**Term list: Informatics Session 8, Cybersecurity, v. 1.0**

Audit trail

A listing or log of user logons and actions. Allows inappropriate or malicious activity to be detected, defined, and tracked.

Authentication

Confirmation of the identity of a user. Generally requires something the user knows (password), something the user has (token or phone), and/or something the user is (biometrics). A requirement for more than one of these for access is stronger than requirement for a single item and is called “multi-factor authentication.”

Authorization

Permission for a user to access specific data or carry out specific tasks. Often role or group-based, depending on the user’s job in the context of the session.

Biometrics

Physical characteristics that can be used to identify a user, such as fingerprint, facial appearance, retinal scan, voice print, or hand vein pattern.

Backup

Copying data and sometimes programs to media separate from the operational system. Backups allow reconstruction of a system to an undamaged state if problems occur. Rotating backups creates multiple backup copies on a cycle, such as daily backups on a rotating weekly (7 backups) or monthly (~30 backups) schedule. Multiple backups are usually stored in several locations (eg, off-site backups) so that a disaster in one location will not damage all backups. Hot backups receive data continuously so there is minimal data loss in case of problems and they are ready to go whenever needed.

Breach

An inappropriate release or use of protected health information.

Breach notification rule

Breach reporting regulations. A covered entity must report breaches to the individuals affected, to media if > 500 individuals, and to the HHS (within 60 d if > 500, annually for < 500). Requirements extend to reporting by business associates.

Brute force

An strategy to gain entry (crack) a system by trying many passwords, based on information about a user, dictionaries, and/or lists of stolen passwords.

Buffer overflow

A type of software vulnerability where it is possible to induce a program to write excessively long data erroneously to computer memory, exceeding the allotted memory space for data and overwriting some of the space allotted for program instuctions. This can allow insertion of malicious code into an area of memory that will allow it to be executed.

Business associate agreement

A contract that specifies the responsibilities of a business associate of a HIPAA covered entity in using PHI or a limited data set from the entity.

Certificate (digital)

A data file provided by a certificate authority that contains a public key and private key for encryption, and identifying information on the owner of the keys that allows the validity of the certificate and public key to be checked programmatically at the certificate authority.

Certificate authority

A business that creates, provides, and registers digital certificates that contain pairs of public/private encryption keys.

Confidential data

Sensitive data that is shared within and restricted to a defined scope of distribution.

Covered entity

A business that is subject to HIPAA regulations. Covered entities include healthcare providers, health plans, and health data clearinghouses.

Cyberattack

A malicious attack on a set of digital systems for the purpose of controlling the systems, disabling the systems, extorting the system owners, or stealing data. Cyberattacks may have both electronic and physical (interpersonal or “social”) elements.

Cross site scripting

A cyberattack in which a vulnerability in a Web site is used to attach a malicious script to a Web page so that it is surreptitiously downloaded and executed on the computers of site visitors. The script sends data from or otherwise opens the user’s computer to the malicious author.

Data at rest

Data that is stored as files or in databases on media such as disk drives. Data at rest that is not encrypted is more vulnerable to access than encrypted data if a computer is compromised or stolen.

Data in motion

Data that is being sent or received over the network. Data in motion is vulnerable to interception and can be read if it is not encrypted.

Data integrity

Accuracy, completeness, and consistency of data.

Data use agreement

A contract that specifies the responsibilities of analysts or researchers in using PHI or a limited data set from a covered entity.

De-identification

Removing identifying information from a healthcare data set so that it cannot be reasonably linked to the individuals from which it is derived. HIPAA allows data to be defined as deidentified if it can be shown to be statistically infeasible to re-identify it, or if all 18 of the set of identifiers defined by HIPAA have been removed.

Denial of service (DoS)

A cyberattack in which a system or network is flooded with data packets so that it is overloaded and cannot be accessed or used. A distributed denial of service (DDoS) attack uses many compromised (“zombie”) computers in multiple locations in a coordinated attack on one organization or system.

Digital signature

A formal digital signature uses public (asymmetric) key encryption. A document is encrypted with the signer’s private key. Successful decryption with the public key proves the document was signed by the signer (authentication and non-repudiation) and that is has not been changed after signing (integrity).

Encryption

Encoding data so that it is not readable but can be decrypted and rendered readable using the appropriate algorithm and encryption key. See data at rest and data in motion.

End to end encryption

Communication of encrypted data from a sender to a recipient without decryption and re-encryption during transit.

Fail-over

Automatic switching to a hot backup or a complete alternate system when a problem develops in a primary system.

Firewall

A hardware device or software program that serves as a barrier to outside network communication. Firewalls may filter external data to determine its origin or content, using rules to exclude suspicious communications, they may implement proxy servers to mediate communications and protect the inside of a network from external probing, and they may track communication sessions such that only responses to outgoing communications are allowed into the network and they are routed only to the computers that initiated the communication.

Good password practices

Eight or more characters, upper and lower case with symbols and numbers, phrases, avoid single words and common number sequences, avoid personal information, do not share or enter in front of observers, keep records in a safe place that is private, use a password manager.

Good system security practices

Workstation screen blanking and locking on inactivity, latency or temporary lockout on failed login, remove retired accounts and pw quickly, use data switches instead of network hubs, review access audit trails, use firewalls and partition networks, use encryption at rest and in motion.

General Data Protection Regulation

European Union regulations related to personal data security. The GDPR covers health data and non-health consumer data.

HIPAA

Health Insurance Portability and Accountability Act. Defines responsibilities for covered entities in managing PHI in the areas of privacy, security, and breach notification.

IRB

Institutional Review Board. Reviews and approves proposed human subjects research projects with a focus on patient safety and equity, medical benefit, scientific merit, and appropriate data use.

Key (encryption)

A large number used to encrypt data. In a single key (symmetric) encryption scheme, the key is used both the encrypt and decrypt the data. Symmetric encryption is less secure than public key (asymmetric) encryption because the one key must be protected but also shared for decryption. For this reason, these keys are sometimes called “secret keys.” Symmetric encryption has advantages, though, because it is more efficient so encryption/decryption is faster. Also, the keys can be larger so the encrypted file is harder to crack without the key. One-time use secret keys are commonly used to secure communications where the initial session setup and secret key sharing is done under asymmetric public key encryption and then ongoing encryption/ decryption uses the secret key. See public key.

Legacy software

Software that is no longer actively maintained or patched by the developer, but is necessary because other software or devices depend on it. Vulnerabilities discovered in legacy software can be problematic because there may not be a mechanism to fix them.

Limited data set

A healthcare data set with the 18 HIPAA identifiers removed, except for dates, exact ages, and addresses to the level of city and zip code.

Malware

Software designed for malicious purposes.

Minimum Necessary Standard

A provision of HIPAA that requires the use of the minimum amount of PHI for any given task.

Multi-factor authentication

 See authetication.

Network switch

A device that connects multiple computers on different wired branches to a network and directs network traffic addressed to a computer only to the branch containing that computer. Switches are more secure than passive hubs because hubs allow all traffic into all branches and therefore one listening device on a branch can copy all communications to all computers on the hub.

Non-repudiation

With digital signatures, a signature that prevents the signer from denying signing the document or arguing that the document has been changed since the signature. Documents encrypted with the author’s private key that can be decrypted with the public key cannot be repudiated.

Packet filtering

The ability to examine a network communication based on its target address, origin address, and/or data content. Firewalls typically implement rules based on packet filtering to determine whether to pass packets they receive into the networks they protect.

Patch (software)

A software correction for a discovered error in a program, i.e., a software update. The error may create a security vulnerability that the patch eliminates.

Penetration testing

A staged cyberattack on a set of systems to assess its vulnerability to real cyberattacks. Penetration testing may include both electronic attacks and social engineering in an attempt to test the target’s resistance to compromise. Penetration testing is usually carried out by companies with substantial system security expertise.

Permissions

The data access and software functions available to a particular user role.

PHI

Protected Health Information. Any individually-identifiable health information. Generally consists if an identifier paired with health information, except for gene sequences which can be both identifiers and health information.

Phishing

Sending fraudulent emails in an attempt to trick users into providing login data or installing malware. Spear phishing is composing highly specific fraudulent emails using personal data to target a user.

Private data

Data that is defined as sensitive and kept secret. This may refer to personal data that individuals do not wish to share.

Privacy rule

The HIPAA privacy rule establishes standards for protecting health information. It defines HIPAA covered entities, protected health information (PHI), limited data sets, and deidentified data, and establishes appropriate and inappropriate uses for health data and required controls, actions, and penalties related to those uses.

Private key

One of two paired keys in a public key (asymmetric) encryption scheme. The private key is held undisclosed by the user while the public key is shared. Data encrypted with the private key can be decrypted using the public key, and vice versa.

Proxy server

A server located outside of a network through which Internet communications are routed. A proxy server often acts in concert with a firewall. The network addresses of all outgoing communications are changed to the proxy server address. The proxy server keeps records of communications so that return messages can be correctly routed if they correspond to sent messages and otherwise pass security rules. Proxy servers prevent systems on their networks from responding directly to unsolicited messages, and thus prevent external systems from probing networks to determine their structure, types of devices, and potential vulnerabilities.

Public key

One of two paired keys in a public key (asymmetric) encryption scheme. The public key is shared while the private key is held undisclosed by the user. Data encrypted with the public key can be decrypted using the private key, and vice versa. See private key.

Ransomware

Malware that encrypts files, databases, and application programs so that they are inaccessible to users until a ransom is paid. Ransomware may spread within a network for a time before activating and take down multiple servers, crippling a business and creating a situation from which recovery by rebuilding all systems at once is impractical.

Roles

Groups of users, usually job-related, that have equivalent permissions.

Router

A network device that controls the flow of data packets based on their addressing. Routers can be used to segment or partition networks into separate regions, and pass packets into those regions only if they are addressed to devices located there. Correctly-configured routers can prevent compromised systems in one part of the network from probing for vulerabilities in other parts and therefore prevent or slow the spread of malware across a network.

Safe harbor

A procedure defined in HIPAA for deidentifying data by removing all of 18 specified identifying data elements. If the safe harbor procedure is used, release of the data is not a data breach. For this reason the procedure is widely used, but it has been shown to be possible to reidentify safe harbor data depending on what other data is available. For this reason the safe harbor procedure may change in the future.

Security (of data)

The ability to maintain the integrity, confidentiality, and availability of data.

Security rule

The HIPAA security rule requires covered entities to protect the confidentiality and integrity electronic health data using appropriate administrative, physical, and technical safeguards.

Social engineering

Non-electronic (i.e., human) means for compromising data systems, for example, inducing users through interpersonal interaction to reveal their passwords or other important credentialing information.

SQL injection

An attack on a database front end such as a Web site that enters structured query language (SQL) code into form fields (such as a login form) in an attempt to extract data from the database.

Switch

 See network switch.

Synthetic data

Concocted data that is shown to be statistically not different from real data and therefore can be used in place of real data for certain applications.

TLS

Transport Layer Security. The communication standard covering encryption of Internet communications. Invoked for Web communications using “https” to start URLs. Evolved from the Secure Sockets Layer (SSL) protocol in the late ‘90s.

Trojan horse

A program that appears innocuous but secretly contains malware. Obtains permission to run from the user under false pretenses, then uses that permission to perform malicious functions.

Two factor authentication

Authentication using something known by the user (eg, a password) plus a second item the user is or has, such as a biometric or token/cell phone.

Vulnerability scan

An electronic survey of an organization’s systems looking for known vulnerabilities and misconfigurations that have security implications. Vulnerability scans are usually carried out by security experts that may be consultants or within an organization’s own IT security group. These scans differ from penetration testing in that they are not actually cyberattacks.

Zero-day exploit or vulnerability

A software vulnerability that is either not generally known or has been discovered but not yet remediated. All computers of the type involved are susceptible to exploit.