CHRONIC [CYBER] PAIN

Exposed & Misconfigured Databases in the Healthcare Industry

INTSIGHTS
Threat Intelligence Realized.

RESEARCH REPORT

Ironically, the healthcare industry isn't being very healthy. There's an epidemic taking place across healthcare organizations, and it's time for a check-up.

We've seen healthcare organizations become increasingly targeted by threat actors over the past few years, a trend that almost every industry is facing these days. There's a lot of talk around medical device vulnerabilities and ransomware attacks against healthcare organizations, but their most sought-after asset is their data. With new data privacy and compliance laws in place, patient data has become increasingly valuable, and increasingly difficult to protect. Furthermore, this data often can't be changed or mitigated once it's exposed. People can't change their Date of Birth or health conditions the way they can change their credit card number, so the fallout from these attacks can go on for many years after the initial breach.

As healthcare organizations attempt to move data online and increase accessibility for authorized users, they've dramatically increased their attack surface, providing cybercriminals with new vectors to steal ePHI. Yet, these organizations have not prioritized investments in cybersecurity tools or procedures. Healthcare budgets are tight, and if there's an opportunity to purchase a new MRI machine versus make a new IT or cybersecurity hire, the new MRI machine often wins out. Healthcare organizations need to carefully balance accessibility and protection.

HEALTHCARE ORGANIZATIONS NEED TO CAREFULLY BALANCE ACCESSIBILITY AND PROTECTION

In this research report, we will explore a key area of the healthcare attack surface, which is often the easiest to avoid—exposed databases. It's not only old or outdated databases that get breached, but also newly established platforms that are vulnerable due to misconfiguration and/or open access.

METHODOLOGY

For this report, we wanted to understand how easy it is for someone to search for and find vulnerable or exposed healthcare databases. While many other industries suffer from similar exposures, healthcare organizations are particularly at risk because of the sensitivity of ePHI and medical data.

For our research, we chose a couple of popular technologies used for handling medical records, including known and widely used commercial databases, legacy services still in use today, and new sites or protocols that try to mitigate some of the vulnerabilities of past methods. We wanted to demonstrate that you can easily find access to sensitive data in each state: at rest, in transit, or on in use.

The tactics we used were pretty simple: Google searches, reading technical documentation of the aforementioned technologies, subdomain enumeration, and some educated guessing about the combination of sites, systems and data. All of the examples presented here were freely accessible, **and required no intrusive methods to obtain**. Simply knowing where to look (like the IP address, name or protocol of the service used) was often enough to access the data.

EXPOSED HEALTHCARE SYSTEMS: KEY FINDINGS

In our research, we found numerous examples of exposed healthcare systems and databases, using a variety of tactics. We tried to focus on systems that hold patient data, as these DBs are the most important to protect.

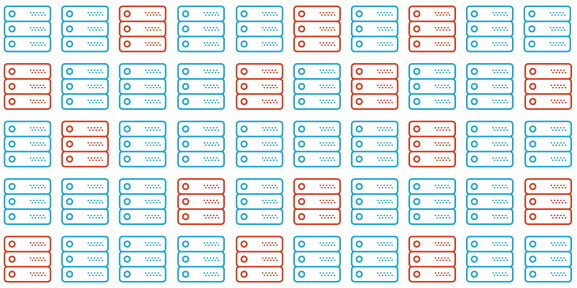
Before we show some examples, let's do a brief summary of our findings using some estimated (and very rough) math.

FINDINGS & ACTIVITY SUMMARY

The following table contains a high-level summary and estimated metrics for our findings from this research to provide context for how hackers may perform similar activities.

Total Research Hours

90 Hours



50 Total Databases Evaluated

15 Total Databases Found Exposed

Estimated Rate of Exposed Databases

30%

Total Exposed Records Discovered

1.5 Million

Records Discovered Per Hour

16,667
Medical Records / Hour

Estimated Black Market Price Per Medical Record

$1

Per Record

Annual Hacker Salary
(40 Hours/Week, 50 Weeks/Year)

$33+Million

Let's provide some commentary for some of these numbers.

ESTIMATED RATE OF EXPOSED DATABASES: 30%

Although our findings were not statistically significant, our rate of 30% is fairly consistent with what we're seeing across all industries for exposed assets. In our research report, DevOps Beware: Your Servers Are Open for Business, we found that over 23% of DevOps servers were openly accessible via the web.

TOTAL EXPOSED RECORDS: 1.5 MILLION

One of the databases we found in our search contained 1.3 million records, which is a relatively large database for this kind of searching. Therefore, our total figure may be a bit exaggerated, but the number is impressive regardless. Even if it is exaggerated, hackers can find a large number of records in just a few hours of work, and this data can be used to make money in a variety of ways.

ANNUAL SALARY: $33+ MILLION

This number assumes that a hacker can continue to find records at a rate of 16,667 per hour and works 40 hour weeks for the entire year. While this number may also be exaggerated, it shows the potential for cybercriminals to make large amounts of money using this data.

It's also important to note that ePHI and medical data is harder to make money with compared to other data, like credit card info. Cybercriminals tend to be lazy, and it's much quicker to try using a stolen credit card to make a fraudulent purchase than to buy ePHI data and run a phishing or extortion campaign. This may lessen the value of ePHI data in the eyes of some cybercriminals; however, ePHI data has a longer shelf-life and can be used for more sophisticated and more successful campaigns.

ATTACKER MOTIVATIONS

Threat actors usually follow the money. Where there is a chance to make a quick buck, cybercriminals will flock. But healthcare organizations don't typically hold any type of currency. Attackers who target the industry usually do it for one of three reasons:

1. **State-Sponsored APTs Targeting Critical Infrastructure:** APTs are more sophisticated and are usually more difficult to stop. They will attempt to infiltrate a network to test tools and techniques to set the stage for a larger, future attack, or to obtain information on a specific individual's medical condition.
2. **Attackers Seeking Personal Data:** Attackers seeking personal data can use it in multiple ways. They can create and sell ePHI lists (called Fullz), they can blackmail individuals or organizations in exchange for the data (see Figure 1), or they can use it as a basis for further fraud, like phishing, Smishing, or scam calls.
3. **Attackers Taking Control of Medical Devices for Ransom:** Attackers targeting vulnerable infrastructure won't usually target healthcare databases, but will target medical IT equipment and infrastructure to spread malware that exploits specific vulnerabilities and demands a ransom to release the infected devices. Since medical devices tend to be updated infrequently (or not at all), this provides a relatively easy target for hackers to take control.

Each of these scenarios pose a risk of its own, and are driven by different motivations. Therefore, hackers will use different tactics in each scenario to accomplish their goal.

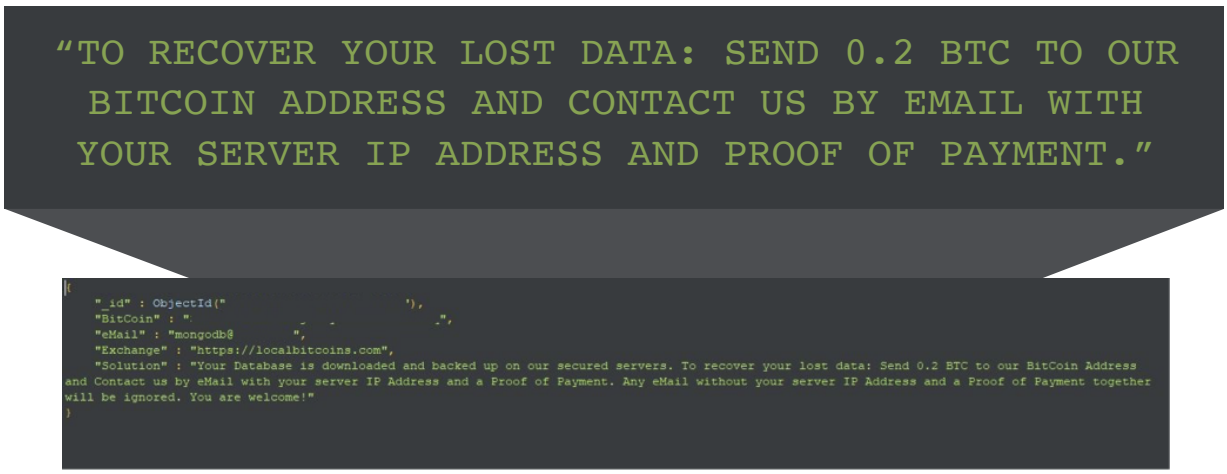


Figure 1: Hacker Note Requesting Ransom to Get DB Back

CHRONIC ISSUES: DATA MANAGEMENT AND COLLABORATION

A common initiative across all healthcare organizations is increasing the level of collaboration and information sharing. Making data accessible between local clinics, physicians, hospitals and health insurance providers helps speed up treatments and gives doctors important background information to make crucial medical decisions.

But with this data-sharing comes a multitude of security issues. As you increase access to data, you increase your exposure to unauthorized access. As we've seen in our research, many healthcare organizations are struggling to manage all the different access levels and exposure points of their databases.

There are two main factors contributing to this challenge.

1. **Varying Software Solutions for Data Storage and Sharing:** Each solution has its own set of access levels, defaults, implementations and configurations. As organizations adopt more technologies, it becomes harder to map your digital exposure (a.k.a your digital footprint). Many of these systems are built using off-the-shelf databases, such as Elasticsearch or MongoDB, and the access to these DBs is often misconfigured. Even if organizations keep the data in a local DB, they still need to allow access to the data through shared protocols or web APIs to collaborate. Managing the access credentials is a big task for an organization that needs to collaborate with dozens, or even hundreds of other organizations. A leak of one API key from one partner organization, can compromise an entire patient DB.
2. **Data Standardization to Increase Ease of Sharing:** Healthcare organizations have begun using known and agreed upon tags, symbols and technology to encourage fluent communication between all these systems. Standardization is a good thing in general, but it makes it easier to understand protocols, databases, and communication. Having a basic understanding of how these systems communicate with each other makes it easy to know what to look for.

These two factors create a wide array of challenges that make it very hard to find the right balance of accessibility and protection.

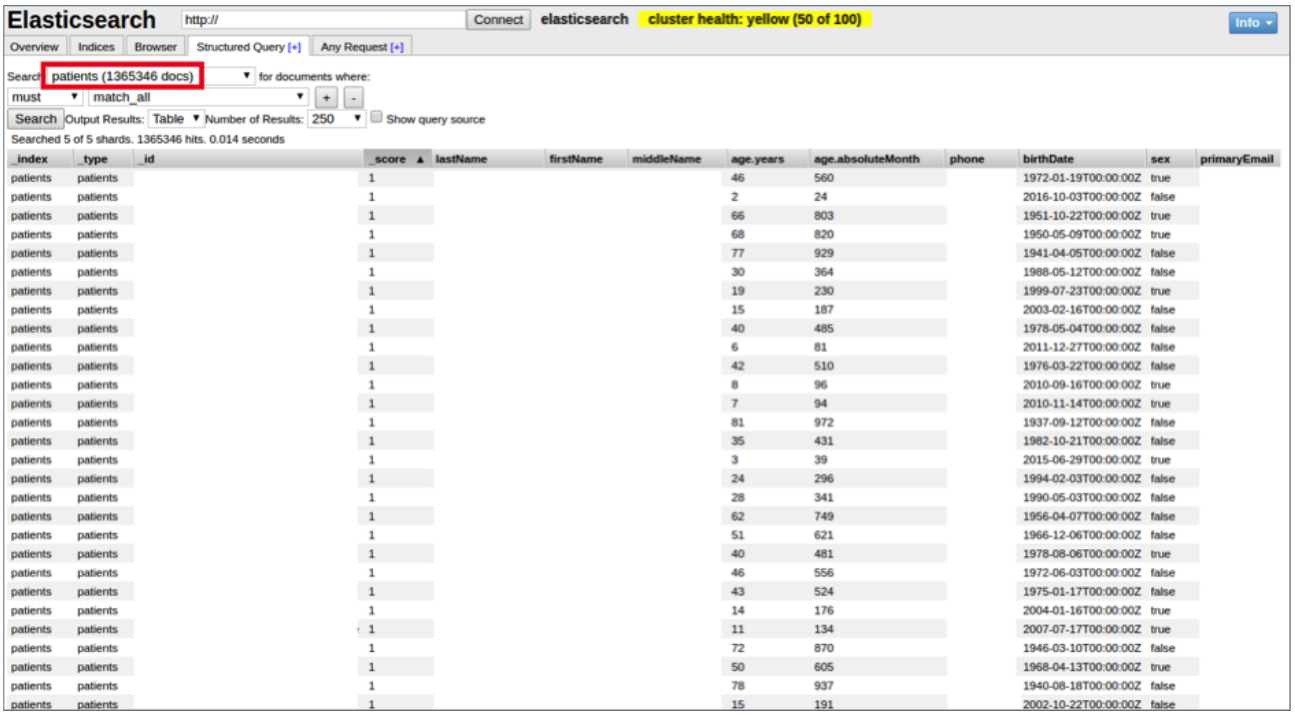
EXPOSED DATABASE EXAMPLES

MAJOR REGIONAL CLINIC: ELASTICSEARCH EXAMPLE

Elasticsearch is a known and widely used database. If its not protected by a third-party security pack or hidden from the public web on an internal network, you can access it just by knowing the IP address of the server.

Tools such as Shodan allow you to search for open databases. Since medical data usually contains specific words or objects, like physician, patient, diagnosis, disease, etc., it is very easy to identify medical DB's from other types of data.

Using Shodan, we searched for Elasticsearch databases and port numbers to try to identify openly accessible healthcare databases. After just a few hours of searching using this technique, we found this database (Figure 2) from a major regional clinic of a big European capital. This database contains over 1.3M patient records.



_index	_type	_id	_score	lastName	firstName	middleName	age.years	age.absoluteMonth	phone	birthDate	sex	primaryEmail
patients	patients		1				46	560		1972-01-19T00:00:00Z	true	
patients	patients		1				2	24		2016-10-03T00:00:00Z	false	
patients	patients		1				66	803		1951-10-22T00:00:00Z	true	
patients	patients		1				68	820		1950-05-09T00:00:00Z	true	
patients	patients		1				77	929		1941-04-05T00:00:00Z	false	
patients	patients		1				30	364		1988-05-12T00:00:00Z	false	
patients	patients		1				19	230		1999-07-23T00:00:00Z	true	
patients	patients		1				15	187		2003-02-16T00:00:00Z	false	
patients	patients		1				40	485		1978-05-04T00:00:00Z	false	
patients	patients		1				6	81		2011-12-27T00:00:00Z	false	
patients	patients		1				42	510		1976-03-22T00:00:00Z	false	
patients	patients		1				8	96		2010-09-16T00:00:00Z	true	
patients	patients		1				7	94		2010-11-14T00:00:00Z	true	
patients	patients		1				81	972		1937-09-12T00:00:00Z	false	
patients	patients		1				35	431		1982-10-21T00:00:00Z	false	
patients	patients		1				3	39		2015-06-29T00:00:00Z	true	
patients	patients		1				24	296		1994-02-03T00:00:00Z	false	
patients	patients		1				28	341		1990-05-03T00:00:00Z	false	
patients	patients		1				62	749		1956-04-07T00:00:00Z	false	
patients	patients		1				51	621		1966-12-06T00:00:00Z	false	
patients	patients		1				40	481		1978-08-06T00:00:00Z	true	
patients	patients		1				46	556		1972-06-03T00:00:00Z	false	
patients	patients		1				43	524		1975-01-17T00:00:00Z	false	
patients	patients		1				14	176		2004-01-16T00:00:00Z	true	
patients	patients		1				11	134		2007-07-17T00:00:00Z	true	
patients	patients		1				72	870		1946-03-10T00:00:00Z	false	
patients	patients		1				50	605		1968-04-13T00:00:00Z	true	
patients	patients		1				78	937		1940-08-18T00:00:00Z	false	
patients	patients		1				15	191		2002-10-22T00:00:00Z	false	

Figure 2: Exposed Elasticsearch Database Containing 1.3M Patient Records

LOCAL CLINIC: MONGODB EXAMPLE

Another popular database program is MongoDB. Just like Elasticsearch, it is a generic DB that's used by various industries, one of them being healthcare organizations. Using search techniques similar to the Elasticsearch example above, we found the following patient database from a Canadian clinic openly accessible (Figure 3).

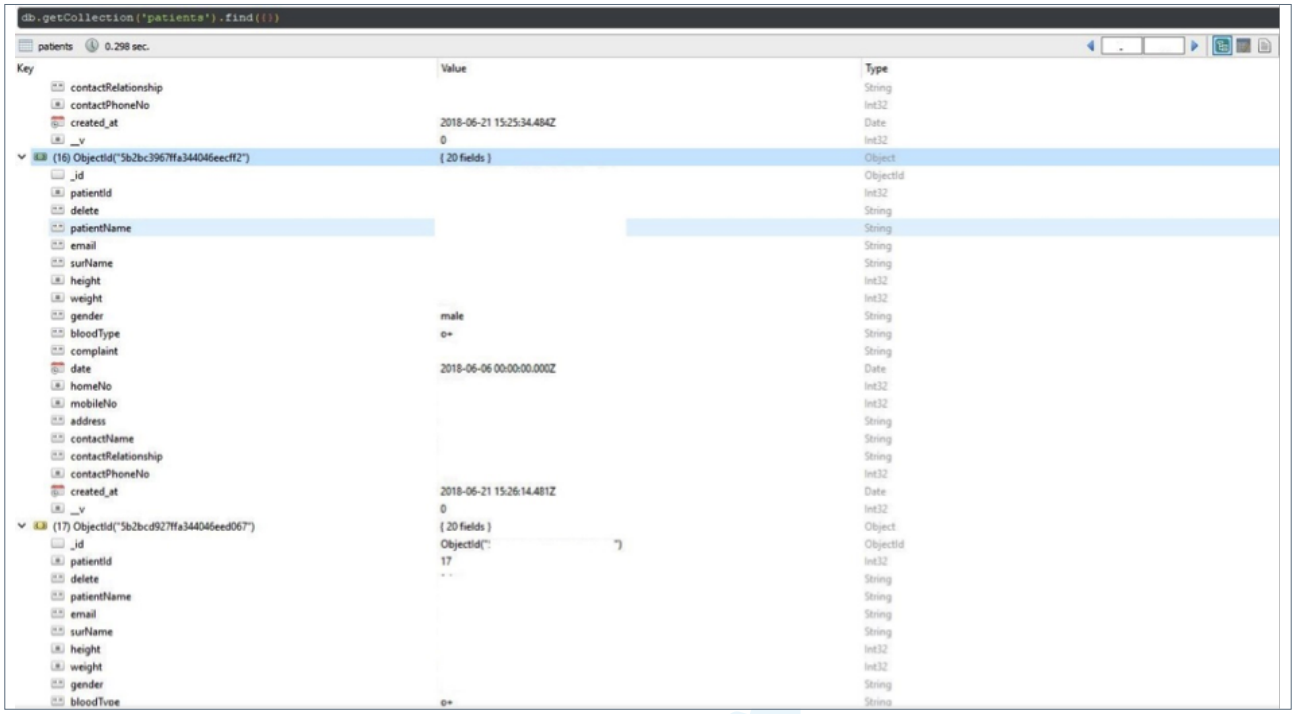


Figure 3: Exposed Canadian Clinic Patient DB Built on MongoDB

HOSPITAL: SMB SERVICE EXAMPLE

SMB Service is a file sharing protocol, usually used to access files within an internal network. It is most-known because of the WannaCry ransomware attacks, which utilized (and still can today) a zero-day flaw in V1 of the protocol to attack thousands of organizations, many of which were in the healthcare industry.

SMB has poor security features and should not be exposed to the public web. With that said, many old backup services use SMB to communicate and transfer files. Many of these services are kept as legacy services and are easily accessible from the web.

It seems that even one and a half years after the initial WannaCry outbreak, organizations are still exposing their SMB service to the web. Here is a list of directories from an Asian hospital that we found using this SMB service (Figure 4). Note: directories such as “Imaging_Xray-CArm”, “Critical_Care-Patients-Pumps-Anesthesia Disk” and “Cardiology_ECG-Defibrillator-TMT” contain very intimate and dangerous information and images.

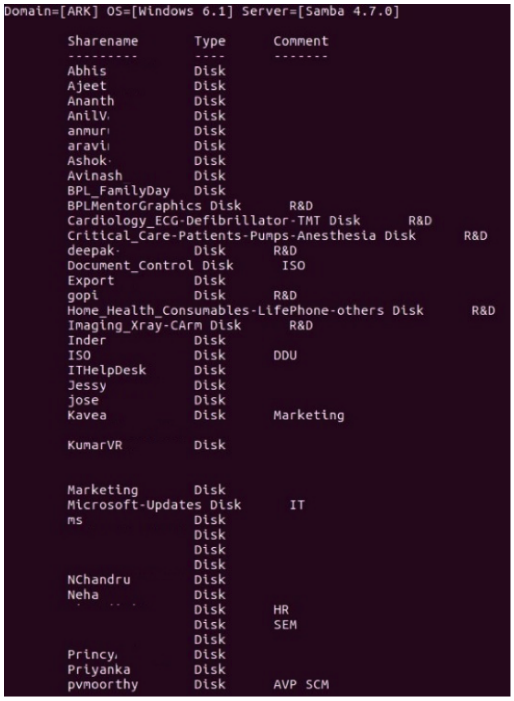


Figure 4: Exposed SMB Service

HOSPITAL: FTP SERVICE EXAMPLE

FTP is a very old and known way to share files across the Internet. It is also a scarcely protected protocol that has no encryption built in, and only asks you for a username and password combination, which can be brute forced or sniffed by network scanners very easily.

Here we found a hospital in the US that has its FTP server exposed (Figure 5). FTP’s usually hold records and backup data, and are kept open to enable backup to a remote site. It could be a neglected backup procedure left open by IT that the hospital doesn’t even know exists.

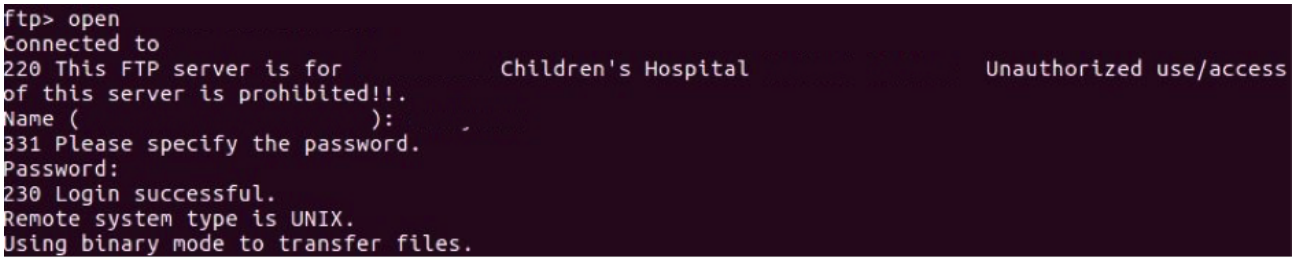


Figure 5: A US Children’s Hospital Open FTP site

CDR INTEGRATION SERVICES: FHIR PROTOCOL EXAMPLE

FHIR stands for “Fast Healthcare Interoperability Resources” and it is a communication protocol and standard for exchanging electronic health records. FHIR is essentially a server that healthcare organizations setup to make their medical records accessible to other organizations in a simple and uniform way.

FHIR has a number of security measures in place, and you do need an API key to access it. But healthcare organizations still misconfigure this service, making their medical records publicly available. Even with security measures in place, sometimes the number of clients that access this service makes it very hard to track who has the API key to access the server.

Another fault is its lack of leveled privileges between different clients, which gives anyone with an API key access to the whole DB, regardless of whether they are allowed to view other patient data or not. This is also a problem as it is hard to track the specific API keys handed out to clients, and it enlarges the exposure of the patients DB to third-party compromises.

API access is simple and uniform. Knowing the standard fields used by the FHIR protocol makes it very easy to scan and extract complete databases with simple HTTP requests. A lot of CDR (Clinical Data Repository) systems uses the FHIR protocol to make data accessible, and a poorly implemented solution can expose whole databases of patient information including: patient name, photo, email, height, weight, diagnosis, medical history, medical conditions, treatments and drug prescriptions, which are all sensitive data points that cannot be changed after they have been exposed (Figure 6).

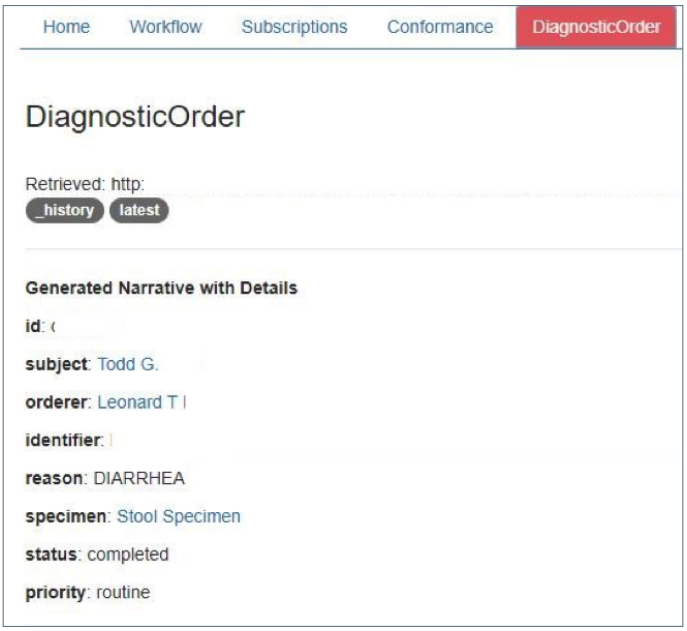


Figure 6: Patient Diagnosis from an Exposed FHIR Server

3RD PARTY PATIENT MANAGEMENT SERVICES EXAMPLE

For smaller clinics that don't have the budget to setup a local EHR system, there are a number of cloud hosted solutions. These solutions allow multiple clinics to login to a site in which they will run their whole operation.

But some of these sites are poorly secured. Some of them use an easy-to-hack username and password combination, without any multi-factor authentication. Furthermore, some of the sites have vulnerabilities that allow simple access to their management console.

Here we have the site of a Texas EHR system, which suffered from a web vulnerability that allows anyone who knows the URL to the admin console access to the entire system database (Figure 7).

The site holds data on thousands of patients and dozens of physicians. The system also holds the data of dozens of clinics, including patient list, physicians list, patient diagnosis, DWC forms, fax and email correspondence, treatments, and billing information (Figure 8).

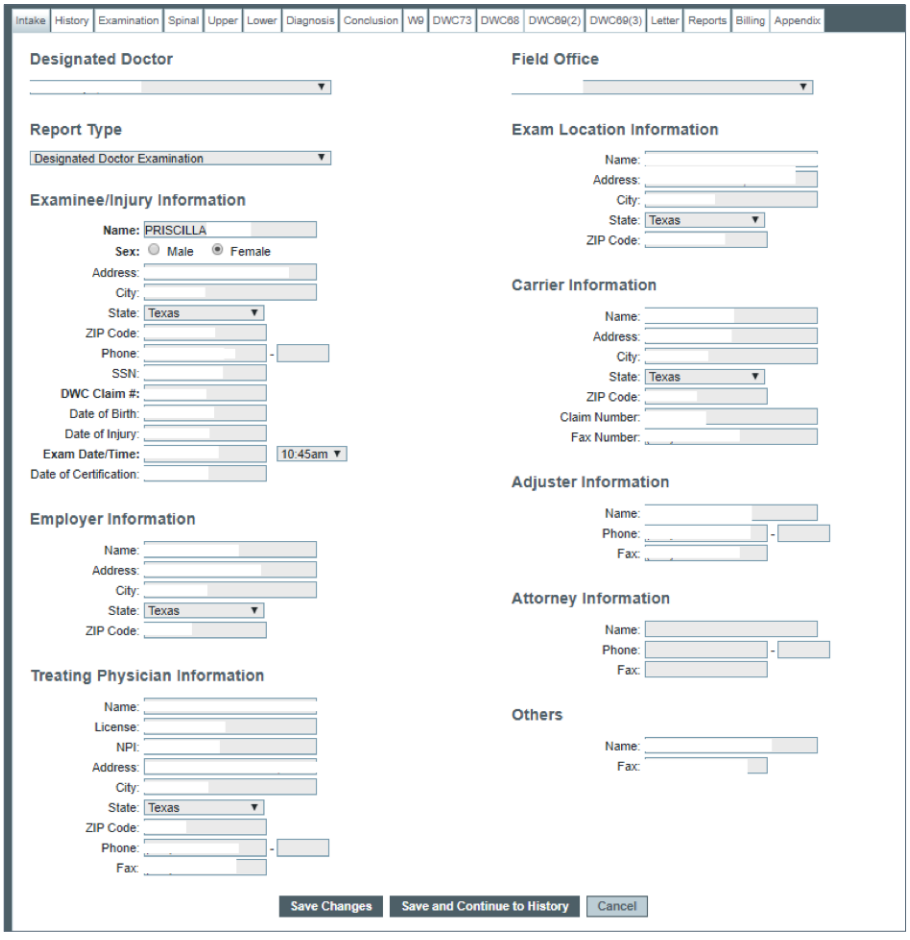


Figure 7: Exposed EHR System Admin Console

UPDATED HISTORY OF PRESENT INJURY/SUMMARY OF RECORDS

The examinee reports that she was mopping and slipped and fell, hurting her right arm, head, shoulder, back, right side of body/trunk and hip.

03/07/16 Texas Workers' Compensation Work Status Report, M.D., has prevented and still prevents the employee from returning to work as of 03/07/16 through 03/28/16.

03/16/16 CT of cervical spine Emergency Center, , M.D. Impression: 1. No acute intracranial abnormality. 2. No acute cervical spine abnormality.

03/16/16 X-ray of shoulder, Emergency Center, M.D. Impression: Intact right shoulder.

06/15/16 MRI of cervical spine, One Step Diagnostic, , M.D. Impression: 1. Straightening of the usual cervical lordosis. This could be an effect of muscle spasm. 2. Hemangioma in the body of C6. 3. Multilevel spondylosis. 4. Central disc protrusions at C5-6 and C7-T1.

06/15/16 MRI of shoulder, One Step Diagnostic, M.D. Impression: 1. Mild acromioclavicular impingement with retracted full-thickness tear of the distal supraspinatus tendon with free fluid in the subacromial and subdeltoid bursa. Mild partial-thickness involvement of the infraspinatus and subscapularis tendons noted. 2. Biceps tendinitis with likely partial-thickness tear at the level of the humeral neck. Moderate effusion noted.

07/25/16 Procedure, Reconstructive Houston, PA, M.D. Procedure: Arthroscopic right distal clavicle resection (Mumford type). Primary right shoulder arthroscopic rotator cuff repair. Primary limited arthroscopic debridement of right shoulder.

11/15/16 Texas Workers' Compensation Work Status Report, will allow the employee to return to work as of 11/15/16 with the restrictions through 12/06/16.

INTERIM HISTORY

Relevant medical records are discussed below in the EXTENT OF COMPENSABLE INJURY section.

UPDATED EXAMINEE COMPLAINTS

The examinee complains of neck pain that radiates down into her right shoulder to about the mid biceps area, but at times down into her hand. Overall she states her response to treatment has been positive and her condition improved. She stated that the stabbing pain is gone but she is still experiencing throbbing pain, especially with increased activity. The pain is worse when driving, mopping, lifting things and the other things she used to do before her injury. Putting her arm in a pillow lessens the pain as well as her lidocaine patches. She rates her current pain as 7 out of 10, with 10 being the highest level of pain; the pain has been consistently 7-8 out of 10 on a daily basis for the past 6 months. She stated with normal activities, such as those she did before her injury, she will experience numbness and weakness from the neck down into her right hand.

Figure 8: Exposed Medical History Record

RECOMMENDATIONS

As we can see from this research, healthcare organizations are not doing a very good job of protecting their patient data. With simple search techniques and a basic understanding of how these systems work, you can find an endless amount of ePHI data.

Healthcare organizations must understand their digital assets and which systems may be exposed. Each system is different, so you should refer to its specific security features and settings to ensure they are secured properly. However, here are a few general best practices for evaluating if your data is exposed and/or at risk.

1. **Use Multi-Factor Authentication for Web Applications:** If you're using a system that only needs a username and password to login, you're making it significantly easier to access. Make sure you have MFA setup to reduce unauthorized access.
2. **Tighter Access Control to Resources:** Limit the number of credentials to each party accessing the database. Additionally, limit specific parties' access to only the information they need. This will minimize your chance of being exploited through a 3rd party, and if you are, will limit the damage of that breach.
3. **Monitor for Big or Unusual Database Reads:** These may be an indication that a hacker or unauthorized party is stealing information. It's a good idea to setup limits on database reads and make sure requests for big database reads involve some sort of manual review or confirmation.
4. **Limit Database Access to Specific IP Ranges:** Mapping out the organizations that need access to your data is not an easy task. But it will give you tighter control on who's accessing your data and enable you to track and identify anomalous activity. You can even tie specific credentials to specific IP ranges to further limit access and track strange behavior more closely.
5. **Use 3rd Party Intelligence and Pen-Testing Services:** Using a hacker's point of view can help you understand where you are vulnerable and weak. These intelligence and testing services enable you to view your organization like an attacker would, so you can prioritize and lock down access to sensitive data.

ABOUT THE ANALYST: ARIEL AINHOREN

Ariel Ainhoren is a Security Researcher at IntSights, focused on discovering new cyber trends, threats, hacker strategies and vulnerabilities. He is a seasoned security professional with over 8 years of experience in the cyber industry, with expertise in computer forensics, malicious programs, vulnerability management and Microsoft Products. Ariel enjoys solving cyber puzzles, preferably byte by byte.

ABOUT INTSIGHTS

IntSights is redefining cyber security with the industry's first and only enterprise threat management platform that transforms tailored threat intelligence into automated security operations. Our ground-breaking data-mining algorithms and unique machine learning capabilities continuously monitor an enterprise's external digital profile across the surface, deep and dark web, categorize and analyze tens of thousands of threats, and automate the risk remediation lifecycle — streamlining workflows, maximizing resources and securing business operations. This has made IntSights' one of the fastest growing cyber security companies in the world. IntSights has offices in Tel Aviv, Amsterdam, New York and Dallas and is backed by Glilot Capital Partners, Tola Capital, Blumberg Capital, Blackstone and Wipro Ventures.

To learn more, visit www.intsights.com.

INTSIGHTS

Threat Intelligence Realized.