Privacy by Design in Federated Identity Management

Interpreting Legal Privacy Requirements for FIM and Comparing Risk Mitigation Models

2015 International Workshop on Privacy Engineering – IWPE'15 - MAY 21, 2015 - SAN JOSE, CA

Rainer Hörbe, presenter
Identinetics GmbH, Austria
rh@identinetics.com

Walter Hötzendorfer, co-author
Centre for Computers and Law
University of Vienna, Austria
walter.hoetzendorfer@univie.ac.at

Overview

FIM usage: why, who, where?

FIM-related privacy risks

Motivation for this project

Approach

Findings

FIM Usage

Why	**Scalability**: registration cost **Interoperability**: attribute semantics, trust policies **Compliance**: Loss of control across many silos
Who	Independent entities with common interests. (Supply chains, government agencies, R&E institutions, enterprise group members, professional networks, markets with roaming agreements.)
Where	eduGAIN, airlines, defense supply chains, government extranets, G2C/G2B services, ..
Edge Cases	Mobile SIM, social networks, centralized (single IDP) federations.

FIM-Related Privacy Risks

Due to FIM:

- **Observability** of behavior by central instances
- **Linkabilty** by introducing common identifiers
- **Impersonation** by Identity/Credential Providers or because of weaknesses in SSO mechanism

Due to the lack of FIM with PbD

- **Linkabilty** by reusing identifying attributes
- **Impersonation** caused by password reuse

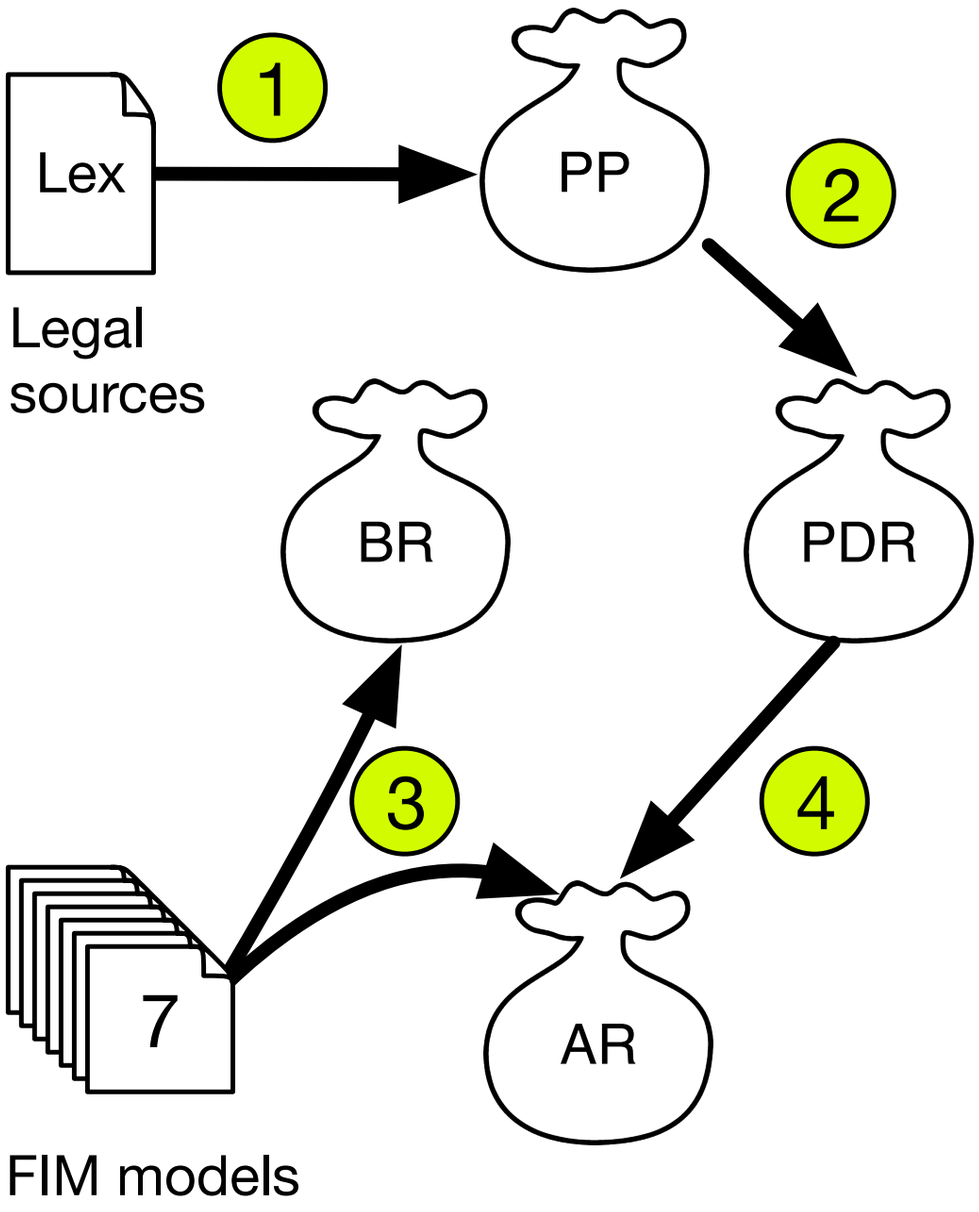
Privacy Risks Unrelated to FIM

Linkability	Identifying contents across services Services integration/large privacy domains
Observability	Device fingerprinting IP-address
Impersonation	Weak endpoint security Poor crypto

Motivation and Scope

- FIM Projects featuring cross-sector federation
  (smart cities, citizen eIDs, B2B across supply chains)
- How to handle the increased privacy risk considering legal requirements, cost, complexity, convenience, feasibility?
- Scope limited on WebSSO use case
  (SAML, OpenID Connect)
- Focus on Observability and Linkability

Approach to Understand Requirements



PP .. Privacy Principles

PDR .. Privacy by Design Requ.

BR .. Business Requirements

AR .. Architectural Requirements

Privacy Principles

PP1 Fairness + lawfulness		↓	↓	↓	
PP2 Final purpose		↓	↓		
PP3 Proportionality	↓	↓	↓		
PP4 Data quality				↓	
PP5 Information security		↓			↓
PP6 Openness + transparency				↓	
PP7 Individual participation				↓	
PP8 Accountability				↓	

Privacy by Design Rules

PDR1 Minimal identification	X				
PDR2 Disclose/need to know		X			
PDR3 Limited Linkability			X		
PDR4 Transparency + user control				X	
PDR5 Information security					X

Privacy by Design Rules

PDR1 Minimal identification	↓	↓					↓	
PDR2 Disclose/need to know	↓		↓			↓		↓
PDR3 Limited Linkability		↓						
PDR4 Transparency + user control				↓				
PDR5 Information security								↓

Architectural Requirements

X								AR1 Limited observability
	X							AR2 Limited linkability
		X						AR3 No unauthorized aggregation
			X					AR4 Constrained linking
				X				AR5 Consent handling
					X			AR6 No supreme instance
						X		AR7 Minimal attribute release
							X	AR8 Unique identification

Existing Implementations

Business Requirements

BR1 Allow limited linking				↑				

Privacy by Design Rules

PDR1 Minimal identification	↓	↓					↓	
PDR2 Disclose/need to know	↓		↓			↓		↓
PDR3 Limited Linkability		↓						
PDR4 Transparency + user control				↓				
PDR5 Information security								↓

Architectural Requirements

X								AR1 Limited observability
	X							AR2 Limited linkability
		X						AR3 No unauthorized aggregation
			X					AR4 Constrained linking
				X				AR5 Consent handling
					X			AR6 No supreme instance
						X		AR7 Minimal attribute release
							X	AR8 Unique identification

Existing Implementations

Business Requirements

BR1 Allow limited linking				↑				

The Problem Children

AR1 Limited observability

AR2 Limited linkability

AR3 No unauthorized aggregation

AR4 Constrained linking

AR5 Consent handling

AR6 No supreme instance

AR7 Minimized attribute release

AR8 Unique identification

Organizational Controls
Attribute-Based Credentials
Late Binding
Proxy Pool
User-based IdPs
Constrained Logging Proxy
Blind Proxy

Models for Limited Observability:
(2) Attribute-Based Credentials

ABCs provide assertions to the RP without the IdP knowing the actual RPs.

Pro: Strong technical control.

Con: (a) No implementation in mainstream products; lack of deployment profiles for SAML or OpenID Connect; (b) IdP business model; (c) performance; (d) Increased complexity.

Models for Limited Observability:

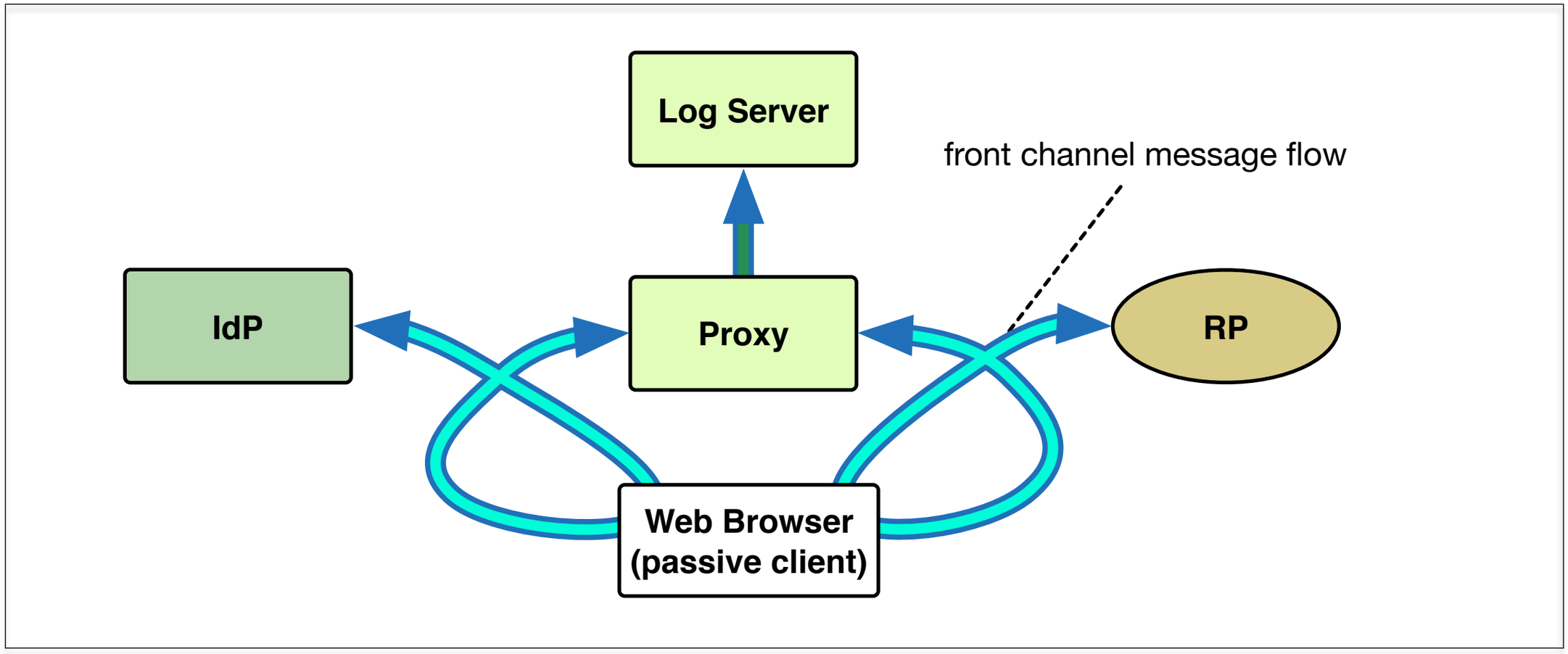
(3) Late Binding/Federated Credentials

Credential-only federation (CSPs with brokers, or U2F tokens) rely on the separation between credential service assurance and identity assurance. Attributes are not released by the IdP, but obtained by the RP.

Pro: Straightforward architecture that goes well with existing technology based on common SAML profiles. Credential providers have only a minor privacy risk.

Con: (a) Less business value because attributes are collected per RP;
(b) Identifying attributes like name, residential and e-mail addresses could enable linking.

Models for Limited Observability:
(6) Constrained Logging Proxy

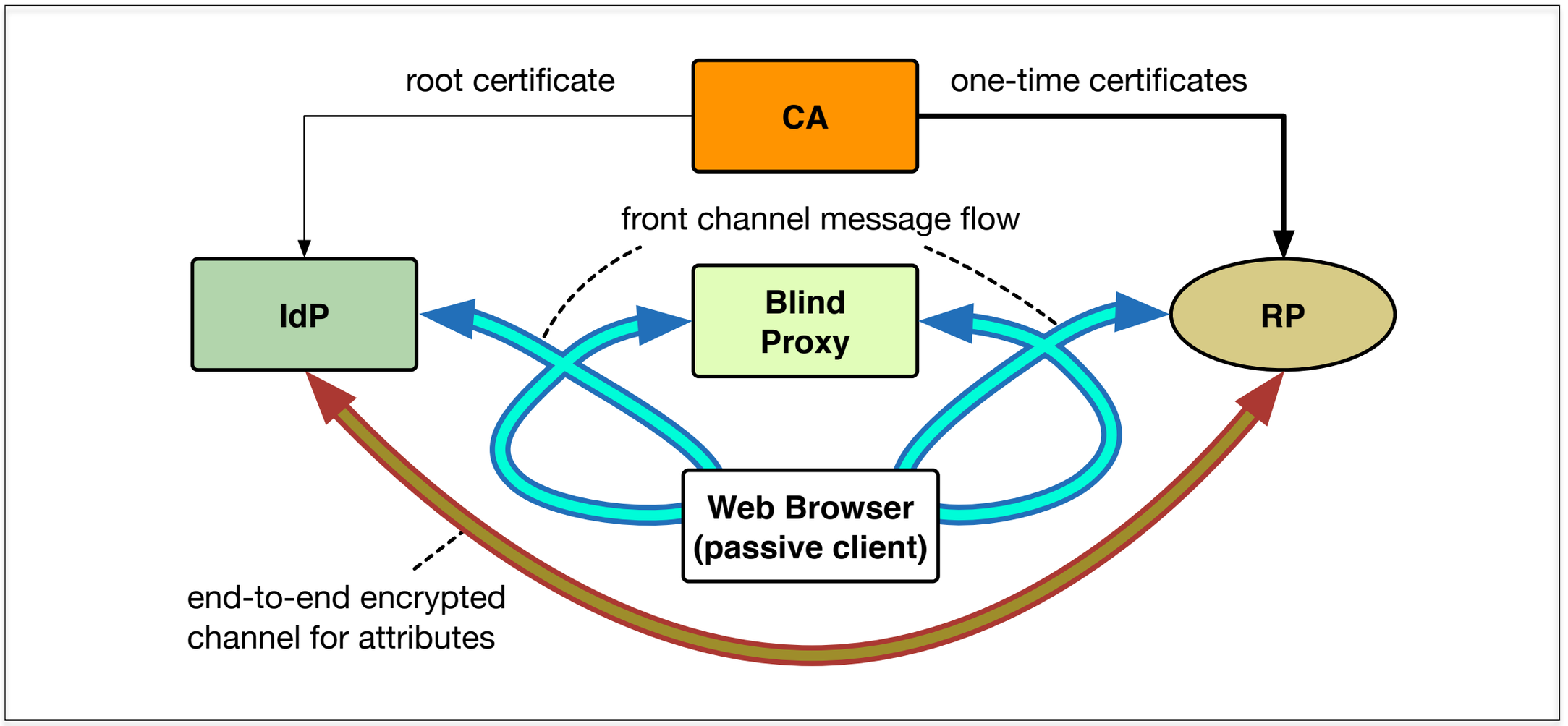


The proxy stores log files only in a separate, well-protected system for a very limited time.

Pro: Has been implemented without changes to FIM protocols.

Con: While an adversary could cause only limited damage with a single data breach, a complete take-over of the proxy would compromise the privacy goal.

Models for Limited Observability:
(7) Blind Proxy



Pro: It proposes reasonably strong technical control, works with any credential technology and is fairly easy to fit into hub-and-spoke federations.

Con: (a) Requires (small) extension to existing SAML and OIDC implementations. (b) It requires RPs to participate in a considerably large anonymity set.

The Problem Children

AR1 Limited observability

AR2 Limited linkability

AR3 No unauthorized aggregation

AR4 Constrained linking

AR5 Consent handling

AR6 No supreme instance

AR7 Minimized attribute release

AR8 Unique identification

Approaches for Limited Linkability Between Privacy Domains

- Unique Identifiers limited in scope:
  - Pairwise identifiers (IDP - RP)
  - Group or sector-specific identifiers
- Proxy attributes for identifying attributes:
  - Blind „reverse proxy“ for e-mail and jabber
  - User-selected pseudonyms for display names
  - Virtual credit cards, crypto-currencies for payments
  - PO-boxes etc. for physical shipment

The Problem Children

AR1 Limited observability

AR2 Limited linkability

AR3 No unauthorized aggregation

AR4 Constrained linking

AR5 Consent handling

AR6 No supreme instance

AR7 Minimized attribute release

AR8 Unique identification

Approaches for Constrained Linking (Between Privacy Domains)

- Types of link constraints:
  - A group of privacy domains (>=2)
  - By direction (i.e. unidirectional)
  - Temporal (e.g. until expiry or revocation)
- Examples:
  - Austrian eID with sector-specific identifiers encrypted for another sector's target application
  - Mediated links in a blind proxy model: All access via proxy is encrypted end-to-end, except the identifier that is mapped by the proxy.

Conclusions

- Increased privacy risks introduced by FIM can be mitigated with technical controls.
- Effort to implement controls for limited observability varies with the strength of the controls.
- Limited likability with pairwise identifiers is current practice. However, identifying attributes are left out of the equation. There is room for improvement with moderate effort.

Blind Proxy Profiles & Implementations

- SAML PEFIM Profile
  https://kantarainitiative.org/confluence/x/-wIxB
- PEFIM Proxy reference implementation
  http://github.com/its-dirg/pefim-proxy_docker/
- PEFIM IDP & SP implementations
  - PySAML2
    https://github.com/its-dirg/pefim_sp
    https://github.com/its-dirg/pefim_idp
  - Shibboleth
    Will be available soon at shibboleth.net
  - OpenAM
    on request from cryptas.com