Federazione e dati personali: quali tutele per utenti e gestori di identità?

Norberto Gavioli

Università dell'Aquila

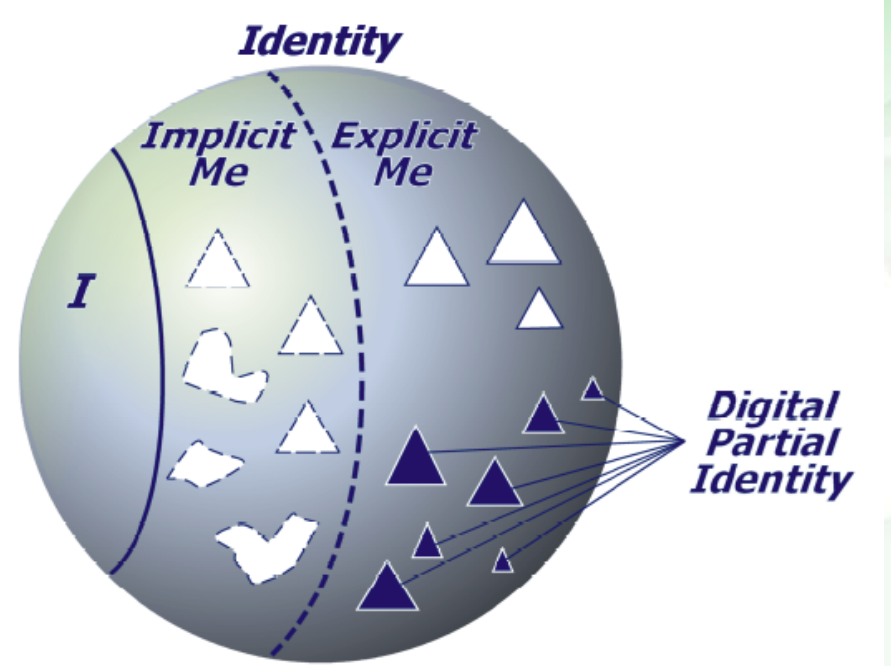
Identità

- può essere definita come l'esclusiva percezione della propria vita, che però è sempre collegata ad un corpo mediante il quale la persona si relaziona con la società.
- I = autopercezione di sé
  Me = quegli attributi di me che sono accessibili agli altri mediante una comunicazione

https://www.datenschutzzentrum.de/idmanage/study/ICPP_SNG_IMS-Study.pdf

Identità <> Identità Digitali Parziali

- Insieme di proprietà (attributi) di una persona che sono tecnicamente, immediatamente e operativamente accessibili
- Tutti quei dati personali che possono essere memorizzati e collegati tramite applicazioni informatiche



https://www.datenschutzzentrum.de/idmanage/study/ICPP_SNG_IMS-Study.pdf

Identità digitale parziale

- Relativa ad un contesto
- In ogni contesto esistono attributi necessari ed altri non necessari
- In certi contesti la persona vuole rimanere anonima, in altri contesti preferisce presentarsi con uno pseudonimo, in altri casi è necessario rivelare l'identità reale
- E' necessario poter collegare tra loro le identità digitali parziali relative ad una stessa persona?

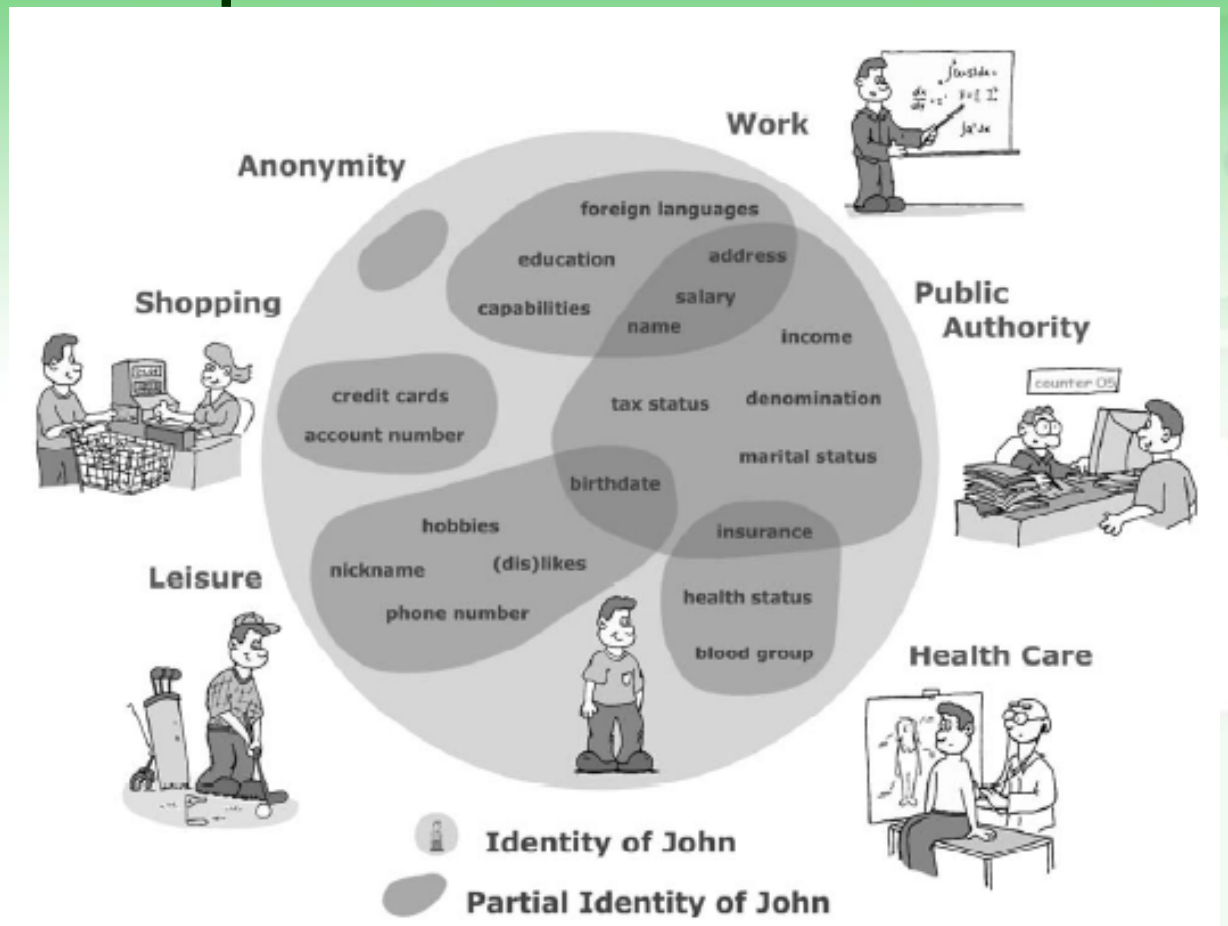


Fig. 1 – Partial identities of John.

(Borcea-Pfitzmann et al., 2006).

Trattamento dati personali

- In Italia il trattamento dei dati personali, quindi anche delle identità digitali parziali, è regolato dal **Decreto legislativo 30 giugno 2003, n. 196 - Codice in materia di protezione dei dati personali** (Testo unico privacy)
- I sistemi di Identity Management devono essere progettati in modo da rispettare la legge, in particolare proteggendo i dati (Art.1).

Trattamento dati personali

- **Art. 3. Principio di necessità nel trattamento dei dati**
  1. I sistemi informativi e i programmi informatici sono configurati riducendo al minimo l'utilizzazione di dati personali e di dati identificativi, in modo da escluderne il trattamento quando le finalità perseguite nei singoli casi possono essere realizzate mediante, rispettivamente, dati anonimi od opportune modalità che permettano di identificare l'interessato solo in caso di necessità. *(n.d.r. pseudonomizzazione)*

Sempre dal testo consolidato in vigore (Art. 7 par. 2)

- L'interessato ha diritto di ottenere l'indicazione:

a) dell'origine dei dati personali;

b) delle finalità e modalità del trattamento;

c) della logica applicata in caso di trattamento effettuato con l'ausilio di strumenti elettronici;

d) ...

Principali aspetti riguardanti la tutela della privacy

- Nello scambio di dati relativi all'accesso di un utente ad un servizio si deve cercare di assicurare *quanto più possibile*:
  - confidenzialità dei dati,
  - anonimato dell'utente.

Confidenzialità

- Per confidenzialità si intende il rendere disponibili i dati scambiati all'interno di una transazione ad un insieme predefinito di entità (audience) e contemporaneamente renderli inaccessibili a terze parti.

Anonimato

- È più complesso dare una definizione esaustiva di anonimato. Si potrebbe pensare all'anonimato come una misura della difficoltà di associare dei dati a chi li possiede.
- L'anonimato garantisce l'utente rendendo difficoltosa la raccolta di informazioni sull'interessato.

Confidenzialità: insufficiente se sola precauzione

- La sola confidenzialità di un dato non assicura la privacy dell'utente. A volte il solo sapere che un utente accede ad una risorsa può essere un'informazione potenzialmente dannosa per l'utente stesso.

L'anonimato come appartenenza ad un insieme

- Essere anonimi rafforza la tutela della privacy
- L'anonimato è da intendersi come la possibilità di individuare i soggetti solo all'interno di un insieme *numeroso* di utenti che accedono ad un altrettanto *numeroso* insieme di risorse

Pseudonimi

- Una possibile via per ottenere buoni livelli di anonimato è quella di impiegare pseudonimi
- Uno pseudonimo è un identificativo unico che permette al solo fornitore di credenziali di risalire all'identità dell'individuo

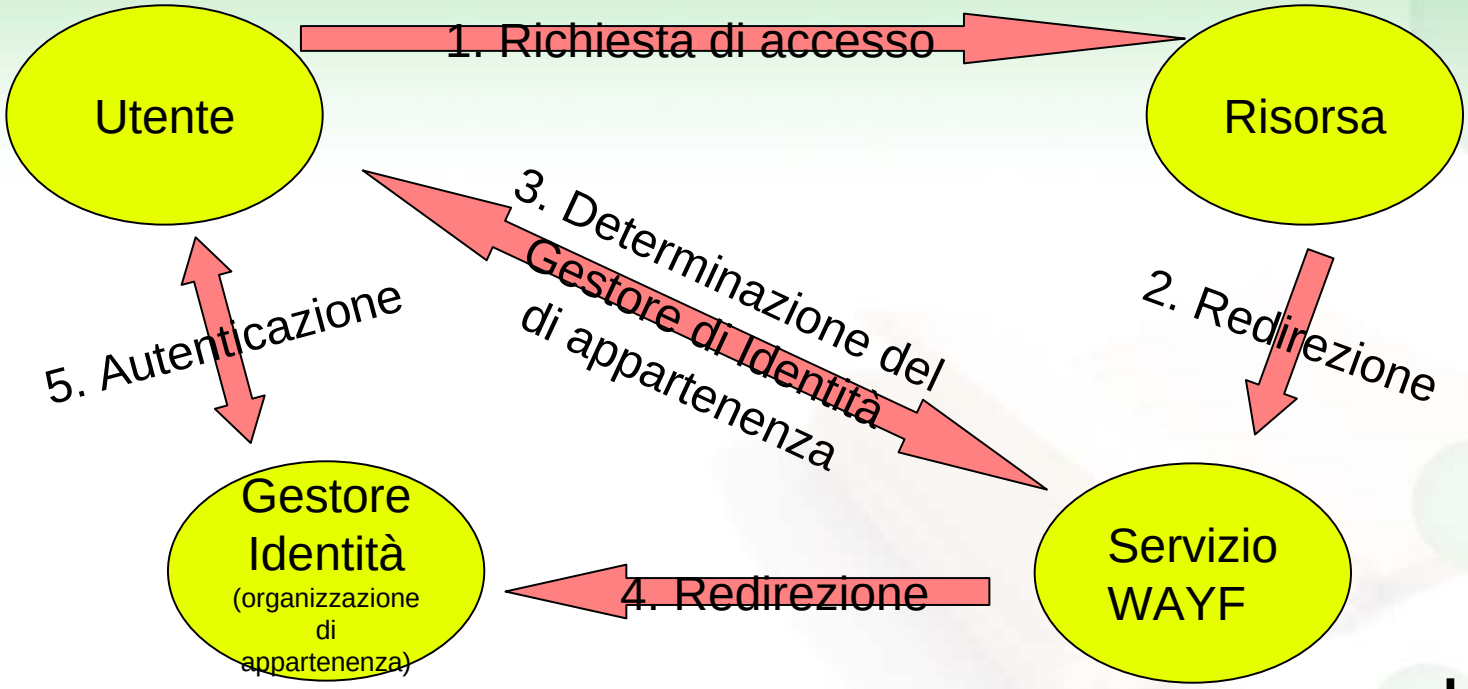
Pseudonimi *one time* e pseudonimi persistenti

- Gli pseudonimi possono avere carattere temporaneo oppure persistente
- Pseudonimi ad uso singolo forniscono un alto livello di anonimato, ma non permettono all'utente e alle risorse di storicizzare le sessioni
- Pseudonimi permanenti sono suscettibili al tracciamento da parte di osservatori

Ruoli ed anonimato

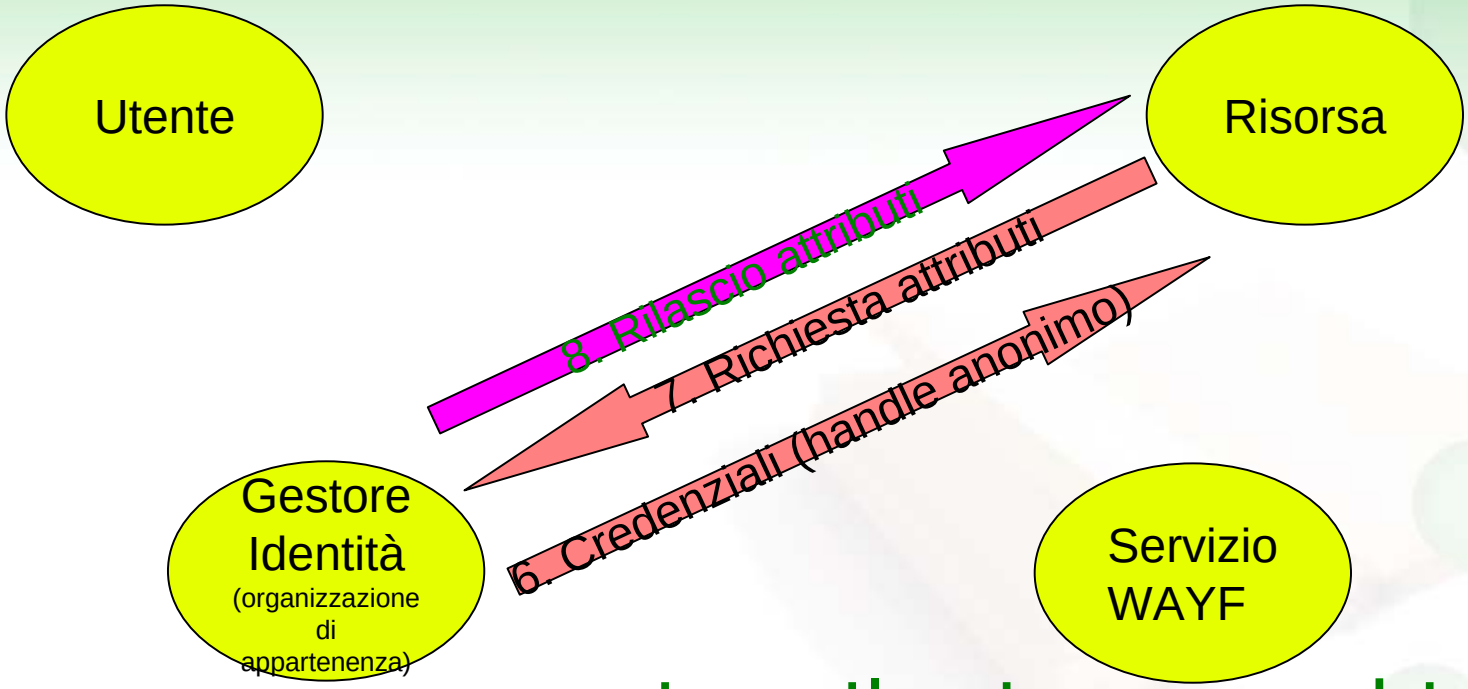
- Un utente può essere facilmente identificato in base al suo ruolo.
  - Se ad esempio il rettore di un'università accede ad una funzionalità specifica di una risorsa in base al rilascio contemporaneo del suo identificativo persistente e del suo ruolo come attributo, verrà successivamente identificato come persona fisica anche per usi diversi della stessa risorsa.
  - Per questi tipi di funzionalità bisogna prevedere servizi specifici con autorizzazione basata esclusivamente su attributi (eduPersonEntitlement) che garantiscono il diritto di accesso alla risorsa.

Schema di transazione Shibboleth



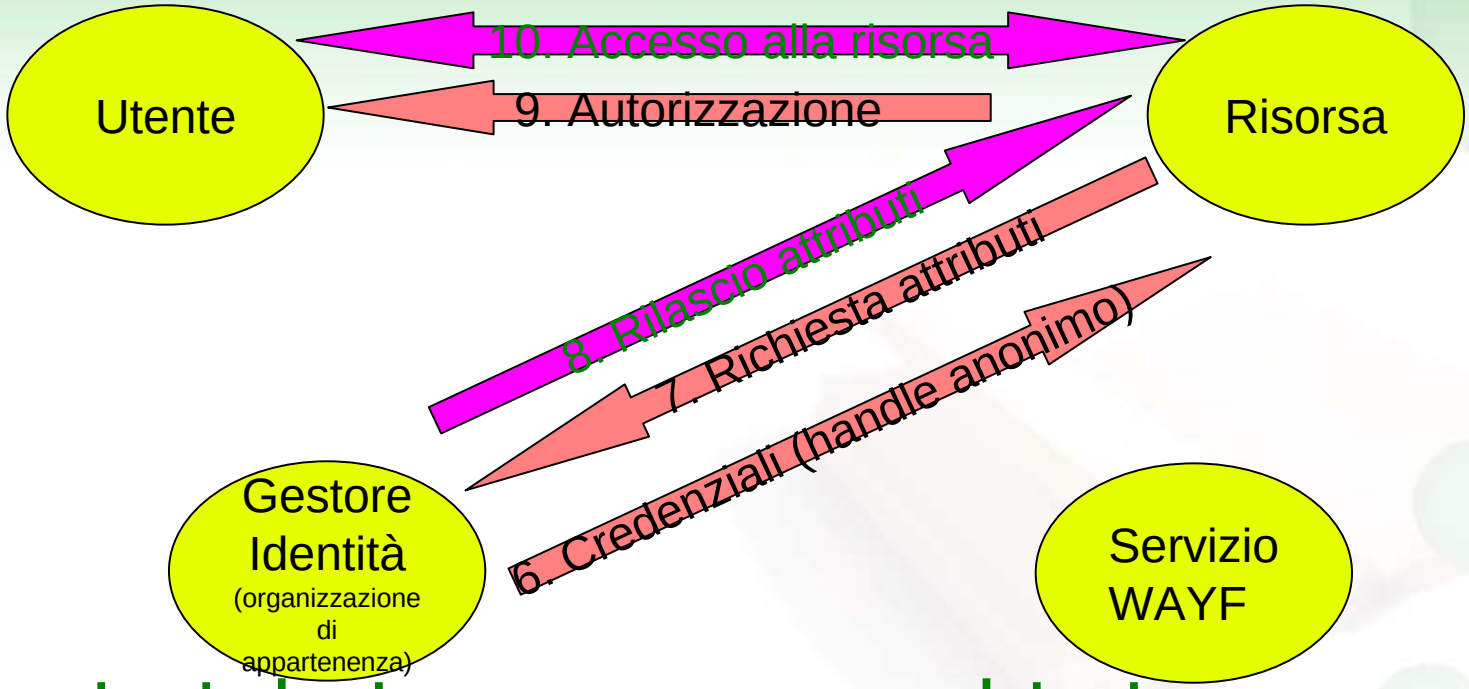
Fino a questo punto non vengono scambiati dati con la risorsa

Schema di transazione Shibboleth



A questo punto avviene il primo scambio di dati, sono coinvolti solo la Risorsa ed il Gestore Id.

Schema di transazione Shibboleth



Ulteriori dati vengono scambiati direttamente tra utente e risorsa

Privacy e Shibboleth

- Fobie da sfatare
  - Non *circolano* password: il meccanismo di autenticazione è locale rispetto alla federazione e coinvolge solo utente e organizzazione di appartenenza.
  - La federazione non gestisce dati personali: eventuali dati personali vengono scambiati solo tra organizzazione di appartenenza e risorsa, oppure tra utente e risorsa e non coinvolgono la federazione.

Privacy e IDEM

- IDEM non è un soggetto che tratta dati personali
- Eventuali trasferimenti di dati personali avvengono solo tra singoli gestori di credenziali (IdP) e risorse fornitrici di servizi (SP), se previsto nei singoli contratti di fornitura, e vanno regolati tra questi in base alla normativa vigente.

Privacy e IDEM

- È a carico del singolo IdP informare i propri utenti sul trattamento dei dati previsto dall'utilizzo dei servizi dei propri fornitori.

Qualità dei dati

- È diritto dell'interessato che i dati oggetto del trattamento siano mantenuti accurati ed aggiornati, ed è anche, e soprattutto, una garanzia di affidabilità del modello di fiducia della federazione
- Con l'impiego di un IdP federato non è necessario aggiornare i dati degli utenti presso i singoli fornitore di servizi

Confidenzialità e Shibboleth

- La confidenzialità in una transazione Shibboleth viene assicurata da
  - Utilizzo di protocolli sicuri (SSL) per comunicazioni e certificati
  - Riconoscimento reciproco tra risorsa, WAYF e gestore di identità (IdP) tramite certificati (descritti nei metadati)
  - Protocollo interno ed esterno utilizzato dall'IdP per il servizio di autenticazione (JAAS)
  - Eventuale protocollo crittografico utilizzato per l'accesso alla risorsa da parte dell'utente
  - Sicurezza del protocollo SAML: asserzioni crittografate e firmate nella fase di invio delle credenziali

Anonimato e Shibboleth

- La prima limitazione dell'anonimato è insita nel fatto che SAML utilizza *autorità* a cui le varie *entità* coinvolte si riferiscono
- Nel caso di Shibboleth l'utente viene subito circoscritto tra quelli che si autenticano nell'ambito di un ben preciso IdP

Anonimato e Shibboleth

- L'invio di ulteriori attributi valorizzati da parte dell'IdP verso la risorsa, restringe ulteriormente l'insieme di appartenza dell'utente
- Sapere, ad esempio, che un utente appartiene ad un dipartimento con tre o quattro membri può (anche insieme ad altre informazioni) inficiare l'anonimato.

Buone pratiche per il rilascio degli attributi

- È importante rilasciare, in dipendenza della risorsa alla quale l'utente intende accedere, i soli attributi necessari alla finalità del servizio fornito dalla risorsa.
- Shibboleth permette di impostare le politiche di rilascio degli attributi tramite filtri
  `IdP_HOME/conf/attribute-filter.xml`
- Adottare uApprove (utente consapevole)
  http://www.switch.ch/aai/support/tools/uApprove.html

Impiego (obbligatorio) di eduPersonTargetedID

- Shibboleth permette di generare un identificativo anonimo persistente ottenuto con un algoritmo (digest) in funzione degli identificativi dell'utente, della risorsa e dell' IdP.
- Tale pseudonimo viene adottato da IDEM tramite l'attributo eduPersonTargetedID (vedi ST per la compilazione e l'uso degli attributi).

Vantaggi di eduPersonTargetedID

- Si tratta di uno pseudonimo che non permette il confronto degli identificativi utente rilasciati a risorse distinte.
- Per le risorse, è possibile storicizzare le sessioni, mantenendo le configurazioni / preferenze degli utenti senza doverne conoscere necessariamente l'identità.

SAML 2.0

- Shibboleth è basato su SAML
- SAML prevede che diverse autorità si scambino messaggi sotto forma di asserzioni
- Ci sono tre tipi fondamentali di asserzioni: Autenticazione, Attributi, Autorizzazione

Caratteristiche riguardanti la privacy in SAML

- Supporto di pseudonimi
- Supporto di identificatori *one time* o transienti
- Supporto di diversi livelli di autenticazione adeguati al tipo di risorsa
- Supporto per l'autenticazione federata

Principali caratteristiche di sicurezza di SAML

- Autenticazione
  - Sessioni attive autenticate in modo non persistente (canale di comunicazione)
  - Autenticazione permanente dei messaggi tramite protocollo di firma digitale (XMLsig)

Principali caratteristiche di sicurezza di SAML

- Confidenzialità
  - Utilizzo di TLS/SSL o Ipsec per il trasferimento
  - Crittografia XML per la codifica dei messaggi (XMLEnc)
- Integrità dei Dati
  - Utilizzo di TLS/SSL o Ipsec per il trasferimento
  - Firma digitale dei messaggi tramite (XMLsig)

Principali caratteristiche di sicurezza di SAML

- Gestione delle chiavi
  - È indispensabile basarsi su un sistema (autorità di certificazione) che assicuri fiducia nella corrispondenza (binding) delle chiavi alle rispettive identità (entità)
  - Il livello di fiducia nella corrispondenza delle identità digitali agli utenti dipende dal processo di accreditamento dei gestori di identità.

Conclusioni

- Il meccanismo di autenticazione federata IDEM non prevede trattamento di dati personali
- Eventuali scambi di dati avvengono tra organizzazione di appartenenza e forntori di servizi
- La sicurezza delle transazioni di autenticazione è garantita da protocolli crittografici standard
- È *possibile* lasciare all'utente ampie marginalità di decisione sui dati rilasciati alle risorse a cui si accede tramite IDEM