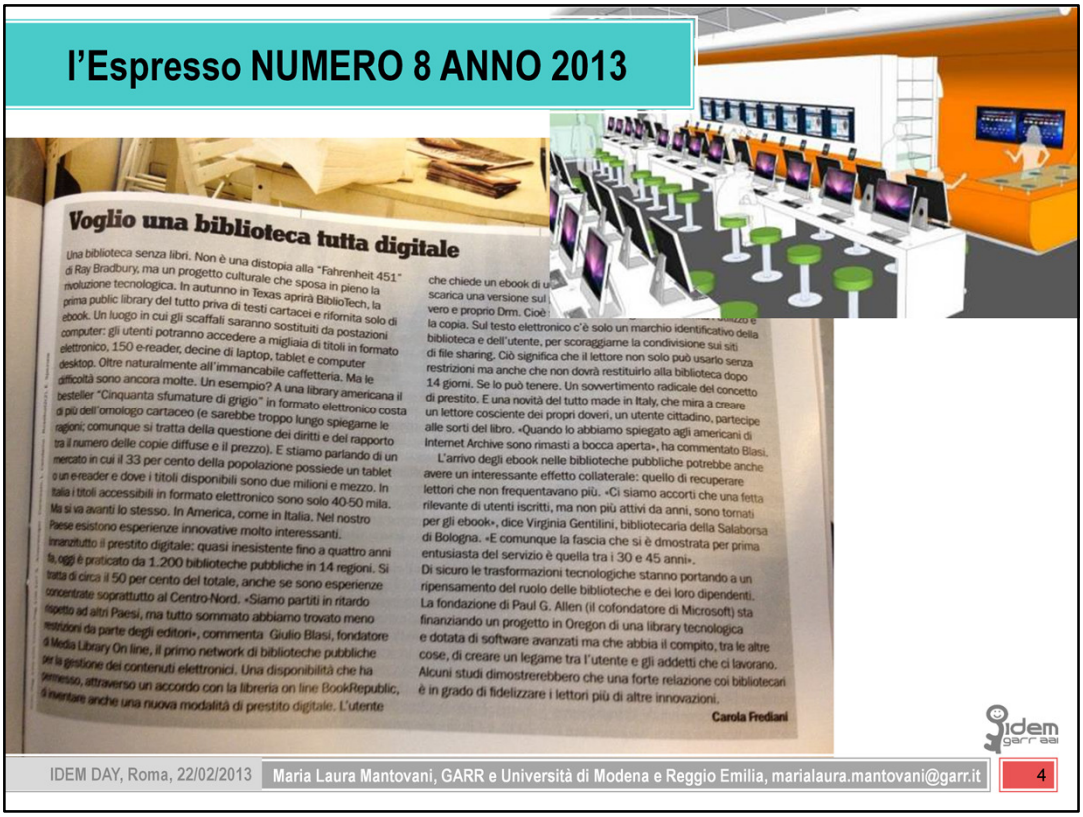
L'identità federata in biblioteca (…e non solo)

IDEM DAY || STANDARD EDITION
ROMA, 26 FEB 2013

Il problema che abbiamo davanti è l'accesso alle informazioni. Come favorirlo?
Le informazioni disponibili aumentano. E non c'è dubbio che Google abbia dato un impulso straordinario relativamente a come le informazioni debbano essere disponibili e reperite.
Le necessità e le preferenze degli utenti delle biblioteche, in alleanza con le nuove tecnologie, ci stanno aiutando a condurre cambiamenti nelle biblioteche. Studenti, ricercatori, insegnanti si aspettano di accedere alle informazioni senza sosta, da qualsiasi parte del mondo, e attraverso un numero sempre crescente e diversificato di attrezzature, dai pc portatili, ai tablet, ai telefoni.

Cosa significa questo per le librerie universitarie come le conosciamo oggi? Cosa saranno le biblioteche tra 10 anni? Esisteranno nella forma fisica come le conosciamo? Che ruolo avranno i bibliotecari nel supportare l'apprendimento e la ricerca nell'era digitale?
Le biblioteche sono a un punto di svolta. Mentre la tecnologia trasforma rapidamente il nostro modo di accedere alle informazioni, e le risorse sono sempre disponibili on-line e in formato digitale, il ruolo consolidato della biblioteca come spazio di scaffali per alloggiamento di libri sembra non essere in grado di stare al passo con le esigenze di studenti e ricercatori. I bibliotecari devono continuare il loro ruolo di facilitatori all'accesso alle informazioni, ma sicuramente in un modo nuovo.

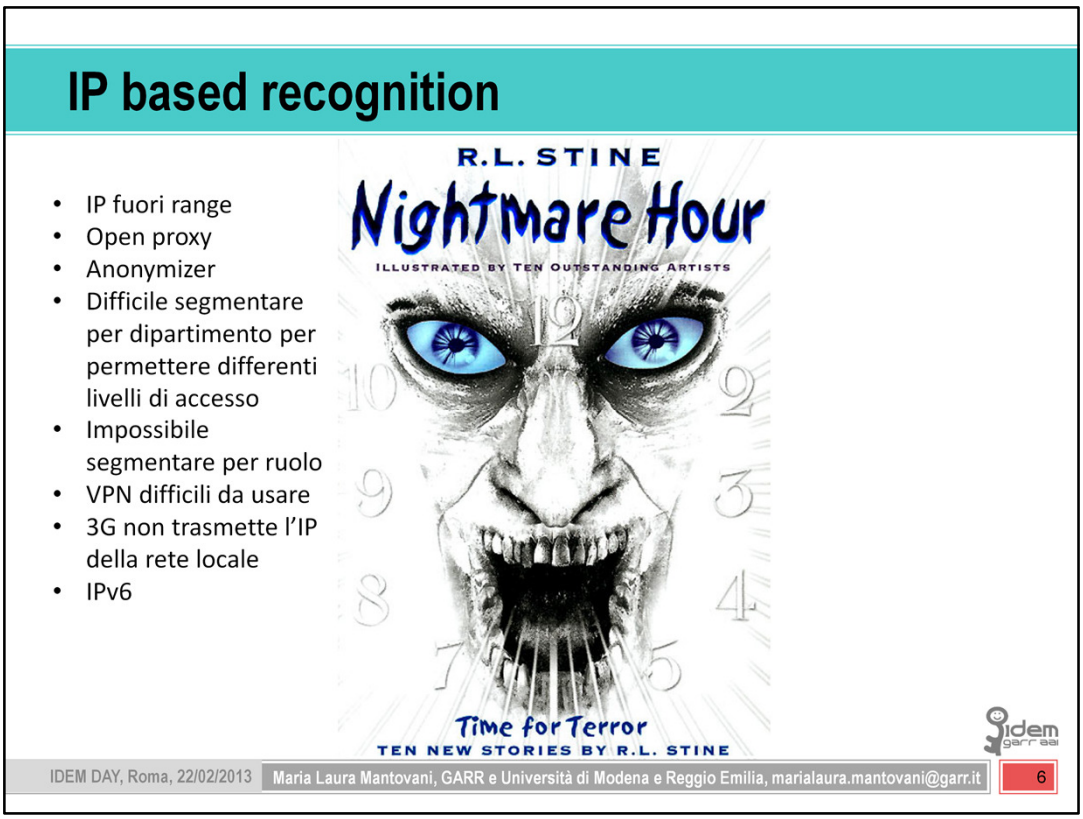


Questo è un articolo apparso sull'ultimo numero dell'espresso. È l'annuncio di una biblioteca tutta digitale che sarà aperta a San Antonio Texas il prossimo autunno e che partirà con una dotazione di 100 e-readers per il prestito, e 50 e-readers per bambini, 50 pc, 25 laptop e 25 tablet nelle sale di lettura [1]. L'articolo prosegue dicendo che anche in italia c'è una esperienza analoga che è già presente in 1200 biblioteche pubbliche di 14 regioni italiane (circa il 50% del totale) che permette già ai propri utenti di scaricarsi in prestito e-book, giornali, film e musica, stando a casa propria, sul proprio PC. Per chi, come me, è utente delle biblioteche della provincia di Modena, l'accesso è già consentito anche tramite autenticazione IDEM.

[1] http://www.pcworld.com/article/2025194/all-digital-library-system-is-planned-for-texas-community.html

Quindi la risposta delle biblioteche alla domanda di cambiamento è stata rapida. La biblioteca si sta già attrezzando per supportare l'accesso a una vasta gamma di risorse online, ampliando la partecipazione, fornendo approcci all'apprendimento più orientati agli studenti, supportando collaborazioni di ricerca innovative, complesse e distribuite. Abbiamo visto organizzazioni (intendo Università italiane [2], ma anche un sistema bibliotecario pubblico come il CEDOC della provincia di Modena [3]) che hanno tratto vantaggio dall'implementazione del SSO. Fare funzionare meglio il SSO sarà la chiave per l'utente per accedere ai contenuti a cui ha diritto. Da una parte molti SP hanno implementato il SSO, ma sarebbe più corretto dire che molti NON l'hanno ancora implementato. Per questo motivo le biblioteche sono costrette ad operare in un ambiente che contiene un mix di sistemi diversi di autenticazione, dove il riconoscimento basato su IP è ancora il metodo più comune

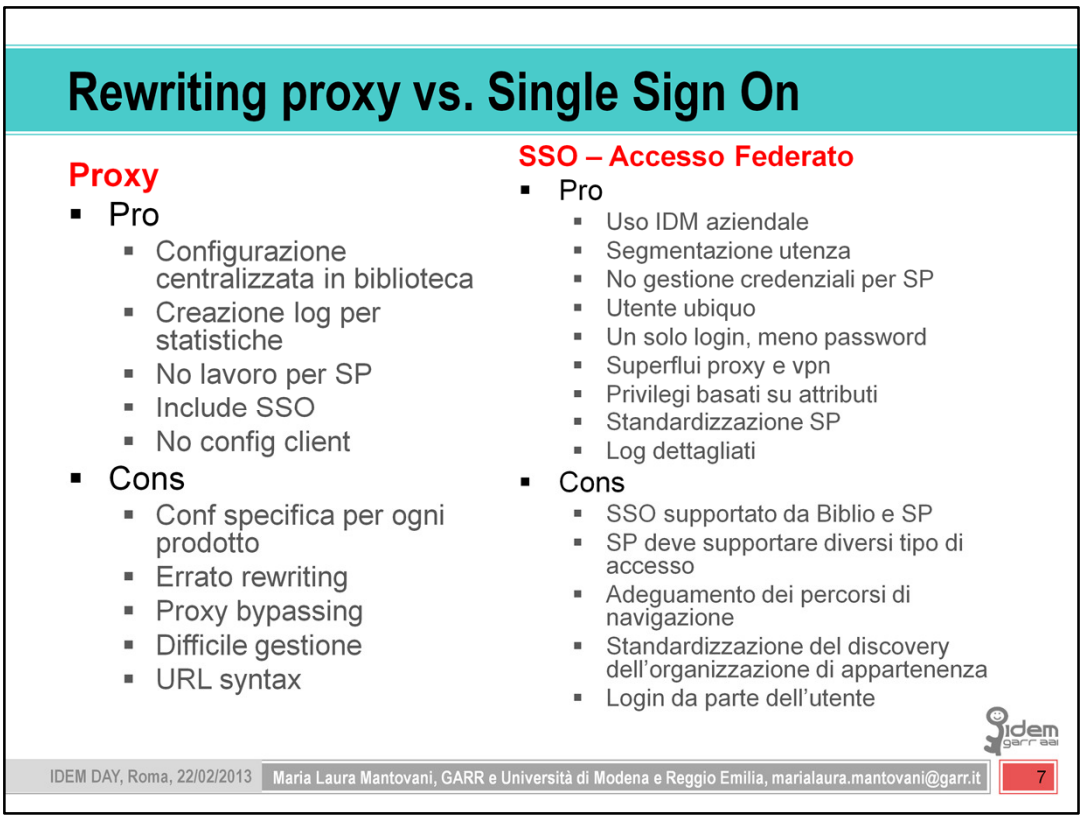
[2] Accesso federato alle risorse bibliografiche on-line; l'esperienza del Sistema Bibliotecario di Ateneo dell'Università degli Studi di Trento,Tullio NICOLUSSI (Università degli Studi di Trento), http://www.garr.it/a/conf11-prog/tab?connectionname=conf_spe&spe=19
[3] IDEM BIBLIO DAY 2012: **Il CeDoc: un'esperienza concreta di sinergia tra enti per migliorare i servizi**– L. Prampolini (CEDOC, Provincia di Modena), https://www.idem.garr.it/it/documenti/doc_download/191-presentazione-cedoc



Le richieste di accesso alle informazioni da parte dell'utenza coinvolgono differenti sistemi e piattaforme, che richiedono autenticazione e autorizzazione (essendo prodotti commerciali). Con la crescita della complessità sia delle licenze che della rete (mi riferisco all'accesso da molteplici e differenziati dispositivi soprattutto mobili) il sistema di controllo dell'accesso basato o sull'IP di provenienza o sull'assegnazione di credenziali fornite dalla piattaforma non funzionano più bene.
La gestione è diventata troppo onerosa, troppo confusa e soggetta ad errori sia da parte dell'utente, che da parte delle biblioteche, che da parte dei fornitori di risorse.

Inconvenienti dell'autenticazione basata su IP

- Il terminale dell'utente non si trova nel range di indirizzi IP permessi (molto frequente con mobile device)
- Una macchina del campus diventa un open proxy permettendo a chiunque di accedere alle risorse licenziate
- Nei casi di abuso come il precedente è difficile identificare il responsabile
- Alcune organizzazioni internazionali usano "anonymous proxy" che oscurano l'indirizzo IP di provenienza
- Alcune gestioni della rete non permettono di segmentare i dipartimenti secondo range di IP ed in tali casi non è possibile individuare un dipartimento che richiede una licenza particolare.
- Poiché tutti gli utenti sono anonimi, è impossibile segmentare le licenze sulla base delle categorie (professori, studenti, …)
- Gli accessi dall'esterno dell'organizzazione forniti tramite VPN sono ostici agli utenti e sono soggetti ad interruzioni e per questi motivi molte organizzazioni hanno rinunciato ad usare questo metodo
- I terminali che usano 3G a volte bypassano l'indirizzo IP anche se sono connessi alla rete locale
- La transizione verso IPv6 renderà la gestione degli IP ancora più difficile e confusa



La situazione attuale di eterogeneità negli approcci all'autenticazione a all'autorizzazione sta diventando da incubo sia per le organizzazioni licenziatarie (biblioteche), sia per i fornitori di risorse, sia per gli utenti, perché oltre al riconoscimento basato su IP sono stato introdotti anche altri sistemi, come proxy, VPN e SSO, Siamo in una fase di transizione verso i nuovi modelli, ma è probabile che per un po' di tempo dovremo ancora lavorare in un ambiente eterogeneo e convivere con molteplici sistemi di identificazione. I problemi legati al riconoscimento basato su IP e legati all'uso di proxy hanno condotto allo sviluppo del SSO: per l'utente: una unica autenticazione per accedere a tutte le risorse. Non è più necessario avere un terminale collegato alla rete della biblioteca, si può essere ovunque; per la biblioteca: una gestione unificata delle identità degli utenti.

Anche i proxy e il SSO hanno pro e contro, ed è necessaria una soluzione efficace per affrontare questo ambiente ibrido e comprendere come fare interoperare al meglio proxy server e SSO. Dato che questa eterogeneità è causa di problemi per l'utente, tutte le organizzazioni sono fortemente incoraggiate a passare ai nuovi modelli al più presto possibile.

Proxy

Vantaggi

Nella maggior parte delle situazioni funzionano bene, perciò sono largamente usati

La configurazione di tutte le risorse viene fatta in un unico posto

Tutte le ricerche e gli accessi avvengono attraverso il proxy, i logfile possono essere usati per statistiche

Per gli SP non c'è lavoro aggiuntivo

Per l'utente l'accesso è simile al SSO perciò la percezione è che sia trasparente

La maggior parte dei proxy non richiedono più la configurazione della postazione dell'utente

Svantaggi

Occorre gestire la configurazione per ogni prodotto licenziato

Se la pagina contiene script o incorpora link molto profondi, la riscrittura della pagina può essere errata
È possibile che l'utente arrivi alla risorsa bypassando il proxy e in questo caso non ottiene l'accesso. Il proxy non partecipa quando vengono usati i motori di ricerca sul web, quando vengono referenziati link protetti in articoli. L'utente che sa che l'organizzazione ha la licenza per quella risorsa rimane confuso, l'utente che non lo sa pensa di non potervi accedere.
Gestire un proxy richiede la gestione di un server e competenze tecniche molto rare.
Gli insegnanti che vogliono aggiungere deep link ai propri corsi dovrebbero conoscere la "sintassi del proxy" per essere sicuri che funzioni

Accesso federato

Vantaggi

L'organizzazione può usare le stesso sistema di IDM che è già presente
Supporta la segmentazione dell'utenza grazie alla asserzione di privilegi
Evita la gestione di credenziali agli SP
Il browser può essere ovunque e l'utente non è più legato al campus
SSO permette di muoversi liberamente attraverso molteplici SP, Il login viene fatto una sola volta all'inizio della sessione
Il numero di password che l'utente deve ricordare si riduce
Proxy e vpn non sono più necessari
L'uso di attributi opzionali potrebbe fornire agli utenti più elevati livelli di servizio
La comunità mondiale delle università/ricerca ha standardizzato la definizione degli attributi, di conseguenza anche le implementazioni degli SP possono essere standardizzate
I log file presso l'organizzazione di appartenenza possono fornire informazioni ad un livello di dettaglio che non è accessibile all'SP

Svantaggi

La biblioteca e l'SP devono entrambi supportare l'autenticazione federata. A causa di questo la situazione ibrida è da gestire
SP potrebbe dover supportare diversi metodi di accesso federato (Shib e Athens)
L'autenticazione federata si deve poter applicare a qualsiasi percorso scelto dall'utente per arrivare alla risorsa
L'interfaccia dei discovery service varia tremendamente da un SP ad un altro e questo confonde l'utente
L'utente deve fare login

A questo proposito nel 2011 è uscita una raccomandazione chiamata espresso emanata da NISO, una importante organizzazione per la standardizzazione, punto di riferimento mondiale per biblioteche, editori e fornitori di piattaforme per la distribuzione delle informazioni

La Raccomandazione intende definire un percorso per passare dal meccanismo di riconoscimento corrente basato sull'IP -> all'autenticazione federata.

Si tratta principalmente di fornire agli utenti una percezione consistente del loro percorso di navigazione quando attraversano una moltitudini di siti e situazioni.

Il processo di login federato dovrebbe essere facile e intuitivo e si deve evitare che l'utente abbandoni la sessione di lavoro a causa di una mancata comprensione di quello che sta succedendo.

Nello specifico si raccomanda che:

Gli SP continuino a supportare autenticazioni multiple per il prossimo futuro, cioè IP, proxy e SSO. Nel periodo di transizione occorre accettare che alcune organizzazioni si muovano più velocemente di altre

SP e IDP dovrebbero muoversi rapidamente per ridurre l'utilizzo del controllo di accesso basato su IP. Ci sono infatti molti problemi di sicurezza e questo metodo non è più adeguato ad un ambiente che sta diventando ubiquitario.

Si dovrebbe il più velocemente possibile eliminare l'autenticazione presso l'SP

SP e IDP dovrebbero il più velocemente possibile implementare l'accesso federato standardizzato

SP e biblioteche dovrebbero usufruire delle potenzialità concesse dalla pseudonomizzazione e dal la richiesta di consenso

E infine, importantissimo: SP e IDP dovrebbero migliorare le loro interfacce utente per permettere all'utente una percezione consistente del percorso di navigazione

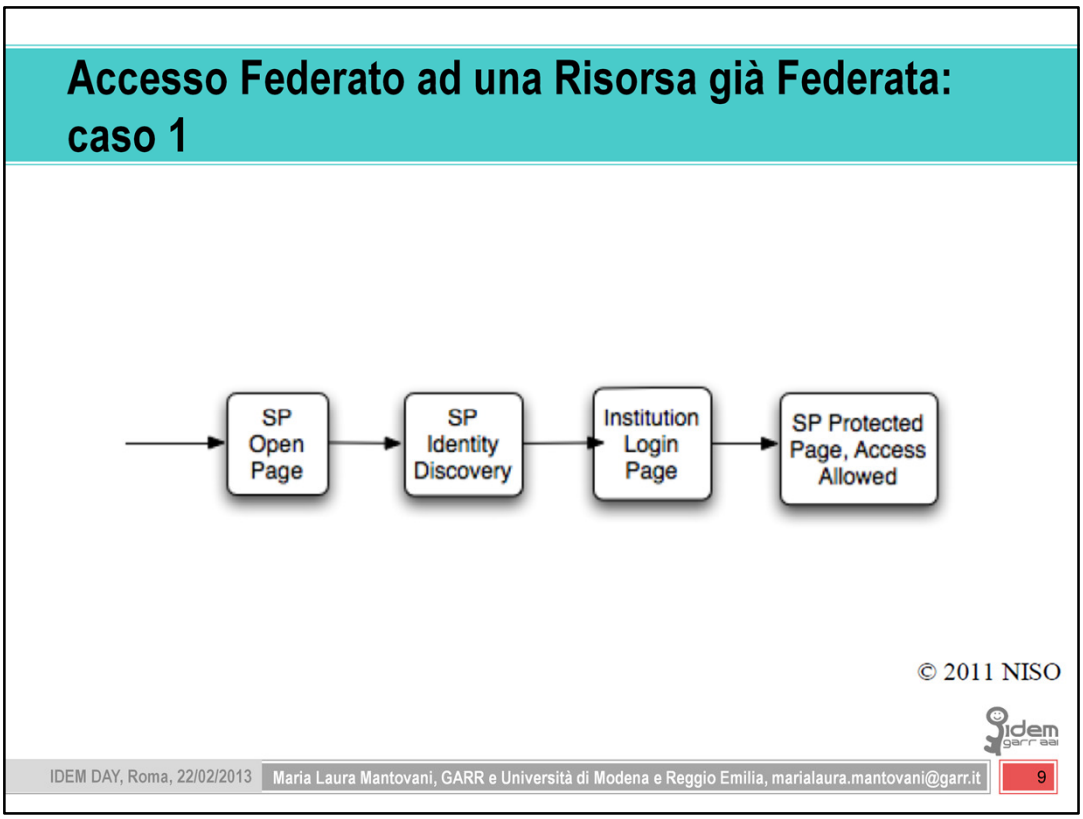
---

[, nello specifico:

SP dovrebbero adottare uno standard nel posizionamento e nell'uso delle parole per

accedere al login dalla pagine pubbliche
SP dovrebbero presentare un identity discovery standard
IDP dovrebbero creare una percezione consistente per l'utente che si muove dall'SP all'IDP all'SP
SP e IDP dovrebbero inserire loghi nei posti appropriati per fornire un feedback visuale che il percorso sta procedendo come atteso
SP dovrebbero offrire uno stesso URL per l'accesso tramite riconoscimento IP e tramite accesso federato

]



Cerchiamo allora di approfondire cosa prescrive la raccomandazione Espresso.
Lavorare in un ambiente ibrido è particolarmente confondente per l'utente. Vari studi riguardanti le problematiche delle interfacce utente hanno focalizzato sui problemi di usabilità e consistenza. Il browser dell'utente potrebbe dover attraversare siti distinti, ma l'utente è alla ricerca di qualche consistenza sulle modalità di identificazione da parte del SP. Nel modello federato l'utente cerca un feedback visuale che gli indichi che la sequenza di pagine che sta attraversando lo sta portando verso la pagina di destinazione desiderata.

Siccome è un metodo di riconoscimento diverso rispetto al riconoscimento in base al'IP, bisogna tuttavia renderlo all'utente il più familiare possibile.

Vediamo un primo esempio di navigazione. Questo è il caso più semplice di accesso ad una risorsa federata e si svolge in 4 passi. Quando un utente conosce l'URL della risorsa a cui vuole accedere e lo scrive direttamente nel suo browser (oppure ha un bookmark). I 4 passi sono:

1. Sito del service provider detto SP Open Page
2. Pagina di scelta dell'organizzazione di appartenenza, detta Identity Discovery
3. Pagina di login dell'organizzazione di appartenenza
4. Pagina dell'SP protetta dall'autenticazione che ora è accessibile perché l'utente si è autenticato

Come deve essere organizzato questo percorso?

Vi presento il MODELLO raccomandato da espresso agli editori ma anche a tutti gli SP e IDP in generale che deve essere adottato per standardizzare il processo di login e facilitare l'utente nella procedura di login federato.

SP Open Page

Presentare un link "Login" posizionato in alto a destra (non ingombrare con altri metodi di autenticazione)

Il link/bottone deve dire solo "Login" (no Shibboleth, no Federation, no Institution, ….)

SP Discovery Page

Il Discovery Service deve essere incorporato nell'SP. Benefici:

- SP può personalizzare la pagina con il proprio logo, colori e grafica, per fornire all'utente un feedback visuale
- Se l'SP è membro di molte federazioni, presentare un unico livello di scelta (non "Scegli prima il tuo paese)
- Evitare di presentare all'utente tutti gli IDP di una federazione
- Permette all'SP di gestire anche le organizzazioni che non sono ancora pronte per Shibboleth

Mostrare le scelte fatte in precedenza

Institution Login Page

Deve riportare nome e possibilmente logo dell'SP da cui è partito l'utente

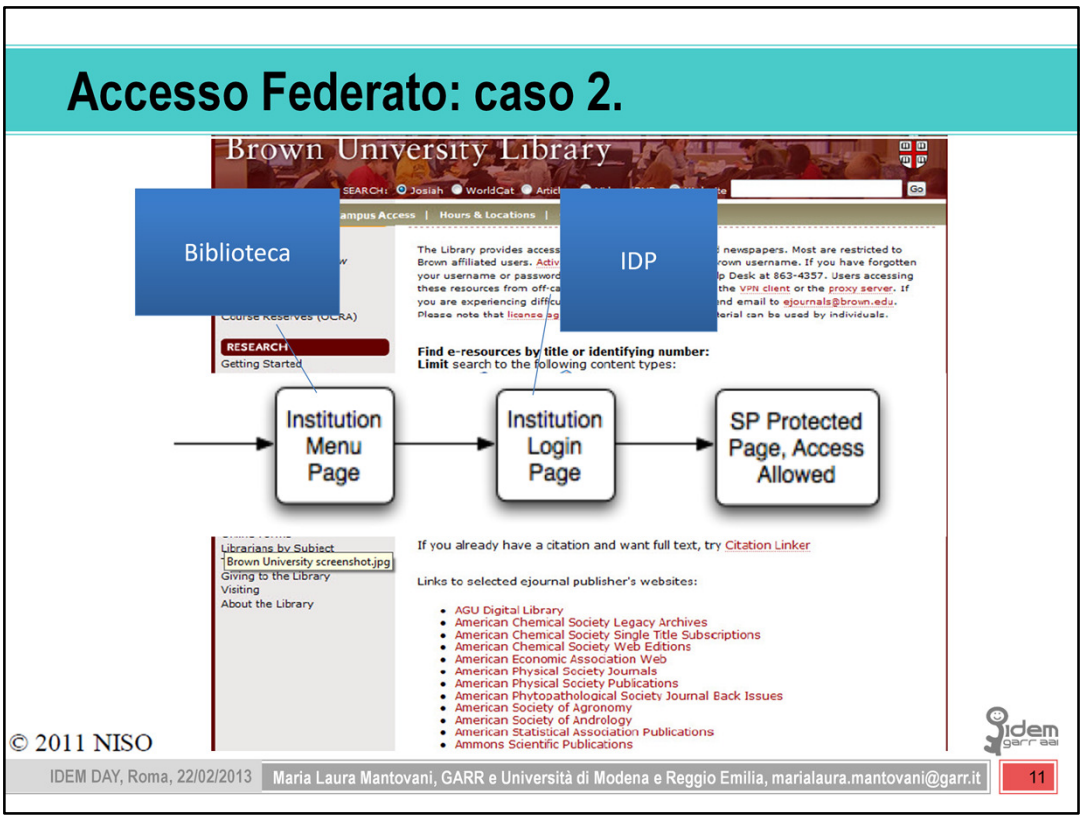
SP protected page

Deve riportare nome e possibilmente logo dell'organizzazione che ha fornito l'accesso

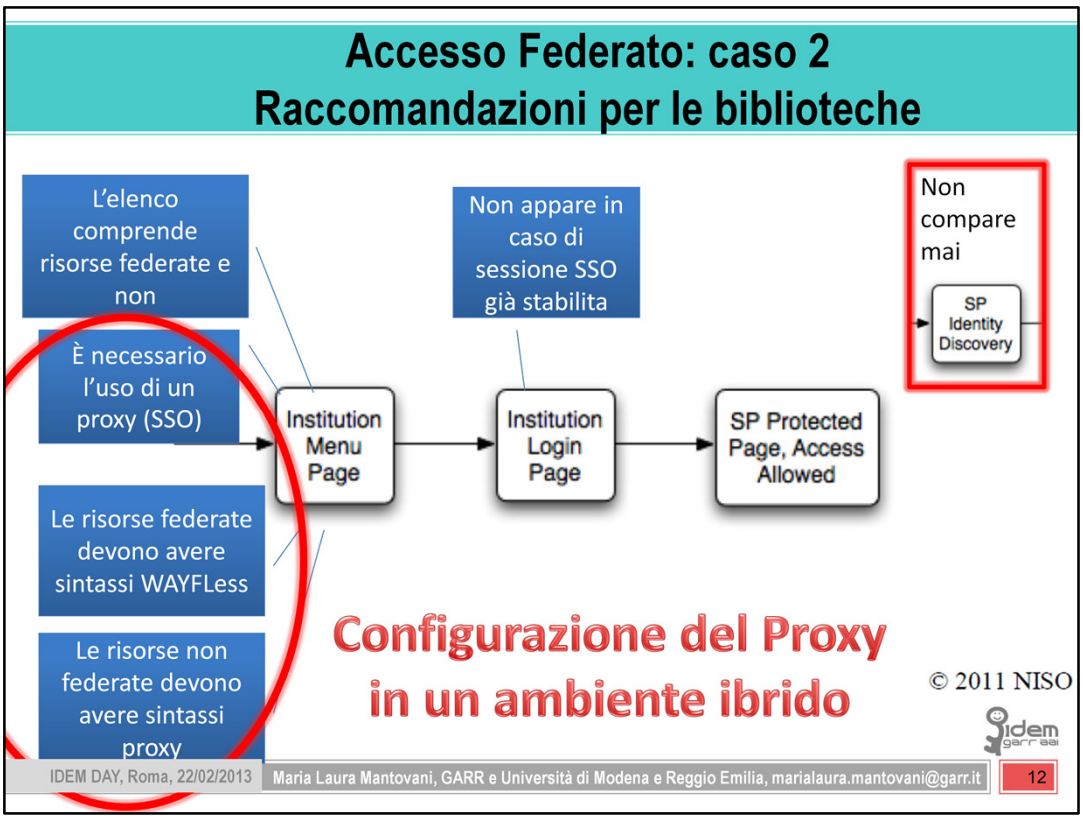
La Open page deve riconoscere se esiste una sessione attiva per l'utente, che può essere:

- Il riconoscimento dell'IP
- Un precedente login SSO

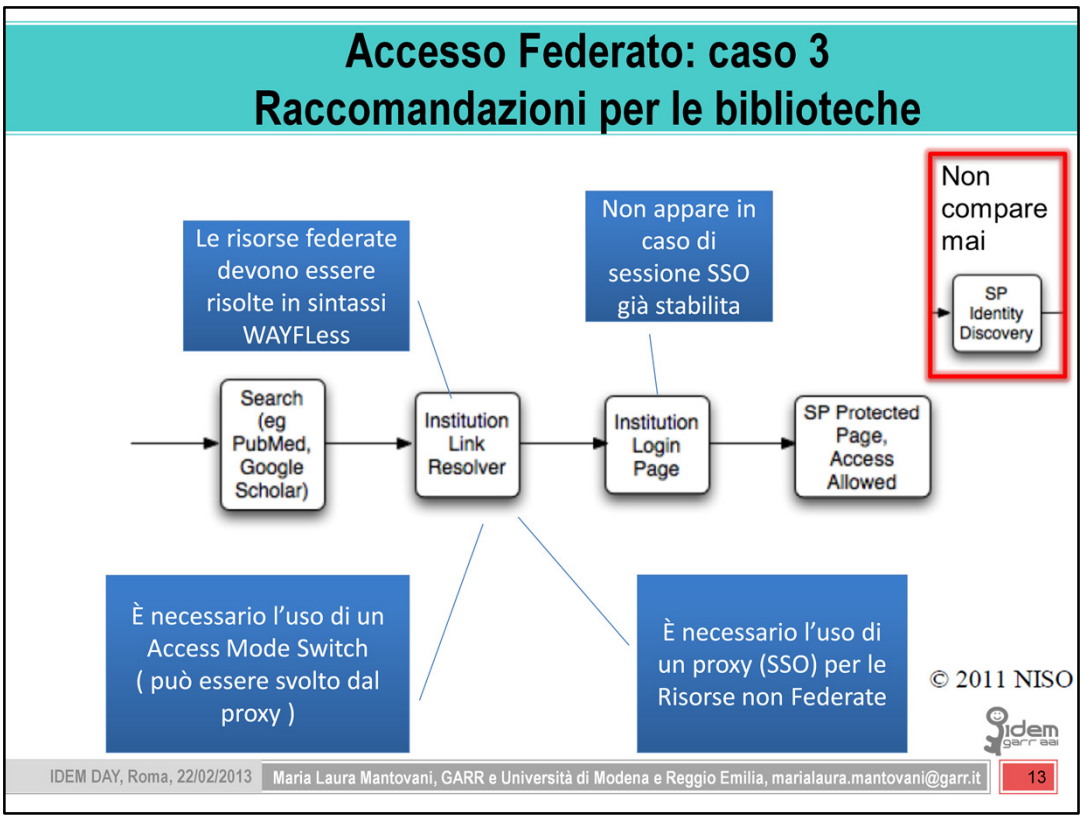
In tal caso , tutti i passaggi da 1 a 3 devono essere evitati e deve essere indicato "Access to XYZ provided via Sample University Library"



La raccomandazione espresso presenta poi 2 specifici percorsi per le biblioteche.
In questo caso la procedura di accesso non è più in 4 passi, ma in 3 passi.
Le biblioteche hanno usualmente un portale che elenca le risorse disponibili e riservate agli utenti della biblioteca.
Questo portale: Institution Menu Page è fornita dalla biblioteca. Qui scelgo la risorsa a cui voglio accedere.
Passo per l'Institution Login Page è fornita dall'IDP
Terzo passo di arrivo alla pagina protetta della risorsa



Ed ecco le raccomandazioni da seguire per questa tioologia di accesso.
Dato che l'elenco nella institution login page comprende risorse federate e non federate, è necessario l'uso di un proxy che supporti il SSO federato.
Nell'elenco le risorse federate devono avere sintassi wayfless. Le risorse non ancora federate devono avere sintassi proxy. In questo modo per l'utente non c'è differenza e il suo percorso è facilitato. In questo percorso non compare mai la pagina di discovery dell'Identity provider, perché è perfettamente superflua. Infatti dato che la pagina iniziale appartiene ad una istituzione, non ci sarà mai la possibilità di scegliere un IDP differente da quello dell'istituzione stessa.



Il terzo caso preso in considerazione dalla raccomandazione espresso è il caso di un utente che vuole accedere ad una risorsa protetta che è stata trovata tramite un motore di ricerca. Oppure tramite l'uso di un discovery tool, oppure l'utente ha un collegamento tramite un riferimento (DOI) oppure un URL elencato da qualche altra risorsa. L'accesso di questo tipo bypassa la Institution Menu Page e anche la pagina di Discovery

In questo caso il percorso è in 4 passi.

1. Search
2. Institution link resolver
3. Institution login page
4. Sp protected page

Anche in questo percorso non compare mai il servizio di discovery dell'IDP (cioè il WAYF), se ne può fare a meno perché abbiamo l'insitution link resolver.
Per fare funzionare correttamente l'institution link resolver con risorse federate e non federate, occorre avere un rewriting proxy che supporti il SSO e configurare opportunamente il proxy perché funzioni come Access Mode Switch. L'access mode switch deve riscrivere gli url delle risorse federate in sintassi wayfless e gli url delle risorse non federate in sintassi proxy. La pagina di login deve apparire solo se non è già stata stabilita precedentemente una sessione di SSO, altrimenti viene saltata. In questo modo l'utente non deve fare click in più del necessario e la sua richiesta arriva direttamente alla pagina protetta, o se necessario gli viene chiesta solo l'autenticazione sull'IDP già predefinito.

Per concludere, dobbiamo accettare questa situazione di eterogeneità in cui ci troviamo, ma
Da una parte dobbiamo spingere perché tutte le risorse attivino il SSO federato
D'altra parte dobbiamo rendere questa situazione eterogenea il più accogliente possibile per l'utente, adottando le interfacce utente indicate dalla raccomandazione espresso ed avvalendoci dell'aiuto che ci puo' dare un rewriting proxy dotato di SSO.
Questo è il modo in cui oggi il bibliotecario è chiamato a facilitare l'accesso alle informazioni a cui l'utente vuole accedere.