

Maria Oelinger

Realisierung einer Website auf der Basis
eines UML-Entwurfes am Beispiel einer
Materialsammlung aus Informatik und
Mathematik

Diplomarbeit

BEI GRIN MACHT SICH IHR WISSEN BEZAHLT



- Wir veröffentlichen Ihre Hausarbeit, Bachelor- und Masterarbeit
- Ihr eigenes eBook und Buch - weltweit in allen wichtigen Shops
- Verdienen Sie an jedem Verkauf

Jetzt bei www.GRIN.com hochladen
und kostenlos publizieren





Realisierung einer Website auf der Basis eines UML-Entwurfes

am Beispiel einer Materialsammlung aus Informatik und Mathematik

Diplomarbeit DI

vorgelegt von Maria Oelinger

VORWORT

An die LeserInnen

Wer die Ergebnisse aus dieser Arbeit als hilfreich empfindet, sollte an die vielen netten Leute denken, die dazu einen Beitrag leisteten. Durch persönliche Anmerkungen, konstruktive Kritik, spontanes Feedback und die Praxiserfahrungen während der Projekte mit KollegInnen konnte diese Diplomarbeit erst entstehen.

Universität Duisburg

Zunächst danke ich Prof. Dr. Hoppe für seine wertvollen Hinweise und Ratschläge während der Bearbeitung des Themas. Er hatte immer ein offenes Ohr für Fragen und Probleme, was zur Qualität dieser Arbeit wesentlich beigetragen hat. Die anderen Dozenten der Informatik haben mich mit ihren Vorlesungen vom ersten Semester an für dieses Fach begeistert und weckten mein Interesse für Datenbanken. Die MitarbeiterInnen des Fachbereichs Informatik diskutierten mit mir und hinterfragten kritisch die einzelnen Funktionalitäten der Website.

Ressourcen

Martin Gerardi und Dirk Holtwick von der spirito GmbH Duisburg danke ich dafür, dass sie mir den Dado Web Application Server und den Webspace zur Verfügung stellen. Sie nahmen sich ausgiebig Zeit, mir meine Fragen zu beantworten. Die MitarbeiterInnen von spirito gaben mir Anregungen und ließen mich die Site in ihren Browsern testen. Viele Ideen konnten so während fachlicher Diskussionen entstehen.

Materialien für die Fachinhalte

Patricia Jung stellt mir den Linux-Kurs zu Verfügung. Andrea Heck von der AFRA GmbH überlässt mir das Skript zum Testmanagement zur Veröffentlichung. Alexander Barth von TIME'S UP erlaubt mir, das Skript zum interaktiven Environment Body Spin zu publizieren. KommilitonInnen tragen mit Seminarunterlagen und Prüfungsprotokollen zur Materialsammlung bei.

Informatik- netzwerk

Ein Dankeschön an die Ladies: Zuerst an Veronika Oechtering, die die Informatica Feminale in Bremen organisiert. Dort begann meine Themensuche. Dann danke ich Brigitte Jellinek, die mich heftig mit dem Informatik-Virus infizierte. Meine Mentorin Dr. Angelika Lukat von der Fraunhofer Gesellschaft hat mich bei der leidigen Themensuche ermuntert und mich mit ihren Erfahrungen unterstützt. Barbara Roth vom Mentoringprojekt MUFFIN 21 hat sich als strenge Korrektorin dieser Arbeit verdient gemacht. Außerdem gehören unbedingt noch hierher: Die Mädels vom admin@-Projekt (Universität Hamburg), die mich in die Geheimnisse von UML einweihen, nämlich Heike Wagner, Beate Orłowski und Irina L. Marinescu.

Unterstützung

Mein Bruder Georg hat mir damals meinen ersten Rechner geschenkt. Mit ihm kann ich über Bits, Bytes und Bugs diskutieren. Er hat im Verlauf der Realisation konstruktiv die Website kritisiert und sie getestet.

Besonderen Dank verdient sich mein Lebensgefährte Sven, der meine gelegentliche Frustration aus nächster Nähe ertragen musste und trotzdem immer davon überzeugt war, dass diese Arbeit einmal fertig sein wird ...

Inhaltsverzeichnis

Vorwort	2
Inhaltsverzeichnis	3
Abbildungsverzeichnis	6
Tabellenverzeichnis	7
Abkürzungen	8
1. Einleitung	9
1.1 Motivation	9
1.1.1 Geschichte des WWW	9
1.1.2 Schlechte und gute Beispiele für Websites	13
1.2 Vorgehensweise	21
1.3 Problembeschreibung	22
2. Anforderungen	25
2.1 Anforderungen und Skalierbarkeit	25
2.1.1 Templates	25
2.1.2 Flexibel erweiterbare Struktur	26
2.1.3 Erstellung und Pflege	27
2.2 Analyse	28
2.2.1 Übersicht über die Bereiche	28
2.2.2 Übersicht über die Fachgebiete	28
2.2.3 Übersicht über die Dateien	29
2.3 Theoretische Grundlage: UML	29
3. Funktionalität	30
3.1 UML-Diagramme	30
3.1.1 Welche Funktionen können mit welchen Diagrammen dargestellt werden?	30
3.1.2 Use-Case-Diagramme	31
3.1.3 Klassendiagramme	34
3.1.4 Sequenzdiagramme	36
3.1.5 Zustandsdiagramm	37
3.1.6 Bewertung und Begründung der Auswahl	38

3.2 Metadaten – Möglichkeiten des Findens	39
3.2.1 Dublin-Core-Standard und Metadatenklassifizierung	39
3.2.2 XMI – XML Metadata Interchange	41
3.2.3 Metatags in HTML	42
3.2.4 Weitere Formate	43
3.2.5 Bewertung und Auswahl	44
3.3 Datenbank und Web Application Server	45
3.3.1 DB-Anbindung	45
3.3.2 Ausgabeinterface	45
3.3.3 Admininterface: Änderungs- und Eingabeinterface	46
3.3.4 Siteadmininterface	49
3.3.5 Siteweites Suchen und Finden	50
3.3.6 Suchergebnisse sammeln	50
3.4 Machbarkeit	50
4. Konzept und Umsetzung	53
4.1 Szenarien	53
4.2 Strukturierungsmöglichkeiten	54
4.2.1 Indexierung	54
4.2.2 Suche	54
4.3 Basissoftware	55
4.3.1 Software für Redakteure	55
4.3.2 Software für Webdesigner	56
4.3.3 Software für Entwickler und Programmierer	58
4.3.4 Software für den Besucher	59
4.4 Dateiformate und ihre Aufgaben	60
4.4.1 Design, Entwicklung und Realisierung	60
4.4.2 Redaktion	63
4.5 Layout und Design	64

5. Realisierung und Dokumentation	67
5.1 Sitestruktur und Inhaltsübersicht	68
5.2 Datenbankstruktur	71
5.3 Funktion und Besonderheiten des Web Application Servers	72
5.4 Styleguide	74
5.5 HTML-Seiten	79
5.6 Datenimport	80
5.6.1 Zugriff gewähren	80
5.6.2 Konvertierung der vorhandenen Materialien	82
6. Fazit	84
6.1 Zusammenfassung	84
6.2 Ausblick	84
6.3 Zukunftstrends	85
Anhang	87
A. Quellenverzeichnis	88
B. Materialien	95
C. Dado-Befehlsverzeichnis	100
D. Datenbankstruktur	110
E. Browser-Screenshots	115
F. Glossar	121

Abbildungsverzeichnis

Abb. 1.1	Lynx – ein textbasierter Browser	10
Abb. 1.2	Aufbau des URL	12
Abb. 1.3	Fehlendes Bild und unbeabsichtigte Hintergrundkachelung	14
Abb. 1.4	Schlichte Startseite	15
Abb. 1.5	Fehlerhafte Darstellung in Opera	16
Abb. 1.6	Farbige Texte sind keine Links	16
Abb. 1.7	Farbige Texte sind Links	17
Abb. 1.8	Redundanz für verschiedene Wahrnehmungstypen	17
Abb. 1.9	Animiertes Steinschlag-Schild	18
Abb. 1.10	Feature Druckversion	19
Abb. 1.11	URL-Angabe: So findet man die Seite leicht wieder	19
Abb. 1.12	Der Besucher kann selbst entscheiden, ob die Information für ihn aktuell ist	20
Abb. 3.1	Use-Case-Diagramm "Engineering"	31
Abb. 3.2	Use-Case-Diagramm "Content"	32
Abb. 3.3	Use-Case-Diagramm "Schnittstelle zwischen Content und Engineering"	33
Abb. 3.4	Klassendiagramm "Interfaces" – Komponenten	34
Abb. 3.5	Klassendiagramm "Materials"	35
Abb. 3.6	Sequenzdiagramm "Edit content"	36
Abb. 3.7	Sequenzdiagramm "Search"	37
Abb. 3.8	Zustandsdiagramm "Search"	38
Abb. 3.9	MCF-Screenshot: 3D-Darstellung von Webinhalt	44
Abb. 3.10	Screenshot der Seite "Skript" in der Ausgabeansicht	46
Abb. 3.11	Editiersymbol im Admininterface	46
Abb. 3.12	Screenshot der Seite "Skript" in der Änderungsansicht	47
Abb. 3.13	Screenshot zur Änderung der Metadaten	48
Abb. 3.14	Screenshot der Datenbanktabelle "dictionary_en2de" für das Wörterbuch	49
Abb. 4.1	Anti-Alias-Effekt	62
Abb. 4.2	jpg-Artefakte	62
Abb. 4.3	Seite "Quellenverzeichnis" mit Fußzeile	64
Abb. 4.4	Design einer Skriptseite	65
Abb. 4.5	Farbwechsel im Design?	66
Abb. 5.1	Unsichtbarer Aufbau der Sitemap	70
Abb. 5.2	Grobe Skizze für das neue Design	74
Abb. 5.3	Banner der Website	77
Abb. 5.4	Zuordnung der Einträge aus der Textdatei zu den Tabellenfeldern der DB	83

Tabellenverzeichnis

Tabelle 1.1	Browser im Jahr 2002 (nicht vollständig)	11
Tabelle 1.2	Generic-Top-Level-Domains	12
Tabelle 3.1	Machbarkeitsübersicht	50
Tabelle 5.1	Ordnerstruktur	69
Tabelle 5.2	Farben der Website	75
Tabelle 5.3	Grafiken zur Navigation	75
Tabelle 5.4	Grafiken zu Grundfunktionen und für das Layout	76
Tabelle 5.5	Grafiken zur Interaktion	76
Tabelle 5.6	Symbole im Admininterface	77
Tabelle 5.7	Style Sheets für die Online-Ausgabe	78
Tabelle 5.8	Style Sheets für die Druckversion	78
Tabelle 5.9	Zugriffsrechte	80
Tabelle 5.10	Benutzer und Gruppen	81
Tabelle 5.11	Gruppen und Verzeichnisse	81

Abkürzungen

Abb.	Abbildung
bzw.	beziehungsweise
d. h.	das heißt
d. i.	das ist
Dado	Dado Application Server
DB	Datenbank
engl.	englisch
etc.	et cetera
evtl.	eventuell
f	folgende
ff	(mehrere) folgende
i. a.	im Allgemeinen
i. d. R.	in der Regel
kB	Kilobyte
o. ä.	oder ähnliches
S.	Seite
s.	siehe
Tab.	Tabelle
u. a.	und andere, unter anderem
usw.	und so weiter
vgl.	vergleiche
vs.	versus (im Gegensatz zu)
z. B.	zum Beispiel
z. Zt.	zur Zeit

1. EINLEITUNG

1.1 Motivation

Ziel der Arbeit

Die vorliegende Arbeit befasst sich mit dem Problem, Lehrmaterialien über das WWW zu publizieren. Die Materialien stammen aus den Bereichen Mathematik, Informatik und Kunst. Die Materialsammlung besteht bereits in Form eines statischen Webangebotes, das im Rahmen dieser Arbeit als Grundlage dient. Eine Dynamisierung soll den weiteren Ausbau der Sammlung erleichtern. Mittels Datenbankanbindung und Web Application Server wird eine dynamische Website erreicht. Die Besucher sollen Inhalte einfach und schnell finden. Einige komfortable Zusatzfunktionen können die Akzeptanz der Website erhöhen. Um ein Bild möglicher Funktionen zu erhalten, werden zunächst die Geschichte des World Wide Web und einige Websites dargestellt.

1.1.1

Geschichte des WWW

Leben mit dem WWW

Das World Wide Web ist mittlerweile selbstverständlich Teil unseres Lebens geworden. Es verbindet die ganze Welt miteinander und macht den Austausch von Informationen zu einem Kinderspiel. Die über das WWW bequem erreichbaren Informationen erleichtern das tägliche Leben. Zum Beispiel spart man durch Routenplanung, Fahrplaninformationen, den Online-Service von Bibliotheken und anderes wirklich viel Zeit. Kurz und gut: Die Vorzüge des WWW sind nicht mehr wegzudenken, schon gar nicht aus der Wissenschaft.

Tim Berners-Lee: Geburt des World Wide Web

1989 hat Tim Berners-Lee den ersten Plan für ein Informationsnetz. Wissenschaftliche Dokumente sollen online sichtbar sein, d. h. Texte müssen formatiert und Grafiken eingebunden werden können. Darüber hinaus möchte Berners-Lee Hypertextfunktionalität haben, eine tatsächliche Vernetzung der Daten. So entsteht der Name WWW – World Wide Web. Die Grundsteine des Projekts werden das Dateiformat HTML (Hypertext Markup Language) und das neue Internet-Protokoll HTTP (HyperText Transfer Protocol). Berners-Lee ersinnt auch das Konzept des URL [dfn-expo 2001]. Er ist zu dieser Zeit Informatiker am CERN (Europäisches Kernforschungszentrum). Der Brite verfeinert das WWW 1990 zusammen mit Robert Cailliau. Im November beginnt Nicola Pellow, eine Studentin am Leicester Polytechnic in Großbritannien, ihre Arbeit am textbasierten Browser [CERN 2001]. Neu bei Berners-Lees System ist, dass Dokumente plattformunabhängig, d. h. unabhängig von verschiedenen Computersystemen und unterschiedlicher Hardware, geöffnet werden können. Online-Betrachtung und Surfen [Rufin u. a. 2000] machen heute einen wesentlichen Teil des WWW aus.

Schon um Weihnachten 1990 demonstriert das CERN eine prototypische Software. Mittels eines Interfaces für das CERN Computer Center und die bekannten Usenet-Newsgroups erreicht man eine schnelle Akzeptanz. Alle Informationen werden auf diese Weise sofort zugänglich, und zwar über einen einfachen WWW-Browser [CERN 2001], [Münz 2001], [Cailliau 1995].

1991 können bereits Universitäten und Forschungseinrichtungen auf das Web zugreifen. Kurz darauf wird der Dienst über das Internet angeboten [CERN 1997].

Browser: Blick ins Netz

Zunächst gibt es zwei Browserarten. Einerseits existiert die Originalversion, die zwar sehr hoch entwickelt, aber nur für NeXT-Maschinen brauchbar ist. Auf der anderen Seite gibt es den textbasierten Browser. Dieser ist leicht zu installieren und läuft auf jeder Plattform. Ein solcher textbasierter Browser ist Lynx, der im Juli 1993 in der Version 2.0 erscheint. An der Universität von Kansas wird er von Lou Montulli, Michael Grobe, Charles Rezac und später auch Garrett Blythe entwickelt. Lynx ist heute noch im Einsatz, er wird u. a. von sehbehinderten und blinden Surfern und Webmastern genutzt. Denn durch die reine Textdarstellung vereinfacht sich der Aufwand, den Inhalt einer Webseite durch geeignete Software vorlesen zu lassen. Lynx unterstützt derzeit den Standard HTML 4.0 [Hänel 2001].

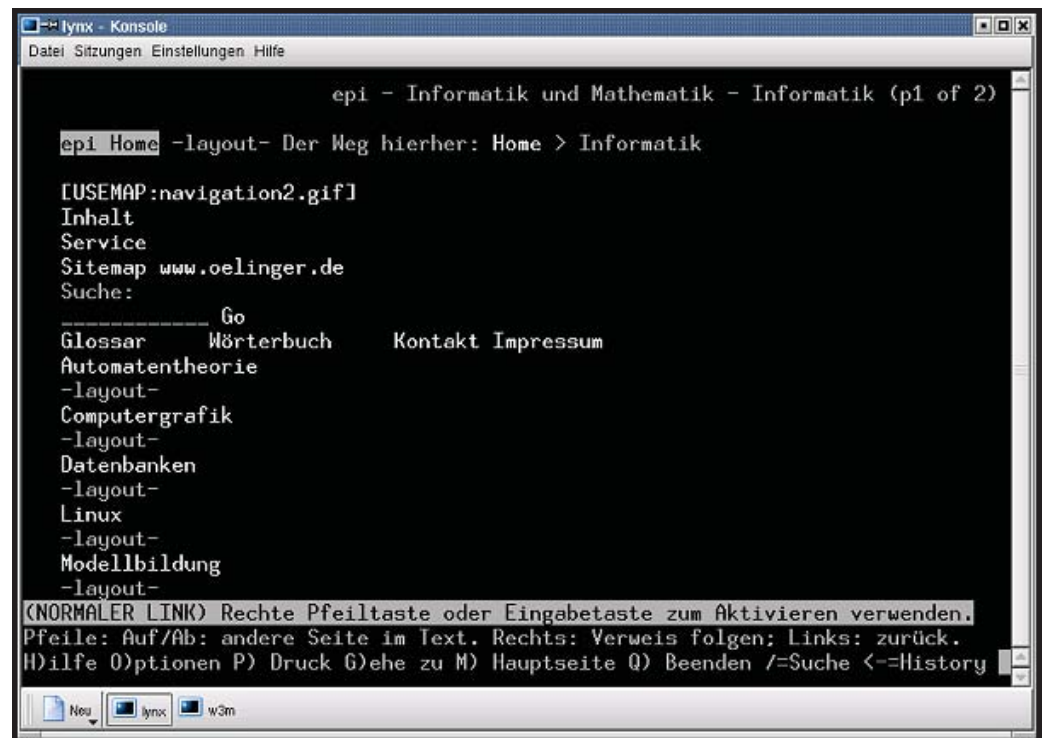


Abb. 1.1: Lynx – ein textbasierter Browser

Im Februar 1993 entwickelt Marc Andreessen am NCSA (National Center for Supercomputing Applications) der University of Illinois eine erste Version seines Mosaic-Browsers. Dessen grafische Oberfläche erleichtert den Umgang mit dem Web. Ein Jahr später wird Mosaic weltweit von über 2 Millionen Usern verwendet.

Im April 1994 geben Jim Clark (Mitgründer von Silicon Graphics) und Marc Andreessen bekannt, dass sie eine Firma namens "Netscape Communications" gründen wollen. Diese soll Software entwickeln, die die Nutzung des WWW vereinfacht. Im Dezember 1994 ist der Netscape Navigator erhältlich [Jung u. a. 1996]. 1996 erscheint der erste Opera-Browser [Opera 2001].

Im September 1995 wird von Microsoft im Zuge der Veröffentlichung von Windows 95 eine neue Browser-Variante entwickelt, der Internet Explorer. Seit November 1996 konkurrieren Netscape und Microsoft um die Spitzenposition auf dem Browsermarkt [Jung u. a. 1996].

Name	Hersteller / Autor	Betriebssysteme
emacs/W3	The GNU Project	Unix, Windows u. a.
Internet Explorer	Microsoft	Windows
Konqueror	KDE e. V.	Unix
Lynx	Lynx Org.	DOS, OS/2, Unix, Windows u. a.
Mozilla	Mozilla Org.	Mac, OS/2, Solaris, Unix, Windows u. a.
Netscape Navigator	Netscape	Mac, Unix, Windows u. a.
Opera	Opera Software	Mac, OS/2, Unix, Windows u. a.
w3m	Akinori Ito	Mac, MS-DOS, OS/2, Unix, Windows

Tabelle 1.1: Browser im Jahr 2002 (nicht vollständig)

Im September 1996 startet Cari D. Burstein die "Viewable With Any Browser"-Kampagne, die der Entwicklung der browserabhängigen HTML- und CSS-Umsetzung entgegenwirken will [Behme 1998]:

"Anyone who slaps a 'this page is best viewed with Browser X' label on a Web page appears to be yearning for the bad old days, before the Web, when you had very little chance of reading a document written on another computer, another word processor, or another network."

(Tim Berners-Lee in Technology Review, Jul. 1996)

Es ist ärgerlich, auf einer Website herauszufinden, dass man erst eingelassen wird, wenn man mit dem Netscape Navigator oder dem Internet Explorer wiederkommt. Genauso unerfreulich ist es, wenn man feststellen muss, dass browserspezifische Tags eingesetzt werden oder die Unterstützung für textorientierte Browser fehlt.

Cari D. Burstein will diesem Trend bewusst entgegentreten. Obwohl Websites nicht in allen Browsern gleich aussehen können, sollten sie in jedem Browser lesbar sein [Burstein 2000].

DENIC und immer mehr Webseiten

1991 gibt es an der Universität Dortmund den ersten freiwilligen Nameserver-Dienst DENIC (Deutsches Network Information Center). 1993 wird der Interessenverbund DENIC gegründet. Daraus wird 1996 die DENIC eG. 1999 benennt die ICANN (Internet Corporation for Assigned Names and Numbers), deren Hauptaufgabe in der Verwaltung der IP-Adressen und Top-Level-Domains liegt, Firmen, die Domain-Namen registrieren dürfen. Im Oktober 1999 übersteigt die Anzahl der registrierten Domains eine Million, im April darauf werden zwei Millionen registrierte Domainnamen erreicht. Im September 2000 erhöht sich die Domainzahl auf drei Millionen, im Februar 2001 auf vier Millionen, schon im November 2001 wird die fünf-millionste Domain vergeben [DENIC 2001]. Weltweit übersteigt die Größe des Web eine Milliarde indizierbare Webseiten. 2001 sind es über zwei Milliarden Seiten [dfn-expo 2001].

URL

Mit dem Konzept des Uniform Resource Locator (URL) ist es möglich, Webseiten eindeutig zu adressieren. Das Konzept baut auf dem Prinzip des

DNS (Domain Name System) von 1984 auf. Dieses wurde von Paul Mockapetris an der University of Southern California in Los Angeles entwickelt [Sillion 2001]. Der URL ermöglicht es, nicht nur auf lokale Dokumente zu verweisen, sondern auf Dateien, die sich irgendwo im Web befinden. Die Struktur ist wie folgt aufgebaut:

http	://	www.informatik.	oelinger	.de	/math/	index.html
Protokoll	Trennzeichen	Zusatz (www oder Subdomain)	Second-Level-Domain	Top-Level-Domain	Verzeichnis	Datei
Servername				Pfad		

Abb. 1.2: Aufbau des URL

Zuerst wird das Protokoll notiert. Für das WWW ist das http. Dann folgen ein Doppelpunkt und zwei Schrägstriche. Hinter diesem Trennzeichen schließt sich der Servername an, bestehend evtl. aus einer Subdomain sowie der Second-Level-Domain und der Top-Level-Domain. Im Servernamen werden die einzelnen Bestandteile jeweils durch einen Punkt getrennt. Zum Schluss findet sich der Pfad, bestehend aus keinem, einem oder mehreren Verzeichnissen und der Datei selbst. Die Trennung wird jeweils durch einen Schrägstrich angezeigt.

Ein weiteres bekanntes Protokoll ist ftp, das File-Transfer-Protocol. Dies wird u. a. dazu eingesetzt, die Daten einer Website (html-Dokumente, Bilder und andere Dateien) vom lokalen Rechner auf einen Server zu übertragen. https steht für die datengeschützte Übertragung mit dem Secure-Socket-Layer-Protokoll.

Alles, was zwischen dem Protokoll und der Second-Level-Domain steht, ist allein Sache des Inhabers der Domain. Der URL zum Fachbereich Informatik an der Universität Duisburg ist z. B. [Rufin u. a. 2000]:

`http://www.informatik.uni-duisburg.de/`

Top-Level-Domains

Top-Level-Domains sind entweder Country-Code-Domains oder Generic-Top-Level-Domains, die die Website einem bestimmten Bereich zuordnen. Eine interessante Top-Level-Domain ist .tv – die Endung wird zwar gern von Fernsehsendern genutzt, doch sie ist eigentlich eine Länderkennung, und zwar die von Tuvalu. Die Länderkennung sagt nichts darüber aus, in welchem Land sich der Server mit der entsprechenden Adresse befindet. Vor dem Jahr 2000 gab es lediglich sieben Generic-Top-Level-Domains. Seit 2000 gibt es noch einmal sieben neue [Sillion 2001].

alt	für	neu	für
.com	commercial – Wirtschaft	.biz	business – Unternehmen
.edu	education – Bildung	.info	(ohne Einschränkung)
.gov	government – Regierungsbehörden	.name	Privatpersonen