Identity Management an deutschen Hochschulen

Peter Gietz

DAASI International GmbH

1 Einleitung

Ein neues Schlagwort geht durch das Land, „Identity Management“. Heterogene Produkte und Produkt-Suiten werden unter diesem Schlagwort vermarktet, neue Bedürfnisse und neue Hoffnungen geweckt. Auch an deutschen Hochschulen sind diese Botschaften angekommen und vielerorts wurden bereits großangelegte Projekte gestartet.

Mit der Zeit haben sich an Hochschulen eine ganze Reihe von Diensten entwickelt, die alle eigene Benutzerverwaltungen voraussetzen und so zu erheblicher Redundanz der Datenpflege geführt haben. Waren ursprünglich die personenbezogenen Daten von Mitarbeitern und Studierenden in nur zwei Datenbanken gespeichert (in der zentralen Verwaltung und in der Universitätsbibliothek), gibt es mittlerweile eine Vielzahl von Datenbanken, in denen die gleichen Daten zu finden sind: E-Mail- und Login-Account-Verwaltungen, Telefondatenbanken, Mitarbeiterverzeichnisse, Adress-Listen, um nur die wichtigsten aufzuzählen. Überall werden Namen und Kontaktdaten erfasst, überall müssen die Daten gepflegt werden. Eine ganze Reihe von neuen Diensten, wie Web-Portale, Rechnen im Netz (Grid-Computing), Telefonieren über das Internet, etc. stehen vor der Tür und verschärfen das Redundanzproblem.

In einer zunehmenden Konkurrenzsituation zwischen den Hochschulen und neuen Erwartungshaltungen der Kunden, also der Studierenden, sind solche modernen Netzdienste wichtige Alleinstellungsmerkmale. Auf der anderen Seite müssen Hochschulen Stellenstreichungen kompensieren, z.B. durch die Vereinfachung von Verwaltungsvorgängen. In einer solchen Situation bekommt Identity Management, welches verspricht Redundanzen aufzuheben und so den Verwaltungsaufwand zu minimieren, eine wichtige Bedeutung in den strategischen Planungen der Hochschulen.

Dieser Beitrag, der die Ergebnisse eines siebenstündigen Tutoriums zum Thema „Identity Management“ zusammenfasst, möchte eine Begriffsdefinition vornehmen, grundsätzliche Architekturen beschreiben, einen kurzen Überblick über die relevanten Technologien geben, sowie mögliche Implementierung an deutschen Hochschulen beschreiben.

2 Grundbegriffe des Identity Managements

Neben „Identity Management“ (im Folgenden IdM) gibt es im gleichen Zusammenhang eine Reihe weiterer neuer Begriffe, die im Folgenden definiert werden sollen: Identität, Gruppe, Rolle, Provisioning, Metadirectory. Zu weiteren Grundbegriffen und Grundkonzepten, die im IdM wichtig sind, hier aber den Rahmen sprengen würden, vgl. [Slone].

2.1 Identität

Im WWW findet man unter anderem folgende Definitionen für Identität:

www.webster-dictionary.org:

> *The state or quality of being identical, or the same; sameness.*
> *„Identity is a relation between our cognitions of a thing, not between things themselves.“ (Sir W. Hamilton)*

Dictionary.com:

> *The collective aspect of the set of characteristics by which a thing is definitively recognizable or known.*
> *The set of behavioral or personal characteristics by which an individual is recognizable as a member of a group.*

www.hyperdictionary.com:

> *The distinct personality of an individual regarded as a persisting entity.*

Zusammengefasst können wir folgende Merkmale von Identität feststellen:

- eine Gleichheit, die in der Wahrnehmung hergestellt wird, durch die Feststellung X ist identisch mit Y
- eine Reihe von Eigenschaften, die für eine Entität charakteristisch sind und wodurch sie als Mitglied einer Gruppe erkennbar wird
- eine unterscheidbare persistente Persönlichkeit

Auch in der EDV machen diese Merkmale von Identität Sinn. Im engeren Sinn ist eine elektronische Identität nichts anderes als ein eindeutiger Name, Nummer oder Login-Id (im Folgenden wird generell von ID gesprochen), die von einem Computersystem als identisch mit einer Person, einem Computer oder einem Computerprogramm – vielleicht unter dem Begriff Dienst-Benutzer zusammenfassbar – angesehen wird. Aber eine elektronische Identität kann auch weitere mit einer solchen ID verbundene Merkmale beinhalten, wie z.B. eine E-Mail-Adresse, Vor- und Nachname, akademischer Titel, Telefonnummer, etc. Zusätzliche zu einer ID gespeicherten Merkmale können Gruppenzugehörigkeiten, wie z.B. Mitgliedschaften in Maillisten oder Rollen, wie z.B. Funktionen innerhalb der Organisation. Im Idealfall kann man auch von der Persistenz einer ID (nicht jedoch unbedingt der einzelnen weiteren Merkmale) sprechen, wenn diese dauerhaft eine Identität herstellt. Andererseits kann eine Person in verschiedenen Zusammenhängen (unterschiedliche Computersysteme, unterschiedliche Rollen bei einem Computersystem) auch verschiedene Identitäten haben. Dies ist ein Grund für die in der Einleitung beschriebene Redundanz.

2.2 Authentifizierung und Autorisierung

Ein Computersystem führt beim Vorgang der Authentifizierung eine Überprüfung der Identität durch. Die BenutzerIn behauptet, sie sei die mit einer ID verbundene Person.

Das Computersystem muss nun Tests durchführen, um diese Behauptung zu beweisen, bzw. zu widerlegen.

Der einfachste und heute immer noch gebräuchlichste Test, ist die Abfrage eines Passworts. Weiß die BenutzerIn das der ID zugewiesene Passwort, ist die Identität bestätigt, ansonsten widerlegt. Dieser Beweis hat natürlich seine Tücken: Passwörter können mithilfe von Wörterbuchattacken erraten oder im Netz abgehört werden. Bestehen Passwörter aus ausreichend vielen Zeichen aber nicht aus ganzen Wörtern und enthalten sie sowohl Klein- und Großbuchstaben als auch Ziffern und Sonderzeichen, dann gelten sie als nur mit sehr hohem Aufwand erratbar. Wird die Netzverbindung über die das Passwort geschickt wird verschlüsselt oder wird nur eine Abwandlung des Passworts (Challenge-Response-Verfahren) geschickt, ist es auch nicht abhörbar. In jedem Fall wird mit dem Passwort ein Wissen als Beweis für die Identität verwendet.

Andere, stärkere Beweise für Identität sind z.B. der Besitz einer Smartcard, ein privater Schlüssel im Rahmen eines asymmetrischen Verschlüsselungsverfahrens (z.B. X.509-PKI), etc., welcher zumeist mit einem Wissensbeweis kombiniert wird (z.B. Smartcard und PIN). Schließlich können auch sogenannte biometrische Merkmale, also Körpereigenschaften, wie das Iris-Muster, ein Fingerabdruck etc. als Beweis herangezogen werden. Bei allen diesen Verfahren ist die vorherige Zuordnung des Beweismittels zu einer Person eine wichtige Voraussetzung und Fehler bei der Zuordnung können, wie abgehörte Passwörter, zu einer Verfälschung des Beweises führen. Wird also z.B. ein X.509-Zertifikat eines öffentlichen Schlüssels, der eindeutig mit einem privaten Schlüssel verbunden ist, ohne vorherige Prüfung der Identität – im Allgemeinen durch Überprüfung des Personalausweises – ausgestellt wird, kann durch den privaten Schlüssel nicht wirklich Identität bewiesen werden.

Sind die Sicherheitsmechanismen zu schwach, erhöht sich die Gefahr des Identitäts-Diebstahls, bei dem ein Mensch eine Identität z.B. mit einem abgehörten Passwort nur vortäuscht um sich bestimmte Rechte eines Anderen anzueignen.

Nachdem ein Benutzer authentifiziert worden ist, kann das Computersystem aufgrund spezifizierter Zugriffsregeln (auch Access Control bzw. Policy genannt) entscheiden, welche Ressourcen dem Benutzer zugänglich gemacht werden, und welche nicht. Dieser Vorgang wird Autorisierung genannt.

2.3 Identity Management

Identitäts-Management (IdM) wird nach [Lee] verstanden als die Verwendung neuerer Technologien zur Verwaltung von Identitäts-Informationen sowie zu der auf solche Identitäten basierende Kontrolle des Zugriffs auf Ressourcen. Ziel von IdM sei es, Produktivität und Sicherheit zu steigern, sowie die Kosten für Verwaltung von Identitätsinformationen zu senken. Da es schon seit frühester Computerzeit Login-Ids und Zugriffskontrolle gegeben hat, handelt es sich bei IdM also nur um neue Technologien für alte Probleme. Der neue Begriff ist neben seiner Tauglichkeit im Marketingbereich v.a. deswegen nützlich, weil er auf ein notwendiges Gesamtkonzept hinzeigt, wodurch in der Tat Redundanzen und damit Kosten vermieden werden können.

2.4 Gruppe versus Rolle

Wie bereits erwähnt gehören gruppenbildende Merkmale zu Identitätsinformationen. Hierfür haben sich zwei verschiedene Konzepte etabliert, das der Gruppe und das der Rolle. Diese beiden Begriffe können folgendermaßen unterschieden werden:

Rolle ist eine Eigenschaft, die bestimmte Verhaltensregeln und -weisen bestimmt, sowie mit bestimmten Rechten und Pflichten zusammenhängt. Im IdM sind dies insbesondere Funktionen innerhalb einer Organisation z.B. „Professor“. Rollen können hierarchisch strukturiert sein, z.B. Mitarbeiter -> Professor

Gruppe ist viel allgemeiner gefasst. Die Übereinstimmung nur eines beliebigen Merkmals reicht aus, um eine Gruppe zu definieren, z.B. „Abbonent von Mailingliste X“. Gruppen können auch durch Ausschluss gebildet werden, z.B. alle nichtpromovierten Dozenten.

2.5 Provisioning und Metadirectory

„Provisioning“ ist ein weiteres neues Schlagwort, welches heutzutage in Produktbeschreibungen und Marketing-Broschüren häufig gebraucht wird. Auch hier sind Definitionen der ursprünglichen Wortbedeutung hilfreich:

- www.webster-dictionary.org:
  The act of providing, or making previous preparation.
  To supply with food.
  That which is stipulated in advance; a condition; a previous agreement.

Es geht also um eine vorsorgliche Bereitstellung, die auf eine Vereinbarung beruht. Der Begriff ist auch im technischen Bereich nicht neu und wurde dort ursprünglich im Telekommunikationswesen[1] für die Bereitstellung von Hard- und Software verwendet, die ein Kunde für die Benutzung eines Telekommunikationsdienstes benötigt. In Fällen, in denen Hard- und Software beim Kunden bereits vorhanden ist, wird unter Provisioning nur noch der Akt des Freischaltens verstanden, also z.B. der Eintrag der Kundendaten in eine Datenbank, bzw. die Verknüpfung solcher Daten mit einem Dienst oder einer Dienstgüte. Als neues Schlagwort wird letztere Bedeutung auf jeglichen Computer-Dienst erweitert. So findet sich als sehr kurze und allgemeine Definition:

- www.webopadiea.com:
  The process of providing users with access to data and technology resources.

Ganz allgemein versteht man also heute unter Provisioning den Prozess, durch den Benutzer Zugriff auf Daten und Ressourcen bekommen. Dies wird gegenwärtig insbesondere das Erstellen einer ID und eines Passworts sein, bzw. die Versorgung einer Benutzerdatenbank mit diesen Daten.

In zentral verwalteten Strukturen, wie sie im IdM aufgebaut werden, geht es also bei Provisioning um nichts anderes als um die Synchronisierung von Daten (insbesondere ID und Passwort) in Richtung der am IdM angeschlossenen Systeme, wie z.B. Email-Account-Verwaltung.

[1] Vgl. die Definition, die unter http://www.atis.org/tg2k/_provisioning.html zu finden ist.

Eine Methode zur Synchronisierung von Daten zwischen verschiedenen Datenbanken wird „Metadirectory“ genannt, wobei ein Verzeichnisdienst („Directory“) als zentraler Datenspeicher benutzt wird um Daten zwischen heterogenen Datenbanken zu synchronisieren. Im Gegensatz zu Provisioning können mit Metadirectories beliebige Daten in beliebige Richtungen synchronisiert werden.

3 Organisatorische Prozesse und Ihre Abbildung im IdM

Die Intention beim Einsatz von IdM ist die organisatorischen Prozesse, die mit der Verwaltung von Personen und Dingen zu tun haben, abzubilden, um einerseits den Verwaltungsaufwand zu minimieren, andererseits zusätzliche produktivitätssteigernde Informationsdienste zur Verfügung zu stellen. Im Einzelnen wollen Personen

- Informationen über sich veröffentlichen, um z.B. kontaktiert werden zu können
- Informationen über andere Personen erhalten
- sich authentifizieren, also ihre Identität beweisen, um Ressourcen und Dienste in Anspruch nehmen zu können

Organisationen wollen

- Identitätsinformationen über Mitarbeiter oder Mitglieder verwalten
- Benutzer ihrer Ressourcen verwalten
- Konsistenz der Identitäten in verschiedenen Informationsspeichern erreichen
- Vortäuschung falscher Identitäten verhindern

Die Prozesse in Organisationen, die im Rahmen von IdM abgebildet werden, werden in [OpenGroup] folgendermaßen zusammengefasst:

Personen:

- werden in Organisationen aufgenommen
- erhalten Rollen und Berechtigungen
- agieren in ihrer Rolle
- wechseln Rollen und Berechtigungen
- verlassen die Organisation

Organisationen bzw. Organisationseinheiten:

- werden gegründet
- agieren in Arbeitsprozessen
- werden zusammengefügt (merge)
- werden aufgeteilt (split)
- werden aufgelöst

Im IdM werden solche organisatorischen Prozesse mittels EDV-Prozessen abgebildet. Solche EDV-Prozesse sind z.B.:

- Identitäten erzeugen

- Identitätsinformationen aktualisieren
- Identitäten löschen
- Identitäten archivieren
- Identitätsinformation anfordern und anzeigen
- Identitäten verifizieren
- mit Identitäten signieren (PKI)
- Zugriffskontrollregeln durchsetzen (Lese- und Schreibrechte)
- Datenbanken für Identitäten aufbauen und pflegen
- Identitätsdatenbanken synchronisieren
- Identitätsdatenbanken aufteilen und zusammenführen

Solche Prozesse haben verschiedene Agenten also zuständige Stellen, die in einzelnen Prozessen agieren. An Universitäten sind dies z.B.:

- Human Resources Abteilung (im Folgenden HR) für Mitarbeiterdaten, die meistens in zentralen Datenbanken verwaltet werden, z.B. für Lohnbuchhaltung und Abrechnung der Urlaubstage, etc.
- Studentenverwaltung oder Studentensekretariat, die ebenfalls die Personendaten in dedizierten Datenbanken verwalten.
- Rechenzentrum, das Mitarbeitern und Studierenden Dienste zur Verfügung stellt, wie z. B. Systemverwaltung (Login-Accounts für Arbeitsrechner und zentrale Server), Mail-Accounts, Abrechnungsdienste für Printing, etc. Alle diese Dienste setzen wiederum voraus, dass Personendaten in Datenbanken gepflegt werden.
- Universitätsbibliotheken, die für jeden Benutzer einen Personendatensatz (Benutzer-Account) pflegen.
- Technisches Betriebsamt, welches für die Telefoneinrichtung und die Abrechnung der Telefongespräche zuständig ist und hierfür wiederum eigene Verwaltungsdatenbanken mit Personendaten pflegt.
- Pressestelle, welche für die Außenansicht der Universität zuständig ist, wozu die Veröffentlichung von Mitarbeiterverzeichnissen und Vorlesungsverzeichnissen gehört. Auch in diesen Verzeichnissen werden Personendaten in meist zwei getrennten Datenbanken gepflegt.
- Postversandstelle, welche u.a. für den postalischen Versand von Mitteilungen zuständig ist, wofür Adressenverwaltungsdatenbanken mit wiederum den gleichen Personendaten gepflegt werden
- Schließlich gibt es in zunehmenden Maße Stellen, die für die Kommunikation mit ehemaligen Studierenden (Alumni) zuständig sind, entweder ein zuständiges Referat in der Verwaltung oder eine der Universität angeschlossener Verein der Freunde der Universität o.ä.. Solche Stellen, die zunehmend für zusätzliche Einnahmequellen wichtiger werden, pflegen ebenfalls Personendaten.

Abbildung 1 zeigt ein Worst-Case-Scenario, in dem eine Mitarbeiterin/Studierende mit allen diesen Stellen in Kontakt treten muss, um in den Genuss der entsprechenden Dienste zu kommen. Bei jedem der folgenden Prozesse werden Interaktionen mit einer oder mehreren dieser Stellen notwendig:

- Immatrikulation
- Studienfachwechsel
- Abschluss / Wechsel der Universität
- Einstellung als DozentIn
- Promotion
- Wechsel der Büroräume
- Namensänderung (z.B. Heirat)
- Änderung des Wohnorts
- Änderung des Arbeitsvertrags
- Beendigung des Arbeitsverhältnisses
- Immatrikulation als GasthörerIn / Seniorenstudium
- Exmatrikulation

Wenn eine Person alle diese Prozesse durchläuft, finden in einem Worst-Case-Scenario, in dem es keinerlei Synchronisierung zwischen den Datenbanken gibt, über 60 Interaktionen mit den datenpflegenden Stellen statt, wobei insgesamt 10 Verwaltungsangestellte involviert sind.

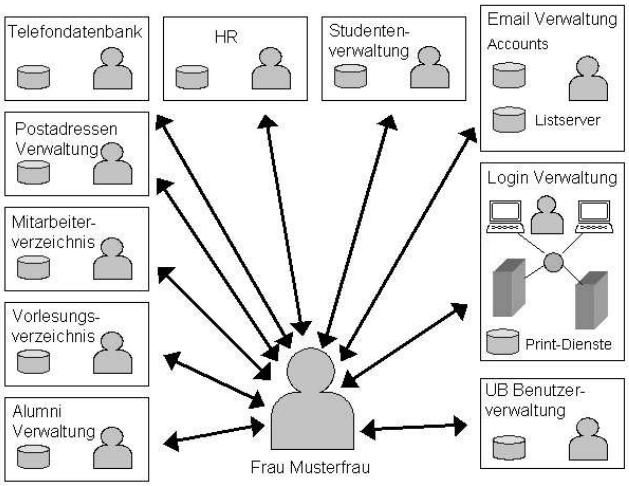

Abbildung 1: Interaktionen mit personendatenpflegenden Stellen ohne IdM

Augenblicklich werden neue Dienste entwickelt, die ebenfalls Personendaten voraussetzen und somit bei redundanter Datenhaltung den Aufwand für die Benutzer und für die Verwaltung weiter erhöhen. Als Beispiele seien hier genannt: Grid Computing Services (verteiltes Rechnen im Netz), Kommunikationsplattformen wie Foren und Portale, Voice

over IP (Telefonieren über das Internet), Public Key Infrastructure (zertifikatsbasierende Infrastruktur für Verschlüsselung und digitaler Signatur).

Ohne IdM bekommen Hochschulen zunehmend Probleme da isolierte, voneinander unabhängige Verzeichnisse/Datenbanken mit den gleichen Identitätsdaten gepflegt werden, die jedoch nicht miteinander interagieren und zwischen denen kein Vertrauen bezüglich der Richtigkeit der Daten besteht. Jede dieser Datensammlungen hat eigene Administratoren, die eigene Datenbanksysteme, Datenschemata, Administrationstools, Authentifizierungs- und Zugriffskontrollmechanismen sowie Regeln und Prozesse verwenden. Jeder Administrator verursacht Verwaltungskosten.

Benutzer brauchen Login-Accounts für jeden Computer und jede Anwendung, die sie nutzen wollen, und müssen sich viele Passwörter merken. Da sie das überfordert müssen immer wieder Administratoren, bzw. Helpdesks kontaktiert werden mit der Bitte, das Passwort zurückzusetzen. Weitere Probleme sind zu langsame Prozesse und Identitätsinformationen, die in verschiedenen Datensammlungen unterschiedlich (Meyer vs. Meier) sind. Die Folge: Benutzer erhalten zu spät Zugriff auf Ressourcen, Zugriffskontrollen werden falsch und das Berichtigen ist wegen notwendiger Kommunikation mit anderen Administratoren aufwendig. Nach dem Ausscheiden eines Mitarbeiters werden u.U. nicht alle Accounts und Berechtigungen gelöscht, wodurch ein generelles Sicherheitsproblem entsteht.

Eine Studie der Metagroup aus dem Jahr 2002 [MetaGroup] hat Unternehmen mit über $500 Millionen Umsatz untersucht, also Unternehmen, die in Bezug auf Mitarbeiterzahlen mit größeren Universitäten vergleichbar sind. Die Ergebnisse zeigten u.a. dass:

- 45% der Help-Desk-Aktivitäten aus dem Rücksetzen von Passwörtern bestehen
- 11% der Mitarbeiter mindestens ein Zugriffsrechte-Problem pro Monat haben
- Provisioning-Vorgänge zwischen 6 und 29 Stunden dauern
- interne Benutzerinformation an 22 verschiedenen Orten gespeichert werden

Eine solche Situation ist das Resultat historisch gewachsener Infrastrukturen und Prozesse, die sich ohne ein von allen Beteiligten akzeptiertes IdM-Konzept nur schwer optimieren lassen. Hinzu kommt ein zunehmender Erwartungsdruck der Mitarbeiter und Kunden (Studierende). Hier zeigt sich, dass IdM keine Aufgabe einer einzelnen Stelle, sei es die Verwaltung oder das Rechenzentrum, sein kann, sondern dass eine IdM-Lösung nur dann implementierbar ist, wenn alle beteiligten Stellen von den Vorteilen überzeugt sind und gemeinsam an der Realisierung arbeiten. Besitzstände, Aufgabenhoheiten, sowie weitere politische und menschliche Faktoren gilt es zu berücksichtigen, soll ein IdM-Projekt Erfolg haben.

4 Grundbausteine des IdM und Technologie-Standards

IdM besteht aus verschiedenen logischen Komponenten. Verschiedene Produkte beinhalten nur Teilmengen dieser Komponenten und aus diesen Komponenten lassen sich verschiedene Architekturen aufbauen. Im Folgenden werden solche Komponenten auf einer generischen Ebene kurz beschrieben.

4.1 Informationsspeicher

Grundlage eines jeden zentralen IdM ist eine autoritative Datenquelle. Idealerweise sollten wenigstens die Grunddaten, also Name, Vorname, akademischer Titel, Art der Organisationszugehörigkeit, postalische Adresse und weitere Kontaktdaten nur in einer autoritativen Datenquelle gepflegt werden. Realistischer ist es anzunehmen, dass es mehrere solche autoritativen Quellen gibt, z.B. die HR-Datenbank für Name, Vorname, etc. und die E-Mail-Account-Verwaltungsdatenbank für die E-Mail-Adresse. Auch in diesem Fall sollten diese Daten in einen Informationsspeicher zusammengeführt werden, welcher dann als einzige das IdM-System mit Daten versorgt. In der Regel werden für diese Informationsquelle entweder relationale Datenbanksysteme (RDBM) mit der standardisierten Abfragesprache SQL verwendet, oder aber Verzeichnisdienste, die auf X.500 bzw. LDAP basieren, wobei sich letztere Technologie zunehmend für IdM-Zwecke durchsetzt. LDAP-basierte Systeme weisen einige Vorteile gegenüber RDBM, die in folgender Tabelle dargestellt werden. Eine Einführung in LDAP bietet z.B. [Gietz 2002].

Aspekt	**RDBM**	**LDAP**
Schema	Es gibt keine standardisierten Tabellenschemata.	Es gibt insbesondere zur Abbildung von Personendaten, inklusive Gruppen und Rollenkonzepte international standardisierte Datenschemata.
Organisation	Entitätsinformationen werden logisch auf verschiedene Tabellen aufgeteilt.	Entitätsinformationen bleiben logisch an einem Platz, nämlich in einem Informationsobjekt, welches einen Knoten in einem hierarchischen Baum darstellt.
Mehrfachwerte	Mehrfachwerte erzwingen eine neue Tabelle (Normalisierung) oder verschiedene Datenfelder wie z.B. Telefonnummer_1, Telefonnummer_2.	Beliebig viele Mehrfachwerte lassen sich problemlos speichern.
Datentypen	Begrenzte Anzahl von Datentypen wie String, Integer, Float, und Datum.	Theoretisch unbegrenzte Anzahl von Datentypen (hier Syntax genannt), de facto eine Vielzahl von Datentypen, insbesondere für Personendaten, z.B. eigene Syntax für Telefonnummern.
Vergleichsregeln	Vergleichsregeln sind nicht Teil des Datenmodells, sondern müssen in den jeweiligen Abfrageprogrammen implementiert werden.	Vergleichsregeln sind Teil des Datenmodells. Man kann also z.B. bei der Schemadefinition bestimmen, ob bei Wertvergleichen Groß- und Kleinschreibung berücksichtigt werden soll oder nicht. Es gibt unterschiedliche Vergleichsregeln für Gesamt- und Teilvergleiche.
Flexibilität	Änderungen des Datenschemas, also der Tabellenstruktur sind nur schwer möglich. Änderungen betreffen die gesamte Datenbank.	Änderungen des Datenschemas einfach möglich: man fügt einem Informationsobjekt eine neue Objektklasse hinzu und kann dann entsprechende neue Attribute speichern. Änderungen betreffen jeweils nur die gewünschten Informationsobjekte.
Netzzugriff	Netzzugriff ursprünglich nicht vorgesehen. Wird meistens über ein Gateway realisiert.	Netzprotokoll ist Hauptteil des LDAP-Standards. eine Verteilung der Daten im Netz ist einfach möglich.
Authentifizierung	Meist nur ein proprietärer Authentifizierungsmechanismus.	Verschiedene standardisierte und über das Netz funktionierende Authentifizierungsmechanismen.

Es soll hier nicht verschwiegen werden, dass RDBMs gegenüber LDAP auch Vorteile aufweisen, wie z.B. verlässliche Transaktionsmechanismen, die in LDAP-Systemen nur über das eigentliche Datenbackend zu realisieren sind, oder höhere Performanz bei Schreibvorgängen. Insgesamt ist LDAP aber gerade für Personendaten konzipiert worden, die wesentlich häufiger gelesen als modifiziert werden, und die Vorteile überwiegen bei Weitem die Nachteile.

Mithilfe von Verzeichnisdiensten lassen sich verschiedene Architekturen aufbauen, wovon hier nur zwei kurz erwähnt werden sollen.

Zum einen ist es möglich einen zentralen Datenspeicher zu implementieren, der dezentral administriert werden kann, also verschiedene Datenbereiche oder Datenfelder von verschiedenen Administratoren gepflegt werden können. Aus diesem zentralen Speicher werden dann automatisch verschiedene Dienste im Netz repliziert, also z.B. ein zentraler Authentifizierungsdienst (Login-Server, der z.B. auch vom IMAP-Mailserver verwendet wird) im Intranet und ein öffentliches Mitarbeiterverzeichnis im Internet. Ein solches Beispiel zeigt Abbildung 2.

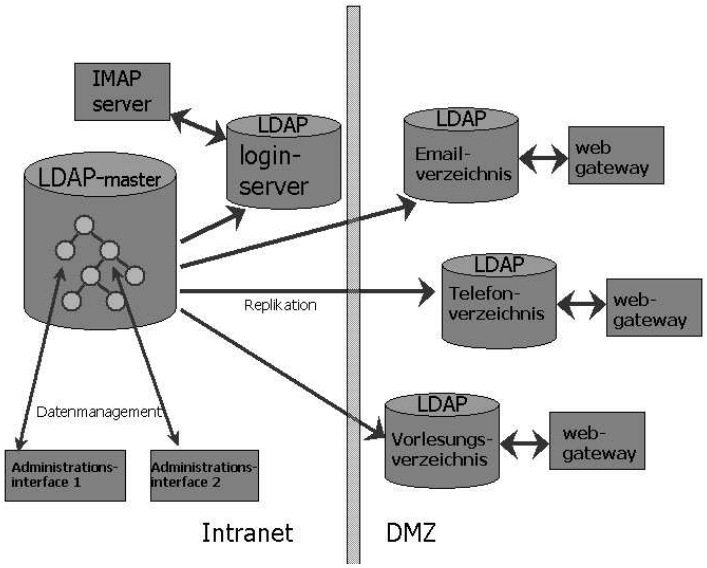

Abbildung 2: Beispiel eines zentralen Verzeichnisdienstes

Eine weitere Möglichkeit LDAP im Rahmen des IdM einzusetzen ist die Implementierung eines Metadirectoy. Hierbei werden Daten aus verschiedenen Datenbanken zusammengeführt und in die verschiedenen Datenbanken zurücksynchronisiert. Das Datenmanagement findet nur in den angeschlossenen Datenbanken statt, wobei verschiedene Datenbanken für verschiedene Datenfelder autoritative Quelle sein können. Das Metadirectory dient einerseit zur Synchronisierung der Daten, andererseits bietet es eine Gesamtsicht der Daten der

Einzeldatenbanken. Letztere Funktionalität bieten auch sogenannte Virtual Directories, welche selbst im Gegensatz zum Metadirectory keine eigenen Daten vorhalten, sondern nur dynamisch Gesamtsichten über die angeschlossenen Datenbanken bieten. In jedem Fall werden sogenannte Konnektoren benötigt, Softwaremodule, welche die Datenkonvertierungen und Datenübertragungen vornehmen. Idealerweise sollten die Konnektorenprozesse bei jeder Datenänderung in den Datenbanken durch sogenannte Trigger angestoßen werden. Abbildung 3 zeigt beispielhaft die Funktionsweise eines Metadirectory, bei dem Name, etc. aus der HR-Datenbank stammen, die Telefon- und Raumnummer aus der Telefonapparatedatenbank. Diese Merkmale werden in die anderen Datenbanken synchronisiert. Zusätzlich wird ein Teil des Metadirectory als Mitarbeiterverzeichnis veröffentlicht.

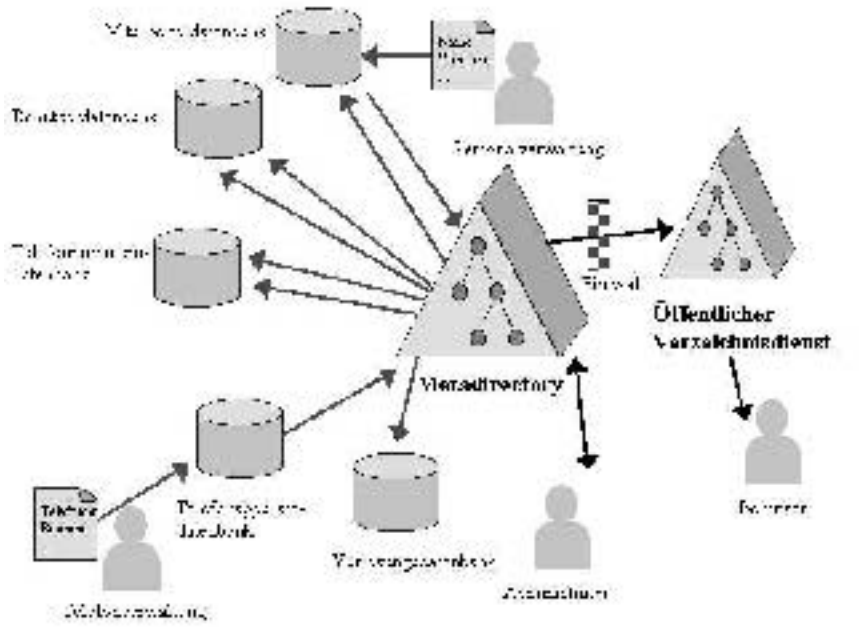

Abbildung 3: Beispiel eines Metadirectory

Angesichts der oben bereits beschriebenen historisch gewachsenen Strukturen, ist es meist realistischer die Metadirectory-Architektur anzuwenden, da diese flexibel an vorhandene Prozesse angepasst werden kann und vorhandene Systeme nur minimal beeinträchtigt. Die Lösung mit einem zentralen Verzeichnisdienst ist jedoch eine konsequentere Zentralisierung und einfacher zu warten. Eine solche Zentralisierung der Infrastruktur muss wohlgemerkt nicht auch eine Zentralisierung der Datenhoheiten und -verantwortlichkeiten bedeuten, setzt aber das Vertrauen der Datenadministratoren zu den Infrastrukturbetreibern voraus.

4.2 Provisioning-Komponente

Eine wesentliche Komponente des IdM ist das Provisioning, welches dafür zuständig ist, dass angeschlossene Systeme Benutzer-Account-Informationen erhalten um Authentifizierungs- und Autorisierungsprozesse durchführen zu können. In beiden im vorigen Abschnitt beschriebenen LDAP-basierten Architekturen wird das Provisioning über LDAP realisiert. Allerdings setzt dies voraus, dass alle angeschlossenen Systeme eine LDAP-Schnittstelle besitzen um Authentifizierungs- und Autorisierungs-Informationen zu erhalten. Man kann zwar sagen, dass bereits viele Systeme (sowohl Betriebssystem-Login-Prozesse, als auch viele Anwendungen, die auf Benutzerdaten zugreifen müssen) LDAP-Schnittstellen für Authentifizierungsprozesse besitzen und es sich hierbei um einen zunehmenden Trend handelt. Authorisierungsinformationen hingegen, also detaillierte Informationen, welche Benutzer, bzw. welche Gruppen oder Rollen auf welche Ressourcen zugreifen dürfen, können zwar ebenfalls im LDAP-Server abgelegt werden, es gibt jedoch nur wenige Anwendungen, die darauf zugreifen können. Eine der wenigen Ausnahmen bildet der Apache-Web-Server, bei dem man Autorisierungsinformation mittels LDAP-Abfragen ablegen kann. Näheres zu diesem Themenkomplex kann man [Gietz 2004] entnehmen.

Die Alternative zu LDAP ist in gängigen IdM bzw. dedizierten Provisioning-Produkten die Implementierung von Spezialkonnektoren für die verschiedenen anzuschließenden Systeme. Da in der Wirtschaft hauptsächlich nur eine begrenzte Anzahl von Systemen verwendet werden, bieten solche Produkte ausreichende Lösungen, zumal meistens auch generische Konnektoren, wie z.B. CSV (Comma Separated Vector) sowie Entwicklungsumgebungen für die Programmierung eigener Konnektoren zur Verfügung gestellt werden.

Allerdings zeigt sich auch im Bereich Provisioning der Bedarf an internationalen Standards. So wurde vom Gremium OASIS (www.oasis-open.org) der Standard SPML (Service Provisioning Markup Language) [Rolls] spezifiziert, der zunehmend von den Produkten unterstützt wird. Dieses XML-Schema-basierende Protokoll dient zum Austausch von Provisioning-Information. Hierbei agieren folgende Komponenten:

- Requesting Authority (RA), eine Software-Komponente welche SPML-Requests an einen SPML Service Point schickt
- Provisioning Service Point (PSP), eine Doftware-Komponente welche SPML Requests erhält und verarbeitet, also das eigentliche Provisionierungs-System
- Provisioning Service Target (PST), eine Software-Komponente welche das Ziel einer Provisioning-Aktion darstellt

SPML enthält ein Binding für SOAP, sodass es sich einfach in moderne Web-Services-Architekturen einbauen lässt, sowie ein Binding für direktes HTTP, um einfachere Architekturen unterstützen zu können. Das Funktionsmodell von SPML enthält Request- und Response-Elemente für folgende Operationen:

- Add
- Delete
- Modify
- Search

- Extended operation

Bei der Definition von SPML wurde Wert darauf gelegt, dass es mit den OASIS-Standards DSMLv1 und v2 kompatibel ist (Directory Service Markup Language)[DSMLv1] [DSMLv2], welche wiederum das XML-Austauschformat für LDAP-Daten bzw. Operationen darstellen. Auch dies ist ein Indikator für die Wichtigkeit von LDAP im IdM.

Ein weiterer in diesem Zusammenhang wichtiger OASIS-Standard ist SAML (Security Assertion Markup Language) womit sicherheitskritische Informationen, Authentifizierungs- und Autorisierungsinformation, sowie weitere Attributierungen in XML beschrieben und ausgetauscht werden können. SAML v.1.1.[SAMLv1.1] definiert Protokolle, durch die Clients sog. Assertions von SAML-Autoritäten anfordern und deren Antworten empfangen können. Die SAML-Autoritäten können diverse Informationsquellen, beispielsweise externe Policy-Repositories oder in der Anfrage bereits enthaltene Assertions, benutzen, um ihre Antworten zu generieren. Assertions werden dann zu einem SAML-Ausweis gebündelt, der zusammen mit der Anforderung an den Policy Enforcement Point (PEP) der angeforderten Ressource geschickt wird. Dieser entscheidet dann über den Zugriff. Der SAML-Ausweis wird digital unterschrieben, um Modifikationen zu verhindern. Deswegen können SAML-Ausweise wiederverwendet werden, ohne eine erneute Authentifizierung erforderlich zu machen. SAML ermöglicht damit Single Sign-On (SSO). Als Authetifizierungsmechanismen werden Passwörter, Kerberos, Zertifikate[2] (X.509, SPKI, XKMS, SSL/TLS) und XML digitale Signaturen unterstützt.

Schließlich wurde ebenfalls von OASIS der Standard XACML (Extended Access Control Markup Language) [XACML] herausgegeben, mittels dessen generische Zugriffsregeln (Policy) anwendungsunabhängig spezifiziert werden können, wobei sich eine Policy auf eine andere an einem anderen Ort im Netz gespeicherte Policy beziehen kann. Es gibt XACML-Profile für SAML und LDAP/DSML.

Mit LDAP, DSML, SPML, SAML und XACML stehen für die wichtigsten Bereiche des IdM stabile und anerkannte Standards zur Verfügung, die auch zunehmend in kommerziellen Produkten implementiert werden. Für die meisten dieser Standards gibt es aber auch Open-Source-Implementierungen, sodass die meisten Bausteine für eine Open-Source-basierte IdM-Implementierung vorhanden sind.

4.3 Weitere Komponenten

4.3.1 Policy Management

Viele IdM-Produkte enthalten eine Oberfläche, mit der man Authorisierungsinformation editieren und verwalten kann. Diese Komponente wird Policy Management genannt. Um Flexibilität und Interoperabilität zu gewährleisten, sollten die Regeln in XACML ablegbar sein.

Zugriffsregeln sollten nicht pro Identität vergeben werden, sondern besser pro Rolle. Hierdurch lässt sich die Anzahl der zu definierenden Regeln wesentlich verringern. Die Rollenattributierung kann bereits im zentralen Datenspeicher vorgenommen werden, indem

[2] Zu einem LDAP-Schema für die Ablage von X.509-Zertifikaten vgl. [Gietz 2003]

einem Personendatensatz zum Beispiel die Rolle „Professor“ hinzufügt wird. In der Zugriffs-Policy braucht man dann nur noch die Berechtigungen eines Professors festzulegen. Durch Hierarchisierung von Rollen (Beispiel: Professor hat die Unterrollen ordentlicher Professor und Gastprofessor mit unterschiedlichen Berechtigungen) kann man trotzdem relativ flexibel Berechtigungen definieren. Als Ausnahme kann man immer noch zusätzlich Einzelregeln für einzelne Identitäten festlegen.

Man sollte nicht zu viele verschiedene Rollen definieren, da sonst der Hauptvorteil des Rollenkonzepts, nämlich die Reduzierung der Regelanzahl, vertan ist. Deshalb muss man in den Daten Rolle und Rollenbindung voneinander getrennt halten. In manchen Fällen möchte man zusätzliche Daten zur Rollenbindung speichern, wie z.B. ein Ablaufdatum, ab dem die Identität nicht mehr die Rolle und deren Berechtigungen erhalten soll. Hier bietet es sich an, drei Informationsobjekte zu verwenden. Eins für das Identitätsobjekt, eins für ein generisches Rollenobjekt und schließlich ein eigenes Objekt, in welchem einerseits die Bindung zwischen Identität und Rolle abgebildet werden kann, andererseits die benötigten Zusatzdaten, wie z.B. Ablaufdatum. Es bedarf allerdings genauerer Untersuchungen und Experimente, um die Performanz eines solchen Datenmodells zu evaluieren.

Auch Gruppen können bei den Zugriffsregeln berücksichtigt werden. Man kann im LDAP-Datenmodell drei Arten von Gruppen unterscheiden:

1. Statische Gruppen: Hierbei werden in einem Gruppeneintrag Verweise auf alle Mitglieder der Gruppe gesammelt. Um also z.B. die E-Mail-Adresse aller Gruppenmitglieder zu erhalten, muss ein Programm zuerst den Gruppeneintrag suchen und dann alle Verweise zu den Mitgliedern verfolgen. Es ergibt sich hierbei allerdings das Problem, referenzielle Integrität aufrecht zu erhalten, da man bei jeder Löschung eines Personeneintrags, suchen muss, ob auf diesen Eintrag von einem Gruppeneintrag aus verwiesen wurde, um diesen Zeiger dann ebenfalls zu löschen.
2. Halbdynamische Gruppen: Hierbei wird genau umgekehrt verfahren, es wird also im Personeneintrag ein Zeiger auf einen Gruppeneintrag gesetzt. Das Programm kann hier als Suchfilter eben diesen Verweis verwenden, kommt also mit einem einzigen Suchvorgang aus. Das Referenzialitätsproblem ergibt sich zwar auch bei halbdynamischen Gruppen, jedoch wesentlich entschärfter: Nur wenn ein Gruppeneintrag gelöscht wird, ist es notwendig alle auf diesen Gruppeneintrag zeigende Verweise zu suchen und entsprechend abzuändern.
3. Dynamische Gruppen: Unter einer Dynamischen Gruppe versteht man eine Gruppenzugehörigkeit, die nur durch ein gemeinsames Merkmal gebildet wird. Alle Einträge, die also bei einem bestimmten LDAP-Filter gefunden werden, gehören zu einer Gruppe. Da auf keinen anderen Eintrag verwiesen wird, gibt es auch kein Referenzialitätsproblem. Allerdings lassen sich keine Zusatzinformationen zu einer Gruppe speichern, im Falle einer Mailingliste also z.B. nicht, ob die Liste moderiert ist, wer für die Liste zuständig ist, oder ähnliches.

4.3.2 Workflow Management

Mit dieser Software-Komponente kann man die Abläufe bei Neueinstellung, Datenänderungen und Aufhebung des Beschäftigungsverhältnisses festlegen. Hierbei wird genau de-

finiert welche Attribute von wo nach wo fließen sollen und was die jeweilige autoritative Datenquelle ist. Es werden mittlerweile verschiedene Workflow-Beschreibungssprachen standardisiert, die für Interoperabilität zwischen verschiedenen Workflow-Managementsystemen sorgen könnten, wobei die beiden wichtigsten Workflow-XML (wf-XML)[3] und Web Services Choreography Markup Language (WS-CML)[4] sind.

4.3.3 Passwort-Management

Auch bei gleichem Passwort müssen die Passwörter oft redundant gespeichert werden da verschiedene Systeme verschiedene Verschlüsselungs- bzw. Hash-Algorithmen für die Ablage von Passwörtern verwenden. Problematisch sind hierbei insbesondere Windows- und Lotus Notes Passwörter. Bei der Synchronisierung von Passwörtern finden deshalb komplexere Prozesse statt. Es gibt hierbei zwei verschiedene Architekturen für das Passwort-Management:

1.) Zentrales Zurücksetzen des Passworts
Hierbei können Passwörter nur über ein zentrales Interface, in der Regel ein Web-Formular, eingetragen, geändert bzw. zurückgesetzt werden. Die Passwörter müssen nur in eine Richtung (jeweils richtig verschlüsselt) synchronisiert werden. Hierbei ist zu beachten, dass lokale Passwortänderungen (also z.B. passwd auf einem Unix-Rechner, bzw. das entsprechende Interface auf Windowsrechnern) deaktiviert werden müssen, was Akzeptanzprobleme bereiten könnte. Andererseits ist die Struktur einfacher und damit übersichtlicher und leichter zu pflegen.

2.) Dezentrale Passwort-Synchronisierung
Hierbei sind Passwortänderungen z.B. sowohl an der Windows-Domäne als auch im Unix-Rechner möglich, sodass sich die Benutzer nicht umstellen müssen. Allerdings ist dieses System komplexer und enthält mögliche Sicherheitslöcher, da das Passwort zunächst unverschlüsselt über das Netz geht, damit das System es dann für das jeweilige Zielsystem verschlüsseln kann.

4.3.4 Auditing

In einem IdM-System finden ständig automatisch gestartete Prozesse statt. Insbesondere da es sich um sicherheitsrelevante Prozesse handelt, ist es wichtig, dass alle Prozesse und Transaktionen in einem zentralen Log-File, bzw. einer zentralen Log-Datenbank geloggt werden. Hierdurch lassen sich alle Änderungen der Account-Daten und Berechtigungen sowie Einbruchsversuche verfolgen.

5 IdM an deutschen Hochschulen

Mittlerweile haben viele Hersteller Metadirectory- bzw. IdM-Produkte auf den Markt gebracht, wie z.B. BMC, Calendra, Computer Associates, IBM, MaxWare, Microsoft, Novell, Siemens und Sun. Allerdings bieten diese Produkte, die eher für den Einsatz in der

[3] Vgl. http://www.WfMC.org.

[4] Vgl. http://www.W3C.org.

Wirtschaft konzipiert wurden, keine Konnektoren zu universitätsspezifischen Datenbanken, wie sie z.B. von der HIS GmbH entwickelt werden. Außerdem bleibt die Realisierung komplex und bedarf einer sehr guten Kenntnis der Produkte. Schließlich sind die Lizenzbedingungen oft für die Universitäten zu kostenintensiv insbesondere wenn die Kosten nach der Anzahl der Identitäten berechnet werden.

Es besteht aber auch die Möglichkeit, ein IdM auf Grundlage des OpenSource-Produkts OpenLDAP zu realisieren, welches schon einige Ansätze zu Metadirectories in verschiedenen Datenbackends verfolgt, mittels derer relativ einfach Konnektoren zu den Datenbanken selbst entwickelt werden können. Für die erwähnten XML-basierten IdM-Standards gibt es ebenfalls Open-Source-Implementierungen bzw. -Funktionsbibliotheken, so dass sich mit überschaubarem Aufwand ein kompletes Open-Source-basierendes IdM-System realisieren ließe.

Einerseits besteht also ein hoher Bedarf an IdM-Lösungen im deutschen Forschungsumfeld, andererseits gibt es eine Vielzahl von Realisierungsmöglichkeiten. Einige Universitäten fangen an, sich individuell mit der Thematik zu beschäftigen, stoßen dabei aber schnell auf Ressourcengrenzen, da es sehr zeitintensiv ist, sich einen guten Überblick über die Möglichkeiten zu verschaffen. Ein solcher Überblick ist jedoch Voraussetzung ist für richtige strategische Entscheidungen.

Hier wäre ein Zusammenschluss der Hochschulen sinnvoll, um gemeinsame Aufgaben mit Synergieeffekten gemeinsam zu lösen. Ansätze solcher Kooperationen lassen sich auf Länderebene bereits feststellen, insbesondere in Thüringen, in Ansätzen auch in NRW, wobei jeweils nur einige der vorliegenden Möglichkeiten evaluiert wurden.

Ein nationales Identity-Management-Kompetenzzentrum könnte in dieser Hinsicht viel mehr leisten. Neben Beratung und Support für Forschungsinstitutionen könnten folgende Projektziele angestrebt werden:

- Studie zum Bedarf von Identity Management an deutschen Universitäten, sowie deren spezifischen Anforderungen
- Herstellerunabhängige Evaluation der kommerziellen IdM-Produkte mit besonderer Berücksichtigung der erhobenen Anforderungen
- Entwicklung von Konnektoren für OpenLDAP zur Realisierung einer OpenSource-IdM-Lösung
- Entwicklung von XML-basierenden Komponenten zur Unterstützung der Standards DSML, SPML, SAML und XACML unter Rückgriff auf bestehende Open-Source-Projekte
- Allgemeine Richtlinien zum Aufbau von Identity-Management (Datenstruktur, Datenschema, Frequenz der Updatemechanismen, etc.)
- Erstellung von „Kochbüchern“ für die einzelnen IdM-Implementierungen, einschließlich der Open-Source-Lösung

Alle Projektergebnisse könnten allen Hochschulen zur Verfügung gestellt werden und diesen bei der Implementierung einer eigenen IdM-Lösung von großem Nutzen sein. Insbesondere für kleinere Universitäten und Fachhochschulen, die in der Regel über zu wenige Mittel verfügen, um die kostspieligen kommerziellen Produkte anzuschaffen, könnte die Open-Source-Lösung die einzige Möglichkeit sein, zu einem IdM zu gelangen.

Literatur

[DSMLv1] Tauber, James; Hay, Todd; Beauvais, Tom; Burati, Mike; Roberts, Andrew: Directory Services Markup Language (DSML), Last updated: 1999-12-02, http://www.oasis-open.org/committees/dsml/docs/dsml.zip

[DSMLv2] Directory Services Markup Language v2.0 December 18, 2001, Approved as an OASIS Standard April 30,2002, http://www.oasis-open.org/committees/dsml/docs/DSMLv2.doc

[Gietz 2002] Gietz, Peter: Verzeichnisdienste für Hochschulen auf Open Source Grundlage. In: v. Knop, Jan, Bode, Friedrich: Tagungsband über die Informations- und Verzeichnisdienste in Hochschulen, Düsseldorf 2002.

[Gietz 2003] Gietz, Peter: Neues Konzept für einen DFN-weiten Zertifikatsserver. In: von Knop, Jan; Haverkamp, Wilhelm; Jessen, Eike [Hrsg.]: Security, E-Learning, E-Services, 17. DFN-Arbeitstagung über Kommunikationsnetze, Düsseldorf Lecture Notes in Informatics (LNI) – Proceedings, Series of the German Informatics Society (GI) P-44, Bonn 2003

[Gietz 2004] Gietz, Peter: LDAP-basierte zentrale Authentifizierungssysteme, In: Thorbrügge, Marco [Hrsg.]: 11. Workshop „Sicherheit in vernetzten Systemen“. 03./04. Februar 2004, Hamburg, DFN-CERT GmbH, Hamburg Publications 2004

[Lee] Lee, Spencer C.: An Introduction to Identity Management. SANS Security Essentials Certification Practical Assignment version 1.4b option 1, March 11, 2003, http://www.sans.org/rr/papers/6/852.pdf

[MetaGroup] The Meta Group: The Value of Identity Management: How securing identity management provides value to the enterprise, August 2002, .

[OpenGroup] The Open Group: Business Scenario: Identity Management, 5. July 2002, http://www.opengroup.org/dif/projects/im-scen/idmbs_1.pdf

[Rolls] Rolls, Daran [ed.]: Service Provisioning Markup Language (SPML) Version 1.0, OASIS Committee Specification, 3 June 2003, http://www.oasis-open.org/committees/download.php/3032/cs-pstc-spml-core-1.0.pdf

[SAMLv1.1] Maler, Eve; Mishra, Prateek; Philpott, Rob: OASIS: Assertions and Protocol for the OASIS Security Assertion Markup Language (SAML) V1.1, OASIS Standard, 2 September 2003, http://www.oasis-open.org/committees/download.php/3406/oasis-sstc-saml-core-1.1.pdf

[Slone] Slone, Skip: Identity Management – A White Paper, The Open Group Identity Management Work Area, March, 2004. Vgl. http://www.opengroup.org/bookstore/catalog/w041.htm

[XACML] Godik, Simon, Moses, Tim [eds.]: eXtensible Access Control Markup Language (XACML) Version 1.0, OASIS Standard, 18 February 2003, http://www.oasis-open.org/committees/download.php/2406/oasis-xacml-1.0.pdf