Space for hospital logo

CYBERMANUAL

SIMPLE DIGITAL SECURITY RULES IN A HEALTHCARE ENTITY

A practical guide for medical and administrative staff, and guidelines for the Management Board.

POLISH ASSOCIATION
OF PRIVATE HOSPITALS

UEHP 35 YEARS
PRIVATE HOSPITALS
SHAPING EUROPE'S HEALTH

TABLE OF CONTENTS

Dear Readers,

Why is cybersecurity everyone's business?

In the era of hospital digitalization, the computer has become just as much of an everyday tool as a stethoscope, a blood pressure monitor, or a syringe. Most of the information regarding our patients' health – test results, medical histories, drug dosages – is stored within IT systems.

We must understand one crucial thing: cybersecurity in a hospital is not just about data protection; first and foremost, it is about patient safety.

When computer systems are locked by hackers:

- A doctor cannot check if a patient is allergic to a given medication.
- Scheduled surgeries must be cancelled.
- The Emergency Room (ER) cannot efficiently admit patients because the registration system is down.
- Laboratory test results do not reach the ward on time.

Each of us – regardless of whether we are a doctor, nurse, orderly, or office worker – is a guardian of this data. This guide will help you understand how simple, everyday habits can protect our facility from paralysis.

Andrzej Sokołowski
President of the Polish Association of Private Hospitals

For staff

CHAPTER 1:
The traps, or how fraudsters try to catch us

Attackers rarely break through the walls of servers. More often, they try to deceive a person, counting on haste, fatigue, or routine.

1.1. Impersonation (phishing)

This is the most common method. A fraudster sends an email or text message that looks like a message from a bank, from the Ministry of Health, from the national health insurance fund, or even from hospital management or the IT department.

What it looks like: you receive a message saying "Your password is about to expire, click here to update it" or "Please find attached an urgent invoice for payment."

What to watch out for:

- **Time pressure:** the message suggests you must act "immediately," otherwise something bad will happen.
- **Errors in names or in institution names:** the sender's email address often differs by a single letter from the genuine one.
- **Suspicious links:** never click on links in messages you were not expecting.

1.2. Digital ransom (ransomware)

This is malware that padlocks all the files on a computer. A message appears on the screen stating that unless the hospital pays a ransom (often running to millions), the data will be destroyed. Such malware most often enters the system when someone opens an infected email attachment.

CHAPTER 2:
Golden rules of safe work at the computer

Security starts with simple actions that require no technical knowledge, just attentiveness.

2.1. Your password, your personal key

A password is like the key to the controlled drugs cabinet: you do not leave it in the lock and you do not give it to anyone.

- **Do not write passwords down on paper:** sticking passwords to your monitor or keeping them under the keyboard is an invitation to a thief.
- **Do not share your account:** if a colleague asks, "Just log me in on your account for a moment, I've forgotten mine," politely refuse. Anything done under your account goes against your name, including any clinical errors.
- **Create a password that is easy to remember and hard to guess:** instead of "password123," use, for example, the first letters of each word in a favorite song or poem, with a number added.

2.2. The "leaving the desk equals locking the screen" rule

This is the single most important habit in a hospital. Whenever you are suddenly called to a patient or step away to grab a coffee:

- **Always lock the screen.** You do not need to log out. Just press Windows key + L at the same time.
- **Why does this matter?** A computer left unattended allows anyone passing by (for example, a visitor) to view private data or to make changes in the system under your name.

2.3. Clean desk and secure printing

Data protection is not just about the computer, it is also about paper.

- **Do not leave printouts behind:** if you are printing test results or patient data, collect them from the printer immediately. Documents left on a tray in the corridor are accessible to anyone.
- **The shredder is your friend:** documents containing personal data that are no longer needed must go into the shredder, not into a regular waste bin.

CHAPTER 3:
Personal devices and unknown media

The internet and modern gadgets bring risks that are easy to overlook.

3.1. The "found USB stick" trap

Never connect to a work computer any USB memory stick that you found in a corridor or that you received as a giveaway at a medical conference.

- **This is a common attacker tactic:** they plant an infected device, counting on someone's curiosity to plug it in. Connecting such a USB stick can infect the entire hospital network within seconds.

3.2. Work is work, private is private

- **Do not check your personal email on a work computer:** your private email account may be less protected than the hospital system. Opening attachments there exposes the hospital network.
- **Do not install your own software:** if you need a clinical calculator or a dictionary, ask the IT department. Software from untrusted sources may spy on what you type on the keyboard.

CHAPTER 4:
What to do when something feels wrong

Most people are afraid to admit a mistake online. In our hospital we want to change that: reporting an incident is a sign of responsibility, not something to be embarrassed about.

4.1. Symptoms that should put you on alert

- The computer has suddenly become very slow or freezes.
- Strange icons or pop-up windows appear on the desktop that were not there before.
- Someone in your contacts has received an email from you that you did not send.
- A website you visit regularly looks different from usual.

4.2. Step by step procedure

1. **Stay calm.**
2. **Stop working on the device.** Do not try to repair the system yourself or restart the computer unless told to do so by IT.
3. **Disconnect from the network** if you can (for example, unplug the cable from the back of the chassis).
4. **Call the IT department** or notify your line manager without delay. Say exactly what happened (e.g. "I clicked on a link in a suspicious email").
5. **Inform colleagues** so they do not open similar messages.

SUMMARY:
Your personal checklist

**Please review these points before each shift.
It only takes a moment.**

Password:
is my password known only to me? Is it not written down where it can be seen?

Screen:
do I lock the computer every time I leave the desk (Windows + L)?

Email:
do I check who the sender is before opening an attachment?

Desk:
are there no loose sheets with patient data lying at my workstation?

Screen:
do I know the IT department's phone number to call in case of trouble?

Your clinical experience and medical knowledge are the foundation of our hospital.
Adding a touch of vigilance in the digital world creates the safest possible environment for our patients.

Thank you for your daily attentiveness!

In the unlikely event of a cyberattack, the next page sets out a **Cyber First Aid Sheet,** prepared specifically for clinical and non-clinical staff. **It is a ready to use document to be printed and posted in duty rooms and at nursing stations.**

CYBER FIRST AID SHEET

Instructions for staff on duty (physicians, nurses, administration)

Attackers often exploit night shifts and time pressure in the Emergency Department. Your quick reaction can save hospital systems from paralysis.

1. Warning signs, react when:

- A message appears on the screen about files being locked, or a ransom demand.
- The computer has suddenly become very slow or clicks by itself.
- You received a suspicious email or text and clicked a link or opened an attachment.
- You are receiving phone prompts to approve a login (MFA push) that you did not initiate.

2. Step by step algorith

1. STOP WORKING: do not click anything else, and do not force close any windows.
2. DISCONNECT FROM THE NETWORK: if you can, unplug the network cable or turn off Wi Fi. Do not unplug the computer from the power supply, IT staff need to preserve evidence held in RAM.
3. CALL IT or SOC at the emergency number:
4. INFORM THE TEAM: warn colleagues on the ward not to open similar messages.

REMEMBER:

reporting an incident just in case is better than silence. An early reaction means less risk for your patients.

3. What you absolutely must NOT do

- Do not fix it yourself: do not install antivirus software downloaded from the internet.
- Do not destroy evidence: do not delete suspicious emails or tidy up the desktop.
- Do not give in to pressure: if someone calls and asks for your password or MFA code because there is an outage, hang up. It is a scam.

4. Quick report to the IT specialist

(Fill in and hand to the IT contact)

- Date and time of the event:
- Computer number or name:
- What happened? (e.g. clicked a link, strange window):
- Did you enter a password or a code from your phone? YES / NO

IT emergency number (24/7):

Person responsible on the ward:

For the management board

CHAPTER 5:
Management Board obligations under the NIS2 Directive

Modern cybersecurity has ceased to be the exclusive domain of IT service providers, or a passive system for protecting servers. With the entry into force of the NIS2 Directive, the scope of entities required to manage cybersecurity systemically has been extended to the healthcare sector, and responsibility for this area has been placed directly on senior management.

5.1. The scale of the threat: data and statistics

The healthcare sector is now one of the prime targets of cybercriminals at EU level, and the picture published by the EU Agency for Cybersecurity (ENISA) leaves no room for complacency.

According to ENISA, health has been the most affected sector for four consecutive years (2020 to 2023) in the significant incidents reported by Member States through ENISA's Cybersecurity Incident Reporting and Analysis System (CIRAS). Within the sector, healthcare providers, and hospitals in particular, take the brunt of the attacks: in ENISA's Threat Landscape for the Health Sector, 53 percent of incidents affected healthcare providers and 42 percent affected hospitals.

The threat mix is dominated by two patterns familiar to every Board:

- **Ransomware.** Across the reporting period covered by ENISA's first health threat landscape, ransomware accounted for 54 percent of cybersecurity threats in the health sector, and it remained one of the prime threats in the 2024 update, responsible for 45 percent of health related incidents.
- **Data breaches.** In the same 2024 update, 28 percent of health related incidents were data breaches, reflecting both the value of patient data on criminal markets and the maturity of breach notification regimes that bring such events to light.

National picture: Poland as an example

National data confirm the EU level trend. In Poland, figures published by Centrum e-Zdrowia (the national e-health centre) show a sharp rise in incidents affecting healthcare entities:

- In 2023, 405 attacks were recorded across Polish healthcare facilities.
- In 2024, that number more than doubled, reaching 1,028 incidents.
- In 2025, the number of attacks rose by a further 40 percent year on year.

Although these figures relate to a single Member State, the dynamic mirrors the EU level pattern observed by ENISA: healthcare providers are increasingly targeted, and the share of high impact incidents (in particular ransomware) is growing year on year.

5.2. The strategic role of the Board

Under NIS2 requirements, the management bodies of healthcare entities play a key role in the protection system. Their direct obligations include:

- **Approving risk management measures:** management must take an active part in selecting and approving the protection strategy.
- **Oversight of implementation:** the Board is responsible for monitoring whether the adopted measures are effectively executed in the hospital's day to day work.
- **Personal liability:** the new rules provide that members of management bodies may incur personal liability for breaches of cybersecurity duties.

5.3. Synergy between the DPO and the IT team

Experience shows that a cyberattack often exposes legal and technical gaps simultaneously. Effective security management requires close cooperation between the Data Protection Officer (DPO) and the IT team. To deliver on NIS2's requirements, management should:

- **Designate a shared area of responsibility:** for example, by an internal directive setting out where the competences of the DPO and the cybersecurity specialists meet.
- **Establish a reporting duty:** introducing regular meetings and a permanent reporting mechanism to the Board makes it possible to course correct the protection strategy on an ongoing basis.
- **Define a joint incident procedure:** this document must clearly set out the obligation for the two functions to inform each other and to respond jointly, in order to minimize the impact of any breach.

5.4. Planning, documentation, and the accountability principle

Implementing NIS2 standards is a long term process and requires the right resources. Cybersecurity must be reflected in:

- **Financial plans:** as a necessary investment in infrastructure and in staff competence.
- **Organizational plans:** by embedding security procedures into the standard processes for running the facility.

A key element is documenting the actions taken. Reliable records of management decisions and protective measures carry significant legal weight: in the event of an incident, they make it possible to demonstrate the due diligence of senior management.

CONCLUSIONS FOR THE BOARD:

NIS2 makes clear that cybersecurity in healthcare has become a strategic matter, not a purely technical one. Resilient IT systems, well prepared staff, and a high level of management awareness will, in the years ahead, be decisive for the uninterrupted operation of every healthcare entity.

CHAPTER 6:
The digital trail, handling, reporting, and cooperation with law enforcement after an attack

In an age of pervasive digitalization, every action in clinical IT systems leaves behind a binary fingerprint. Understanding the nature of this evidence and the procedures for preserving it is critical both for the legal protection of the healthcare entity and for the effectiveness of action by prosecutors, police cybercrime units, and incident response teams (CSIRTs).

6.1. The nature and location of the digital trail

A digital trail is a binary record, a sequence of zeros and ones, produced as a fragment of code by hardware and software, documenting every operation in the system.

The foundation for preserving evidence is the systematic collection of logs (event records) from IT systems.

Key logs that should be collected:

- Logs of all authentication operations and remote access events.
- traffic logs, both at the perimeter and within the LAN.
- Logs from security systems.
- Logs from key servers.
- Logs from clinical and business applications.
- Logs from the backup system.

All logs should be sent to a dedicated repository (for example, a log collector). It is also recommended to collect and monitor logs from the log collection system itself.

Key logs that should be collected:

- **Clinical systems:** activity records in the Electronic Health Record (EHR) and in imaging systems (X ray, CT, MRI, PACS).
- **Network infrastructure:** server logs, connection histories, and records held in cloud services.
- **Communications:** email server histories and user login records.

It is important to remember that digital traces can be volatile: data held in RAM may be lost irretrievably once the computer is powered off or restarted.

6.2. Legal obligations of the Board

In the event of an incident with criminal characteristics, management has reporting duties under national criminal procedure law. In most EU jurisdictions these obligations follow a similar pattern:

- **Citizen's duty:** anyone who becomes aware of a publicly prosecuted offence has a civic duty to notify the prosecutor's office or the police.
- **Institutional duty:** public institutions and equivalent entities are obliged to notify the relevant authorities of an offence without delay, and to take steps to prevent the destruction of evidence until law enforcement arrives.

In practice, this means a legal duty to report a suspected intrusion, data theft, or destruction of records, and to actively protect the digital scene of the incident.

6.3. Cooperation with the CSIRT and incident reporting

Every incident in healthcare should be reported without delay to the relevant incident response team (CSIRT). Under NIS2, reports go to the national CSIRT and, where one exists, to the sectoral CSIRT for health.

- Reporting channels (to be completed with country specific contact details):
- National CSIRT incident reporting portal.

Sectoral health CSIRT, web form or dedicated email address.

It is important to file the report as soon as possible, even if you do not yet have all the information. Additional details can be added later, and a fast reaction limits the scale of damage.

6.4. Rules of digital hygiene after an attack is detected

For a digital trace to serve as evidence, it must not be altered by inappropriate action by staff. Where possible, the following is recommended:

- **No DIY repairs:** do not run antivirus tools or cleaners on a compromised device, they may overwrite logs or irreversibly damage useful files.
- **Isolate, do not erase:** disconnect the device from the network (cable or Wi Fi), but leave it powered on to preserve the contents of RAM.
- **No file modification:** do not open, copy, or move files affected by the attack. The correct technical action is to take a forensic (bit for bit) image.
- **Document the event:** write a note immediately, recording the time the anomaly was noticed and the wording of any messages, as details fade with time.

6.5. Incident reporting form (template for staff)

Date and time noticed	
Ward or organisational unit	
Workstation or PC number	
System or application in use	EHR / PACS / Email / Other:
Description of the event (what happened?)	
Was a link or attachment clicked?	YES / NO / DON'T KNOW
Was a password or MFA code disclosed?	YES / NO / DON'T KNOW
Visible messages (e.g. ransom note)	
Person reporting	

NOTE:

when this form is completed by the IT department, it should be supplemented with the technical Indicators of Compromise (IoCs): hostname, IP address, asset or inventory number, subnet, and event type. These are the data points that the CSIRT is most likely to ask for.

CONCLUSIONS FOR THE BOARD:

The digital trail is the silent witness of the offence. Reliable preservation of evidence is the best defence against any allegation of failing to fulfil duties under criminal procedure law and NIS2. For CSIRT teams, logs are the only source of precise information that allows the attack vector to be analysed and the set of compromised devices to be narrowed down. In the absence of logs, the only safe recommendation is often to reinstall every host on the network, which dramatically extends the period of paralysis and increases the cost of restoring continuity.

CHAPTER 7: Cyber insurance, a financial and operational shield for the Board

As people responsible for running a healthcare entity, you know that risk cannot be eliminated to zero, it can only be managed. In a Modern Hospital 2026 strategy, cyber insurance is no longer just a policy in a drawer. It becomes an operational mechanism for crisis response, supporting the hospital through the most difficult first hours after an attack.

7.1. More than indemnity: operational readiness

A common mistake in thinking about cyber insurance is to treat it purely as a financial instrument. In reality, a modern policy provides access to incident response (IR) teams. At the moment of an attack, management does not have to scramble to find digital forensics specialists or specialist lawyers, the insurer provides them on a 24/7 basis.

7.2. The response mechanism, step by step

In line with security standards, the incident handling process under an insurance cover is divided into five key stages:

1. **Detection and reporting:** even if the scale of the attack is not yet clear, the critical step is to notify the help centre (24/7 hotline) without delay. An incident may be a ransomware lockout, but it may equally be a suspected leak of patient data or a failure of diagnostic systems.
2. **Immediate support (Incident Manager):** the insurer assigns a dedicated coordinator. The Board gains a partner who takes on the burden of crisis management.
3. **Expert support:** the hospital receives help in three areas:
   - IT and digital forensics: stopping the attack, securing evidence, and analysing how the intrusion occurred.
   - Legal support: experts assess notification obligations to the data protection authority and to sectoral regulators.
   - Crisis communications: PR specialists help prepare messaging for patients, the media, and staff, protecting the facility's reputation.
- **Containment and continuity restoration:** the goal is the fastest possible return to providing clinical care and limiting the financial impact of downtime.
- **Claims handling:** once the crisis is contained, a formal settlement of costs and any third party claims (for example, by patients whose data has been leaked) follows.

7.3. What risks does the cover address?

A common mistake in thinking about cyber insurance is to treat it

6.2. Legal obligations of the Board

For the facility's safety, it is essential that the cover address the most common threats in the healthcare sector:

- **Ransomware attacks:** costs of data recovery and negotiations.
- **Data breaches (GDPR):** costs of administrative penalties (where insurable under applicable law) and of legal defence.
- **Electronic fraud:** attempts to extort funds from the hospital.
- **Human error:** unintentional security breaches by staff (for example, leaving a computer unlocked).

CONCLUSIONS FOR THE BOARD:

Cyber insurance is an investment in business continuity. With attacks on hospitals increasing, having access to specialist experts within tens of minutes of detecting a problem is the only way to avoid prolonged paralysis of the facility and major reputational losses.

Most cybersecurity standards were written for large enterprises, then handed down to hospitals with the expectation that they would somehow fit. They rarely do. Clinical staff cannot pause a resuscitation to read a 200-page policy. A medium-sized hospital does not have a 24/7 security operations center on standby. And NIS2 does not arrive with a step-by-step implementation manual attached.

ShieldNet is built differently. It is Europe's first cybersecurity standard designed specifically for small and medium-sized organizations, with a dedicated sector annex for healthcare providers. It translates the language of regulators (NIS2, GDPR, EHDS, ENISA guidance) into something a hospital can actually execute: concrete controls, ready-to-use artefacts, and a maturity path that matches the resources you really have.

Five Action Zones, one logical journey

- **ShieldNet** organizes cybersecurity around five Action Zones that mirror how hospitals respond to risk in practice:
- **KNOW & PLAN** - understand your assets, your risks, and your obligations under NIS2.
- **SHIELD & BUILD** - put in place the protective controls that matter most for clinical continuity.
- **WATCH & PATCH** - keep systems current and detect anomalies before they become incidents.
- **REACT & RECOVER** - respond when something goes wrong, with playbooks your staff can follow at 3 a.m.
- **LEARN & LIFT** - turn every near-miss into measurable improvement and Board-level evidence of due diligence.

What the Healthcare Sector Annex adds

The sector annex extends the core standard with controls and templates calibrated to clinical realities, including:

- Protection of **electronic health records** (EHR) and **medical imaging** (PACS) environments.
- Operational continuity playbooks **for Emergency Departments**, operating theatres, and laboratories.
- Medical-device segmentation guidance aligned with **MDR** and EHDS.
- NIS2 accountability templates that map directly to **Board-level reporting** obligations.

What it delivers

For the Management Board, ShieldNet provides documented evidence of due diligence under NIS2, the kind of paper trail that protects management personally when an auditor or regulator asks the hard questions. For IT and security teams, it eliminates the blank-page problem: every control comes with a ready artefact (policy, checklist, register, procedure) that can be adapted in hours, not weeks. For clinical staff, it means rules expressed in plain language, because cybersecurity that nobody understands is cybersecurity that nobody follows.

Patient safety begins with system safety. ShieldNet makes both achievable.

„You've probably seen predictive text on smartphones or computers. You type a word, and the device suggests the most-likely next word. Large language models let chatbots do something similar.. but on a far biggest scale."
- a simple explanation of LLMs at the Computer History Museum in Mountain View, Silicon Valley, where it all began.

The Physician and AI: How to Safely Use LLMs in Daily Practice?

Fundamentally, understanding what an LLM[1] is is not difficult for a physician. You simply need to pause for a moment and disconnect from an infosphere overloaded with clickbait and narratives driven by various interests. Although today's transformer-based models are trained on billions and trillions of words, they fundamentally rely on a method of statistically guessing the next most probable word based on the known context (previously inputted words). Thanks to high efficiency and immense computing power, the statistical tool that is an LLM performs quite effectively in several areas of text processing.
These include:

- Text generation.
- Text editing (paraphrasing, summarizing, changing style, combining).
- Summarization.
- Translation from one language to another, or translation within the same language (phrasing things differently or more simply).
- Information retrieval, including from distributed sources.
- Literature reviews and gap analysis.
- The combined application of several of the above, e.g., converting a patient's spoken words into text (speech-to-text), structuring text (organizing the medical history from a subjective examination), or generating proposed recommendations for the physician's approval (such as dietary and rehabilitation guidelines).

For these reasons, modern medicine is turning to Large Language Models (LLMs), which are revolutionizing how we work with text and data. It must always be emphasized that these tools do not replace diagnosis, but rather act as advanced administrative or research assistants.

Understanding that an LLM is merely a method of the statistical "drawing" of words makes it easy to spot the first danger, known as an LLM hallucination: regardless of the circumstances, the model will always "draw" a combination of words and arrange it into a grammatically correct structure. Unfortunately, the combination provided by the model can be completely clinically incorrect. In most cases, the "draw" turns out right, and the model generates useful results. However, the internet provides a growing number of examples of clinically flawed recommendations from ChatGPT and other LLMs. To build caution and properly understand the phenomenon of LLM hallucinations, it is worth following the *Epic fAIls* plebiscite hosted by the Director of the Center for Trustworthy AI, Prof. Przemysław Biecek, PhD, DSc, Eng.. In the recent *Epic fAIls 2025*, an example was given of a "dietary recommendation" that suggested replacing table salt with another, toxic salt[2].

Therefore, the fundamental and first rule for safely using LLMs - as defined by the ethical standard IEEE/ISO/IEC 24748-7000-2022[3] - is the physician's **ACCOUNTABILITY**. You can safely use LLM tools, it is worthwhile, and it is simply happening - but you must apply the "trust, but verify" principle, evaluating the LLM's output every single

time, just as you would evaluate the work of an assistant or an intern (who, by the way, can also sometimes be prone to hallucinations). AI does not replace diagnosis; it can support administrative processes.

Secondly, LLM language models are most often available online. Hence, the cited standard clearly sets out the second recommendation, which is **PRIVACY**. We must never enter personal data onto the internet: first names, last names, national identification numbers (like Personal ID or Passport number), or unique case descriptions that allow a patient to be identified. We also must not do this in ChatGPT or Gemini via a web browser or mobile phone. When using an LLM within one of the 7 use cases mentioned above, you can safely describe symptoms using general terms: e.g., "patient, male, 45 years old". Even in such a case, however, you must remember that this data "feeds the model" - data entered into public models can be used for their further training. Consequently, the same LLM, when asked by someone else, might randomly generate an answer by combining the words and sentences we previously inputted. Due to the design of this mathematical machine, it can connect information in unintended ways, identifying "patient M, with advanced obesity disease, skin lesions, and discoloration around the neck" in a small town - much like gossip works, by guessing who it might be about.

Physicians have perfectly mastered the ability to deal with these types of risks. Simply put, LLMs need to be demystified and treated like an assistant who collaborates with everyone in town; sometimes, they might indiscreetly let something slip.... In the case of scientific work, one should also not input anything groundbreaking or constituting intellectual property into LLMs. It is, however, worth "hiring" an LLM to conduct a literature review and comparing the result with your own research.

The most important guidelines for the safe use of LLM models in a physician's daily practice are **ACCOUNTABILITY** and **PRIVACY**.

Below is a short "Before You Click Send" checklist:

- Have I removed personal data?
- Have I verified the result using another source?
- Does the model have medical certification (if applicable)?
- Would I have acted the same way (diagnosed, recommended, concluded) without the use of an LLM?

Finally, it's worth noting that today it is easy to deepen one's knowledge about the safe use of LLMs. When doing so, you should rely on trusted authorities within the clinical community, avoiding celebrities forecasting the end of the world or the advent of universal "LLM medicine". Both extremes are false. For those interested in advanced security guidelines regarding large language models, the OWASP project publication titled *"2025 Top 10 Risk & Mitigations for LLMs and Gen AI Apps"*[4] is available.

In daily practice, adhering to a physician's standard accountability (due diligence) and confidentiality (ensuring privacy and data protection) are the best security practices. It is enough to know that - as the curators of the Silicon Valley Computer History Museum exhibition write - LLMs are word-prediction tools based on advanced statistics. While highly useful, this tool, like any other, has its limitations. exhibition write - LLMs are word-prediction tools based on advanced statistics. While highly useful, this tool, like any other, has its limitations.

Tomasz Jaworski,
IEEE CertifAIEd Authorized Assessor in AI Ethics,
Healthcare Cybersecurity Director
at Palo Alto Networks Poland

1 Examples of contemporary LLMs include tools offered by global providers such as ChatGPT (OpenAI), Gemini (Google), and Claude (Anthropic). There are also strong national LLMs developed by engineers from the same Polish math school like most of OpenAI funders: Bielik (a project by SpeakLeash and Cyfronet AGH) and PLLUM (created by a scientific consortium commissioned by the state) .
2 https://tinyurl.com/33fufvar
3 IEEE/ISO/IEC 24748-7000-2022 "IEEE/ISO/IEC International Standard - Systems and software engineering - Life cycle management - Part 7000: Standard model process for addressing ethical concerns during system design"
4 https://genai.owasp.org/llm-top-10/

Glossary of Abbreviations and Terms

ORGANISATIONS, DIRECTIVES AND ROLES

CIRAS (Cybersecurity Incident Reporting and Analysis System): A system managed by ENISA for reporting and analysing significant cybersecurity incidents across Member States.
CSIRT (Computer Security Incident Response Team): A specialised team dedicated to responding to and managing computer security incidents (e.g. a national or sectoral health CSIRT).
Cybercrime Units: Specialised law enforcement departments or police units that collaborate in the investigation and prosecution of cyberattack perpetrators.
DORA (Digital Operational Resilience Act): An EU regulation concerning the digital operational resilience of the financial sector, often cited as a benchmark for protective service compliance.
DPO (Data Protection Officer): The individual within a healthcare facility responsible for overseeing data security and ensuring compliance with privacy regulations.
E-Health Centre / National Health Authority: A strategic body responsible for maintaining central IT systems and implementation of e-health projects, often managing a sectoral CSIRT for medical entities.
ENISA (European Union Agency for Cybersecurity): The EU agency that provides threat landscape reports, statistics, and cybersecurity guidelines for Member States.
GDPR (General Data Protection Regulation): The primary regulation governing personal data protection, breach notifications, and patient data management.
NIS2: An EU directive requiring the healthcare sector to implement systemic cybersecurity management and introducing personal liability for management boards.

TECHNOLOGIES AND MEDICAL SYSTEMS

Asset / Inventory Number: A unique identification number assigned to hardware (e.g. PCs, servers) to help IT staff locate equipment during an incident.
EHR (Electronic Health Record): A digital version of a patient's medical history and clinical documentation used for treatment.
HIS (Hospital Information System): A comprehensive IT system used for managing overall hospital operations.
LAN (Local Area Network): The internal computer network within the hospital that connects devices in a specific area or building.
LLM (Large Language Models): Artificial Intelligence tools (e.g. ChatGPT, Gemini) used for text processing based on statistical patterns.
Log Collector / Repository: A dedicated, secure system where event records (logs) from various sources are sent for storage and analysis.
MDR (Managed Detection and Response): An advanced, outsourced security service providing 24/7 active monitoring and threat response.
MFA (Multi-Factor Authentication): A security process requiring multiple verification methods to grant access (e.g. a password plus a mobile code).
MFA Push: A notification sent to a mobile device as part of MFA, requiring the user to manually approve a login attempt.
PACS (Picture Archiving and Communication System): Technology used for storing and transmitting digital medical images (e.g. X-rays, CTs, MRIs).
RAM (Random Access Memory): Volatile computer memory; data is lost if the device is turned off, which is why hardware should not be restarted after an attack without IT instruction.
Remote Access: The ability to connect to hospital systems from an external location, requiring strict monitoring and logging.
SOC (Security Operations Centre): A centralised unit or team that monitors an organisation's security posture and manages incident response.

THREATS AND INCIDENTS

Attack Vector: The specific path or method used by an attacker to gain access to the hospital's systems.
Digital Forensics: The scientific process of recovering and investigating data from digital devices to determine the course of an attack and preserve evidence.
Digital Trail / Binary Fingerprint: A sequence of binary records and code fragments generated by hardware and software that documents user activity.
Forensic Image: An exact, bit-for-bit copy of a data carrier (e.g. a hard drive) used for investigation without altering the original evidence.
Incident Response (IR): The organised process of reacting to a security breach to limit damage, remove the threat, and restore continuity.
IoC (Indicators of Compromise): Technical evidence of a security breach, such as specific IP addresses, hostnames, or file names.
LLM Hallucination: A phenomenon where an AI model generates grammatically correct but factually or clinically incorrect content.
Logs: Event registers; digital records that document every operation or activity performed within an ICT system.
Phishing: A fraudulent method of impersonating institutions in emails or messages to steal passwords or infect systems.
Ransomware: Malicious software that locks files or systems and demands a ransom payment for their release.