.cert.jp serves web pages without HTTPS, is that a critical infrastructure vulnerability? Perhaps it is a choice to prioritize availability over integrity. Perhaps that choice will allow an attacker to tamper with the page as it is viewed, and substitute in a fake phone number that a victim of an attack will then call. If all their pages are always covered by TLS, is TLS 1.2 acceptable or a problem? Security experts agree 1.3 is better, but not all national standards reflect that. If they are following a standard which has not been updated, how should we score that choice? This example and these questions bring us to recommendation P2:

Recommendation P2: Evaluate national standards for security advice and the consistency and character of that advice.

5. HEALTH METAPHORS: MEDICINE AND PUBLIC HEALTH

BACKGROUND

A physical health checkup is a medical activity: its goal is focused on the health of an individual. Medicine's focus on the health of individuals is complemented by public health's focus on the health of communities.

Public health considers both communicable and non-communicable diseases (such as heart disease or diabetes). Many non-communicable diseases relate to lifestyle choices, genetics, environmental factors like pollution, or a combination of all three. Public health officials do not prescribe fixes and walk away; they carefully study how well those fixes work and the factors that inhibit patients from taking them, including patients not knowing about a treatment, feeling that the cure is worse than the disease, or feeling that the cure is too expensive to get or too difficult to remember.

Public data on deaths informs public health investments: we spend more money on cancer than on flesh-eating bacteria because we know that cancer kills more people. Importantly, many countries collect data on causes of death and hospital admittance, and use that data to inform public health investments including fundamental research, applied research, ongoing training for medical professionals, and public health campaigns. Large countries spend money crafting and aligning their data collection, including the criteria for what constitutes a disease.

METAPHORS

One goal of this work has been to create a framework similar to physical health related aspects where yearly health checkups result in indicator data measurements (e.g. cholesterol levels, creatinine levels, blood sugar levels) and the results are used in a diagnostic process to assess certain health risks.

> Diagnosis has been described as both a process and a classification scheme, or a 'pre-existing set of categories agreed upon by the medical profession to designate a specific condition'. When a diagnosis is accurate and made in a timely manner, a patient has the best opportunity for a positive health outcome because clinical decision making will be tailored to a correct understanding of the patient's health problem. In addition, public policy decisions are often influenced by diagnostic information, such as setting payment policies, resource allocation decisions, and research priorities.[6]

Much of the work on enterprise risk management is analogous to medicine; there has been a great deal of metaphor and very little discussion of how public health really applies in the digital world. We would like people to "show hygiene" and come across as scolds when they do not.

PUBLIC HEALTH COMPLEMENTS MEDICINE; CYBER PUBLIC HEALTH COMPLEMENTS ENTERPRISE RISK MANAGEMENT

Even if we assume businesses are motivated to protect themselves for reasons of self-interest and can do so effectively, a society of secure businesses is not a secure society. There are many potential problems that societies band together to address. For example, thieves running down the street or onto someone else's property, pollution, businesses that are losing money and cannot afford security, and businesses that sell adulterated or unsafe food. There are many choices of metaphor which we could use. The COVID-19 pandemic makes public health an obvious choice. Further, from computer viruses to cyber hygiene, the metaphor is in frequent use. More importantly, the metaphor provides us with a few useful characteristics that other metaphors do not.

6 National Academy of Sciences, Engineering, and Medicine 2015

First, it allows us to focus on harms of various sorts. This is also true of police metaphors, but those are generally focused on a list of crimes, and require agreement between countries on what counts as a crime. This is a contentious issue, debated as extradition treaties are created. So, unlike police metaphors, public health does not require us to define crimes, or to select a given country's list of crimes. Third, public health allows us to look at things which impact an individual (lack of exercise), other, specific, people (communicable disease), or communities (pollution).

Another way to think of this is that enterprise security is about your health, while cyber public health includes the health of others and the unhygienic conditions which allow other problems to thrive.

APPLICATION

Both the medical (physical healthcare diagnostic model) and the public health models inform the IIHMF. A set of categories are presented to designate a specific cybersecurity risk condition. Measurements are made for specific indicators and metrics formulae are used to normalize the measured data and characterize a symptom. The resulting metric is then used to designate a specific cybersecurity health condition (i.e. the cybersecurity risk factor). The risk factor can then be used to ascertain which mitigation strategies to use to reduce the risk, similar to recommended treatments once a specific health risk is diagnosed.

We lack data about what causes problems in the digital world. In the first half of 2018, there were 945 data breaches reported to regulators worldwide, while in the second half, 41,502 breaches were reported.[7] There was no real change in security; what changed was that GDPR went into effect and mandated reporting of breaches to regulators. However, we lack ways of nuanced or mature ways of characterizing these breaches and their proximate or root causes, or the effectiveness of mitigations. Information about the effectiveness of mitigations would allow future public health measures to prioritize recommendations. This leads us to recommendation P3:

Recommendation P3: Mandate that companies reporting privacy breaches include information about the controls in place and their efficacy relative to the breach.

Mitigation strategies will take many forms, including ones affecting procedural, operational, technical and educational aspects. Breaking down the Internet infrastructure cybersecurity risk factors to identify and characterize the specific symptoms will help ascertain which mitigation strategies should be implemented and where modifications to existing mitigation strategies need to be made to be more effective.

Many of the reasons people do not get healthcare may have digital equivalents. For example, people fail to learn about patches, or they are worried that the patch will require them to learn a changed user interface or perhaps break something. They may feel that multi-factor authentication is too expensive or inconvenient, or worry that they will lose a USB dongle

7 Shastri, Wasserman, and Chidambaram 2021

that allows them to access their bank account. This study begins to systematically survey such challenges.

It is tempting to conclude that some economic incentives and regulations may need to be considered, but in order to have a better understanding of the challenges, we need to drill down with our investigation. Reasons for inaction may be different across regions. For example, in the ASEAN region and East Asia, development stage may play a factor in addition to resources. This leads us to recommendation E2:

Recommendation E2: Investigate reasons that organizations are not acting on security advice.

Such a list will inform further work in many forms, including product improvements, toolkits, and software to help configure and manage systems. Ideally these will change the way the infrastructure is managed, and that will be visible through consistent periodic measurements.

Consistent periodic measurements and comparing trends over time will give systematic indications on whether the risk factors are reducing, staying the same, or increasing. If these measures do not change, ideally, at least the reasons that organizations are not acting will change, and we can look to further improve the situation. It may also be interesting to consider innovative approaches to these problems, and whether a scorecard could track the amount of innovation that we observe as we measure.

Such consistent measurement requires guidance on what to measure and how. For example, in measuring the cost of downtime, one organization might include legal fees paid as retainer, while another might call that an ongoing legal cost. One company may offer a money back guarantee, while another does not. Legitimate choices can result in dramatically different costs and accounting for costs.

It is important to note that the goal of information gathering in the diagnostic process is to reduce diagnostic uncertainty enough to make optimal decisions for subsequent care. Similarly, for the IIHMF there is no illusion of precision and the metrics created from the individual indicator measurements are constructs which will indicate whether a specific cybersecurity risk factor is improving, getting worse or staying the same.

6. RISK MODELS

There are many ways of modeling risk, and many ways of integrating that into planning and assessment activities. In this section, we review some common ways of looking at technical enterprise risk, then look at how public health approaches differ.

TECHNICAL ENTERPRISE RISK APPROACHES

Before delving into a public health model, the technical risks that are relevant to enterprises should be understood. In a technical environment, "risk is usually expressed in terms of risk sources, potential events, their consequences, and their likelihood.[8] Assessing risk from a cybersecurity perspective has been challenging due to the many uncertainties in a complex environment where a specific event can have multiple causes and lead to multiple consequences. The consequences also may not be immediately known but may accumulate over time.

There are two types of risk analysis models: quantitative risk analysis and qualitative risk analysis. A quantitative risk analysis often seems more objective, scientific or data-driven. A qualitative risk analysis makes no attempt to put hard numbers on things. A qualitative risk analysis will also include the appropriate categorization of the risks, either source-based or effect-based.

Standards such as the ISO 31010 *Risk Management – Risk Assessment Techniques* and the NIST Special Publication 800-39 *Managing Information Security Risk* have been developed to help organizations manage risk in enterprise networks. However, there are no existing standards to assess cybersecurity risk in an Internet Infrastructure Public Health context.

A first pass at adapting the enterprise model to critical Internet services in a country results in a list such as:

- **Data Confidentiality**: having secret information accessed by unauthorized individuals,
- **Data and System Integrity**: not being able to trust that the data came from authorized source and was not modified in transit,
- **Data and System Availability**: not having data or systems be accessible,
- **Vulnerability Exposure**: exposing infrastructure components to known protocol or implementation vulnerabilities.

The first three are very standard, and are explained, for reference, in Appendix 2. Vulnerability exposure was added to the list because, even without being attacked, exposure to vulnerability seems relevant to a public health model. However, in attempting to tie these to public health, we discovered that these traditional approaches seem very focused on the enterprise, not on a society, and started investigating alternative approaches that would tie more closely to public health.

FRAMING TECHNICAL RISKS TO PUBLIC HEALTH

This IIHMF is designed to align the technical risks and mitigations to commonly understood public-health concepts.

We have explored a number of analysis tools, including different models focused on goals, diseases, or impacts. One difficulty with each of these models is that attacks often have both primary and knock-on effects. A great many attacks on computers (rather than infrastructure) have as their effect "the attacker can run code of their choosing." That model, somewhat obviously, does not offer a lot of granularity or distinction.

8 International Organization for Standardization 2018

As part of this work, we crafted three models which connect computer security issues to public health. Each is focused on the impact of an activity, and thus we call them Impact Model 1, Impact Model 2, and Impact Model 3. They are each useful for their direct value in doing the analytic work, and as subjects of study. We present each model in turn, to help readers understand our journey, and then present some lessons learned. This work is still evolving, and we realize that there are tremendous depths to be plumbed. For this phase of this project, we are using Impact Model 3.

IMPACT MODEL 1

Impact Model 1 draws on a characterization of areas of concern to public health:

- Disease
  - Communicable (the flu)
  - Non-communicable (heart disease)
- Environmental factors
  - Pollution
- Lifestyle factors
  - Exercise
  - Diet

Perhaps we can analogize, and say that spam is like pollution, and a failure to patch is a failure to exercise. But as we go deeper, there are few practices that are not like "lifestyle". The presence of malware may be like disease, but it is difficult for remote scanners without credentials to detect, much like someone generally needs to go to a doctor and be examined before they are diagnosed with a disease.

We organized the outline above into tree form, with rounded rectangles representing public health concepts, and parallelograms representing computer security concepts that seemed to our team to relate closely to those public health concepts.

The "run code" node is darker, because a great many computer security problems involve a state change where an attacker gains the ability to run code on a new computer or with new account privileges. For example, ransomware is installed and can then run code on a computer to encrypt the local data. SQL Injection, Cross Site Scripting (XSS), and buffer overflows all give the attacker the ability to run code. Similarly, phishing gets an attacker the ability to run code as the victim. That code is constrained by the bank, but the attacker has new abilities[9]. Running code is often informally labelled "taking control of." We avoid that language because the original owner of the system may or may not lose control as a result of the attacker gaining the ability to run code. We consider "running code" to be the problem

9 This is a slightly unusual perspective on phishing, as the code the attacker is now running is hidden behind buttons with labels like "transfer funds", and that button runs code. Sometimes, these are thought of as credential theft.

because it enables the attacker to do things that may result in the system owner losing control.

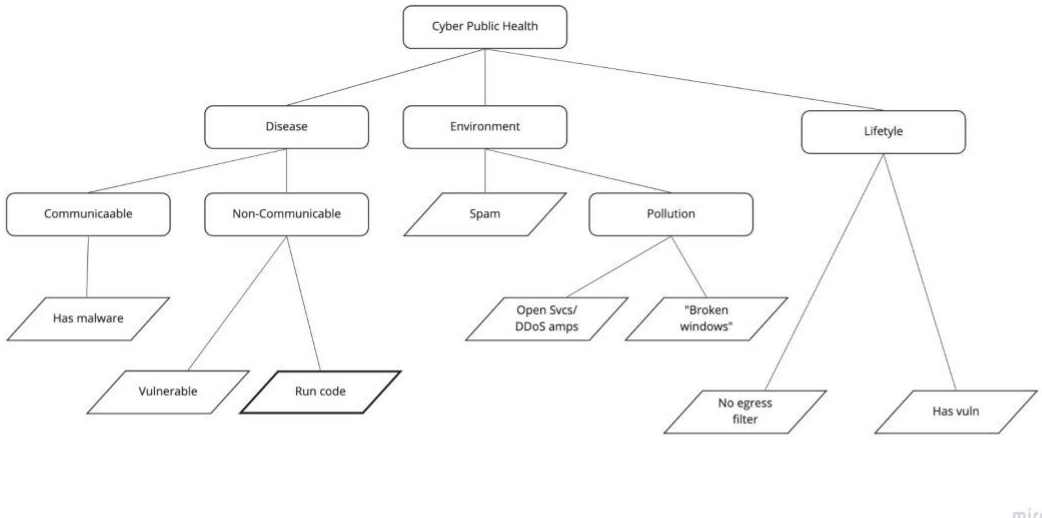

Figure 2: Impact Model 1 tree

Impact Model 1 initially seemed interesting, but it turned out to be difficult to use in practice. For example, we discovered that things like "close open services to prevent DDoS Amplification" could count as "pollution reduction" or "lifestyle."

IMPACT MODEL 2

Given the challenges with Impact Model 1 and project timelines, we chose to drill down to a few interesting nodes to allow us to categorize proposed measurements. Our process involved reducing the number of nodes and in doing so, we limited our ability to consider certain problems, and recategorized others. For example, in Model 1, "lifestyle" included not egress filtering, which is something one does primarily to protect others, and so its location in the model is changed in Model 2. Model 2 also gives up the ability to categorize "has malware", because we have a goal of protecting oneself, and addressing a disease is not a goal in Model 2. This is not an argument that these issues are unimportant, but a recognition that we can make progress on a subset of the problem. This brings us to recommendation E3:

Recommendation E3: Develop a fuller model of the mapping between cybersecurity issues and public health issues.

Impact Model 2 is focused on where the primary benefit accumulates:

Goal (Primary)	Categorization	Example
Protect oneself	Lifestyle	Patch vulnerabilities
Protect others	Environmental	Egress filtering

Table 2: Impact Model 2

Lifestyle: These are choices made in mitigation where if you are not deploying then you are impacting initially only yourself and/or a very small subset of Internet infrastructure. Over time this may become more impactful to others.

Environmental: These are choices made in mitigation where if you are not taking action, the impact falls on others. Environmental factors evolved into a more specific "harm to others" in Impact Model 3.

Impact Model 2 is quite simple. Its simplicity is a strength, enabling direct technical ties between Internet scanning measures and public health of Internet infrastructure. It's simplicity is also provocative: is that all we need? Can we get more value from more detailed or nuanced models? For example, it lacks any characterization of disease or communicability.

Disease: These are choices where there is an active problem, and you are choosing to not be treated. For example: In March 2021, 25,000 German organizations did not take advice to rapidly patch a flaw in Microsoft Exchange.[10] Microsoft and BSI actively urged organizations to patch saying, "now, these exploits are being deployed at mass scale against thousands of targets - apparently worldwide." Perhaps having vulnerable software is like having a disease, or perhaps having malware installed is like having a disease. Or perhaps one disease, relatively easily treated, can lead to complications, harder to treat. Perhaps the choice to not take the Microsoft-recommended treatment of installing a patch is like not taking an aspirin, or perhaps it is more like not getting chemotherapy. There is potentially useful research to be done in adopting reasons people refuse medical treatment to computer security. For example, perhaps waiting on vaccines is like waiting on patches: we wish to see if others have bad side effects.

Communicability is also a potentially important property. "Traditional" Internet worms, such as the Morris worm of 1988 or Slammer and Blaster of the 2000s, propagated by exploiting vulnerable services. Those services were programs that were acting as servers, awaiting connections. They were vulnerable in that an attacker could exploit a code flaw to run code, and the code which they ran was a new copy of the worm. The communicability property is that each worm scans other computers for the service, then attempts to exploit them. That takes time. Each computer can only send so many packets in a second. The more vulnerable servers which exist, the faster the worm will propagate. The higher the proportion of vulnerable servers to IP addresses, the faster the worm will propagate. This has an obvious analogy in the now-famed R0 of public health, the rate of spread of a disease.

10 Rinke 2021

Impact Models 1 and 2 are useful outcomes of this work despite their flaws, or perhaps even because of those flaws. They enable a more nuanced discussion of strengths and weaknesses of these approaches.

IMPACT MODEL 3

Impact Model 3 simplifies the "environmental" factors of Model 2 into a "Harm to Others" category. This is useful as a refinement to "environmental" factors which focuses attention on a reason that system owners might not bother to address a problem, which is that the problem does not hurt them. Compared to the difference between Model 1 and Model 2, the change from Model 2 to Model 3 is fairly small. Model 3 is focused on the harm, rather than the action someone takes. In Model 2, actions could be by system owners or by system designers. The actions could be to set a default or to change a setting away from the system default.

Impact Model 3 is simply that a problem has, as its most obvious outcome, either harm to self or harm to others. Out of date software that would allow an attacker to run code on my computer is counted as harm to me.[11] A service that allows network amplification of denial of service is counted as harm to others.

In this model, while we work to understand the risk indicators and whether the harm is primarily to me or others, we also understand there are other (or "side") effects that may result from a risk indicator's presence. Those other effects sometimes appear different from the primary effect, and lead to apparent confusion. As an example, in Table 3, we note that "Out of date software" is primarily a harm to me, but that an attacker could also use that to install a bot and attack others.

We have chosen other effects which may give the impression of being circular. Seeing them as circular would be a mistake. Computer security experts see the primary impact of a vulnerable service as harm to the system and its owner. Similarly, a small increase in network fees is not the first thing we think of when we hear about DDoS amplification.

Problem	**Primary harm** (Impact Model 3 Categorization)	**Explanation**	**Other effects**
Out of date software	Harm to self	Attacker runs code on my computer;	Attacker installs a bot used to attack others
Misconfigured	Harm to self	attacker reroutes my	

11 The term "me" is used synonymously with "self", because sometimes using self in a sentence obscures the point. Both are intended to refer to the operator of the system. That is likely an organization of some form, and refers to harm to any part of the organization.

software		network packets because of a lack of ROA.	
Open port (amplification)	Harm to others	Attacker uses my computer for DDoS amplification	I spend more on network fees

Table 3: Impact Model 3

We chose to use Impact Model 3 to guide our work in Appendixes 3 and 4, because of its simplicity and apparent ties to public health.

Impact model 3 is very simple, and there are parts of our measurement plan that illustrate its boundaries. For example, is poorly configured TLS on a website a harm to the site owner or their customers?

OTHER IMPACT MODELS

We used other, ad-hoc models in this work, such as:

- **Reputation**: The impact of a problem resulting from the measured indicator would be impact to the reputation of the system owner or operator
- **Cost**: The impact of a problem would be costly to the operator. Which costs are attributable to the problem, and how to handle indirect costs both lead to complexity.
- **Availability**: May be either the system being measured, in that it goes offline and becomes unavailable, or contribution to denial of service attacks.
- **Lack of confidence:** We lack technical measures of confidence.
- **Adverse effects:** This is very broad.
- **Confidentiality/Integrity/Availability (CIA), augmented with Vulnerability Exposure:** This traditional model of enterprise risk does not directly describe the impacts to others that are associated with public health.

Tying these other impact models to public health was challenging, and so they were replaced as the project progressed. The CIA(V) model was most developed, and is discussed in Appendix 2.

INDICATORS TO MEASURE COMPONENT HEALTH

Each of the areas defined in the *Components of Internet Infrastructure* subsection of this report have specific indicators to be measured. Appendix 4 provides more detail on each individual indicator.

Compilation of this index was a reiterative process. Our earliest iteration of this document was a much longer collection of Internet security problems that we perceived to be relevant and important (and still do). But as our research progressed, and in trying to gather as much information as we could regarding these indicators, we understood the need to implement and use a list of selection criteria that would help us navigate the decision-making process of keeping or eliminating certain indicators. These included the notion that everything we include must be measurable and the data should either already be collected at a single source or its collection should be simple enough for us to do on our own. The full list of selection criteria can be found in Appendix 4 in the "Selection Criteria - Indicators" section and the list of indicators we chose not to include are listed in the "Tier 2" section.

As we progressed with framing our model to public health using the Impact Models, our correlation of indicators to certain aspects of those different models also evolved. As a result, we have archived certain columns in Appendix 1 that were associated with Impact Models 1 and 2, in favor of keeping only those relevant to Impact Model 3 to avoid confusion.

As cybersecurity and the risks associated with it evolve, so too will our selection of indicators. This Framework is meant to be adaptive in that way. Moreover, we can better assess which measurements are most effective through outcome data. This brings us to recommendation P4:

Recommendation P4: Create standards for incident and near miss reporting and investigation which show which controls were in place, and which functioned as intended.

ROLES AND LIFESTYLE CHOICES IN THE INTERNET INFRASTRUCTURE ECOSYSTEM

We investigated a persona-driven model, assessing how a variety of participants contribute to or inhibit public health through their activities. There is a great deal we could investigate further here, but in the interest of getting to measurements, we have temporarily set this investigation aside.

That model is shown in Appendix 1.

7. MEASURING INFRASTRUCTURE INDICATORS

We will first talk about measurement, to set some definitions, and then talk about public health and Internet infrastructure, and finally, tie them together.

WHAT IS MEASUREMENT?

Let's start from the basics of measurement. When we measure, we assess a thing, possibly on a scale. We might measure apples, and say we have three apples. With a balance, we can say these apples weigh more than those. With a scale, we can weigh them and say they weigh 100 grams each. Doing that requires that we have a measure, grams, and can produce gram measures to some level of accuracy. We can say that Yurie has more apples than Adam this year, and that Adam has more apples than Yurie this week, and both can be true. It can also be true that Yurie has more apples (by weight) than Adam, even if Adam has more apples. Being more accurate comes at a cost: it may be easy to measure to within an inch, but measuring to 1/128th of an inch requires precisely placing and reading the yardstick. There can be similar types of issues with measurement in Internet health. For example, a transient break may lead to a route being unavailable for a few moments. Should we plan to detect that?

There are things that can be directly measured, such as mass or temperature, and measures we can derive from those direct measures. We can say something is heating or cooling with measures of temperature and time, and we can surmise that something is evaporating if over time it is both cooling and getting lighter. We can measure a great many things, some of which are inherent properties of the apples, even though, like weight, they may change with time. Some, like velocity, are of temporary interest ("why is my apple shipment stuck?" or "holy cow there's an apple flying at my head!") Many things which we could measure are not interesting to most apple consumers. There are other things, like the prevalence of bacteria on the apple's surface, which are interesting and hard to measure.

We can either measure things ourselves, or acquire data from others. When measures are complex to gather or interpret, acquiring data allows us to leverage the work of those originators, and avoid issues created by having different measurement systems. We may use the data gathered by others as primary data, or combine it with data we've gathered to enrich our understanding. For example, we might gather IP reputation information at some future point.

We also have a set of combined measures, such as "route problems." Route problems is a combined measure, in contrast to a derived measure. By combined measure, we mean that we can take several direct measures, such as "Route misorigination by a direct customer" and "bogon prefixes by the AS" and combine them. Above, we consider measures such as "per autonomous system," which is a measure that's tricky to use when one of our direct measures is "bogon prefixes by the AS." But it's harder to measure bogon prefixes by IP address - so the combined measures may only lead to a subset of the derived measures.

Evaluation will be an important part of this project. Some of the things we directly measure require an evaluation step or steps. For example, if we want to look at TLS configuration of websites, we can evaluate the version of TLS and the cipher choices for which the server is configured. For example, we might get a response like "New, TLSv1/SSLv3, Cipher is ECDHE-RSA-AES128-GCM-SHA256." In order to assess whether that is a reasonable choice, we need to evaluate it against some criteria. Sometimes criteria will disagree. For example, some guidance states "Use TLS 1.3 only" while others allow earlier versions in some circumstances. We will need to define what evaluation criteria we are using and why. This brings us to recommendation E7:

Recommendation E7: Create a set of evaluation criteria that allow this project and its successors to evaluate observed measures.

CORRELATION, CAUSATION AND COMPLEX SYSTEMS

Moving away from the apples, we might observe that people who are wealthy own Maseratis, and, hoping to make people wealthy, buy Maseratis for them. Of course, this is counter-productive, as they are exceptionally expensive to maintain, and have low resale value as a side effect. So even if the recipients can resell them in a possibly flooded market, we would do better to not buy Maseratis. In the world of Internet infrastructure, there are many things we hope are associated with "health", and the relationship between them is complex, nuanced, and can be tightly interwoven.

As discussed in section 5, and broadly defined, public health as a discipline concerns itself with the health of populations, including disease, elements of lifestyle that either extend or shorten life, and environmental factors, all of which can play into public health. The success or failure of public health interventions can be measured both narrowly (are fewer cigarettes being sold per month per 1000 residents of King County, WA?) and broadly (are the people of King County dying faster this year than they did last year? What is the average age at death, or the life expectancy at birth?) The discipline of public health also looks at the ease and efficacy of applying treatment, and gives serious consideration to possible side effects.

We can analogize these things at a human level, and hope for a cyber public health. As we delve in, it turns out that computers are not people, networks are not societies, and so the precise things which we measure are not the same. There are many different things which have been analogized as computers having a disease. Those include being infected with malware, having a vulnerable configuration (which can be "patched", making the computer "immune" to a given attack. There are technical states ("out of date") which may be analogous to human aging – an aged computer is less likely to resist the latest attacks.[12] There are things that operators do, such as configure security systems, that may be akin to exercise. There are things that operators do, such as leave certain defaults in place, that put others at risk of spam or denial of service attacks. Perhaps these are analogous to polluting.

We lack many things here. Those include tools for consistent measurement, ways to categorize those measures, and ways to tie them consistently and reliably to either a nation or a measure of health. We lack assessments of ease or side effects of treatments. We lack information about the reliability of treatments.

AN INTERNET INFRASTRUCTURE SCORECARD

We seek to measure a set of things which we believe are crucial to an assessment of public health of Internet infrastructure. Having measured those, we can put them into a "scorecard" which brings us to recommendation E5:

---

12 Older computers may lack hardware, such as a Trusted Platform Module (TPM), Arm TrustZone, or Apple's T2 chip. They may be unable to run the latest operating systems, and modern operating systems are incorporating many new defenses. Operators might not have upgraded computers which can run more modern software.

Recommendation E5: Create a formula for an Internet infrastructure health scorecard, and engage with local and international civil society on its content and uses.

We start with direct measurements, such as a list of open port 22 in some IP range, or route authority assignments published. From this we can craft a variety of derived measures, such as open port 22 per 1000 responding devices, or per autonomous system, or per country. We have started to pull together a preliminary list of direct and derived measures in Appendix 4. As we pull in third party data, these measures will become more clear and complete.

We plan to measure different things, and those things will be measured in different "spaces," including the space of IP addresses, DNS domains, and networking Autonomous Systems. For example, in measuring SPF (the Sender Policy Framework for email security), the measure must be by domain; thus we measure for cybergreen.com and cybergreen.com gets one measurement. It may be a domain with just a few computers or millions. When measuring by domain, we do not generally attempt to translate those into computer or IP address counts. Similarly, we do not attempt to go from Autonomous System to a count of IP addresses. That count may fluctuate wildly with time, or a system with a small count of routable IP addresses may have many behind a NAT system.

These derived measures will be used, along with any other normalization factors and metrics, to create scorecards. Scorecards will also rely on a system of weighting and scoring: we might count the number of open CHARGEN ports as more or less important than the absence of TLS on a webserver. The scoring and weighting criteria will need to be defined.

Each scorecard will include a set of measures in some time period of a month or a quarter. We recommend starting with a quarter because it is easier, and we expect the scorecards to change as we gather and address feedback. We also propose that the first few scorecards be seen as samples, so that we are not forever committed to the methodologies used to gather them.

Recommendation E6: Run a pilot to measure internet infrastructure health and engage with the questions raised by preliminary data collection, analysis and comparison.

To create these first scorecards, we can select the measures for each scorecard by available grouping, such as Autonomous System or country, and thus create a scorecard by country. By arranging such scorecards by country over time, we can enable a variety of analyses.

We can also create a scorecard which arranges indicators by if the issues are harm to self or harm to others. An additional type of derived measure could be "is there more harm to self or more harm to others?" We could create a scorecard with countries or autonomous systems so arranged.

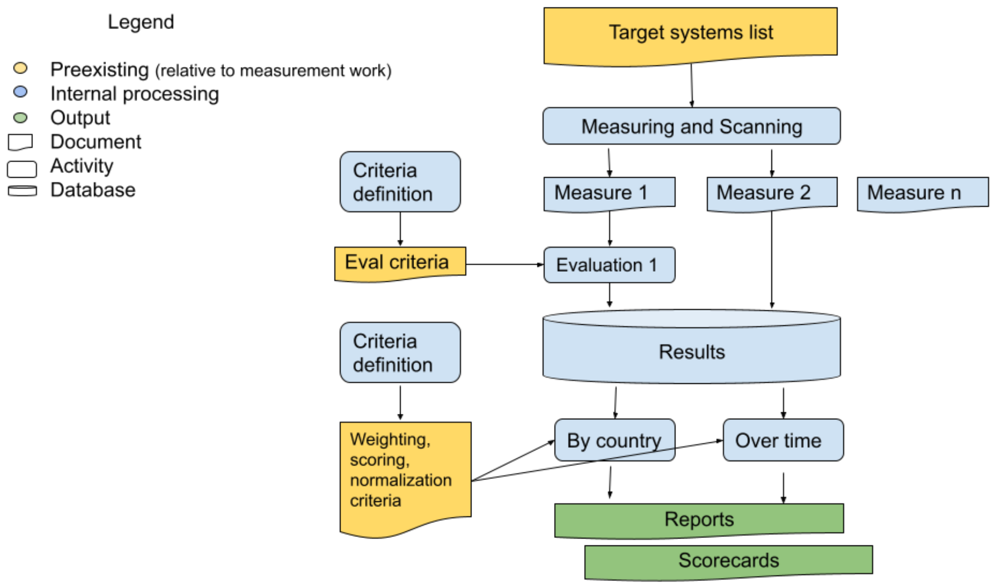

Figure 3: Overall process of scoring system

This report explains Figure 3 and especially the reasoning behind it in some detail. To summarize:

1. We compile a list of components and indicators, based on selection criteria which includes being externally visible and measurable.
2. We define a list of targeted systems by IP address, domain or other qualifier.
3. We perform some set of measurement activity, and record direct output of measure 1, measure 2, etc.
4. We conduct evaluations by applying criteria to the output. For example, one criterion might be that only TLS 1.3 or keys longer than 1025 bits are acceptable.
5. For some measures, we can simply say "there is an open port 19" and, knowing that port 19 can be used in attacks, continue. For other measures, we need to evaluate what we see (is a certificate still valid?). In each case, the measures are recorded in a results database.
6. With those results and a set of weighting, scoring and normalization choices, we can select data either or both by country or over time, and produce reports or scorecards.

8. IMPROVING INTERNET INFRASTRUCTURE HEALTH

In improving the health of a population, we can intervene at the level of manufacturers by, say, requiring seatbelts or pollution controls in cars. Similarly, we can assess and demand certain security features from manufacturers of home routers. As outlined in Appendix 1, there are many players. We focus here on operators of Internet infrastructure, rather than its producers or consumers.

In the interest of achieving measurable progress, we are also focused on externally detectable deployment of specific, effective technical mitigations to explainable problems. External detectability means that we can observe it, in some cases with a scanner, from an arbitrary point on the Internet, and in other cases, by looking at a database, for example, one of routing authority announcements. We focus on effective technical mitigations because we hope to focus attention on those specific improvements.

In this set of recommended indicators to measure, we focus on point in time measurements. This is not to imply that the point in time is the only measurement, but for most measures, such as the number of Routable IP Addresses covered by a valid ROA, we can derive other measures. For brevity, we'll call this RIPAROA. Given a census of IP addresses in a country (IPSPACE), we can measure RIPAROA/IPSPACE. We can then compare RIPAROA/IPSPACE between countries. We can measure RIPAROA repeatedly over time, and report on which IPSPACE (as a proxy for countries) is getting better or worse. A given country might expand its IP space and thus get apparently worse. These important issues will be dealt with in follow on work.

INTERNET INFRASTRUCTURE RISK MITIGATION

Let us imagine two states for a country's Internet infrastructure. In one, 3 of the country's 7 big ISPs run route filtering. In the other, 6 of the 7 do. We have little doubt that one state is less risky than the other, even if we may disagree on how to quantify that risk, which will be the subject of future research.

Recommendation E1: Invest in models and datasets that illuminate risk and connect it to indicators which can be studied externally.

Our immediate goal is to use specifics of what we can measure to allow us to conduct thought experiments to understand what we can learn, how we can analogize, and what measures we might derive. We also hope to perform measurements to provoke further discussion and understanding with concrete numbers.

There may be many technical reasons why deployment of controls or adjustment of operations is not accomplished, and some of these will indicate larger systemic reasons which policy makers will need to consider. Some factors that may be in play include:

- Administrators may not be considering risk holistically.
- Compensating controls may be too expensive to consider in some cases.
- Control advice may be overwhelming or contradictory
- Organizations may be deploying technologies without considering security.
- There may not be enough pressure from policy makers to drive change.
- Organizations may simply not know the best course of action for driving security.

There will need to be further study to determine if those underlying theories are the causes of Internet health issues. That research should attempt to determine the true cause and focus on determining the specific mitigation approaches which policy makers can implement over time to drive down the risks involved, keeping in mind that these issues may be perpetual and iterative studies will be required as technology and culture changes. Further, measuring and normalizing the data associated with the key areas of focus will help to highlight and prioritize the concerns based on the risk each area exhibits.

To be clear, we do not mean to imply this would be a single study. Interesting work has been done[13], and much like public health, changes in the world can result in changes in the health of Internet infrastructures, and so the work will be ongoing.

9. CONCLUSION

Many people see the dangers and problems of the Internet getting worse. Many of these dangers, such as ransomware, are problems for systems operated by private firms or individuals. Others, like disinformation campaigns, operate on commercial systems. It has been hard to discern if problems like contributions to DDoS attacks are getting better or worse.

This report is amongst the first to grapple with precisely defining public health for internet infrastructural systems.

In defining public health for Internet infrastructure, we have created a new opportunity to focus on prevention and mitigation on a global scale. Many problems faced by public and private sector entities are symptoms of unhealthy technical practices, contributors to an unhealthy Internet ecosystem or both. A collective effort to target such underlying causes of systemic cyber risk (risk factors), rather than merely treating its symptoms, will have a far-reaching impact in establishing confidence in the safety and resiliency of the global Internet ecosystem.

These advances are important in and of themselves, in helping us think more clearly about how to apply public health thinking to cybersecurity, and for their practical usefulness to policymakers.

In conclusion, the use of the IIHMF allows states to measure their overall risk, understand how it changes over time, and compare to other states. Policies cannot be developed without a clear understanding of the problems, and those policies' effectiveness cannot be measured without continuous understanding of changes that happen over time. More research must be done to uncover the right metrics, measurements, and normalization techniques needed to tell the story in the proper context and enable thoughtful peer comparison. The scorecard will enable states to understand and contextualize the state of their internet infrastructure in a public health framework.

REFERENCES

"About MANRS." n.d. MANRS Observatory. Accessed 2020. https://observatory.manrs.org/#/about.

13 For example, Tiefenau et al. 2020

Beverly, Robert, Arthur Berger, Young Hyun, and k Claffy. 2009. "Understanding the Efficacy of Deployed Internet Source Address Validation Filtering." *ACM*.

Bischoff, Paul. 2019. "Which Countries Have the Worst (and Best) Cybersecurity?" Comparitech. February 6, 2019. https://www.comparitech.com/blog/vpn-privacy/cybersecurity-by-country/.
Data updated regularly.

"Calculating the Cost of Downtime." n.d. Atlassian. Accessed May 2021. https://www.atlassian.com/incident-management/kpis/cost-of-downtime.

Chung, Taejoong, Emile Aben, Tim Bruijnzeels, Balakrishnan Chandrasekaran, David Choffnes, Dave Levin, Bruce M. Maggs, et al. 2019. "RPKI Is Coming of Age." *ACM*.

"Data Collection | The Shadowserver Foundation." n.d. The Shadowserver Foundation. Accessed 2020. https://www.shadowserver.org/what-we-do/data-collection/.

Felt, Adrienne Porter, Richard Barnes, April King, Chris Palmer, Chris Bentzel, and Parisa Tabriz. 2017. "Measuring HTTPS Adoption on the Web."

Fontugne, Romain. 2020. "The Internet Health Report | RIPE Labs." RIPE Labs. July 27, 2020. https://labs.ripe.net/author/romain_fontugne/the-internet-health-report/.

Huston, Geoff. 2020. "Measuring the Effectiveness of Route Origin Validation Filtering via Drop Invalids from the Perspective of the End User Using a Technique of Broad Scale Reachability Measurement," July.

International Organization for Standardization. 2018. "ISO 31000:2018(En) Risk Management — Guidelines."

———. 2019. "ISO.IEC 31010:2019 – Risk Management – Risk Assessment Techniques."

Jacobs, Jay, Sasha Romanosky, Ben Edwards, Michael Roytman, and Idris Adjerid. 2019. "Exploit Prediction Scoring System (EPSS)."

Kühne, Mirjam. 2016. "RIPE Atlas: RIPE NCC Analyses | RIPE Labs." RIPE Labs. June 30, 2016. https://labs.ripe.net/atlas/user-experiences/ripe-ncc-analyses.

Kührer, Matt, Thomas Hupperich, Christian Rossow, and Thorsten Holz. 2014. "Exit from Hell? Reducing the Impact of Amplification DDoS Attacks."

Lévy-Bencheton, Cédric, Louis Marinos, Rossella Mattioli, Thomas King, Christoph Dietzel, and Jan Stumpf. 2015. "Threat Landscape and Good Practice Guide for Internet Infrastructure." *ENISA*, January.

Lian, Wilson, Eric Rescorla, Hovav Shacham, and Stefan Savage. n.d. "Measuring the Practical Impact of DNSSEC Deployment." Accessed 2020.

National Academy of Sciences, Engineering, and Medicine. 2015. *Improving Diagnosis in Health Care*. Washington, DC: The National Academies Press.

National Institute of Standards and Technology. 2011. "Managing Information Security Risk: Organization, Mission, and Information System View." *NIST Special Publication 800-39*.

Nicholson, Paul. 2020. "Five Most Famous DDoS Attacks and Then Some ." A10 Networks. July 27, 2020. http://www.a10networks.com/blog/5-most-famous-ddos-attacks/. Updated: October 30, 2020

Osterweil, Eric, Michael Ryan, Dan Massey, and Lixia Zhang. 2008. "Quantifying the Operational Status of the DNSSEC Deployment."

Rinke, Andreas. 2021. "Up to 60,000 Computer Systems Exposed in Germany to Microsoft Flaw." Reuters. March 10, 2021. https://www.reuters.com/article/us-usa-cyber-microsoft-germany-idUSKBN2B2262?taid=60493b8cb7d77200018e3357&utm_campaign=trueAnthem:+Trending+Content&utm_medium=trueAnthem&utm_source=twitter.

Rossow, Christian. 2014. "Amplification Hell: Revisiting Network Protocols for DDoS Abuse."

Shastri, Supreeth, Melissa Wasserman, and Vijay Chidambaram. 2021. "GDPR Anti-Patterns." *Communications of the ACM*, February.

"SSL Server Test (Powered by Qualys SSL Labs)." 2020. Qualys SSL Labs. 2020. https://www.ssllabs.com/ssltest/.

The Spamhaus Project Webteam. n.d. "The Spamhaus Project." The Spamhaus Project. Accessed 2020. https://www.spamhaus.org.
Tiefenau, Christian, Maximilian Häring, Katharina Krombholz, and Emanuel von Zezschwitz. 2020. "Security, Availability, and Multiple Information Sources: Exploring Update Behavior of System Administrators." 2020. https://www.usenix.org/conference/soups2020/presentation/tiefenau.

Trimintzios, Panagiotis. 2011. "Measurement Frameworks and Metrics for Resilient Networks and Services - Technical Report." *ENISA*, February.

"What Is the Internet's Infrastructure? [Video] ." 2019. I2Coalition. February 6, 2019. https://www.i2coalition.com/what-is-the-internets-infrastructure-video/.

APPENDIX 1: PARTICIPANTS IN THE INTERNET INFRASTRUCTURE ECOSYSTEM

The Internet Infrastructure ecosystem is comprised of distinct roles, each of which has the responsibility of ensuring they make informed cybersecurity risk mitigation decisions in the area of the Internet infrastructure over which they have authority. The actions or inactions taken by any entity reflects a lifestyle choice. Similar to public health, some lifestyle choices do not impact anyone else, such as wearing a helmet when riding a motorcycle. However other lifestyle choices, such as smoking, can affect other people's health.

The following table shows what each role is responsible for and what lifestyle choices an entity should consider. This list is not exhaustive, and is the product of several brainstorming sessions across expert working groups.

Role	Description of Role	Lifestyle Choice
End User (Individual/sysadmin)	The individual utilizing digital systems and applications	- Should full data disk encryption be enabled (instead of relying on cryptographically protecting only sensitive files)? - Is multi-factor authentication mechanism used? For what percentage of application and system logins? - What percentage of passwords are reused on multiple systems? - Is there an inventory of all devices under their control? - How often is the inventory updated? - How often is the inventory validated? - How prevalent is the use of default parameters? - Is there an inventory or listing of connections to external parties? - How often do they patch? - How quickly do they patch vulnerabilities after public announcement? (useful to know but too many unknown factors since also pertains to what can be automated, whether they have privileged early notifications, whether custom code involved from vendors that needs evaluation or added development cycles to include patch, etc) - How quickly is a workaround deployed to mitigate vendor vulnerability risk if need to wait for a patch from the vendor?
Providers (Software, Services, Platform, Content)	The vendors who create devices and/or digital services (including cloud)	- Is data in transit cryptographically protected? - Is data at rest cryptographically protected? - What percentage of software/service/platform/content provider user authentication has multi-factor authentication capabilities? - What percentage of customers have multi-factor authentication configured? - How secure are the Software Development Lifecycle (SDLC) processes? - Has there been a comprehensive threat

		modeling performed? - Are there mechanisms in place for receiving, evaluating, and addressing exploitable vulnerabilities?
Service Providers (Communications Infrastructure)	The ISPs, mobile phone companies and new infrastructure providers	- What percentage user authentication has multi-factor authentication capabilities? - What percentage of customers have multi-factor authentication configured? - Is data in transit cryptographically protected? [this can relate to RPKI, DNSSEC] - Is data at rest cryptographically protected? - Can special use and reserved routing prefixes be propagated? - Are mechanisms in place to prevent routing leaks? - Are mechanisms in place to ensure only traffic with valid source IP addresses are forwarded? - Can DNS cache spoofing happen? - Are 3rd party pen tests performed against infrastructure devices?
CSIRTs	The computer security incident response teams that are at the national level and coordinate incident response between government and private business entities	- Access to / relationships with critical infrastructure / service providers / LE - Research resources and agenda - Mechanism to communicate findings, their severity, etc.
Government Agencies		- Are processes in place for setting up secured communications when necessary? - Existence, adequate funding and adequate staffing of a national CSIRT? - Policy about/around cyber risk mitigation and incident response? - Ability to staff
Device Manufacturer		- Ability to be remotely reset / restarted and updated - Ability to alert users (directly or via a signal to a SIEM) to suspect activity - Ability to alert users (directly or indirectly) to alteration - Ability to alert users to conditions that may put CIA in jeopardy - What other insecure protocols exist as default on in vendor devices [ftp, tftp, SNMPv1/2, etc]
Protocol Developer		- Do IPsec related security parameters matter? - Is list of SSL/TLS parameters of interest which have known vulnerabilities and exploits? - Should we look at how prevalent Telnet availability is (given that it was how Mirai was instantiated)

APPENDIX 2: ENTERPRISE RISK DEFINITIONS

Data Confidentiality

Data confidentiality refers to protecting data against unintentional, unlawful, or unauthorized access, disclosure, or theft. Most often, access to confidential data starts with gaining access to passwords and other credentials to gain unauthorized privileged access and carrying out actions that may lead to the misuse or abuse of a system. Nefarious actors with malicious intent often follow a progressive pattern of activity designed to give them comprehensive insight into a system's design and configuration in order to establish some level of persistence against being detected and ejected from a system for the duration of a system's utility to the attacker. The following attack objectives create a data confidentiality risk for Internet infrastructure assets:

- *Eavesdropping:* A passive attack where someone is stealthily listening to the private conversation or communications of others without their consent in order to gather information.

- *Interception:* A passive attack whereby information is gathered to ascertain the network topology or specific device information, which can be further used to exploit known vulnerabilities and/or lead to espionage.

Data and System Integrity

Data and system integrity refers to protecting the data and the systems that the Internet Infrastructure relies on against improper maintenance, modification or alteration. It also includes data authenticity where data must come from specific, trusted sources. This risk encompasses areas where data must remain accurate and uncorrupted, with modification only by certain people under certain conditions. The following attack objectives create a data and system integrity risk:

- *Message Forgery:* This active attack is often instantiated in falsification of messages such as emails to create phishing campaigns that seek to make recipients of the email disclose and or verify sensitive information.

- *Message Diversion/Deletion*: An active attack where legitimate messages are removed before they can reach the desired recipient or are redirected to a network segment that is normally not part of the data path.

- *Message Modification*: This active attack is one where a previous message has been captured and modified before being retransmitted. The message can be captured using a man-in-the-middle attack or message diversion.

Data and System Availability

- *DDoS Amplification Attack*: An active attack where an initial small query turns into a much larger payload, targeted at a specific victim.

Vulnerability Exposure

A *flaw* is an error in the implementation or design of a system that can cause a deviation from state; a *fault* is a deviation from state occurring due to chance interaction with a flaw, and a *failure* is a deviation from mission due to a flaw. A *vulnerability* is a condition in a system, whether by design or arising through a flaw, that permits a violation of the implicit or explicit security guarantees of a system. An *exploit* is an explicitly engineered artifact which uses the vulnerability. The difference between a fault-tolerant and vulnerability-tolerant system is that the former accounts for chance, and the latter for interest.

- *Protocol Vulnerability Exploitation*: An active attack that takes advantage of known protocol vulnerabilities, due to design or implementation flaws, to cause inappropriate behavior.

There is not always a one to one mapping between attack vectors and consequences. For example, a routing integrity issue may lead to one or more of:

- Information disclosure when the contents of packets are read
- Use of the disclosed information to plan further attacks
- Denial of service when a system is overwhelmed by traffic

We can consider a threat the possibility of future violence, and a risk to be a threat with quantified likelihood and impact (at some level of precision). Obviously, understanding the impact to an organization or others is a part of what motivates us to take action. It may be that the lack of easily understood causality makes not acting easier.

In the mitigations section, we list some possible consequences. We are hopeful that clarity of impact will drive action.

APPENDIX 3: MITIGATION RECOMMENDATIONS

The following tables provide a preliminary summary of the harms associated with the six components chosen for the IIHMF, how measurements can be taken, and mitigation recommendations.

Open Services	
Harm to self	The system locations, in great enough quantity represent a **reputational** risk for the country or region. Additional **costs** due to bandwidth utilization and reduced **availability** of Internet resources may generate an impact if great enough.
Harm to others	Attackers could use to perform amplification of DDoS attacks, expose system information (for example via SNMP) or take advantage of exploitable vulnerabilities to gain system-level access
Indicator Measurement	In this report, this indicator is largely focused on open services that can be used in DoS amplification, including DNS, NTP, SNMP, SSDP, and others. Open services can be determined by scanning against the devices

Mitigation	System developers: default configurations should not ship with services open unless needed. Deployment: ● Systems operators should have processes for checking deployed configurations and turning off un-needed services. ● Mechanisms should detect potential abuse such as traffic volume monitoring to detect abnormal behavior.

Routing	
Harm to self	Whether introduced through human error or malicious threat actor activities, traffic could be misdirected and either lost completely or redirected to malicious endpoints.
Harm to others	Routing problems can lead to a **lack of confidence** in the infrastructure of a country if this happens too often within a specific area. Additionally, critical infrastructure that requires robust communication can be **adversely affected**.
Indicator Measurement	We plan to rely on ISOC and other internet observatory data, as they are working hard on ensuring that there's good data and the challenges with it are addressed.

Mitigation	This may not be addressed due to lack of reporting, lack of measurement of redirection, or difficulty in deploying the proper countermeasures. Technical: Routing infrastructure devices should be capable of filtering routes and be capable of Route Origin Validation. Operational: • Internet service provider operators should have filters in place to make sure they only accept the correct prefixes from their customers. Prefixes exchanged between BGP peers should be controlled with inbound and outbound filters that can match on IP prefixes, AS paths or any other attributes of a BGP prefix. • The operators should also ensure that the Internet Routing Registry (IRR) database is up to date with their information. This database contains the Internet routing information used by network operators to register their assigned network resources. Many providers utilize the IRR information and existing tools capable of retrieving information from the registry to build a list of originated or transited prefixes (IP address block information) that can then be utilized to create automated filtering rules. • Policies and procedures should be in place to encourage the use of good filtering practices, keeping the IRR up to date and deploying RPKI.

Domain Name System (DNS)

Harm to self	Availability risks are introduced when modification of DNS query/response paths or DNS protocol data can lead to traffic being blackholed and not reaching its intended or being used as a DoS vector. Availability could be compromised when DNS requests are sent that causes larger than expected responses and results in a DoS attack. While there are multiple ways that DDoS attacks can be instantiated in the DNS, one of which is the DNS open service vector discussed in the open services section. In this Framework we will restrict the attacks to areas that can be measured and mitigated. DNS is a core component of Internet infrastructure, yet many DNS health risks that affect integrity only impact a subset of the DNS ecosystem. The exception is large-scale DDoS attacks that have more significant and wide-spread negative effects. There can be significant wider-scale impact of the top-level domains (i.e., gTLDs and ccTLDs) are impacted but for now this health risk and impact is identified as a lifestyle risk.
Harm to others	Integrity can be affected by forging DNS responses and causing traffic to be misrouted to malicious servers. DNS integrity can also be compromised using cache poisoning attacks where legitimate DNS queries receive falsified responses. This is also sometimes referred to as DNS spoofing since the DNS responses are "spoofed" or altered to redirect traffic to an attacker's chosen destination.
Indicator Measurement	The indicator measurements for DNS provide information to ascertain the level of Domain Name System Security Extensions (DNSSEC) deployment readiness. Indicators would also include open DNS sources determined during the open services research, as well as understanding the total number of DNS server deployments in a region.

Mitigation	<u>Technical:</u> DNS infrastructure devices should be capable of DNSSEC. <u>Operational:</u> DNSSEC should be deployed by DNS operators. <u>Procedural:</u> Processes and procedures should be in place to effectively deploy DNSSEC and to ensure that periodic digital key updates are performed.

Email	
Harm to self	Email security is critical for all organizations. Business Email Compromise results in high numbers of malware-infected systems and fraud cases. Email risks can allow for impersonation, forgery, and unauthorized access to messages which could be used to glean sensitive information.
Harm to others	The effects of email-related attacks can be far reaching if unauthorized access or impersonation is used with an authoritative domain (e.g., government, critical infrastructure) to glean sensitive information from others. Further, compromised email accounts can be used to send real emails to other related organizations to further spread compromise.
Indicator Measurement	To measure indicators related to email, a list of domains associated with email addresses associated with a particular focus (i.e., critical infrastructure, government) would need to be known. Scanning for DMARC and SPF implementation could then be accomplished. Other aspects of email security regarding the controls and protections would have to be researched as a matter of policy.

Mitigation	<u>Technical:</u> Implement numerous Email Security Mechanisms. ● DMARC which is a policy created and defined in DNS and enables verification on an email security gateway. ● SPF provides authorization by defining which systems can send messages using the organization's domain name and is created and defined in DNS. ● DKIM provides authentication by adding a digital signature to all messages. The DKIM record is created and defined in DNS. If third parties are used to send messages, then they may need to be involved in the creation of the record. ● DANE and MTS-STS provide authentication and confidentiality as both need to have some form of certificate for authentication to create a secure connection between email servers. DANE is DNS based (requires DNSSEC) and needs assistance from domain registrars to be created. MTA-STS requires a website with a valid SSL certificate in place. DNS is also required to create records for validation. <u>Operational:</u> In order to encourage broad adoption of email security mechanisms, governments should lead the way and deploy across all public domains. This allows the government to claim the private sector should do the same, because the government has already proven it is possible and effective. Governments can further encourage private sector adoption by: ● Identifying or creating tools to help the private sector deploy. ● Requiring or paying a supplemental amount for government suppliers/contractors to deploy. ● Purchasing cloud services for email that include these mechanisms by default and making those same contractual arrangements more broadly available. ● Providing funds to implement these mechanisms or requiring its use, such as in regulated industries.

Certificates

Harm to self	With the easy availability of free TLS certificates, not using a certificate may cause customer concern, brand damage, or substitution of a more secure vendor. The impact to the organization can be reputational, loss of availability, and loss of revenue.
Harm to others	With the theft of digital security certificates, attackers can misrepresent themselves as a legitimate organization. They can thus cause reputational risk, reducing trust in the overall infrastructure. Additionally, availability is clearly affected because the services are no longer directly available.
Indicator Measurement	Certificates are, by design, served up when a web or email (SMTP) connection is made, and that enables collection, analysis and evaluation.
Mitigation	Technical: Ensure that there is a certificate management process, and that your certificates are of appropriate length. Administrative: Monitor the certificates in use for compliance with appropriate standards.

Security Protocols & Services	
Harm to self	Protecting security protocols and services such as VPN, SSL/TLS, and IPsec is critical for ensuring access to people and processes, while preventing unauthorized access to others. When these protocols and services are compromised, it can be difficult to identify the breach, and determine further what other damages may have occurred in the environment.

Harm to others	Harm to others from these services seems hard to find. It's possible that others might need to keep less secure older software around, and use it by accident.
Indicator Measurement	To measure these indicators, certain scanning can be used to determine the existence of the protocol in use but may not be enough to understand the controls in place regarding the protocol. Therefore, measurement would require an understanding of how prevalent the technologies are within a given space (i.e., critical infrastructure) and further research would be required through data gathering to determine if the controls in place for those discovered protocols are meeting appropriate control standards.
Mitigation	Technical: Limit the use of certain protocols such as SSH and use proper controls to ensure appropriate access restrictions. Ensure VPN systems are appropriately set up with end-to-end encryption that meets appropriate standards and uses MFA. Track SSL/TLS certificates using current best practices and ensure the safety of the private certificates.

APPENDIX 4: INTERNET HEALTH INDICATORS

The following tables show our thought process in the selection of indicators to measure and some initial measurement characteristics which might be used to create a scorecard.

Component Descriptions

Component	Description
Open Services	These services should be measurable over either IPv4 and/or IPv6 transport and overall give an indication of whether they can be utilized for amplification attacks. These are typically services with specific ports that were in the past few years used for large DDoS attacks and/or critical services that are ubiquitously used in unmanaged open configurations everywhere such as NTP, DNS and SNMP and are often vulnerable to amplification attacks.
Routing	Indicators to ascertain routing infrastructure health can be classified as actively measured indicators vs. indicators from observed data. Outage incidents and data regarding BGP leaks and/or hijacks are useful indicators to get definitive behavior information. Actively measured data results need context regarding traffic engineering choices to make more accurate conclusions. Since architectural context is difficult to obtain, this limits what can be actively measured.
Domain Name Service	The domain name system is a globally distributed, loosely coherent dynamic database of information. It maps names to IP addresses and is also used for other types of information dissemination. It is a fundamental service that must be reliable, available and trusted.
Email	Email is a critical component of digital communications and it is imperative that this communication can be trusted and relied upon. Technical solutions exist which can minimize email fraud that comes from a fraudulent or impersonated organization.
Certificates	Asserting one's identity through use of digital certificates
Security protocols & services	An important part of trusted Internet infrastructure is fundamental security services and protocols that are utilized. This includes commonly utilized Virtual Private Networking (VPN) protocols such as SSL/TLS and IPsec. Weaknesses that lead to added risk factors include cryptographically weak protocol support, supported of weak ciphers and insecure key lengths, as well as unverified digital certificate parameters (some of which were introduced as indicators in the previous 'Credentials' section. There are also dependencies based on whether intermediary certificate chain is incorrect, whether an intermediary certificate is missing in the certificate chain and whether self-signed certificates are used.

Operating systems/software versions (in "Tier 2")	Fingerprinting a system can help in determining attention paid to common OS vulnerabilities and whether systematic and timely patching is performed. It will also help determine the outdatedness of systems and software.
Legacy Protocols (in "Tier 2")	These services should be measurable over either IPv4 and/or IPv6 transport and overall give an indication of whether they can be a vector for credential compromise because credentials are sent in cleartext rather than being cryptographically protected.

Table Header Descriptions

Header	**Description**
Indicator	What the indicator is
What indicator tells us	Our hope for why we would measure it
Why useful	How the measure ties to a problem, indicates action or inaction, ties to risk, or ties to a problem for others.
Useful added context	Additional context that we wanted to record
Data sources	Where we might be able to get data
Indicator units	What the units of counting are. (Count, list) For some, these are easy, like count of IP addresses or # hosts responding on a port. For others, they're lists (what TLS versions are supported) There are also group indicators where we bring a set of measures into a higher level assessment - for example, "bad routes" are a grouping of things covered in routing in depth
Impact Model 3 (primary harm)	Is the presence of the risk indicator primarily a harm to self or harm to others?
Impact Model 3 (side effects)	Does the presence of the risk indicator have side effects which extend harm beyond the primary recipient? We use "none" to indicate that there are not obvious, direct, or common side effects, not to indicate that there could never be side effects.
Ease of measuring	Notes on how easy it will be to measure. Informal, not yet on a scale. For example, DMARC is easy to gather - it's in DNS. The use of a secure email gateway is unclear how to reliably gather. For ease of measuring IP space issues, we assume IP4 ease - IP6 scanning is slow.
Measurement characterization	What we're measuring. For example, the absolute count of IPs or the amplification available. We may have multiple measures for a given indicator.

First derived measures	Things we can derive from the primary measures, such as IPs per country, or routes not authenticated per AS
Second derived measures	Things that involve a first derived measure in their calculation, so IPs per country being high, average, or low relative to country for some indicator
Notes	Notes for the reader

Selection Criteria - Indicators

Criteria	**Description**
Data availability	It is easy to scan for open CHARGEN ports. It is hard to test to see if an email security gateway is in use.
Data quality including how accurately and consistently it can be gathered	Can we gather the data accurately and consistently? Take for example “default passwords in use.” Assuming we’ve solved important ethical questions of scanning for this data, testing to see if a login is successful can be difficult. For example, perhaps we’re scanning for ssh logins, and to test, we’re looking for a shell prompt, ending in a “%” character. A banner comes back %STOP SCANNING US% and we see a false positive. We might connect to a honeypot, designed to accept all credentials. Our scans might be filtered by defenders. We may be able to gather operating system version information to the resolution of Windows 7 vs 8 vs 10, but we cannot determine reliably which version of Windows 10 someone has deployed from nmap-style fingerprinting.
Data quantifiability	Some data, like open CHARGEN ports, is easily quantified. Other data, like “list of TLS ciphers” is not directly quantifiable. In fact, different standards might declare various sub-values to be acceptable or not. DMARC can be set to (essentially) 4 values: none, report, quarantine and reject. To the best of our knowledge, there is no standard that says to go to “reject". See also the discussion of DNS within “links to a specific country.” Is the collection opt-in, such as the CAIDA anti-spoof tool?
How it links to health – whats the story? Can we draw a causal link?	We have drafted a variety of tools to help us link indicator measures to health metaphors. That work is early, and the data tables are providing very helpful and illustrative test cases, and we expect to do substantial work to refine the linkage tools.
How it links to a specific country (IP, domains, AS)	There are three main families of data we can gather regarding internet infrastructure: IP4 data, Domain data, and AS data. In each case, there’s complexity that we will need to manage.

Redundancy – does the data show something unique?	Our first data tables included both open CHARGEN and open qotd ports within open services. While it is likely that if we collected data, we would get different numbers for each, each is a legacy service that enables the amplification of denial of service attacks, and there is no clear benefit to analyzing the measure of each as we use the possibility of measuring each to help us understand the factors presented in this list, or as we investigate how we would derive meaning from our measurements. As such, we can accelerate the pre-measurement work by assessing only CHARGEN, and later considering if CHARGEN analogs, including but not limited to qotd, are worth measuring because of identified links to health.
Illustrativeness – does the data point illustrate a point about public health or infrastructure?	After cutting for redundancy, “open services” still has 4 services to look at: -Open CHARGEN. A nearly useless service for debugging. -Open DNS. There are “best practice” docs that call for open DNS resolvers as a reliability tool. This is a contested recommendation. -Open SNMP. We are not aware of advice to allow anyone to query SNMP. Additionally, it may reveal confidential information about the network operations of the operator. -Open SSDP. SSDP is a service discovery protocol, and it may be required to configure various devices. Each of these enables DDoS amplification, but the logic for keeping them around and other impacts may differ.
Is it likely to change rapidly or slowly?	Data about routing may change rapidly. DMARC policy is likely to change slowly.
Does a point in-time-measure give us what we need?	The indicator "For each domain with a DNSKEY RR, the number of DNSKEY RRs" is intended to show dilligence in updating RR keys. To really measure that, we need to repeatedly sample the RR key and asssess frequency of changes. The existence of multiple keys could indicate either lots of changes, or that old keys are not properly retired.
Does it illustrate something about actions taken/not taken by the system owner?	This turns out to be surprisingly complex. For example, if the 2015 version of router-manufacturer-1's OS has oen snmp on all ports, then snmp being open would show inaction by the owner, while if they closed it in 2016's defaults, we would need to assess which OS was running, the defaults shipped on that version, and then compare.
Analytic ease	In discussing DMARC, we discussed how there are four states. However, DMARC records indicate the domain’s suggestion for unsigned emails. Is a DMARC policy of “reject” the same measure of health when applied to all of a domain’s emails and when the SPF policy says the domain should never send email at all?
Likelihood of detection	There are some countermeasures or issues which may be so rare in practice that analyzing them tells us little about public health. (This is somewhat circular logic: it is conceivable that they

	are more prevalent than we believe, and so our decision to focus our efforts elsewhere could backfire.)
Countermeasure factors	There are a variety of factors that make defenses hard to implement, discussed elsewhere. It is worth keeping indicators that help us identify such factors.
Cost of data gathering	If the data involves knowing various operating system defaults, then we need to gather a long list, and setting up and maintaining the tests will be more expensive than a simple port scan. If the data comes from a partner who wants to charge, then the cost goes up.
Legal issue might cause	Testing for default passwords might look like a break-in attempt
Global data	Is data available across countries?
Cross comparability	Is the way the data is gathered similar enough to allow us to compare and contrast between sources?
Baselining challenges	What is the base data that out there? What security configuration was it set to, for example, for CHARGEN? When did it change from default on to default off? Is a version from two years ago still in support? This is easy for a small handful of systems, but it gets hard as we ask, for example, What is the acceptable version of ZyxOS? Where do we go for that data, and how do we check it?

Open Services

Indicator	**What indicator tells us**	**Why useful**	**Useful added context**	**Data sources**	**Indicator units**	**Impact model 3 (primary harm)**	**Impact model 3 (side effects)**	**Ease of measuring**	**Measurement characterization**	**First derived measures**	**Second derived measures**	**Notes**
Open CHARGEN	Number of CHARGEN open ports (UDP19)	Gives information on legacy protocol use and risk of amplification attack	- IPv4 vs IPv6 - % of all IP space in country	CyberGreen	IP4 addresses with open port	Harm to others	None	Very easy	A = absolute # ("count")	B = A/per million hosts C = A/ per AS D = A/ Change over time (count)	G = B over time H = B relative to other countries* J = C relative to other AS * Measure H requires consistent allocation of host to coutry	Does anything ship with this open after, say 2010? When do we cutoff when judging if someone took action? Is it "supported"? "new"?

Open DNS	Number of DNS recursive resolvers that answer to any query (UDP 53)	Can ascertain risk of utilizing DNS for amplification attack	- IPv4 vs IPv6 - % of all DNS recursive resolvers in country	CyberGreen	IP4 addresses with open port	Harm to others	None	Very easy	A = Absolute count K = amplification available	As M3 + P = K/million hosts,Q =K/AS	P, Q over time	There may be a risk or percieved risk that customers depend on current configuratio n which would inhibit change
Open SNMP	Number of SNMP servers that answer to any query (UDP 161)	Can ascertain risk of utilizing SNMP for amplification attack	- IPv4 vs IPv6 - % of potential SNMP servers in country - SNMP1 vs SNMPv2 vs SNMPv3	CyberGreen	IP4 addresses with open port	Harm to others	Harm to self	Very easy	L = host enumeration M = host enum, IP6	As open CHARGEN	as open CHARGEN	Harm to self is disclosure; For impact model II, this one has both harm to others (ddos amplify) and harm to self (lifestyle)
Open SSDP	Number of SSDP servers that answer to any query (UDP 1900)	Can ascertain risk of utilizing SSDP for amplification attack	- IPv4 vs IPv6 - % of all SSDP servers in country	CyberGreen	IP4 addresses broadcastin g	Harm to others	Harm to self	Unclear - see notes	As open DNS (line 4)	As open DNS	As open DNS	Ease of measure: I think of SSDP as a UDP broadcast, and so it's not clear what the expectation is wrt egress filtering, etc. For impact model II, I assume the primary harm is ddos amplification , not discovery of further vulns

Routing

Indicator	What Indicator Tells Us	Why Useful	Useful Added Context	Data Source	Indicator units	Impact model 3 (primary harm)	Impact model 3 (side effects)	Ease of measuring	Measurement characterization	First Derived measures	Second derived measures	Notes
# of ROA	Are they using, managing ROA	ROA protects internet routing		ISOC (RPKI Validator?)	Routes	Harm to others	Harm to self (packets won't go)	Pull from database	A = count of routes, B = # of routes for the AS	C = A/B indicates coverage	TBD	
Bad ROA payloads	There's a problem if someone is issuing bad ROA	Indicates process problems	See "Routing in depth" tab	TBD	By AS	Harm to self	Harm to self (packets won't go)	Medium - what's "bad"?	D= # of bad ROA, E = elapsed time before replacement	F= D/AS, G = E/AS	F, G over time. F, G by AS scope/size	This could be a compound measure of bad = signature fail, or it could be IPs for which they're not responsible
Invalid routes	Number of routes originated by the AS that are invalidated by a corresponding ROA	Can ascertain degree of RPKI deployment	-IPv4 vs IPv6 -% of routes originated by AS	ISOC (RPKI Validator)	By AS	Harm to others	Harm to self (packets won't go)	Complex - defer to ISOC	Per ISOC	Data over time, by country	TBD	Could be both invalidated by ROA and operationally respected by other routers; this measures how well the AS is doing, and the odds that a route hijack will work
Not registered routes	Number of routes originated by the AS that are not registered in an IRR as route objects.	Can determine attention to detail for automated filtering (?)	-IPv4 vs IPv6 -% of routes originated by the AS	ISOC (RIPEstat)	By AS	Harm to others	Harm to self (packets won't go)	Complex - defer to ISOC	Per ISOC	Data over time, by country	TBD	As invalid routes

Route problems	Can the ISPs manage their routing with a reasonable degree of competence ?	Indicates process problems	See "Routing in depth" tab	ISOC Bgpstream aggregation	By AS	Harm to others	Harm to self (packets won't go)	Easy to take ISOC data, complexity in managing	A= count of problems per AS	B = count of problems per day; C = count of problems per IP	B, C over time and/or relative to other ISPs.	

Routing in depth

Indicator	What Indicator Tells Us	Routing problem grouping	Why Useful	Useful Added Context	Data Source	Indicator units	Impact model 3 (primary harm)	Impact model 3 (side effects)	Ease of measuring	Measurement characterization	First Derived measures	Second derived measures	Notes
Prefix covered by ROA	Number of AS prefixes that could utilize RPKI	Bad ROA Payload	Can ascertain to what extent RPKI is deployed for a specific AS	Total number of ASs in country and prefixes per AS (IPv4 vs IPv6)	[Need Source] (utilizing RIR data and open source RPKI validator tools)	AS prefixes	Harm to self	Harm to self (packets won't go)	Medium - what's "bad"?	Derived measure, raw measure not in list This is a % of AS	Comparison between countries or comparison over time	Comparison between countries over time	Timing: is an issue; if ISP A provides ROA by day and misses a day, is that equal to ISP B who does it monthly and misses a month?
Validity of ROA Payload	Whether payload is valid or not	Bad ROA Payload	If payload is invalid and has error it can point to misconfigurations	% of valid and/or invalid payloads compared to total number	[Need Source] (utilizing RIR data and open source RPKI validator tools)	Days with 100% validity?	Harm to self	Harm to self (packets won't go)	Medium - what's "bad"?	derived measure	Comparison between countries or comparison over time	Comparison between countries over time	Is a bad ROA worse than no ROA?
Not registered ROAs	Number of routes originated by the AS that are not covered by	Bad ROA Payload	Can ascertain degree of RPKI deployment (and compare	-IPv4 vs IPv6 -% of routes originated by AS	ISOC (RPKI Validator)	IP addresses days	Harm to self	Harm to self (packets won't go)	Medium - what's "bad"?	Count	Comparison between countries or comparison over time	Comparison between countries over time	If we measure days with 100% validity, then is this = 100-line

	any ROA in RPKI		registered vs non-registered ROAs to ensure it adds up to complete number of prefixes an AS originates)										above? Can be legit, but there are more secure approaches
Route Leak By AS	Number of incidents where the AS was the culprit of BGP leakage events.	Route problem	Helps determine degree of filtering	- IPv4 vs IPv6 - Total global route leaks	ISOC (bgpstream)	# of events	Harm to others	Harm to self (packets won't go)	Easy to take ISOC data, complexity in managing	Absolute count	Relative to other AS	Comparison between countries over time	
Route misorigination by the AS	Number of incidents where the AS was the culprit of BGP misorigination (hijacking) events.	Route problem	Helps determine degree of filtering	- IPv4 vs IPv6 - Total global route misoriginations	ISOC (bgpstream)	# of events	Harm to others	Harm to self (packets won't go)	Easy to take ISOC data, complexity in managing	Absolute count	Impacted IP in real AS	Comparison between countries over time	Adam spent a lot of time on the indicator type question for “Route misorigination by the AS”: it’s nearly indistinguishable (malice vs incompetence?) from an attack. I’ve come down on a new branch of running code, “propagate bad network” following this logic: It’s an

													action with direct security consequences for a downstream host that's trusting its upstream to route properly, and as a result of that trust is not crypto-protecting its communications. In that, it's like other direct attacks. It's not a communicable disease, there's no propagation of the attack code or control. It's not environmental, in that it doesn't impact the internet overall, and it' not lifestyle.
Route leak by a direct customer	Number of incidents where the AS was an accomplice (the	Route problem	Helps determine degree of filtering	- IPv4 vs IPv6 - Total global route leaks	ISOC (bgpstream)	# of events	Harm to others	Harm to self (packets won't go)	Easy to take ISOC data, complexity in managing	Absolute count	A = #/day; % of ISPs that do this on a given day	B = A / countries. C = change in A over time	

	misoriginating AS was present in the AS-PATH) to BGP leakage events.			by direct customers									
Route misorigination by a direct customer	Number of incidents where the AS was an accomplice (the leaking AS was present in the AS-PATH) to BGP hijack events.	Route problem	Helps determine degree of filtering	- IPv4 vs IPv6 - Total global route misoriginations by direct customer	ISOC (bgpstream)	# of events	Harm to others	Harm to self (packets won't go)	Easy to take ISOC data, complexity in managing	Absolute count	#/day; % of ISPs that do this on a given day	B = A / countries. C = change in A over time	
Bogon prefixes by the AS.	Number of incidents where the AS originated bogon address space.	Route problem	Helps determine degree of filtering	- IPv4 vs IPv6 - Total global incidents where AS originated bogons	ISOC (CIDR Report)	# of events	Harm to others	Harm to self (packets won't go)	Easy to take ISOC data, complexity in managing	Absolute count	#/day; % of ISPs that do this on a given day	B = A / countries. C = change in A over time	
Bogon prefixes propagated by the AS.	Number of incidents where the AS propagated bogon address space announcements received from its peers.	Route problem	Helps determine degree of filtering	- IPv4 vs IPv6 - Total global incidents where AS propagated bogons	ISOC (CIDR Report)	Incidents	Harm to others	Harm to self (packets won't go)	Easy to take ISOC data, complexity in managing	Count	#/day; % of ISPs that do this on a given day	B = A / countries. C = change in A over time	
Bogon ASNs by the AS	Number of incidents where the AS announced bogon	Route problem	Helps determine degree of filtering	- IPv4 vs IPv6 - Total global incidents where an	ISOC (CIDR Report)	Incidents	Harm to others	Harm to self (packets won't go)	Easy to take ISOC data, complexity in managing	Count	TBD	TBD	

	ASNs as adjacency.			AS announced bogon AS as adjacency									
Bogon ASNs propagated by the AS	Number of incidents where the AS propagated bogon ASNs announcements it received from its peers.	Route problem	Helps determine degree of filtering	- IPv4 vs IPv6 - Total global incidents where an AS propagated bogon ASNs announcements it received from its peers.	ISOC (CIDR Report)	Incidents	Harm to others	Harm to self (packets won't go)	Easy to take ISOC data, complexity in managing	Count	TBD	TBD	
Reserved IP Prefixes propagated by an AS	Measures whether filtering done effectively on reserved address space.	Route problem	Specific subset of anti-spoofing	IPv4 vs IPv6	Test to Be Constructed	Incidents	Harm to others	Harm to self (packets won't go)	TBD	TBD	TBD	TBD	

Domain Name Service (DNS)

Indicator	What indicator tells us	Why useful	Useful added context	Data sources	Indicator units	Impact model 3 (primary harm)	Impact model 3 (side effects)	Ease of measuring	Measurement characterization	First derived measures	Second derived measures	Notes
No of domains with DNSKEY Resource Records	Number of zones that have a public/private key pair associated with it	Ascertain level of DNSSEC deployment	Total number of domains in a country	SecSpider?	domains	Harm to self	Harm to others (can't validate DNS)	Easy	Count	Count by country?	Count by country over time	Is country by cc tld?

For each domain with a DNSKEY RR, the number of DNSKEY RRs	Whether multiple keys are valid and in use	Possibly ascertain how diligent key changes are, but that really requires time series data		SecSpider?	Keys	Harm to self	Harm to others (can't validate DNS)	Medium (need time series data)	A = # of keys in use. a[date1, date2, date 3]	a[date1, date2, date 3]	Analysis of array and rate of change	Adam comments: justification of "Can ascertain how diligent key changes are" but "1" might mean that they change the key every month, and then delete the old keys.
Key sizes and Algorithms used per public/private key pair	Key sizes and algorithms in prevalent use	Can ascertain whether outdated and insecure keying parameters used		SecSpider?	List of key sizes	Harm to self	Harm to self (slow down dns processing)	Easy	List of key size, algorithm	B = per key "BCP compliant" or not	B per domain; B per domain over time	How often do the BCPs change? How do we deal with keys that were ok last week and are not good this week?
No of domains with Resource Record Signature (RRSIG) Resource Records	How many domains are signed	Ascertain level of DNSSEC deployment	Total number of domains in a country	SecSpider?	As DNSKEY	Harm to self	None	Easy	Count	Count over time	None	
DNS authoritative and recursive services on separate devices	Whether an authoritative server also acts as recursive server	Susceptibility to fate sharing and can help ascertain lack of attention to good DNS hygiene practices		TBD	IP addresses	Harm to self	None	Hard (devices might be multi-homed)	Similarity of IP1 to IP2	Yes/no	# of matched bits?	
NSec records in use	DNSSec discipline			TBD	NSEC records	Harm to self	None	Easy	Count (NSEC 0 or 3)	Count by country		Combining 2 measures

Number of Lame Delegations	Will inform if a nameserver is delegated responsibility for providing nameservice for a zone (via NS records) but is not performing nameservice for that zone	Can help ascertain lack of attention to good DNS hygiene practices		TBD	yes/no per name server	Harm to self	None	Medium (need time series data)	A = # of NS records, B = reachable NS servers providing nameservice for that domain	C = Change in A,B per domain over time. D = B/A per domain	E = % of domains per country where D≠ 1. F = Avg D per country	(1) A domain can have a lot of NS records. If it has 2 and one is bad, that seems different than it having 4 and 2 being bad, even though both are 50%. (2) B combines two problems into one measure.

Email

Indicator	What indicator tells us	Why useful	Useful added context	Data sources	Indicator units	Impact model 3 (primary harm)	Impact model 3 (side effects)	Ease of measuring	Measurement characterization	First derived measures	Second derived measures	Notes
DMARC Implemented	To what extent domain has implemented DMARC (if at all). This determines ability to authenticate the authenticity of an email message.	Allows us to understand the % of domains using DMARC as a safeguard to prevent phishing/other scams	The list of domains at GCA is not exhaustive, so we would be relying on a sample	GCA	List (domain, DMARC string available)	Harm to others	Harm to self (break mail deliverability risk)	Easy with a list of domains	Yes/no	% domains in a "TLD"	Compare across "TLDs"	Measure by domain, not IP Is this p=none? Counting as harm to others because they can't validate your mail, turn up dmarc reliance
DMARC policy	Policies that pass "implemente	"none" is essentially a test state,	The list of domains at GCA is not	GCA	List (domain,	Harm to others	Harm to self (break mail	Easy with a list of domains	(unset, set, invalid); (set, ok,	% at each level	% at each level within a TLD	What about marketing domains

	d" include "none", "quarantine" and "reject"	and we shouldn't confuse it with q/p.	exhaustive, so we would be relying on a sample		DMARC status)		deliverability risk)		good, invalid)			without email outbound?
Servers that support STARTTLS	TLS - enabled email			Step 1. for each domain (domains) {mxs = gethostbyname(domain, mx) foreach mx (mxs) { starttls[domain, mx] = connect(mx, 587)}}	Mail server IPs	Harm to self	None	Easy	Count of servers with Start TLS; count of MX servers	Is it =100%?	% of domains in a TLD that support STARTTLS	Harm to self = your email isn't encrypted
SPF Implemented	Whether a domain is using SPF (yes/no) and if there are any errors associated with its implementation that need attention	Allows us to understand the % of domains using SPF as means of defining authorized senders	The list of domains at CGA is not exhaustive, so we would be relying on a sample	GCA	DNS domains	Harm to others	Harm to self (break mail deliverability risk)	Easy	Is there an SPF record?	% SPF in a TLD	% SPF in a TLD over time	Not clear what the starting point for measurement would be - are we starting from IP? a list of domains? SPF records can include something like +all which basically invalidates SPF. We should check to see if such things are common in the field.
SPF Errors	if there are any errors associated with its implementation that	Errors in SPF records can be detected by a variety of automated tools; not	The analytic tools vary, and what we detect will be influenced by various	TBD	DNS domains	Harm to others	Harm to self (break mail deliverability risk)	Collecting the SPF settings is easy. Deciding what's an error is	We can select one (or more) and simply declare that all metrics within it are	B = average score of domains within a TLD C = Variance of	D = compare between TLD	

	need attention	using those tools is probably correlated with other issues	tools. (See ease of measurement column)					harder. For example, compare the lists in https://mxtoolbox.com/spf.aspx https://dmarcian.com/spf-survey/ - the short answer is they're not the same.	even, and so assess "this domain meets 12 out of 15" items in the mxtoolbox list.	scores within a TLD		
MTA-STS	Whether a domain is using MTA-STS (yes./no). MTA-STS allows for security communication between mail servers (prevent man-in-the-middle type attacks).	Allows us to understand the % of domains using MTA-STA as a means of secure communications.	The list of domains at CGA is not exhaustive, so we would be relying on a sample	GCA	DNS domains	Harm to self	Harm to self (break mail deliverability risk)	Easy	Yes/no	as SPF	as SPF	

Certificates

Indicator	What indicator tells us	Why useful	Useful added context	Data sources	Indicator units	Impact model 3 (primary harm)	Impact model 3 (side effects)	Ease of measuring	Measurement characterization	First derived measures	Second derived measures	Notes
Digital certificate: % of certificates that expired and validity needed to be updated	Whether digital certificates which instantiate identity or give authorization are used while being invalid	Can inform where there is unpatched sw or where there are process gaps	How many invalid digital certificates are still utilized by user/application by 'trusting' it despite invalidity	Scanning	Count	Harm to self	None	Possibly hard - are we talking at instant of collection? How long invalid? --- AML: You'd have to directly challenge	List of invalid certs by advertised DNS name	A= invalid certs on www.domain B = Invalid certs as % of domain	A, B over time; A, B over TLD	Counting as a non-communicable disease because while it's worrisome, maybe the site is out of date and replaced,

								each cert in an area				and was left up, unmaintained, and so no cert renewal is needed? That's potentially also leaving vulnerable code running, but it might be not needed, maintained at a different level, etc. Counting as "harm to self" because it blocks people from reaching their website
Digital certificate: algorithm used to generate key pair	Key generating algorithms in prevalent use	Can ascertain whether outdated and insecure keying parameters used	Best current practice docs	Scanning	As DNSalgo	Harm to self	None	Mostly ok - what do we do with long lived keys? Evaluate anticipated state against current guidance?	As DNSalgo	As DNSalgo	As DNSalgo	
Digital certificate: key lengths used	Key sizes in prevalent use	Can ascertain whether outdated and insecure keying parameters used	Best current practice docs	Scanning	As DNSalgo	Harm to self	None	Mostly ok - what do we do with long lived keys? Evaluate anticipated state against current guidance?	As DNSalgo	As DNSalgo	As DNSalgo	
SSL/TLS Cert –	Whether SSL/TLS certificates	Identifies whether there are	None	Test public web servers with Qualys	Count	Harm to self	None	Easy	A = list of sites with	B = A/country	C=B/time; D = B relative	Assuming this is revocation

Expired Validity	which instantiate identity or give authorization are used while being invalid	poor certificate renewal practices and insecure implementations in use		like SSL tests?					TLS expired certs		to other countries	at the CA Revoked certs lead to people being used to revocation, and so this is like pollution. Marking as harm to self because the org is 'shooting themselves in the foot' because browsers will block access to their site.
SSL/TLS Cert – Self Signed	Whether the SSL/TLS certificate is self-signed	It is typically believed that 3rd party certificates are more trusted but that is debatable	None	Test public web servers with Qualys like SSL tests?	Count	Harm to self	None	Easy	As invalid certs	As invalid certs	As invalid certs	As invalid certs

Security protocols & services

Indicator	**What indicator tells us**	**Why useful**	**Useful added context**	**Data sources**	**Indicator units**	**Impact model 3 (primary harm)**	**Impact model 3 (side effects)**	**Ease of measuring**	**Measurement characterization**	**First derived measures**	**Second derived measures**	**Notes**
SSL / TLS protocol versions accepted for negotiation	Which version of protocol is accepted for use	Points to risk in potentially accepting outdated and insecure	Best current practice doc	Test public web servers with Qualys like SSL tests?	List	Harm to self	None	Easy to measure, harder to evaluate	For list (standards) acceptable/ not	A = % which are 100% ok B = domains which are 100% ok	A per domain, A over time	Potentially, there's conflicting advice out there ("is TLS 1.2 still acceptable?

		SSL/TLS versions										") NIST still accepts (properly configured) TLS 1.1 **https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-52r1.pdf** and good advice is to go to 1.3, but what do we test against?
SSL/TLS Cipher Suite Support	Which algorithms are supported in automated negotiations	Identifies whether vulnerable cipher suites can be negotiated for connection	Best current practice doc	Test public web servers with Qualys like SSL tests?	List	Harm to self	None	Getting a list, easy. Evaluating it requires a BCP evaluation	As protocol version	As protocol version	As protocol version	None; RC4
% HTTPS enabled web servers	How many web servers use cryptographically protected VPN access	Prevalence of cryptographically securing web access		scan based on DNS resolution	IPs	Harm to self	None (perception of speed risk)	Easy to measure, harder to evaluate	A=yes/no per server	As protocol version	As protocol version	Maybe it's a brochureware site, and https doesn't matter? Do we want enabled, or https-only as our measure?
SSH Version	Which secure shell is most prevalently used	Can ascertain whether outdated security software is used	BCP docs, version lists for things other than openssh	Scan IP4	IP, SSH version strings	Harm to self	None	Medium	A = List of known unsafe servers B = list of servers by domain	C = A/B is % of unsafe servers by domain	C over time, C by TLD	Can we get version in all cases? Can we interpret? For example, my mac is serving up SSH-2.0-OpenSSH_8.1 , my linux

												box is on SSH-2.0-OpenSSH_8.2p1 Ubuntu-4ubuntu0.2 ; is the Mac ok?

Tier 2

These are components/indicators that we ultimately chose not to include to measure due to certain selection criteria not being met

Open services															
Indicator	**What Indicator Tells Us**	**Why Useful**	**Useful Added Context**	**Data Sources**	**Impact Model II Indicator type**	**Indicator type**	**Ease of measuring**	**Measurement characterization**	**First Derived measures**	**Second derived measures**	**Complex or difficult to implement?** **Big tradeoff or risk in implementation?**	**Benefit/Harm to self/others**	**Action/Inaction**	**Notes**	
QOTD	Number of QOTD open port	Gives information on legacy protocol use and risk of amplification attack	- IPv4 vs IPv6 - % of all IP space in country	Shadow server	Pollution	*as CHARGEN*	Very easy				no	harm others			
Open NTP	Number of NTP servers that answer to any query	Can ascertain risk of utilizing NTP for amplifica	- IPv4 vs IPv6 - % of all NTP servers in country	CyberGreen	Pollution	*as DNS*	Very easy				no	harm others		like open DNS	

	(UDP 123)	tion attack													
LDAP	Number of LDAP servers supporting Connectionless LDAP [CLDAP]	Can ascertain risk of utilizing LDAP for amplification attack	-IPv4 vs IPv6 - Spoofing ability - Zero day patch	Shadow server	Pollution		very easy	N = account enumeration						Again, for impact model II, harm to others via ddos	
Email															
Indicator	**What Indicator Tells Us**	**Why Useful**	**Useful Added Context**	**Data Sources**	**Indicator units**	**Indicator type (Impact Model II)**	**Indicator type**	**Ease of measuring**	**Measurement characterization**	**First Derived measures**	**Second derived measures**	**Notes**	**Complex or difficult to implement?** **Big tradeoff or risk in implementation?**	**Benefit/Harm to self/others**	**Action/Inaction**
Secure Email Gateway Implemented								very hard				This seems hard to detect			
2FA Implemented								very hard				MFA for POP or IMAP could be scanned for but might require authenticated scans, but client implementation issues make it			

												hard to require, and			
Encryption & Digital Signatures								meaning unclear				Not clear what this refers to. DKIM signed? PGP signed/			
/* Email Archiving ? */															
/* A-V Scanning ? */															
DANE	Whether a domain is using DANE (yes/no). DANE allows for security communication between mail servers (prevent man-in-the-middle type attacks).	Allows us to understand the % of domains using DANE as a means of secure communications.	The list of domains at CGA is not exhaustive, so we would be relying on a sample	GCA (possible Internet.nl could help)											
TLS-RPT	Whether a domain is using MTA-STS (yes./no). TLS-RPT is a reporting mechanism for TLS and MTA-STS.	Allows us to understand the % of domains using TLS-RPT to review the status of mechanis	The list of domains at CGA is not exhaustive, so we would be relying on a sample	GCA				harm to self							

		ms that use TLS (such as MTA-STS)													
Credentials															
Indicator	**What Indicator Tells Us**	**Why Useful**	**Useful Added Context**	**Data Sources**	**Indicator units**	**Indicator type (Impact Analysis II)**	**Ease of measuring**	**Measurement characterization**	**First Derived measures**	**Second derived measures**	**Notes**	**Complex or difficult to implement?** **Big tradeoff or risk in implementation?**	**Benefit/Harm to self/others**	**Action/Inaction**	**Notes**
% of default passwords in use	How many Internet infrastructure devices use default passwords				count	lifestyle	Very hard without being intrusive, slow	passwords in list which show login success			*Do we have a plan to measure honeypots? (Or is this that plan? :))*				
% of multifactor authentication enabled	Uptake of multifactor authentication	Can ascertain susceptibility to impersonation		Can this be measured?			How to measure unclear?								
Username/Password Combinations		Known cred leakage per domain used in email section		Dark Web											
Security protocols/services															

Indicator	**What Indictor Tells Us**	**Why Useful**	**Useful Added Context**	**Data Sources**	**Indicator units**	**Indicator type**	**Measurement characterization**	**First Derived measures**	**Second derived measures**	**Notes**	**Complex or difficult to implement?** **Big tradeoff or risk in implementation?**	**Benefit/Harm to self/others**	**Action/Inaction**	**Notes**	
SSH service															
IPsec – IKEv2 vs IKEv1	Whether outdated key negotiation algorithms used	Can ascertain whether outdated security software is used													
SSL/TLS certificate revocation information exists	That there is an existing mechanism in place to revoke certificate	If a credential is lost to access secret key to create a revocation ability then a certificate may never get revoked		?? Test public web servers with Qualys like SSL tests?	debatable ciphers supported										
Legacy Protocols															

Indicator	What Indicator Tells Us	Why Useful	Useful Added Context	Data Sources											
Telnet	Number of devices with Insecure protocol enabled that can leak cleartext passwords (TCP 23)	Helps ascertain risk of using outdated protocol for leaking cleartext passwords	- IPv4 vs IPv6 - % of all IP address space in country	Shadowserver											
Ftp	Number of devices with Insecure protocol enabled that can leak cleartext passwords (TCP 21)	Helps ascertain risk of using outdated protocol for leaking cleartext passwords. Also whether used in anonymous mode.	- IPv4 vs IPv6 - % of all IP address space in country	Shadowserver											
TFTP	Number of devices with Insecure protocol enabled that can leak cleartext passwords (UDP 69)	Helps ascertain risk of using outdated protocol for leaking cleartext passwords	- IPv4 vs IPv6 - % of all IP address space in country	Shadowserver											
Operating systems/soft															

ware versions															
Indicator	**What Indicator Tells Us**	**Why Useful**	**Useful Added Context**	**Data Sources**											
DNS Recursive Resolver OS versions															
DNS Authoritative Server OS versions															
Router OS versions															
Email Server OS versions															
NTP Server OS versions															
RADIUS OS versions				Most likely not publicly measurable											
TACACS versions				Most likely not publicly measurable											

Network															
Anti-spoof configured	Active anti-spoof capability	Can determine degree of attention to anti-spoofing	-IPv4 vs IPv6 -Total number of IP space in country	Spoofer Project (CAIDA)	Count										
Network Disconnections	Measures when ASs are disconnected from global routing table														
# of country level IXs	Whether routed traffic stays local where possible		Can determine readiness of a catastrophic Internet cut-off		Observed thru Euro-IX database										
DNS															
Hash functions used in RRSIGs	Cryptographic signing functions in prevalent use	Can ascertain whether outdated and insecure functions used		SecSpider?	As keysize	lifestyle									
No of domains with Delegation Signer (DS) Resource Records	Whether signed zone is linked to established chain of trust	Ascertain level of DNSSEC deployment	Total number of domains in a country	SecSpider?	As DNSKEY	lifestyle									
DNSKEY set available	That a resolver can obtain a	Can create metric for availabilit	How many varying resolvers	SecSpider?	As DNSKEY	lifestyle									

	zone's DNSSKEY set.	y re DNSSEC and whether zone's key set is available via any authoritative server	can reach a zone's key-set												
	NSEC0	The NSEC record is used to prove that something really does not exist, by providing the name before it, and the name after it.	Allows for a proof of non-existence for record types. If you ask a signed zone for a name that exists but for a record type that doesn't (for that name), the signed NSEC record returned lists all of the record types that do exist for the requested domain name.				lifestyle								
	NSEC3	The NSEC record is used to prove that something really	Same as NSEC0 but adds some privacy aspects since the				lifestyle								

		does not exist, by providing the name before it, and the name after it.	names are run through a one-way hash, before giving it out, so the recipients can verify the non-existence, without any knowledge of the actual names.												
# of Root server instances in a country	Whether there is root server coverage in a country	Can determine readiness of a catastrophic Internet cut-off			Count?		healthy lifestyle								

ACKNOWLEDGEMENTS

This report was prepared by the CyberGreen Institute.

CyberGreen is a global non-profit and collaborative organization that serves the global public benefit by supporting a more resilient and healthier global Internet Ecosystem. CyberGreen is a trusted player in that Ecosystem following transparent ways of working, and identifying sources of risk and best practices for the community. We are committed to evidence-driven metrics and measurements.

Contributors and authors:
Merike Kaeo
Adam Shostack

We would like to express our gratitude to the following individuals for their input and advice:

Lurong Chen
Michael Collins
Leslie Daigle
Matt Ford
Shawn Hernan
Adam Lange
Franziska Lichtblau
Yoshinobu Matsuzaki
Shehzad Mirza
Eric Osterweil
Phil Reitinger
Andrei Robachevsky
Vyas Sekar
Joe St Sauver
Michael Tanji