



CORONA
EMERGENCY
LECTURE

Security of Distributed Software

Prof. Dr.-Ing. Martin Gaedke

Chemnitz University of Technology

Department of Computer Science

Professorship of Distributed and Self-organizing
Systems

<http://vsr.informatik.tu-chemnitz.de>



Part III

TRUSTWORTHY SOFTWARE ENGINEERING



Trustworthy Software Engineering

- Trustworthy Software
 - In <http://cordis.europa.eu/fp7/ict/security/docs/ict-wp0910.pdf> defined as
 - Trustworthiness can be seen as software and infrastructure that is secure, reliable and resilient to attacks and operational failures; guaranteeing quality of service; protecting user data; ensuring privacy and providing usable and trusted tools to support the user in his/her security management.
 - Trustworthiness needs to be considered from the outset rather than being addressed as add-on feature.
- So, we focus on: Identity & Security by Design (SBD)
 - Who is it for?
 - Why does it matter?
 - What is it all about?
 - Where does it apply?
 - When to apply?
 - How to apply?



Chapter 1

IDENTITY



Introduction

■ Fundamental question

- Internet as a Danger Zone in terms of Identity
- What exactly needs to be protected?
- What should one orient towards?
- Which data is exceptionally worthy of protection?

■ Security vs. Identity

- For starters: Keynote by Dick Hardt at WWW 2007 on „Identity 2.0“
- Speech on Identity by Kim Cameron [<http://www.identityblog.com>]



Video

- Keynote by Dick Hardt
at WWW 2007 on "Identity 2.0"
 - <https://www.youtube.com/watch?v=RrpajcAgR1E>

Problem

- **Kim Cameron:** "The Internet was built without a way to know who and what you are connecting to."
- **Initial situation:**
 - Internet Services are left on their own
 - Must provide security → isolated identity solutions
 - Criminalization of Internet
 - Leads to loss of Internet's credibility, for example, drawback for e-businesses
 - Identity layers are complex
 - Successful attempts, such as SSL and Kerberos – however, overall too many different scenarios are required, so agreement is difficult
- **Possible solution: Identity Metasystem**
 - Such a system provides confidential support to ensure who is connecting to whom/what on the Internet
 - Many questions: Who holds the data? Who trusts whom? What scales? How does one realize openness to new developments that do not yet exist?



Identity

- Identity description
 - Not simple – there exist numerous attempts and different forms, see Wikipedia
 - Lecture is based on Kim Cameron's definition
 - Interesting trends: FOAF and Semantic Web
- **Definition Identity** – Digital identity is a set of *claims*, which are made by a *digital Subject* about self or other subjects.
 - **Definition Digital Subject** – person or thing (referred or real) in a digital realm that is described or with which one is dealing
 - “with which one is dealing” – often in the context with request/response model
 - Example digital subjects: real persons, devices, resources, rules/policies and relationships between digital subjects
 - Discussions of the ‘subject’ term extend into the philosophy (Oxford English Dictionary for Subject: “central substance or core of a thing as opposed to its attributes.”)
 - See Kim Cameron on Entity, Thing und Subject
 - **Definition Claim** – Claim suggests that something is true, typically something that seems to be controversial or questionable.
 - Remark: Claim is a relationship between a certain instance, a digital subject and an identity attribute



Understanding Identity

- We must be able to **structure our understanding** of digital identity
 - We need a way to avoid returning to the **Empty Page** every time we talk about digital identity
 - We need to inform peoples' thinking by teasing apart the factors and dynamics explaining the successes and failures of identity systems since the 1970s
 - We need to develop hypotheses – resulting from observation – that are testable and can be disproved
 - Our goals must be pragmatic, bounding our inquiry, with the aim of defining the characteristics of an unifying identity metasystem
 - The Laws of Identity offer a “good way” to express this thought
 - Beyond mere conversation, the Blogosphere offers us **a crucible**. The concept has been to employ this crucible to *harden and deepen the laws*.
- Identity and Subject and Claim etc....
- These definitions embrace Kerberos, X.509, SAML. They take this problem of the evaluation of the usefulness of a digital identity up to a higher level in the systems sense of multiple layers. These definitions separate the layer of where stuff is communicated from the layer where evaluations are done – a very important step forward.



Laws of Identity

- **1. User Control and Consent**
- **2. Minimal Disclosure for a Constrained Use**
- **3. Justifiable Parties**
- **4. Directed Identity**
- **5. Pluralism of Operators and Technologies**
- **6. Human Integration**
- **7. Consistent Experience Across Contexts**



Laws of Identity (Law 1)

- **1. User Control and Consent**
- *Digital identity systems must only reveal information identifying a user with the user's consent*
 - Systems need to appeal in their convenience and simplicity
 - Constantly care about users' confidence
 - Requires a holistic commitment
 - User must be central to control with respect to which identities are used and which data is made public
 - System must protect from deception (for example, web-site location and misuse)
 - System must inform the user of possible consequences upon certain action (data sharing, login, etc.)
 - The holistic approach must be used as a paradigm in all contexts (e.g. when logging into a company or a private blog it should always be clear that the user consents to the release of certain data)



Laws of Identity (Law 2)

- **2. Minimal Disclosure for a Constrained Use**
- *The solution that discloses the least identifying information and best limits its use is the most stable long term solution*
 - One should assume that data/information violations are unavoidable
 - To reduce risks, information use should be checked with respect to 2 strategies: “must be obtained” or “must be saved”
 - Less information implies less value implies less risk
 - “As little as possible identification information” means:
 - Reduction of linkable information
 - Use of claim-transformations. See 01.01.1950 vs. Over 18?
 - Avoid unnecessary information storage for “possible future” use (Why should a credit card be stored by the shop?)
 - This law is closely related to information disasters



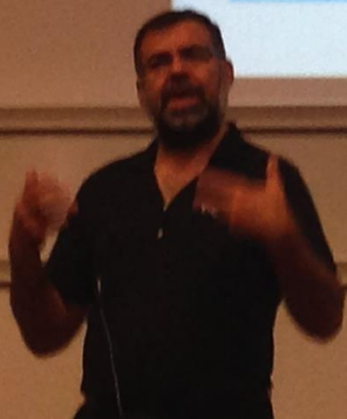


Regions from Pictures

Hover over a tag region to see images for that tag.



Current Tag: barcelona



Y! Risks of Privacy: AOL Query Logs Release Incident

A Face Is Exposed for AOL Searcher No. 4417749,

By MICHAEL BARBARO and TOM ZELLER Jr.

The New York Times, Aug 9 2006



No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men".

Other queries: "landscapers in Lilburn, Ga," several people with the last name Arnold and "homes sold in shadow lake subdivision gwinnett county georgia."

Data trail led to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three dogs.

Laws of Identity (Law 3)

■ 3. Justifiable Parties

- *Digital identity systems must limit disclosure of identifying information to parties having a necessary and justifiable place in a given identity relationship*
 - User has to have a clear understanding whom the information is/will be exchanged with
 - System itself may not draw conclusions about relationships between subject and parties
 - Example Microsoft Passport – is very successful if I'd like to log into MSN, but why should Passport know if I log into eBay or Google
 - In which situations are regulatory identities required?
 - Same holds for intermediaries (what should they know to achieve their goal)
 - All participants must submit statements of how the information will be used



Laws of Identity (Law 4)

■ 4. Directed Identity

- *A unifying identity metasystem must support both "omni-directional" identifiers for public entities and "unidirectional" identifiers for private entities*
 - Digital Identity should always be viewed in the context of another Identity or a set of Identities
 - OMNI-DIRECTIONAL: Public entities require "beacons" (publicly known Identifier or URI)
 - Example: Web sites (URLs) or public devices
 - UNI-DIRECTIONAL: Private entities (people) require an ability not to be turned into a beacon
 - They require a unidirectional identifier, which can be used in combination with a trusted beacon (no correlation, e.g. user-bank interaction)
 - Negative examples: Bluetooth and RFID, partially WLAN



Laws of Identity (Law 5)

- **5. Pluralism of Operators and Technologies**
- *A unifying identity metasystem must channel and enable the inter-working of multiple identity technologies run by multiple identity providers*
 - System may be ideal with respect to one characteristic, but not with respect to another
 - Example: Authority vs Employer vs Individual (as customer or simply person)
 - Old and new technologies must be used and can co-exist – identity system must not be in competition with technology, but must use it
 - Technologies may have more growth than others (identity ecology)



Laws of Identity (Law 6)

■ 6. Human Integration

- *A unifying identity metasystem must define the human user as a component integrated through protected and unambiguous human-machine communications*
 - Communication can be completely secure, but what about the last two meters (off the screen and into the eyes of the viewer) – Does the user really know who it is he's communication with?
 - Phishing attacks are a good example
 - Protocol for use of safety issues has to become a ceremony, absolutely predictable and controlled
 - Example: communication in the cockpit (channel 9 on United Airlines)
 - How does one achieve such reliability?



Laws of Identity (Law 7)

- **7. Consistent Experience Across Contexts**
- *A unifying identity metasystem must provide a simple consistent experience while enabling separation of contexts through multiple operators and technologies*
 - Simplicity and clarity are the main goal – Identities have to be used in a similar fashion to all other things on the desktop
 - User must be able to see, verify, add and remove Identities
 - Which type of Identity is acceptable in which context?
 - Properties of such candidates are defined by the using parties
 - Users must be able to recover the Identity in the given context and understand which information is associated with it
 - Person (human/legal) could possibly accept different types of Identities
 - User must be able to choose the best Identity in his opinion



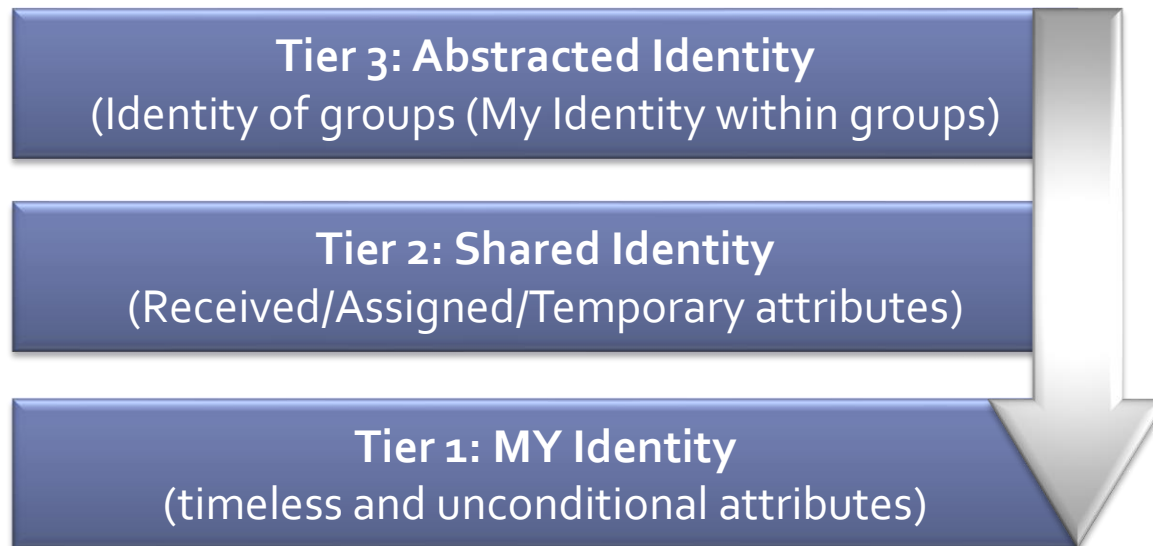
Chapter 2

IDENTITY IN THE LIGHT OF PRIVACY, SECURITY AND TRUST



Identity in Context

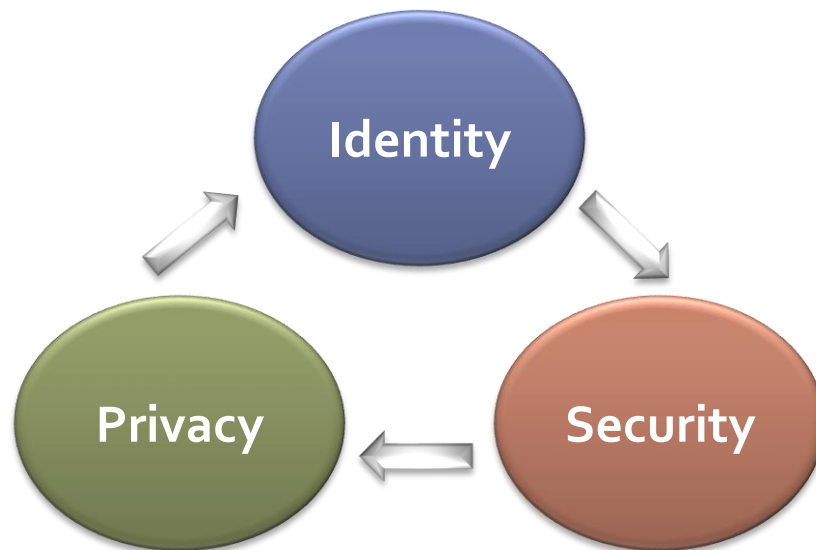
- 7 Laws of Identity define requirements of dealing with Identities
 - First focus on conceptual / basic understanding
- Identity in global context has to comply with different levels
 - Layered approach of Identity management



Based on
"Digital Identity",
Phillip J. Windley
and Ping Identity Corp.

Identity - Security - Privacy

- Identity (in a digital setting) is often “only” closely linked to security - Identity is more!
 - Security – Protect data from unauthorized access, removal, tampering
 - Privacy – Protect attributes, preferences, etc., which are associated with Identity, against unnecessary use by subject
 - Identity is in relation to others → Attributes realize trust relationships



Aus "Digital Identity", Phillip J. Windley

Identity & Trust

■ Trust [Source: Wikipedia]

- In a social context, **trust** has several connotations. Definitions of trust typically refer to a situation characterised by the following aspects: One party (trustor) is willing to rely on the actions of another party (trustee); the situation is directed to the future. In addition, the trustor (voluntarily or forcedly) abandons control over the actions performed by the trustee. As a consequence, the trustor is uncertain about the outcome of the other's actions; he can only develop and evaluate expectations. The uncertainty involves the risk of failure or harm to the trustor if the trustee will not behave as desired.
- **Trust – Conviction and belief in the sincerity, honesty and good intentions of another party with respect to a risk-prone action.**

Trust Examples

- Shopping with a Credit Card – Which trust relationships and risks exist?
 - Identity and Credit Card company
 - Identity and service
 - Identity/service and card reader
 - Identity and cash register
 - Identity/service and money
 - ...
- → Trust is always associated with risk
- Trust is something one connects to a person
 - One can not enforce trust for another person ("Give me your trust")

Trust Properties (1)

- Trust is rarely transitive
 - Example: I trust Anne's taste in music, she, in turn, trusts Peter's – therefore, I would, possibly, trust Peter in selecting music for my Birthday party.
- Trust can not be shared
 - Example: A trusts B, A trusts C does not imply that B and C trust each other.
- Trust is not symmetric
 - Example: If I trust you, you don't necessarily trust me in return.
- Trustworthiness can not be self-declared
 - "Trust me!"



Trust Properties (2)

- Trust is a value closely related to evidence
 - I buy a computer, which is 10 EUR more expensive, since I trust the brand.
 - Computer allows access upon login, since the provided evidence (login/password) serve as proof.
 - What are:
HTTPS and certificates in this context?
Do you also know the secret?
- Trust is hard to quantify
 - I trust Anna more than Peter – What does that mean?
 - In business context trust can be evaluated against risks (given obvious risk levels)
 - Otherwise, contract as a basis: Analysis is required, risks are evaluated and, thereby, contractual relationships are defined. Leads to Service Level Agreements (SLA) between providers and users.

Trust Properties (3)

- Trust by reputation
 - Trust in a person can develop from other people's statements about him/her (Communities of Trust)
 - Examples:
 - All security experts advise caution when traveling in the following countries.
 - eBay: One buys a product from a handler he doesn't know, but which has a high reputation (good reviews)

Identity & Privacy (1)

- Privacy is an important and complicated topic (tightly coupled with data protection)
- Identity und Privacy are closely related
 - What does privacy mean for a person?
 - Generally: Private data shouldn't become public
 - However, often: Private data disclosure is ok if it yields considerable benefits
 - Privacy must be observed in context
 - Example: Discount systems: Provide us your address and date of birth and we'll give you a 15% discount
 - What does Privacy have to do with trust?

Identity & Privacy (2)

- Privacy is partially legally regulated
 - Example: Introduction of an electronic phonebook within the company GM (lasted 2 years, due to statutory regulations)
 - Examples
 - Federal Data Protection Act (FDPA / BDSG). Excerpt: "The purpose of this Act is to protect individuals from being compromised in his personal rights through handling of their personal data..."
 - European Data Protection Directive (for EU companies with respect to customer data storage and utilization)
 - Patriot Act (for financial institutions in the USA, enforcing the collection of customer data and cross-checking with government agencies as anti-terrorism measures)

Identity & Privacy (3)

- Conclusion in legal context: Own applications and systems must take Identity and Privacy into account. (see Laws of Identity)
 - Embed the concepts of Identity und Privacy in design
 - Use of Identity und Privacy-relevant information must be comprehensible, verifiable and reportable at any point in time
 - Identity Management System or an Identity Metasystem must be able to answer questions on Identity-Privacy terms
 - Legal requirements force system operators to testify on privacy policy
 - Example: Web-Shop sends Cookies to customers
 - What should the privacy policy say? Examples: We use Cookies. The shopping cart will not work without them.

Identity & Privacy (4)

- Privacy principle – respect privacy
 - Accountability
 - Identifying purposes
 - Requirement of affected person's consent
 - Minimal privacy data collection (time limit)
 - Limitation of use
 - Data collection accuracy
 - Protection
 - Access to personal data (to the owner)
 - Comprehensible regulations

[Based on "Digital Identity", P.J. Windley]



Chapter 7

IDENTITY MANAGEMENT SYSTEMS



Introduction

- **What is needed for Identity implementation?**
 - Some kind of *Identity Metasystem* → contains 3 certain roles (can be more)
- **Identity provider**
 - Person or an organization, which creates digital identities, either for themselves or on behalf...
 - Examples:
 - Online-Shop could create identities for customers
 - Authorities provide identities for their employees
 - Company handling token creation for on-line age verification
- **Relying Party (human/legal person)**
 - Person or organization, which requires digital identity before allowing entry/access
 - Example:
 - Users willing to revoke a contract – the Relying Party defines which claims are required in order to execute cancellation, as well as which formats and credentials are accepted
- **Digital subject**
 - Individual or entity for which claims are made



Definition Identity Management

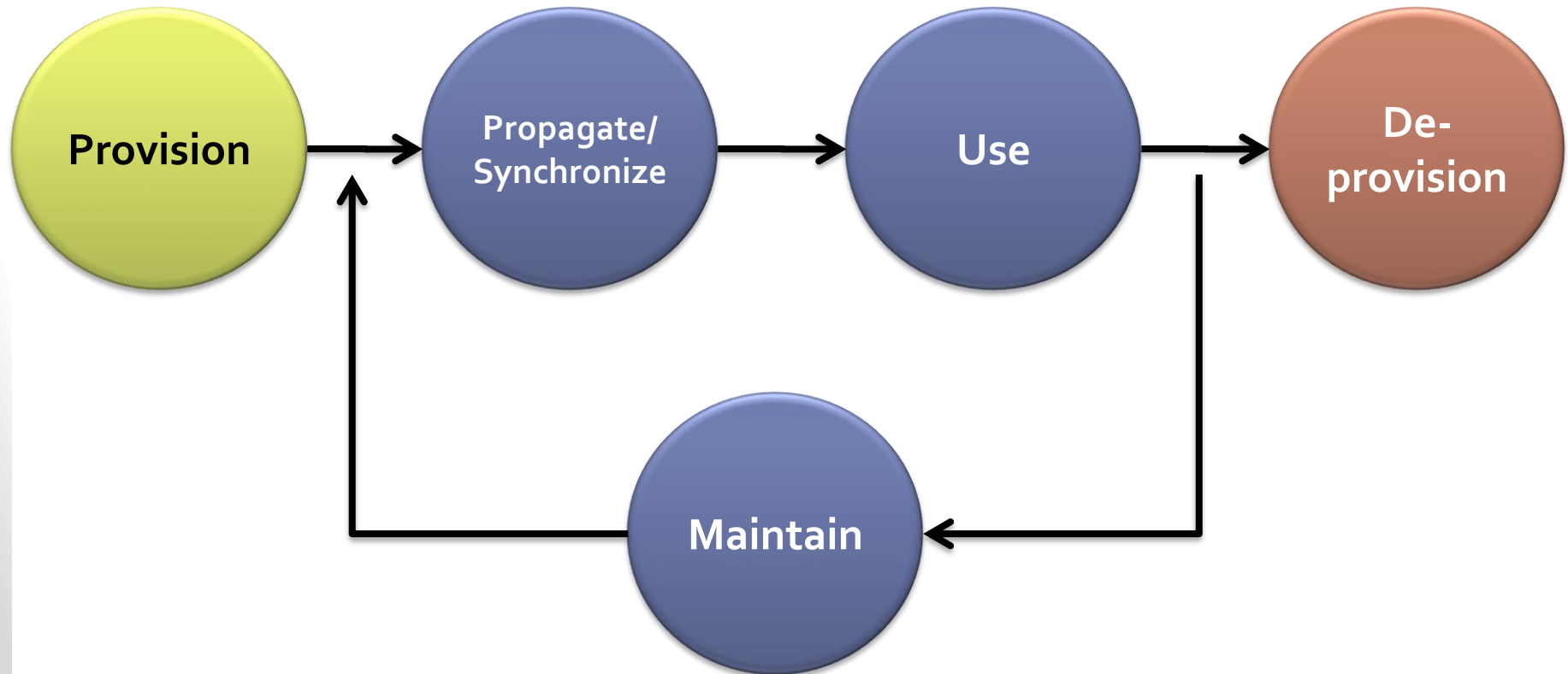
Definition:

“Identity management is the set of processes, tools and social contracts surrounding the creation, maintenance and termination of a digital Identity for people or, more generally, for systems and services to enable secure access to an expanding set of systems and applications.”

(HP Whitepaper, Identity Management: the drive to federation, 2003)



Identity Management Lifecycle



(In Anlehnung an Phillip. J Windly, Digital Identity)

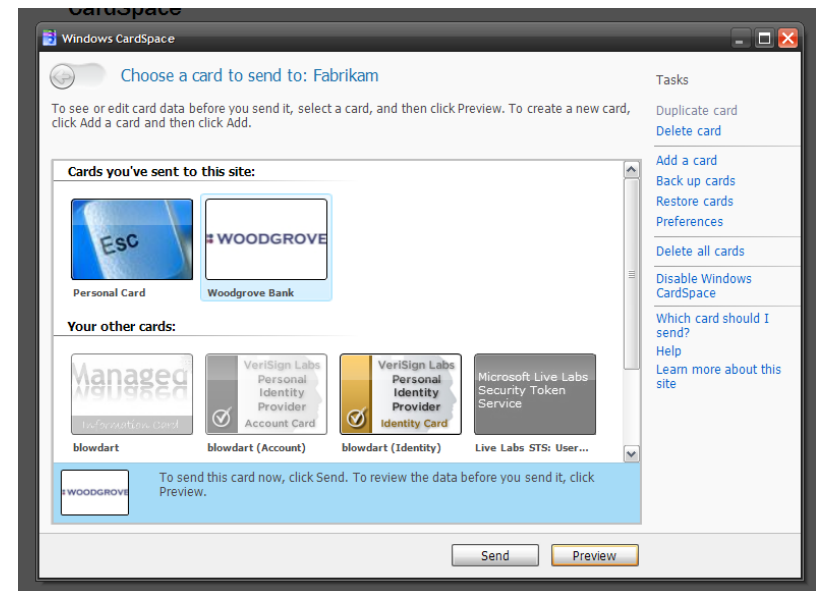
Identity Management Levels

- Personal identity management
- Organization-related identity management
- Federated identity management



Personal Identity Management

- Entity-perspective
 - Management of different identities (different accounts for different systems)
 - Management and control of which information is provided to a service (z.B. Email, phone number etc.)
- Exemplary products
 - MS Passport
 - MS Cardspace



Org.-related Identity Management

- Organizational perspective
 - Management of identities of an organization
 - Different services of an organization are provided with and updated by identity information.
 - Traceability of data flows and data accesses
 - Management of privileges and roles within the organization
 - Definition of organization's policies as to the entities i.e. which data can be accessed.
- Exemplary products
 - SUN Identity Management Suite (SUN Identity Manager)
 - Microsoft Identity Integration Server
 - IBM Tivoli Identity Manager



Organizational Questions

- What kind of identities exist?
- Which attributes are required to describe an identity?
- What are the sources and goals of individual attributes?
- Who is legally responsible for which attribute?
- Which values should individual attributes possess?
- Who approves distribution paths?
- Who can access which attributes?



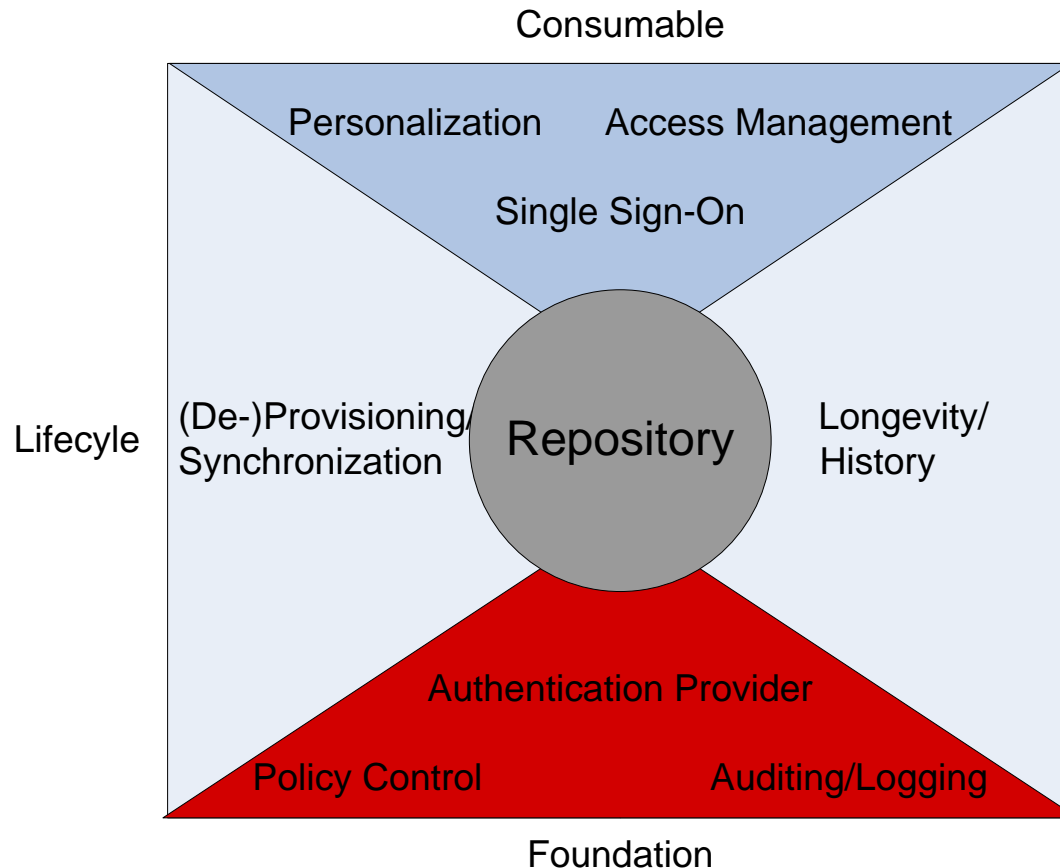
Anticipated Benefits of IDM systems

- Reduced management overhead
- Better optimization/automatization of business processes
- Reduced time required for providing a new employee with access rights to resources
- Reduced risk of a former employee accessing resources
- Policy and legal requirement compliance support (privacy)
- Data consistency (data matching, modification checks, ...)
- Standard interfaces (APIs, standards ...) to data/services/resources



Components of IDM systems

„The focus of identity management is on *user provisioning* — the creation, maintenance, and termination of user accounts and management of credentials in support of *authentication* and *access control*.“ (Hurwitz Group, 2001)



IDM Basic Components (1)

■ Repository

- Repository represents the core component for many identity management systems
- It is a **logical data storage** (i.e. database, directory service), in which identity information, guidelines and other organization information can be stored

■ Propagation

- Depending on the system in use, an identity entry could need to be transferred from the current repository to another one

■ Authentication Provider/Identity Provider

- Is responsible for primary identity authentication
- Often issues a credential, which can be used for further authentication and authorization (z.B. SAML Token)
- Provides multiple interfaces (z.B. LDAP, Kerberos), by means of which service can perform authentication

IDM Basic Components (2)

■ Policy Control

- Policy control governs rules of information usage, disclosure and logging
- Authorization policies determine which identity can access and manipulate which information
- Policy control monitors the defined guidelines, creates events to be audited and signalled according to certain rules (for example, security warnings)

■ Auditing, Monitoring

- Auditing provides necessary mechanisms for information detection and storage
- That information normally contains access protocols and data operations (specifically in the repository)
- It forms a basis for tracking whether the policies are being adhered to and is used for subsequent security checks



IDM Lifecycle Components

- **(De-)Provisioning and propagation/synchronization**
 - Applies automation of all the procedures and tools to manage the identity lifecycle.
 - This Lifecycle is split into initial provisioning, synchronization and de-provisioning phases.
 - In the initial provisioning phase the according service is supplied with the necessary identity information such that the new identity can use the service (provisioning process).
 - In the synchronization phase identity information is updated and compared between services (synchronization and propagation process).
 - In the de-provisioning phase all the identity information is removed (de-provisioning process).
- **History, Longevity**
 - History and longevity tools create historical records, by means of which one can examine evolution of an identity overtime (i.e. creation, activation, locking, new status, removal).
 - These components provide means for such activities as investigating whether or not a certain identity exists in the system and which changes it underwent.

IDM Usage Components

■ Single Sign-on

- Single Sign-on enables an identity to perform its initial authentication and access numerous services and data without further re-authentication.
- Initial authentication is typically performed by an associated Identity Provider, which issues a credential.
- That Credential is then used to authenticate to other systems.

■ Personalization

- Personalization and preference management tools provide the identity an ability to set up individual settings for applications/services bound to that identity.

■ Access Management

- Similar to policy control
- Identity can define policies as to which identity can access/modify which information.



Federated Identity Management

- Federation perspective
 - Association of organizational units, organizations or even nations
 - Shared use of resources and services of Federation partners
 - Cross-organizational business processes within the Federation
 - Modeling and definition of trust relationships
 - Federative services are then made available according to the defined trust relationships providing ease of access to resources/data (i.e. Single Sign-on)

- Exemplary products/projects/approaches
 - Liberty Alliance Projekt (SAML 2.0)
 - WS-Federation specification
 - SUN Identity Management Suite (SUN Federation/Access Manager)
 - Ping-ID, PingFederate
 - Shibboleth
 - FOAF+SSL



Chapter 8

MODELLING TRUSTWORTHY SYSTEMS



Introduction

- What is meant by “Federation”?

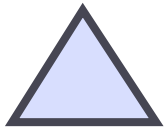
“Federation is an association of independent organizational units, which have a trust relationship.”
- Among the latest developements in the field of IdM.
- Is driven both by the state and industry
 - Common and simplified resource access
 - Cmplx problems/business processes and a high level of specialization require cooperation.
 - Harmonization of business pocesses
 - Cost savings with respect to administration and resource use
- Frequently used technologies
 - SAML (Security Assertion Markup Language)
 - XML (Schema, Encryption, Signature etc.)
 - Web Service interfaces

WAM (WebComposition Architecture Model)

- WebComposition Architecture Model
 - Developed in 2005 by Gaedke and Meinecke
 - Uses different layered models, e.g. Federation model
 - Uses UML notation in combination with OCL
 - Provides overview of Web Architecture, Federation and other Realm related relationships
- Six Core Entities
 - Connected by Bindings (described in profiles)
 - Entities are assigned to Zones/Realms
 - Connections with *Labels* are provided to complement the properties in addition (detailed) to graphical notation
 - Labels und their descriptions are stored in separate databases and support reuse in further projects



WAM Core Entities (1)



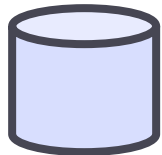
■ Service

- Represents the system's distributed (atomic or composite) components
- E.g. SOAP Web service



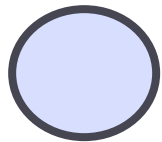
■ Application

- Allows users to interact with the overall system
- E.g. Web applications or portals



■ Data Provider

- Distinguish between the services and the underlying systems that serve as the actual data sources
- Connected to service or application with undirected line



■ Process Unit

- Connected systems that perform functionality beyond data management
- E.g. software that performs computations or triggers events

WAM Core Entities (2)

■ Security Realm

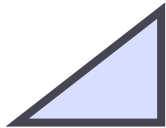
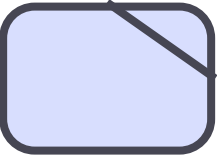
- Envelopes applications, services, data provider and process entities as organizational zones of control – as such functions as identity and access management context
- E.g. defines set of roles and permissions
- Realms might be nested
- Implemented e.g. as a Security Token Service

■ Identity provider

- Store for accounts/identities (of known users as well as applications)
- Allow to authenticate the members of the realm – issues security tokens
- E.g. through login forms or Web service interfaces

■ Name Label

- These labels represent a naming context for each entity
- Naming-Labels might be used as shortcut for a detailed description of these entities



Name Label

Slide excerpts of the lecture EVS

WAM Core Entities (3)

←
[i-label]

- Invocation (Communication Profile)
 - Potential accesses on services and applications
 - Labels indicate the designated communication protocols (label acts as a shortcut for a detailed description of the communication relationship)
 - E.g. SOAP via HTTP, SOAP via SMTP, WS-I compliance etc.

←
[t-label]

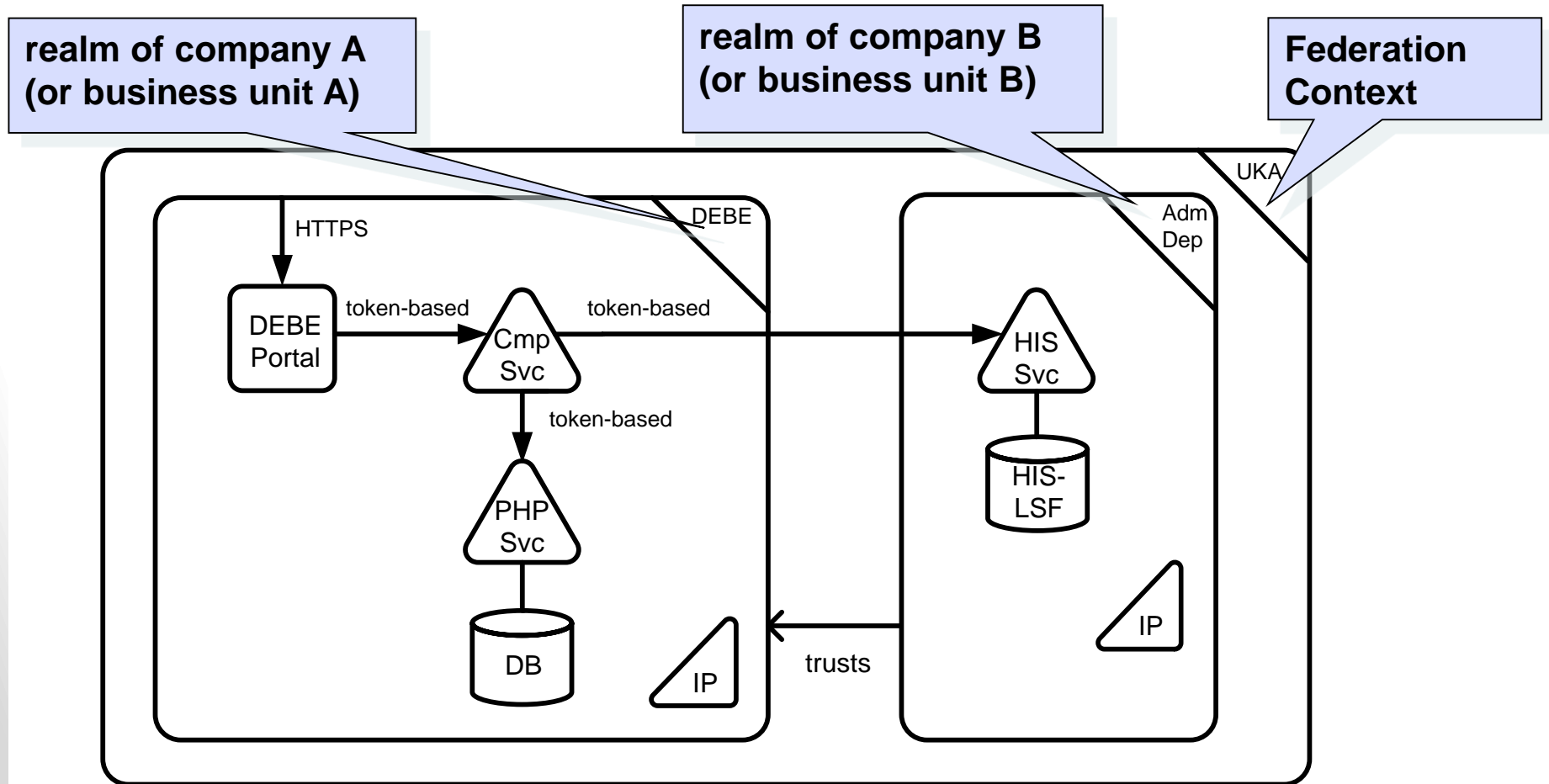
- Trust (Trust Profile)
 - Trust-label separate realms that form a federation
 - STS of the trusting realm accepts the tokens originating from the trusted realm (label acts as a shortcut for a detailed description of the trust relationship)
 - Identities of the foreign requestors can be mapped to tokens that are locally valid – these relationships are defined for the trusts labels

—
[f-label]

- Functionality (Functionality Profile)
 - Links Web Service technology with functionality
 - E.g. technology in use for calling process unit or data provider
- OCL might be used to describe details of Invocation and Trust

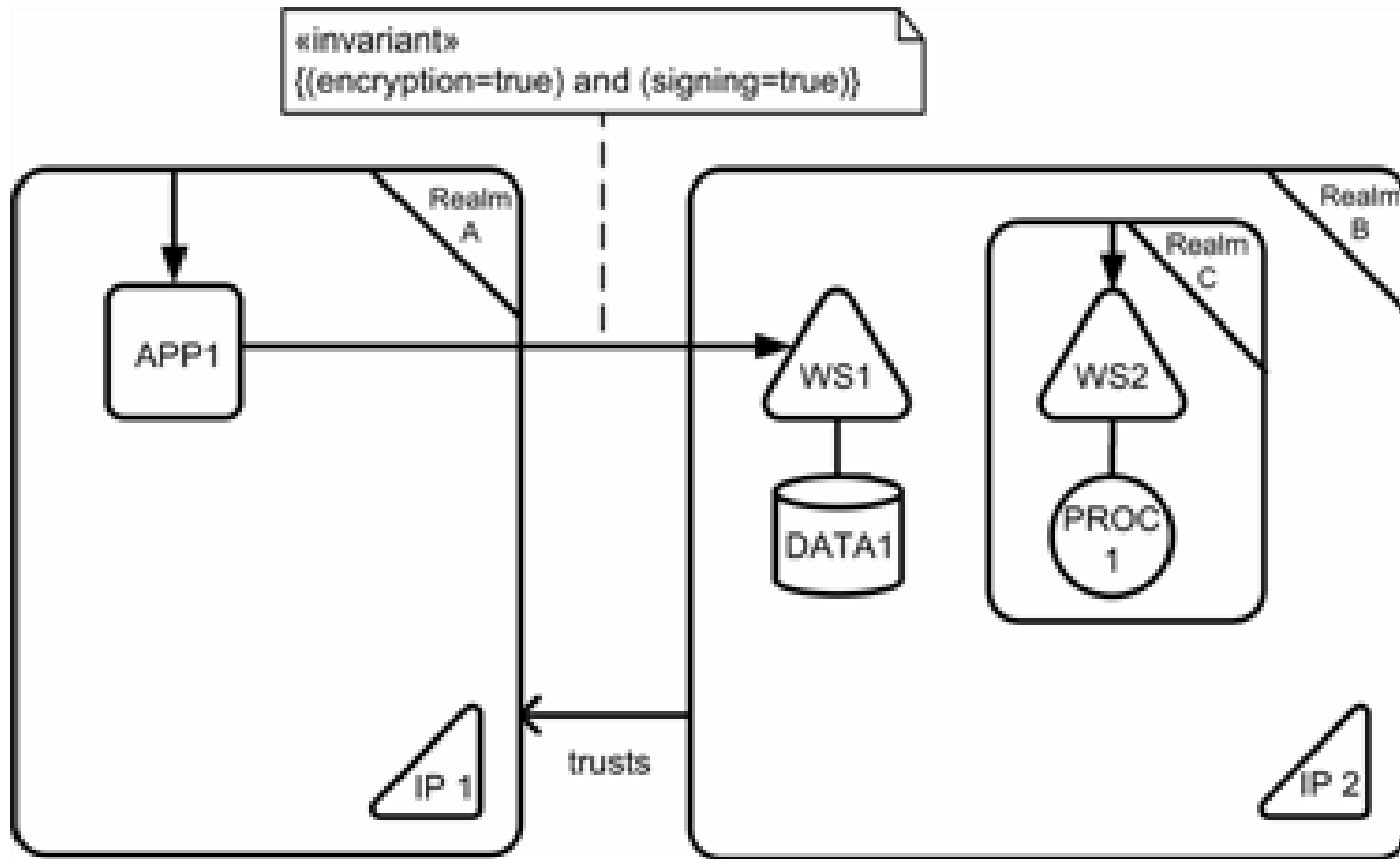
Slide excerpts of the lecture EVS

WAM-Modeling Example*



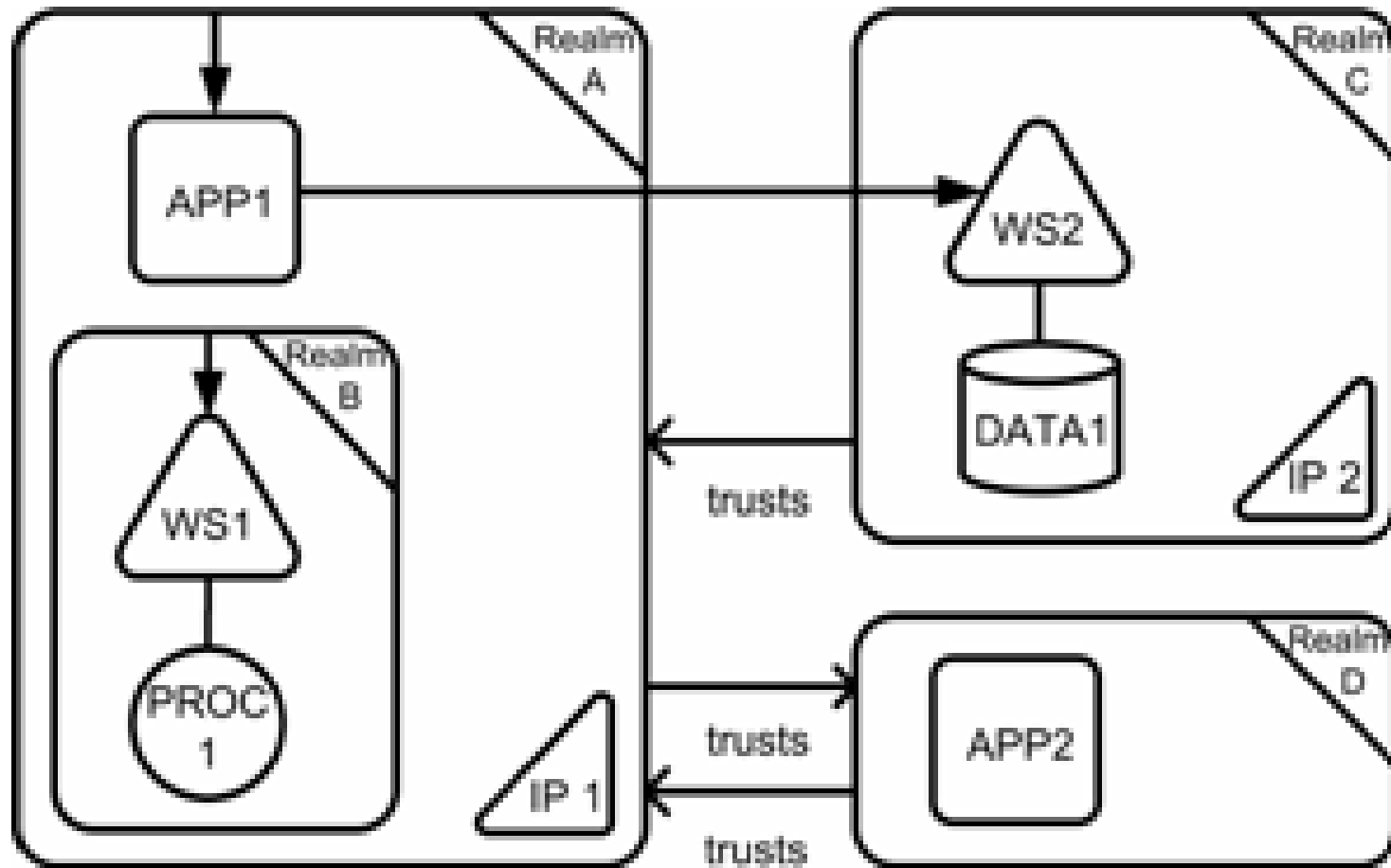
Slide excerpts of the lecture EVS

WAM Example – Use of OCL



Slide excerpts of the lecture EVS

WAM Example – Trust Relationships



**Federation
relationship**

Slide excerpts of the lecture EVS

WAM und WS-Federation

- WS-Federation DEMO
- <http://webcomposition.net/idfs/>



idFS – Flow

