Centre universitari adscrit a la
Universitat Pompeu Fabra Barcelona

Grau en Enginyeria Informàtica de Gestió i Sistemes d'Informació

ANÀLISI DE RISCOS DE CIBERSEGURETAT EN L'ENTORN MÈDIC

Memòria

Jordi Caylà Bayona

Tutor: Dr. Pere Tuset-Peiró

Curs 2024/2025

TecnoCampus
Mataró-Maresme

Dedicatòria

Diuen que quan algú mor, tornarà a néixer després d'haver comptat totes les estrelles del firmament. No sé quantes en deus portar, pare, però mentrestant, aquest treball va per tu.

Agraïments

Al Dr. Pere Tuset, per una tutoria excepcional. Al Dr. Michael Pilgermann i al seu equip d'estudiants, per participar en aquest projecte. A la meva mare, per ser la llum quan només hi havia tempesta. Al meu germà, per ser la tempesta (però la d'estiu, la que apaga els focs). A l'Anna per corregir aquest treball. A la Gisela, el Pol, el Rems, la Lea, la Mar i tots els meus companys, per aquests 4 anys plens d'aprenentatge. A l'Aya, pel suport incondicional.

Abstract

This project analyzes the state of cybersecurity in medical environments using OSINT techniques, focusing on medical devices connected to the Internet. An automated pipeline was developed for data collection, processing and analysis from Censys and Shodan, applied across twenty European countries. The results reveal thousands of Internet exposed devices with critical vulnerabilities, some of which are publicly accessible without authentication. The project highlights the lack of security measures in medical systems and proposes actions to enhance the protection of these essential infrastructures.

Resum

Aquest projecte analitza l'estat de la ciberseguretat en entorns mèdics a través de tècniques OSINT, centrant-se en dispositius mèdics connectats a Internet. S'ha desenvolupat una *pipeline* automatitzada per a l'obtenció, tractament i anàlisi de dades de Censys i Shodan, aplicant-la a vint països europeus. Els resultats evidencien milers de dispositius exposats a Internet amb vulnerabilitats crítiques, alguns dels quals són accessibles públicament sense autenticació. El projecte destaca la manca de mesures de seguretat en sistemes mèdics i proposa línies d'actuació per millorar la protecció d'aquestes infraestructures essencials.

Resumen

Este proyecto analiza el estado de la ciberseguridad en entornos médicos mediante técnicas OSINT, centrándose en dispositivos médicos conectados a Internet. Se ha desarrollado una *pipeline* automatizada para la obtención, tratamiento y análisis de datos de Censys y Shodan, aplicándola a veinte países europeos. Los resultados evidencian miles de dispositivos expuestos a Internet con vulnerabilidades críticas, algunos de ellos accesibles públicamente sin autenticación. El proyecto pone de relieve la falta de medidas de seguridad en los sistemas médicos y propone líneas de actuación para mejorar la protección de estas infraestructuras esenciales.

Índex

Índex de figures

Glossari de termes

ARPANET Advanced Research Projects Agency Network (Xarxa de l'Agència de Projectes d'Investigació Avançada)

API Application Programming Interface (Interfície de Programació d'Aplicacions)

Banner Informació que un servei proporciona quan es connecta un client, sovint usada per identificar serveis i versions.

CVE Common Vulnerabilities and Exposures (Vulnerabilitats i Exposicions Comunes)

CVSS Common Vulnerability Scoring System (Sistema de Puntuació de Vulnerabilitats Comunes)

DICOM Digital Imaging and Communications in Medicine (Imatges i Comunicacions Digitals en Medicina)

DNS Domain Name System.

DPing DICOM Ping

EMR/EHR Electronic Medical Records/Electronic Health Records (Registres Mèdics Electrònics/Registres de Salut Electrònics)

FPDF Free PDF (Llibreria per a la creació de documents PDF)

GDPR General Data Protection Regulation (Reglament General de Protecció de Dades)

Honeypot	Parany virtual dissenyat per detectar i prevenir activitats malicioses en xarxes informàtiques
HTML	HyperText Markup Language (Llenguatge de Marcat d'Hipertext)
HTTPS	HyperText Transfer Protocol Secure (Protocol de Transferència d'Hipertext Segur)
HTTP/2	HyperText Transfer Protocol versió 2 (Protocol de Transferència d'Hipertext versió 2)
IDH	Índex de Desenvolupament Humà
IKE	Internet Key Exchange (Intercanvi de Claus d'Internet)
IPv4	Internet Protocol version 4 (Protocol d'Internet versió 4)
IT	Information Technology (Tecnologia de la Informació)
JSON	JavaScript Object Notation (Notació d'Objectes de JavaScript)
Landing page	Pàgina de destí on arriba un usuari després de clicar en un enllaç
LGPD	Llei General de Protecció de Dades
Log-in	Procés d'autenticació per accedir a un sistema informàtic
Malware	Programari maliciós dissenyat per danyar o obtenir accés no autoritzat a un sistema informàtic
MT	Medical Technology (Tecnologia Mèdica)
NAT	Network Address Translation (Traducció d'Adreces de Xarxa)
NMAP	Network Mapper (Mapeig de Xarxa)

Orthanc	Sistema de codi obert per a l'emmagatzematge i gestió d'imatges mèdiques DICOM
OSINT	Open Source Intelligence (Intel·ligència de Fonts Obertes)
PACS	Picture Archiving and Communication System (Sistema d'Arxiu i Comunicació d'Imatges)
PHP	Hypertext Preprocessor (Pre-processador d'Hipertext)
PIB	Producte Interior Brut
Pipeline	Seqüència d'operacions on la sortida d'una operació és l'entrada de la següent
Ransomware	Programari maliciós que xifra les dades d'un sistema
Reverse DNS	Procés per esbrinar el nom de domini associat a una adreça IP.
RGPD	Reglament General de Protecció de Dades
RIPE NCC	Organització que administra l'assignació de recursos d'internet a Europa, com adreces IP i ASNs.
RSA	Rivest–Shamir–Adleman (Algoritme/Conferència)
Scrapping web	Tècnica d'extracció automatitzada de dades de llocs web
SOPHOS	Empresa de seguretat informàtica
TCP/IP	Transmission Control Protocol/Internet Protocol (Protocol de Control de Transmissió/Protocol d'Internet)
TLS	Transport Layer Security (Seguretat de la Capa de Transport)

VPN	Virtual Private Network (Xarxa Privada Virtual)
ZGrab	Eina d'escaneig de serveis a Internet
ZMap	Eina d'escaneig massiu d'Internet

1. Objecte del projecte

1.1 Introducció

Des de principis dels anys 90, Internet ha esdevingut una eina essencial d'ús quotidià. La seva integració als diversos sectors de la indústria ha incrementat notablement la productivitat, essent el sector mèdic un dels més beneficiats per aquesta evolució tecnològica. La digitalització de les històries clíniques, els sistemes de gestió hospitalària, la telemedicina i els dispositius mèdics connectats han transformat radicalment la manera com es proporciona l'atenció sanitària.

No obstant això, la seguretat durant la creació dels primers protocols i serveis d'internet no era una prioritat. Els fonaments d'Internet es van desenvolupar en un entorn de confiança entre institucions acadèmiques i de recerca, sense anticipar l'expansió massiva i la diversitat d'usos que tindria en el futur. Aquesta manca de consideració de la seguretat en el disseny inicial ha provocat una evolució paral·lela de riscos cibernètics que avui dia representen una amenaça significativa.

El sector mèdic resulta especialment vulnerable a aquestes amenaces per diverses raons: la sensibilitat de les dades que gestiona, la criticitat dels sistemes per a la vida dels pacients, i la progressiva interconnexió de dispositius mèdics que tradicionalment funcionaven aïllats. Les amenaces en aquest àmbit no es limiten a les fuites de dades personals i confidencials dels pacients, sinó que també inclouen la possibilitat de sabotatge a equips mèdics vitals que podrien posar en risc directe la salut o la vida dels pacients.

Tot i conèixer els greus problemes que comportaria l'accés no autoritzat a aquests serveis, diversos estudis han revelat vulnerabilitats alarmants. Al 2016, un grup d'investigadors van trobar 2.774 servidors de comunicació d'imatges mèdiques exposats directament a Internet arreu del món sense les mesures de seguretat adequades [1]. Més recentment, el 2019, investigadors de Greenbone Networks van demostrar que molts dels servidors que comuniquen dispositius mèdics entre si eren accessibles obertament des d'Internet sense cap mesura efectiva de protecció [2], exposant milions d'imatges mèdiques i dades personals de pacients.

1.2 Motivació

La revelació d'aquestes vulnerabilitats manifesta que, malgrat que en el paradigma actual la seguretat a la xarxa sigui més present que als anys 90, les mesures implementades no són prou contundents per protegir infraestructures crítiques com les sanitàries. Resulta especialment preocupant l'estat d'Internet i dels serveis per on circula informació mèdica confidencial, així com la facilitat amb què es pot accedir a dispositius mèdics essencials per al funcionament d'una infraestructura hospitalària.

El més alarmant és que aquestes vulnerabilitats sovint passen desapercebudes per a la majoria dels veritables afectats: pacients i professionals sanitaris. Existeix una bretxa important entre la percepció de seguretat que tenen els usuaris d'aquests sistemes i la realitat de les seves vulnerabilitats. Aquesta situació crea un fals sentit de seguretat que pot agreujar el problema, ja que la manca de conscienciació dificulta l'adopció de mesures preventives adequades.

La motivació principal d'aquest projecte sorgeix, doncs, de la necessitat urgent d'avaluar i visibilitzar l'estat real de la seguretat en els entorns mèdics digitals, especialment considerant que aquestes vulnerabilitats poden tenir conseqüències directes sobre la privacitat dels pacients i, en casos extrems, sobre la seva salut i benestar.

També es vol respondre a les preguntes que sorgeixen sobre l'estat d'Internet proposat. Per què aquests servidors estan exposats? Ho haurien d'estar? És negligència o falta de pressupost per a tenir una infraestructura segura? En són conscients els administradors de les respectives xarxes mèdiques dels problemes que pot comportar tenir aquestes bretxes de seguretat? En qualsevol cas, s'espera obtenir informació per a respondre totes aquestes preguntes i treure'n conclusions.

1.3 Objectius del projecte

Aquest projecte té com a objectiu principal crear una *pipeline* per a la obtenció, tractament i anàlisi de dades de fonts d'informació pública. Aquesta metodologia permetrà la investigació de vulnerabilitats en els sistemes mèdics connectats a Internet. Els objectius específics són els següents:

1. Quantificar i classificar les vulnerabilitats més comunes en entorns mèdics digitals, amb especial atenció a aquelles que podrien comprometre la confidencialitat de les dades dels pacients o la integritat dels dispositius mèdics.
2. Analitzar patrons i tendències en la distribució geogràfica i institucional de les vulnerabilitats detectades.
3. Contribuir a la sensibilització sobre la importància de la seguretat informàtica en entorns mèdics, proporcionant evidències empíriques sobre l'estat actual de la seguretat.

El propòsit final és contribuir a aconseguir una xarxa més segura i evitar possibles entrades malintencionades a serveis crítics dins les infraestructures sanitàries, exemplificant el que no hauria de ser i informant a les entitats afectades sobre el perill de tenir escletxes en els seus sistemes informàtics.

1.4 Abast del projecte

Inicialment, el projecte centrarà els esforços en l'anàlisi i tractament de dades públiques recollides mitjançant eines OSINT d'Espanya i Alemanya. Aquesta delimitació geogràfica permetrà validar el correcte funcionament de la *pipeline* desenvolupada i obtenir resultats comparatius entre dos països europeus amb diferents marcs reguladors i nivells d'adopció tecnològica en l'àmbit sanitari.

L'abast tècnic del projecte inclou:

1. El desenvolupament d'una *pipeline* automatitzada per a la recol·lecció, processament i anàlisi de dades sobre dispositius mèdics connectats a Internet.
2. La identificació i classificació dels tipus de dispositius mèdics exposats, les seves vulnerabilitats més comunes i les configuracions no segures detectades.
3. L'anàlisi estadística i visualització de les dades obtingudes per identificar patrons i tendències.
4. La documentació exhaustiva de la metodologia emprada per facilitar la seva replicació i escalabilitat.

És important destacar que, tot i estar orientat a la detecció de vulnerabilitats, durant el desenvolupament d'aquest projecte no s'utilitzarà en cap moment *hacking* actiu que pogués comprometre la integritat o funcionament dels sistemes analitzats. Totes les tècniques utilitzades es basen exclusivament en la recol·lecció i anàlisi d'informació disponible públicament.

A més, seguint principis ètics fonamentals en la recerca de seguretat, totes les entitats afectades per alguna vulnerabilitat en el seu sistema seran informades degudament abans de la publicació d'aquest treball, proporcionant-los temps suficient per implementar les mesures correctores necessàries.

La metodologia desenvolupada està dissenyada amb la idea de poder-la escalar per a l'anàlisi de qualsevol país del món de manera senzilla en futures ampliacions del projecte.

2. Estat de l'art

2.1 Marc històric

A finals dels anys 60, l'Agència de Projectes d'Investigació Avançada (ARPA) del Pentàgon dels Estats Units van proposar una solució tecnològica per a connectar els ordinadors de diversos centres d'estudi a grans distàncies i així poder compartir informació entre els investigadors. Abans, els ordinadors només es podien comunicar entre ells si estaven a la mateixa zona i tot i així, la capacitat per intercanviar dades era limitada. Amb aquest nou sistema de xarxa, conegut com a "commutació de paquets", es va aconseguir enviar informació entre els centres d'estudi de la Universitat de Califòrnia i l'Institut d'Investigació de Stanford mitjançant dos ordinadors. Aquesta nova xarxa revolucionària es coneix com a ARPANET, la precursora de l'Internet actual.

Només dos anys després de l'enviament del primer missatge per aquesta xarxa, l'investigador Bob Thomas Morris va crear el primer *malware*: CREEPER (en català, planta trepadora). Aquest era un programa informàtic que es transmetia entre els dispositius connectats a la xarxa i es replicava entre ells, mostrant per pantalla el missatge "*I'm the creeper: catch me if you can!*". Per lluitar en contra d'aquest *malware*, l'investigador Ray Tomlinson va crear el que es podria considerar com el primer antivirus de la història: REAPER (en català, la segadora). Aquest perseguia i eliminava el virus del sistema informàtic.

Entre els pioners en els atacs informàtics destaca Kevin Mitnick. El jove estudiant de secundària va piratejar *The Ark*, un sistema informàtic per al desenvolupament de sistemes operatius. Mitnick, utilitzant una tècnica que ara porta com a nom "enginyeria social", va fer-se passar per un enginyer de software amb alt càrrec a la companyia DEC (la creadora de *The Ark*) mitjançant una trucada telefònica a qui havien bloquejat el seu compte. Amb les dades d'accés que aquest treballador li va proporcionar, va poder accedir sense autorització a grans quantitats de dades confidencials de l'empresa. [3]

Més endavant, durant la dècada dels 80, quan l'ús dels ordinadors ja s'havia estès a l'administració pública, la societat ja n'era més conscient dels riscs cibernètics que anaven apareixent ja que l'amenaça cibernètica creixia a l'alça. Amb la creació dels nous virus informàtics neix la creació dels nous softwares antivirus. N'és un exemple el *virus Vienna*, un programa que es replicava automàticament entre els dispositius i corrompia les dades d'aquests. Bernd Fix, després que el seu ordinador resultés infectat amb aquest virus, va codificar un software antivirus modern que localitzava i eliminava aquest *malware*. Aquest tipus de software no va tardar en comercialitzar-se per protegir la informació dels usuaris d'Internet. [3]

És evident, doncs, que la seguretat informàtica tendia a anar a remolc dels delinqüents, ja que en una època on es destinaven els esforços a desenvolupar aquesta nova xarxa creixent que era Internet i a estandarditzar l'ús dels ordinadors personals, no es podia vetllar gaire per la protecció. Això ha fet créixer la seguretat informàtica amb una coixesa que s'està vetllant per solucionar.

2.2 Estat actual de la seguretat a la xarxa

En el panorama tecnològic actual, han sorgit noves amenaces cibernètiques a causa de l'evolució constant de les diferents branques tecnològiques. La Intel·ligència Artificial, per exemple, s'enfronta a atacs sofisticats que comprometen els seus models d'aprenentatge. Durant la conferència RSA de 2021, una de les més prestigioses en seguretat informàtica a nivell mundial, diversos investigadors van demostrar com es podia enganyar el sistema de pilot automàtic de vehicles Tesla i Mobileye simplement projectant imatges fantasma [4]. Aquest incident evidencia la necessitat no només de desenvolupar nova tecnologia, sinó de fer-ho garantint la seva seguretat.

Malgrat l'avenç tecnològic, els atacs tradicionals continuen vigents, tot i que amb tècniques més perfeccionades, contra infraestructures modernes. Els atacs de tipus *malware* descrits anteriorment en aquest estudi segueixen representant una amenaça significativa. Un cas particularment rellevant és el de *LockBit*, un dels *malware* d'encriptació de dades més efectius i destructius dels darrers temps.

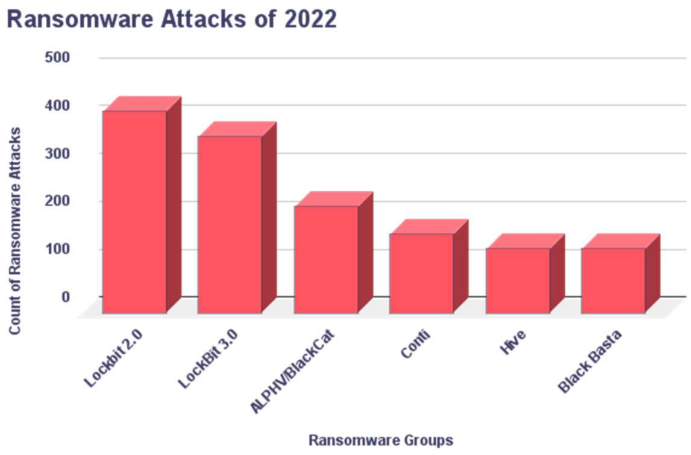


Fig. 2.2.1 Atacs de *Ransomware* durant el 2022. Font: [5]

LockBit és un atac d'encriptació de dades (*Ransomware*) distribuït mitjançant un *malware* que, un cop infecta el dispositiu, xifra totes les dades i el deixa completament inutilitzat. Els ciberdelinqüents solen exigir un rescat econòmic a canvi de la clau de desencriptació per recuperar la informació. [6]

Aquest tipus d'atac representa actualment un dels majors desafiaments per a la seguretat informàtica global. Les estadístiques són alarmants: el 70% dels atacs cibernètics impliquen l'encriptació de dades i, segons un estudi de SOPHOS realitzat a 5.000 organitzacions, el 59% d'aquestes han patit algun atac de *Ransomware*, amb un cost mitjà d'aproximadament quatre milions de dòlars per incident. [7]

2.3 Seguretat informàtica en entorns mèdics

No hi ha gaire informació pública al respecte sobre com es gestionen o s'estructuren els processos de negoci que esdevenen als centres mèdics. En general, però, podem separar els diferents processos en:

Les Tecnologies de la Informació (IT) en un centre mèdic comprenen tots els processos relacionats amb la pròpia gestió del centre i el tractament i emmagatzematge de la informació

que s'hi genera. En són exemples les dades de factures, historial clínic de pacients, funcions administratives com la programació, facturació i inventari d'estoc, etc.

La protecció d'aquests sistemes és clau per al correcte funcionament del centre, però també afecta directament a la salut dels pacients. Els tractaments que proposen els professionals es basen en les dades que s'han obtingut prèviament del pacient; si es falsegen les dades d'un pacient o s'intercanvien per les d'un altre, el tractament no seria l'indicat i podria arribar a ser mortal.

El correcte funcionament del centre mèdic també depèn d'aquest tipus de sistemes. Com a qualsevol empresa, els seus pacients també són els seus clients. Sense un control d'estoc de material o la impossibilitat d'efectuar factures, el centre entraria en bancarrota. No només això, sinó que l'incompliment de la llei general de protecció de dades comporta unes sancions d'entre deu i vint milions d'euros depenent de la magnitud del problema, uns preus que moltes de les clíniques petites i mitjanes no podrien arribar a pagar.

Les Tecnologies Mèdiques (MT) són els sistemes mèdics que comprenen les màquines que estan en contacte directe amb el pacient. Aquestes són altament sensibles ja que qualsevol mal funcionament pot comportar un risc de mort directe per al pacient. Comprenen des de les màquines més superficials com els escàners de raigs X o els monitors de signes vitals, fins als equips més sofisticats com braços robots quirúrgics, incubadores intel·ligents per a nadons o fins i tot els respiradors que mantenen amb vida a alguns pacients.

És evident, doncs, que la complexitat tecnològica als hospitals ha crescut en gran mesura i que això introdueix un seguit de desafiaments per a la seguretat informàtica d'aquests sistemes. Per això, detectar i corregir vulnerabilitats dels sistemes mèdics ha esdevingut una prioritat per al personal dels centres, però els equips encarregats d'aquesta tasca habitualment són inexistents o estan sobrecarregats i això comporta que aquestes tasques quedin endarrerides.

Fonts com *Check Point Research* exposen que actualment, el sector mèdic està entre els sectors més atacats.

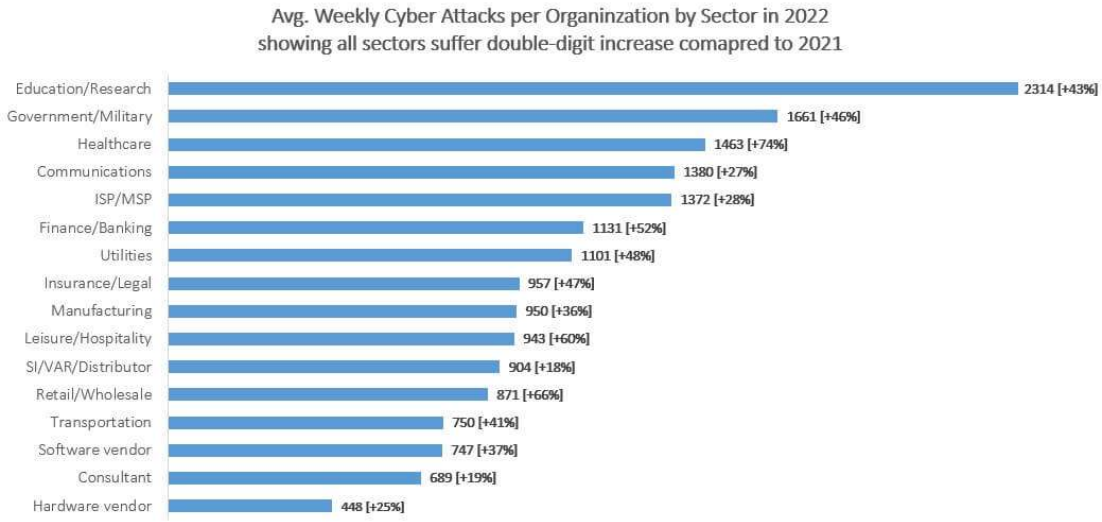


Fig. 2.3.1 Increment dels atacs setmanals per sector durant el 2022. Font: [8]

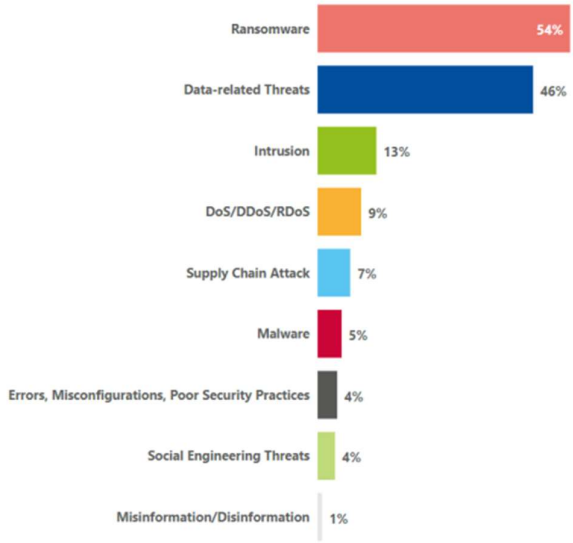


Fig. 2.3.2 Tipus d'atacs a centres mèdics de Gener 2021 a Març 2023. Font: [9]

I com es veu a la Fig. 2.3.2 el principal atac que reben els centres mèdics són els de *RansomWare* impulsats pel desig d'aconseguir un benefici econòmic.

2.4 Estudis similars

Com s'ha mencionat breument en aquest document, un estudi publicat per Mark Stites i Oleg S. Pianykh a la prestigiosa *American Journal of Roentgenology* el 2016 prenia com a objectiu analitzar l'estat de la seguretat dels servidors d'imatge mèdica mundials. Buscaven determinar com de vulnerables eren els serveis de radiologia a potencials bretxes de seguretat. Per això, van crear una aplicació especialitzada anomenada "DICOM Ping" (DPing) que identificava servidors de radiologia no segurs enviant sol·licituds d'enllaç DICOM estàndard. A diferència d'estudis anteriors que només comprovaven els ports de connexió oberts, la seva aplicació intentava iniciar la comunicació DICOM amb servidors remots. Per fer això, van escanejar quasi tot el rang d'IPv4, unes 3.7 bilions d'adreces, utilitzant servidors de computació Amazon EC2 amb vint màquines virtuals corrent en paral·lel.

Categoritzaven les respostes dels servidors en dues possibilitats: connexió rebutjada, cosa que mostrava una seguretat bàsica del servidor tot i que no hauria d'estar exposat a Internet, i connexió acceptada, indicant una vulnerabilitat molt important al servidor. Els resultats van ser alarmants; es van identificar 2774 servidors d'imatge mèdica desprotegits a Internet i d'aquests, 719 (un 26% aproximadament) estaven completament oberts a connexions externes, és a dir, permetien la modificació i descàrrega d'informació dels pacients. [1]

Seguint amb la investigació de l'empresa Greenbone sobre l'estat de la seguretat informàtica en servidors mèdics ja mencionada en aquest document, es va publicar un segon informe de la mateixa companyia comparant els resultats 60 dies després de la publicació del primer, on s'havien trobat més de 730 milions d'imatges mèdiques exposades a Internet. En aquest segon informe s'explica que hi ha tres respostes diferents entre els països afectats:

- 11 països van desconnectar d'Internet els servidors exposats trobats en el primer informe.
- 45 països van millorar lleugerament o no van empitjorar la situació que tenien en el primer informe.
- 5 països van empitjorar notablement la situació.

Aquest informe destaca un important problema global de ciberseguretat que afecta la privadesa dels pacients i els sistemes sanitaris, amb greus implicacions per al compliment de la seguretat de les dades i possibles danys financers.

I relacionat amb eines OSINT hi trobem l'estudi que el propi Censys va dur a terme per analitzar l'estat de la seguretat dels sistemes sanitaris exposats a Internet. Utilitzant la seva pròpia eina d'escaneig d'Internet van identificar interfícies i serveis sanitaris accessibles públicament. Van analitzar els sistemes exposats per tipus de producte, proveïdor, ubicació geogràfica i presència a la xarxa i a més van enviar notificacions als contactes de seguretat de cada dispositiu exposat com a part del seu procés de divulgació responsable.

Com a resultat, es van trobar prop de 14000 adreces IP que exposaven dispositius sanitaris i sistemes de dades que contenien informació mèdica sensible. La categoria més gran de sistemes exposats eren servidors gestors d'imatges mèdiques (36%), seguits de sistemes EMR/EHR (28%) que emmagatzemen registres complets de pacients. Similar als altres estudis, els investigadors també van descobrir que molts d'aquests sistemes no tenien mesures de seguretat adequades, i alguns permetien accés sense autenticació a bases de dades d'imatges mèdiques. Val a destacar que totes aquestes troballes es van efectuar amb informació pública, simplement fent un escaneig massiu d'Internet sense tenir en compte explotació de vulnerabilitats o tècniques de *hacking* més avançades.

3. Desenvolupament

En aquest punt s'explicarà la metodologia utilitzada i les eines emprades per aconseguir els resultats finals, que s'avaluaran durant el punt **4. Avaluació**. Aquest treball no se centra en el desenvolupament tecnològic però les eines creades durant els processos d'obtenció, emmagatzematge i anàlisi de resultats han fet possible el tractament d'un volum tan gran de dades.

3.1 Eines emprades

Per a l'obtenció, tractament i anàlisi de dades s'han hagut de crear un seguit de programes que automatitzin el procés. L'exploració de tot el rang d'IPv4 manualment utilitzant eines com NMAP no és factible degut al gran nombre d'hores que requeriria. Una opció seria utilitzar servidors en paral·lel per fer aquesta feina, però segueix sense ser factible degut al cost econòmic que suposaria llogar aquests equips i la quantitat de dades que s'haurien de tractar. Per això, els programes que s'han creat en aquest projecte s'ajuden d'eines d'escaneig massiu d'Internet ja creades, com Censys i Shodan (explicades més endavant). Aquestes estan dedicades a la recol·lecció de dades sobre els dispositius exposats a Internet i actualitzen les seves bases de dades una vegada a la setmana, després d'escanejar tot el rang complert d'IPs.

Per a la creació d'aquest programes, s'ha utilitzat el llenguatge *Python* donat que és senzill i molt flexible. A més, compta amb una gran quantitat de llibreries que es poden integrar per facilitar la implementació de diferents funcionalitats. *Python* és conegut com un llenguatge poc ràpid i això podria afectar a la velocitat de processament de les dades obtingudes, però en les darreres actualitzacions que ha rebut, s'ha optimitzat gran part de la seva lògica fent-lo així més competitiu en aquest aspecte. [10]

Juntament amb *Python*, s'han utilitzat diverses llibreries que implementen funcionalitats necessàries per al correcte funcionament dels programes. Les principals són:

- FPDF. És una llibreria que permet crear documents PDF fàcilment. S'ha utilitzat per a la creació automàtica d'informes sobre els servidors analitzats.

- PYMongo. És el controlador oficial de MongoDB, el proveïdor de bases de dades utilitzat, per a *Python*. Permet la comunicació amb les bases de dades de MongoDB d'una manera senzilla i àgil i s'ha utilitzat per emmagatzemar les dades de manera ordenada per a poder analitzar-les posteriorment.

- Matplotlib. Aquesta llibreria crea gràfics a partir de les dades d'entrada. S'ha utilitzat per mostrar de manera automàtica, ordenada i visual els gràfics que es representen més endavant en aquest document.

- Logging. És un mòdul de *Python* que facilita la creació i actualització d'un registre d'accions en el codi. Els programes que s'han creat estan dissenyats per funcionar durant molta estona seguida i és important guardar tot el procés i l'estat en què es trobava per si hi ha algun error imprevist.

Quant a l'emmagatzematge de les dades s'ha utilitzat les bases de dades no relacionals de MongoDB. S'ha escollit aquesta opció donada la seva senzillesa a l'hora d'implementar-les en local. Les dades que s'han de guardar no estan relacionades entre elles i la interfície gràfica que presenta aquesta eina és suficient per poder visualitzar les taules que s'hi emmagatzemen en ella. A part de la seva completa integració amb *Python* utilitzant la seva llibreria, presenta un estil d'emmagatzematge tipus JSON, una estructura de dades àmpliament coneguda i fàcil de modelar. MongoDB respon molt bé davant grans volums de dades com els que s'han treballat en aquest projecte i implementa un sistema de consulta ràpida per a més velocitat en les transaccions.

Però les eines principals en aquest projecte han estat els programes OSINT Censys i Shodan. OSINT (*Open Source Intelligence*) es refereix a la recol·lecció i anàlisi d'informació obtinguda de fonts públiques i obertes. Hi ha moltes eines que utilitzen l'estil OSINT per crear grans bases de dades d'informació pública. *The Harvester* [11] és una eina de

recol·lecció d'informació que amb un domini o IP pública pot trobar noms, subdominis, correus electrònics i molt més, de fonts obertes i públiques.

Censys i Shodan, tot i no funcionar exactament d'igual manera, tenen el mateix objectiu: recol·lectar informació sobre els servidors connectats a Internet fent un reconeixement complert de tots els dispositius i tots els serveis exposats. Les dues eines utilitzen sondes repartides per la xarxa per recol·lectar aquesta informació i, posteriorment, les emmagatzemen per poder-les consultar a través de la seva API web.

3.1.1 Censys

Censys va ser desenvolupat a partir d'investigacions acadèmiques realitzades per científics de la Universitat de Michigan (i de la Universitat d'Illinois Urbana-Champaign en alguns casos), on es va concebre inicialment per aprofitar el potencial d'escaneig d'Internet i millorar la seguretat en línia. Al començament, el projecte es va basar en eines com ZMap i ZGrab, que permeten escanejar i recopilar informació sobre el rang d'adreces IPv4 i les configuracions dels serveis exposats a Internet [12].

Entre les funcionalitats principals de Censys s'hi destaquen la capacitat de cerca i consulta a través de la seva API o la seva interfície web (https://search.censys.io), l'anàlisi de la superfície d'atac d'una organització, el que permet comprendre les vulnerabilitats que estan exposades a Internet a temps real, i la integració amb diversos llenguatges de programació que ofereixen les seves llibreries, que facilita la automatització de processos [13].

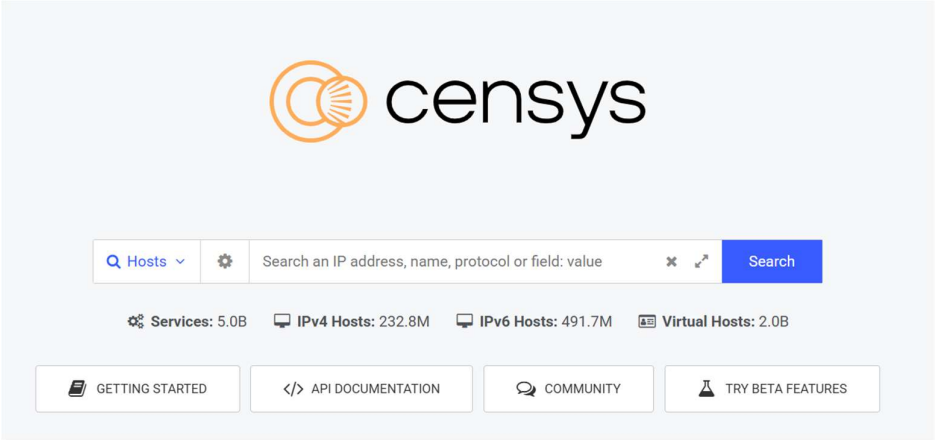


Fig. 3.1.1.1 Interfície Web de l'eina Censys. Font: [14].

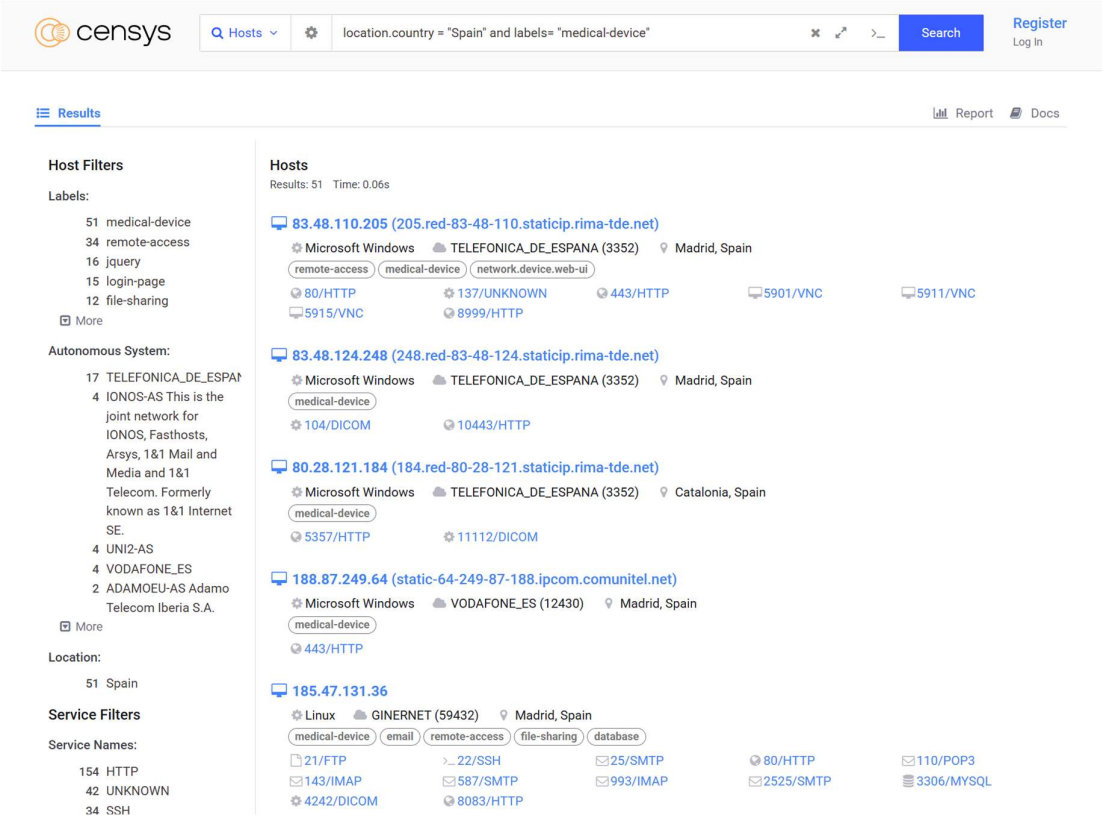


Fig. 3.1.1.2 Exemple de consulta a la base de dades de Censys mitjançant la seva interfície web. Font [14].

L'eina també permet l'accés a les dades històriques dels *hosts*. És útil per saber si un servei ha estat creat recentment i/o rep actualitzacions constants o és un servei antic que s'ha deixat d'actualitzar. Amb aquestes dades es pot determinar si són serveis de prova o porten operatius ja de fa temps i tenen un ús real.

3.1.2 Shodan

Shodan va ser fundat per John Matherly el 2009. El seu objectiu principal és indexar i oferir informació sobre dispositius exposats a la xarxa.

L'eina realitza escanejos en múltiples ports i protocols per recopilar *banners*, respostes i metadades dels dispositius que responen, obtenint informació detallada sobre el sistema operatiu, programari, versions i configuracions. Aquesta informació podria revelar potencials vulnerabilitats. Una vegada recopilada, aquesta informació s'indexa en una base de dades centralitzada. Així mateix, Shodan ofereix una interfície web intuïtiva i una API que facilita als desenvolupadors integrar les seves funcionalitats en aplicacions i sistemes de monitorització de seguretat, permetent l'automatització a la recerca d'amenaces [15].

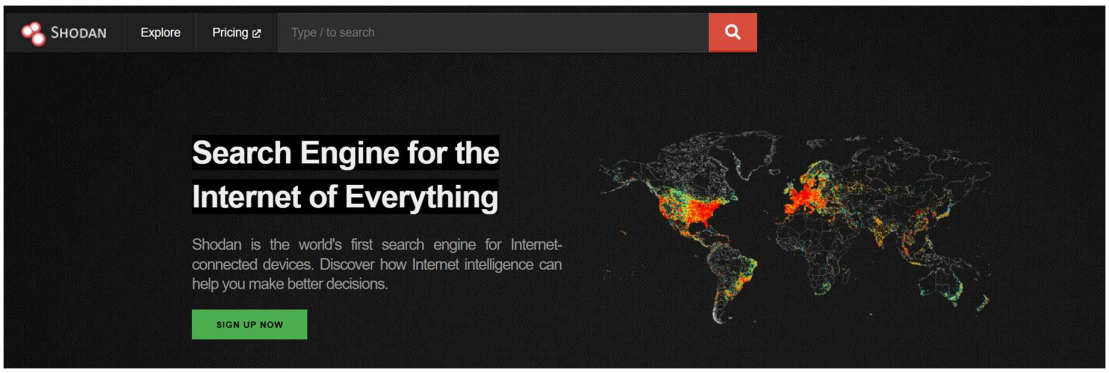


Fig. 3.1.2.1 Interfície Web de l'eina Shodan. Font: [15].

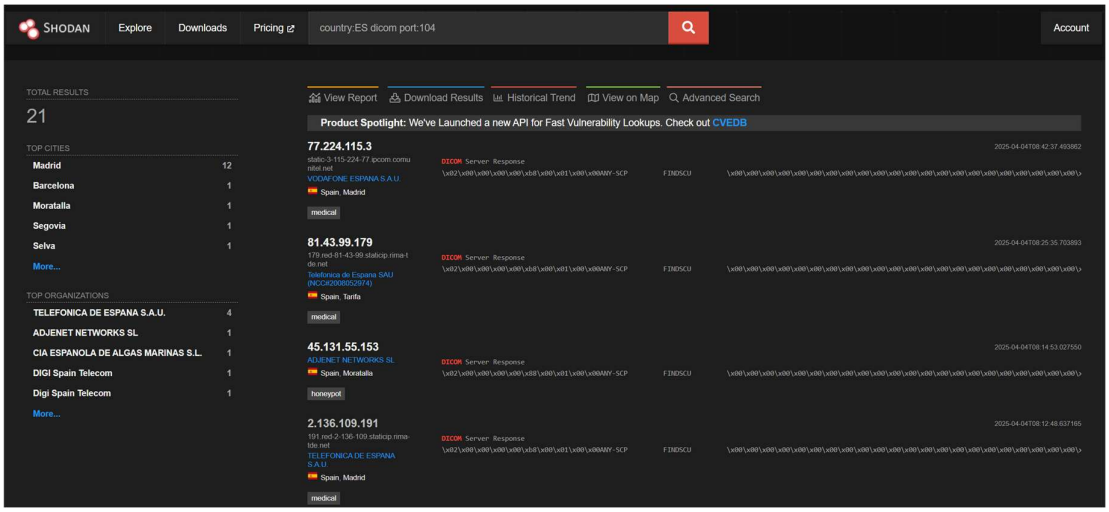


Fig. 3.1.2.2 Exemple de consulta a la base de dades de Shodan mitjançant la seva interfície web. Font [15].

Els dos motors de cerca tenen una retribució econòmica basada en subscripció que proporciona un nombre diferent de *tokens* segons el nivell que es pagui. Cada *token* representa una consulta a l'API. Per a aquest projecte s'ha comptat amb l'estatus d'investigador per a l'eina Censys, el que ha augmentat el nombre de *tokens* disponibles de 100 *tokens*/mes a 25.000 *tokens*/mes. Com que per a Shodan només es comptava amb 100 *tokens*/mes, s'ha optat per consultar dispositius mèdics a Censys i, després d'obtenir el llistat d'IPs, demanar informació sobre cada *host* a Shodan, ja que les consultes individuals no consumeixen *tokens*.

3.2 Metodologia

Per poder obtenir les dades de les APIs de Censys i Shodan i després poder tractar-les per extreure'n informació, s'ha dissenyat una *pipeline* perquè aquestes dades segueixin un camí i es vagin modelant per arribar a treure estadístiques i gràfics. Aquesta consta de 3 blocs en quant a les dades: obtenció, tractament i anàlisi. Cada bloc està diferenciat amb un programa diferent fet en *Python*. Existeix la possibilitat d'unir els tres blocs en un sol programa, però la complexitat d'aquest no permetria la flexibilitat que s'ha necessitat per resoldre els problemes que han anat sorgint.

Per això, s'ha plantejat una obtenció de dades des d'un primer programa que recol·lecta informació de Censys i Shodan i l'emmagatzema en una carpeta, amb el nom del país on pertany, en fitxers de text amb extensió ".censys" o ".shodan" respectivament. Una vegada obtingudes aquestes carpetes, un altre programa les recorre i emmagatzema les dades en una base de dades de MongoDB. L'últim pas és consultar les dades necessàries que ara ja han estat endreçades per a fer-ne gràfics i extreure'n conclusions. La Fig. 3.2.1 mostra el procés que està detallat més endavant.

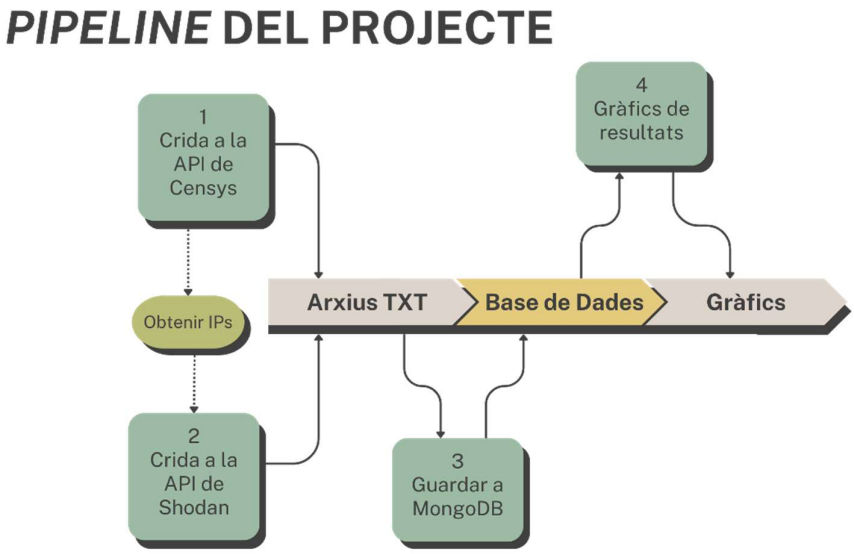


Fig. 3.2.1. Representació gràfica de la *pipeline*. Font: Elaboració pròpia.

3.2.1 Cerca de dispositius mèdics a Internet

Per començar, aquest estudi està enfocat únicament en les vulnerabilitats dels serveis mèdics i per tant s'han de buscar els servidors que siguin utilitzats com a tal. Una primera aproximació podria haver estat la d'escanejar tot el rang d'IPv4 en busca de servidors que semblessin serveis mèdics. Els dos motius que van fer descartar aquesta aproximació van ser la quantitat de servidors i el nombre de ports que poden tenir oberts aquests serveis que s'allotgen al rang d'IPv4 (2^{32} = 4.294.967.296 servidors i 2^{16} = 65.536 ports per servidor. Uns 281.474.976.710.656 escanejos) i que tot i que no sigui directament il·legal fer un

escaneig de ports, pot ser considerat una manera de *hacking* actiu i una amenaça per a les institucions. Per això es va decidir utilitzar les eines mencionades anteriorment que mitjançant sondes distribuïdes per la xarxa d'Internet, proporcionen informació nova dels seus escaneigs recurrents.

3.2.1.1 Enfocament A-priori

Un primer enfocament que ha estat testat durant aquest treball és el que s'ha anomenat com a "*A-Priori*". Aquest consisteix en obtenir el rang d'IPs d'un hospital en concret mitjançant consultes *reverse-DNS* sobre la pàgina web d'un hospital. Si el servidor de la pàgina web de l'hospital està localitzat dins les instal·lacions de l'hospital i utilitza el rang d'IP que se li ha proporcionat al centre, a les hores podrem obtenir aquest sub-rang d'IP en que treballen tots els dispositius mèdics. Aquesta no és una pràctica habitual; les pàgines web són sempre una porta de connexió entre la xarxa informàtica del centre i els usuaris i, per tant, presenta un risc més elevat tenir-lo directament en connexió amb el sistema informàtic del centre mèdic. Per aquest motiu, molts centres de salut opten per allotjar-la sobre serveis externs i l'aproximació "*A-Priori*" només funciona en determinats casos.

Tot i això, en els casos en els quals es pot obtenir el rang d'IPs de l'hospital es genera una consulta a les dues eines mencionades, Censys i Shodan, per tot el rang d'IPs. Normalment, per a hospitals de mitjana capacitat o centres de salut més petits se solen utilitzar màscares de fins a 4.094 *hosts* (o un /20), que permet generar consultes en un temps suficientment ràpid.

3.2.1.2 Enfocament Zero-Knowledge

El segon enfocament que es va proposar va ser el "*Zero-Knowledge*", que consisteix en una consulta directa a ambdues eines sobre direccions IP de serveis mèdics. El buscador d'aquestes eines proporciona la possibilitat de filtrar els servidors per una gran varietat de informació: rang IP, port, serveis que hi corren, localització, etiquetes, etc. S'ha provat amb diferents metodologies com el filtratge per serveis que solen fer servir els dispositius mèdics, com DICOM, el filtratge per ports, com el 443 o el 11112 (HTTPS i DICOM) o el filtratge per etiquetes.

És aquesta última la que s'ha fet servir en aquest projecte. A part d'escanejar els serveis que corren els servidors, aquestes eines etiqueten cada servidor seguint diferents normes. S'ha contactat amb un treballador d'una d'aquestes dues companyies per a demanar informació sobre els criteris que es tenen a l'hora d'etiquetar els servidors, però han volgut mantenir ocult aquests criteris, tot i que en el seu missatge anomena la possibilitat de filtrar els servidors per els *banners HTML*. Tot i això, l'etiquetatge que proporcionen ambdues eines és prou correcte com per poder-s'hi fiar. S'ha provat d'entrar manualment a les pàgines HTML, que exposaven prop d'un centenar de serveis que l'eina havia etiquetat com a "*Medical Device*" i en el 100% dels casos, la pàgina mostrava signes inequívocs que es tractava d'un aparell relacionat amb l'àmbit mèdic (normalment humà, però també de veterinàries) com mostra la Fig. 3.2.1.1.

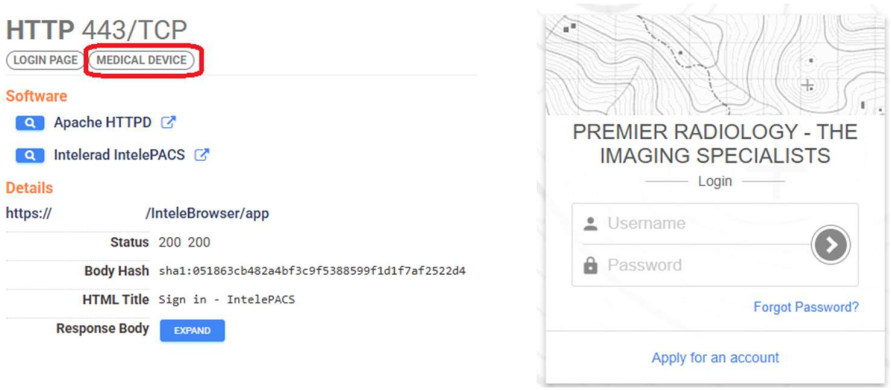


Fig. 3.2.1.1. Exemple de servidor amb l'etiqueta "*Medical Device*" i el seu contingut HTML. Font: Elaboració pròpia.

3.2.1.3 Hackathon

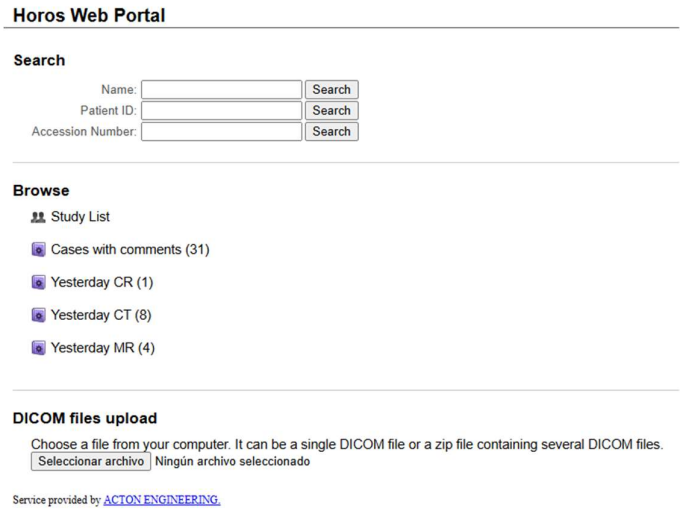
Per decidir la metodologia a utilitzar es va organitzar una *hackathon* que va comptar amb la participació de diversos estudiants amb diferents nivells de coneixement en la matèria, des d'estudiants de grau fins a graduats cursant màsters. Aquesta activitat es va desenvolupar dins la càtedra de la Universitat Politècnica de Catalunya (UPC) anomenada "Càtedra Carismàtica de Ciberseguretat" [16], amb el finançament de l'Institut Nacional de Ciberseguretat (INCIBE).

Fig. 3.2.1.3.1. Imatges de la *hackathon*. Font: Benedikt Michaelis.

Durant 4 dies, en esprints de 2h, es va estar treballant amb ambdós mètodes per saber quin obtenia millors resultats. Després de valorar els resultats obtinguts es va decidir utilitzar el mètode "*Zero-Knowledge*" ja que com s'ha explicat abans, l'etiquetatge que fan les dues eines és prou correcte per a ser un criteri a seguir. Durant aquesta *hackathon* però, no només es va determinar el procediment a seguir sinó que es van elaborar eines en forma de codis que obtenien la informació, la tractaven i l'organitzaven perquè l'ull humà pogués anar veient què s'hi amagava darrere aquella IP.

Un dels resultats més sorprenents va ser el d'una IP a l'Índia, que pretén amagar una base de dades d'un servei de radiologia i escaneig rere un usuari i contrasenya. La aplicació web, però, està mal dissenyada i sense haver de fer un *login*, es pot veure els resultats. Es pot pensar que es podria tractar d'un *honeypot*, però la ubicació on està registrat aquest servidor detectat per Censys mostra una clínica si s'utilitza el Google Streat View sobre les mateixes coordenades. La Fig. 3.2.1.3.2 com a exemple.

Fig. 3.2.1.3.2. Portal web de la clínica índia. Font: Elaboració pròpia.



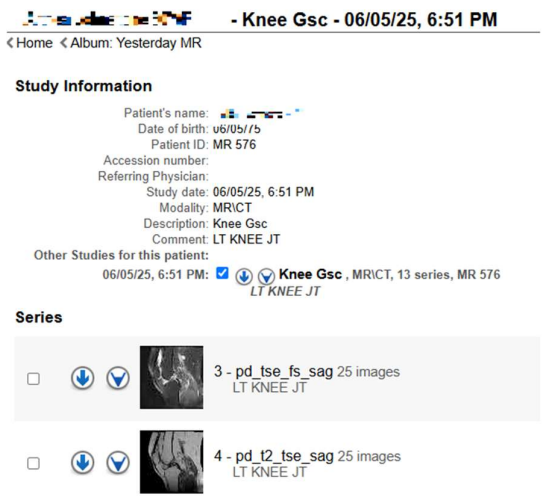


Fig. 3.2.1.3.3. Portal web de la clínica índia després de prémer el botó "Login". Font: Elaboració pròpia.

Fig. 3.2.1.3.4. Centre mèdic del servidor. Font: Elaboració pròpia.

3.2.2 Adquisició de les dades

Una vegada s'ha decidit la metodologia per al triatge de *hosts*, s'ha creat una eina en forma de programa en *Python* que fa una primera cerca a *Censys* per a preguntar per tots els servidors marcats com a "*Medical Device*" d'un país en concret. Una vegada obtingut aquest objecte de tipus cerca, s'itera per a totes les IPs obtingudes emmagatzemant la informació en una carpeta amb el nom del país, i un fitxer de tipus text amb la extensió ".censys" amb

la IP del dispositiu com a nom del fitxer. Totes aquestes IPs són guardades en una variable de tipus llista per poder consultar una per una a la plataforma Shodan.

S'ha de seguir aquest procediment ja que per fer una consulta d'aquestes característiques Censys gasta només un crèdit (un crèdit és una consulta que es fa a la API), però Shodan gasta un crèdit per a cada IP. Els crèdits són limitats i tot i que s'ha comptat amb comptes estudiantils a ambdues plataformes, no són suficients per la quantitat de consultes que s'han hagut de fer. Per tant, es procedeix de manera diferent a les dues plataformes: a Censys es demana per tots els serveis amb la etiqueta "*Medical Device*" d'un país (això gasta un crèdit, per a Censys s'ha pogut obtenir 250.000 crèdits mensuals) i a Shodan es pregunta per informació sobre una IP en concret (això no gasta cap crèdit).

Una vegada la informació de Shodan ha arribat, s'emmagatzema de la mateixa manera que la de Censys però amb una extensió ".shodan" a la mateixa carpeta.

Cal destacar que s'obtenen les dades de les dues plataformes ja que cadascuna retorna informació diferent. Les dues ho fan en format JSON; un diccionari principal amb diverses claus i entrades que contenen llistes o altres diccionaris, però Censys retorna informació més detallada sobre els serveis del *host* (la resposta d'un HTTP, els *banners*, la data de la resposta), metre que Shodan retorna una informació més detallada sobre les vulnerabilitats d'aquest *host* i dels seus serveis en diversos diccionaris, apuntant els CVE als quals és vulnerable.

Una vegada les dades són emmagatzemades en carpetes de fitxers de text un altre programa se n'encarrega de guardar la informació que s'utilitzarà per a una posterior anàlisi en una base de dades no relacional. Dins d'aquesta i separat en carpetes per país s'hi troben els objectes *host* amb la IP com a identificador. Dins de cada objecte s'hi emmagatzema informació d'aquell dispositiu com la localització, els serveis exposats amb els ports i les vulnerabilitats.

Seguint aquesta metodologia poden sorgir dos dubtes:

1) Per què no s'ha emmagatzemat directament la informació a la base de dades sense passar per als fitxers? És una manera de fer més flexible el procés. Tallar la *pipeline* d'aquesta manera permet fer còpies de seguretat als fitxers de manera aïllada sense dependre del correcte funcionament de la base de dades. A més, a la base de dades no s'emmagatzema

tota la informació, sinó la més important. Això permet fer consultes més veloces però si es vol treure conclusions sobre informació no emmagatzemada, s'hauria de canviar el codi per a obtenir de nou la informació i això suposa fer encara més consultes. Per tant, això funciona a mode de tallafocs i per si es vol afegir informació a les gràfiques de més endavant.

2) Per què s'ha utilitzat una base de dades i no s'ha programat per agafar la informació des dels fitxers de text directament? Per dos motius. El primer és que fer una consulta manual a les dades és més senzill des de la interfície gràfica que té la base de dades de MongoDB (MongoDB Compass). El segon és que fer una consulta automatitzada mitjançant codi a la base de dades és més senzill i fins a un 500% més ràpid que consultar fitxer per fitxer. Això és important quan tens prop de 2.000 fitxers diferents.

3.2.3 Tractament de les dades

Després de l'obtenció de les dades i de l'ordenació d'aquestes, per a poder treure uns resultats analítics, s'ha creat un últim conjunt d'eines que automatitzen els processos d'anàlisi, computació i visualitzat de gràfics.

Es diferencien dos tipus d'objectius principals dels programes: un primer bloc que pretén estructurar les dades trobades per a poder contactar més fàcilment als afectats, i un segon bloc que pretén analitzar les dades trobades per a extreure conclusions de l'estat de la seguretat als centres mèdics.

Per al primer bloc, s'ha creat una eina que consulta els fitxers de dades obtinguts (no la base de dades) per a buscar informació sobre la entitat afectada. Aquesta se sol trobar dins el paràmetre "TLS → Certificate → Names", però també es pot trobar resolent el *reverse* DNS de la IP. Aquesta eina, doncs, recull la informació sobre l'afectat i alguns camps més de context, com el port, el servei que opera o el que hi ha instal·lat, perquè amb una senzilla ullada es pugui buscar informació sobre la entitat per poder avisar. Cal remarcar que Censys proporciona una direcció electrònica de contacte, però en cap cas és la de l'afectat, ni tan sols la de l'instal·lador del servidor vulnerable. És la de l'operari registrat a la base de dades de RIPE NCC que ha adjudicat i administra aquella direcció IP. Tot i això, a vegades s'hi pot trobar informació sobre el centre afectat a l'apartat de direcció, com es mostra a la Fig. 3.2.3.1.

Fig. 3.2.3.1 Contacte RIPE d'un afectat. La direcció correspon al centre de salut. Font: Elaboració pròpia.

Per al segon bloc, el d'extreure conclusions sobre l'estat de la seguretat informàtica als centres mèdics, s'han creat diversos programes amb aquest objectiu. Cadascun té una finalitat diferent i s'utilitzen per crear gràfics diferents. Com que aquests programes només són utilitzats per a obtenir un sol gràfic del qual treure-hi conclusions, el temps que es triga executant-lo un per un no és important. Prima més la senzillesa del codi a l'hora de poder corregir sobre possibles errors i per això s'han dividit les tasques en arxius Python i no en mètodes.

3.3 Eines construïdes

S'han construït 4 tipus de programes que automatitzen els processos de descarregar les dades, emmagatzemar-les en una base de dades, mostrar estadístiques sobre aquestes i crear informes per poder contactar amb els afectats. Són programes senzills creats amb Python i cadascun té una única funcionalitat, com s'ha explicat anteriorment, per facilitar el *debugging* i l'escalabilitat d'aquests. En aquest apartat s'explicaran només els més importants o interessants ja que aquest projecte no pretén tenir un enfocament de desenvolupament tecnològic sinó més d'anàlisi de dades, tot i que s'han desenvolupat eines tecnològiques per a l'automatització de processos.

1) Aquest és una part del procés d'emmagatzematge de les dades en una base de dades local de MongoDB. Es pot veure a la Fig. 3.3.1 com està dividit el procés entre els arxius de Shodan i els arxius de Censys, ja que de cadascun s'emmagatzema informació diferent.

```python
def process_country_folder(country_folder, db=None):
    """Process all files in a country folder"""
    country_name = os.path.basename(country_folder)
    logger.info(f"Processing country folder: {country_folder} (Country: {country_name})")

    collection = None
    if not DRY_RUN and db is not None:
        # Create or get collection for this country
        collection = db[country_name]
        # Create indexes
        collection.create_index("ip")
        collection.create_index("country")
        collection.create_index("ports")
        collection.create_index("services.service_name")
        collection.create_index("tags")
        # Add indexes for vulnerability lookup
        collection.create_index("vulnerabilities.cve_list")
        collection.create_index("services.vulnerabilities")
        logger.info(f"Using collection '{country_name}' for country data")

    # Get all .shodan and .censys files
    shodan_files = glob.glob(os.path.join(country_folder, "*.shodan"))
    censys_files = glob.glob(os.path.join(country_folder, "*.censys"))

    # Process Shodan files
    for shodan_file in shodan_files:
        success = process_shodan_file(shodan_file, country_name, collection)
        update_progress(success)

    # Process Censys files
    for censys_file in censys_files:
        success = process_censys_file(censys_file, country_name, collection)
        update_progress(success)

    logger.info(f"Completed processing for country: {country_name}")
```

Fig. 3.3.1. Mètode principal del codi. Font: Elaboració pròpia.

2) El següent és un dels codis que proporcionen un gràfic. Aquest en concret mostra el percentatge de servidors amb X ports. La utilitat d'aquest gràfic està explicada més endavant a l'apartat **4.2 Neteja de dades**.

```
# Use MongoDB aggregation to count ports more efficiently
pipeline = [
    {
        "$project": {
            "port_count": {
                "$cond": [
                    {"$isArray": "$ports"},
                    {"$size": "$ports"},
                    0
                ]
            }
        }
    },
    {
        "$group": {
            "_id": "$port_count",
            "count": {"$sum": 1}
        }
    },
    {
        "$sort": {"_id": 1}
    }
]

result = list(collection.aggregate(pipeline))

# Process aggregation results
for item in result:
    num_ports = item["_id"]
    count = item["count"]

    # Limit to maximum ports for the graph
    if num_ports > max_ports:
        continue

    port_counts[num_ports] += count
    total_hosts += count
```

Fig. 3.3.2. Lògica principal del programa. Font: Elaboració pròpia.

3) El tercer, mostra en terminal estadístiques sobre les dades emmagatzemades a la base de dades. Aquest s'ha utilitzat en diversos casos en els que es volia comprovar la correctesa dels gràfics. També per fer llistes llargues combinant noms de països i xifres poder després copiar i enganxar-ho sobre aquest document. Així s'ha millorat la velocitat i la exactitud per mencionar percentatges concrets de països. A la Fig. 3.3.3 s'hi veu part de la funció *main* del programa on connecta amb la base de dades de MongoDB, extreu les dades necessàries (percentatges de nombres de ports amb mitjana, mediana i desviació estàndard) i després mostra per terminal els valors exactes.

```python
# Connect to MongoDB
client, db = connect_to_mongodb(args.mongo_uri, args.database)
if client is None or db is None:
    logger.error("Failed to connect to MongoDB. Exiting.")
    sys.exit(1)

try:
    # Analyze port statistics by country
    stats_by_country = analyze_ports_by_country(db, args.max_ports)

    # Print the table
    print_table(stats_by_country, args.max_ports, args.sort_by, not args.ascending)

except Exception as e:
    logger.error(f"Error analyzing port statistics: {e}")
    if args.verbose:
        import traceback
        logger.error(traceback.format_exc())
```

Fig. 3.3.3. Funció *main* del programa. Font: Elaboració pròpia.

4) L'últim és un programa que automatitza la creació d'informes per a agilitzar la cerca d'informació sobre un *host* per poder contactar amb l'afectat. Aquest crea un informe en format PDF amb la IP, el país, la ciutat i les coordenades (Fig. 3.3.4). Després per cada port obert mostra el número de port, el servei, el venedor i el producte i la versió. La part més interessant és el nom del certificat, on normalment es troba informació sobre el centre on pertany aquest dispositiu. A més, el programa (utilitzant un altre fil) pren captures de pantalla dels serveis HTTP i HTTPS del *host* per veure què s'hi amaga darrere i emmagatzema aquestes en format PNG a la mateixa carpeta de l'informe (Fig. 3.3.5).

Host Information

Property	
IP Address	[redacted]
Country	Germany
City	Stuttgart
Coordinates	[redacted]

Services Information

Property	
Port	80
Service Name	HTTP
Vendor	N/A
Product	OpenSSL
Version	3.0.7

Property	
Port	443
Service Name	HTTPS
Vendor	Apache
Product	HTTPD
Version	N/A
Certificate Name 1	portal.[redacted].de

Fig. 3.3.4. Informe automatitzat amb dades de l'afectat. Font: Elaboració pròpia.

```
try:
    # Create the reports directory if it doesn't exist
    os.makedirs(directory, exist_ok=True)

    # Load input data
    with open(input_file, 'r', encoding='utf-8') as f:
        hosts_data = json.load(f)

    # Set total hosts for progress tracking
    total_hosts = len(hosts_data)
    logger.info(f"Starting processing of {total_hosts} hosts with {max_workers} workers")

    # Initialize browser pool
    initialize_browser_pool(max_workers)

    # Start screenshot worker threads
    screenshot_workers = []
    for i in range(max_workers):
        worker = threading.Thread(target=take_screenshot_worker, name=f"Screenshot-Worker-{i+1}")
        worker.daemon = True
        worker.start()
        screenshot_workers.append(worker)
    logger.info(f"Started {len(screenshot_workers)} screenshot worker threads")

    # Process hosts in parallel
    with concurrent.futures.ThreadPoolExecutor(max_workers=max_workers) as executor:
        futures = [executor.submit(process_host, host, directory) for host in hosts_data]
        for i, future in enumerate(concurrent.futures.as_completed(futures)):
            try:
                future.result()  # This will raise any exception that occurred during processing
            except Exception as e:
                logger.error(f"Error in host processing thread {i}: {e}")

    logger.info("All hosts processing complete. Waiting for remaining screenshots...")
    # Signal screenshot workers to stop and wait for them to finish
    for _ in range(max_workers):
        screenshot_queue.put(None)
    for worker in screenshot_workers:
        worker.join(timeout=30)  # Wait up to 30 seconds for each worker
```

Fig. 3.3.5. Mètode principal on s'inicialitzen els processos de captura d'imatge i es creen els informes. Font: Elaboració pròpia.

4. Avaluació

4.1 Dades obtingudes

El procés d'obtenció de dades de Censys i Shodan ha estat realitzat durant el primer trimestre de l'any 2025. Amb la finalitat de fer viable aquest estudi amb el marge de temps estipulat s'han obtingut les dades de 20 països europeus amb diferents PIB (Producte Interior Brut) i diferents IDH (Índex de Desenvolupament Humà) per poder comparar l'estat de seguretat informàtica entre ells. El llistat, de major a menor nombre de servidors obtinguts, és el següent: Alemanya (1449), França (1114), Regne Unit (1047), Irlanda (1000), Suècia (840), Itàlia (150), Països Baixos (82), Espanya (54), Turquia (52), Suïssa (44), Grècia (36), Polònia (27), Bèlgica (24), Finlàndia (24), Hongria (13), Portugal (12), Romania (10), Àustria (10), Dinamarca (6) i Eslovàquia (2). En total 5.996 servidors entre 20 països com mostra la Fig. 4.1.2.

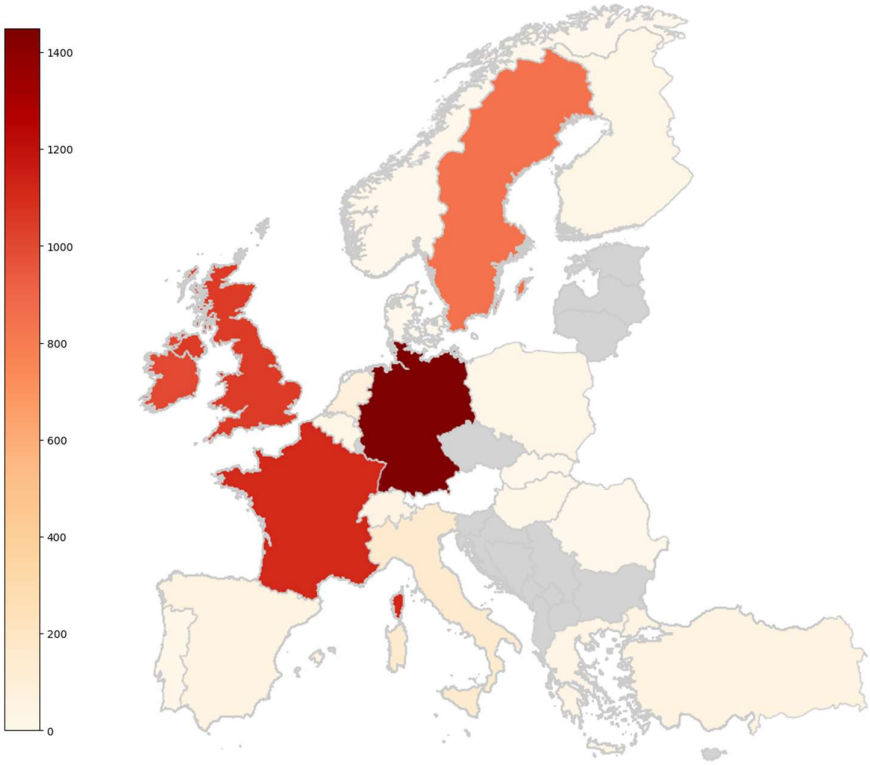


Fig. 4.1.1. Distribució dels *hosts* en el mapa europeu. Font: Elaboració pròpia.

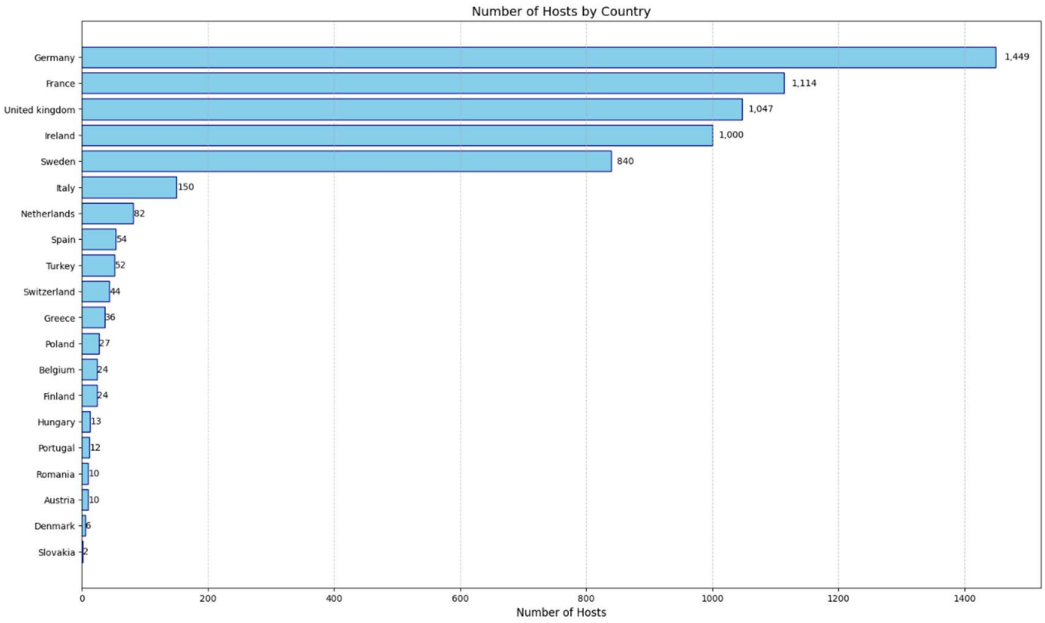


Fig. 4.1.2. Distribució del nombre de *hosts* per país. Font: Elaboració pròpia.

Aquest resultats s'han relativitzat a la superfície del país per buscar la densitat de servidors trobats per cada 1000 km^2. Per fer això s'ha consultat a la API *restcountries.com* sobre la superfície de cada país [17]. Com mostra la Fig. 4.1.3 destaca el gran nombre de servidors per cada 1000 km^2 d'Irlanda (14.2) amb quasi quatre vegades més que el segon país (4.3). En segon i tercer lloc hi apareixen Regne Unit i Alemanya respectivament. Els dos destaquen sobre les següents posicions en el gràfic amb més del doble de densitat (4.3 i 4.05 respectivament, sobre 2.01).

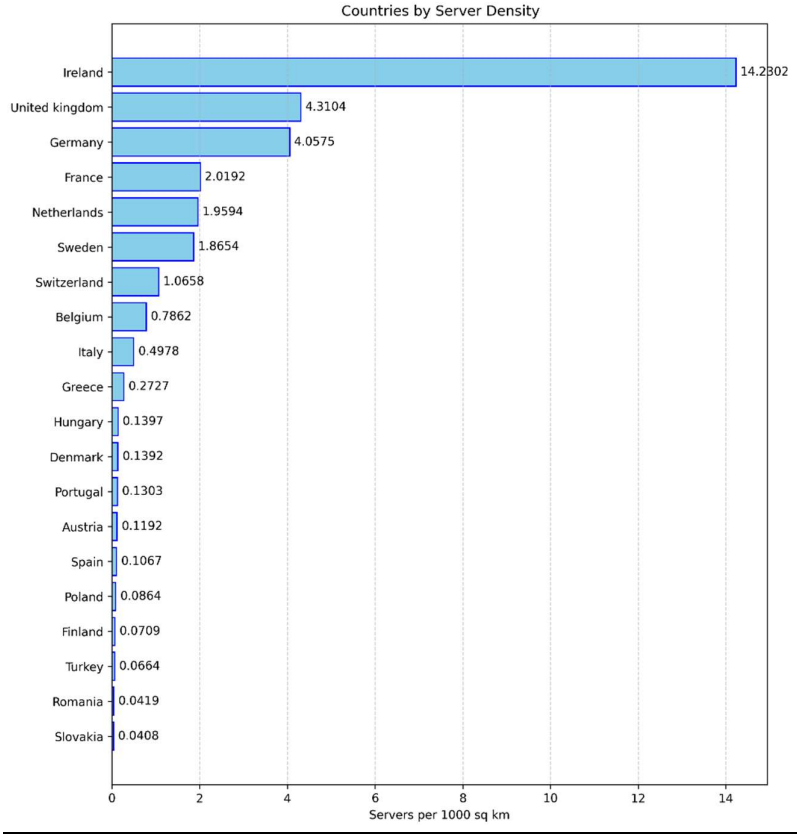


Fig. 4.1.3. Densitat de servidors per cada 1000 km^2. Font: Elaboració pròpia.

També s'ha relacionat el nombre de servidors obtinguts amb el valor del PIB d'aquell país. La hipòtesi rere aquest anàlisi és que la relació entre PIB i nombre de servidors exposats hauria de ser inversament proporcional. A major PIB, major hauria de ser la inversió en tecnologia i menor hauria de ser el nombre de servidors exposats a Internet. Cal aclarir que el fet que un servidor estigui exposat a Internet no és directament perillós, al contrari, és un fet comú. Però quan Censys o Shodan marquen un servidor com a "relacionat amb l'entorn mèdic" no sol ser un servidor que hauria d'estar exposat. Els serveis HTTP que publiquen aquests *hosts* solen mostrar pàgines web on es veu clarament que es tracta d'un dispositiu mèdic (com una càmera de vigilància) o d'un servidor mèdic (com un PACS). La Fig. 4.1.4, però, mostra tots els *hosts* obtinguts sense el filtratge de *honeypots*.

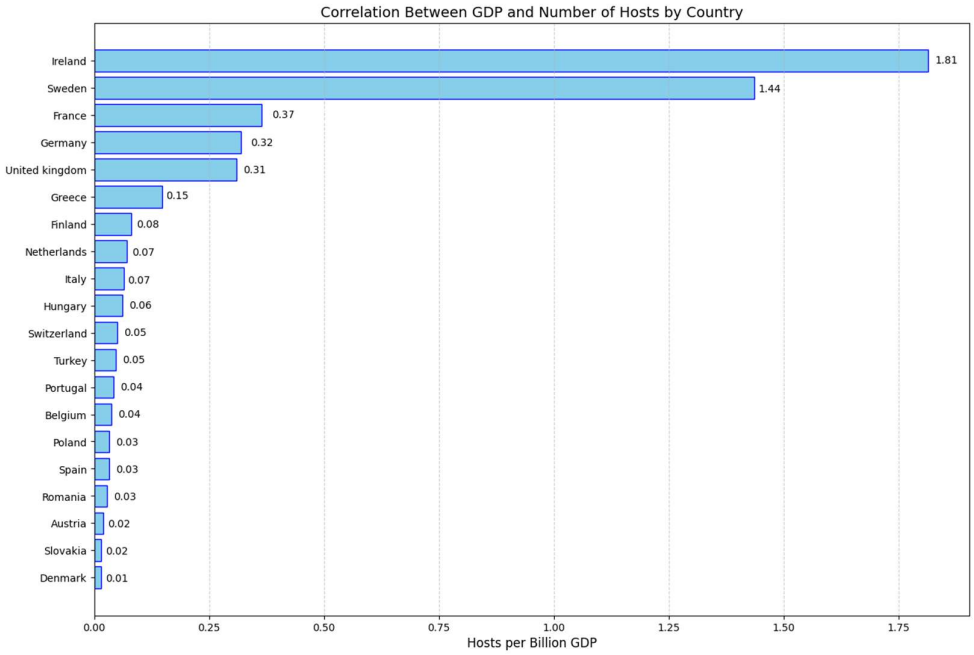


Fig. 4.1.4. Relació entre el PIB d'un país i el nombre de *hosts* exposats. Font: Elaboració Pròpia

En el gràfic s'hi pot apreciar com Irlanda presenta major proporció pel que fa a la relació entre PIB i nombre de servidors exposats (1.81 *hosts* per cada bilió d'euros), tot i no ser una de les principals potències econòmiques europees (la número 13, Octubre 2024) [18]. Un dels possibles factors que fa que Irlanda encapçali els dos gràfics seria que aquest país s'ha establert com a capital tecnològica europea albergant les seus europees de moltes multinacionals tecnològiques com Google, Facebook, Apple, Microsoft o Intel entre d'altres. S'ha establert com a capital tecnològica per diversos factors com les polítiques fiscals favorables, la bona connexió entre Europa i Amèrica o fins i tot el clima, que en ser temperat redueix els costos de refrigeració dels grans centres de dades d'aquestes empreses.

En aquest gràfic també hi destaca Suècia (que no ho feia a l'anterior) amb una relació d'1.44 *hosts* per cada bilió d'euros. Suècia tampoc és una de les principals potències econòmiques europees (la número 12, Octubre 2024) [18], però si una de les que més inverteix en recerca I+D (prop d'un 3.4% del PIB anual, 2023) [19]. Aquest fet, juntament amb que és un dels principals referents quant a la digitalització del sistema sanitari, explicaria els valors obtinguts en aquest gràfic.

Per últim, s'ha comparat el nombre de servidors amb la puntuació obtinguda de cada país a l'índex de seguretat informàtica adjudicada per la Unió Internacional de Telecomunicacions l'any 2024 [20]. Aquest índex categoritza cada país dins de cinc nivells de seguretat informàtica segons l'estat d'aquesta. Es valoren cinc categories sobre una puntuació màxima de 20 punts per obtenir la mitjana final total: mesures legals, mesures tècniques, mesures organitzatives, capacitat de desenvolupament i mesures cooperatives. La hipòtesi rere aquest gràfic és poder comparar la suposada seguretat de cada país amb el nombre de *hosts* que hi ha exposats. La Fig. 4.1.5 mostra aquesta relació.

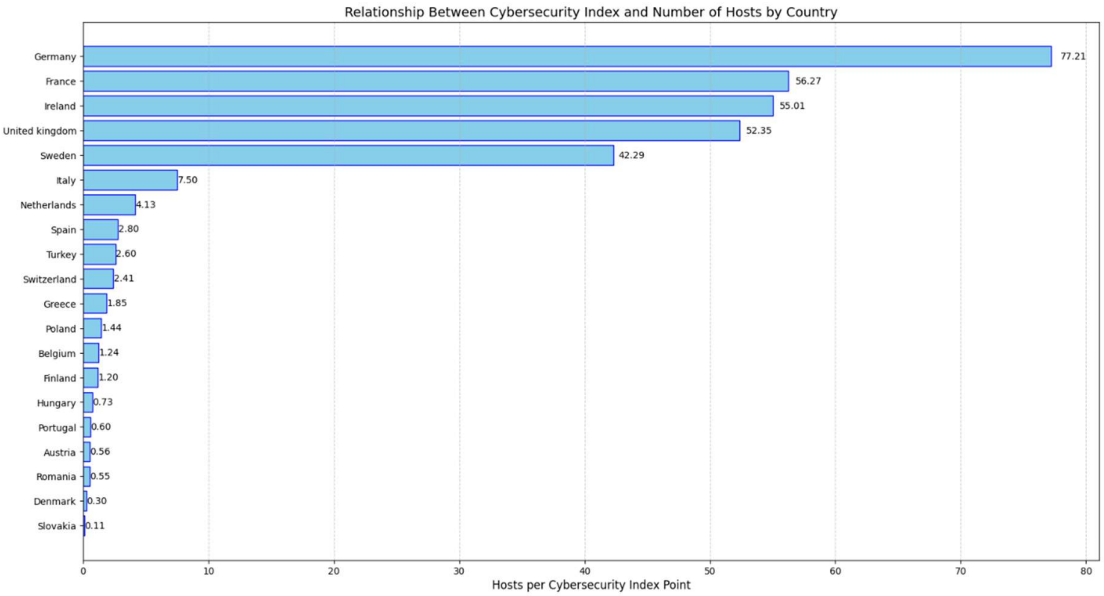


Fig. 4.1.5. Relació entre l'índex de Ciberseguretat i el nombre de *hosts* exposats. Font: Elaboració Pròpia.

En aquest gràfic s'hi observa una semblança molt directa amb la Fig. 4.1.2, on hi apareix el nombre de *hosts* per país. Això passa perquè els valors obtinguts per cada país a l'índex de seguretat informàtica són molt semblants, sent 20 el valor més alt (Regne Unit, Itàlia, Turquia, Dinamarca i Finlàndia) i 17,746 el valor més baix (Hongria). Això mostra el nivell de ciberseguretat a Europa en comparació amb països d'Àsia Central com Turkmenistan (5,17), Orient Mitjà com el Líban (6,474) o fins i tot Sud Amèrica amb Argentina (10,302).

Tot i això, no tots els països europeus tenen puntuacions tan altes (6,752. Bòsnia i Hercegovina).

4.2 Neteja de les dades

Després de consultar les dades obtingudes sobre els *hosts* s'ha detectat que la mitjana és de 57 ports oberts per *host*. Això és sospitós ja que és poc probable que tots els servidors tinguin de mitja 57 serveis per donar cobertura a 57 accions diferents. El fet que aquest nombre sigui tan elevat s'ha relacionat amb dues possibilitats: 1) La primera seria que es tractés de diversos dispositius que comparteixen IP pública utilitzant un DNAT. 2) La segona seria l'existència de *honeypots* entre aquestes dades. Un *honeypot* és un servidor desplegat per l'administrador de sistemes de la infraestructura que pretén simular un servidor real utilitzat en el centre mèdic, però amb moltes bretxes de seguretat, amb la finalitat que els atacants utilitzin les seves eines per explotar-les i els encarregats de la seguretat del centre puguin preparar els servidors reals front aquestes amenaces.

Fig. 4.2.1. Pàgina HTML que exposa un servidor marcat com a "*Medical Device*". Font: Elaboració pròpia.

Per confirmar la hipòtesi de la existència de *honeypots* s'han estudiat els patrons de configuració i estat d'un servidor real front un *honeypot* i s'ha determinat que: 1) El nombre de ports ha de ser raonable. 2) És comú que el servidor real es trobi dins la infraestructura de xarxa del centre mèdic i el *honeypot* utilitzi un servidor extern per ser desplegat. 3) El

nombre de DNS *records* per servidor hauria de ser proporcional (en major o menor mesura) al nombre de ports oberts.

Un cop determinats aquests patrons de comportament s'han fet els gràfics per confirmar la hipòtesi.

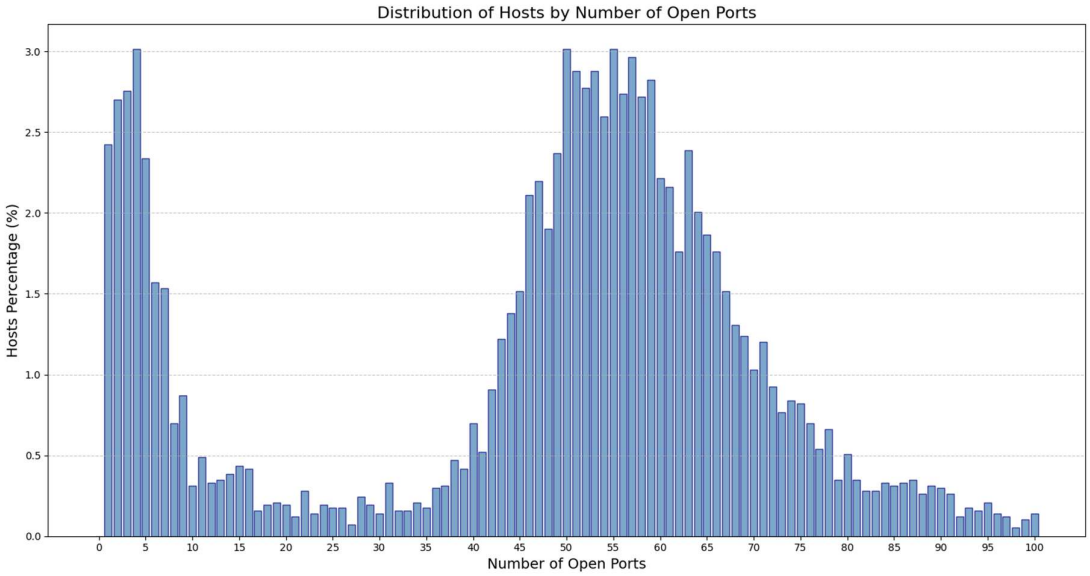


Fig. 4.2.2. Nombre de ports oberts per servidor. Font: Elaboració pròpia.

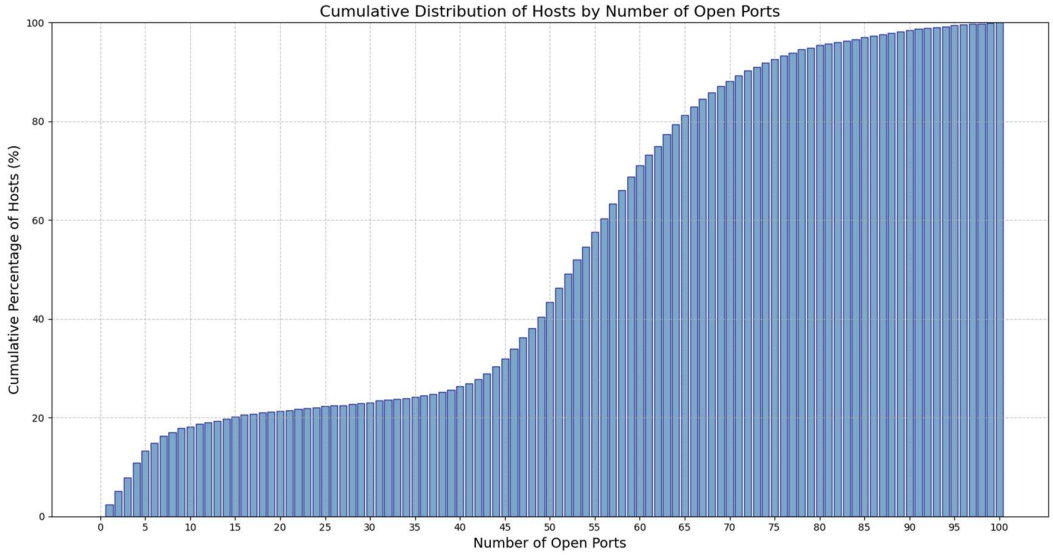


Fig. 4.2.3. Nombre de ports oberts per servidor (percentatges acumulats). Font: Elaboració pròpia.

En la Fig. 4.2.2 s'hi pot veure la distribució dels servidors per nombre de ports oberts. S'ha fet un tall en el valor 100 de ports oberts ja que es troba inviable que un servidor tingui una xifra tan elevada de serveis (hi ha servidors on la xifra augmenta fins als 3000 ports oberts). En aquest gràfic hi apreciem dues agrupacions de servidors. A la part esquerra, s'hi pot observar que un nombre elevat de servidors té un nombre raonable de ports oberts (d'1 a 10). A la part més a la dreta s'hi aprecia un altre gran grup de servidors que té un nombre ja sospitós de ports oberts (de 35 a 80), que es consideren excessius per ser un únic servidor.

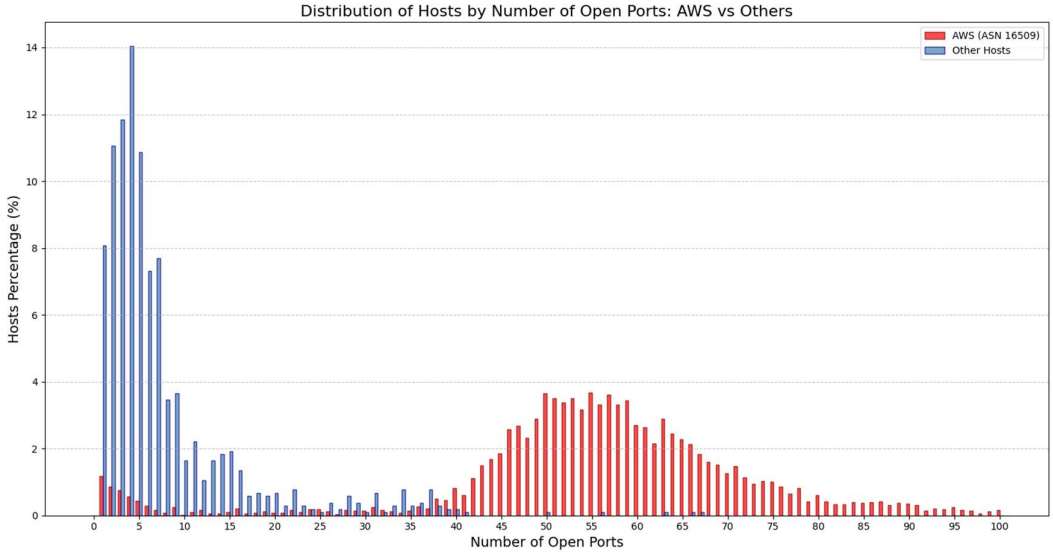


Fig. 4.2.4. Nombre de ports oberts per servidor diferenciat per Sistema Autònom. Font: Elaboració pròpia.

A la Fig. 4.2.4 s'hi pot apreciar una sospitosa coincidència amb la Fig. 4.2.2 després de filtrar els equips entre el número de sistema autònom al que pertanyen. En vermell apareixen els dispositius pertanyents al nombre de sistema autònom 16509 (AWS), en blau tots els altres sistemes autònoms. El sector més a l'esquerra de *hosts* amb un menor nombre de ports oberts està situat clarament en sistemes autònoms no relacionats amb AWS (Amazon Web Services), una de les principals empreses de computació en el núvol. A la part dreta del gràfic s'hi observa (sobre el mateixos valors que a la Fig. 4.2.2) una altra agrupació de *hosts*. Aquests, però, pertanyen al sistema autònom (AS) d'AWS.

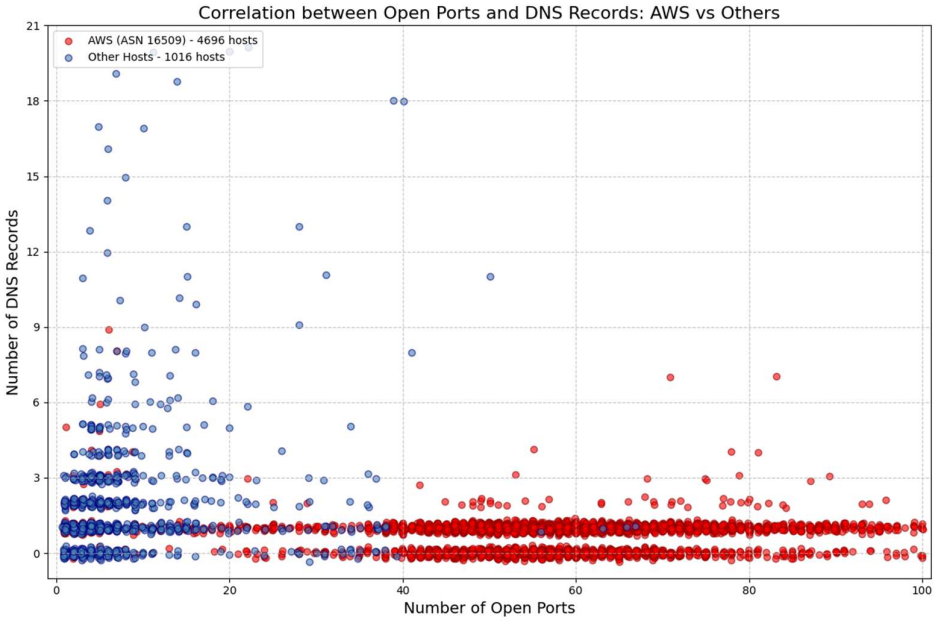


Fig. 4.2.5. Relació entre el nombre de ports oberts i els DNS *records* que apunten al servei. Font: Elaboració pròpia.

En la tercera Fig. 4.2.5, podem comprovar que la tercera afirmació sobre els servidors reals no es compleix per un grup de *hosts*. A l'esquerra s'hi aprecia un grup de servidors amb un nombre inferior de ports oberts que tenen entre 1 i 6 DNS apuntant als seus serveis. A la part esquerra, però, s'hi aprecia com un gran grup de servidors que formen part del AS d'AWS només tenen 1 o cap DNS apuntant als seus serveis, cosa que estranya ja que a major nombre de serveis, major hauria de ser el nombre de DNS apuntant-hi, tot i que per aquesta afirmació hi trobem diverses excepcions: 1) Per als servidors *cloud* d'AWS es solen utilitzar estructures on un sol registre DNS pot dirigir el trànsit a múltiples serveis, utilitzant balancejadors de càrrega, API *gateways* o servidors *proxy*. 2) Es podria tractar d'estratègies de seguretat el fet de no exposar un gran nombre de DNS per passar més desapercebut. 3) AWS utilitza altres mètodes a part del DNS per descobrir serveis. Els servidors dins de la xarxa AWS que només s'han de connectar amb altres servidors de la xarxa AWS ho fan de manera privada a través d'IPs privades o de DNS que assigna automàticament AWS als seus *hosts*.

Amb aquestes tres demostracions provem com a vàlida la hipòtesi sobre que hi ha molts d'aquests servidors que són *honeypots*. Aquests no haurien de sortir en el recompte global d'estadístiques per analitzar l'estat de la seguretat informàtica als centres mèdics correctament, ja que no es tracta de servidors d'ús real. Per aquest motiu, s'ha fet un garbell sobre tots els *hosts* i s'han eliminat de les estadístiques els que tinguessin més de 30 ports oberts, que és el límit relatiu que es mostra en les tres figures.

Dels 5996 servidors totals abans del garbell, ara n'han quedat 1317 amb menys de 30 ports oberts (un 21,96%). La Fig. 4.2.6 mostra la distribució per país.

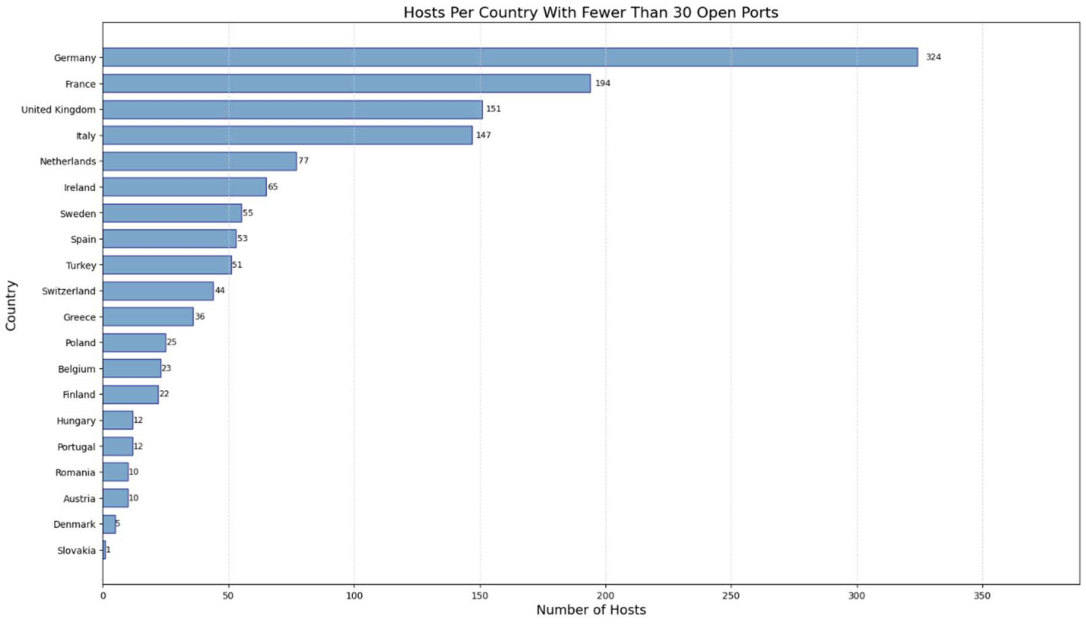


Fig. 4.2.6. Nombre de ports oberts per servidor per país amb el filtratge de *honeypots*. Font: Elaboració pròpia.

Com es veu a la figura, el nombre de *hosts* s'ha reduït considerablement. A més, s'hi aprecia com els valors entre els països ara són molt més propers i no s'accentua tant la diferència entre els quatre primers del rànquing amb la resta de països. El filtre més gran ha estat el d'Irlanda, on només s'han mantingut un 6,5% dels *hosts* originals, seguit de Suècia, on s'han mantingut un 6,55% dels *hosts* originals. Per altra banda, països com Itàlia o Espanya que

partien amb un volum de *hosts* de nivell mitjà, s'han quedat en aquest volum ja que s'han mantingut un 98% i un 98,15% respectivament dels *hosts*.

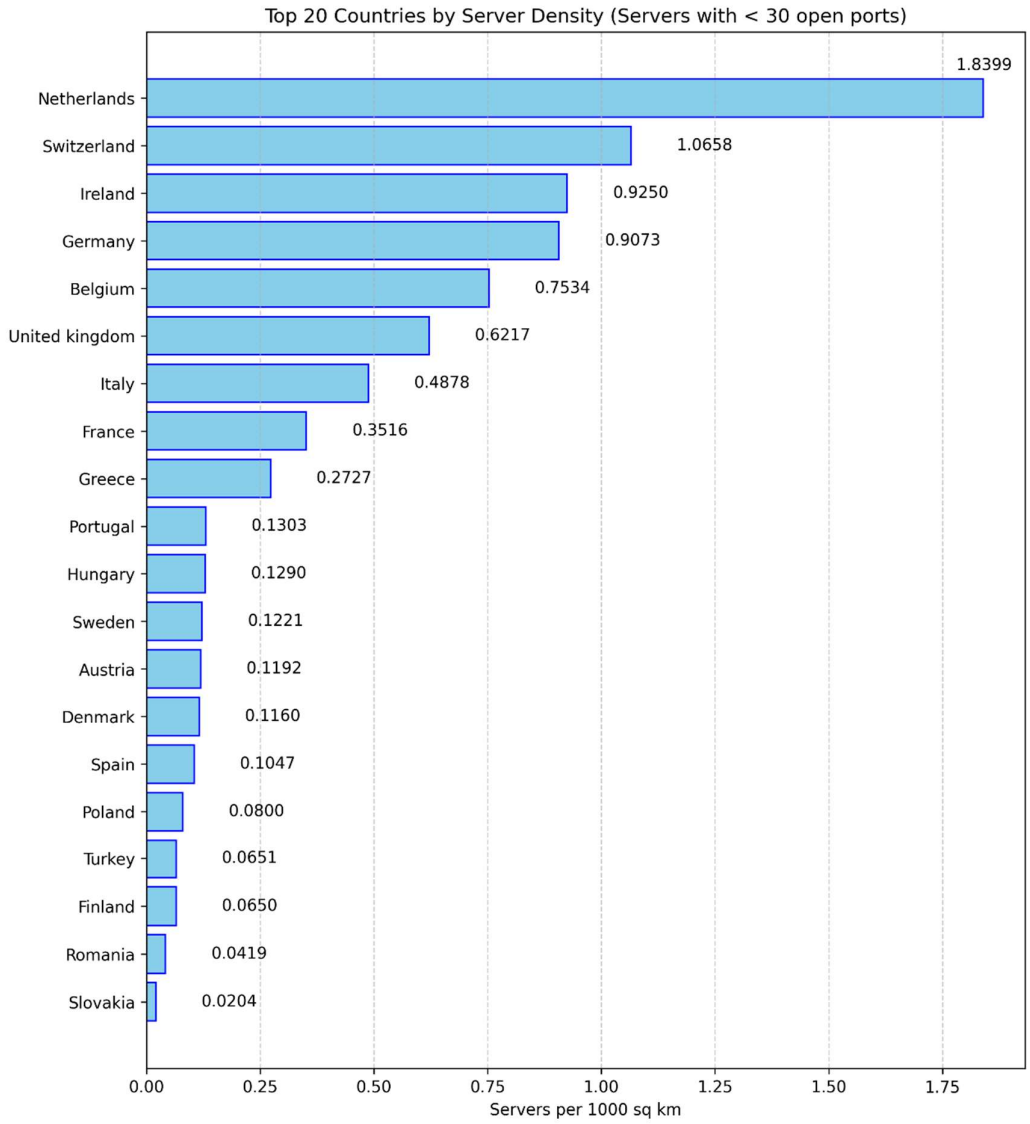


Fig. 4.2.7. Densitat de servidors per cada 1000 km² filtrat per *honeypots*. Font: Elaboració pròpia.

A la Fig. 4.2.7 s'hi pot observar la densitat de servidors com a la Fig. 4.1.3, ara filtrant amb el criteri de *honeypots*. Si es comparen els dos gràfics s'observa un gran canvi en quant al rànquing de països. En aquest segon, Irlanda no surt destacat. Es mostra una estructura més

homogènia amb Països Baixos com a país amb més servidors per cada 1000 km^2, però amb menys del doble de densitat sobre el segon país, Suïssa. La hipòtesi que s'ha plantejat anteriorment sobre el motiu pel qual Irlanda presentava una densitat de servidors tan elevada, agafa força. Si es comparen els dos gràfics es pot veure com Irlanda presenta una gran densitat de servidors en el seu territori però no de servidors mèdics reals. En canvi ara Països Baixos i Suïssa han agafat protagonisme.

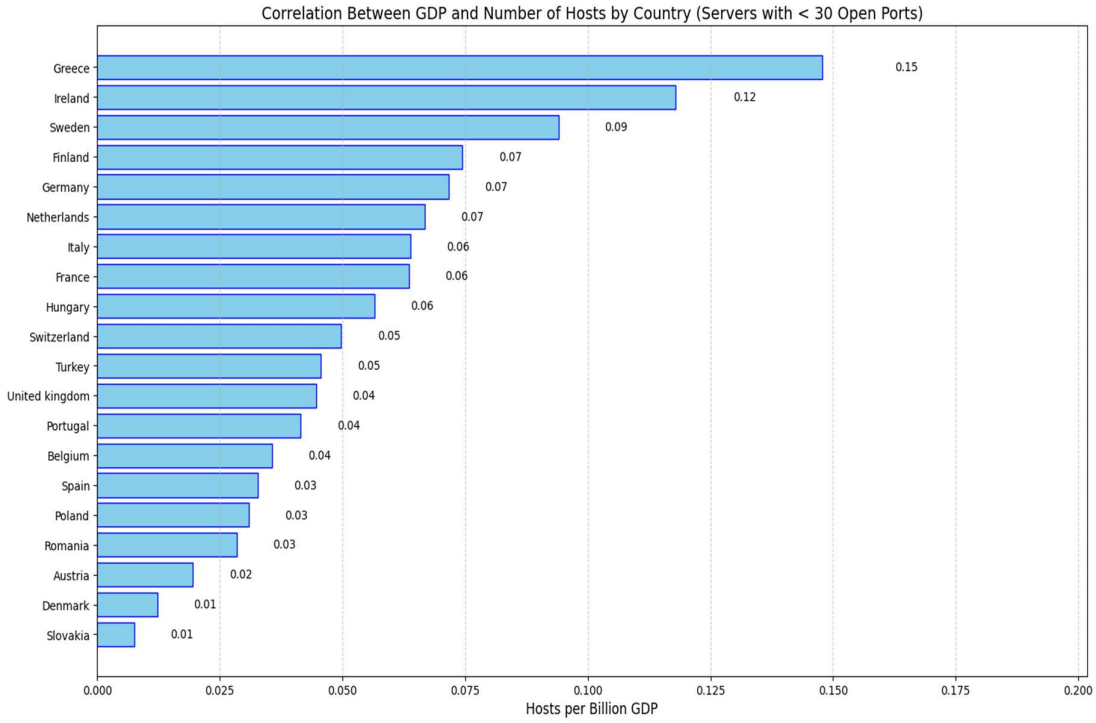


Fig. 4.2.8. Relació entre el PIB del país i el nombre de *hosts* exposats filtrats per *honeypots*. Font: Elaboració pròpia.

A la Fig. 4.2.8 s'hi pot observar la relació entre el PIB del país i el nombre de dispositius mèdics exposats a Internet. S'hi observa una gran diferència amb la Fig. 4.1.4, que mostra la mateixa relació sense filtrar. En aquesta última s'hi observava una clara distància entre Irlanda, Suècia i la resta de països. Ara es mostra Grècia al capdamunt d'aquest gràfic. Una de les principals causes podria ser el baix PIB del país comparat amb un desplegament de xarxa informàtica comú entre els països europeus. La hipòtesi rere aquesta afirmació seria

que Grècia ha tingut una transformació digital accelerada gràcies al finançament de la Unió Europea com a part del programa de recuperació "Greece 2.0" [21]. A part, els països nòrdics com Suècia o Finlàndia també surten representats a la part alta del gràfic, enfortint la hipòtesi sobre la inversió d'aquests països en la digitalització dels centres de salut.

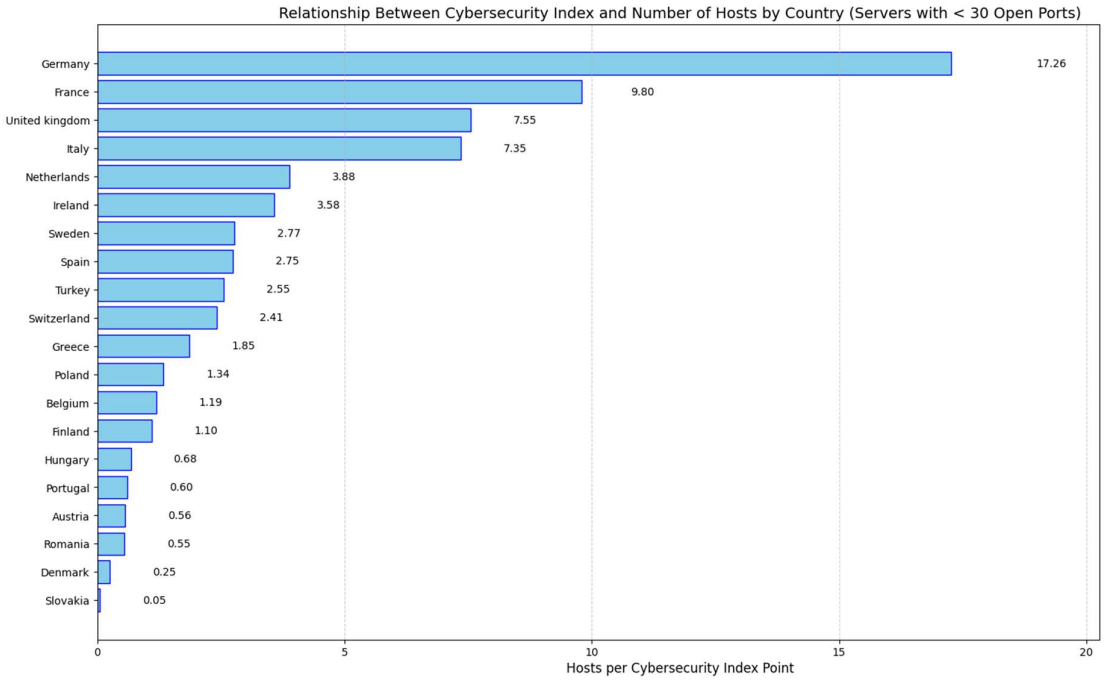


Fig. 4.2.9. Relació entre l'índex de Ciberseguretat i el nombre de *hosts* exposats filtrats per *honeypots*. Font: Elaboració Pròpia.

En aquesta última Fig. 4.2.9, s'hi mostra la relació entre la puntuació obtinguda per cada país en l'Índex de Ciberseguretat amb el nombre de servidors mèdics exposats a Internet. En aquest cas s'hi aprecia una diferència notable front la Fig. 4.1.5, però no respecte la Fig. 4.2.6. Igual que en el cas de la Fig. 4.1.5 el nombre de servidors exposats i la relació entre aquest nombre i la puntuació a l'Índex de Ciberseguretat està molt lligat ja que per als països seleccionats en aquest estudi, els valors a l'índex son propers a la perfecció i no s'aprecien desviacions considerables.

4.3 Anàlisi dels resultats

En aquest apartat es diferencia dos grups ens els resultats obtinguts: un primer grup on es mostra l'anàlisi quant a serveis exposats i sistemes operatius dels *hosts*, i un segon grup més interessant que mostra els diversos CVE amb el seu respectiu CVSS que s'han trobat.

Un CVE (*Common Vulnerabilities and Exposures*) constitueix un identificador estandarditzat (format per "CVE" seguit de l'any de publicació i un número seqüencial) assignat per una CNA (CVE *Numbering Authority*) a vulnerabilitats de seguretat documentades en components de software. Serveix per a que el personal de seguretat informàtica pugui reaccionar ràpidament quan es troba una vulnerabilitat en un software.

El CVSS (*Common Vulnerability Scoring System*) uneix un valor numèric a aquest CVE que proporciona informació sobre la gravetat d'aquell forat de seguretat. Mitjançant mètriques base (vector d'atac, complexitat, privilegis requerits, interacció d'usuari, abast, confidencialitat, integritat i disponibilitat), mètriques temporals (maduresa de l'explotació, nivell de dificultat a l'hora de corregir l'error, confiança de l'informe) i mètriques ambientals (modificadors específics del context), resulta en una puntuació decimal entre 0 i 10 que classifica la gravetat i permet la priorització objectiva dels esforços de mitigació en entorns de seguretat informàtica.

Els CVE poden ser utilitzats per explotar vulnerabilitats als servidors que no tinguin els softwares actualitzats a la última versió. Per tant, que algun *host* mèdic presenti CVEs coneguts, és un problema greu de seguretat

4.3.1 Anàlisi de característiques dels *host*

S'ha avaluat la informació que presenta cada *host* vinculada a la ciberseguretat. La primera figura (Fig. 4.3.1) mostra la distribució de serveis proporcionats per als *hosts* una vegada filtrats per nombre de ports i per país.

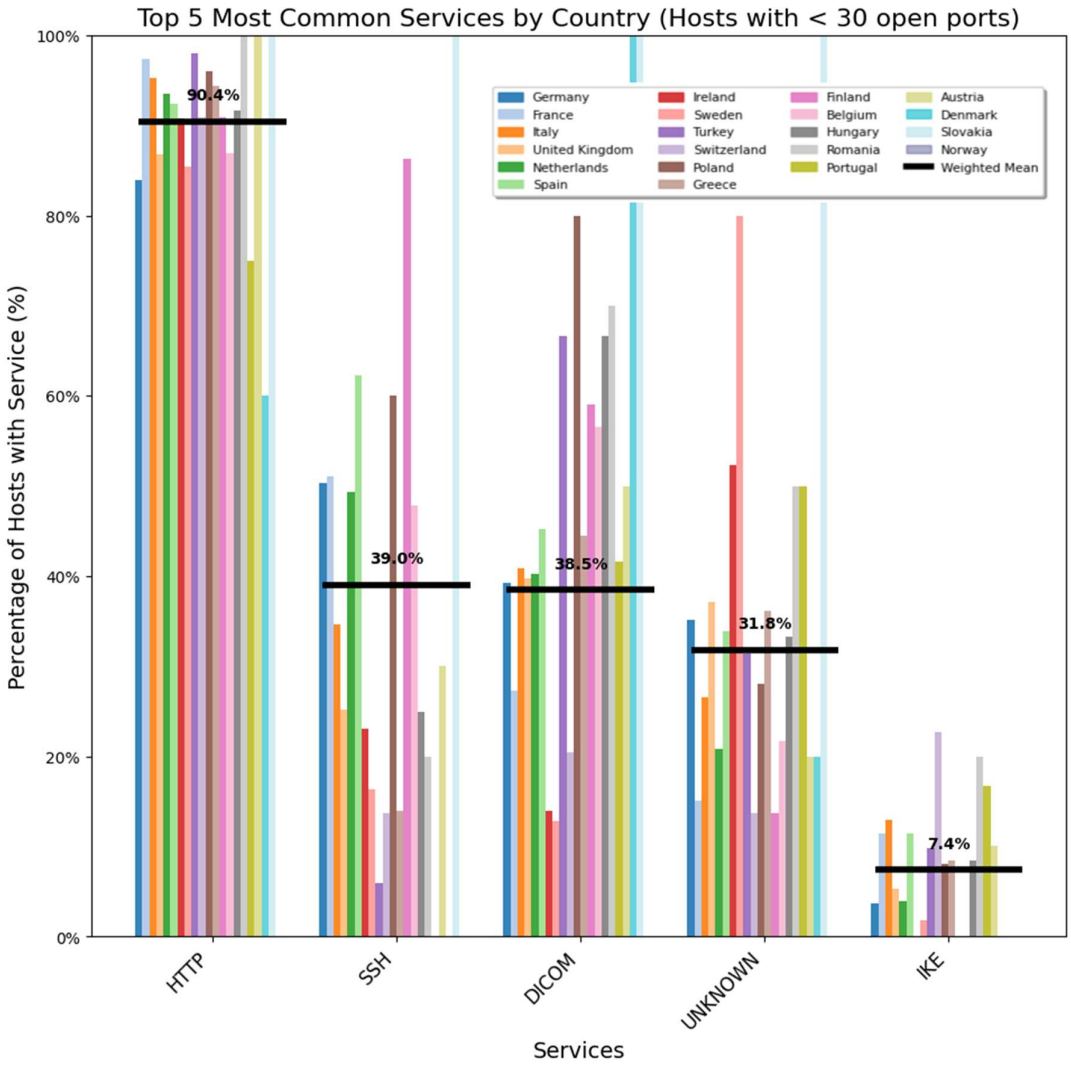


Fig. 4.3.1. Els serveis més comuns que presenten els *hosts* filtrats per *honeypots*. Font: Elaboració pròpia.

S'ha calculat la mitjana ponderada per a cada servei ja que hi ha països que presenten molt poc nombre de *hosts* i el percentatge d'aparició d'un servei en aquells països és molt escandalós. És el cas d'Eslovàquia amb el servei SSH que presenta un 100% d'ús ja que només té un únic *host*. Amb aquestes dades podem observar que els 5 serveis més utilitzats són els d'HTTP, SSH, DICOM i IKE. Hi ha una part de serveis dels quals no s'ha pogut obtenir informació. També s'han trobat servidors utilitzant serveis com MYSQL, POP3 o SMTP però en una molt menor mesura. S'hi pot observar que hi ha una gran paritat en els percentatges relacionats amb l'HTTP, ja que aquest és un protocol molt comú en general.

Però la paritat desapareix als percentatges relacionats amb el protocol SSH, ja que aquest és més específic i només es presenta als servidors que es vulguin controlar remotament o en els que es vulgui accedir als fitxers de manera remota utilitzant un navegador web. És evident també l'aparició del protocol DICOM al tractar-se de dispositius mèdics, ja que aquest és utilitzat per a l'emmagatzematge i la recuperació d'imatges mèdiques. També s'hi troba el protocol IKE (*Internet Key Exchange*) que s'utilitza per establir connexions VPN entre dos *hosts* a través d'Internet. Això suggereix que una part dels servidors trobats permeten una connexió VPN remota per tal d'accedir als equips de manera segura. La última part a comentar són els serveis que ni Censys ni Shodan han estat capaços d'ubicar. Probablement el motiu rau en què aquest grup de protocols siguin protocols dedicats i no siguin estàndards. Els dos motors de cerca assumeixen el tipus de servei mitjançant el port que utilitza i els *banners* d'aquest protocol, però si no són estàndards no poden identificar correctament de quin servei es tracta.

La següent figura (Fig. 4.3.2) mostra la distribució que presenten els sistemes operatius dels *hosts*.

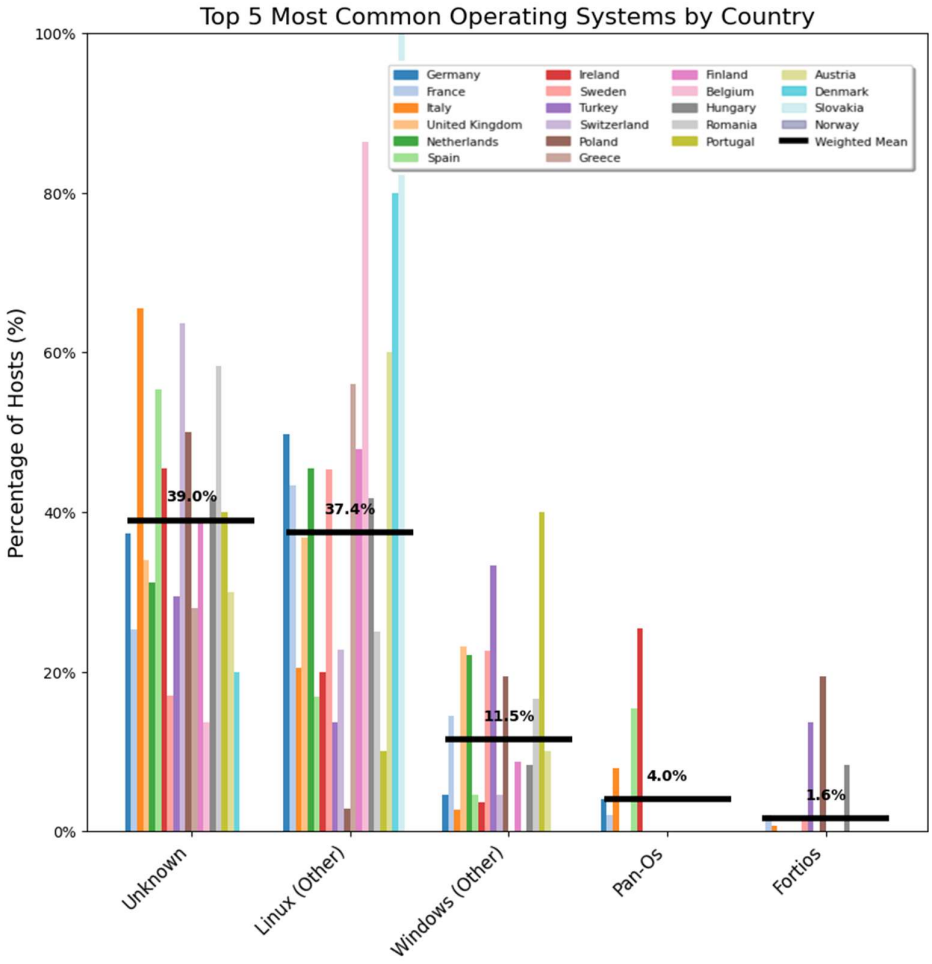


Fig. 4.3.2. Distribució dels sistemes operatius en els *hosts* filtrats per *honeypots*. Font: Elaboració pròpia.

Com s'hi pot observar hi ha una gran varietat de percentatge d'aparició per país, però els sistemes operatius predominants són Linux i Windows en diverses versions. Hi ha una petita part que presenta Pan-Os i Forti-Os, segurament *firewalls* de Palo Alto Networks i Fortinet, respectivament, que apareixen com a dispositius mèdics possiblement perquè es troben en els recintes d'un centre mèdic. L'últim gran grup seria el de sistemes operatius no detectats. Seguint la línia dels protocols podria ser que fossin sistemes operatius dedicats i molt personalitzats i que les firmes d'aquests no coincidissin amb les bases de dades dels motors de cerca. Una altre hipòtesi seria l'ocultació deliberada; molts dispositius estan configurats per no revelar informació sobre el seu sistema operatiu com a mesura de seguretat front a atacs externs.

4.3.2 Anàlisi de vulnerabilitats

Com s'ha explicat en apartats anteriors, Shodan presenta la possibilitat de detectar els CVE que presenta cada *host*. La següent figura (Fig. 4.3.3) mostra els més comuns entre tots els països.

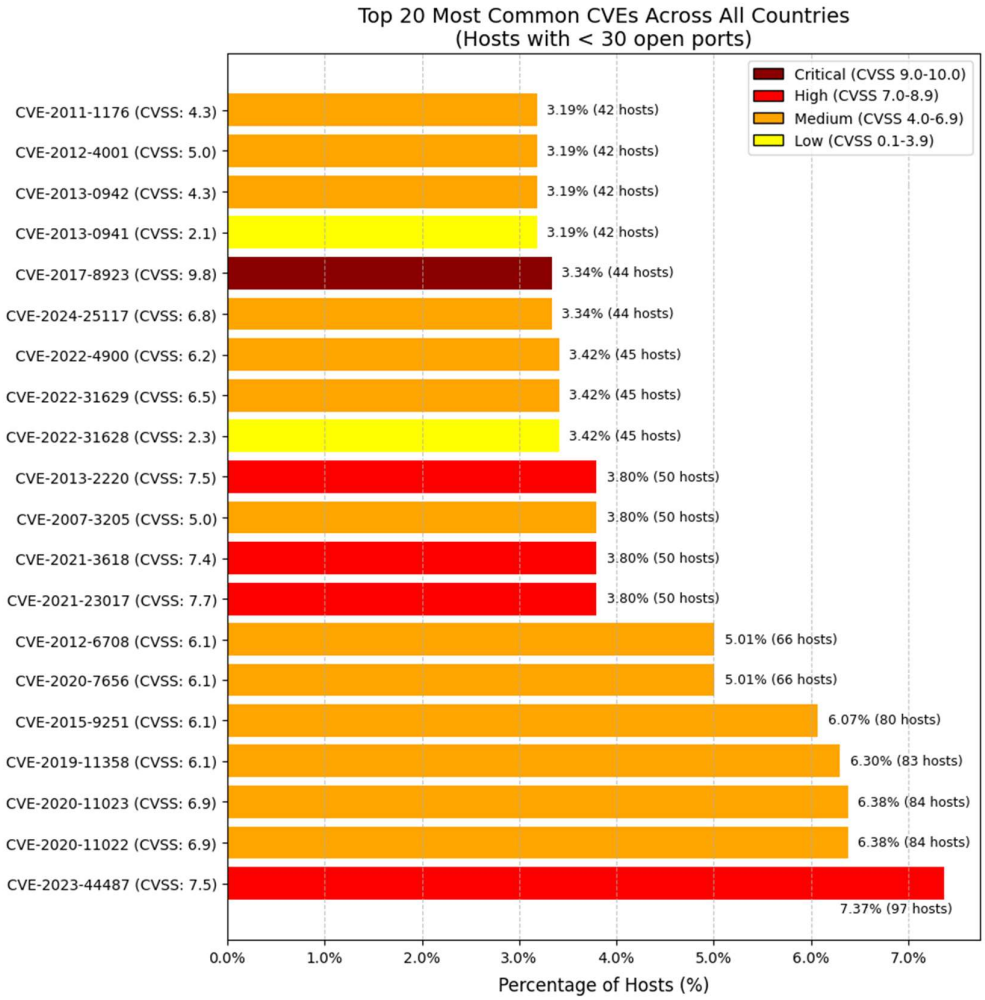


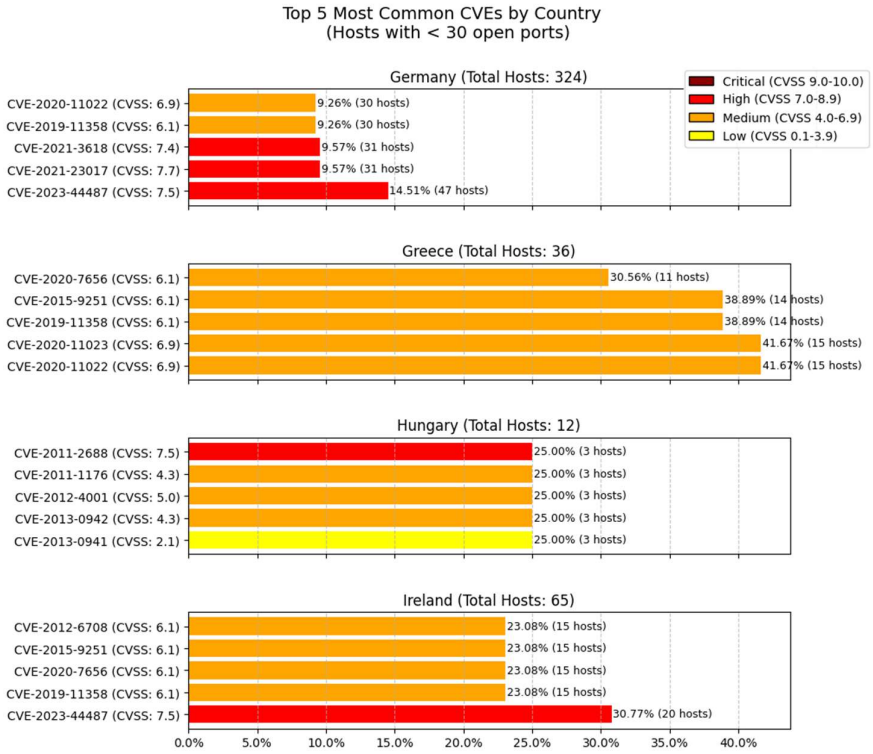
Fig. 4.3.3. CVEs més comuns per tots els països filtrat per *honeypots*. Font: Elaboració pròpia.

En aquesta figura es mostra els CVE més comuns per a tots els països amb un rang de colors que defineix el crític que és. S'hi pot observar una gran varietat de CVE, la gran majoria amb una puntuació de perillositat mitjana-alta. El més comú és el CVE-2023-44487 que es

tracta d'una explotació de vulnerabilitat sobre el protocol HTTP/2. Aquest protocol implementa una característica anomenada "multiplexació de fluxos" i pot ser utilitzat per crear una sobrecàrrega al *host* després d'iniciar una gran quantitat de fluxos i cancel·lar-los immediatament [22]. Aquest atac de denegació de servei és una gran amenaça per als servidors mèdics que necessiten estar funcionant a totes hores.

El CVE més crític, però, és el CVE-2017-8923 amb una puntuació de 9.8 degut a la seva senzillesa d'execució. Es tracta també d'un atac de denegació de servei que explota una vulnerabilitat de les versions anteriors a la 7.1.5 de PHP on encadenant cadenes molt llargues de caràcters, resulta en un error de desbordament (*out of bounds*) detenint així el funcionament de l'aplicació [23].

La Fig. 4.3.4 representa els CVEs més comuns distribuïts per països. Degut al gran nombre de vulnerabilitats diferents, només es comentaran les més importants.



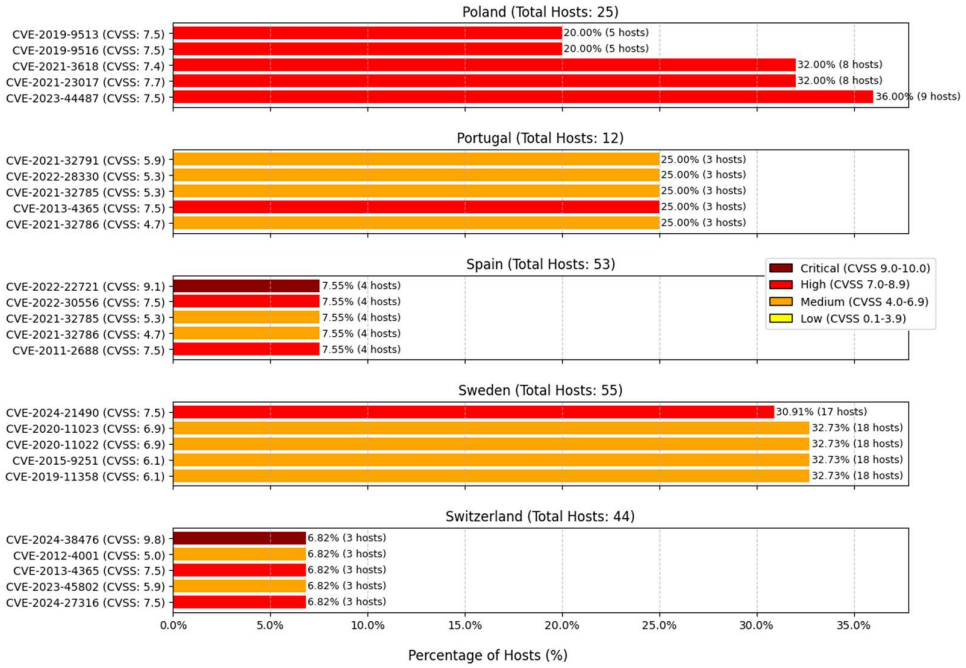


Fig. 4.3.4. CVEs més comuns distribuïts per països i filtrats per *honeypots*. Font: Elaboració pròpia.

Com es pot observar a la figura, hi ha molts CVE diferents ja que cada dispositiu presenta un software instal·lat i unes versions de característiques similars però no iguals. El més crític de tots és el CVE-2024-38476 que presenten 3 *hosts* a Suïssa. Es tracta d'un atac sobre el nucli d'Apache HTTP Server que permet a l'atacant realitzar divulgació d'informació. Aquest és un CVE relativament nou, però destaca la antiguitat que tenen els CVE d'Hongria. Els 5 més comuns van ser trobats entre 2011 i 2013, més de 10 anys des de la publicació d'aquest estudi. El CVE-2011-1176 explota una vulnerabilitat de la versió 2.2.11-01 i 2.2.11-02 del mòdul per a servidors web d'Apache "mpm-itk". Aquest mòdul ha rebut diverses actualitzacions, essent la versió actual la 2.4.7-04 llançada el 14 de Febrer de 2016. Aquesta situació posa de manifest el poc manteniment que han tingut aquests servidors.

Per aquest motiu s'ha volgut avaluar la distribució de longevitat dels CVE. La Fig. 4.3.5 mostra els resultats.

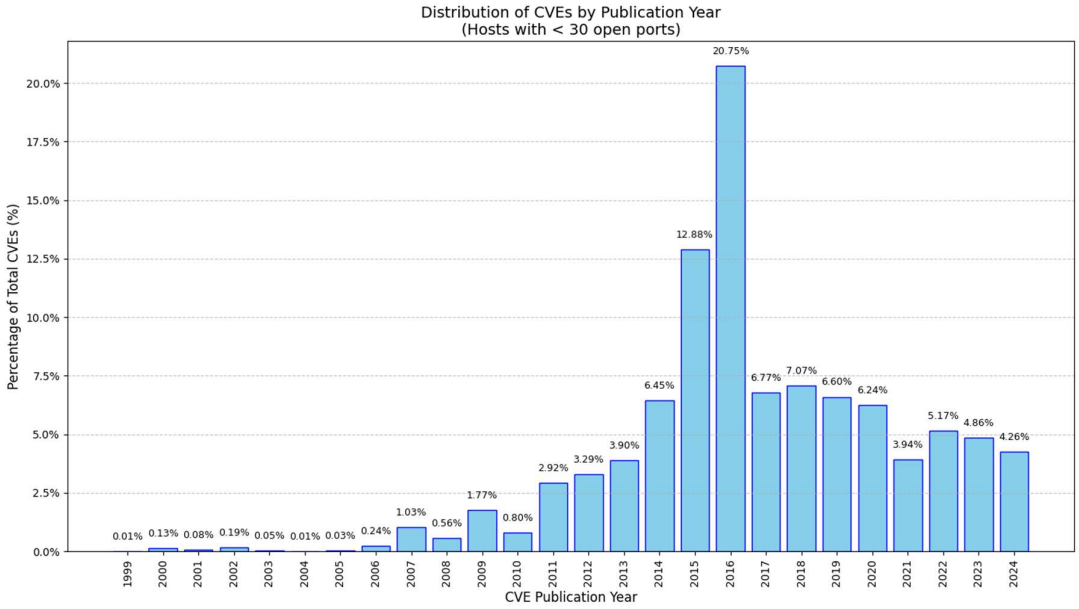


Fig. 4.3.5. Distribució dels CVE quant a any de publicació filtrat per *honeypots*. Font: Elaboració pròpia.

Com es pot observar, hi ha una gran quantitat de servidors que presenten vulnerabilitats antigues. Destaca el percentatge de CVE descoberts entre 2015 i 2016, possiblement relacionat amb la retirada del suport tècnic sobre Windows XP a l'abril de 2014 [24], un sistema operatiu àmpliament utilitzat en servidors mèdics. La tendència en els dispositius mèdics és la de tenir cicles de vida extremadament llargs; a diferència d'equips de consum, aquest tipus de dispositius poden tenir una vida operativa de 15-20 anys de durada. Si no s'actualitzen regularment, aquests equips queden obsolets i cada vegada més vulnerables a amenaces.

Quant a la perillositat de les amenaces, la Fig. 4.3.6 mostra la puntuació CVSS mitjana. S'han representat només les puntuacions entre el 5 i el 10 per simplificar el gràfic. És important destacar que els percentatges sumen més del 100% per país ja que un mateix *host* pot presentar diversos CVE amb diversos CVSS. La mitjana ponderada està calculada per a tots els països i sí que està calculada sobre un 100%.

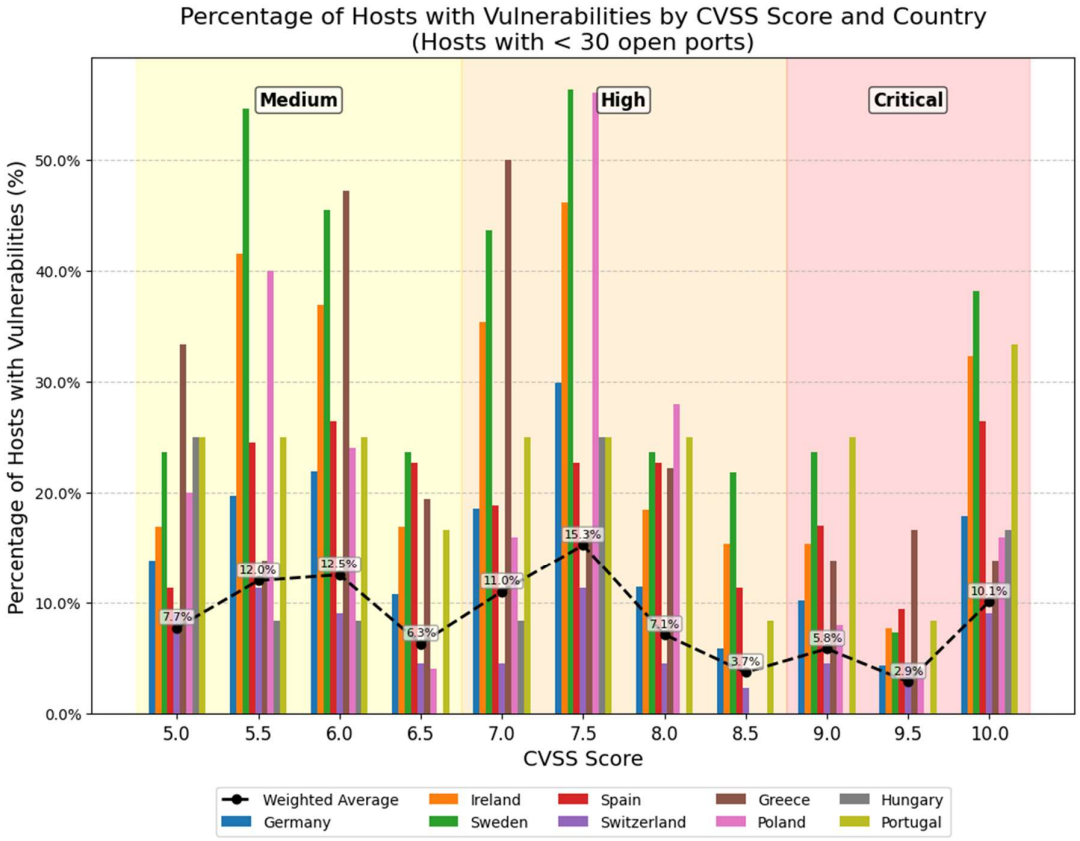


Fig. 4.3.6. Percentatge de *hosts* amb x nivell de puntuació per país filtrat per *honeypots*. Font: Elaboració pròpia.

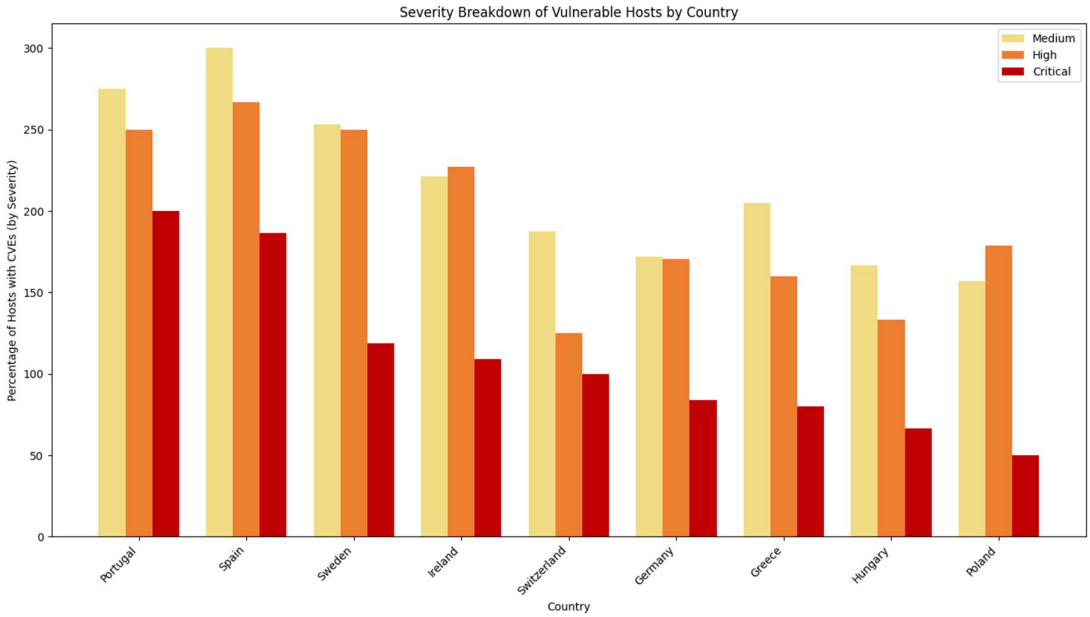


Fig. 4.3.7. Percentatge dels *hosts* amb vulnerabilitat segons el rang de gravetat per país. Font: Elaboració pròpia.

El gràfic mostra la distribució de la puntuació de perillositat dels CVE per a cada país. No es mostra una tendència clara però si que s'hi pot veure una major representació per als valors entre el 5 i el 7,5. Tot i això, hi ha un pic important al voltant del 10, la puntuació màxima. Aquest valor és alarmant ja que les vulnerabilitats amb un CVSS de puntuació al voltant de 10 solen ser senzilles d'executar i presenten una amenaça major.

5. Conclusions i treball futur

En aquest apartat es detallen les conclusions extretes de la realització del treball i s'exposen les línies de treball futur.

5.1 Conclusions sobre els serveis que proporcionen els *hosts*

Els servidors exposats presenten un patró molt clar pel que fa als serveis que ofereixen. Hi ha un seguit de solucions tecnològiques en format software que, un cop instal·lades al servidor, faciliten el desplegament d'un PACS. *Orthanc* [25], *Mirth Connect* [26] o Philips PACS [27] entre d'altres, són alguns dels softwares utilitzats pels centres de salut que busquen desplegar un servidor que gestioni les imatges produïdes a les cabines de ressonància o de raigs X. Els propis softwares obren els ports necessaris per al funcionament i la comunicació entre equips però no s'hauria de quedar aquí. Aquest tipus de servidors haurien d'estar protegits rere un *firewall* corporatiu i no exposats directament a Internet. L'accés extern hauria de realitzar-se mitjançant VPN segura o altres mecanismes d'accés remot controlat, amb autenticació robusta i xifratge *end-to-end*.

El mes comú és que la informació que emmagatzema el servidor (o fins i tot el control d'aquest) estiguin protegits rere un usuari i contrasenya en una *landing page* sobre HTTP i estiguin directament exposats a Internet. Aquesta pràctica és considerada una negligència per part de l'instal·lador ja que no és protecció suficient per aquests dispositius d'alt risc.

5.2 Conclusions sobre la informació que emmagatzemen els servidors

Tot i que el patró més comú sigui el d'emmagatzemar informació rere un usuari i contrasenya, s'han trobat servidors on la informació està exposada directament. Com s'ha comentat anteriorment en aquest treball s'ha trobat un servidor a l'Índia que exposa directament les dades a Internet, però aquest no ha estat l'únic. S'han trobat servidors a Alemanya on després de la instal·lació d'un dels softwares anomenats anteriorment, no s'ha procedit a canviar ninguna configuració addicional. Això provoca que la informació sigui pública sense necessitat de *log-in* com mostra la Fig. 5.2.1 i Fig. 5.2.2.

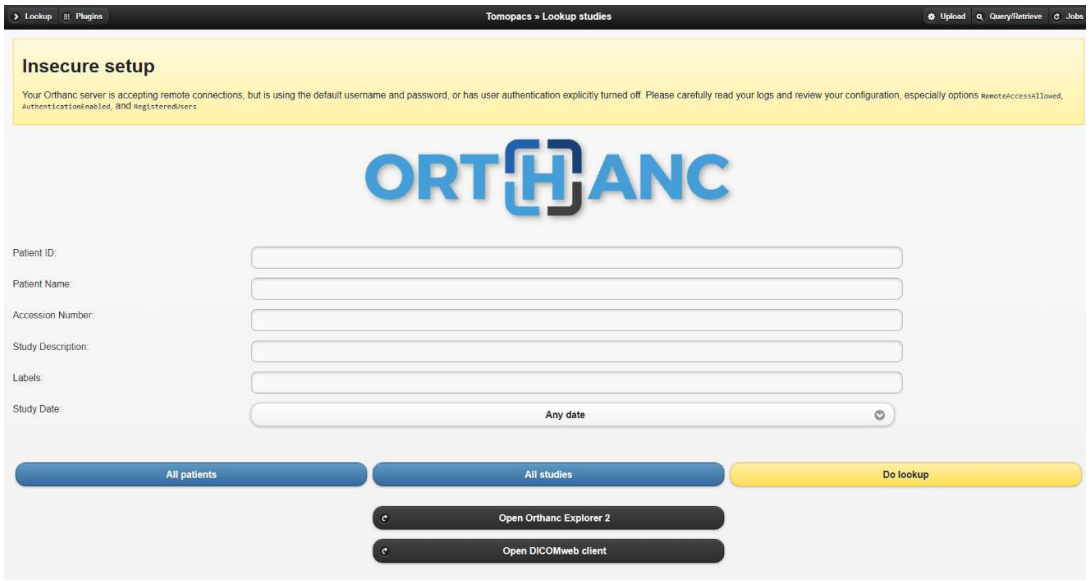


Fig. 5.2.1. Desplegament del software *Orthanc* amb un avís de configuració insegura. Font: Elaboració pròpia.

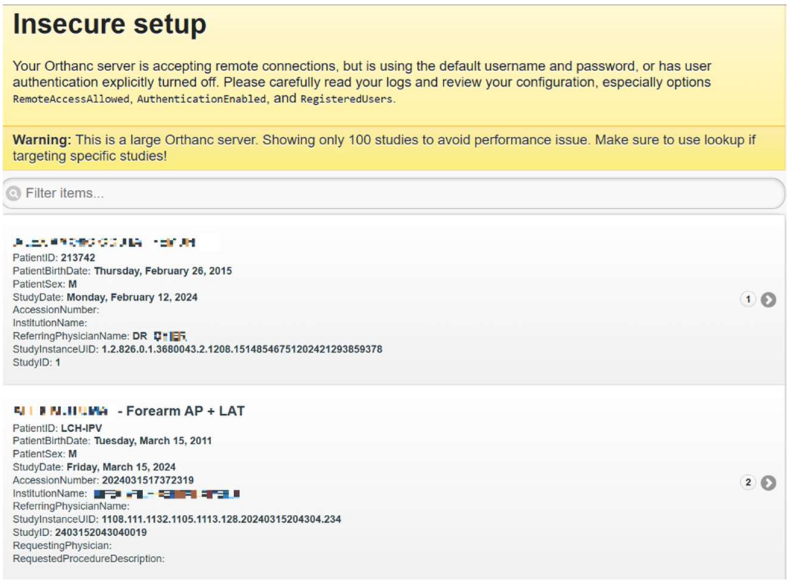


Fig. 5.2.2. Informació que es mostra després de prémer el botó "*All studies*". Font: Elaboració pròpia.

Cal destacar que no es creu que aquest servidor estigui en ús actualment però si que es pretengui utilitzar en un futur, ja que la resolució del DNS mostra un subdomini dins el domini del centre mèdic anomenat "*demo*". A més, els noms dels suposats pacients no són alemanys, tot i que aquest servidor (i el centre mèdic), estiguin físicament a Alemanya.

Encara que sigui un servidor encara no desplegat per al seu funcionament, s'ha fet un seguiment d'aquest i amb les actualitzacions que ha anat rebent des de Gener de 2025 fins a Maig de 2025 encara no s'ha solucionat el problema de visibilitat a Internet.

5.3 Conclusions sobre les vulnerabilitats dels servidors

S'han trobat servidors, com l'últim mencionat, que han estat en funcionament des de fa menys d'un any des de la publicació de l'estudi, però d'altres porten anys en funcionament i les actualitzacions que han rebut no han solucionat els problemes de seguretat. Com mostrava la Fig. 4.3.5 hi ha un gran nombre de *hosts* que presenten vulnerabilitats trobades fa més de 10 anys des de la publicació d'aquest estudi, posant de manifest que no reben actualitzacions completes de software ni auditories de seguretat.

És possible que molts centres mèdics no comptin amb un equip de seguretat informàtica degut al poc pressupost que solen tenir. Hospitals mitjans o grans poden permetre's tenir equips especialitzats en ciberseguretat, però altres centres més petits, com centres de podologia o dentistes de pobles més aïllats, només s'han pogut permetre un únic pagament a l'equip de desplegament de xarxa i no tenen un manteniment recurrent. Es proposa una regulació sobre el desplegament de servidors de tipus mèdic, per a que qualsevol dispositiu d'aquestes característiques hagi de tenir una configuració de seguretat bàsica abans d'entrar en funcionament. A més, es proposa impulsar campanyes de sensibilització sobre els riscos cibernètics dirigides al personal sanitari, subvencionades amb recursos públics dins de les polítiques d'ajudes a la digitalització sanitària.

5.4 Conclusions finals

En aquest estudi s'ha investigat l'estat de la ciberseguretat en entorns mèdics utilitzant una *pipeline* que obté informació d'eines OSINT, l'emmagatzema a una base de dades i es

consulta per fer-ne una anàlisi. Després dels gràfics obtinguts es pot concloure que la seguretat informàtica als servidors mèdics no és suficient. Es recomana als hospitals i centres de salut: 1) una inversió més alta en seguretat informàtica. 2) Millorar els esforços per a la protecció de dades dels usuaris. 3) Actualitzar periòdicament els softwares dels servidors. 4) Implementar serveis VPN per a la connexió remota a aquests servidors. 5) Separar serveis d'informació de caràcter general de serveis d'informació mèdics. 6) Monitoritzar la xarxa informàtica, fins i tot utilitzant eines OSINT, a la recerca de possibles vulnerabilitats. 7) Tot i implementar softwares ja desenvolupats, configurar correctament aquests. 8) Amagar tota la infraestructura de xarxa que no hauria d'estar accedida pels usuaris rere un *firewall* amb NAT.

Si no es segueixen aquestes recomanacions, els centres mèdics seran objectius fàcils per a atacs informàtics que poden resultar en robatori massiu de dades mèdiques sensibles, atacs de *ransomware* que paralitzin completament l'activitat assistencial i multes milionàries per incompliment del RGPD. Això comportaria la pèrdua de confiança dels pacients en el sistema sanitari digital i un retrocés significatiu en la digitalització mèdica. Els costos de recuperació, les demandes judicials i l'impacte a la reputació podrien arribar a comprometre la viabilitat econòmica dels centres afectats.

5.5 Possibles ampliacions

Com a possibles ampliacions relacionades amb aquest estudi es proposen dues branques principals:

1) Continuar amb el desenvolupament de les eines per al contacte automatitzat amb els afectats. Amb la informació que proporcionen les eines OSINT, es pot trobar informació sobre el centre mèdic encarregat del servidor. Utilitzant eines d'*scrapping* web o altres eines OSINT amb bases de dades de noms i informació de contacte d'empreses, es pot automatitzar el procés de la creació d'informes amb les dades de contacte dels centres mèdics i les vulnerabilitats que presenten els seus dispositius. Així es pot fer un seguiment de la resposta obtinguda i qualificar la velocitat de resposta front una vulnerabilitat.

2) Augmentar el rang de països sobre els que s'ha fet l'estudi i comparar l'estat de la seguretat entre continents. És una ampliació interessant ja que permet comparar un rang més

ampli de situacions geopolítiques com l'estat de seguretat en països tercermundistes o de països en situació de guerra.

6. Bibliografia

[1] M. S. a. O. Pianykh, «How Secure Is Your Radiology Department? Mapping Digital Radiology Adoption and Security Worldwide,» American Journal of Roentgenology, 02 Març 2016. [En línia]. Available: https://www.ajronline.org/doi/10.2214/AJR.15.15283. [Últim accés: 26 Març 2025].

[2] G. N. GmbH, «Information Security Report - Unprotected patient data in the Internet,» Greenbone, 17 Novembre 2019. [En línia]. Available: https://www.greenbone.net/wp-content/uploads/Greenbone_Security_Report_Unprotected_Patient_Data_a_Review.pdf. [Últim accés: 26 Març 2025].

[3] L. O. Sánchez, «La historia de la ciberseguridad,» NordVPN, 25 Desembre 2022. [En línia]. Available: https://nordvpn.com/es/blog/historia-ciberseguridad/. [Últim accés: 30 Març 2025].

[4] A. Moiseev, «Cómo engañar a los pilotos automáticos de Tesla y Mobileye,» Kaspersky Daily, 27 Maig 2021. [En línia]. Available: https://www.kaspersky.es/blog/rsa2021-tesla-mobileye-perception-gap/25378/?srsltid=AfmBOoqaKxKwDUq8aix0hqjiUszDwTzXMC51CZqXJqcKY4pcw6ZVUBe2. [Últim accés: 30 Març 2025].

[5] S. Research, «Dark Web Profile: Hive Ransomware Group,» SOC Radar, 26 Gener 2023. [En línia]. Available: https://socradar.io/dark-web-profile-hive-ransomware-group/. [Últim accés: 30 Març 2025].

[6] INCIBE, «LockBit: acciones de respuesta y recuperación,» Instituto Nacional de Ciberseguridad, 14 Març 2024. [En línia]. Available: https://www.incibe.es/incibe-cert/blog/lockbit-acciones-de-respuesta-y-recuperacion. [Últim accés: 30 Març 2025].

[7] SOPHOS, «El estado del Ransomware 2024,» SOPHOS, Abril 2024. [En línia]. Available: https://assets.sophos.com/X24WTUEQ/at/pzm7pw4k5ghvxmfbtcx57mr/sophos-state-of-ransomware-2024-wpes.pdf. [Últim accés: 30 Març 2025].

[8] C. P. R. Team, «Check Point Research Reports a 38% Increase in 2022 Global Cyberattacks,» Check Point, 5 Gener 2023. [En línia]. Available: https://blog.checkpoint.com/2023/01/05/38-increase-in-2022-global-cyberattacks/. [Últim accés: 1 Abril 2025].

[9] ENISA, «ENISA THREAT LANDSCAPE: HEALTH SECTOR,» ENISA, Juliol 2023. [En línia]. Available: https://enisa.europa.eu/sites/default/files/publications/Health%20Threat%20Landscape.pdf. [Últim accés: 1 Abril 2025].

[10] Sam, «Top 10 Fastest Programming Languages In 2024,» Medium, 23 Gener 2024. [En línia]. Available: https://medium.com/@learnwithakshay/top-10-fastest-programming-languages-in-2024-a135df564592. [Últim accés: 02 Març 2025].

[11] OSINTUX, «The Harvester,» OSINTUX, [En línia]. Available: https://www.osintux.org/documentacion/the-harvester. [Últim accés: 1 Abril 2025].

[12] E. IT, «Censys, el motor de búsqueda para descubrir vulnerabilidades de seguridad en Internet,» Educación IT, [En línia]. Available: https://blog.educacionit.com/censys-el-motor-de-busqueda-para-descubrir-vulnerabilidades-de-seguridad-en-internet/. [Últim accés: 2025 Abril 4].

[13] Censys, «Censys Official Web Page,» [En línia]. Available: https://censys.com/. [Últim accés: 2025 Abril 4].

[14] Censys, «Web Interface,» [En línia]. Available: https://search.censys.io/. [Últim accés: 2025 Abril 4].

[15] Shodan, «Shodan Web Page,» [En línia]. Available: https://www.shodan.io/. [Últim accés: 2025 Abril 4].

[16] C. Carismàtica, «WEB OFICIAL DE LA CÀTEDRA CARISMÀTICA DE CIBERSEGURETAT,» [En línia]. Available: https://carismatica.upc.edu/ca/. [Últim accés: 22 Maig 2025].

[17] D. Autors, «Rest Countries,» Rest Countries, [En línia]. Available: https://restcountries.com/. [Últim accés: 12 05 2025].

[18] I. M. Fund, «World Economic Outlook Database,» Octubre 2024. [En línia]. Available: https://www.imf.org/en/Publications/WEO/weo-database/2024/October/weo-report. [Últim accés: 12 Maig 2025].

[19] D. Neufeld, «Ranked: Countries investing the most in R&D,» Visual capitalist, 17 Abril 2025. [En línia]. Available: https://www.visualcapitalist.com/rd-investment-by-country/. [Últim accés: 12 Maig 2025].

[20] I. T. Union, «Global Cybersecurity Index,» ITU Publications, 2024.

[21] I. T. Administration, «Greece - Information and Communication Technology,» International Trade Administration, 28 Desembre 2023. [En línia]. Available: https://www.trade.gov/country-commercial-guides/greece-information-and-communications-technology. [Últim accés: 13 Maig 2025].

[22] C. &. I. S. Agency, «HTTP/2 Rapid Reset Vulnerability,» 13 Octubre 2023. [En línia]. Available: https://www.cisa.gov/news-events/alerts/2023/10/10/http2-rapid-reset-vulnerability-cve-2023-44487. [Últim accés: 15 Maig 2025].

[23] INCIBE, «CVE-2017-8923,» 12 Maig 2017. [En línia]. Available: https://www.incibe.es/en/incibe-cert/early-warning/vulnerabilities/cve-2017-8923. [Últim accés: 15 Maig 2025].

[24] Microsoft, «El soporte técnico para Windows XP, Office 2003 y Exchange 2003 ha finalizado,» Microsoft, 03 Setembre 2023. [En línia]. Available: https://learn.microsoft.com/es-es/lifecycle/announcements/windows-xp-office-exchange-2003-end-of-support. [Últim accés: 15 Maig 2025].

[25] «Orthanc WebSite,» [En línia]. Available: https://www.orthanc-server.com/. [Últim accés: 17 Maig 2025].

[26] «Mirth Connect,» [En línia]. Available: https://www.nextgen.com/solutions/interoperability/mirth-integration-engine. [Últim accés: 17 Maig 2025].

[27] «Philips PACS,» [En línia]. Available: https://www.philips.es/healthcare/solutions/diagnostic-informatics. [Últim accés: 17 Maig 2025].