Security, Privacy and Confidentiality

Ashish Sharma
Department of Biomedical Informatics

Based on NIH Computer and Information Security Course
and an introductory lecture on Security presented by Dr. Tahsin Kurc

Information Security

- Clinical information systems sensitive information linked to patients (or participants in a study)
- **Sensitive information** and **patient's/participant's privacy** have to be protected
- There are institutional, state, and federal rules governing what information can be disclosed and to whom the information can be disclosed
- Individual / Institutional **intellectual property** has to be protected
- Clinical informatics systems and informaticians have to provide support for protecting sensitive information

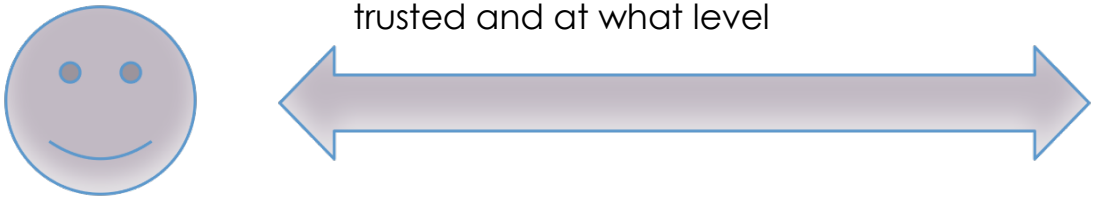
Security Concepts

Authorization and Access Control

Authorization: *Determining whether an entity can perform an operation on a resource*

Access Control: *Determining what an entity can do on a resource*

Trust Management: Managing whether an individual's credential should be trusted and at what level



Identification &Authentication

Identity Management: *Being able to uniquely and verifiably identify entities in an environment*

Authentication: *Being able to verify that an entity is who/what the entity claims to be*

Secure Information Storage and Exchange

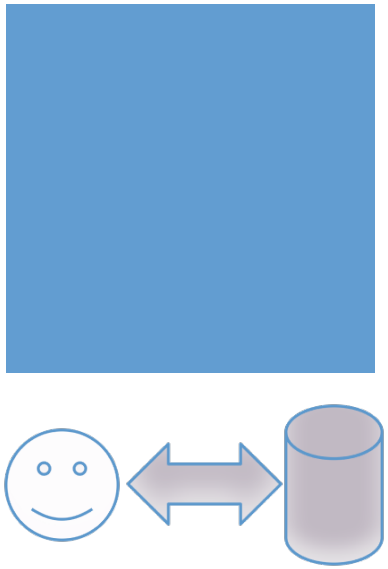
Data Integrity: *Encryption, Cryptographic hashing*

Communication Integrity: *Digital signatures*

Information/Systems Security Goals

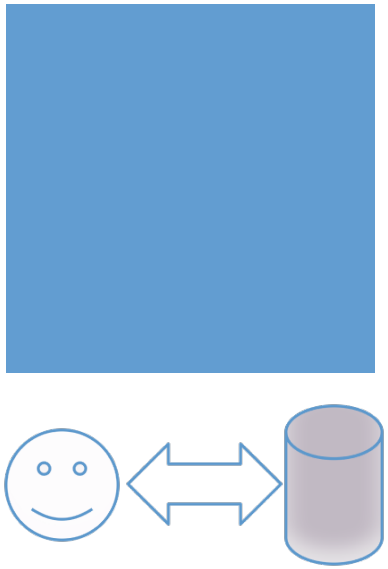
- **Confidentiality:** A resource is protected from unauthorized access and disclosure
- **Integrity:** Resource accessed, retrieved, or exchanged is not corrupted.
  - **Authenticity:** Verifiable assertion that content of a resource has not been altered by some other party
  - **Non-repudiation:** ensuring that when a message is exchanged between parties, the sender of a message cannot later deny having sent the message and that the recipient cannot deny having received the message.
  - **Accountability:** Actions performed on a resource can be tracked back to the entity that performed the actions
- **Availability:** A resource is available to authorized parties when needed

Secure Information Storage and Exchange



1. Cryptographic Hashes for **Integrity**
2. Encryption for **Privacy**
   - Symmetric Key Encryption
   - Public Key Infrastructure
3. Digital Signatures for **Authenticity**

1. Cryptographic Hashes



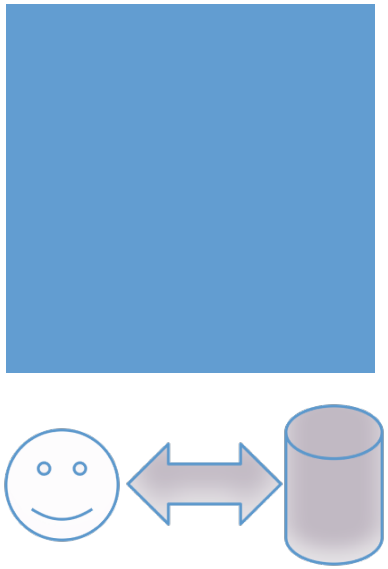
- A deterministic procedure that takes an arbitrary block of data and returns a fixed-size bit string
  - HashFunction(data) → Hash value
  - Must not be reversible
  - Should be designed to withstand brute-force attacks
- An accidental or intentional change to the data will change the hash value
  - Should produce results that have no or very low possibility of collision
- Examples
  - MD5: 128 bits output e.g. /etc/password in RHEL
  - SHA-1: 160 bits
  - SHA-2: 256 bits (Federally mandated for certificate issuers operated or funded by federal agencies)

Applications

- Password storage
  - Passwords are mapped to a hash value and stored in the system
  - When a password is to be checked, it is mapped to a hash value by the same method and compared to stored values
- File and message integrity checking
  - A check-sum value is generated for file or message
  - When a message is received, it is processed by the same hash function to generate a check-sum value
  - The two check-sum values are compared
- Secure Certificate Generation, Digital certificates

2. Encryption

- Re-arranging data temporarily into a form that is unreadable and unintelligible

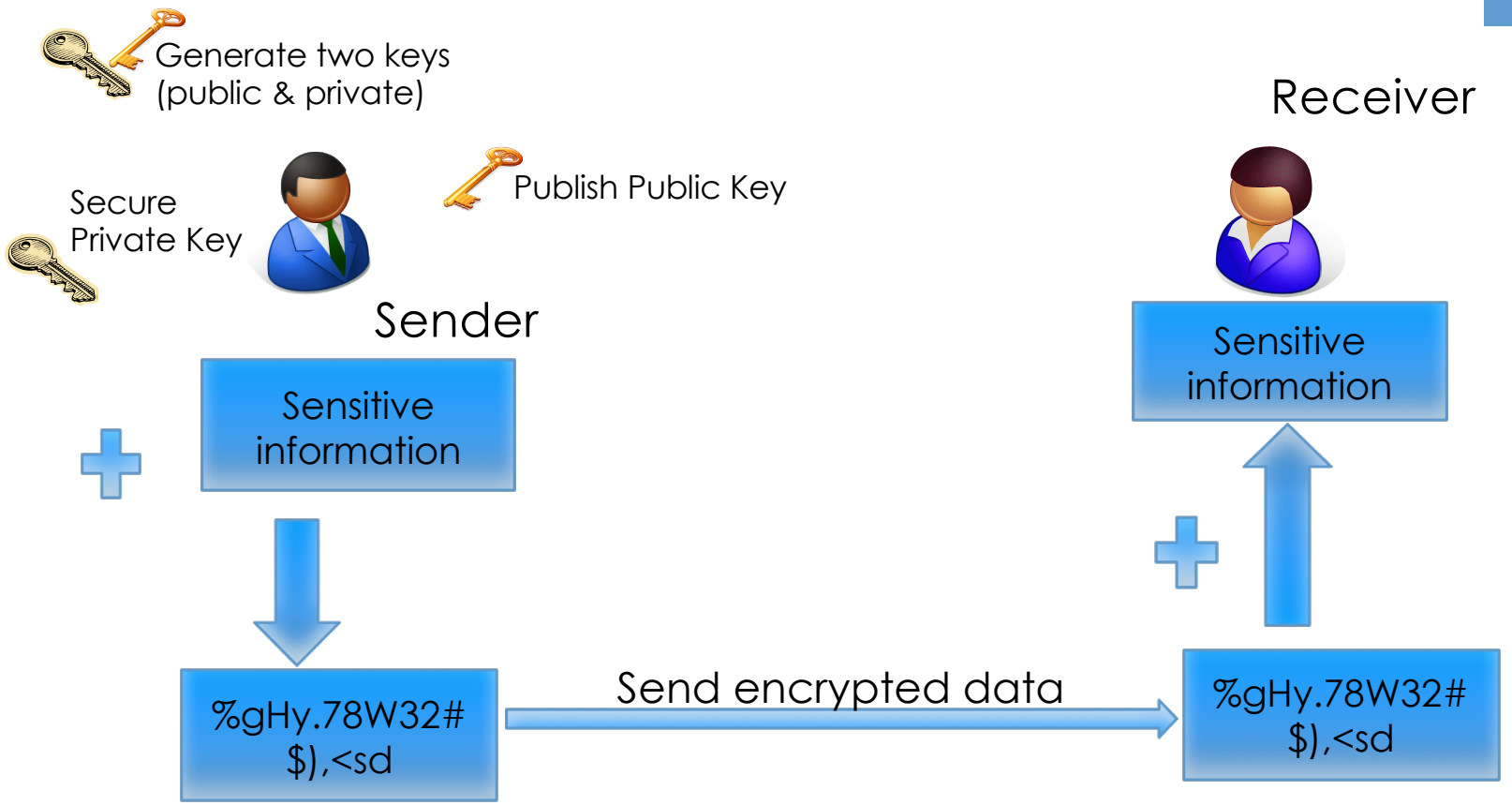


- Example Algorithms
  - AES (128, 192, 256 bit), Serpent, Blowfish, Twofish
- Software Based
  - TrueCrypt, PGP, FileVault
- Hardware Based

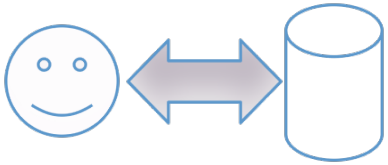
Symmetric-key vs Public Key Encryption

- One key is used in symmetric-key encryption
  - Key needs to be private
  - Key distribution problem
- Public Key Encryption
  - Asymmetric keys
  - One private and one public
  - Use one key for encryption and the other for decryption
  - Alleviates key distribution problem

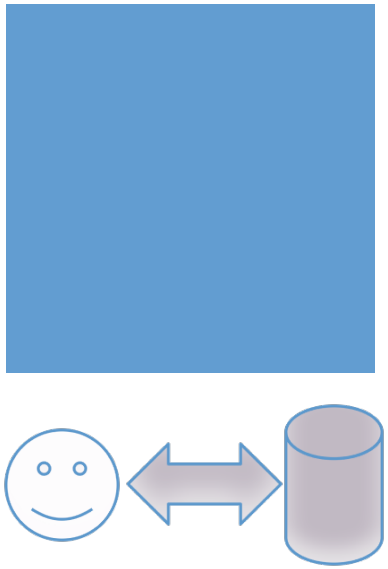
Public Key Encryption



How do I ensure that the data integrity is not compromised over the communication channel?

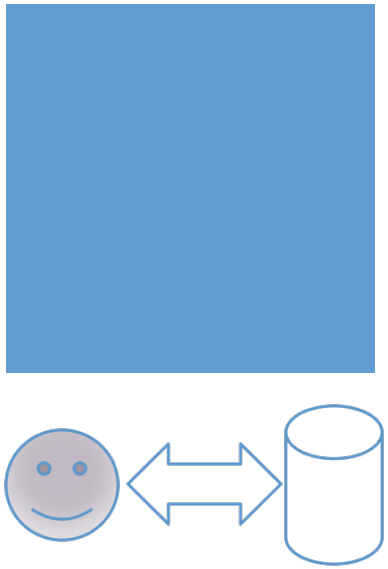


Secure Information Storage and Exchange



- ✓ Cryptographic Hashes for **Integrity**
- ✓ Encryption for **Privacy**
  - ✓ Symmetric Key Encryption
  - ✓ Public Key Infrastructure
- ✓ Digital Signatures for **Authenticity**

Identification and Authentication



- Uniquely and verifiably identifying a user in the environment
- Determining if the user is who he/she claims to be
- Common informatics methods
  - User name and password
  - Certificates (using Public Key Infrastructure)
- Identity Management System Examples
  - OpenLDAP
  - Microsoft Active Directory
  - SourceID
  - caGrid Dorian
  - OpenID

Some Relevant Standards

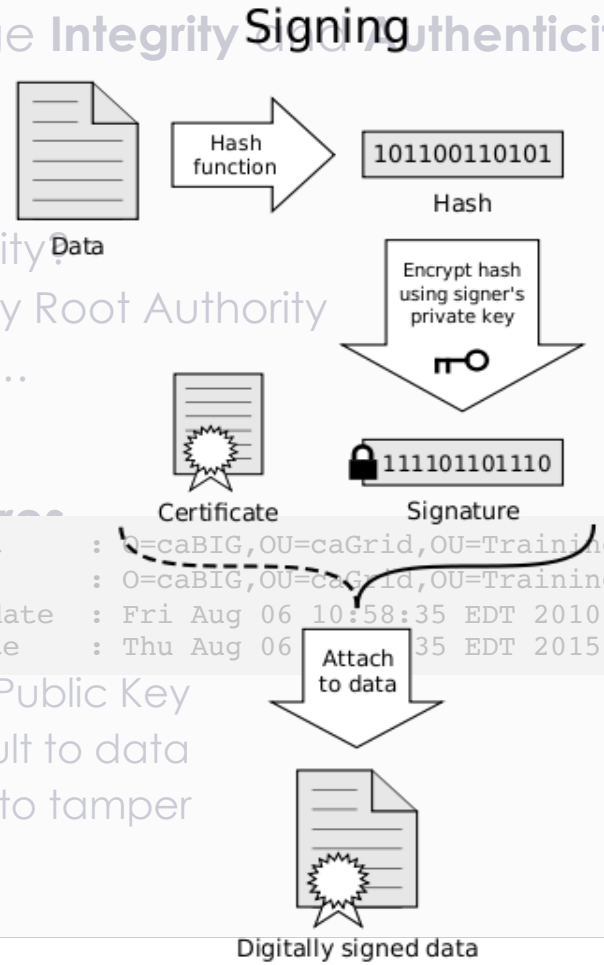
- X.509
  - A widely adopted standard for defining digital certificates
- SAML (Security Assertion Markup Language)
  - An XML-based framework for communicating user authentication, entitlement, and attribute information.
  - Entities make assertions regarding the identity, attributes, and entitlements of a subject (an entity that is often a human user) to other entities
- Shibboleth
  - A system for identity federation and secure exchange of authorization information (e.g., user attributes)

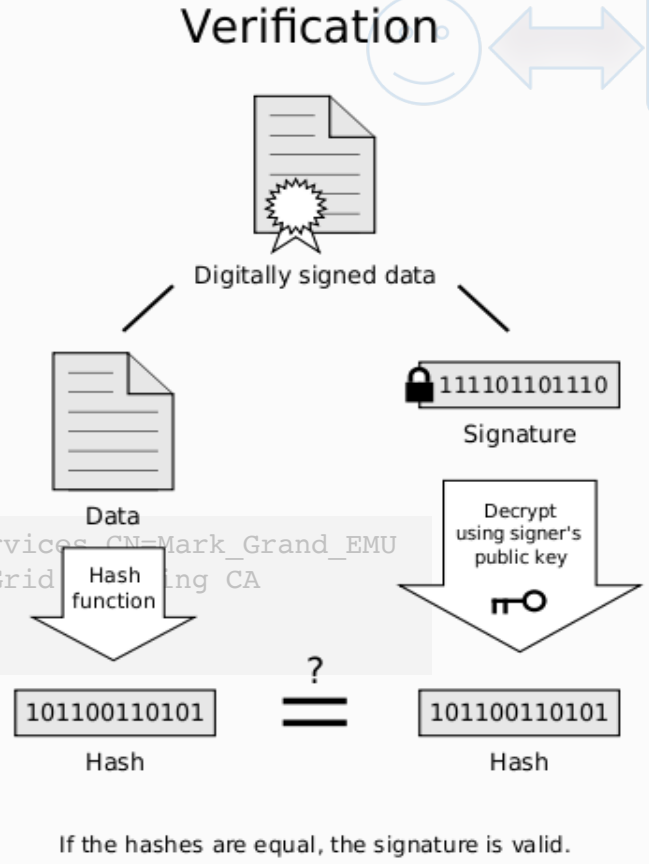
Authentication using Certificates

- Public Key Certificate = Public Key + Information about Identity
- Relies on Public Key Infrastructure (PKI)
  - Public key has information about user
  - Stored as X-509 certificate
  - Certificates obtained from a certificate authority
    - Self Signed
    - Endorsed — Verisign
- Proxy certificates created from permanent certificates
- Proxy certificates are used to interact with resources

Certificates and Digital Signatures

- Assure Message **Integrity** and **Authenticity**
- **Certificates**
  - Identify User
  - Certificate Validity?
    - Signed by Root Authority
    - Verisign …
- **Digital Signatures**
  - **Minimize**
    - Calcula
    - Encrypt it with Public Key
    - Attach the result to data
    - Data is harder to tamper

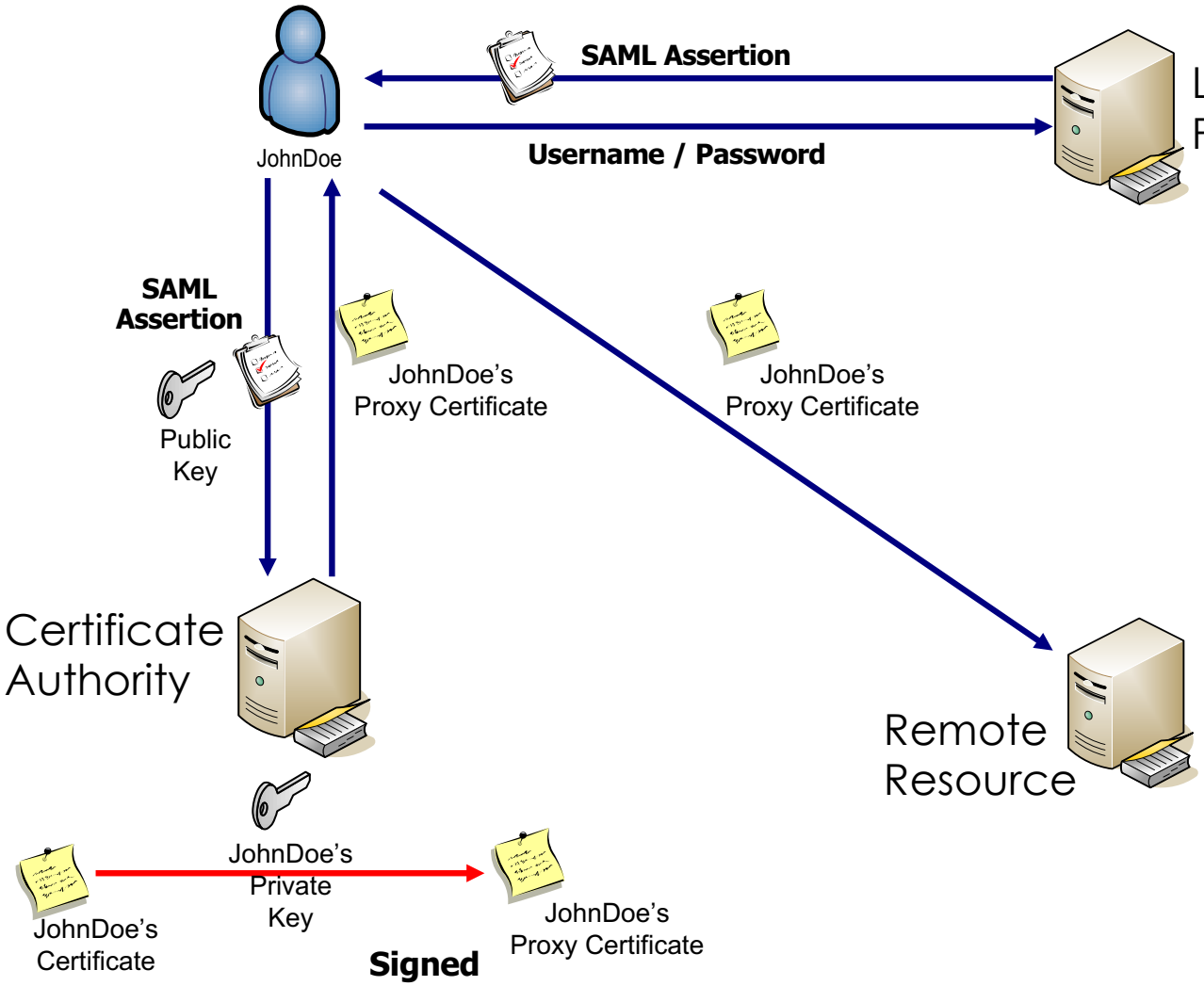




How do I know that the data that appears t have come from you, did in fact come from you?

http://en.wikipedia.org/wiki/File:Digital_Signature_diagram.svg

Authentication using Certificates



User Credentials

- Long term X.509 Certificate and Private Key.
- Clients authenticate using a proxy
- Proxy consists of a private key and proxy certificate signed by the user's long term private key.
- Short Lifetime
- Asserts Identity of users and services.
- Enables single sign-on
- Delegation

Authorization and Access Control

- Authorization determines whether a client is allowed to access and/or perform an operation on a given resource
- Access Control is closely related to Authentication and Authorization
  - Defining and enforcing who can access what and perform which operations

Early Access Control

- Secure Network not Resource
  - Firewalls, DMZ, Audit Trails
  - Assume control and ownership of network and physical resources
  - Centralized security management
- Access patterns are client-server
- Limited number of users
- Static, coarsely grained
  - E.g., User, Admin, Guest
  - No consideration of operational *contexts*

Access Control Models

- Access Control Lists
- Role-based Access Control (RBAC)
  - Users are assigned one or more roles
  - Access to a resource and permission to perform an operation are granted based on these roles
- Attribute-based Access Control (ABAC)
  - Access is granted based on the attributes of the requestor and the attributes of the resource.
  - ABAC policy specifies claims on what attributes should be satisfied and proved/verified to grant access to a resource.

XACML (eXtensible Access Control and Markup Language)

- **XML like**: you can actually read and write XACML with your favorite text editor
- **human-readable** and verbose enough for users to get an understanding of what it's doing
- belongs to the OASIS family of standards.
- **eXtensible**: you can add profiles to cater for specific scenarios
  - a profile for hierarchical resources,
  - role-based access control
  - export control…
- is about **access control**: authorizing who can do what when and how
- implements **ABAC**, attribute-based access control

XACML Architecture

- **Policy Decision Point (PDP)**: core of the architecture where policies are evaluated and decisions are reached.
  - PDP ≈ Supreme Court
- **Policy Enforcement Point (PEP)**: integration items which can be anywhere in an application architecture. PEPs are extremely versatile depending on where they enforce access control.
  - PEP ≈ Police, Judges…
- **Policy Retrieval Point (PRP)**: point through which policies are read from the policy repository. Policy retrieval points ensure the independence of XACML from specific storage mechanisms.
  - PRP ≈ place where law is written and maintained.
- **Policy Information Point (PIP)**: Attributes or Protection Elements. Describe users, services, resources, actions, and the environment. Policy Information Points are attribute stores. They can be any format and located anywhere.
  - PIPs ≈ police records, census bureau, etc… are to a nation and its citizens.
- **Policy Administration Point (PAP)**: this is where you manage your policies.

http://www.webfarmr.eu/2010/09/xacml-101-a-quick-intro-to-attribute-based-access-control-with-xacml/

XACML: Standard for representing access control policies

- e**X**tensible **A**ccess **C**ontrol **M**arkup **L**anguage
- Policy & PolicySet – combining of applicable policies using CombiningAlgorithm
  - Contains: Description, Target, Rules, Obligations, Rule Combining Algorithm
- Target – Rapidly index to find applicable Policies or Rules
  - Normally use Subject or Resource
  - Matches against value
- Rules – Smallest unit of administration
  - Contains: Description, Target, Condition, Effect
- Effect – Permit or "Deny"
- Combining Algorithms: Deny-overrides, Permit-overrides, First-applicable, Only-one-applicable
  - Applicable for Policies
- Support for various datatypes

Source: XACML and Federated Identity, NASA Goddard Space Flight Center, http://sewpsc.sewp.nasa.gov/documents/Lockhart_IdM.ppt

Regulations

HIPAA, PHI…

Security and Privacy

- Federal, state, and local laws govern access to and control of health record information, particularly:
  - Who can have access
  - What should be done to protect the data
  - How long the records should be kept
  - Whom to notify and what to do if a breach is discovered

HIPAA

- Health Insurance Portability and Accountability Act
- One part of it deals with insurance coverage
- Another part regulates usage and sharing of health information
  - Privacy Rules
  - Transactions
  - Security Rules
  - Enforcement Rules
  - UID Rules

Security and Privacy: HIPAA

- HIPAA = Health Insurance Portability and Accountability Act of 1996
- One part deals with insurance coverage while the second deals with usage of information
  - Protected Health Information (PHI) includes any health information that:
    - Explicitly identifies an individual
    - *Could reasonably be expected to allow* individual identification.
  - Excludes PHI in education records covered by Family Educational Rights and Privacy Act (FERPA), employment records.

Security and Privacy: HIPAA (cont'd)

18 identifiers recognized as providing identifiable links to individuals.

- Name, address, ZIP code
- Dates (birth dates, discharge dates, etc.)
- Contact info, including email, web URLs
- Social Security Number or record numbers
- Account numbers of any sort
- License number, license plates, ID numbers
- Device identifiers, IP addresses
- Full face photos, finger prints, recognizable markings

(Summary of the HIPAA Privacy Rule, n.d.)

Who does it affect

- Health insurance companies, HMOs, corporate health plans, Medicare and Medicaid
- Most healthcare providers — hospitals, clinics, doctors, psychologists, chiropractors, pharmacies, nursing homes…
- Healthcare Clearinghouse — work with healthcare data → transcription, coding...

Covered Entity

What is HIPAA Privacy?

- Federal law governing privacy of patients' medical records and other health information maintained by covered entities including:
  - Health plans, including Veterans Health Administration, Medicare, and Medicaid
  - Most doctors & hospitals
  - Healthcare clearinghouses
- Gives patients access to records and significant control over use and disclosure.
- Compliance required since April 2003.

(Summary of the HIPAA Privacy Rule, n.d.)

HIPAA Privacy Rule

- When can PHI be disclosed by a covered entity
- It is the responsibility of the covered entity to make **reasonable** efforts to limit the use or disclosure of PHI to achieve the stated goals of data access.
- You, as a patient, could waive the privacy rule
  - *Read the papers you sign before you see your doctor*

HIPAA Privacy Rule

- Privacy and security complaints
  - All investigated by Office of Civil Rights (OCR) of Dept. of Health and Human Services (HHS), as of 2009.
  - 35,386 complaints received (as of July 2016), of which 24,331 required corrective actions.
  - 143M individuals affected (as of July 2015)
  - Steep fines for validated complaints.
  - Entities needing the most corrective actions:
    - Private health care practices
    - General hospitals
    - Pharmacies
    - Outpatient facilities
    - Group health plans

(HIPAA Enforcement Highlights, 2012; Numbers at a Glance, n.d.; Poremba, 2008; Hamilton, 2009)

HIPAA Privacy Rule (cont'd)

Violations investigated most often:

1. Impermissible uses and disclosures of protected health information (PHI)
2. Lack of safeguards of PHI
3. Lack of patient access to their PHI
4. Uses or disclosures of more than the minimum necessary PHI
5. Complaints to the covered entity

(HIPAA Enforcement Highlights, 2012; Numbers at a Glance, n.d.; Poremba, 2008; Hamilton, 2009)

HIPAA Security Rule

- Established standards for securing electronic protected health information (ePHI) created, received, maintained, or transmitted.

- Entities required to:
  - Ensure confidentiality, integrity, availability of all ePHI
  - Identify and protect against reasonably anticipated threats to the security or integrity of the information.
  - Protect against reasonably anticipated, impermissible uses or disclosures.
  - Ensure compliance by workforce.
- Works in tandem with Privacy Rule.

(Summary of the HIPAA Security Rule, n.d.)

What is Required by HIPAA Security Rule?

Categories:

1. Administrative safeguards
2. Physical safeguards
3. Technical safeguards

(Summary of the HIPAA Security Rule, n.d.)

Administrative Safeguards

- Address process of security management in your organization.
- Risk analysis
  - Evaluating likelihood and impact of potential risks to ePHI
  - Implementing appropriate security measures to address identified risks
  - Documenting security measures chosen, with rationale
  - Maintaining continuous, reasonable, appropriate protections
- Ongoing process, with regular reviews.

(Summary of the HIPAA Security Rule, n.d.)

Administrative Safeguards (cont'd): Access policy

- Policies & procedures for authorizing access to ePHI only when appropriate for one's role (role-based access).
  - Who gets access to ePHI data?
  - What level of access is needed?
  - Who is the agent authorizing the access?
  - Is this authorization adequately documented?
  - Is the access periodically reviewed?
  - Is there a process for rescinding access when no longer needed?

(Summary of the HIPAA Security Rule, n.d.)

Physical Safeguards: Access

- Limit physical access to facilities, while ensuring that authorized access is allowed.
  - Server rooms where ePHI is stored
  - Work areas where ePHI is accessed
  - Back-up media storage potentially containing ePHI
- Inventory hardware and software.
  - Know where inventory is kept.
  - Know value of hardware, software, equipment.

(Summary of the HIPAA Security Rule, n.d.)

Physical Safeguards (cont'd): Device Security

- Policies and procedures for proper use of & access to workstations & electronic media, including transfer, removal, disposal, re-use.
  - Lock down publicly-accessible systems potentially containing ePHI.
  - Strong passwords
  - At least 256-bit encryption, especially for wireless, backups, & offsite data
  - Media thoroughly wiped and rendered inaccessible

(Summary of the HIPAA Security Rule, n.d.)