

Datenbank- und Informationsmanagement in der Historischen Sozialforschung: eine praxisorientierte Einführung

Pierau, Karl

Veröffentlichungsversion / Published Version

Themenheft / topical issue

Empfohlene Zitierung / Suggested Citation:

Pierau, K. (2002). Datenbank- und Informationsmanagement in der Historischen Sozialforschung: eine praxisorientierte Einführung. *Historical Social Research, Supplement*, 14, 1-221. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-285890>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:
<https://creativecommons.org/licenses/by/4.0>

HSR Supplement / Beiheft Nr. 14 (2002)

Karl Pierau

**Datenbank- und
Informationsmanagement in der
Historischen Sozialforschung.**

Eine praxisorientierte Einführung

Köln:
Zentrum für Historische Sozialforschung
2002

Inhaltsverzeichnis

Editorial	7
Vorwort.....	11
1. Grundbegriffe des Datenmanagements	13
1.1 Was sind Daten?	13
1.2 Was ist eine Datenbank?.....	15
1.3 Verhältnis der Datenbank zu anderer Software	17
1.4 Phasenkonzept des Datenbankentwurfs	19
1.4.1 Spezifikation/Anforderungsanalyse (Etappe 1)	20
1.4.2 Konzeptueller Entwurf (Etappe 2).....	20
1.4.2.1 Drei-Ebenen-Architektur	20
1.4.3 Logischer Entwurf (Etappe 3).....	21
1.4.4 Datendefinition (Etappe 4).....	22
1.4.5 Physischer Entwurf (Etappe 5)	22
1.4.6 Implementierung und Wartung (Etappe 6)	22
1.5 Schematische Darstellung des Phasenkonzepts	23
2. Anforderungsanalyse / Informationsbedarfsanalyse	26
2.1 Ziele und Aufgaben der Anforderungsanalyse	26
2.2 Funktioneller Entwurf.....	28
2.3 Die erste Phase des Datenbankentwurfsprozesses	29
2.4 BIOSOP-Normalbiographie.....	31
2.5. Studienbeschreibung: BIOSOP-Projekt	41
3. Datenbankmodelle	43
3.1 Grundlagen von Datenbankmodellen	43
3.2 Das relationale Datenmodell.....	45
3.2.1 Grundlagen des relationalen Modells	45
3.2.1.1 Was ist eine 'Relation'?	46
3.2.1.2 Definition „Relationale Datenbank“	48
3.2.1.3 Integritätsbedingungen.....	49
3.2.1.4 Zusammenfassung der Eigenschaften von Relationen:.....	51
3.2.2 Relationale Algebra	51
3.2.2.1 Operation SELECTION – Auswahl von Zeilen aus der Tabelle	51
3.2.2.2 Operation PROJECTION – Auswahl von Spalten der Tabelle	52
3.2.2.3 Operation JOIN – Verbund.....	53
3.3 Zusammenfassung	54
4. Entity Relationship Model	56
4.1 Grundkonzepte des klassischen ER-Modells	56
4.1.1 Werte	57

4.1.2 Entities	57
4.1.3 Beziehungen	57
4.1.4 Attribute.....	58
4.1.5 Zweistellige versus mehrstellige Notation	59
4.2 Weitere Konzepte im ERM.....	60
4.2.2 Die IST-Beziehung	61
4.2.3 Kardinalitäten	61
5. Darstellung der Parlamentarier durch Konzepte des ERM	63
5.1 Was ist unsere 'Miniwelt'?	63
5.2 Entity-Typ Reichstag	63
5.3 Entity-Typ Landtag.....	66
5.4 Entity-Typ Parlament.....	67
5.5 Entity-Typ Berufsausbildung.....	67
5.6 Entity-Typ Publikationen.....	69
5.7 Entity-Typ Parteitagsteilnahme	69
5.8 Entity-Typ Werdegang	70
5.9 Entity-Typ Personen	71
5.10 Weitere Beziehungen im Überblick.....	72
6. Abbildung des ER-Diagramms auf das relationale Modell.....	76
6.1 Entity-Typen	77
6.2 Beziehungstypen.....	79
6.3 Zusammenfassung von Tabellen bei zwingenden Beziehungen	79
7. Normalformen.....	87
7.1 1. Normalform	89
7.2 2. Normalform	89
7.3 3. Normalform	90
8. Einführung in ACCESS	92
8.1 Einleitung	92
8.2 Arbeiten mit ACCESS.....	95
8.2.1 ACCESS starten und beenden	95
8.2.2 ACCESS lässt sich nicht beenden.....	96
8.3 Das Anwendungsfenster	96
8.3.1 Vorhandene Datenbank öffnen	97
8.3.2 Weitere Möglichkeiten eine Datenbank zu öffnen.....	98
8.4 Das Datenbankfenster	98
8.5 Tabellen	100
8.5.1 Neue Tabelle erstellen	100
8.5.2 Felder in die Tabelle einfügen	102
8.5.3 Übersicht über die wichtigsten Feldeigenschaften.....	103
8.5.4 Primärschlüssel festlegen.....	104
8.5.4.1 Selbst den Primärschlüssel festlegen	104

8.5.5	Tabelle schließen und speichern	104
8.6	Formulare	107
8.6.1	Formular öffnen	107
8.6.2	Wozu dienen Formulare?.....	108
8.6.3	Vorteile von Formularen.....	108
8.6.4	Cursorsteuerung in einem Formular mit der Tastatur	109
8.6.5	Erstellen eines Formulars.....	109
8.6.6	Verändern eines bestehenden Formulars	114
8.7	Suchen, Sortieren, Filtern	120
8.7.1	Suchen und Ersetzen.....	120
8.7.2	Sortieren.....	121
8.7.3	Filter	122
8.7.3.1	Auswahlbasierter Filter.....	122
8.7.3.2	Auswahlausschließender Filter	123
8.7.3.3	Formularbasierter Filter	123
8.8	Abfragen	125
8.8.1	Auswahlabfrage manuell erstellen	126
8.8.1.1	Felder in die Abfrage aufnehmen.....	128
8.8.1.2	Sortierreihenfolge der Abfrage festlegen	128
8.8.1.3	Felder aus der Abfrage entfernen.....	129
8.8.2	Abfrage ausführen	129
8.8.3	Abfrage speichern	130
8.8.4	Tabellen verknüpfen	130
8.8.4.1	Eine Verbindung zwischen den Tabellen herstellen	131
8.8.4.2	Vordefinierte Beziehungen	133
8.8.5	UND- und ODER-Verknüpfungen	134
8.8.6	Wichtige Regeln zur Formulierung von Abfragen.....	137
8.8.7	Logische Operatoren.....	140
8.8.8	Berechnete Felder	140
8.8.9	Mathematische Operatoren	141
8.8.10	Anzeigeformat einstellen	142
8.8.11	Textfelder aneinanderfügen	142
8.9	Komplexe Beispiele.....	143
8.9.1	Der Ausdrucksgenerator	147
9.	Auswertung der Datenbank.....	153
9.1	Übung 1: Einfache Abfragen – Trennen von Informationen	153
9.2	Übung 2: Verteilung des höchsten Bildungsabschlusses	162
9.3	Übung 3: Soziale Herkunft der Abgeordneten.....	170
9.4	Übung 4: Ausgeübte Berufe vor Mandatsantritt	172
9.4.1	Ergebnistabellen	175
9.5	Übung 5: Mandatsdauer in Abhängigkeit vom Bildungsniveau und von der sozialen Herkunft.....	179
9.5.1	Mandatsdauer und Bildungsniveau.....	182
9.5.2	Mandatsdauer und soziale Herkunft	184

9.6 Übung 6: Kreuztabelle – Bildungsniveau und soziale Herkunft.....	186
9.6.1 Alter bei Erstmandat und Geschlecht.....	187
9.6.2 Soziale Herkunft und Bildungsniveau	188
Abkürzungen.....	193
Referenzliteratur	194
Anhang A: Datenbankstruktur der Datenbank „Sozialdemokratische Parlamentarier in den deutschen Reichs- und Landtagen“ – Parl-SPD.....	195
1. Tabellen, die nur zu Übungszwecken erstellt wurden.....	195
2. Tabellen der Datenbank „Sozialdemokratische Parlamentarier in den Deutschen Reichs- und Landtagen“	197
3. Beziehungen zwischen den Tabellen	213
Anhang B: Übersicht: Übungsaufgaben im Text	217
Abbildungs- und Tabellenverzeichnis	221

Editorial

Zu den Serviceaufgaben des Zentrums für Historische Sozialforschung (ZHSF) zählt auftragsgemäß die Vermittlung von Grundlagenwissen zur Historischen Sozialforschung für graduierte Forscher (weniger für Studenten) verschiedenster Ausbildungs- und Qualifikationsniveaus. Dies geschieht am ZHSF seit 1980 im Rahmen der Herbstseminar-Grundkurse. Darüber hinaus werden in den Aufbaukursen Fortbildungsangebote zur Vertiefung der Fertigkeiten für die Absolventen der Grundkurse oder für Forscher mit entsprechenden Grundkenntnissen angeboten. Das ZHSF-Herbstseminar „Methoden der Historischen Sozialforschung“ hat nicht zuletzt durch den Deutschen Wissenschaftsrat bei den bisherigen Evaluationen höchste Anerkennung gefunden. Inzwischen haben weit mehr als 1.000 Teilnehmer das Herbstseminar erfolgreich absolviert. Insgesamt sind im Zusammenhang mit dem Herbstseminar bislang 16 Skripten in Form von HSR-Supplementheften oder HSR-Schwerpunktheften in einer Gesamtauflage von mehr als 20.000 Exemplaren erschienen. In dieser bewährten Tradition steht auch das vorliegende Skript „Datenbank- und Informationsmanagement in der Historischen Sozialforschung“ von Karl Pierau. Das Skript entstand im *Kontext der Herbstseminare 2001 und 2002*, die neben den Vor-Ort-Modulen auch erstmals virtuelle Module anboten. Zum Stellenwert des Skripts im Rahmen des teil-virtualisierten Herbstseminars bedarf es einiger Erläuterungen.

Die Herbstseminare sind mittlerweile eine feste Institution geworden und werden seit Jahren gut nachgefragt. Der Erfolg beruht zu einem erheblichen Teil auf dem *Workshop-Charakter der Seminare* vor Ort in Köln. Zwei Wochen intensiver, gemeinsamer Beschäftigung mit den Themen Methoden der Historischen Sozialforschung, Statistik, Datenbanken und SPSS gelten gleichsam als ein „Bildungs-Event“ und besitzen somit auch eine starke soziale Komponente.

Ein Hauptproblem, mit dem fast jedes Herbstseminar zu kämpfen hatte, war das *unterschiedliche Vorwissen der Teilnehmer*. Dieses reichte von Null bis hin zu einem hohen Niveau bei Teilnehmern, die schon umfangreiche, quantitativ orientierte Projekte bearbeitet hatten. Um überhaupt die praktischen Übungen des Workshops sinnvoll durchführen zu können, ist es also unabdingbar, die Teilnehmer zumindest auf ein minimales gemeinsames Kenntnisniveau zu bringen. Dies geschah bislang aus Zeitmangel vorwiegend durch Frontalunterricht vor Ort, der im Extremfall die Anfänger überforderte und die Semi-Profis langweilte. Zudem beanspruchten die Grundvorlesungen mindestens ein Drittel der Seminarzeit, was deutlich zu Lasten der Praxisübungen im Computer-Pool ging. In den Seminaren (z.B. in den Aufbaukursen) jedoch, in denen die Teil-

nehmer schon ein Mindestmaß an Vorkenntnissen mitbrachten, und daher die relativ kurzen Einführungsvorlesungen lediglich repetitiven Charakter hatten, entwickelte sich in der Regel schnell eine kreative Workshop-Arbeitsatmosphäre.

Ein weiteres Hauptproblem stellt die *lange Zeit von zwei Wochen Workshop* dar. Denn das Hauptklientel der Herbstseminare ist durch zwei Charakteristika geprägt: Zeit- und nicht zuletzt Geldmangel. Studenten, Doktoranden und Wissenschaftliche Hilfskräfte zählen nicht gerade zu den Großverdienern in der Akademischen Landschaft und sind zudem zwischen Familie, Beruf und Examen oder Promotion extrem eingespannt. Die Kosten, die durch die Unterbringung in Köln verursacht werden und der zeitliche Aufwand von zwei Wochen schreckte viele potentiellen Teilnehmer schon im Vorfeld ab.

Durch den *Einsatz der virtuellen Lehre* lassen sich die genannten Hauptprobleme zumindest teilweise lösen. Insbesondere die Zeit- und Ortsunabhängigkeit des Lernens sowie die individuelle Betreuung durch E-Mail und Diskussionslisten, die zugleich Feedback für die Dozenten sind, können dazu genutzt werden, den Workshop effektiver zu gestalten. Denn der enorme Vorteil des sozialen Miteinanders im Workshop, der eine große Stärke des Herbstseminars ist, soll auf jeden Fall erhalten bleiben. Dies ist in einem vollständig virtuellen Seminar schlechterdings unmöglich.

Ein „*Virtuelles Herbstseminar*“* ist deshalb nur sinnvoll als *Ergänzung zum Herbstseminar „vor Ort“*. Es soll den Vorlesungsteil nach Möglichkeit ersetzen und somit als Propädeutikum / Repetitorium zur Historischen Sozialforschung dienen und das Herbstseminar entlasten, indem es die Teilnehmer dazu auffordert, sich schon vor Seminarbeginn mit dem Thema systematisch auseinanderzusetzen sowie erste Übungsaufgaben virtuell zu bearbeiten. Sie müssen dies jedoch nicht vollständig isoliert tun, sondern haben die Möglichkeit, sich per E-Mail und Diskussionsforum oder Chat an die anderen Teilnehmer und auch an die Dozenten zu wenden.

Dadurch soll das bisher sehr unterschiedliche Basiswissen zur Historischen Sozialforschung, das die Teilnehmer mitbringen, auf ein gemeinsames Niveau gebracht werden. Den Teilnehmern wird so auch ermöglicht, eine realistische Erwartungshaltung für das Herbstseminar zu entwickeln. Die Dozenten sollen durch das Virtuelle Seminar von der Aufgabe entlastet werden, die elementaren Grundlagen von Datenbanken, Statistikanalyse und Forschungsmethodik immer wieder vortragen zu müssen. Diese Zeit können sie besser dazu einsetzen, die Workshop-Teilnehmer beim Selbstlernen im Propädeutikum zu unterstützen und mehr Coach als Dozent zu sein. Auf diese Weise soll das Herbstseminar

* Ein erster Erfahrungsbericht zum virtuellen Herbstseminar findet sich im Beitrag von Jan Schminder (HSR 26, 2001, Heft 1, S. 124-144; oder im Internet unter: <http://www.zhsf.uni-koeln.de/>).

nar vor Ort wieder mehr an „Workshop-Charakter“ gewinnen, da dies für Teilnehmer und Dozenten fruchtbarer ist als ein dichtgedrängtes Programm von Grundvorlesungen.

Das *Gesamtbasismodul des ZHSF-Herbstseminars 2002* umfasste demgemäß einen zweitägigen Vor-Ort-Workshop zum Einstieg, ein dreimonatiges virtuelles Basismodul, zwei viertägige Vor-Ort-Basismodule und ein zweimonatiges virtuelles Follow-Up zur Nachbereitung. Dabei wurden folgende Lerninhalte (die Inhalte bezogen auf dieses Supplementheft sind kursiv hervorgehoben) vermittelt:

Virtuelles Basismodul I: „Allgemeine und Anwendungsbezogene Forschungsstrategie“

- Historische Sozialforschung: Wissenschaftlicher Standort; Theorie-, Hypothesen-, Begriffsbildung; Operationalisierung, Indikatorenbildung; Methodenauswahl; Quellenrecherche, Quellenauswahl, Historische Methode; Quellenverarbeitung; Quellenkritik Internet; Primäranalyse; Sekundäranalyse.
- Wissenschaftlicher Kontext: Empirische Geschichts- und Sozialforschung; Historische Fachinformatik.
- Exemplarische Forschungsmethoden: Generationenforschung und Kohortenanalyse; Lexikalische und Kollektive Biographik; Lebenslaufforschung; Retrospektive Interviews; Biographische Methode; Narrative Interviews; Oral History; Methoden der Eliten-, Parteien- und Wahlforschung.
- Exemplarische Forschung: Kollektive Biographik von Parlamentarierbiographien.
- *Propädeutikum „Informations- und Datenmanagement“.*
- Propädeutikum „Deskriptive Statistik“.

Vorort-Basismodul II: „Informations- und Datenmanagement“

- *Qualitative Datenverarbeitung;*
- *Datenbankentwurf, relationelles Datenmodell, Entity-Relationship-Modell, Datenerfassung;*
- *Anwendung des Datenbankmanagementsystems ACCESS.*

Vorort-Basismodul III: „Deskriptive Statistik“

- Quantitative Datenverarbeitung;
- uni- und bivariate Häufigkeitsverteilungen;
- Maßzahlen zur Beschreibung univariater Verteilungen;
- Korrelation, lineare Einfachregression;
- Varianzanalyse;
- Stichprobenverfahren;

- Drittvariablenkontrolle in der Tabellenanalyse;
- Anwendung des Statistikanalysepakets SPSS.

Eines der *Hauptlernziele* des Herbstseminars besteht darin: den theoretischen, empirischen und technischen Ablauf eines Forschungsprojektes allgemein und *exemplarisch* von der Theorie bis hin zur Analyse und Darstellung umzusetzen. Als *exemplarisches Forschungsprojekt* diente in den Basismodulen des Herbstseminars 2002 das im HSR-Supplementheft 11 (W. Schröder / W. Weege / M. Zech; Köln 2000) u.a. vorgestellte BIOSOP-Forschungsprojekt („Sozialdemokratische Parlamentarier in den Deutschen Reichs- und Landtagen 1867-1933“). Zu diesem BIOSOP-Projekt lagen im Herbstseminar vor:

- ein gedrucktes Handbuch (mit Biographien, Chronik, Wahldokumentation; Düsseldorf 1995),
- eine ACCESS-Datenbank (mit qualitativen biographischen Texten) und
- ein SPSS-Datensatz (mit quantitativen biographischen Daten).

Diese drei Projektprodukte bildeten die *Informations- und Datenbasis für das Herbstseminar* und standen den Teilnehmern eingeschränkt online / virtuell und vollständig in der Vor-Ort-Veranstaltung in Köln zur Nutzung zur Verfügung.

Den Herbstseminarteilnehmern stand virtuell jeweils eine umfassendes Skript zu den drei Bereichen „Methoden“, „Datenbanken“ und „Statistik“ zur Verfügung; alle drei *virtuellen Skripten* sind praxisorientierte Einführungen vornehmlich am Beispiel des BIOSOP-Projektes. Während zumindest in herkömmlicher Form für die Bereiche „Methoden“ und „Statistik“ gedruckte Skripten aus früheren Herbstseminaren schon vorliegen, fehlte bislang für den Bereich „Datenbanken“ ein gedrucktes Skript. Das vorliegende Supplementheft von Karl Pierau füllt nun diese Lücke, ist aber zugleich auch das erste Printprodukt aus dem Bestand der vorhandenen virtuellen Skripten. Mit den Supplementheften 11, 12 und 13 wurde begonnen, umfängliche Materialien / Texte zum jeweiligen Heft ergänzend im Internet anzubieten. Diese Tradition setzt das Supplementheft 14 notwendig fort, denn nur so lässt sich z.B. auf die exemplarische BIOSOP-Datenbank für Übungszwecke zurückgreifen. HSR-Supplement als Printprodukt und HSR-Transition (<http://www.hsr-trans.de>) als Online-Produkt ergänzen sich dabei sinnvollerweise.

Köln, im Oktober 2002

Wilhelm Heinz Schröder

Vorwort

Die Datenbankproblematik wird von Max Vetter (1989) als das Jahrhundertproblem der Informatik bezeichnet. Er versteht darunter die Bewältigung der Datenflut, die von jeder modernen Gesellschaft produziert wird. Zur Lösung dieser Aufgabe fordert er von der Informatik die Schaffung von methodischen Grundlagen, die auch von Nichtinformatikern beherrscht werden können. Datenbanksysteme sind mittlerweile selbstverständliche Hilfsmittel in Wirtschaft, Verwaltung und Wissenschaft geworden. Ihre Anwendung setzt eine datenorientierte Vorgehensweise und Beschreibung relevanter Sachverhalte voraus.

Seit 1980 führt das ZHSF Seminare zu Methoden der Historischen Sozialforschung durch. Es entwickelten sich daraus die Kompaktveranstaltungen des ZHSF-Herbstseminars, die in mehreren Modulen für Anfänger und Fortgeschrittene in konzentrierter Form Kenntnisse vermitteln. Die Kursteilnehmer sollen die Lerninhalte am Beispiel von historischen Quellen und (im ZHSF archivierten) Datenbeständen aus der Historischen Sozialforschung in Arbeitsgruppen forschungspraktisch umsetzen, die Übungsdaten selbständig unter Einsatz von EDV auswerten und die Analyseergebnisse schriftlich darstellen.

Seit 1987 fanden auch Veranstaltungen zur Datenbankmethodik, seit 1996 in Form von mehrtägigen Workshops – begleitet von praktischen Computerübungen – statt, die inzwischen als Basismodul „Informations- und Datenmanagement“ fest in das Herbstseminar eingebunden sind. Parallel zu Vorträgen zur Vermittlung grundlegender theoretischer Aspekte wurden praktische Übungen mit dem Datenbankmanagementsystem ACCESS durchgeführt.

Das vorliegende Skript bildete die Grundlage für das Basismodul II „Informations- und Datenmanagement“ der Methodenausbildung im Rahmen des ZHSF-Herbstseminars. Praxisnah wird an einem Beispiel aus der Historischen Sozialforschung (Parlamentarismusforschung) an Methoden von Datenbankentwurf und Datenbankmanagement herangeführt. Notwendigerweise muss mit einem Datenbankmanagementprogramm gearbeitet werden. In den Herbstseminarkursen wurde MS ACCESS eingesetzt, weil es einen hohen Verbreitungsgrad besitzt und relativ leicht zu erlernen ist. Ausdrücklich möchte ich aber betonen, dass ich keine Einführung in ACCESS anbieten wollen, sondern das vorliegende Skript als eine Säule der Einführung in die Historische Sozialforschung sehe.

Die Beispiele im Skript beziehen sich im Wesentlichen auf das Projekt „Sozialdemokratische Parlamentarier in den deutschen Reichs- und Landtagen 1867 - 1933“ (BIOSOP). Es wird ein Forschungsprojekt simuliert, das sich die Aufgabe stellt, einen Beitrag zur Erforschung gesellschaftlicher Führungsgruppen zu leisten. Dabei geht es um die Bereitstellung einer Datenbasis, die als ein flexibel zu handhabendes Instrument für die weitere empirische Forschung dienen kann.

Dieses Skript folgt aus nahe liegenden Gründen nicht der Chronologie der Stoffvermittlung im Herbstseminar, sondern der chronologisch-methodischen Abfolge des Aufbaus einer Datenbank. Es empfiehlt sich nach dem Durcharbeiten des ersten Kapitels, „Grundbegriffe des Datenmanagement“, das Kapitel 8, „Einführung in ACCESS“ parallel zu den weiteren Kapiteln zu lesen. Die verfügbare Übungsdatenbank „Sozialdemokratische Parlamentarier“ kann von der Webseite von HSR-TRANS (<http://www.hsr-trans.de>) herunter geladen werden.

Als Ergänzung zu diesem Skript empfehle ich dem interessierten Leser die anspruchsvolle Einführung in das Thema von Heuer / Saake (1995). Auf die umfangreiche Literatur zu ACCESS soll nicht näher eingegangen werden. Ich möchte nur auf die Computereinführungen, die vom Herdt-Verlag für Bildungsmedien GmbH (<http://www.herd.de>) gemeinsam mit dem RRZN der Universität Hannover herausgegeben werden, hinweisen. Sie sind beim Herdt-Verlag für etwa 12 € oder in den Rechenzentren der meisten Universitäten für etwa 5 € zu beziehen.

Danken möchte ich an dieser Stelle noch allen Freunden und Kollegen, deren Hinweise und Tipps Berücksichtigung fanden. Mein besonderer Dank gilt Prof. Dr. Wilhelm H. Schröder und Prof. Dr. Horst Hollatz, die beide mir mit kritischen Anregungen zur Seite standen. Schließlich darf ich mich auch bedanken bei Philip Janssen, dem die gründliche Durchsicht des Heftes oblag, und Conny Baddack, die die sorgfältige Drucklegung des Heftes übernahm.

Köln, im Oktober 2002

Karl Pierau

1. Grundbegriffe des Datenmanagements

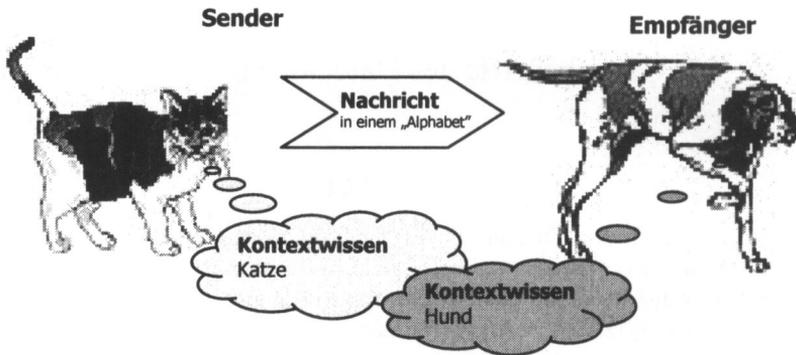
1.1 Was sind Daten?

Die Begriffe *Datum* (Mehrzahl Daten) und *Information* werden umgangssprachlich häufig synonym gebraucht. Leider findet man auch in wissenschaftlichen Publikationen irreführende oder widersprüchliche Definitionen. Daher sollen die Begriffe hier kurz erläutert werden.

Eine *Nachricht* setzt immer einen Sender der Nachricht und einen Empfänger voraus. Jede Nachricht ist in einem gewissen 'Alphabet' abgefasst, das dem Sender und dem Empfänger bekannt sein muss. Das können unser lateinisches Alphabet, Folgen von bits oder Bytes oder auch nur Gesten, Bilder oder sonstige Zeichen sein. Nachrichten sind lediglich eine Aneinanderreihung von Zeichen nach bestimmten syntaktischen Regeln.

Um die Nachricht zu verstehen, muss bei Sender und Empfänger ein sich deckendes *Kontextwissen* vorhanden sein. Durch Interpretation der Nachricht unter Berücksichtigung gewisser Regeln (Kontextwissen), die der Nachricht ihren bestimmten Sinn und Inhalt zuordnen, liefert die Nachricht eine Information. Eine Menge von Regeln und Vereinbarungen, die jeder syntaktisch zulässigen Nachricht ihre Bedeutung zuordnet bezeichnet man als *Semantik*. Nachricht und Information verhalten sich wie Form und Inhalt.

Bei natürlichsprachig abgefassten Nachrichten treten gelegentlich Missverständnisse auf. Das ist der Fall, wenn ein und dieselbe Nachricht durch unterschiedliche Interpretation verschiedene Informationen liefert: Das Kontextwissen von Sender und Empfänger der Nachricht stimmen nicht überein. Ein Beispiel aus der Biologie für die unterschiedliche Interpretation von Nachrichten (Gesten, Körperhaltungen) bietet die sprichwörtliche Feindschaft zwischen Hund und Katze. Die Körpersprache von Hund und Katze unterscheidet sich erheblich: Anstarren oder Schwanzwedeln hat bei beiden eine unterschiedliche Bedeutung. Katzen starren ihren Gegenüber an, was vom Hund als Dominanzgebärde gedeutet wird. Hunde verfügen als Rudeltiere über eine wesentlich differenziertere Körpersprache als Katzen. Für sie drückt Schwanzwedeln nicht nur Freundlichkeit aus. Das Schwanzpeitschen der Katze ist ein deutliches Angriffszeichen. Das sprichwörtliche „wie Hund und Katze“ findet zumindest teilweise eine informationstechnische Erklärung.



Zu ein und derselben Information können verschiedene, sie übertragende Nachrichten existieren, z.B. wenn die Nachrichten in unterschiedlicher natürlicher Sprache verbreitet werden.

Daten sind eine spezifische Form von Nachrichten, die vom Computer verwaltbar, verarbeitbar und transportierbar sind. Sie stellen als solche keinen Sinn dar, sondern sind nur formale endliche Folgen von Zeichen eines Alphabets, die durch Auftragen bestimmte Positionen eines Trägermediums (Papier, elektronisches Medium) nach bestimmten Regeln (Syntax) verändern. Wenn das zu Grunde liegende 'Alphabet' bekannt ist und ein Zusammenhang zu einem realen Sachverhalt hergestellt wird (Semantik), liefern die Daten auch eine *Information*.

Aufgabe 1.1: Versuchen Sie, sich den Unterschied zwischen den Begriffen 'Nachricht' und 'Information' anhand des folgenden Satzes klarzumachen:
 „Diese Nachricht gibt mir keinerlei Information.“
 Notieren Sie Ihre Gedanken.

1.2 Was ist eine Datenbank?

Datenbankmanagement-system	Unter einem Datenbankmanagementsystem (DBMS) versteht man ein Softwarepaket, das in der Lage ist, Anwenderdaten in einem Computer zu verwalten.
----------------------------	---

Bei einer solch einfachen Definition, wie man sie auch häufig in der Literatur findet, erhebt sich sofort die Frage, worin ein DBMS sich z.B. von einem Textverarbeitungssystem unterscheidet. Auch dort werden Anwenderdaten verwaltet, um sie in einer bestimmten Weise zu verarbeiten. Der entscheidende Punkt ist, was unter 'verwalten' verstanden wird. DBMS können Daten hinzufügen, verändern und löschen. Auch das wird von Textverarbeitungssystemen geleistet. Was DBMS diesen aber voraus haben, ist, dass sie auf wesentlich effektivere Weise Informationen verknüpfen und bereitstellen können. Sie sind dahingehend optimiert, dass sie dies auch mit sehr großen, langfristig zu haltenden Datenbeständen (persistente Daten) können.

Große Datenbestände stellen häufig einen beträchtlichen Wert dar. Sie müssen sicher zum Schutz vor einer versehentlichen Beschädigung verwaltet werden können (Datensicherheit) und vor mutwilliger Zerstörung oder dem Ausspähen geschützt sein (Datenschutz). Auch hier leisten DBMS wesentlich mehr als Textverarbeitungssysteme oder Kalkulationsprogramme.

Relationales Datenbankmanagement-system	Unter einem relationalen Datenbankmanagementsystem (RDBMS) versteht man außerdem, dass dieses DBMS für die Verwaltung der Daten intern eine bestimmte Technologie benutzt und dass es bestimmten Anforderungen genügt. Es verwaltet die Daten in einfachen Tabellen.
---	--

Die Daten selbst in ihrer verarbeitbaren Struktur stellen die *Datenbank (DB)* dar. Howe (1983) schrieb 1983 dazu: „Eine Datenbank (DB) ist eine Sammlung von nichtredundanten Daten, die von mehreren Applikationen benutzt werden.“

Datenbank-system	Ein Datenbanksystem (DBS) ist die Kombination von Datenbank und DBMS, also Daten und Programm gemeinsam.
------------------	--

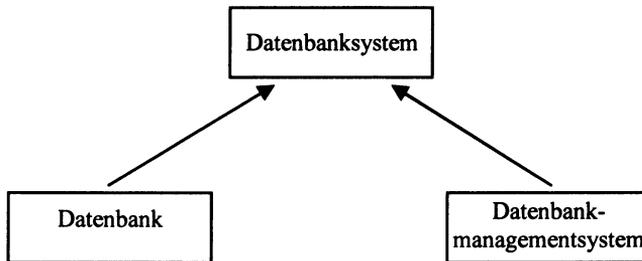


Abb. 1.1: Datenbankbegriff

Eine Datenbank ist die Beschreibung eines Sachverhaltes mit Daten. Gelegentlich findet man auch die Vorstellung, dass eine Datenbank ein Modell eines Ausschnitts der Realwelt sei. Der Modellbegriff wird in einem sehr vielfältigen Sinn gebraucht. Mathematische Modelle in den Naturwissenschaften besitzen die Möglichkeit der Erklärung. Darin unterscheidet sich der Modellbegriff der Softwaretechnik; sie implizieren keine erklärende Komponente, enthalten auch nicht notwendig strukturelle Analogien zum Referenzsystem.

Ebenfalls nicht haltbar ist die Vorstellung, Objekte bzw. Phänomene des Referenzsystems würden auf 'Objekte' des beschreibenden Systems abgebildet. Der Datenbankdesigner steht nicht unvoreingenommen einer realen Welt gegenüber, die es abzubilden gilt, sondern er hat es mit Begriffen, Vorstellungen, fachwissenschaftlichen Theorien zu tun. Gegenstände sind sprachrelativ konstituiert. Ihre Identität ist durch Bezugnahme innerhalb eines öffentlichen raumzeitlichen Bezugssystems gegeben.¹

Eine Beschreibung kann nicht wahr oder falsch, sondern nur mehr oder weniger angemessen sein. Ob unsere Datenbank den Erwartungen entspricht, hängt damit zusammen, wie es uns gelingt, unter Berücksichtigung der Aufgabenstellung mittels einer Anforderungsanalyse die Datenbank zu spezifizieren. In einer Folge von Schritten wird die Spezifikation, Resultat der Analyse, als eine noch informale Beschreibung, in eine formalisierte Darstellung, die Datenbank, überführt.

¹ Vgl. P. Schefe, Informatik Spektrum 22, 122-135.

1.3 Verhältnis der Datenbank zu anderer Software

Eine moderne Sicht auf Datenbanken ist in Heuer/Saake (1995) dargestellt. Software wird klassifiziert nach ihrer Bedeutung für den allgemeinen Betriebsablauf in einer Datenverarbeitungsanlage.

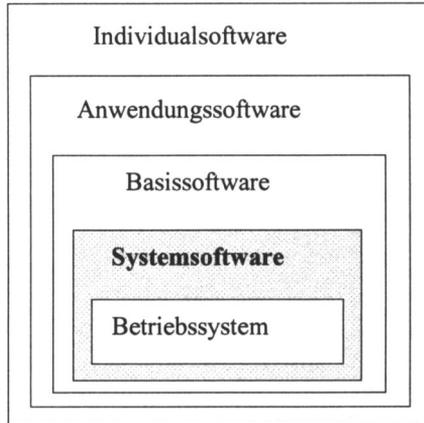


Abb. 1.2: Aufteilung in Software-Schichten nach Heuer/Saake (1995)

Typische Betriebssysteme sind etwa:

- Windows 98 und Windows ME, die beide auf dem Betriebssystem MS-DOS aus den frühen 80er Jahren basieren;
- Windows NT und dessen Weiterentwicklung Windows 2000, als 32-bit-Betriebssysteme von Microsoft für professionelle PC-Anwendungen entwickelt; der Kern dieses Systems ist als Nachfolger für die auf MS-DOS basierenden Windowssysteme vorgesehen;
- OS/2, ursprünglich gemeinsam von IBM und Microsoft konzipiert als Nachfolger für MS-DOS (bzw. PC-DOS, wie das von IBM im Kern gleiche Betriebssystem für IBM-Rechner bezeichnet wurde); eine Weiterentwicklung wurde inzwischen eingestellt;
- UNIX, ein professionelles System der mittleren Datentechnik; Linux ein frei benutzbares Unix-ähnliches System für PCs;
- Das Mainframe-Betriebssystem VMS.

Zur Systemsoftware, die direkt auf dem Betriebssystem aufbaut, zählen Datenbanksysteme und Benutzerschnittstellen (wie MS-Windows 3.1 oder so genannte X-Produkte – grafische Oberflächen für Unix-Systeme). Bekannte Datenbanksysteme (mit den entsprechenden Betriebssystemen) sind

- Oracle (Unix, Windows),
- Sybase (Unix),
- Ingres (Unix),
- Informix (Unix),
- Adabas (Mainframe, Unix, Windows),
- MS ACCESS (Windows),
- dBase (Windows),
- Foxpro (Windows),
- Paradox (Windows),
- MySQL (Linux, Windows), ein kostenfreies, empfehlenswertes Datenbanksystem.

Zur Basissoftware gehören Grafiksysteme und Textverarbeitungssysteme. Anwendungssoftware ist auf bestimmte Anwendungsklassen hin zugeschnitten, z.B. CAD-Systeme für Konstruktion, Buchhaltungssysteme und Lagerhaltungssysteme.

Datenbanken entwickeln sich zu einer Art „Über-Software“, die nicht nur Daten im herkömmlichen Sinne verwaltet, also Daten als Beschreibung von Objekten, sondern auch Informationen über die Beschreibungen vorhalten. Kurz gesagt handelt es sich um Daten über Daten, so genannte *Metadaten*. Beispielsweise könnten alle Textdokumente und Graphiken (allgemeiner: Dateien) durch ihre Metadaten in Datenbanken verwaltet werden. Sollen die eigentlichen Daten bearbeitet werden, übergibt die Datenbank die Steuerung an die entsprechende Basissoftware (Graphiksystem, Textverarbeitung).

Beispiel: *Stellung von Datenbanken im DV-System*

Für das BIOSOP-Projekt liegt folgendes Szenario vor:

Nach konzeptionellen Vorarbeiten beginnt eine Phase der intensiven Quellenrecherche. Wir nehmen an, es liegt eine Liste mit Namen von Abgeordneten vor. Diese durch ein DBMS verwaltete Liste wird fortwährend ergänzt. Mehrere Projektbearbeiter schreiben Anfragen an staatliche Archive, Zeitungsarchive, Meldeämter und private Informationsquellen, um Auskünfte über Abgeordnete zu bekommen. Es treffen Antworten ein, die direkt im Computer verwaltet werden (z.B. E-Mails) oder schriftliche Dokumente, die archiviert und datenbankgestützt katalogisiert werden. Weitere Anfragen können aus den eingegangenen Antworten resultieren. Alle diese Vorgänge, ihre Zusammenhänge, spiegeln sich in der Datenbank wider. Es kann nachvollzogen werden, wer wann an wen was geschrieben hat, zu welchen Abgeordneten von wem Informationen eingingen, bzw. wo Informationen aus welchen Gründen nicht bereitgestellt wurden.

Dabei steuert das DBMS andere Programme, z.B. indem über Hy-

perlinks in der Datenbank die Textverarbeitung aufgerufen wird, um bestimmte Dokumente anzuzeigen, oder es wird die Serienbrieffunktion aktiviert, um automatisch Briefe zu generieren. Die Abbildung 1.3 zeigt die Funktionen des DBMS schematisch.

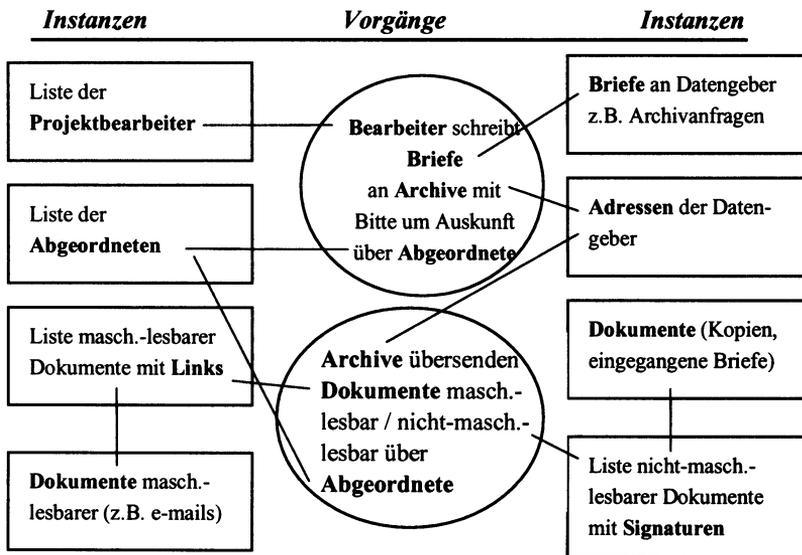


Abb. 1.3: Arbeitsabläufe und ihre Widerspiegelung in einer Datenbank

1.4 Phasenkonzept des Datenbankentwurfs

Der Weg von der theorieinduzierten Betrachtung des relevanten Ausschnitts der realen Welt zu einer Datenbank wird als Entwurfsprozess bezeichnet. Anfang der 70er Jahre befasste sich die Informatikforschung mit methodischen Aspekten der Softwareentwicklung. Eine sehr wesentliche Erkenntnis war, dass man eine Folge von Etappen bei der Herstellung von Software durchlaufen sollte, die nacheinander abzuarbeiten und parallel dazu zu dokumentieren sind. Diese Etappen sind unter dem Begriff *Phasenkonzept* in die Literatur eingegangen. Genauer müssen wir sagen, dass es sich nicht um ein Phasenkonzept, sondern um eine Klasse von ähnlichen Konzepten handelt. Für jede Phase wurden Methodiken und Arbeitsschritte entworfen.

Folgt der Entwurfsprozess diesem Phasenkonzept, so liefert er bis zur tatsächlichen Datenbankrealisierung zu jeder Phase ein Entwurfsdokument. Es

erfolgt eine Abbildung des Dokuments des vorigen Schrittes in das Dokument des aktuellen Schrittes.

Wir fordern, dass die Entwurfsschritte zwei Eigenschaften haben:

- Informationserhalt: Die folgende Datenbankbeschreibung enthält alle Informationen der vorhergehenden.
- Konsistenzerhaltung: Alle Regeln und Einschränkungen bleiben erhalten und keine werden hinzugefügt.

1.4.1 Spezifikation/Anforderungsanalyse (Etappe 1)

Die Datenbank soll einen Ausschnitt der realen Welt, die so genannte *'Miniwelt'* (*universe of discourse, Diskurs*) im Computer abbilden, d.h. durch Daten beschreiben. Zuerst muss der Teil der realen Welt, der im Sinne der Aufgabenstellung wichtig ist, informell beschrieben werden.

Aufgabe der Spezifikation/Anforderungsanalyse ist es

- Informationen,
- auf ihnen definierte Strukturen,
- Verarbeitungsanforderungen der auf den Informationsstrukturen existierenden und potentiellen DB-Anwendungen

zu sammeln und zu spezifizieren. Ergebnis dieses Schrittes ist eine informale Beschreibung des Fachproblems.

1.4.2 Konzeptueller Entwurf (Etappe 2)

Als Ergebnis des konzeptuellen Entwurfs erhält man eine formalisierte Beschreibung der Miniwelt, die noch frei von konkreten EDV-Systemen ist. Mitte der 70er Jahre wurde vom American National Standards Institute (ANSI) eine Gruppe von führenden amerikanischen Informatikern beauftragt, Standards für relationale Datenbankmanagementsysteme zu entwickeln. Sie arbeiteten einige Jahre an diesem Problem und kamen zu der Schlussfolgerung, dass es nicht sinnvoll ist, Standards für DBMS zu entwickeln. Sie schlugen jedoch in ihrem Bericht 1975 eine ganz bestimmte Herangehensweise beim Aufbau von Datenbanken vor.

1.4.2.1 Drei-Ebenen-Architektur

Wegen der Komplexität einer Datenbank ist es danach sinnvoll, die Daten in drei Schichten zu betrachten: externe, konzeptuelle und interne *Schicht*.

- Nutzer und Anwendungsprogrammierer sehen die Daten der Datenbank in nach außen gerichteten externen Schemata. Zusammengenommen bilden diese die externe Sicht. Jeder Nutzer hat eine individuelle, ihn interessierende Teilsicht (Benutzersichten).

- Die ganzheitliche Betrachtung der Struktur der gesamten Miniwelt bildet das konzeptuelle Schema. Es ist unabhängig davon, ob die Daten computergestützt verwaltet werden oder nicht (logische Gesamtsicht).
- Die interne Schicht behandelt die Datenschemata der Dateien, Datensätze und Zugriffspfade (physische Sicht).

Der Vorteil dieser Drei-Ebenen-Architektur besteht darin, dass kleine Änderungen von Benutzersichten keine oder nur geringe Veränderungen der physischen Datenspeicherung nach sich ziehen, die dann sehr aufwändig mit allen anderen Sichten abgestimmt werden müssten. Andererseits soll die Änderung der physischen Speicherung keinen Einfluss auf die Nutzersicht haben.

Das konzeptuelle Schema als logische Gesamtsicht, als soft- und hardware-unabhängige, mit mathematisch-kybernetischen Formalismen entworfene Sicht, ist relativ stabil gegenüber Veränderungen der Miniwelt. Veränderungen der Nutzersichten auf das konzeptuelle Schema zu übertragen, hat sich als relativ einfach erwiesen. Sind Modelländerungen notwendig, so lassen sie sich vom konzeptuellen Schema mit einfachen Regeln auf die interne Schicht übertragen.

Es gibt mehrere Ansätze zur Erzeugung einer solchen Gesamtsicht. Der bekannteste ist das *Entity Relationship Model* (ERM, Gegenstands-Beziehungs-Modell), das wir zur Herleitung einer geeigneten Datenstruktur für Parlamentarierdaten benutzen werden. Die Zusammenhänge werden durch ER-Diagramme dargestellt.

Es geht also darum, ausgehend von der Anforderungsanalyse, eine formale Beschreibung des interessierenden Gegenstandsbereichs vorzunehmen. Die Informationsstruktur ist in einem abstrakten und formalen Modell, z.B. in einem semantischen Datenmodell wie dem erweiterten Entity Relationship Model (EERM), vorzunehmen.

1.4.3 Logischer Entwurf (Etappe 3)

Das in der vorangehenden Phase entwickelte semantische Datenmodell ist in das Datenmodell des benutzten DBMS zu übertragen. Dabei sind programm-spezifische Besonderheiten des DBMS nicht zu berücksichtigen.

1. Schritt: Transformation des konzeptuellen Schemas (ER-Modell) in das Zieldatenbankmodell (z.B. relationales Modell). Dieser Schritt ist automatisierbar, d.h. es gibt geeignete Software, die die Entity-Relationship-Diagramme in Datenbankstrukturen umsetzen.
2. Schritt: Verbesserung des relationalen Schemas, z.B. durch Minimierung der Redundanzen (Normalisieren).

Als Ergebnis erhält man ein logisches Schema, eine Sammlung von Relationen, falls man ein relationales DBMS gewählt hat. Das ist eine Datenbankbeschreibung in einem Datenbankmodell, für das ein DBMS zur Verfügung steht.

1.4.4 Datendefinition (Etappe 4)

Anwendung des konkreten DBMS auf das logische Schema unter Verwendung der Datendefinitionssprache des DBMS (selbst SQL-Systeme unterscheiden sich in diversen Erweiterungen). Für das relationale Modell bedeutet das: Definieren der Wertebereiche, Relationen und Sichten.

1.4.5 Physischer Entwurf (Etappe 5)

Definition der internen Ebene mit einer Speicherstruktursprache, die z.B. festlegt, ob Zugriff auf eine Relation in Baumstruktur oder Hash-Tabelle organisiert ist, Definition von Indexen, usw. Das sind Aufgaben, die für PC-Datenbanken nicht relevant sind.

1.4.6 Implementierung und Wartung (Etappe 6)

Jedes über einen längeren Zeitraum existierende Softwareprodukt unterliegt Veränderungen, hervorgerufen durch veränderte Nutzungsbedingungen (Umgebungsbedingungen, Aufgabenstellung). Wenn man diese Änderungen zuerst in den konzeptuellen Entwurf einarbeitet und dann erneut die darauf folgenden Phasen durchlaufen werden, sind sie relativ gut beherrschbar.

1.5 Schematische Darstellung des Phasenkonzepts

Der Datenbankentwurfsprozess ist ein iterativer Vorgang, der nicht in strenger Folge die Etappen von 1 bis 6 durchläuft. Werden in einer späteren Etappe veränderte Nutzungsbedingungen festgestellt oder frühere Entwurfsentscheidungen verworfen, so wird zur Anforderungsanalyse zurückgesprungen und der Prozess erneut durchlaufen. Nur so ist zu sichern, dass von Etappe zu Etappe die Informationskapazität erhalten bleibt und nur das Abstraktionsniveau der Informationsdarstellung erhöht wird.

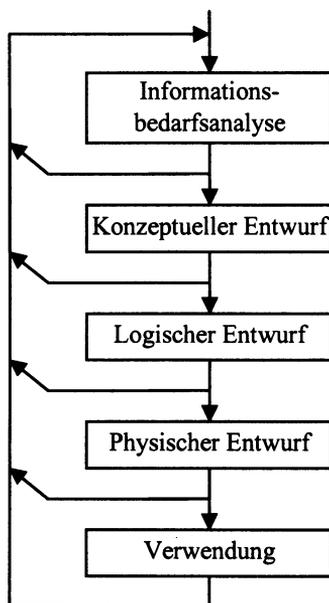


Abb. 1.4: Phasenkonzept nach Lockemann/Schmidt (1987)

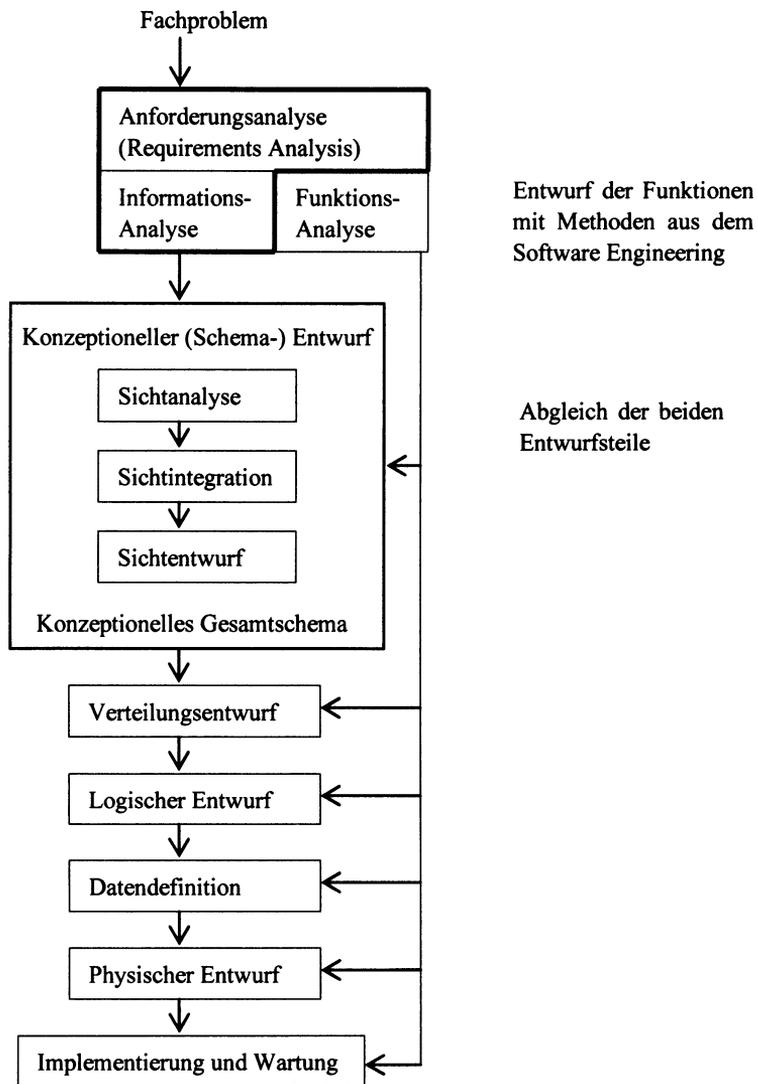


Abb. 1.5: Phasenkonzept nach Heuer/Saake (1995)

Aufgabe 1.2: Was ist ein Datenbankmanagement?
a. ein Datensatz
b. ein Computerprogramm
c. eine Datenbank mit dem dazugehörigen Datenbankverwaltungsprogramm

Aufgabe 1.3: Welche dieser Systeme klassifizieren Sie als Datenbankmanagementsystem?
a. Oracle
b. ACCESS
c. dBase
d. Excel

Aufgabe 1.4: Was sind Meta-Informationen?
a. Spezifische Informationen
b. Ein Maß für Informationsdichte
c. Informationen über Informationen
d. Fehlerhafte Daten

Wählen Sie die für Sie richtige Antwort aus und begründen Sie Ihre Entscheidung.

2. Anforderungsanalyse / Informationsbedarfsanalyse

2.1 Ziele und Aufgaben der Anforderungsanalyse

Der Entwurf einer Datenbank setzt eine wohldefinierte Aufgabenstellung voraus. Nicht weil zufälligerweise Daten zur Verfügung stehen, sondern weil ein bestimmtes, wie auch immer geartetes Interesse vorliegt, einen Ausschnitt der realen Welt zu beschreiben, ist die Datenbank anzulegen. In der DV-Praxis hat sich gezeigt, dass die Datenbank nach der Übergabe an den Auftraggeber häufig nicht die in sie gesetzten Erwartungen erfüllt. Daher sollte immer von einer schriftlich fixierten Beschreibung des Endprodukts ausgegangen werden. In der ersten Entwicklungsphase werden alle Objekte und Prozesse, die in der zu erstellenden Datenbank abgebildet werden sollen, informal beschrieben. Das setzt spezifische Fachkenntnisse über die realen Gegebenheiten voraus. Um alle für den Datenbankentwurf relevanten Zusammenhänge vollständig und ohne unnötige Nebensächlichkeiten zu erfassen, bedarf es grundlegender Kenntnisse der Datenbankentwurfsmethodik. Es ist daher angebracht, von Anbeginn eine enge Zusammenarbeit von Fach- und Datenbankspezialisten anzustreben.

Beschrieben werden soll die so genannte 'Miniwelt', gedachte oder reale Gegebenheiten einer empirisch wahrgenommenen Realwelt. Keinesfalls sollten lediglich die Datenquellen in ihrer gegebenen Struktur den Ausgangspunkt bilden. Die durch den Historiker vorgefundenen Quellen sind im Allgemeinen zu einem anderen Zweck angelegt worden. Die Sicht auf die Realwelt war eine andere, andere Konzepte lagen zu Grunde. Es liegt u.U. eine für die aktuelle Sicht nicht angemessene Informationsstruktur und damit eine nicht angemessene formale Beschreibung durch Daten vor. Das kann zu Informationslücken führen, die später nur mit hohem Aufwand geschlossen werden können.

Als Beispiel seien hier Kirchenbücher mit dem Geburtsregister, dem Heiratsregister und dem Sterberegister erwähnt, die sehr häufig als Datenbasis für genealogische Untersuchungen verwendet werden. Genealogische Ketten lassen sich relativ einfach nachvollziehen, wenn von der gegebenen Struktur abstrahiert wird und nur *Personen* und *Partnerschaften* betrachtet werden.

Wählt man die Tabellenform, so ergibt sich folgende Darstellung für eine Familie über drei Generationen (ID steht für Identifikator):

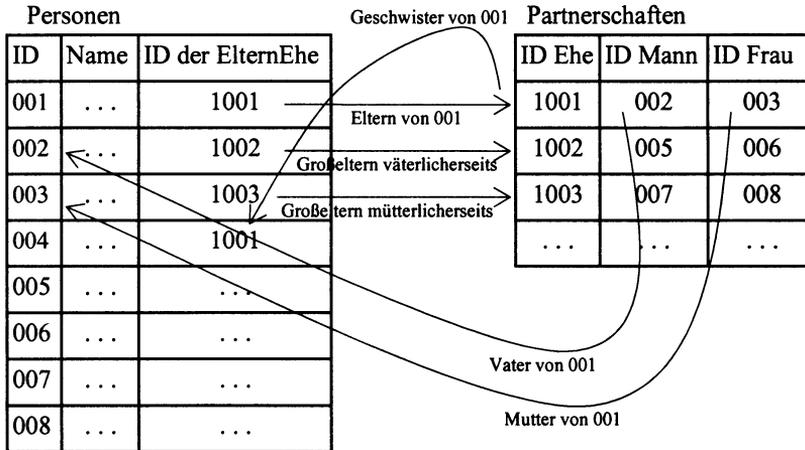


Abb. 2.1: Rekonstruktion von Verwandtschaftsbeziehungen

Wie man sich leicht selbst überlegen kann, sind aus diesen beiden Tabellen alle Verwandtschaftsstrukturen (Eltern, Großeltern, Geschwister, Onkel, Tanten, alle Ehen einer Person usw.) rekonstruierbar.

Für alle Entwurfsphasen, auch für die Informationsbedarfsanalyse, wurden Methoden und Verfahren entwickelt, die den Entwurf nicht als Kunst, die nur von Experten zu beherrschen ist, erscheinen lassen, sondern ein praxisbezogenes Herangehen gestatten. Die durch eine Datenbank darzustellenden Gegebenheiten – die Miniwelt – sind in dieser ersten Phase mit informellen Mitteln zu beschreiben. Die Grundidee einer solchen Vorgehensweise soll kurz skizziert werden.

In der betrieblichen Praxis oder im kommerziellen Bereich geht man so vor, dass man im ersten Schritt Organisationseinheiten identifiziert. Die Miniwelt wird auf logische und ggf. auch auf reale Organisationseinheiten untersucht, die homogen sind bzgl. ihrer Aufgaben und ihrer Sprache, d.h. der in ihr üblichen Terminologie. In der Regel wird es sich dabei um die einzelnen Abteilungen des betrachteten Unternehmens handeln, z.B. die Bereiche Verkauf, technischer Betrieb und Einkauf, Labor und Geschäftsführung. In einem nächsten Schritt werden zu unterstützende Aufgaben und die damit befassten Organisationseinheiten identifiziert. Durch weitere Verfeinerung der Betrachtung, durch Befragung von Fachexperten, Befragung der unmittelbar mit den Aufgaben befassten Mitarbeiter und durch Studium bereits vorhandener Formulare zum Betriebsablauf werden verschiedenen Sichten erarbeitet, die nach einem Abgleich zu einer Gesamtsicht integriert werden. Als Endprodukt dieses Prozesses erhält

man Tabellen und Übersichten, die die Miniwelt mit noch weitgehend niedrigen Abstraktionsmitteln beschreiben.

Prinzipiell wird beim Entwurf von Datenbanken für wissenschaftliche Anwendungen ebenso vorgegangen. Im wissenschaftlichen Bereich sind für den Datenbankentwurf nicht 'Geschäftsprozesse' zu analysieren, sondern es sind wissenschaftliche Zusammenhänge, Theorien und Methoden zu berücksichtigen. Als wesentliche wissenschaftliche Vorleistung muss man die theoretischen Konzepte, die mit dem Forschungsgegenstand verbunden sind, zur Kenntnis genommen haben.

Der Datenbankentwurf gliedert sich in den Entwurf der zu realisierenden Operationen (funktioneller Entwurf) und die Modellierung der möglichen Zustände der Miniwelt und ihre Zusammensetzung.

2.2 Funktioneller Entwurf

Unsere Aufgabe ist es, für das Projekt „Sozialdemokratische Parlamentarier in den deutschen Reichs- und Landtagen 1867 - 1933“ (BIOSOP) eine Datenbank aufzubauen. Die funktionelle Seite erschließt sich unmittelbar aus der für die Auswertung des Datenbestandes vorgesehene Analysemethode, der Methode der Kollektiven Biographie.

BIOSOP versteht sich als Beitrag zur Erforschung politischer Eliten im Rahmen interdisziplinären Wahl-, Parlamentarismus-, Parteien- und Verbändeforschung. Hauptquelle der Erforschung politischer Eliten stellen ihre Lebensläufe dar, die durch die Methode der 'Kollektiven Biographie' analysiert werden.

Bei der Eliteforschung geht es darum

- durch Untersuchung der Zusammensetzung, der Rekrutierung, der Verflechtung, der Transformation von Eliten Erkenntnisse zu gewinnen über Sozialstruktur, Schichtgefüge, Mobilitätsprozesse der Gesellschaft und den damit verbundenen Wandel.
- durch Analyse der Herkunft, der Wertvorstellung, der Kohärenz, der Zirkulation, der Machtstellung von Eliten Erkenntnisse zu gewinnen über soziale Grundlagen und Bedingungen sozialer Prozesse und über Determinanten des politischen Handelns und Verhaltens von Eliten.

In diesem Zusammenhang werden Parlamentarische Eliten als Funktionseliten verstanden.

Funktionseliten	„Funktionseliten sind die mehr oder weniger geschlossenen sozialen und politischen Einflussgruppen, welche sich aus den breiten Schichten der Gesellschaft und ihrer größeren und kleineren Gruppen auf dem Wege der Delegation oder der Konkurrenz herauslösen, um in der sozialen oder der politischen Organisation eine bestimmte Funktion zu übernehmen.“
------------------------	---

Schröder (2000) gibt folgende Definition der Kollektiven Biographie:

Kollektive Biographie	Kollektive Biographie ist die „theoretisch und methodisch reflektierte, empirische, besonders auch quantitativ gestützte Erforschung eines historischen Personenkollektivs in seinem gesellschaftlichen Kontext anhand einer vergleichenden Analyse der individuellen Lebensläufe der Kollektivmitglieder.“
------------------------------	---

Als bekannteste Anwendung der Kollektiven Biographie kann der karrieretheoretische Ansatz gelten: Die Rekrutierung von politischem Führungspersonal wird als kollektiver soziopolitischer Prozess gesehen. Dabei verstehen wir Karriere im engeren Sinne als Sequenz von Positionen, die Personen typischerweise auf ihrem Weg zur Spitze durchlaufen. Karriere im weiteren Sinne wird verstanden als eine nach der Zeit geordnete Sequenz von Konfigurationspositionen zur Beobachtung eines einzelnen Phänomens im Lebenslauf.

Die funktionale Analyse bedient sich der Mittel des Software-Engineering, auf die hier nicht eingegangen werden soll.

2.3 Die erste Phase des Datenbankentwurfsprozesses

Die Anforderungsanalyse soll nur exemplarisch an einigen wenigen Beispielen demonstriert werden.

Die Analyse sucht Objekte und Prozesse zu identifizieren und Regeln zu erkennen, die für die Objekte bzw. für die Beziehungen zwischen diesen Objekten gelten. In Form von eindeutigen, kurzen und einfach zu handhabenden Regeln (in der Literatur findet man den Begriff „Geschäftsregeln“) sind relevante Zusammenhänge zu formulieren.

Parlamentarier mit ihren standardisierten Lebensläufen stehen im Mittelpunkt unserer Untersuchung. Zu klären ist, was wir unter einem Parlamentarier verstehen wollen. Eine pauschale Festlegung wie „alle Parlamentarier“ ist nicht geeignet, da sie der jeweiligen historischen Situation nicht Rechnung trägt. Auf Grund des unterschiedlichen Entwicklungsstandes des Parlamentarismus in Deutschland in der Vor- und Nachkriegszeit ist ein differenziertes Herangehen geboten.

Folgende Kriterien für die einzelnen Epochen kamen z.B. bei den Reichstagsabgeordneten zur Anwendung:

1867 - 1918	<ul style="list-style-type: none"> • Aufgenommen wurden Abgeordnete, die in den Haupt-, Nach- und Ersatzwahlen gewählt worden sind. • Selektionskriterium war der Eintrag der Mandatsfeststellung in den Akten des Reichstages.
1919 - 1933	<ul style="list-style-type: none"> • Nur solche Personen wurden aufgenommen, die als Abgeordnete direkt gewählt bzw. als Mandatsnachrücker offiziell durch das Reichstagspräsidium festgestellt wurden. • Selektionskriterium war der Eintrag der Mandatsfeststellung in den Akten des Reichstages. <p>Die Kriterien waren für Landtagsabgeordnete (insbesondere für die Zeit des Kaiserreichs) teilweise anders.</p>

Gemäß den Vorgaben der Kollektiven Biographik und aufgrund der historischen Quellenlage, enthalten die Biographien der Parlamentarier folgende relevante Informationssegmente:

Personennamen:

- Familienname, Vorname(n), weitere Namen, Namenszusätze.

Personenstandsangaben:

- Geburtsdatum, Geburtsort,
- Sterbedatum, Sterbeort,
- Beruf des Vaters (ggf. der Mutter),
- Religionsbekenntnis (ggf. Bekenntniswechsel bzw. Kirchenaustritt),
- Familienstand (ggf. mit Angabe des Heiratsjahres bzw. der Heiratsjahre),
- Gesamtzahl der Kinder (nur bei Frauen).

Sozialisation:

- Dauer, Art und Ort der Schul-, Weiter- und Hochschulbildung (ggf. Zugehörigkeit zu einer studentischen Vereinigung),
- Dauer des Militärdienstes,
- Wanderschaft (nur für Handwerker, Angabe der durchwanderten Länder).

Berufstätigkeit:

- Dauer, Art und Ort aller hauptamtlichen Berufspositionen bzw. Positionen, die überwiegend zum Haupterwerb dienen,
- Dauer, Art und Ort von relevanten nebenamtlichen Berufspositionen.

Funktionen in Politik, Parteien, Verbänden, öffentlicher Verwaltung, Wirtschaft:

- Dauer, Art und Ort/Raum aller relevanten Funktionen in den genannten Bereichen im Ehren- oder Nebenamt (im Hauptamt vgl. Berufstätigkeiten),

Parlamentarische Ämter und Mandate:

- Dauer, Art und Ort/Raum aller relevanten parlamentarischen Ämter,
- Dauer und Art (ggf. Ort/Raum) aller relevanten parlamentarischen Mandate.

Sonstiges:

- Nennung von persönlichen Besonderheiten (z.B. Verfasser einer Autobiographie oder anderer Publikationen, Lebensschicksal in der NS-Zeit, längere politische Haftstrafen.

Auf der Grundlage dieser noch sehr allgemeinen Informationssegmente sind Informationselemente abzuleiten sowie Einschränkungen vorzunehmen, Regeln und Wertebereiche zu bestimmen. Im folgenden Abschnitt wird die in der Kollektiven Biographik entwickelte 'Normalbiographie' vorgestellt, die bereits sehr detailliert Informationselemente enthält.

Aus dem Kontext Parlamentarismus sind weitere Regeln und Einschränkungen herzuleiten. Das soll beispielhaft für das Reichstagswahlrecht als Rahmenbedingung für eine Reichstagskandidatur erfolgen. Es lassen sich einige für den Datenbankentwurf relevante Fakten ableiten, die hier in Form von 'Geschäftsregeln' aufgelistet werden:

- Wahlkreise sind nicht immer direkt Verwaltungseinheiten zuzuordnen.
- Wahlkreise sind nicht immer im Untersuchungszeitraum konstante Territorialeinheiten.
- Legislaturperioden waren nicht von gleicher Dauer. Vor dem ersten Weltkrieg erstreckten sie sich ca. über 3 Jahre, ab 1918 über 4 Jahre, wobei in beiden Zeitabschnitten Abweichungen davon (vor allem durch vorzeitige Parlamentsauflösungen) auftraten.
- Ein Reichstagskandidat konnte sich im mehreren Wahlkreisen zur Wahl stellen, aber nur für einen Wahlkreis im Parlament vertreten sein.
- Frauen erhielten das Wahlrecht erst ab Januar 1919.
- Bis zum Januar 1919 musste man mindestens 25 Jahre alt sein, um gewählt werden zu können. Danach waren Männer und Frauen ab dem 20. Lebensjahr wählbar.

Wie wir später bei der Modellierung der Miniwelt sehen werden, sind für den Datenbankentwurf eher solche Aussagen von Interesse, die sich auf Typen von Objekten beziehen. Entscheidend ist, ob identifizierte Typen von Einheiten in ihrer zeitlichen Entwicklung als homogen anzusehen sind oder wie sich ihre Beziehungen zu anderen Einheiten gestalten. Die Aussagen lassen sich häufig, wie oben angedeutet, in Kernsätze oder Regeln fassen.

2.4 BIOSOP-Normalbiographie

Im Rahmen der Kollektiven Biographik wurde eine standardisierte Form des Lebenslaufes entwickelt, die alle relevanten Informationen des durchschnittli-

chen Lebenslaufes enthält. Die folgenden Tabellen enthalten kurze Beschreibungen der Informationseinheiten dieser Normalbiographie für das BIOSOP-Projekt.

Der Begriff ‘Kardinalität’ bezeichnet in diesem Zusammenhang die maximal mögliche Anzahl der Objekt- bzw. Merkmalsausprägungen eines Objekt-(Merkmals-) Typs, die sich je einem Vertreter eines anderen Objekttyps zuordnen lassen. Beispielsweise kann ein Abgeordneter mehrere akademische Titel erworben haben. In der Spalte ‘Kard.’ wurde ‘1’ für höchstens eine Ausprägung oder ‘m’ für beliebig viele Ausprägungen eingetragen.

Teil 1: Personenstandsangaben und Sozialisation

Merkmal	Kard.	Beschreibung
Nachname	1	Nachname
Vorname	1	Als Vorname erscheint der Rufname, sonst der erste Name.
Alias	m	weitere Namen
Alias-Hinweise		Folgende Hinweise auf Aliasnamen (ausschließlich): ‘geb.’ ‘verh.’ ‘wiederverh.’ ‘gen.’ (genannt)
Adelsprädikate	1	quellengetreu erfasst
Akademische Titel	m	Akademische Titel mit Fakultätsangabe
Geburtsjahr	1	Format: tt.mm.jjjj, auch unvollständig mm.jjjj oder jjjj
Geburtsort	1	Ortsangabe, Kennzeichnung des Geburtsortes durch die nächst höhere politisch-geographische Einheit
Sterbejahr	1	Format: tt.mm.jjjj, auch unvollständig mm.jjjj oder jjjj
Sterbeort	1	Ortsangabe, Kennzeichnung des Sterbeortes durch die nächst höhere politisch-geographische Einheit
Beruf(e) des Vaters, ggf. der Mutter	1	Berufsbezeichnung wird in der Regel quellenah wiedergegeben
Beziehung zu anderen Abgeordneten	m	Steht der betreffende Abgeordnete in einem Verwandtschaftsverhältnis zu einem anderen Abgeordneten, wird der Name und das Verwandtschaftsverhältnis genannt.

Merkmal	Kard.	Beschreibung
Konfessions- bzw. Religionszugehörigkeit	1	Nicht das Bekenntnis zum Zeitpunkt des Parlamenteintritts, sondern eine Zusammenfassung aller Bekenntnisse und Wechsel. Es ist im Laufe der Erfassung ein Abkürzungsverzeichnis anzulegen, das alle vorkommenden Ausprägungen enthält. Nur diese standardisierten Bezeichnungen sind anzugeben: 'ev.' 'ev., sp. diss.' ...
Schul- und Hochschul- ausbildung	m	Strukturiert nach Zeit (von - bis), Schultyp / Studienfächer, Schulort
Lfd. Nr. der Schulmaß- nahme		Da evtl. nicht alle Schulmaßnahmen mit Jahreszahlen überliefert sind, die Reihenfolge aber häufig bekannt ist, muss eine 'laufende Nr.' der Schulmaßnahme erhoben werden, um die Abfolge korrekt wiedergeben zu können.
Schulmaßnahme von		Anfangsjahr
Schulmaßnahme bis		Endjahr
Schultyp / Fach		Die Bezeichnung des Schultyps wird in der Regel quellennah übernommen; einschließlich Studienfächer; Promotion mit Fakultätsangabe; Erwachsenen-Weiterbildung; Korporationszugehörigkeit.
Schulort		Orte werden ggf. durch eine zusätzliche geographische Spezifikation gekennzeichnet.
Berufsausbildung Art („erlernter Beruf“)	m	Einschließlich berufsbegleitender Schulbildung; bei fehlender Berufsausbildung (z.B. bei ungelerten Arbeitern) wird nur das Anfangsjahr der Berufstätigkeit genannt; die Bezeichnung des erlernten Berufs wird weitgehend quellennah übernommen.
Berufsausbildung von		Anfangsjahr
Berufsausbildung bis		Endjahr
Ort		Ortsangabe; Kennzeichnung des Ortes durch die Angabe der Provinz- oder Landeszugehörigkeit

Merkmal	Kard.	Beschreibung
Wanderschaft	m	Nur für handwerklich gelernte Arbeiter; eine zeitliche Präzisierung wird nicht vorgenommen; Angabe des jeweilige Nationalitätskennzeichen (internationale Autokennzeichen) der durchwanderten Länder; es sind auch Angaben zu erfassen wie „Ausland“ oder „ohne Ortsangabe“.
Militärdienst	1	Es sind die Angaben 'ja' 'nein' 'unbekannt' möglich.
Militärdienst von		Jahresangaben
Militärdienst bis		Jahresangaben
Familienstand	m	Eingetragen werden 'v' für verheiratet und 'l' für ledig; die Angabe „ledig“ bezieht sich nicht nur auf den Zeitraum des Mandats, sondern auf die gesamte Lebenszeit.
Jahr der Eheschließung		Jahr der Eheschließung falls Familienstand 'v'
Bemerkung		Das Bemerkungsfeld enthält mehrere nichtformalisierte Informationen und ist damit nicht atomar; ist die/der betreffende Abgeordnete mit einer/einem andere/anderen Abgeordneten verheiratet, wird der Name der Ehegattin/des Ehegatten erwähnt; bei mehreren Eheschließungen wird die laufende Nummer der Ehe genannt (z.B. „1. Ehe“); bei Frauen werden folgende Zusatzangaben gemacht: Gesamtzahl der nachweisbaren eigenen Kinder, ggf. Angabe der Verwitwung und der Scheidung mit nachfolgender Jahreszahl.
Name der Organisation des ersten Partei- und Gewerkschaftsbeitritt	2	Die Angabe bezieht sich nur auf den ersten Beitritt zur deutschen „Sozialdemokratie“ (ADAV, SDAP, SAPD, SPD, USPD) bzw. zu den „freien“ Gewerkschaften; bei den Gewerkschaftsbeitritten wird der konkrete Einzelverband nicht genannt (Gew.).
Jahr des ersten Partei- und Gewerkschaftsbeitritt		Beitrittsjahr
Hinweis Jahr		Hinweis darauf, ob die Jahresangabe gesichert ist.

Merkmal	Kard.	Beschreibung
Art(en) der unselbständigen Ausübung des erlernten Berufs bzw. der Berufstätigkeit	m	Berufsangabe quellennah; dauert die unselbständige Berufsausübung noch bei Mandatsantritt an, entfällt die Angabe (wird im 2. Teil aufgeführt); ebenso entfallen die Angaben zu den Tätigkeitsorten und zum jeweiligen Zeitraum von Einzeltätigkeiten (bei mehreren ausgeübten Tätigkeiten); ebenso entfällt die Angabe bei Abgeordneten mit akademischen Berufen (wird im Teil 2 aufgeführt).
Beginn der Berufsausübung		Jahresangabe
Ende der Berufsausübung		Jahresangabe
Ort der Berufsausübung		Ortsangabe; Kennzeichnung des Ortes durch die Angabe der Provinz- oder Landeszugehörigkeit
Lfd. Nr.		Da evtl. nicht alle Berufsausübungen mit Jahreszahlen überliefert sind, muss eine 'laufende Nr.' der Berufsausübung erhoben werden, um die Abfolge korrekt wiedergeben zu können.
Bemerkung zur Berufsausübung		Hier erscheinen alle anderen Informationen.

Teil 2: Berufliche und politische Karriere

Merkmal	Kard.	Beschreibung
Art(en) der hauptamtlichen Berufstätigkeit(en) bzw. von Positionen, die überwiegend zum Haupterwerb dienen	m	Berufstätigkeit(en) in derselben Institution werden zusammenfassend dargestellt.
Laufende Nr.		Laufende Nummer der Berufsausübung in chronologischer Reihenfolge
Jahr und Monat des Beginns der hauptamtlichen Berufstätigkeit		Angabe von Jahr und Monat soweit bekannt.
Verlauf ab		Hinweis über den Verlauf der Tätigkeit: „seit“, „um“, ...

Merkmal	Kard.	Beschreibung
Art(en) der hauptamtlichen Berufstätigkeit(en) bzw. von Positionen, die überwiegend zum Haupterwerb dienen	m	Berufstätigkeit(en) in derselben Institution werden zusammenfassend dargestellt.
Verbürgt ab		Die Zeitangaben sind nicht immer genau zu eruieren; daher wird eine Information aufgenommen über die Verbürgtheit des Zeitpunktes.
Jahr und Monat des Endes der hauptamtlichen Berufstätigkeit		Angabe von Jahr und Monat soweit bekannt.
Verlauf bis		Hinweis über den Verlauf der Tätigkeit: „bis“, „um“, ...
verbürgt bis		Die Zeitangaben sind nicht immer genau zu eruieren; daher wird eine Information aufgenommen über die Verbürgtheit des Zeitpunktes.
Ort der hauptamtlichen Berufstätigkeit		Orte werden ggf. durch eine zusätzliche geographische Spezifikation gekennzeichnet.
Kriegsteilnahme	1	'Ja' 'Nein'-Information
Zeitpunkt von		Die Monatsangabe erfolgt nur bei Abweichung von den allgemeinen Mobilisierungs- und Demobilisierungsdaten; nur wenn Kriegsteilname den Wert 'Ja'.
Zeitpunkt bis		Die Monatsangabe erfolgt nur bei Abweichung von den allgemeinen Mobilisierungs- und Demobilisierungsdaten; nur wenn Kriegsteilname den Wert 'Ja'.
Bemerkung		Ggf. wird der zuletzt erreichte Dienstgrad und ggf. schwere Kriegsbeschädigungen erfasst.
Parteispezifikation bei Partei- und Fraktionswechsel bzw. des Parteiausschlusses / -Austritts		Die Eintragungen erfolgen auf der Basis einer standardisierten Liste der Parteibezeichnungen.
Datum des Partei- und Fraktionswechsel bzw. des Parteiausschlusses / -Austritts		Beim Datum des Austritts / Ausschlusses / Wechsels werden nach Möglichkeit das Jahr und der Monat genannt.

Merkmal	Kard.	Beschreibung
Art(en) der hauptamtlichen Berufstätigkeit(en) bzw. von Positionen, die überwiegend zum Haupterwerb dienen	m	Berufstätigkeit(en) in derselben Institution werden zusammenfassend dargestellt.
Art(en) aller relevanten (halb-, neben- und ehrenamtlichen) Ämter / Funktionen in den Bereichen Politik / Parlament / Öffentliche Verwaltung / Verbände / Wirtschaft	m	Quellennahe Bezeichnung
Dauer aller Ämter		Angaben der Dauer mit Anfangsjahr - Endjahr, ggf. bei besonders relevanten Funktionen Monatsangaben
Ort(e)/Bereich(e)		Orte werden ggf. durch eine zusätzliche geographische Spezifikation gekennzeichnet.
Art aller relevanter Ämter / Funktionen innerhalb der Arbeiterbewegung	m	Quellennahe Amtsbezeichnung
Beginn des relevanten Amtes innerhalb der Arbeiterbewegung		Anfangsjahr, ggf. bei besonders relevanten Funktionen Monatsangaben
Ende des relevanten Amtes innerhalb der Arbeiterbewegung		Endjahr, ggf. bei besonders relevanten Funktionen Monatsangaben
Art von längeren Haftstrafen	m	Mit Ausnahme der Haftstrafen während der NS-Zeit (siehe dort); ggf. bei mehreren kleineren Haftstrafen für ähnliche Vergehen: Zeitraum, Gründe und Gesamthaftzeit.
Dauer der längeren Haftstrafe		Die Angabe der Dauer erfolgt mit Anfangsmonat.Anfangsjahr - Endmonat.Endjahr der Haft.
Ort(e) der längeren Haftstrafe		Orte werden ggf. durch eine zusätzliche geographische Spezifikation gekennzeichnet.
Gründe von längeren Haftstrafe		Freier Text

Merkmale	Kard.	Beschreibung
Art(en) der hauptamtlichen Berufstätigkeit(en) bzw. von Positionen, die überwiegend zum Haupterwerb dienen	m	Berufstätigkeit(en) in derselben Institution werden zusammenfassend dargestellt.
Art von Haftstrafen in der NS-Zeit	m	'Zuchthaus' 'Untersuchungsgefängnis' 'KZ' 'Gefängnis'
Beginn der Haftstrafen in der NS-Zeit		Angabe von Anfangsmonat.Anfangsjahr
Ende der Haftstrafen in der NS-Zeit		Angabe von Endmonat.Endjahr
Ort von Haftstrafen in der NS-Zeit		Orte werden ggf. durch eine zusätzliche geographische Spezifikation gekennzeichnet.
Land der Emigration in der NS-Zeit	m	Aufenthaltsland der Emigration, Länderkürzel
Jahr der Emigration in der NS-Zeit		Anfangsjahr der Emigration
Bemerkung		ggf. Jahr und Zielort der dauerhaften Rückkehr nach Deutschland oder ggf. „keine dauerhafte Rückkehr nach Deutschland“
Wohnort(e) / Raum bei Personen ohne Berufstätigkeiten	m	Orte werden ggf. durch eine zusätzliche geographische Spezifikation gekennzeichnet.
Dauer bei Personen ohne Berufstätigkeiten		Angabe der Dauer mit Anfangsjahr - Endjahr, ggf. Monatsangaben
Nennung von sonstigen (außergewöhnlichen und allgemein bekannten) Besonderheiten	1	Freier Text
Verleihung einer Ehrendoktorwürde / Ehrenbürgerschaft	m	Fakultät / Universität der Verleihung einer Ehrendoktorwürde; Ort der Verleihung einer Ehrenbürgerschaft.
Jahr der Ehrung		(ggf. Monat.)Jahr

Merkmal	Kard.	Beschreibung
Art(en) der hauptamtlichen Berufstätigkeit(en) bzw. von Positionen, die überwiegend zum Haupterwerb dienen	m	Berufstätigkeit(en) in derselben Institution werden zusammenfassend dargestellt.
Publizistische Tätigkeit	1	Nur exemplarische Nennung der Tätigkeit als Verfasser, Mitverfasser, Herausgeber, Mitherausgeber von Aufsätzen und Büchern, die vorrangig in Bezug zur politischen Tätigkeit stehen oder als Standardwerke auf ihrem Gebiet anzusehen sind; Nennung von autobiographischen Werken. Die Publikationen werden pauschal charakterisiert und in nachgestellter Klammer wird/werden ggf. ein oder mehrere exemplarische Werke mit Titel genannt.

Teil 3: Parlamentarische Tätigkeit

Merkmal	Kard.	Beschreibung
MdR: Wahlkreisname	m	Der Wahlkreisname wird bis 1918 durch die jeweilige offizielle Kennziffer innerhalb des Landes, der Provinz, des Regierungsbezirks etc. bezeichnet (z.B. „Elsass-Lothringen 10“, „Hannover 18“). Der Wahlkreisname besteht seit 1919 aus der Kennziffer des Großwahlkreises und dem in Klammern nachgestellten Wahlkreisnamen – z.B. „WK 24 (Oberbayern)“, „KPD-Reichswahlvorschlag“).
MdR: Mandatsdauer		Angabe der Mandatsdauer mit Eintrittsmonat.Eintrittsjahr - Austrittsmonat.Austrittsjahr; an anderer Stelle werden Wahldaten erfasst, um so nachträgliches Eintreten oder vorzeitiges Ausscheiden feststellen zu können.
MDB: ggf. Fraktionszugehörigkeit (seit 1949)	m	Die Nennung der Fraktionszugehörigkeit muss nur bei einem Wechsel zu einer anderen Partei erfolgen, bei fortdauernder Zugehörigkeit zur SPD kann eine Standardannahme vorgenommen werden.
MDB: Mandatsdauer (seit 1949)		Angabe der Mandatsdauer mit Wahlmonat.Wahljahr - Wahlmonat.Wahljahr

Merkmal	Kard.	Beschreibung
MDV: Mandatsdauer (seit 1949)	m	Angabe der Mandatsdauer mit Anfangsjahr–Endjahr
MDL: Mandatsland (1871 - 1933 und seit 1946)	m	Angabe des Landesnamen
MDL: Mandatsdauer (1871 - 1933 und seit 1946)		Angabe der Mandatsdauer mit Anfangsjahr(e) - Endjahr(e)
Reichstagskandidat: Wahlkreis (1867 - 1933)	m	Es werden nur erfolglose Wahlbewerbungen genannt, sonst gehört die Information in die MdR-Liste.
Reichstagskandidat: Wahljahr (1867 - 1933)		Bei Hauptwahlen nur Angabe des Wahljahres, bei Ersatz- und Nachwahlen auch Wahltag und Wahlmonat.
Parteitage Partei	m	Delegierung zu den nationalen „sozialdemokratischen“ Parteitag (1863 - 1931); Frauen: Jahr von Delegierungen zu den nationalen Frauenkonferenzen (1900 - 1927)
Parteitage Ort		Ort
Parteitage Jahr und Monat		Jahr, Monat von Delegierungen zu den nationalen „sozialdemokratischen“ Parteitag (1863 - 1931)
Gewerkschaftskongresse	m	Delegierungen zu den allgemeinen nationalen Kongressen der freien Gewerkschaften bzw. des Allgemeinen Deutschen Gewerkschaftsbundes und Nennung der delegierenden Berufsgruppe/Gewerkschaft (1892 - 1931)
Jahr des Gew.-Kongresses		Jahr von Delegierungen zu den allgemeinen nationalen Kongressen der freien Gewerkschaften
Berufsgruppe des Verbandstages	m	Delegierungen zu den nationalen Verbandstagen von freien Einzelgewerkschaften und Nennung der Berufsgruppe (1890 - 1914)
Jahr des Verbandstages		Jahr von Delegierungen zu den nationalen Verbandstagen von freien Einzelgewerkschaften und Nennung der Berufsgruppe (1890 - 1914)

Merkmal	Kard.	Beschreibung
Jahr der Teilnahme an Internationale	m	Jahre von Delegierungen zu den Kongressen der II. Internationalen (1889 - 1912) und der Sozialistischen Arbeiter-Internationalen (1923 - 1933)

2.5. Studienbeschreibung: BIOSOP-Projekt

Die Forschungen zu den Biographien sozialdemokratischer Parlamentarier durch W. H. Schröder reichen bis zu Beginn der 1970er Jahre zurück. Diesen Forschungen ist es zu verdanken, dass dem ZHSF eine umfangreiche Quellensammlung in Form von Personendossiers und anderen Quellen zur Parlamentarismusforschung zur Verfügung steht. Auf Fragestellungen, Methoden, Quellen, Erschließung, Auswertung etc. kann hier nicht eingegangen werden. Hier sei nur zur knappen Information die gekürzte Studienbeschreibung wiedergegeben, wie sie sich im ZA-Datenbestandskatalog zum BIOSOP-Datensatz, den man beim ZA zur Sekundäranalyse bestellen kann, findet.

ZA-Studiennummer: 8011: Sozialdemokratische Parlamentarier in den deutschen Reichs- und Landtagen 1867 bis 1933 (BIOSOP)

Inhalt:

Kollektivbiographische Untersuchung der Lebensläufe sozialdemokratischer Parlamentarier im Allgemeinen (vor allem im Hinblick auf berufliche und politische Karrieremuster) und der Struktur sozialdemokratischer Fraktionen im Reichstag und in den Landtagen im Besonderen.

Themen:

Geburt, Heirat, Alter und Tod; soziale und regionale Herkunft; schulische und berufliche Ausbildung; Positionen der Berufskarriere; Ämter in Politik / Öffentlicher Verwaltung / Verbände und in der Arbeiterbewegung (Parteiorganisation, Publizistik, Gewerkschaften, Genossenschaften, Krankenkassen); Modalitäten des parlamentarischen Mandats.

Grundgesamtheit und Auswahl:

Untersuchungsgebiet

a) Norddeutscher Bund, Deutsches Reich; b) deutsche Einzelstaaten.

Grundgesamtheit:

Totalerhebung aller (ca. 7000) Mandate der (2427) sozialdemokratischen Parlamentarier a) im Norddeutschen Reichstag, im Deutschen Reichstag des Kaiserreichs, in der Deutschen Nationalversammlung und in den Reichstagen der

Weimarer Republik und b) in den Landtagen (Bürgerschaft, Abgeordnetenhaus, Zweite Kammer, Volkskammer etc.) der 26 Bundesstaaten des Kaiserreichs bzw. der (zunächst 25, dann) 18 bzw. 17 Einzelstaaten der Weimarer Republik, einschließlich der Mandate des Danziger Volkstages (1920-1937).

Erhebungsverfahren:

a) Systematische Erschließung von gedruckten Quellen:

- publizierte amtliche Quellen (u.a. Wahlpublikationen, Verhandlungsprotokolle des Reichstags und der Landtage, Staats- und Parlamentshandbücher, Verzeichnisse unterschiedlicher Art);
- publizierte nichtamtliche Quellen (Primär- und Sekundärliteratur mit Biographien bzw. biographischen Annotationen oder Listen bzw. Verzeichnissen von sozialdemokratischen Funktionären);
- allgemeine biographische Nachschlagewerke und Sammlungen sowie sonstige allgemeine Quellen (u.a. Zeitschriften, Tageszeitungen, Jahrbücher, Almanache, Berichte, Protokolle etc. der Arbeiterbewegung).

b) Systematische Erschließung archivalischer Quellen:

- Bestände zum Personenstandswesen (Kirchenbücher, Standesamts- und Melderegister etc.);
- Bestände zum Gerichts- und Polizeiwesen (Akten über politische Strafsachen / Deliquenten, Fahndungs- und Ausweisungsakten, Akten verdächtiger und überwachter Personen bei der Politischen Polizei etc.);
- Bestände zum Regierungs-, Verwaltungs- und Vereinswesen (Berichte der Landräte, Akten bzw. Materialien zu den Reichs- und Landtagswahlen etc.).

c) Punktuelle Erschließung biographischer Quellen:

- (Insgesamt weit mehr als 9000) postalische Um- und Anfragen;
- Befragung von Archiven, Standesämtern, Meldebehörden, arbeitgebenden Einrichtungen und
- Befragung der Nachkommen der Parlamentarier.

Eine Darstellung von Quellen und Methoden findet sich ausführlich im BIOSOP-Handbuch (1995) und verkürzt im Supplementheft 11 (2000).

3. Datenbankmodelle

3.1 Grundlagen von Datenbankmodellen

Modell	Ein Modell ist die vereinfachte Nachbildung eines originalen (realen oder hypothetischen) Untersuchungsobjektes, die im Hinblick auf ein gegebenes Untersuchungsziel alle relevanten Eigenschaften des Originals aufweist. Es entsteht im Ergebnis einer Tätigkeit, die Modellierung oder Modellbildung genannt wird. Datenentwurf im weitesten Sinne heißt, einen Ausschnitt der Umwelt zu modellieren. Ein Datenmodell stellt ein Instrumentarium zur Verfügung, das als Modellierungsvorschrift beim Datenentwurf zur Anwendung gebracht wird.
--------	---

Die Begriffe „Datenmodell“ und „Modell“ sind zu unterscheiden. Es sind historisch gewachsene Begriffe, die etwas unglücklich gewählt wurden. „Datenmodell“ ist, wie oben hervorgehoben, als Modellierungsvorschrift zu sehen. Es stellt ein System von Konzepten zur Beschreibung relevanter Daten dar. Ein „Modell“ ist ein Modellierungsergebnis. Das Ergebnis der Modellierung, die nach den Vorschriften des Datenmodells durchgeführt wird, ist ein Modell des betrachteten Realitätsausschnittes.

Die formalisierte Darstellung eines Modells, das nach den Vorschriften eines Datenmodells abgeleitet einen Ausschnitt der Umwelt abbildet, wird als Schema bezeichnet.

Datenbank-Modell	Ein Datenbankmodell ist ein System von Konzepten zur Beschreibung von Datenbanken. Es legt Syntax und Semantik von Datenbankbeschreibungen für ein Datenbanksystem, den so genannten Datenbankschemata, fest.
------------------	---

Datenbank-Schema	Ein Datenbankschema ist eine Datenbankbeschreibung, in einer 'Sprache' formuliert, die ausgewählte Abstraktionskonzepte zur Beschreibung bereitstellt.
------------------	--

Eine Datenbank ist also eine Ausprägung einer Beschreibung, ein Datenbankmodell eine Auswahl an Abstraktionskonzepten. Zu den klassischen Datenbankmodellen gehören:

- abstrakte Datenmodelle, die speziell zum Datenbankentwurf geeignet sind, z.B. das *Entity Relationship Model* (ERM),
- konkrete Datenbankmodelle, die in einem DBMS umgesetzt wurden, z.B. das *relationale Modell*.

In Abbildung 3.1 sind bekannte Datenmodelle in ihrer historischen Entwicklung dargestellt. Das *hierarchische Modell* wird heute wegen seiner eingeschränkten Anwendbarkeit kaum noch in DBMS umgesetzt. Es wurde durch das flexiblere *Netzwerkmodell* (NWM) vollständig verdrängt. Das NWM wird für sehr große Datenmengen in zeitkritischen Prozessen im Großrechnerbereich benutzt. Das heute verbreitetste Modell ist das relationale. Es ist einfach zu handhaben und theoretisch gut aufbereitet. Sowohl die meisten Anwendungen als auch die meisten Programmsysteme basieren auf diesem Konzept.

Die rasante Entwicklung der Computertechnik ermöglicht die Umsetzung moderner Konzepte wie das *objektorientierte Datenbankmodell*. Seine Bedeutung wird in den nächsten Jahren zunehmen.

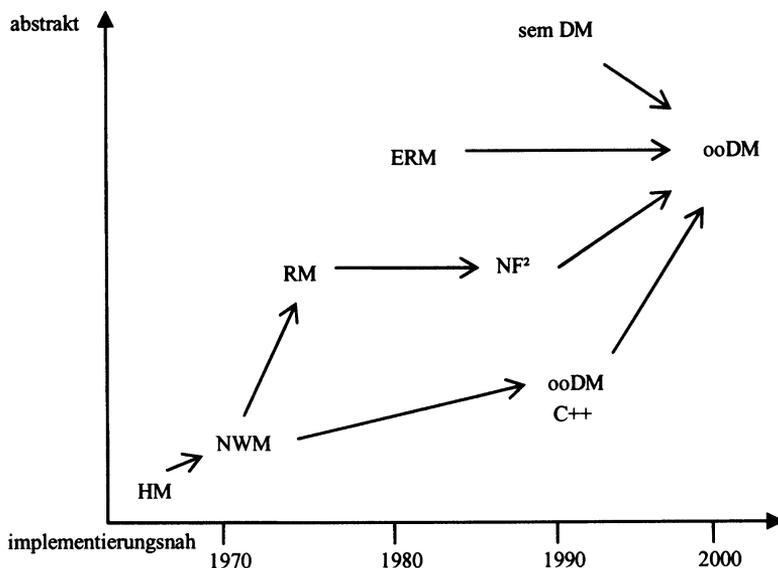


Abb. 3.1 Historische Einordnung der Datenbankmodelle nach Heuer/Saake (1995)

Legende:

- HM: Hierarchisches Modell
- NWM: Netzwerkmodell

RM:	Relationenmodell
NF ² :	Modell der geschachtelten Relationen (Non-First-Normal-Form)
ER:	Entity Relationship Modell
Sem. DM:	Semantische Datenmodelle
OODM C++:	Objektorientierte Datenmodelle auf der Basis von objektorientierten Programmiersprachen wie C++
OODM:	Objektorientierte Datenbankmodelle

In einer Datenbank werden Werte verwaltet. Darunter versteht man elementare Datenelemente, z.B. eine Zahl oder eine Zeichenfolge. In der Informatik hat sich der Begriff des abstrakten Datentyps als Konzept für Wertebereiche herausgebildet. Das sind die in den verbreiteten Programmiersprachen verwendeten Datentypen wie Ganzzahl (integer), Kommazahl (real), Zeichen (character) bzw. Zeichenfolgen (string) und logische Größen (nach dem Mathematiker Boole auch boolesche Werte oder Booleans genannt).

3.2 Das relationale Datenmodell

Edgar F. Codd unterbreitete 1970 in seinem Artikel „A Relational Model of Data for Large Shared Data Banks“ den Vorschlag, die Relationentheorie, eine algebraische Theorie, als Grundprinzip für die Datenspeicherung zu nutzen. Damit war keineswegs bewiesen, dass es auf die von ihm vorgeschlagene Weise sinnvoll möglich ist, ein Datenbanksystem zu entwickeln. Man versuchte Anfang der 70er Jahre in den Forschungslabors von IBM in San Jose, Kalifornien, diese Idee umzusetzen, und entwickelte mit dem ‘System R’ das erste relationale Datenbankmanagementsystem. Die Untersuchung war verbunden mit der praktischen Umsetzung der relationalen Algebra. Es wurde die Sprache SQUARE entwickelt, die heute keine Bedeutung mehr hat.

3.2.1 Grundlagen des relationalen Modells

Das relationale Modell besteht aus der Definition von

1. Objekten (Tabellen),
2. Operatoren (‘Rechenoperationen’, um die Tabellen zu verknüpfen),
3. Regeln (Bedingungen an die Werte bestimmter Spalten).

Operatoren definieren eine (relationale) Algebra, mit der Objekte unter Beachtung bestimmter Regeln bearbeitet werden können. Relationale Datenbanksysteme sind von moderneren DBMS am weitesten verbreitet. Ihnen liegt ein sehr einfaches Datenstrukturierungsmodell zugrunde. Mit SQL (Structured Query Language) steht eine Datenbanksprache zur Verfügung, die weitgehend standardisiert ist.

3.2.1.1 Was ist eine 'Relation'?

Wir wollen den Begriff der Relation an einem einfachen Beispiel erläutern. Ein bestimmtes Personenkollektiv, z.B. sozialdemokratische Reichstagsabgeordnete, die von Januar 1919 bis November 1920 im Reichstag vertreten waren, soll durch ein geeignetes Instrumentarium beschrieben werden. Wir wollen zunächst nur zwei Merkmale betrachten: Nachname und Vorname der Abgeordneten. Formal gesehen sind Namen eine Aneinanderreihung von endlich vielen Buchstaben (ggf. auch von Sonderzeichen). Wir haben damit die Menge aller endlichen Zeichenketten und ein konkreter Name ist ein Element aus dieser Menge. Wir wollen das aber konkreter sehen. Es gibt eine Menge von Nachnamen und eine Menge von Vornamen, z.B. die folgenden:

Nachname: 'Mauerer', 'Meerfeld', 'Merges', 'Michelsen', 'Molkenbuhr', 'Müller'

Vorname: 'August', 'Georg', 'Hermann', 'Johannes', 'Peter'

Kombiniert man die 6 Nachnamen mit den 5 Vornamen, so könnte man 30 (5*6) Personen ohne Verwechslung Namen zuordnen.

Nachname	Vorname
Mauerer	August
Mauerer	Georg
Mauerer	Hermann
Mauerer	Johannes
Mauerer	Peter
...	...
Müller	August
Müller	Georg
Müller	Hermann
Müller	Johannes
Müller	Peter

Eine solche Kombination von Elementen aus Mengen nennt man eine Relation. Es müssen zu einer Relation jedoch nicht alle möglichen Kombinationen gehören. Man könnte z.B. einige Abgeordnete, es sind die, deren Nachname mit 'M' beginnt, erfassen und durch eine Relation beschreiben.

Nachname	Vorname
Mauerer	Georg
Meerfeld	Johannes
Merges	August
Michelsen	Peter
Molkenbuhr	Hermann
Müller	Hermann

Wir erkennen, dass eine Relation geeignet ist, verschiedene Gegebenheiten, in diesem Falle Personen, durch bestimmte Merkmale zu beschreiben. Wenn man diese Kombination der Merkmale in Tabellenform darstellt, entspricht jeweils eine Zeile einer Gegebenheit.

In der ersten Tabelle, die alle Kombinationen enthält, ist keine Zeile doppelt. Das muss auch so sein, denn Nachname und Vorname sind Mengen und wie aus dem Mathematikunterricht bekannt, besteht eine Menge aus *unterschiedlichen* Elementen. Dann muss auch die Kombination zweier Mengen aus voneinander verschiedenen Paaren bestehen und ebenso jede Teilmenge der Paare.

Was ist aber, wenn wir in den Quellen einen weiteren Abgeordneten entdecken, der ebenfalls Herman Müller heißt?

Nachname	Vorname
Mauerer	Georg
Meerfeld	Johannes
Merges	August
Michelsen	Peter
Molkenbuhr	Hermann
Müller	Hermann
Müller	Hermann



Das ist keine Relation mehr, da zwei Zeilen gleich sind. Auch hat diese Tabelle den praktischen Nachteil, dass man die Personen nicht alle voneinander unterscheiden kann, sie ist also ein ungeeignetes Beschreibungsmittel für diese Abgeordneten! Wollen wir sie in der Liste unterscheiden, müssen wir ein weiteres Merkmal hinzunehmen, beispielsweise das Geburtsdatum.

Nachname	Vorname	Geburtstag
Mauerer	Georg	26.11.1868
Meerfeld	Johannes	16.10.1871
Merges	August	03.03.1870
Michelsen	Peter	12.12.1866
Molkenbuhr	Hermann	11.09.1851
Müller	Hermann	18.05.1876
Müller	Hermann	10.02.1868

Da die beiden Abgeordneten mit Namen 'Herman Müller' nicht am gleichen Tag geboren wurden, unterscheiden sich nun wieder alle Zeilen voneinander. Damit ist diese Tabelle eine Relation. Das Problem ließe sich auch durch ein 'künstliches' Merkmal lösen, durch Hinzufügen einer laufenden Nummer für jede Person.

Für mathematisch Interessierte sei angemerkt, dass eine Relation das Kreuzprodukt von Mengen ist, mit der Einschränkung, dass bei Relationen die Reihenfolge der Spalten (Mengen) nicht festgelegt ist.

3.2.1.2 Definition „Relationale Datenbank“

Eine relationale Datenbank ist:

- konzeptionell eine Ansammlung von Tabellen;
- mathematisch steht hinter den Tabellen die Idee der Relation.

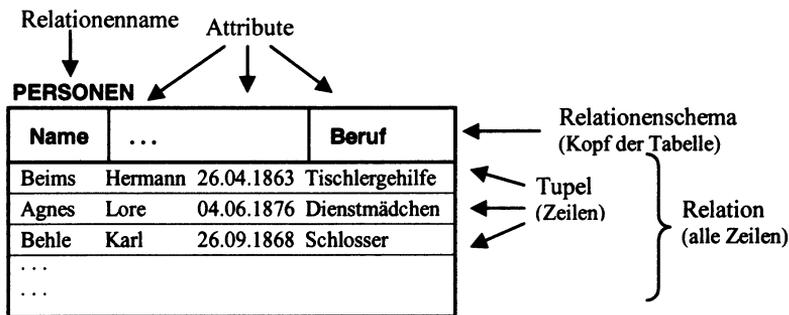


Abb. 3.2: Relationale Datenbank

Bezeichnungen:

erste Zeile (Struktur der Tabelle): Relationenschema (Benennung der Spalten)

alle weiteren Zeilen: Relation zu diesem Schema

eine Zeile: Tupel

Spaltenüberschriften: Attribut (Namen)

Domäne	Wertebereich eines Attributs (tatsächlich in Datenbank auftretende Werte). Ein Wert einer Domäne ist atomar (kann nicht sinnvoll zerteilt werden). Beispiel für eine Domäne: 'rot', 'grün', 'blau', 'gelb' für das Attribut Farbe.
Degree (Ausdehnungsgrad)	Anzahl der Attribute, d.h. Anzahl der Spalten der Tabelle.
Kardinalität	Anzahl der Tupel (Zeilen der Tabelle).

Schlüsselkandidat (Candidate Key)	<p>Eine Attributmenge wird Schlüsselkandidat genannt, wenn alle Werte dieser Menge eindeutig jedes Tupel identifizieren.</p> <p><u>Beispiel:</u> Häufig kann man davon ausgehen, dass Name, Vorname, Geburtsdatum eindeutig eine Person identifizieren.</p>
Primärschlüssel (Primary-Key, Alternate Key, Foreign Key)	<p>Genau ein Schlüssel in jeder Relation wird als Primärschlüssel (Primary Key) ausgewählt zur eindeutigen Identifizierung der Zeilen (Tupel). Er ist ausgewählt aus der Menge der möglichen Schlüssel (Candidate Keys). Aus Performancegründen werden in manchen Systemen zusätzliche Keys definiert, die als Alternate Keys oder Secondary Keys bezeichnet werden.</p> <p><u>Beispiel:</u> Man wählt aus den Attributkombinationen, die eine Person identifizieren, eine aus. Für Personen hätte man die Schlüssel Name, Vorname, Geburtsdatum oder die Versicherungsnummer. Da Versicherungsnummer kürzer und einfacher ist, wählen wir diese als Primärschlüssel.</p> <p>Als Faustregel gilt, dass der Primärschlüssel</p> <ul style="list-style-type: none"> - möglichst einfach sein sollte (geringe Anzahl von Attributen) und - besser eine ganze Zahl als eine Zeichenkette sein sollte.
NULL-Werte	<p>Wenn Werte nicht bekannt oder nicht zutreffend sind, dann sagt man sie seien NULL. Es muss in einer Domäne explizit angegeben sein, dass NULL zu dieser Domäne dazugehört. Kein Attribut des Primärschlüssels darf den NULL-Wert enthalten.</p>

3.2.1.3 Integritätsbedingungen

Integrität bedeutet, dass die Datenbank keine widersprüchlichen Daten enthält. Lokale Integritätsbedingungen gelten jeweils für eine Tabelle.

Beispiel: *Identifikation der Zeilen der Tabelle Reichstagslegislaturperiode*
Legislatur ist Primärschlüssel für Tabelle Reichstagslegislaturperiode, d.h. die Legislaturperiode muss eindeutig identifiziert werden, die Werte dürfen also nur einmal in dieser Tabelle vorhanden sein.

Reichstagslegislaturperiode

Legislatur	Beginn-Jahr	End-Jahr
12	1898	1903
13	1903	1907
14	1907	1912
15	1912	1918
16	1919	1920
17	1920	1924

↑ jeder Wert
nur einmal

Globale Integritätsbedingungen (referentielle Integrität) bedeutet, dass mehrere Tabellen betroffen sind.

Beispiel: Globale Integrität – Parlamentszugehörigkeit und Reichstagslegislaturperioden

Die erste Tabelle enthält Reichstagsmandate in verschiedenen Legislaturperioden. Jede Legislatur aus der Tabelle Reichstagsmandate muss auch tatsächlich als Legislaturperiode in Reichstagslegislaturperiode enthalten sein. Die Spalte Legislatur ist in Tabelle Reichstagsmandate Fremdschlüssel bzgl. Reichstagslegislaturperiode, d.h. Schlüssel einer anderen Tabelle. Die Integritäts-Bedingung ist, dass alle Werte der Spalte Legislatur in Tabelle Reichstagslegislaturperiode auch in Tabelle Reichstagsmandate enthalten sind. (Man kann nur einem Parlament angehört haben, das auch existierte).

Reichstagsmandate

Legislatur	Name	Vorname
16	Aderhold	Karl
16	Agnes	Lore
12	Agster	Alfred
17	Aderhold	Karl

Reichstagslegislaturperiode

Legislatur	Beginn-Jahr	End-Jahr
12	1898	1903
13	1903	1907
14	1907	1912
15	1912	1918
16	1919	1920
17	1920	1924



Die Integrität hat das relationale DBMS sicherzustellen, nicht der Anwender. Am Beginn dieses Kapitels wurde bereits darauf hingewiesen, dass das relationale Modell aus Objekten, Operatoren und Regeln besteht. Das heißt, wenn ein DBMS relational ist, dann muss es die relationalen Regeln durchsetzen können, wenn der Anwender es wünscht. Zu einem Relationenschema gehören nicht nur die Relationen selbst, sondern auch immer die Integritätsbedingungen, d.h. die Verweise zwischen den Relationen müssen korrekt sein, jeder Wert eines Primärschlüssels taucht nur einmal auf, Fremdschlüssel-Werte sind tatsächlich vorhanden.

3.2.1.4 Zusammenfassung der Eigenschaften von Relationen:

1. Es tritt kein doppeltes Tupel auf, d.h. zu keinem Zeitpunkt gibt es zwei Tupel, die den gleichen Wert haben.
2. Die Tupel-Reihenfolge ist nicht definiert, d.h. man darf sich nicht auf eine bestimmte Reihenfolge verlassen.
3. Die Attributreihenfolge ist nicht definiert (Reihenfolge der Spalten).
4. Attributwerte sind atomar. Die Werte eines Attributs unterliegen einer Domäne, d.h. sie entstammen einem Wertebereich. Da alle Elemente einer Domäne atomar sind, gilt dies auch für Attribute.

3.2.2 Relationale Algebra

Die Relationentheorie hat ihren Ursprung in der Algebra. In der Algebra wird gerechnet. Das lässt vermuten, dass man mit den Relationen auch rechnen kann. Tatsächlich gibt es einige „Rechenoperationen“, die man auf Relationen anwenden kann. Ein relationales DBMS muss diese Rechenoperationen mit Relationen auch unterstützen, sonst kann es sich nicht relationales System nennen.

Die wichtigsten Operationen mit Relationen (Tabellen) sollen im Folgenden vorgestellt werden:

3.2.2.1 Operation SELECTION – Auswahl von Zeilen aus der Tabelle

Notation: `SEL(<relation>, <bedingung>)`

Beispiel: `SEL(RTKANDIDATUR, NAME = 'Agnes')`

Es werden aus der Relation RTKANDIDATUR alle Zeilen ausgewählt und zu einer neuen Relation zusammengestellt, bei denen in der Spalte NAME der Wert 'Agnes' auftritt.

RTKANDIDATUR

Identnr	Name	Jahr
10050	Aderhold	1919
10060	Agnes	1919
10070	Agster	1898
10080	Ansorge	1920
10060	Agnes	1920
10060	Agnes	1924

SEL(rtkandidatur, name='Agnes')

Identnr	Name	Jahr
10060	Agnes	1919
10060	Agnes	1920
10060	Agnes	1924

3.2.2.2 Operation PROJECTION – Auswahl von Spalten der Tabelle

Notation: PROJ(relation, attr₁, attr₂, ..., attr_n)

Beispiel: PROJ (LTMandat, LAND)

Aus der Tabelle LTMANDAT (Landtagsmandate) wird die Spalte Land ausgewählt und daraus eine neue Relation gebildet.

LTMANDAT

Identnr	Name	Jahr	Land
10010	Abelmann	1919	Schaumburg-Lippe
10010	Abelmann	1922	Schaumburg-Lippe
10010	Abelmann	1925	Schaumburg-Lippe
10020	Ackermann	1919	Bayern
10030	Adams	1924	Hamburg
10040	Adelung	1902	Hessen
10040	Adelung	1905	Hessen
10040	Adelung	1911	Hessen

PROJ(LTMANDAT, LAND)

Land
Schaumburg-Lippe
Bayern
Hamburg
Hessen

Bemerkung:

Entfernen von doppelten Tupeln geschieht automatisch. (Warum: in der Mathematik sind Relationen definiert als Menge von Tupeln. Eine Menge enthält per Definition keine doppelten Elemente.)

3.2.2.3 Operation JOIN – Verbund

Notation: JOIN (MANDATE, LANDTAGE)

Der JOIN verknüpft zwei Tabellen über Spalten mit gleichen Werten, indem jeweils zwei Tupel verschmolzen werden, falls dort gleiche Werte auftreten.

Beispiel: *Verbinden von Mandaten und Landtagslegislaturperioden*

LANDTAGE			MANDATE	
Land	Jahr	LandtagsNr	LandtagsNr	Name
Baden	1921	24	24	Engler
Württemberg	1920	27	24	Freidhof
Baden	1921	24	24	Gehweiler
Baden	1921	24	24	Graf
Württemberg	1920	27	27	Kinkel
Württemberg	1920	27	27	Müller
Württemberg	1920	27	27	Schneck
Württemberg	1920	27	27	Steinmayer

JOIN (MANDATE, LANDTAGE)

Land	Jahr	LandtagsNr	LandtagsNr	Name
Baden	1921	24	24	Engler
Baden	1921	24	24	Freidhof
Baden	1921	24	24	Gehweiler
Baden	1921	24	24	Graf
Württemberg	1920	27	27	Kinkel
Württemberg	1920	27	27	Müller
Württemberg	1920	27	27	Schneck
Württemberg	1920	27	27	Steinmayer

Die Operation JOIN besitzt einige Besonderheiten. Die dargestellte Form nennt man auch INNER-JOIN. Wenn für einen Wert in der einen Tabelle kein passender Wert in der anderen Tabelle auftritt, wird die entsprechende Zeile nicht in die Ergebnisrelation übernommen. Im Beispiel kommt kein Mandat für das Land Bayern vor. Bayern tritt damit nicht in der Ergebnisrelation auf.

Neben einigen anderen Varianten unterscheidet man noch den im Datenbankmanagementsystem ACCESS verwendeten OUTER-JOIN. Der OUTER-JOIN erzeugt mindestens ein Tupel aus den Zeilen einer der beiden Relationen. Der OUTER-JOIN wird in der Praxis oft benötigt, ist aber nicht in der Daten-

bankabfragesprache ANSI-SQL (als Standard akzeptierter SQL-Dialekt) definiert.

Beispiel: *OUTER-JOIN*, gebildet aus der Sichtweise der Relation LANDTAGE

JOIN (LANDTAGE, MANDATE)

Land	Jahr	LandtagsNr	LandtagsNr	Name
Baden	1921	24	24	Engler
Baden	1921	24	24	Freidhof
Baden	1921	24	24	Gehweiler
Baden	1921	24	24	Graf
Württemberg	1920	27	27	Kinkel
Württemberg	1920	27	27	Müller
Württemberg	1920	27	27	Schneck
Württemberg	1920	27	27	Steinmayer
Bayern	1920	NULL	NULL	NULL

Da für den Landtag von Bayern 1920 (Landtagsnummer 33) keine sozialdemokratischen Landtagsmandate aufgelistet sind, müssen in die entsprechenden Spalten NULL-Werte eingetragen werden.

Der OUTER-JOIN aus der Sichtweise der Mandate liefert in diesem Falle das gleiche Ergebnis wie der INNER-JOIN. Würde die Tabelle MANDATE Mandate enthalten, für die noch keine Landtagsnummer angegeben ist oder ein Landtag der nicht in der Liste der Landtage enthalten ist, so würde das Mandat in der Ergebnistabelle erscheinen, aber mit NULL-Werten für Land und Jahr.

3.3 Zusammenfassung

Vereinfacht ausgedrückt, ist ein relationales Datenbankmanagementsystem ein System, das das relationale Modell umsetzt, d.h. es muss

- Daten in Form von Tabellen verwalten,
- die Operationen der relationale Algebra ermöglichen und
- Integritätsbedingungen durchsetzen.

Das können wir als Mindestfunktionalität von MS ACCESS erwarten. Es entspräche jedoch nicht modernem Programmdesign, müsste man die Verknüpfungsoperationen für Relationen in der hier behandelten Form eingeben. Die Operationen und Bedingungen lassen sich menügesteuert durch Abfragen realisieren. Sie wurden an dieser Stelle so ausführlich behandelt, um die Bedeutung von Abfragen besser verstehen zu können.

Aufgabe 3.1 Ist jede Tabelle eine Relation?

- a. Ja
- b. Nein

Aufgabe 3.2 Was sind NULL-Werte?

- a. die Zahl 0
- b. die Buchstabenfolge 'NULL'
- c. nicht bekannte oder nicht zutreffende Werte

Aufgabe 3.3 Welchen Primärschlüssel für eine Tabelle mit den Grunddaten für Häftlinge halten Sie für geeignet?

- a. das Geburtsdatum
- b. Nachname und Vorname
- c. eine laufende Nummer
- d. keine von diesen

Wählen Sie die für Sie richtige Antwort aus und begründen Sie Ihre Entscheidung.

4. Entity Relationship Model

Eine Datenbank soll einen Ausschnitt der realen Welt abbilden. Das relationale Modell gestattet es, auf einfache Weise Daten zu verwalten. Es ist aber als Beschreibungs- und Analyseinstrument realer Zusammenhänge nicht sehr geeignet. Der Abstraktionsschritt von realen Strukturen zu Relationen ist zu groß, um komplizierte Zusammenhänge erfassen zu können. Daher hat man sich Beschreibungsmittel geschaffen, die geeigneter sind, reale Zusammenhänge und Strukturen darzustellen und die sich mit formalen Mitteln, d.h. durch einfache Regeln in ein Relationenschema überführen lassen. Das ER-Modell geht auf Peter P. S. Chen (1976) zurück. Es gilt als Standardmodell für den Datenbankentwurf.

4.1 Grundkonzepte des klassischen ER-Modells

Wir wünschen uns ein semantisches Datenmodell (SDM), das uns helfen soll, die abzubildende Miniwelt besser zu durchdringen und präziser zu beschreiben. Es soll frei von Realisierungsdetails, unabhängig von konkreten Datenbankmanagementsystemen und deren Datendefinitionssprachen (DDL) sein.

In der Literatur ist eine Vielzahl von SDM beschrieben worden. Vielen gemeinsam ist, dass sie im Kern auf dem allgemeinen Systembegriff der Kybernetik beruhen.

Wir wollen hier den Systembegriff folgendermaßen fassen:

System	Ein System ist eine Menge von Objekten unserer Anschauung oder unseres Denkens, zwischen denen wohldefinierte Beziehungen bestehen.
--------	---

Bekannteste Klasse der SDM ist das Entity Relationship Model (ERM). Entsprechende deutschsprachige Begriffe für Entity und Relationship sind Objekt, Gegenstand, Ding, Entität bzw. Beziehung, Relation. Man sollte diese deutschen Begriffe vermeiden, da sie teilweise mit einer anderen Bedeutung belegt sind. Z.B. hat der Begriff 'Relation' im Zusammenhang mit relationalen Datenbanken eine andere Bedeutung als hier.

Das ERM basiert auf 3 Konzepten:

- Entities
- Attribute
- Relationships (Beziehungen)

In Datenbanken sind Werte gespeichert. Datenbankmodelle beinhalten also Konzepte und Regeln zur Speicherung von Datenwerten. Auch das ER-Modell ist ein Instrument zur Beschreibung von Daten.

4.1.1 Werte

Werte sind Datenelemente, die direkt angegeben werden können, wie z.B. eine Zahl für ein Jahr oder eine Zeichenfolge als ein Personennamen. Die Wertemengen entsprechen den elementaren Datentypen, die neben den Werten auch darauf definierte Grundoperationen charakterisieren. Im ERM sind nur Standarddatentypen definiert, wie sie aus Programmiersprachen bekannt sind, z.B. Ganzzahlen (integer), Kommazahlen (real), Datum, Zeichenkette (string) und logische Größen (Booleans) mit den jeweils dazugehörenden Rechenoperationen.

4.1.2 Entities

Entities sind zu repräsentierende Informationseinheiten. Sie sind nicht direkt darstellbar, sondern nur über ihre Eigenschaften beobachtbar.

Beispiel: *Entity-Typen*

Ein Mandatsträger, eine Legislaturperiode, ein Buch oder ein Parlamentsausschuss können Entities sein. Auch Ereignisse wie Reichstagswahlen können Objekte im Sinne des ER-Modells sein.

Es werden nicht die einzelnen Gegenstände, sondern jeweils Typen (Klassen) von Gegenständen betrachtet (Generalisation mittels Mengen). Das gilt analog bei Beziehungen.



Abb. 4.1: Darstellung eines Entity-Typs

4.1.3 Beziehungen

Beziehungen zwischen Entities werden zu Beziehungstypen zusammengefasst. Die Beziehung wird durch eine Raute symbolisiert. Eine beliebige Anzahl n (≥ 2) von Entity-Typen kann an einer Beziehung beteiligt sein. R sei eine Beziehung (Relationship) und E_1, E_2, \dots, E_n Mengen von n Entity-Typen. Eine n -stellige Beziehung wird in folgender Weise dargestellt:

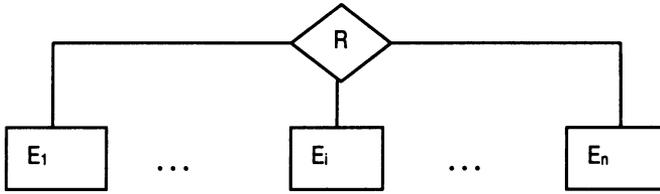


Abb. 4.2: Darstellung von Relationships

Beispiel: *Zweistellige Beziehung*
 ‘Eine Person kandidiert für einen Landtag.’ Das Wort ‘kandidiert’ drückt eine Beziehung aus zwischen Person und Landtag.



Abb. 4.3: Landtagskandidaten

Die Anzahl der beteiligten Entity-Typen ist hier $n = 2$.

Wenn ein und derselbe Entity-Typ mehrfach an einem Beziehungstyp beteiligt ist, dann müssen Rollennamen vergeben werden.

Beispiel: Verheiratet (Frau: Person, Mann: Person)



Abb. 4.4: Rollennamen

Bei der textuellen Notation kann der Rollenname durch die Parameterposition ersetzt werden, z.B. soll Mann immer zuerst stehen. Bei der graphischen Darstellung ist die Vergabe von Rollennamen zwingend.

4.1.4 Attribute

Attribute modellieren Eigenschaften von Entities oder Beziehungen. Alle Entities eines Entity-Typs haben dieselbe Art von Attributen, z.B. der Name oder das Alter einer Person, der Titel einer Vorlesung, oder das Semester, in dem

eine Vorlesung gehalten wird. Graphisch werden Attribute durch Rechtecke mit abgerundeten Ecken dargestellt.

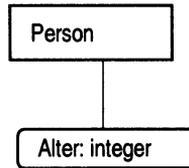


Abb. 4.5: Darstellung von Attributen

Alter ist der Attributname, integer der Wertebereich. Wenn der Wertebereich durch das Attribut indirekt vorgegeben ist, wird er gewöhnlich weggelassen. Im klassischen ERM sind nur Standarddatentypen (Ganzzahlen, Kommazahlen, Datum, Zeichenkette, logisch) zugelassen.

4.1.5 Zweistellige versus mehrstellige Notation

Häufig wird die Einschränkung vorgenommen, dass nur zweistellige Relationships zugelassen sind. Auch wir wollen diese Einschränkung vornehmen. Das bedeutet allerdings eine Einschränkung der Modellierung. Dazu folgendes Beispiel:

Beispiel: *Professor hält Vorlesung und empfiehlt zur Vorlesung ein Buch.* Realitätskonform ist die Darstellung in Abb. 4.6. Die Darstellung in Abb. 4.7 ist bedenklich. Nicht die Vorlesung empfiehlt, sondern der Professor. Er empfiehlt aber nicht schlechthin, sondern nur im Zusammenhang mit der Vorlesung. Würde ein anderer Professor die Vorlesung halten, würde dieser evtl. ein anderes Buch empfehlen (Dreistellige Beziehung). Die Attribute wurden wegen der besseren Übersichtlichkeit weggelassen.

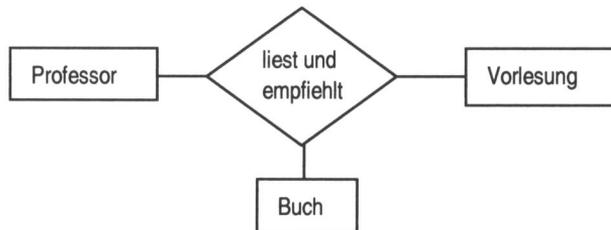


Abb. 4.6: Dreistellige Beziehung zwischen Professor, Vorlesung und Buch

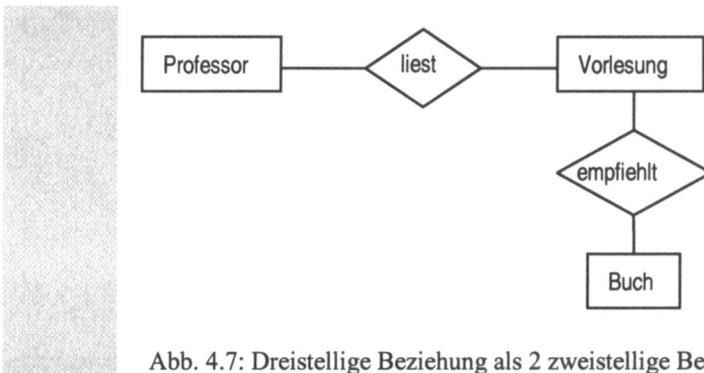


Abb. 4.7: Dreistellige Beziehung als 2 zweistellige Beziehungen
 Man kann zeigen, dass eine Zerlegung der Dreier-Beziehung in drei Zweier-Beziehungen nicht den gleichen Informationsgehalt besitzt.

4.2 Weitere Konzepte im ERM

Mit dieser klassischen Form des ERM sind nur einfache Zusammenhänge zufrieden stellend zu modellieren. Um kompliziertere Zusammenhänge adäquat darzustellen, wurden weitere Modellierungskonzepte vorgeschlagen:

- Schlüsselattribute,
- Kardinalitäten und Optionalitäten von Attributen und Beziehungen,
- funktionale Beziehungen,
- Spezialisierung mittels Ist-Beziehung,
- abhängige Entities,
- Erweiterung des ERM um höhere Konzepte bezüglich der Beziehungen, der Entities und der Wertebereiche.

Eingegangen wird an dieser Stelle auf die ersten vier Konzepte. Sie sind ohne tiefere Kenntnisse der Informatik zu verstehen. Viele in der Praxis des Sozialwissenschaftlers auftretende Zusammenhänge lassen sich mit ihnen noch hinreichend gut beschreiben.

4.2.1 Identifizierung durch Schlüssel

Oft identifizieren einige Attribute mit ihren Werten die Entities eines Entity-Typs.

Beispiel: Identifizierung durch Schlüssel

- Parlamentarier durch eine Nummer (ID),
- Studenten durch ihre Matrikelnummer,
- Bücher durch ISBN,
- Angestellte durch Name, Vorname und Geburtsdatum.

Eine solche Menge von Attributen wird analog zum Relationenmodell als Schlüssel bezeichnet. Bei der textuellen und der graphischen Notation werden die Schlüsselattribute unterstrichen. Treten verschiedene mögliche Schlüssel auf, so nennt man sie wie im relationalen Modell Schlüsselkandidaten. Ein ausgewählter wird Primärschlüssel.

4.2.2 Die IST-Beziehung

Eine häufig auftretende Beziehung zwischen Entity-Typen ist die Spezialisierungs-/Generalisierungsbeziehung, auch IST-Beziehung genannt. Dabei handelt es sich um eine Teilmenge, die zu einer umfassenderen Menge gehört.

Beispiel: *IST – Beziehung Parlamentsausschussmitglied*
 Im Parlament sitzen Mandatsträger. Einige Mandatsträger sind Ausschussmitglieder in bestimmten Ausschüssen.

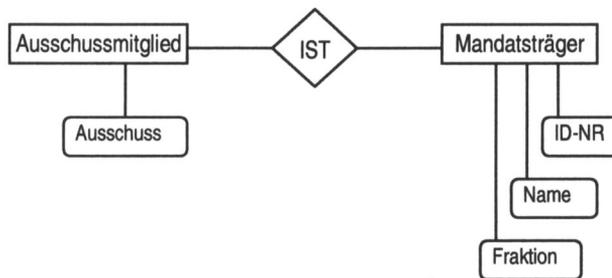


Abb. 4.8: Mandatsträger als Ausschussmitglieder

4.2.3 Kardinalitäten

Mit Kardinalität bezeichnet man die Anzahl der möglichen Beteiligung eines Entity-Typs an einem Beziehungstyp.

Beispiel: *Zugehörigkeit der Abgeordneten zu Fraktionen.*
 Jeder Abgeordnete ist in der Regel einer Fraktion zuzuordnen, jeder Fraktion können theoretisch maximal 647 Mitglieder angehören (Reichstagswahl 5.3.1933, Anzahl der Abgeordneten total 647).

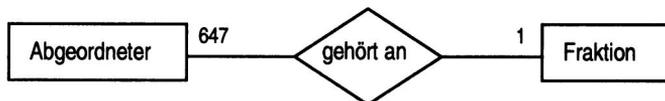


Abb. 4.9: Optionale Beziehung

Man gibt im Allgemeinen nicht die tatsächliche maximale Anzahl an (hier 647), sondern schreibt „n“ oder „*“ für mehr als eine Beteiligung.

Wenn alle Abgeordneten eindeutig Fraktionen zugeordnet werden können, spricht man auch von einem funktionalen Zusammenhang. Der Zusatz ‘in der Regel’ soll zum Ausdruck bringen, dass es evtl. auch Abgeordnete geben kann, die keiner Fraktion angehören. Wenn, wie in diesem Fall, nicht alle Entities an der Beziehung beteiligt sind, spricht man von einer partiellen funktionalen Beziehung (auch optionale Beziehung).

Beispiel: *Betreuung der Fraktionen durch Mitarbeiter (jedem Entity A ist genau ein Entity Typ B zugeordnet)*
Jeder Mitarbeiter ist genau einer Fraktion zugeordnet, für die er zu arbeiten hat.



Abb. 4.10: Totale funktionale Beziehung

Lesart: Jeder Mitarbeiter ist genau einer Fraktion zugeordnet, einer Fraktion können mehrere Mitarbeiter zugeordnet sein.

Auch dies ist ein funktionaler Zusammenhang. Da aber alle Mitarbeiter Fraktionen zugeordnet sind (eine Fraktion kann auch mehrere haben), spricht man von einer totalen funktionalen Beziehung (auch zwingende Beziehung).

5. Darstellung der Parlamentarier durch Konzepte des ERM

Komplizierte Zusammenhänge lassen sich mit den Grundkonzepten des ERM nicht ohne Einschränkungen und ohne Informationsverlust darstellen. Daher wurden viele Erweiterungen vorgeschlagen, die aber für unseren Forschungsgegenstand „Parlamentarismus“, nicht zwingend notwendig sind. Wir wollen die in unserer Datenbank zu beschreibende ‚Mini-Welt‘, die Sozialdemokratischen Parlamentarier in den deutschen Reichs- und Landtagen 1867 - 1933, mit dem klassischen Entity Relationship Model (ERM) beschreiben. Wir untersuchen Struktur und Wandel der deutschen Arbeiterbewegung im Hinblick auf ihr politisch-parlamentarisches Personal. Gesucht werden Lebensverläufe, Gemeinsamkeiten und Veränderungen der nach verschiedenen Gesichtspunkten gruppierten Parlamentarier.

5.1 Was ist unsere ‚Miniwelt‘?

In der Anforderungsanalyse haben wir Informationen aus der Forschungspraxis gesammelt und tabellarisch zusammengestellt. Diese Tabellen sollen als Ausgangspunkt für eine formale Beschreibung der Parlamentarier dienen. Als augenfälliger Entity-Typ zeichnen sich Personen ab. Sie treten in unserer Betrachtung zeitweilig als Reichstagskandidaten und Parlamentarier auf. Da wir auch erfolglose Kandidaturen erfassen wollen, ist es sinnvoll, den generalisierenden Begriff ‚Personen‘ zu wählen und diesen als Entity-Typ zu behandeln. Eine Differenzierung zwischen Personen, die für Landtage und solche die für Reichstage kandidiert haben, scheint nicht sinnvoll, da einige sowohl im Reichstag als auch in Landtagen vertreten waren. Dem Entity-Typ Personen lassen sich viele Attribute zuordnen, was wir aber der Übersichtlichkeit halber erst später vornehmen wollen.

5.2 Entity-Typ Reichstag

Neben den Entity-Typ Personen tritt scheinbar ein zweiter Typ: Parlament. Unser Interesse gilt jedoch verschiedenen Erscheinungsformen von Parlamenten, die in Bezug auf das Forschungsziel eine sehr unterschiedliche Stellung besitzen:

- Reichstage,
- Landtage,
- Bundestage,
- Volkskammer.

Die Beschränkung auf den Typ Parlament erlaubt nur eine sehr unscharfe Betrachtung. Es interessieren nicht Parlamente schlechthin, sondern Zugehörigkeit zu einem bestimmten Parlament in einer Legislaturperiode. Es interessiert also nicht der Reichstag als Institution, sondern Reichstage als für jeweils eine Legislaturperiode gewählte Parlamente. Das gilt analog für Landtage.

Das Forschungsziel umfasst nicht primär eine Analyse des Werdegangs nach 1945. Für uns ist lediglich von Interesse, ob der Abgeordnete nach 1945 noch politisch hervortrat. Als Indikator für politische Aktivität dient uns eine Mitgliedschaft im Deutschen Bundestag, in einem Landtag oder in der Volkskammer der DDR nach 1945.

Eine Reichstagskandidatur führt nicht notwendig zu einer Reichstagsmitgliedschaft (Reichstagsmandat). Es soll unterschieden werden zwischen Reichstagsmitgliedschaft und erfolgloser Reichstagskandidatur (im Folgenden soll unter Reichstagskandidatur die erfolglose Kandidatur verstanden werden). Das ER-Diagramm für Reichstagskandidaturen soll schrittweise aufgebaut werden. Aus der einfachen Aussage 'eine Person kandidiert für den Reichstag' lässt sich im ersten Schritt folgendes ER-Diagramm ableiten:



Abb. 5.1: Personen und Reichstage

Mehrere Bewerber (n) konkurrieren für einen Reichstag und ein Bewerber kann für mehrere Reichstage (m) kandidiert haben.

Die Kandidatur erfolgt in einem oder mehreren Wahlkreisen, in einem Wahlkreis gibt es für eine Wahl mehrere Bewerber. Die Person kandidiert in diesem Wahlkreis für einen bestimmten Reichstag. Man könnte Wahlkreise als Attribute der Beziehung Kandidatur sehen, d.h. nicht als selbständige Objekte, als Entities. Attribute sind, wie in Kapitel 4 ausgeführt, atomar. Sie bestehen daher jeweils nur aus einem einzelnen Wert eines Datentyps, hier aus einer Zeichenfolge (string) für den Wahlkreisnamen. Selbst wenn die Namen eindeutig sind und damit jeder Wahlkreis identifiziert werden kann, bleibt der Nachteil bestehen, dass kein überregionaler Bezug hergestellt werden kann. Will man beispielsweise alle bayerischen Wahlkreise selektieren, wäre das nicht möglich.

Betrachtet man Wahlkreise als Entity-Typ, so kann man diesem Wahlkreisname, überregionale Verwaltungseinheit (Land, ggf. auch Provinz) und evtl. ein Bemerkungsfeld mit weiteren regionalen Besonderheiten als Attribute zuordnen.

Wahlkreise sind im Untersuchungszeitraum keine konstanten territorialen Gebilde. Neben den starken Veränderungen zu Beginn der Weimarer Republik, gab es weitere kleinere Veränderungen. Daher wurde eine Nummer als Identifikator über alle Legislaturperioden hinweg eingeführt. Betrachtet man Wahlkreise als eigenständige Objekte, so folgt daraus eine Beziehung zwischen den drei Entity-Typen Person, Wahlkreis und Reichstag.

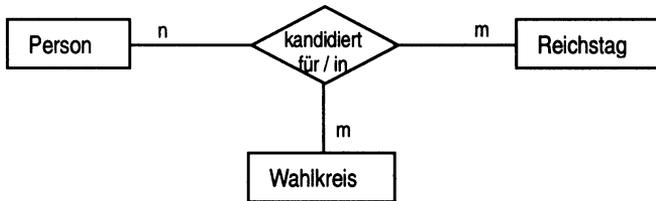


Abb. 5.2: Personen, Reichstage und Wahlkreise

Es sind noch die Attribute anzufügen. Die Wahl erfolgte an einem Wahldatum. Das Wahldatum ist am direktesten mit 'kandidiert' verbunden, ist damit Attribut der Beziehung kandidiert. Die Attribute für Reichstag sind die Nummer der Legislaturperiode (NrLP), die beiden Zeitangaben von der Konstituierung des Reichstages bis zu seiner Auflösung, bzw. bis zur Konstituierung des nächsten Reichstages. Der Wahlkreis wird durch die Attribute Wahlkreisnummer (WahlkreisNR), Wahlkreisname, und weiteren geographisch-territorialen Beschreibungen, zusammengefasst als GeoSpez bezeichnet.

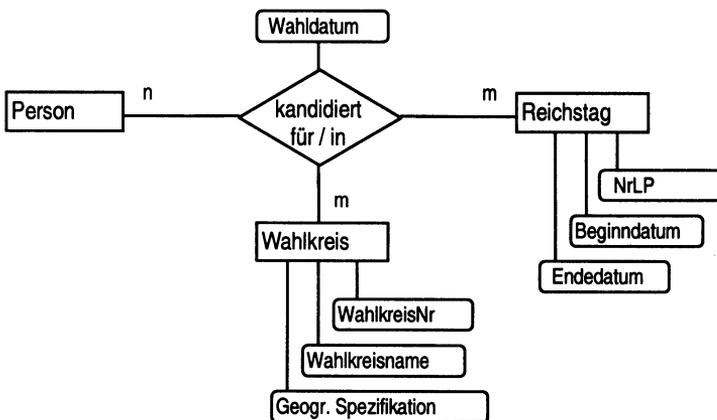


Abb. 5.3: Personen kandidieren in einem oder mehreren Wahlkreisen für Reichstage.

Die Mitgliedschaft im Reichstag setzt eine erfolgreiche Kandidatur voraus. Die Kandidatur in der Datenbank zusätzlich aufzuzeichnen, bringt nur geringen Informationsgewinn.

Die persönliche Mitgliedschaft eines Mandatsträgers im Reichstag muss nicht über die volle Legislaturperiode gehen. Es ist eine individuelle Zeitspanne, die der Beziehung zwischen Person und Reichstag zuzuordnen ist.

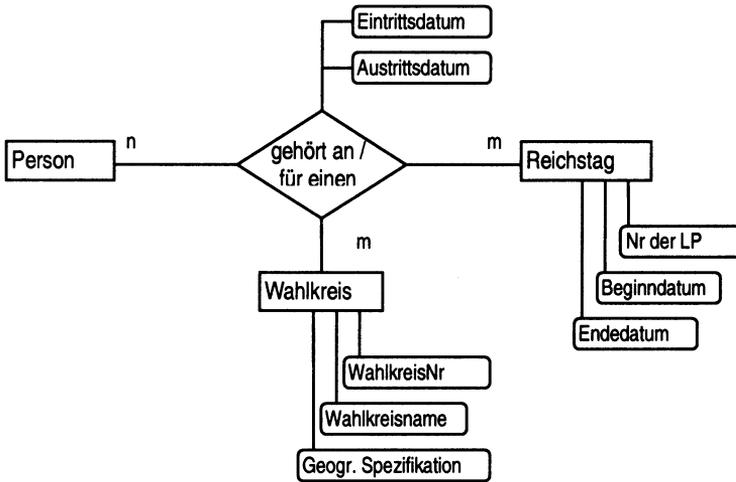


Abb. 5.4: Personen nehmen Reichstagsmandat wahr.

5.3 Entity-Typ Landtag

Die Autoren der Datenbank BIOSOP legten fest, dass für Landtagswahlen keine Wahlkreise zu erfassen sind. Sie erachteten die Angabe des Landes als ausreichend genauen territorialen Bezug. Daher gestaltet sich dieses ER-Diagramm etwas einfacher.

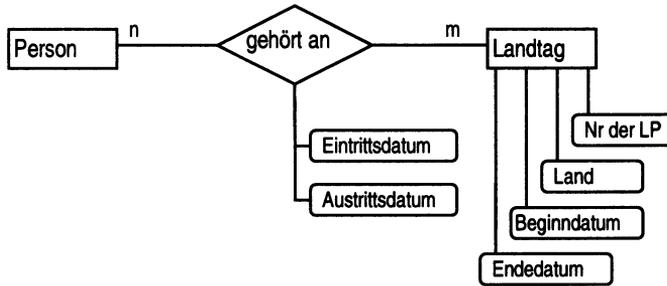


Abb. 5.5: Personen gehören Landtagen an.

5.4 Entity-Typ Parlament

Die Zugehörigkeit zu einem Parlament nach 1945 – Bundestag, Länderparlament oder Volkskammer – beschreibt die folgende Abbildung:

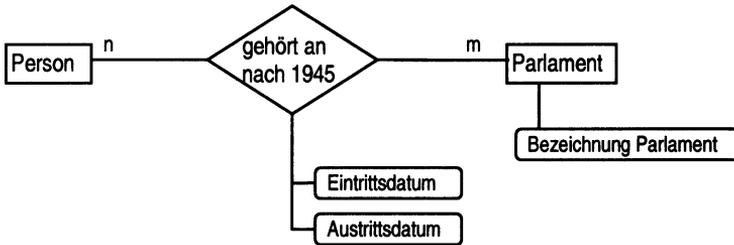


Abb. 5.6: Personen gehören nach 1945 Parlamenten an.

5.5 Entity-Typ Berufsausbildung

Die Bedeutung der Berufsangaben wurde in der Anforderungsanalyse hervorgehoben. Sie werden nicht als Attribut des Entity-Typs Person, sondern als eigenständiger Entity-Typ gesehen. Sie haben nicht nur das Attribut Berufsbezeichnung, sondern ggf. auch mehrere alternative Berufsschlüssel.

Eine Person durchläuft eine Ausbildung zu einem oder mehreren Berufen. Die Ausbildung zu einem Beruf kann an mehreren Orten erfolgen. Erfasst wird die Gesamtausbildungszeit zu diesem Beruf.

Um in der gesamten Datenbank nach Orten recherchieren zu können, werden Orte (Geburts-, Sterbe-, Ausbildungs-, Tätigkeitsorte) als eigenständiger Entity-Typ erfasst.

Bei der Modellierung mit den Konzepten des ERM muss berücksichtigt werden, dass ein Attribut nur jeweils einen einzigen Wert besitzen darf. Ein Wert war ein Element aus einer Grundmenge wie 'ganze Zahl', 'Kommazahl', 'Zeichenkette' oder 'logische Größe'. Wenn man Beruf als Attribut von Person sieht, dann hat das folgende Konsequenzen:

- jede Person hat genau einen Beruf, d.h. wenn jemand Schlosser lernte und danach Maschinenbauer, dann ist das mit einem Attribut nicht darstellbar;
- wenn man zum Beruf einen Berufsschlüssel führen will, muss dieser Schlüssel als Attribut der Person geführt werden und steht nicht direkt, sondern nur über Person mit dem Beruf in Beziehung.

Um eine realitätskonforme Darstellung zu gewährleisten, muss man Informationsgruppen als eigenständigen Entity-Typ behandeln. Für den Fall der Berufsausbildung hieße das, die Informationsgruppe Berufsbezeichnung, Beginn der Berufsausbildung, Ende der Berufsausbildung, Berufsschlüssel und evtl. alternativer Berufsschlüssel sind Attribute des Entities Beruf.

Da auch Orte bei genauerer Betrachtung nicht Attribute mit Werten aus dem Wertebereich Zeichenkette sind, sondern Informationsgruppen, bestehend aus (amtlicher) Ortsbezeichnung, geographischem Ortsschlüssel und übergeordneter Einheit, müssen auch sie als Entity-Typ gesehen werden. Nur so hat man die Möglichkeit z.B. einem Beruf mehrere Berufsausbildungsorte zuzuweisen.

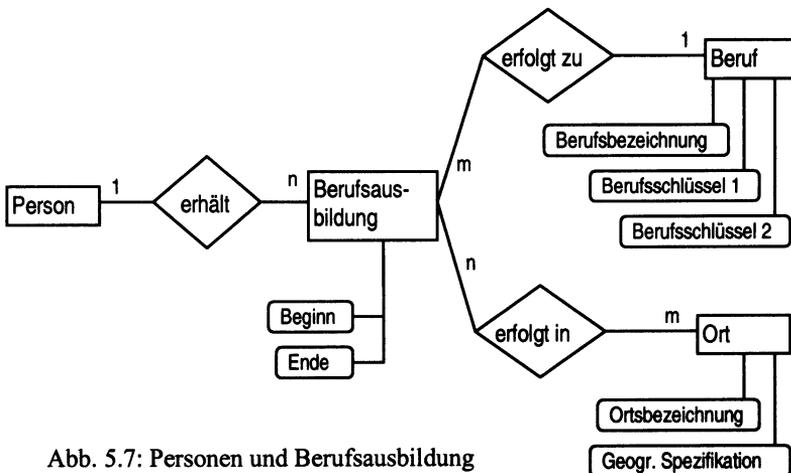


Abb. 5.7: Personen und Berufsausbildung

Damit sind die kompliziertesten Zusammenhänge bereits dargestellt. Exemplarisch soll noch die Modellierung der Publikationstätigkeit, Parteitagsteilnahmen und der Werdegang der Abgeordneten beschrieben werden.

5.6 Entity-Typ Publikationen

Abgeordnete waren in vielfältiger Weise publizistisch tätig. Einige Abgeordnete verfassten mehrere Werke, von anderen ist nichts überliefert. Es ist auch vorstellbar, dass an einer Publikation mehrere Abgeordnete beteiligt waren. Dieser letzte Aspekt soll uns nicht interessieren. Wir verzichten darauf zu recherchieren, welche Abgeordneten gemeinsam publiziert haben. Nicht alle Titel sollen erfasst werden. Bei kleineren Veröffentlichungen soll nur darauf hingewiesen werden, dass der Abgeordnete publizistisch tätig war. Das ER-Diagramm hat dann folgendes Aussehen:

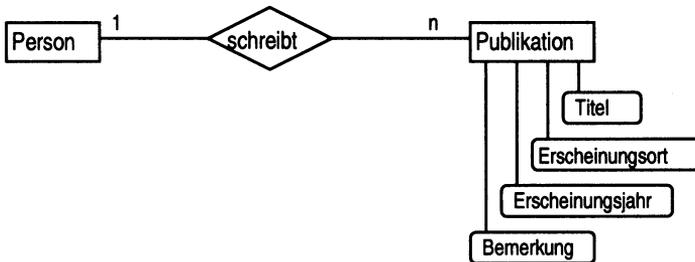


Abb. 5.8: Personen und Publikationstätigkeit

5.7 Entity-Typ Parteitagsteilnahme

Parteitage waren Ereignisse, die zu einem bestimmten Datum (Monat, Jahr) an einem Ort von einer Arbeiterpartei abgehalten wurden. Ein Parteimitglied kann zu mehreren Parteitagen delegiert worden sein. An einem Parteitag nahmen viele Abgeordnete teil.

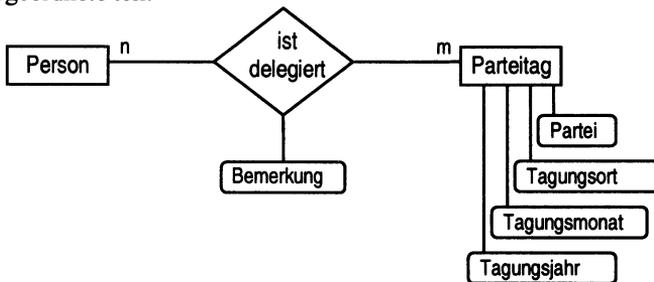


Abb. 5.9: Personen und Parteitage

5.8 Entity-Typ Werdegang

Der Werdegang einer Person wird durch eine Folge von einzelnen Berufspositionen beschrieben. Jede Berufsposition wird in einem Zeitabschnitt ausgeübt. Häufig sind diese Zeitabschnitte unsichere Informationen oder sie sind mit Zusätzen wie 'seit' oder 'bis' versehen. In manchen Fällen ist nur die Reihenfolge der Positionen, aber kein Datum bekannt. Die Datumsangaben interessieren nur als Monats- und Jahresangaben. Eine Position kann an mehreren Orten ausgeübt worden sein. Daraus ergeben sich für Werdegang folgende Attribute:

Attribut	Wert	Bedeutung
lfdnr	ganze Zahl	fortlaufende Nummerierung der Tätigkeit
vonMonat	ganze Zahl < 12	Monat, ab der die Berufsposition innegehabt wurde
vonJahr	1850 < ganze Zahl < 1950	Jahr, ab der die Position innegehabt wurde
vonSicher	Ja Nein	Hinweis darauf, ob die Datumsangabe ab der die Position innegehabt wurde, gesichert ist
von	'nach' 'seit'	Hinweis zum Zeitverlauf, ab der die Position innegehabt wurde
bisMonat	ganze Zahl < 12	Monat, bis zu der die Position innegehabt wurde
bisJahr	1850 < ganze Zahl < 1950	Jahr, bis zu der die Position innegehabt wurde
bisSicher	Ja Nein	Hinweis darauf, ob die Datumsangabe bis zu der die Position innegehabt wurde, gesichert ist
bis	'bis'	Hinweis zum Zeitverlauf, bis zu der die Position innegehabt wurde
Hinweis	Zeichenkette	die Berufsposition in freiem Text

Die Tätigkeiten werden als gesonderter Entity-Typ aufgefasst, um mit einem Tätigkeitsschlüssel, analog zu Berufsschlüssel, Klassifikationen durchführen zu können. Auch Tätigkeitsorte werden analog den Berufsarten als Entity-Typ geführt.

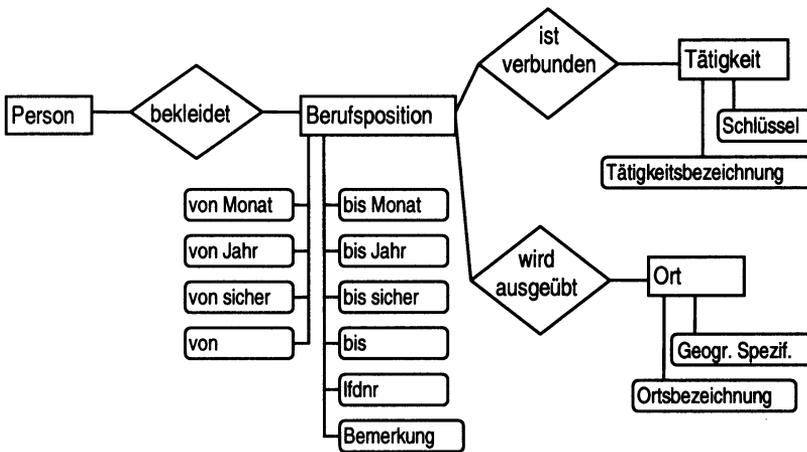


Abb. 5.10: Werdegang von Abgeordneten

5.9 Entity-Typ Personen

Die folgende Tabelle enthält die Attribute des Entity-Typs Person.

Attribut	Wert	Beschreibung
ID	Ganze Zahl	Nummer zur eindeutigen Identifikation
Nachname	Zeichenfolge	Nachname
Vorname	Zeichenfolge	Vorname
Geschlecht	'm' 'w'	Geschlecht; nur 'm' oder 'w' ist möglich
Geburtstag	0 < ganze Zahl <= 31	Geburtstag des Abgeordneten; Nichtbelegung möglich
Geburtsmonat	0 < ganze Zahl <= 12	Geburtsmonat des Abgeordneten; Nichtbelegung möglich
Geburtsjahr	1800 < ganze Zahl < 1915	Geburtsjahr; Nichtbelegung möglich
Sterbetag	0 < ganze Zahl <= 31	Todestag; Nichtbelegung möglich
Sterbemonat	0 < ganze Zahl <= 12	Todesmonat; Nichtbelegung möglich
Sterbejahr	1863 < ganze Zahl < 2000	Todesjahr; Nichtbelegung möglich

5.10 Weitere Beziehungen im Überblick

Ohne weitere Erläuterungen seien folgende Beziehungen angeführt:

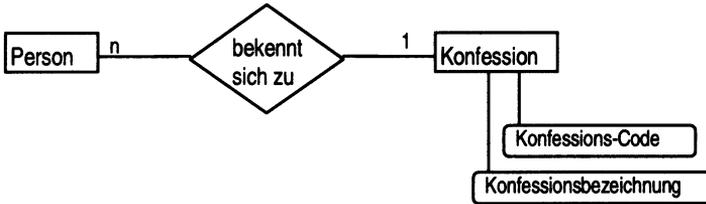


Abb. 5.11: Konfessionsbekenntnis

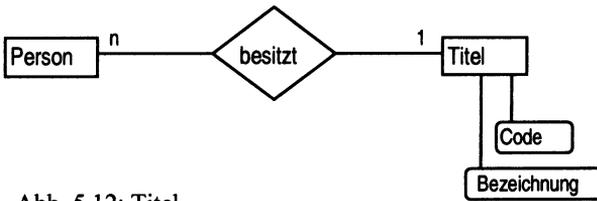


Abb. 5.12: Titel

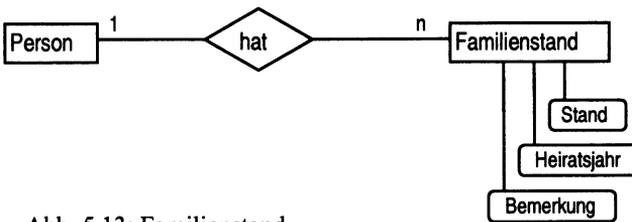


Abb. 5.13: Familienstand

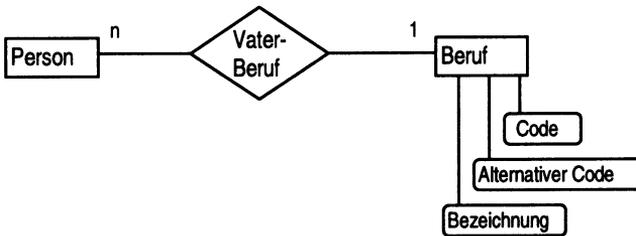


Abb. 5.14: Beruf des Vaters

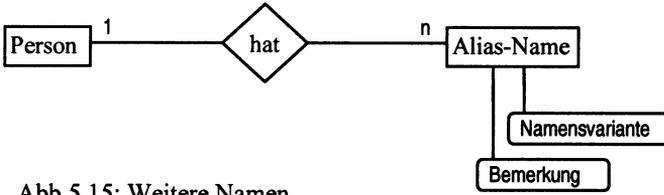


Abb.5.15: Weitere Namen

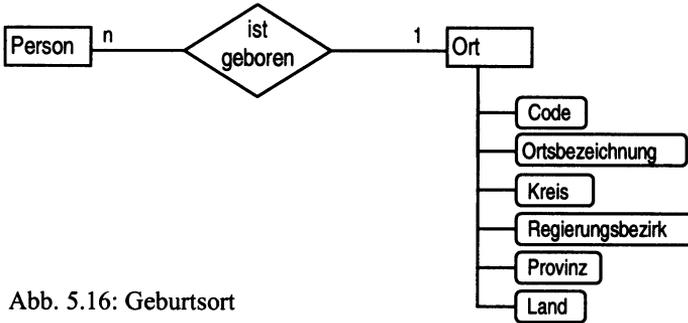


Abb. 5.16: Geburtsort

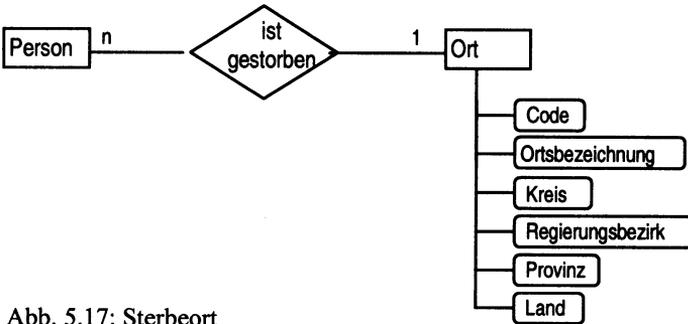


Abb. 5.17: Sterbeort

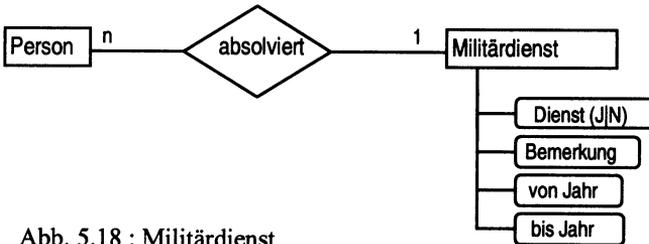


Abb. 5.18 : Militärdienst

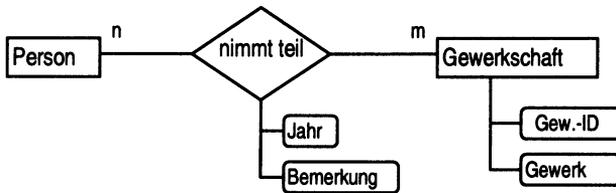


Abb. 5.19: Teilnahme an Gewerkschaftskongressen

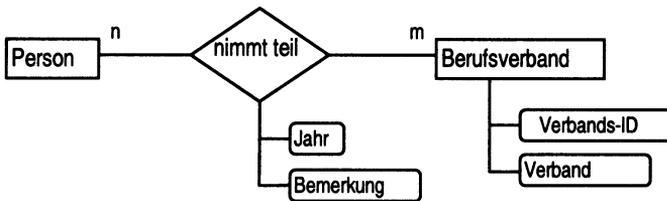


Abb. 5.20: Teilnahme an Berufsverbandstagen

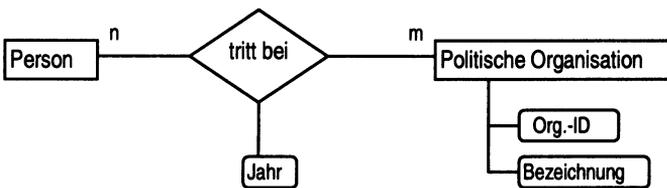


Abb. 5.21: Ersteintritt in eine sozialdemokratische Partei

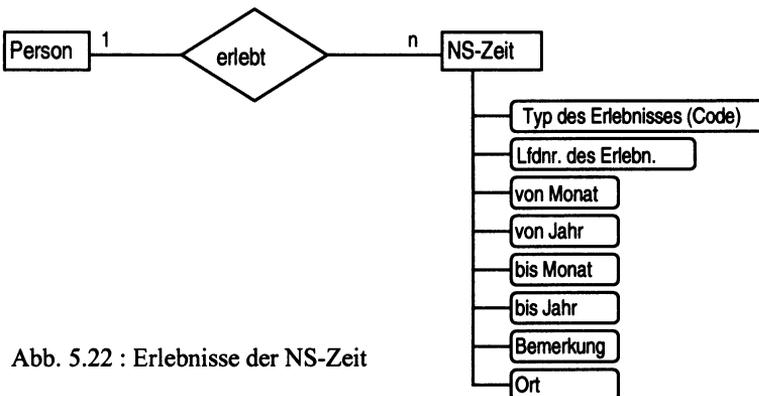


Abb. 5.22 : Erlebnisse der NS-Zeit

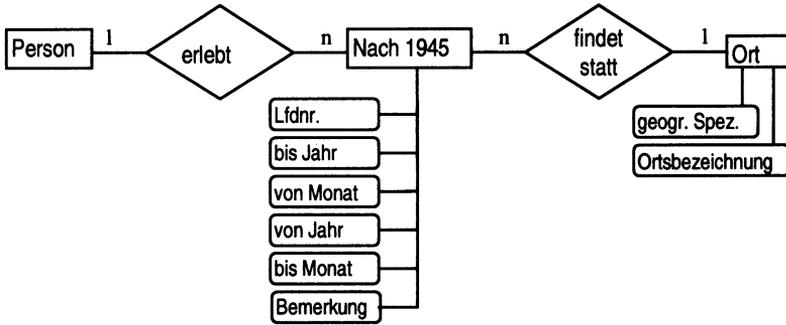


Abb. 5.23: Entwicklung nach 1945

Aufgabe 5.1: Modellieren Sie die Publikationstätigkeit der Parlamentarier für den Fall, dass mehrere Parlamentarier gemeinsam als Buchautoren auftreten können.

6. Abbildung des ER-Diagramms auf das relationale Modell

Der erste Schritt des logischen Entwurfs ist die Übertragung des konzeptuellen Schemas in das Schema des Zielmodells. Konzeptuelles Schema ist die Zergliederung, die wir durch Anwendung des ERM erhalten haben. Zielmodell ist das Datenbankmodell, das dem benutzten Datenbankmanagementsystem zugrunde liegt. Zielmodelle können sein:

- relationales Modell,
- Netzwerkmodell,
- hierarchisches Modell,
- objektorientierte Modelle.

Wir beschränken uns an dieser Stelle auf das relationale Modell. Es bieten sich zwei Wege für diese Übertragung an:

- Transformation nach Faustregeln, manuell. Dieser Weg lässt dem Entwerfer Spielraum.
- Automatische Transformation. Für die automatische Transformation werden so genannte CASE-Tools (Computer Aided Software Engineering) benutzt, die für einige DBMS (z.B. Oracle) entwickelt wurden.

Ziel der Übertragung ist der Erhalt der Informationskapazität. Alle Informationen des ER-Diagramms sollen im Ziel-Datenbankschema erhalten bleiben, es sollen aber auch keine hinzukommen.

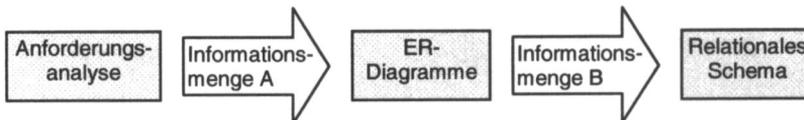


Abb. 6.1: Erhalt der Informationskapazität: $A = B$

Die Informationsmengen A und B sollen identisch sein. Es gelten folgende Grundprinzipien der Übertragung:

- Entity-Typen und Beziehungstypen werden jeweils mit allen Attributen auf eine Relation abgebildet. Die Schlüssel werden übernommen.
- Kardinalitäten werden durch Wahl der Schlüssel in den Relationen ausgedrückt.
- In einigen Fällen können Relationenschemata von Entity- und Beziehungstypen miteinander verschmolzen werden.
- Zwischen den Relationenschemata werden diverse Fremdschlüsselbedingungen eingeführt.

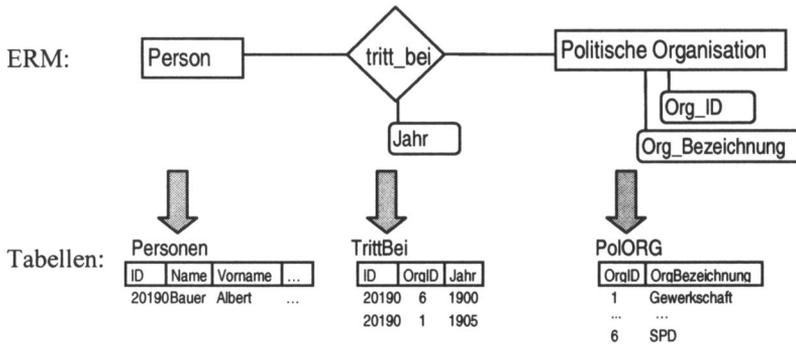


Abb. 6.2: Transformation des ERM in Tabellen

Die folgende Tabelle stellt den ersten Schritt der Übertragung schematisch dar:

<i>Entity Relationship Model</i>		<i>Relationales Modell</i>
Entity-Typ E	wird	Tabelle E
Attribute des Entity-Typs	werden	Spalten der Tabelle
Primärschlüssel des Entity-Typs	wird	Primärschlüssel der Tabelle
Beziehungstyp R	wird	Tabelle R
Primärschlüssel der beteiligten Entity-Typs	werden	Spalten der Tabelle
Attribute der Beziehung	werden	Spalten der Tabelle

Tabelle 6.1: Transformationsschema

Weitere Vereinfachungen der Tabellenstruktur werden in den folgenden Abschnitten beschrieben.

6.1 Entity-Typen

Jeder Entity-Typ ergibt ein relationales Schema mit allen Attributen. Der Schlüssel wird in das Relationenschema übernommen. Wenn mehrere Schlüssel möglich sind, ist einer als Primärschlüssel auszuwählen.

Primärschlüssel sollten möglichst einfach gewählt werden. Besser ist es, ein Attribut auszuwählen statt mehrere, besser eine Ganzzahl als eine Zeichenkette. Einige Datenbankmanagementsysteme bieten Unterstützung an bei der Wahl eines Primärschlüssels. ACCESS fragt, wenn kein Schlüssel deklariert wurde, ob ein Primärschlüssel eingefügt werden soll. Lässt man das zu, wird ein Feld mit Namen ID vom Typ long integer (ganze Zahlen von -2.147.483.648 bis 2.147.483.648 in 4 Byte), eingefügt, dessen Werte das System selbst verwaltet. ACCESS sorgt dafür, dass jeder Wert nur einmal vergeben wird. Der Nutzer

hat keine direkte Möglichkeit, den Wert des Schlüssels zu beeinflussen. Dieser Datentyp wird in ACCESS „AutoWert“ genannt.

Beispiel

Entity-Typ Person

Der Entity-Typ Person hat zwei mögliche Schlüssel: Eine vom Autor vergebene sechsstellige Ziffernfolge zur Identifikation der Abgeordneten und die Attributmenge Vorname, Name, Geburtstag, Geburtsmonat, Geburtsjahr (wir nehmen an, diese Angaben sind für alle Personen der Parlamentarierdatenbank bekannt und auch eindeutig)

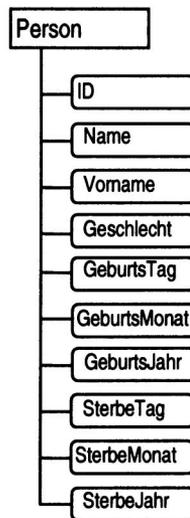


Abb. 6.3: Personen mit ihren Attributen

Relation:

PERSON (ID, Name, Vorname, Geschlecht, GeburtsTag, GeburtsMonat, GeburtsJahr, SterbeTag, SterbeMonat, SterbeJahr)

oder als Tabelle:

Person									
ID	Name	Vorname	Sex	Geburts-			Sterbe-		
				Tag	Mon	Jahr	Tag	Mon	Jahr
010010	Abelmann	Karl	m	9	12	1877	6	1	1928
010020	Ackermann	Friedrich	m	25	5	1876	8	10	1949
010030	Adams	Kurt	m	15	12	1889	7	10	1944
010040	Adelung	Bernhard	m	30	11	1876	24	2	1943
010050	Aderhold	Karl	m	30	6	1884	20	6	1921
010060	Agnes	Lore	w	4	6	1876	9	6	1953

6.2 Beziehungstypen

Jeder Beziehungstyp ergibt ein Relationenschema mit allen Attributen des Beziehungstyps und zusätzlich allen Primärschlüsseln der an der Beziehung beteiligten Entity-Typen. Der Primärschlüssel dieser Relation wird wie folgt gewählt:

Kardinalität	Wahl des Schlüssels
m:n-Beziehung	Beide Primärschlüssel der Entity-Typen werden gemeinsam Primärschlüssel im neuen Relationenschema.
1:n-Beziehung	Der Primärschlüssel der n-Seite wird Primärschlüssel im neuen Relationenschema.
1:1-Beziehung	Beide Primärschlüssel werden je ein Schlüssel im neuen Relationenschema, der Primärschlüssel wird aus diesen beiden ausgewählt.

Tabelle 6.2: Primärschlüssel der Tabellen von Beziehungstypen

Sind die Beziehungen *optional*, so ist das relationale Schema fertig. Bei *zwingenden* Beziehungen können Relationen verschmolzen werden.

6.3 Zusammenfassung von Tabellen bei zwingenden Beziehungen

Kardinalität	Mögliche Vereinfachung
1:n-Beziehung	Das Relationenschema der n-Seite kann in das Relationenschema der Beziehung integriert werden.
1:1-Beziehung	Beide Relationenschemata können in das Relationenschema der Beziehung integriert werden.

Tabelle 6.3: Vereinfachungen bei zwingenden Beziehungen

Da in den Parlamentarierdaten keine geeignete 1:1-Beziehung enthalten ist, wollen wir ein einfaches Beispiel konstruieren. Nehmen wir an, jedem Abgeordneten steht genau ein Schreibtisch zur Verfügung.

Beispiel: 1:1-Beziehung

Zuordnung von Abgeordneten und Schreibtischen als Beispiel für eine 1:1-Beziehung

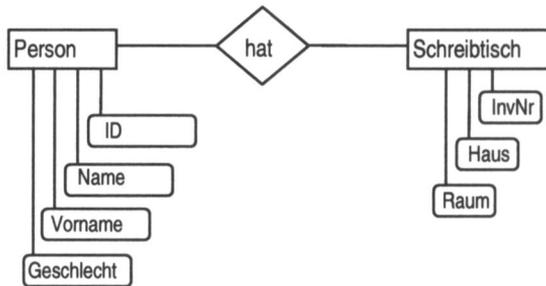


Abb. 6.4: 1:1-Beziehung

Nach der ursprünglichen Vorschrift werden 3 Relationen erzeugt:

- PERSON mit Attributen ID, Name, Vorname und sex,
- Schreibtische mit den Attributen InvNR, Haus und Raum,
- Hat_Schreibtisch mit den Primärschlüsseln der beteiligten Entity-Typen, also ID und InvNR.

Relationen:

- Person (ID, Name, Vorname, Sex)
- Schreibtisch (InvNR, Haus, Raum)
- hat_Schreibtisch (ID, InvNR)

Relationen als Tabellen dargestellt:

Person

ID	Name	Vorname	Sex
10010	Abelmann	Karl	m
10020	Ackermann	Friedrich	m
10030	Adams	Kurt	M

Schreibtisch

InvNR	Haus	Raum
2001	7	20
2002	7	21
2003	8	7

hat Schreibtisch

ID	InvNR
10010	2001
10020	2002
10030	2003

Beispiel: *zwingende 1:1-Beziehung*

Wenn jede Person einen Schreibtisch hat und jeder Schreibtisch durch eine Person besetzt ist, dann liegt eine 1:1-Beziehung vor (zwingend) und es kann zusammengefasst werden.

Person

ID	InvNR	Name	Vorname	Sex	Haus	Raum
10010	2001	Abelmann	Karl	m	7	20
10020	2002	Ackermann	Friedrich	m	7	21
10030	2003	Adams	Kurt	m	8	7

Beispiel: *1:1-Beziehung, eine Seite obligatorisch*

Wenn mehr Schreibtische als Personen vorhanden sind und diese unbesetzt bleiben, dann liegen für bestimmte Spalten keine Werte vor. Es müssen so genannte Nullwerte, in der Tabelle durch '⊥' symbolisiert, im Wertebereich definiert sein.

Person

ID	InvNR	Name	Vorname	Sex	Haus	Raum
10010	2001	Abelmann	Karl	m	7	20
10020	2002	Ackermann	Friedrich	m	7	21
10030	2003	Adams	Kurt	m	8	7
⊥	2004	⊥	⊥	⊥	7	20

In diesem Fall sind zwei Relationenschemata sinnvoll:

Person (ID, Name, Vorname, sex, InvNR)

Schreibtisch (InvNR, Haus, Raum)

oder als Tabellen:

Person

ID	Name	Vorname	Sex	InvNR
10010	Abelmann	Karl	m	2001
10020	Ackermann	Friedrich	m	2002
10030	Adams	Kurt	m	2003

Schreibtisch

InvNR	Haus	Raum
2001	7	20
2002	7	21
2003	8	7
2004	7	20

Beispiel: *1:1-Beziehung, beide Seiten obligatorisch*

Falls nicht jede Person einen Schreibtisch haben muss, dann sind drei Tabellen sinnvoll.

Relationen als Tabellen dargestellt:

Person

ID	Name	Vorname	Sex
10010	Abelmann	Karl	m
10020	Ackermann	Friedrich	m
10030	Adams	Kurt	m
10040	Adelung	Bernhard	m
10060	Agnes	Lore	w

Schreibtisch

InvNr	Haus	Raum
2001	7	20
2002	7	21
2003	8	7
2004	7	20

Hat_Schreibtisch

ID	InvNr
10010	2001
10060	2003

Beispiel: *n:m-Beziehung*

Personen kandidieren für Reichstage (n:m-Beziehung)

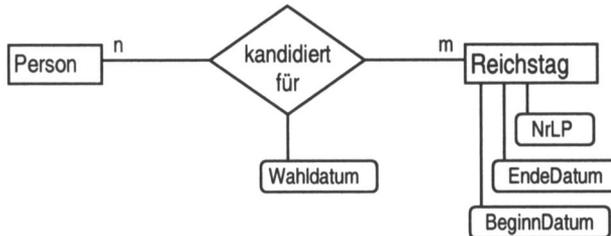


Abb. 6.5: n:m-Abbildung Person und Reichstag

Nach Abbildungsvorschrift werden drei Relationenschemata angelegt: Person mit Attributen ID, Geburts- und Sterbedaten, Reichstag mit den Attributen NrLP, Beginn und Ende der Legislaturperiode, Kandidiert_Für mit den Primärschlüsseln der beteiligten Entity-Typen, {ID, NrLP} und Wahldatum.

Person (ID, Name, Vorname,...)
 Reichstag (NrLP, BeginnDatum, EndeDatum)
 Kandidiert Für (ID, NrLP, Wahldatum)

Beide Attribute ID, NrLP sind in kandidiert Für Fremdschlüssel.

Person

ID	Name	Vorname	sex	Geburts-			Sterbe-		
				Tag	Mon	Jahr	Tag	Mon	Jahr
10020	Ackermann	Friedrich	m	25	5	1876	8	10	1949
10030	Adams	Kurt	m	15	12	1889	7	10	1944
10040	Adelung	Bernhard	m	30	11	1876	24	2	1943
10070	Agster	Alfred	m	12	4	1858	10	1	1904
10100	Albrecht	Adolf	m	14	7	1855	18	4	1930

Reichstag

NrLP	Beginn-		Ende-	
	Monat	Jahr	Monat	Jahr
10	2	1890	6	1893
11	6	1893	6	1898
13	6	1903	1	1907
15	1	1912	11	1918
16	1	1919	6	1920
18	5	1924	12	1924
19	12	1924	5	1928
20	5	1928	9	1930

Kandidiert Für

ID	NrLP	Monat	Jahr
10020	16		
10030	18		
10030	19		
10030	20		
10030	21		
10040	15		
10070	10		
10070	11		
10100	13		

Beispiel:

1:n-Beziehung zwischen Personen und Orten durch Geburt

Eine Person wird in genau einem Ort geboren. In diesem Ort können aber mehrere Personen geboren worden sein.

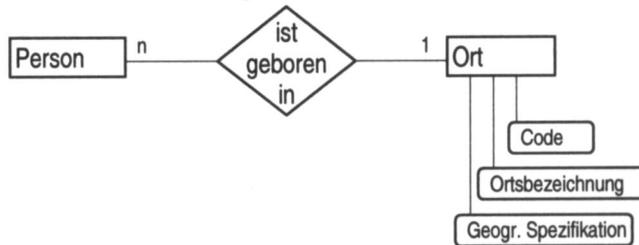


Abb. 6.6: 1:n-Beziehung zwischen Person und Ort

Vorerst sind drei Relationenschemata anzulegen:

- Ort mit den Attributen Ortsschlüssel und geographische Spezifikation,
- Person mit Attributen ID und Name, Vorname, sex,
- ist geboren in mit Primärschlüssel der n-Seite Ort als Primärschlüssel

Jede Person ist in genau einem Ort geboren worden (zwingende Beziehung); daher kann das Relationenschema Person und das Relationenschema ist geboren in verschmolzen werden.

Person (ID, Name, Vorname, sex)

Ort (Ortsschlüssel, geographische Spezifikation)

ist geboren in (ID, Ortsschlüssel)

Person

ID	Name	Vorname	sex
10010	Abelmann	Karl	m
10020	Ackermann	Friedrich	m
10030	Adams	Kurt	m
10040	Adelung	Bernhard	m
10113	Amelung	Friedrich	m
10117	Amhoff	Paul	m
10200	Appelbaum	Karl	m
10220	Arendsee	Martha	w

Ort

Code	Ortsbezeichnung	Geographische Spezifikation
1126	Hamburg	Hamburg
2180	Bremen	Bremen
5077	Edenkoben	BA Landau / Pfalz
54818	Berlin	Berlin
57419	Vehlen	Schaumburg-Lippe
99999	Keine Angabe	-

Ist geboren in

ID	Code
10010	57419
10020	5077
10030	1126
10040	2180
10113	2180
10117	99999
10200	1126
10220	54818

Verschmelzen von Person und 'ist geboren in' führt zu folgenden zwei Relationen:

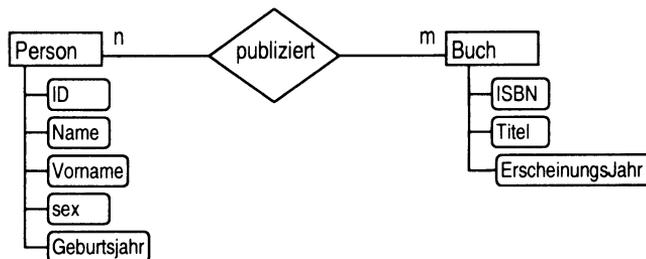
Person

ID	Name	Vorname	sex	Code
10010	Abelmann	Karl	m	57419
10020	Ackermann	Friedrich	m	5077
10030	Adams	Kurt	m	1126
10040	Adelung	Bernhard	m	2180
10113	Amelung	Friedrich	m	2180
10117	Amhoff	Paul	m	999999
10200	Appelbaum	Karl	m	1126
10220	Arendsee	Martha	w	54818

Ort

Code	Ortsbezeichnung	Geograf. Spezifik.
1126	Hamburg	Hamburg
2180	Bremen	Bremen
5077	Edenkoben	BA Landau/ Pfalz
54818	Berlin	Berlin
57419	Vehlen	Schaumburg-Lippe
99999	Keine Angabe	-

Aufgabe 6.1: Das folgende ER-Diagramm stellt eine Lösung der Aufgabe 5.1 dar:



Überführen Sie das Diagramm in ein Relationenschema.

7. Normalformen

Normalisieren verfolgt den Zweck, eine gegebene Tabellenstruktur so zu verändern, dass sie in einem gewissen Sinne günstigere Eigenschaften erhält.

Unter günstigen Eigenschaften versteht man

- redundanzfreie Informationsdarstellung (bis auf Schlüssel), z.B. ist die Adresse einer Person nur einmal im Datenbestand gespeichert.
- Beseitigung von Abfrage- und Änderungsanomalien, z.B. ist es sehr schwierig bei mehrfach gespeicherten Informationen Änderungen so durchzuführen, dass sie überall erfolgen.
- Zugriff zu atomaren Informationseinheiten.

Da man von einer Tabelle ausgeht, ist bereits eine gewisse Struktur der Informationsdarstellung gegeben. Sie kann das Ergebnis einer vorangegangenen Analyse des Forschungsgegenstandes oder eine in den Quellen vorgefundene Informationsstruktur sein. Diese Informationsstruktur kann im Sinne der Relationentheorie ungünstige Eigenschaften besitzen oder gegen relationale Prinzipien verstoßen. Schrittweise, von Normalform zu Normalform, lassen sich diese Mängel beseitigen.

Mit der Normalisierung lernen wir eine zweite Datenbankentwurfsmethode kennen. Normalisieren ist relativ einfach. Daher wird dieses Herangehen besonders von Datenbank-Anfängern bevorzugt. Wir wollen die Unterschiede zwischen beiden aufzeigen.

Das ERM geht von einer so genannten 'Mini-Welt', einer durch die Aufgabenstellung und unser Vorwissen geprägten Sicht auf die Realwelt aus. Man versucht dort Typen mit einer gewissen Eigenständigkeit zu erkennen, die durch Zergliederung und Generalisierung ein immer genaueres Bild der realen Welt liefern. Es wird mehr der Weg von einer Gesamtsicht zur detaillierten Betrachtung beschritten. Das entspricht der von der Softwareentwicklung her bekannten 'Top-Down'-Strategie: Bildlich gesprochen erhebt man sich wie ein Vogel weit über den Untersuchungsgegenstand, so dass Details verschwinden und nur noch grobe Konturen sichtbar sind. Von diesen ausgehend erfolgt eine schrittweise Zergliederung. Das ERM dient nicht nur der Begründung eines geeigneten Relationenschemas, sondern kann ganz allgemein für die Analyse eines zu untersuchenden Sachverhalts herangezogen werden.

Normalisieren geht nicht von einer ganzheitlichen Sicht auf die Untersuchungsobjekte aus, sondern von einer gegebenen Zergliederung. Es wird nicht hinterfragt, woher diese Gliederung kommt, sondern sie ist a priori vorhanden. In bestimmten Fällen, im Allgemeinen nur bei sehr komplexen Problemstellungen, kann das zu ungeeigneten Datenstrukturen führen.

Normalisieren wird der so genannten 'Bottom-Up'-Strategie zugerechnet, da sie von einer Feingliederung ausgeht.

Es ist empfehlenswert, zuerst eine ER-Modellierung vorzunehmen. Die Normalisierung sollte anschließend nur zur Kontrolle der Tabellenstruktur erfolgen. Bei korrekter Ausführung der ER-Modellierung sind die Tabellen aber bereits in der dritten Normalform.

Es werden an dieser Stelle nur die ersten drei Normalformen, die auch die wichtigsten sind, behandelt.

Beispiel: *Parlamentstätigkeit und Aktivität in der Partei*

Es liegt eine Tabelle vor, die für einige Parlamentsausschüsse die SPD-Parlamentarier, die in diesen Ausschüssen mitwirken, ihre dort bekleideten Funktionen und die letztmalige Teilnahme dieser Parlamentarier an einem Parteitag, enthält. Die Funktionen und die Funktionsinhaber wurden je Ausschuss in ein Feld eingetragen. Damit enthalten diese Felder mehrere Informationseinheiten. Im relationalen Schema darf aber in jedem Feld nur eine elementare Information stehen.

Der Ursprung dieser Tabelle und ihre Gliederung soll nicht Gegenstand der Diskussion sein. Nehmen wir an, ein Mitarbeiter hat Informationen im Archiv gefunden und der Struktur der Quelle folgend, sie in dieser Form zusammengestellt:

PARLAMENTSAUSSCHÜSSE

AUSSCHUSS	FUNKTION	NAME	PT-DATUM	PT-ORT
Haushalt	Vorsitzender, Protokollant	Heinze, Abel	1919, 1921	Weimar, Görlitz
Inneres	Stellv. Vors., Sekretär, Sekretär	Belz, Abel, Heinze	1922A, 1921, 1919	Augsburg, Görlitz, Weimar
Sozialfürsorge	Vorsitzender, Sekretär	Greiner, Heinze	1922A, 1919	Augsburg, Weimar

7.1 1. Normalform

Eine Relation ist in der ersten Normalform, wenn alle ihre Felder elementar (oder auch atomar) sind.

PARLAMENTSAUSSCHÜSSE

A#	AUSSCHUSS
1	Haushalt
2	Inneres
3	Sozialfürsorge

FUNKTION

A#	FUNKTION	NAME	PT-DATUM	PT-ORT
1	Vorsitzender	Heinze	1919	Weimar
1	Protokollant	Abel	1921	Görlitz
2	Stellv. Vors.	Belz	1922A	Augsburg
2	Sekretär	Abel	1921	Görlitz
2	Sekretär	Heinze	1919	Weimar
3	Vorsitzender	Greiner	1922A	Augsburg
3	Sekretär	Heinze	1919	Weimar

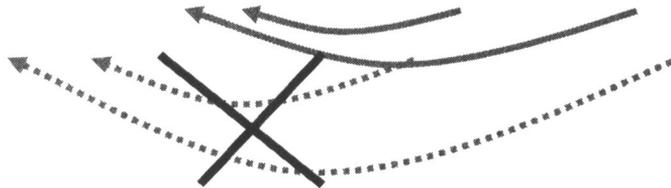
7.2 2. Normalform

Eine Relation befindet sich in der 2. Normalform, wenn sie in der 1. Normalform ist und jedes Nichtschlüssel­feld voll funktional abhängig ist von allen Schlüssel­feldern.

In der Tabelle FUNKTION sind die Felder PT-DATUM und PT-ORT vom Mandats­träger (NAME), und nicht von dessen Funktion abhängig. Daher muss diese Tabelle in zwei Tabellen aufgespalten werden.

FUNKTION

A#	FUNKTION	NAME	PT-DATUM	PT-ORT
----	----------	------	----------	--------



FUNKTION

A#	FUNKTION	P#
1	Vorsitzender	12
1	Protokollant	14
2	Stellv. Vors.	20
2	Sekretär	14
2	Sekretär	12
3	Vorsitzender	30
3	Sekretär	12

MANDATSTRÄGER

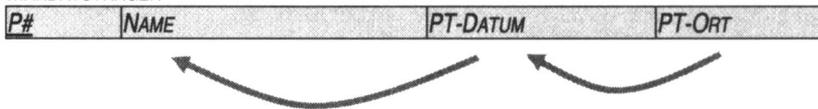
P#	NAME	PT-DATUM	PT-ORT
12	Heinze	1919	Weimar
14	Abel	1921	Görlitz
20	Belz	1922A	Augsburg
30	Greiner	1922A	Augsburg

7.3 3. Normalform

Eine Relation ist in der dritten Normalform, wenn sie sich in der 2. Normalform befindet und jedes Nichtschlüssel­feld nicht transitiv abhängig ist von jedem Schlüssel­feld.

Das Parteitage­datum (Jahr und Monatshinweis, 'A' steht für „August“) identifizieren die Parteitage. Daher ist PT-ORT über PT-DATUM von NAME abhängig. Das ist eine Transitivität.

MANDATSTRÄGER



PARLAMENTS-AUSSCHÜSSE

A#	AUSSCHUSS
1	Haushalt
2	Inneres
3	Sozialfürsorge

FUNKTION

A#	FUNKTION	P#
1	Vorsitzender	12
1	Protokollant	14
2	Stellv. Vors.	20
2	Sekretär	14
2	Sekretär	12
3	Vorsitzender	30
3	Sekretär	12

MANDATSTRÄGER

P#	NAME	PT-DATUM
12	Heinze	1919
14	Abel	1921
20	Belz	1922A
30	Greiner	1922A

PARTEITAGE

PT-DATUM	PT-ORT
1919	Weimar
1921	Görlitz
1922A	Augsburg

8. Einführung in ACCESS

8.1 Einleitung

ACCESS ist

- ein relationales DatenBankManagementSystem (DBMS)
- Bestandteil von Microsoft Office
- sehr benutzerfreundlich

ACCESS ist nicht

- objektorientiert, auch wenn im Folgenden von 'Objekten' gesprochen wird
- für sehr umfangreiche Datenbestände geeignet
- das Nonplusultra der Datensicherheit

Was muss ein Datenbankmanagementsystem können?

- Datenbanken erstellen
- Informationen der Datenbank hinzufügen
- Daten beliebig bearbeiten und verändern
- Daten sortieren und nach bestimmten Kriterien filtern
- Mit Hilfe von Abfragen Daten auswerten
- Daten mit Hilfe von Berichten übersichtlich ausdrucken
- Diagramme zur Darstellung der Daten einsetzen

Was bedeutet 'relational'?

1. Tabellenartige Struktur der Datenspeicherung. Eine Datenbank kann mehrere Tabellen umfassen.
2. Jede Tabelle besteht aus Datenfeldern, die in Zeilen und Spalten angeordnet sind.
3. Jede Tabelle hat einen Namen, der sie von anderen Tabellen eindeutig unterscheidet.
4. In einer Spalte stehen jeweils Daten gleichen Typs. Eine Übersicht der wichtigsten Datentypen bietet die folgende Tabelle:

Datentypen von ACCESS	Dateninhalt
Text	Zeichenfolge bis zu einer Länge von maximal 255 Zeichen
Zahl	numerische Größe, wahlweise mit Kommastellen oder ohne zu definieren
Memo	sehr langer Text, der nicht in der Tabelle, sondern außerhalb gespeichert wird (aber noch in der Datenbank). In der Tabelle wird nur ein Verweis auf den Text gespeichert.
Datum/Zeit	gültige Datumsangaben und/oder Zeitangaben
Ja/Nein-Felder	so genannte Boole'sche Werte, die nur zwei Ausprägungen annehmen können: 'ja' oder 'nein' ('wahr' oder 'falsch', 'true' oder 'false')
Hyperlinks	Internet-Adressen
OLE-Objekte	OLE-Objekte (object linking and embedding) sind binäre Dateien außerhalb der Datenbank, wie z.B. Bilder oder Tonsequenzen. In der Tabelle wird nur ein Verweis auf das Objekt gespeichert.

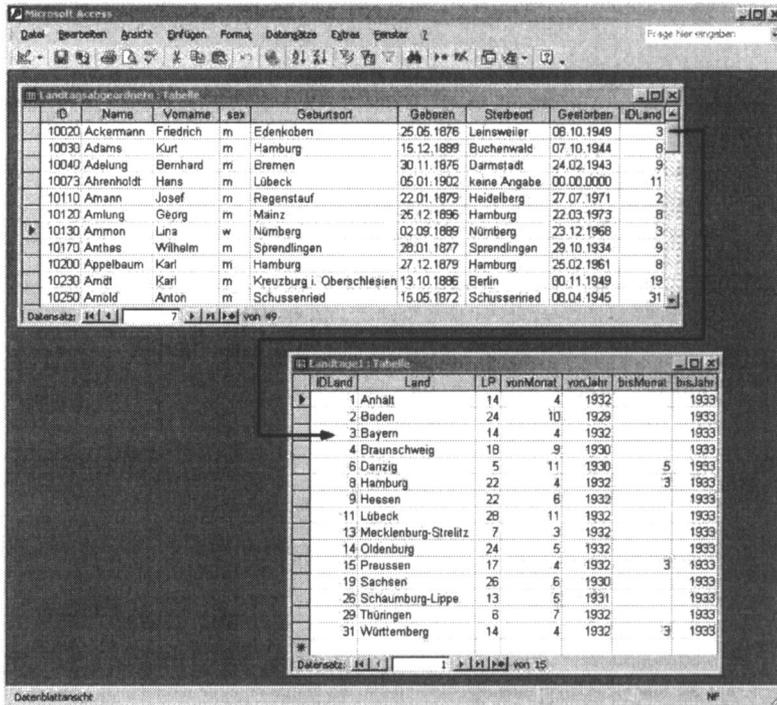
5. Mehrere Tabellen lassen sich durch die Dateninhalte einer oder mehrerer Spalten verknüpfen.

Beispiel: *Tabelle der Abgeordneten und Tabelle der Landtage*

In der Tabelle der Abgeordneten ist für jeden Landtagsabgeordneten eine Zeile vorgesehen. Neben den 'Abgeordneten-Daten' steht nicht der Landtag mit seinen Merkmalen in der Abgeordnetentabelle, sondern nur die Landtagsnummer. Über diese Nummer kann in der Landtagstabelle die entsprechende Zeile für den Landtag gefunden werden.

Man kann sich das bildlich so vorstellen: Ein Bearbeiter hat eine Liste mit Abgeordneten und eine Liste mit Landtagen. Er geht die Liste der Abgeordneten durch, indem er den Zeigefinger der einen Hand benutzt, um über die Zeilen hinweg in einer Spalte abwärts zu gleiten. Dabei zeigt er jeweils auf eine Zeile. Zu bestimmten Abgeordneten notiert er die Landtagsnummer und fährt mit der anderen Hand in der Tabelle Landtage durch die Spalte Landtags-Nummer, bis er den gesuchten Landtag gefunden hat. In analoger Weise arbeitet ACCESS: für jede Tabelle gibt es genau einen Zeiger. Dieser Zeiger kann auf eine Zeile zeigen oder vor der ersten Zeile oder hinter der letzten stehen.

6. Im folgenden Bild zeigt der Zeiger der Tabelle Landtagsabgeordnete auf die Zeile mit der ID 10130, der der Tabelle Landtage auf die erste Zeile (in ACCESS durch das schwarze Dreieck symbolisiert)



7. Auswahl von Spalten: nur Name, Vorname und Abgeordnetenummer (ID) sind relevant

ID	Name	Vorname
10020	Ackermann	Friedrich
10030	Adams	Kurt
10040	Adelung	Bernhard
10073	Ahrenholdt	Hans
10110	Ahmann	Josef
10120	Amlung	Georg
10130	Ammon	Lina
10170	Anthes	Wilhelm

8. Auswahl von Zeilen: alle Abgeordneten mit Geburtsort Hamburg sind von Interesse

ID	Name	Vorname	Geburtsort	sex	Geboren	Sterbeort	Gestorben
1000	Adams	Kurt	Hamburg	m	15.12.1889	Buchenwald	07.10.1944
1004	Appelbaum	Karl	Hamburg	m	27.12.1879	Hamburg	25.02.1961
1007	Ballerstädt	Richard	Hamburg	m	03.03.1873	Hamburg	15.01.1953

9. Ein relationales System unterstützt Sie auch bei der Durchsetzung bestimmter Regeln und logischer Zusammenhänge zwischen den Tabellen (lokale und globale Datenintegrität). Wünschen Sie z.B., dass jeder Landtagsabgeordnete eine eindeutige Abgeordneten-ID haben soll und einem Landtag angehören muss, dann setzt ACCESS diese Forderung selbständig durch.

8.2 Arbeiten mit ACCESS

8.2.1 ACCESS starten und beenden

Der **Start** von ACCESS erfolgt wie bei jedem anderen Windows-Programm auch:

- ◆ Klicken auf die -Schaltfläche
 - ◆ Zeigen auf Menüpunkt PROGRAMME
 - ◆ Klicken auf das Untermenü MICROSOFT ACCESS
- oder
- ◆ Klicken auf das Symbol von ACCESS oder auf eine ACCESS-Datenbank

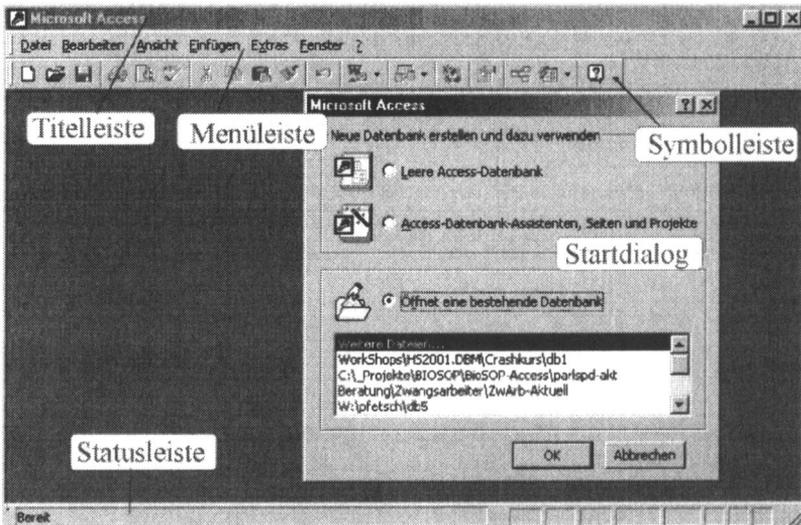
Beenden von ACCESS:

- ◆ Menüpunkt DATEI → BEENDEN
- oder
- ◆ Klicken auf das Schließfeld von ACCESS 
- oder
- ◆ betätigen der Tastenkombination [ALT] [F4]

8.2.2 ACCESS lässt sich nicht beenden

Sollte es vorkommen, dass ACCESS nicht mehr auf Eingaben reagiert, (das Programm hat sich aufgehängt), betätigen Sie die Tastenkombination [STRG] [ALT] [ENTE] und folgen dem Dialog.

8.3 Das Anwendungsfenster



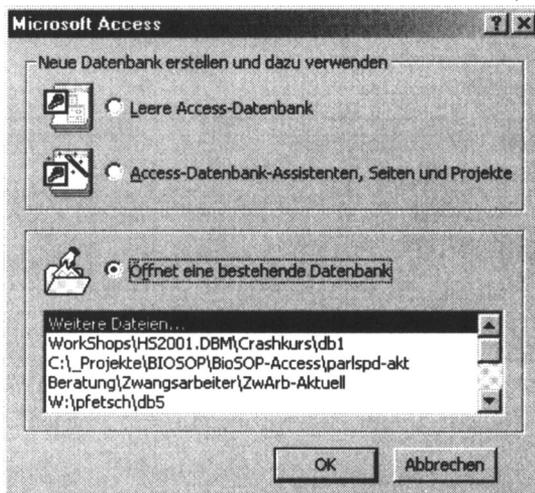
Titelleiste	Die Titelleiste enthält den Programmnamen und die Windows-Standardelemente Minimierfeld, Voll- bzw. Teilbildfeld und Schließfeld.
Minimierfeld 	reduziert das Fenster auf die Schaltfläche in der Windows-Task-Leiste.
Voll-, Teilbildfeld 	Umschalten zwischen Vollbilddarstellung und verkleinerter Fenstergröße
Schließfeld 	schließt Fenster bzw. Programm.
Startdialog	Nach dem Start von ACCESS wird automatisch der Startdialog eingeblendet. Sie können wählen, ob eine neue Datenbank erstellt oder eine vorhandene geöffnet werden soll.

Menüleiste	ACCESS stellt mehrere Menüs zur Verfügung, in denen die verschiedenen Funktionen des Programms abgerufen werden können. Hinter jedem Menü verbirgt sich ein Pull-down-Menü, das bei Anklicken des Menüs herunterklappt und die zugeordneten Menüpunkte (Befehle) enthält.
Statusleiste	Sie erhalten Hinweise zum Programmstatus
Symbolleiste	Die Symbolleisten stellen häufig benutzte Funktionen zur schnellen Mausklickausführung zur Verfügung. Wenn Sie mit der Maus auf ein Symbol zeigen, erscheint kurz danach neben dem Mauszeiger eine Kurzbeschreibung der entsprechenden Funktion, die so genannte Quickinfo.

Die Bedienelemente von ACCESS verhalten sich kontextsensitiv. Wenn eine Funktion zum gegebenen Zeitpunkt nicht sinnvoll anzuwenden ist, wird sie im Menü ausgeblendet. Sie steht erst dann wieder zur Verfügung, wenn ein entsprechender Kontext vorliegt.

8.3.1 Vorhandene Datenbank öffnen

- ◆ Starten Sie ACCESS.
- ◆ Wählen Sie die Option **ÖFFNET EINE BESTEHENDE DATENBANK**.
- ◆ Markieren Sie in der Liste die gewünschte Datenbank oder markieren Sie **WEITERE DATEIEN** und bestätigen Sie mit **OK**.



8.3.2 Weitere Möglichkeiten eine Datenbank zu öffnen

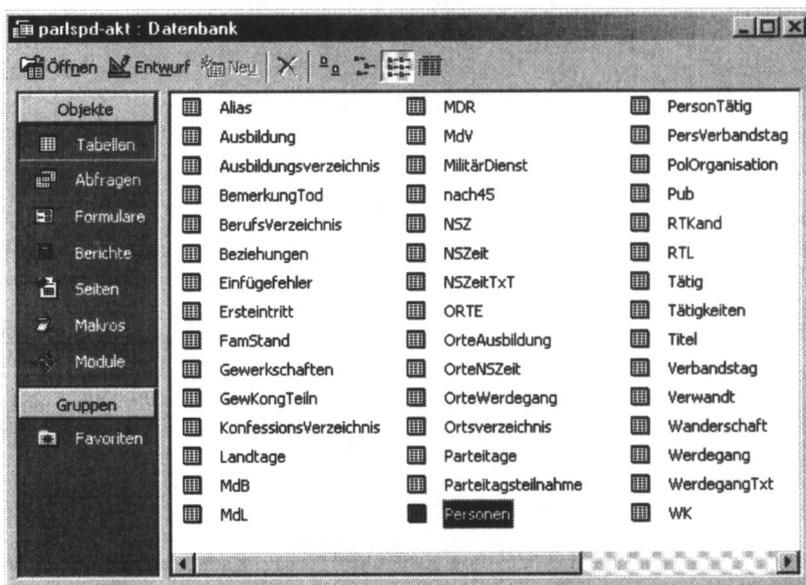
Sie können auch den Startdialog durch **ABBRECHEN** beenden und über den Menüpunkt **DATEI** → **ÖFFNEN** auf eine Datenbank zugreifen. Im Menüpunkt **DATEI** werden die vier zuletzt geöffneten Datenbanken in der unter Windows üblichen Weise angeboten.

Sie können im Explorer oder auf dem Desktop auf das Icon einer Datenbank doppelklicken. **ACCESS** wird mit der gewünschten Datenbank im Zugriff geöffnet.

8.4 Das Datenbankfenster

Eine Datenbank kann aus vielen Objekten bestehen. Jedes Objekt gehört einem gewissen Objekttyp an, z.B. sind die von uns angelegten Datentabellen vom Objekttyp „Tabelle“. Auch die Datenbank selbst ist ein Objekt vom Objekttyp „Datenbank“.

Wenn wir eine Datenbank öffnen, werden die uns am häufigsten begegnenden Objekttypen von **ACCESS** im Datenbankfenster angezeigt.



Am linken Rand werden Ihnen die folgenden Objekttypen angeboten:

Tabellen	Eine Tabelle enthält eine Sammlung von Daten zu einem bestimmten Thema. Jede Zeile einer Tabelle entspricht einem Datensatz; sie repräsentiert die Daten einer „Karteikarte“ (Der Begriff Datensatz wird hier, wie in der Informatik üblich, als Zeile einer Tabelle verstanden).
Abfragen	Mit einer Abfrage können Daten einer Tabelle ausgewertet werden, beispielsweise können alle Personen, die in einem bestimmten Ort geboren wurden, ermittelt werden. Mit Hilfe von Abfragen werden relationale Operationen durchgeführt.
Formulare	Formulare dienen zur Ansicht, Eingabe und Bearbeitung von Daten aus Tabellen oder Abfragen.
Berichte	In Berichten werden Daten aus Tabellen oder Abfragen zusammengefasst und anschaulich dargestellt. Beispielsweise kann eine Liste aller Abgeordneten mit Geburts- und Sterbedaten erstellt und ausgedruckt werden.
Seiten	Seiten (Abkürzung für Datenzugriffsseiten) sind spezielle Webseiten, um Daten der ACCESS-Datenbank im Internet oder Intranet darstellen zu können.
Makros	Mit Hilfe von Makros können wiederkehrende Arbeitsabläufe in Datenbanken automatisiert werden. Diese Arbeit übernimmt in der Regel ein Datenbank-Entwickler.
Module	ACCESS beinhaltet die Programmiersprache Visual Basic for Applications (VBA), mit der ACCESS programmiert werden kann. Diese Arbeit übernimmt in der Regel ebenfalls der Datenbank-Entwickler.

Je nachdem, welchen Typ Sie anklicken, werden die zur Datenbank vorhandenen Objekte aufgelistet. Im Beispiel ist TABELLEN aktiviert. Es werden alle zur Datenbank parlspe-akt gehörenden Tabellen angezeigt.

Im oberen Teil des Datenbankfensters befindet sich eine zusätzliche Symbolleiste, die unter anderem die Schaltflächen ÖFFNEN, ENTWURF und NEU enthält, mit denen Sie Objekte bearbeiten, verändern oder anlegen können.

Jedes Objekt verfügt über gewisse Eigenschaften. Diese werden vom System intern in Systemtabellen verwaltet. Man kann sich die Eigenschaften ansehen und verändern.

Zu den Eigenschaften von Objekten gehören beispielsweise Farbe und Größe, verwendeter Schriftsatz und Schriftgröße, ob das betreffende Objekt geöffnet oder geschlossen ist, ob es geöffnet und auch sichtbar oder unsichtbar ist und vieles mehr.

Über diese Eigenschaften lassen sich wichtige Entscheidungen über das Aussehen, Verhalten und Zusammenspiel von Objekten treffen. Die Tabelle der Eigenschaften zu einem Objekt erreicht man, indem man mit der rechten

Maustaste auf das Objekt klickt und im sich öffnenden Kontextmenü EIGENSCHAFTEN wählt.

Aufgabe 8.1: Starten Sie ACCESS und legen Sie eine leere Datenbank mit dem Namen **Parlamentarier** an.

8.5 Tabellen

8.5.1 Neue Tabelle erstellen

Sie können eine Tabelle auf verschiedene Weise erzeugen. Wir wollen hier die Entwurfsansicht wählen.

- ◆ Aktivieren Sie im Datenbankfenster das Objekt TABELLEN und klicken sie doppelt auf den ersten Eintrag ERSTELLT EINE TABELLE IN DER ENTWURFSANSICHT.
- oder Klicken Sie im Datenbankfenster auf das Symbol  Neu und wählen Sie den Eintrag ENTWURFSANSICHT.
- ◆ Bestätigen Sie mit Ok.

Eine neue leere Tabelle wird in der Entwurfsansicht geöffnet. Wir werden den folgenden Tabellenentwurf mit den angegebenen Feldern umsetzen:

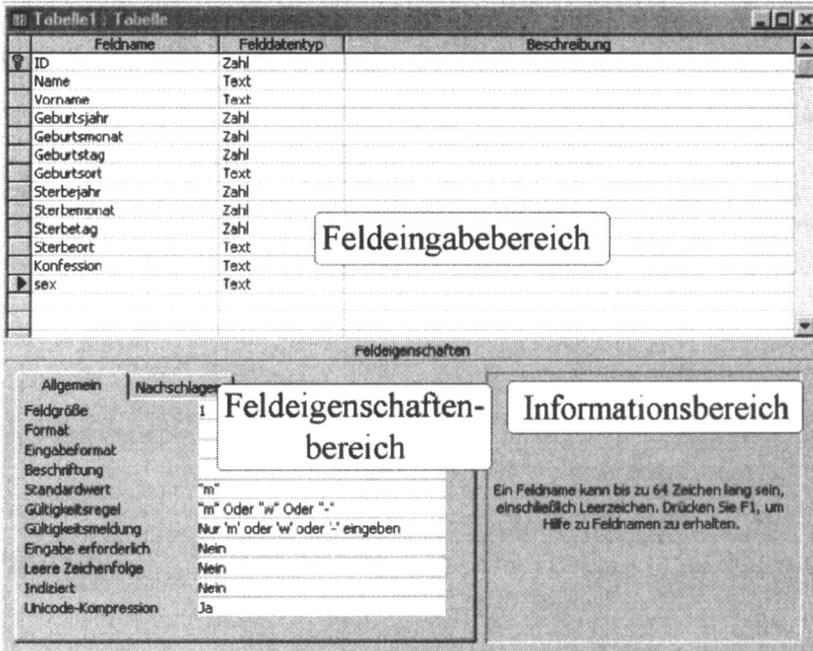
Feldname	Felddatentyp	Feldgröße	Beschreibung
ID	Numerisch	Long Integer	eine vom Bearbeiter vergebene Nummer, die jeden Abgeordneten identifiziert
Name	Zeichen	50	Familienname des Abgeordneten
Vorname	Zeichen	50	Vorname des Abgeordneten
Geburtsjahr	Numerisch	Integer	Geburtsjahr als Zahl
Geburtsmonat	Numerisch	Integer	Geburtsmonat
Geburtstag	Numerisch	Integer	Geburtstag
Geburtsort	Zeichen	50	Ortsname des Geburtsortes
Sterbejahr	Numerisch	Integer	Sterbejahr
Sterbemonat	Numerisch	Integer	Sterbemonat
Sterbetag	Numerisch	Integer	Sterbetag
Sterbeort	Zeichen	50	Ortsname des Sterbeortes
Konfession	Zeichen	50	Konfession als Text; keine Kodierung

Sex	Zeichen	1	es sollen nur die Zeichen 'm' für männlich, 'w' für weiblich und '-' für nicht bekannt eingegeben werden können
Vaterberuf	Zeichen	50	der Beruf des Vaters des Abgeordneten

ACCESS kennt die folgenden Felddatentypen

Felddatentyp	Bedeutung	Speicherbedarf
Zeichen	Zeichenfolge bis zu einer maximalen Länge von 255 Zeichen	max. 255 Byte
Memo	Texte bis maximal etwa 65.000 Zeichen	max. 65.535 Byte
Numerisch	Zahlen der Form:	
Byte	ganze Zahl zwischen Null und 255	1 Byte
Integer	ganze Zahlen zwischen -32.768 und +32.767	2 Byte
Long Integer	ganze Zahlen zwischen -2.147.483.648 und +2.147.483.647	4 Byte
Double	Kommazahlen mit 15-stelliger Genauigkeit	4 Byte
Single	Kommazahlen mit 7-stelliger Genauigkeit	8 Byte
Replikations-ID	dient zur eindeutigen Identifizierung von Objekten in verschiedenen Datenbanken.	16 Byte
Datum/Uhrzeit	Datum und Uhrzeit	8 Byte
Währung	Zahlen im Währungsformat	8 Byte
AutoWert	Zahlen, die ACCESS automatisch verwaltet; es ist nur ein AutoWert-Feld je Tabelle erlaubt.	4 Byte
Ja/Nein	Logische Werte (Boole'sche Werte)	1 bit
OLE-Objekte	Objekte bis maximal 1 Gigabyte. OLE-Objekte können unter anderem Bilder, Musiksequenzen oder Texte sein.	max. 1 Gigabyte

Hyperlink	Texte, die ACCESS als Hyperlink-Adressen verwendet. Eine Hyperlink-Adresse besteht maximal aus Anzeigetext, Adresse und Unteradresse.	max. 2.048 Byte je Bestandteil (Anzeigetext, Adresse, Sub- adresse)
-----------	---	---



8.5.2 Felder in die Tabelle einfügen

- ◆ Klicken Sie in das erste Feld der Spalte Feldnamen und geben sie den ersten Namen ein: *ID*
- ◆ Betätigen Sie die Tabulatortaste, um zur nächsten Spalte zu wechseln.
- ◆ Legen Sie mit Hilfe des Listenfeldes den Datentyp fest oder geben Sie den ersten Buchstaben des Datentyps ein: 'n' für numerisch.
- ◆ Geben Sie in der Spalte Beschreibung eine kurze Erläuterung ein (nicht zwingend): „*Identifikations-Nr. für Abgeordnete*“.

- ◆ Geben Sie weitere Felder ein.
- ◆ Im unteren Teil können Sie zusätzliche Feldeigenschaften angeben. Geben Sie für das Feld 'sex' eine Gültigkeitsregel an:
Die Spalte soll nur die Werte
 'm' für männlich,
 'w' für weiblich,
 '-' für nicht in den Quellen zu ermitteln
enthalten dürfen.
Da die meisten Abgeordneten männlichen Geschlechts sind, erleichtert der Standardwert 'm' die spätere Eingabe.
Geben Sie auch einen Text für die Gültigkeitsmeldung ein.

ACCESS zeigt die Feldeigenschaften jeweils nur für das aktuelle Datenfeld an. Das aktuelle Datenfeld erkennen Sie an dem nach rechts zeigenden Pfeil vor der ersten Spalte.

Die Auswahlmöglichkeiten der Feldeigenschaften sind von dem jeweiligen Feldtyp abhängig. Der Informationsbereich gibt Ihnen eine kurze Beschreibung der aktivierten Feldeigenschaft.

8.5.3 Übersicht über die wichtigsten Feldeigenschaften

Eigenschaft	Beschreibung
Feldgröße	Anzahl der Zeichen, die maximal in das Feld eingegeben werden können
Format	Späteres Anzeigeformat für den Feldwert
Eingabeformat	Definiert ein Muster für alle Daten, die später in das Feld eingegeben werden
Dezimalstellenanzeige	Festlegung der Anzahl Dezimalstellen nach dem Komma bei Zahlen- und Währungsfeldern
Beschriftung	Feldbezeichnung in Tabellen, Berichten und Formulare; kann abweichend von dem im oberen Bereich angegebenen Feldnamen sein
Standardwert	Mit welchem Inhalt soll das Feld bei neuen Datensätzen vorbelegt sein?

Gültigkeitsregel	Regel zur Überprüfung der korrekten Eingabe
Gültigkeitsmeldung	Fehlermeldung bei falscher Eingabe aufgrund der Gültigkeitsregel
Eingabe erforderlich	Bestimmt, ob beim späteren Ausfüllen eines Datensatzes eine Eingabe in das Datenfeld erforderlich ist
Leere Zeichenfolge	Hier wird bestimmt, ob eine leere Zeichenfolge (doppelte Anführungszeichen hintereinander "") als gültiger Eintrag akzeptiert wird

8.5.4 Primärschlüssel festlegen

Ein Primärschlüssel bestimmt eindeutig jede Zeile einer Tabelle. Im vorangegangenen Beispiel ist das Merkmal ID ein eindeutiger Identifikator für jeden Parlamentarier. Die Tabelle enthält für jeden Parlamentarier genau eine Zeile. Damit kann ID als Primärschlüssel für diese Tabelle ausgewählt werden. Ein Primärschlüssel kann auch aus mehreren Datenfeldern bestehen, z.B. aus Name, Vorname, Geburtsdatum.

Es ist nicht zwingend notwendig, einen Primärschlüssel anzugeben. Es ist jedoch sehr empfehlenswert, da über Schlüssel die Verbindung zu anderen Tabellen hergestellt wird. Die Festlegung eines Primärschlüssels erhöht wesentlich die Verarbeitungsgeschwindigkeit.

Wenn Sie die Vergabe der Primärschlüssel und deren Verwaltung und Kontrolle ACCESS übertragen möchten, wählen Sie als Felddatentyp AutoWert. AutoWert erzeugt Zahlen vom Typ Long Integer.

8.5.4.1 Selbst den Primärschlüssel festlegen

- ◆ Klicken Sie die Zeile an, die Primärschlüssel werden soll, im Beispiel ID.
- ◆ Klicken Sie auf das Schlüsselsymbol in der Symbolleiste 

8.5.5 Tabelle schließen und speichern

- ◆ Schließen Sie die Tabelle durch Klick auf das Schließfeld
oder DATEI – SCHLIESSEN
oder [STRG] [W]

- ◆ Bestätigen Sie die Rückfrage, ob Sie den Tabellenentwurf speichern möchten mit JA.
- ◆ Geben Sie einen Namen für die Tabelle ein und bestätigen Sie mit OK.
- ◆ Falls noch kein Primärschlüssel definiert ist, bestätigen Sie die Rückfrage mit JA.

Sie können auch nachträglich Tabellen ändern. Dazu aktivieren Sie das Objekt TABELLEN, klicken auf ENTWURF und fügen neue Felder an oder verändern die vorhandenen Einstellungen. Auch die Reihenfolge der Zeilen in der Entwurfsansicht lässt sich per Drag & Drop verändern.

Aufgabe 8.2: Erstellen Sie eine Tabelle Personen mit den in 8.5.1 beschriebenen Datenfeldern.

Aufgabe 8.3: Geben Sie die folgenden Daten (s. nächste Seite) in die Tabelle Personen ein.

Aufgabe 8.4: Definieren Sie einen geeigneten Primärschlüssel für die Tabelle Personen.

	Beruf des Vaters	Bergmann	Berarbeiter	Keine Angabe	Schneidermeister	Maurer	Schiffseigentümer	Schriftsetzer	Weber	Keine Angabe	Keine Angabe	Metzgermeister
	Konfession	ev., später diss	diss.	diss.	ev., später diss	kath., später diss	ev., später diss	diss.	ev., später diss	diss.	kath.	kath.
	Sterbejahr	1921	1953	1904	1930	1955	1924	1953	1957	1953	1945	1907
	Sterbemonat	6	6	1	4	7	7	5	9	5	3	4
	Sterbetag	20	9	10	18	11	7	22	11	22	20	10
	Sterbeort	Hannover	Köln	Degerloch	Halle (Saale)	Dorsten	Braunschweig	Berlin	Magdeburg	Wiesbaden	Giegen	Berlin
	Geburtsjahr	1884	1876	1888	1855	1880	1888	1885	1887	1880	1874	1846
	Geburtsmonat	6	6	4	7	12	11	3	4	10	12	4
	Geburtstag	30	4	12	14	15	24	29	19	9	22	19
	Geburtsort	Slobberg (Harz)	Bochum	Ilsefeld	Reudnitz	Löchtau	Larabsberg/W.	Berlin	Bramsche	Reichertsdorf/Vogtl.	Dornmelstadt	Dornmelstadt
	Sex	m	w	m	m	w	m	w	w	m	m	m
	Vorname	Karl	Lore	Alfred	Adolf	Marie	Otto	Martha	Marie	Arthur	Erhard	Ignaz
	Name	Aderhold	Agnes	Agster	Albrecht	Ansohn	Annick	Avensee	Anning	Arzt	Auer	Auer
	ID	10050	10060	10070	10100	10160	10180	10220	10240	10280	10310	10320

8.6 Formulare

8.6.1 Formular öffnen

Ein Formular wird verwendet, um Datensätze anzuzeigen oder neue Datensätze einzugeben. Am einfachsten öffnet man ein Formular, indem man

- ◆ im Datenbankfenster FORMULARE aktiviert,
- ◆ doppelt auf das gewünschte Formular klickt
oder
einmal auf das Objekt und dann auf ÖFFNEN klickt.

Das Formular wird in einem neuen Fenster angezeigt. Das abgebildete Formular zeigt die Daten eines Datensatzes, also eine Zeile der Tabelle an. Hier entspricht jeweils ein Datenfeld einem Wert einer Spalte der Tabelle.

The screenshot shows a window titled 'Personen' with a form for 'Sozialdemokratische Parlamentarier in deutschen Parlamenten von 1867 bis 1933'. The form is titled 'Geburtsdaten:' and contains the following fields:

ID	10010
Name	Abelmann
Vorname	Karl
Geschlecht	m
Geburtstag	9 . 12 . 1877
Sterbetag	6 . 1 . 1928

At the bottom, there is a navigation bar for the dataset: 'Datensatz: 1 von 2424'.

Es werden Name und Vorname, das Geburts- und Sterbedatum sowie die Identifikationsnummer der sozialdemokratischen Parlamentarier angezeigt.

Mit der Datensatznavigationsleiste am unteren Rand des Formulars können Sie den Datenbestand durchblättern. Am oberen Rand rechts befinden sich die bekannten Schaltknöpfe zum Schließen und zum Einstellen von Minimal- oder Maximalgröße. Das Formular kann verschoben werden, durch Ziehen des Ran-

des vergrößert oder verkleinert werden. Durch Anklicken eines anderen Fensters, z.B. des Datenbankfensters, steht es nicht mehr im Zugriff, es verliert den Fokus.

8.6.2 Wozu dienen Formulare?

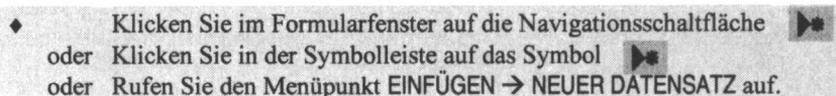
Sie können eine Tabelle öffnen und die darin enthaltenen Daten und Datensätze anzeigen lassen. Es ist jedoch viel komfortabler, Datensätze in Formularen anzusehen und zu bearbeiten. In einem Formular können die gesamten Informationen eines Datensatzes oder auch nur ein Teil davon dargestellt werden. Formulare können auch Informationen aus verschiedenen Tabellen anzeigen, vorausgesetzt, dass bestimmte Formularfelder mit diesen Tabellen verknüpft sind.

8.6.3 Vorteile von Formularen

Vorteile von Formularen sind:

- Einfache, unkomplizierte Dateneingabe,
- Vermeiden von fehlerhaften Eingaben,
- Benutzung von Steuerelementen zur Datensaatzauswahl.

Im Formular Personen enthält jeder Datensatz Informationen zu einzelnen Parlamentariern. Um neue Datensätze aufzunehmen, gehen Sie folgendermaßen vor:



Es erscheint ein leeres Formular. In dieses Formular können Sie die Daten des neuen Datensatzes eingeben. Neue Datensätze werden immer an das Ende der dem Formular zugrunde liegenden Tabelle geschrieben.

8.6.4 Cursorsteuerung in einem Formular mit der Tastatur

Taste	Wirkung
	springt zum nächsten Feld.
 UMSCHALTEN 	springt zum vorherigen Feld.
RETURN 	schließt die Eingabe in ein Feld ab und springt zum nächsten Feld.
POS1	springt bei einem Textfeld an den Anfang des Feldes.
ENDE	springt bei einem Textfeld an das Ende des Feldes.
F2	markiert den gesamten Inhalt eines Feldes.
ENTF	löscht das Zeichen rechts vom Cursor.
 KORREKTUR	löscht das Zeichen links vom Cursor.
ESC	bricht die Eingabe in ein Feld ab.

8.6.5 Erstellen eines Formulars

Um das Zusammenwirken von Formularen und Tabellen zu verstehen, stellen Sie sich ein Formular als ein Blatt Papier vor, durch das Sie auf die Tabelle schauen. An bestimmten Stellen im Papier sind Löcher hinein geschnitten, durch die Sie die Datenfelder der Tabelle betrachten. Sie können am Papier (Formular) einstellen, dass Sie die Daten der Tabelle in einem bestimmten Format sehen möchten. Solche Formatvorgaben lassen sich auch für die Eingabe festlegen. Wenn Sie zu einem anderen Datensatz gehen, übergibt ACCESS die geänderten / neuen Daten an die Tabelle und versucht sie zu speichern. Bei Datenkonflikten, wenn Sie z.B. versucht haben, Buchstaben in ein numerisches Feld zu speichern, weist ACCESS die Daten zurück.

Sie haben an zwei Stellen die Möglichkeit, die Dateneingabe zu kontrollieren: mit Hilfe des Formulars, indem Sie dort vereinbaren, dass Informationen nur in einer von Ihnen festgelegten Art und Weise entgegengenommen werden sollen, und beim Speichern in der Tabelle, wie bereits am Beispiel des Feldes 'sex' (Abschnitt 8.5.2) gezeigt.

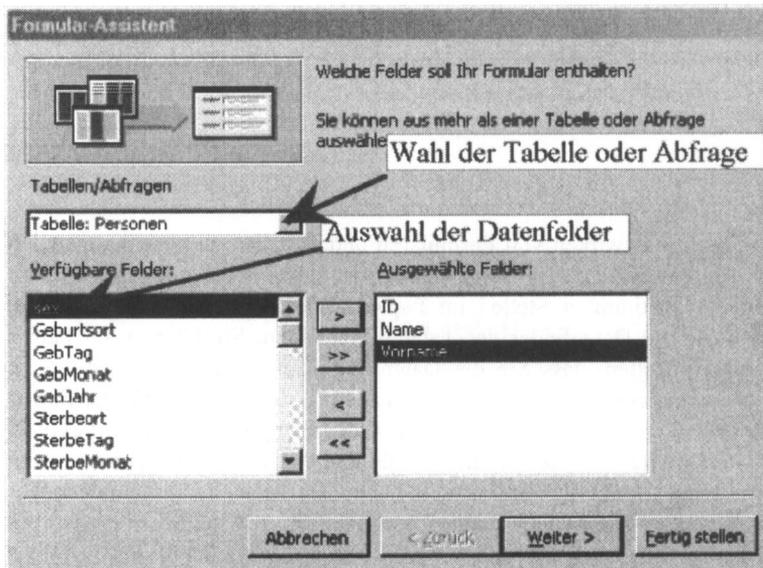
Sie haben zwei Möglichkeiten, ein Formular zu erstellen:

- Erstellen eines leeren Formulars und manuelle Zuweisung der Tabellenfelder zu den Feldern des Formulars.
- Benutzung des Formularassistenten.

Hier soll kurz die Arbeit mit dem Formularassistenten dargestellt werden.

- ◆ Im Datenbankfenster FORMULARE aktivieren
 - ◆ Doppelt auf ERSTELLT EIN FORMULAR UNTER VERWENDUNG DES ASSISTENTEN klicken
- oder
- ◆ Einmal auf NEU klicken, im folgenden Listenfeld FORMULAR-ASSISTENT einstellen und OK klicken

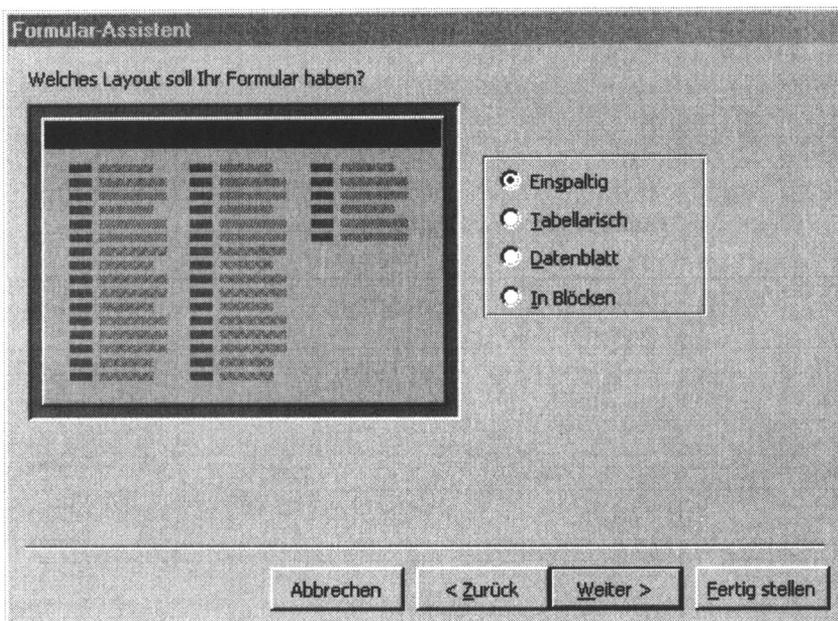
Zuerst wählen Sie im Listenfeld TABELLEN/ABFRAGEN die gewünschte Tabelle oder Abfrage aus. Die zu dieser Tabelle oder Abfrage gehörenden Datenfelder werden im Listenfeld VERFÜGBARE FELDER angeboten. Jetzt können Sie die interessierenden Datenfelder auswählen.



Wirkung der Steuerfelder:

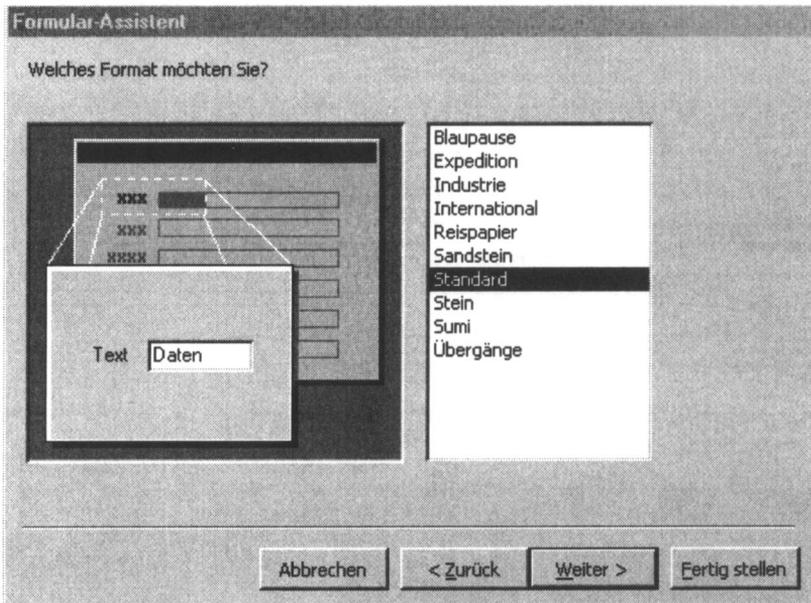
Taste	Wirkung
	wählt das markierte Feld aus:
	wählt alle Felder aus:
	hebt die Auswahl des markierten Feldes auf:
	hebt die gesamte Auswahl auf:

Sind alle gewünschten Felder übernommen worden, klicken Sie auf WEITER. Es erscheint das folgende Fenster:



Wenn Sie mit WEITER bestätigen, erhalten Sie ein Formular, das jeweils genau einen Satz anzeigt und die Felder untereinander in einer Spalte anordnet. Auf die anderen Ansichten soll hier nicht eingegangen werden.

Sie können im folgenden Fenster den Hintergrund des Formulars gestalten.



Das fertige Formular hat das folgende Aussehen:

Personen

ID	10050	Konfession	ev., später diss.
Name	Aderhold	Beruf	Bergmann
Vorname	Karl		
sex	m		
Geburtsort	Stolberg (Harz)		
GebTag	30		
GebMonat	6		
GebJahr	1884		
Sterbeort	Hannover		
SterbeTag	20		
SterbeMonat	6		
SterbeJahr	1921		

Datensatz: 1 von 12

Speichern Sie das Formular und schließen Sie es.

Aufgabe 8.5: Erstellen Sie das Formular mit dem Formularassistenten.

Mit dem Formularassistenten erstellte Formulare können nachträglich verändert werden. Es vereinfacht den Entwurf eines Formulars erheblich, wenn Sie erst den Formularassistenten benutzen und anschließend den Feinschliff von Hand ausführen.

Nehmen wir an, das Formular Personen soll zur Eingabe der Grunddaten weiterer Parlamentarier dienen. Da sowohl die Datenerfassung mit geringst möglichem Zeitaufwand erfolgen soll als auch die Fehlerquote gering zu halten ist, empfiehlt es sich, Maßnahmen zu treffen, die beiden Aspekten gerecht werden. Exemplarisch soll das an den Beispielen Konfession und Geburtsdatum demonstriert werden.

Die in den Quellen gemachten Angaben zur Konfession sind in geringem Maße formalisiert. Für einen evangelischen Abgeordneten traten in den Quellen folgende Angaben auf:

- ev
- ev.
- evang
- evang.
- evangel
- evangel.
- evangelisch

Die Konfession kodiert einzugeben, ist ebenfalls fehleranfällig. Vergibt man beispielsweise für 'evangelisch' den Code 4, so ist ein Tippfehler, wenn man statt '4' eine '5' eingibt, nur mit sehr hohem Aufwand zu finden.

Besser geeignet sind Kombinationsfelder. Man wählt die gewünschte Eingabe aus einer Liste vordefinierter Werte. Ist der einzugebende Wert nicht in der Liste enthalten, tippt man ihn ein.

Diese Liste kann direkt beim Kombinationsfeld verwaltet werden oder auch in einer Tabelle abgelegt sein. Wenn die Werte in einer Tabelle gespeichert sind, kann jede Eingabe eines neuen Wertes zur Liste hinzugefügt werden und für weitere Eingaben zur Verfügung stehen.

In diesem Beispiel sollen die Werte im Kombinationsfeld direkt gespeichert werden. Für die Eingabe des Geburtsdatums musste man 3 Felder wählen. Da ACCESS für den Felddatentyp Datum nur vollständige korrekte Datumsangaben akzeptiert, für den Personenkreis der Parlamentarier aber teilweise nur Geburtsjahr bzw. Sterbejahr zu recherchieren waren, muss das Datum in drei Datenfeldern als Tag, Monat und Jahr gespeichert werden. Die Datenprüfung

soll nicht das Formular übernehmen, sondern sie soll beim Speichern der Werte in die Tabelle erfolgen.

- ◆ Im Datenbankfenster TABELLEN aktivieren
- ◆ Tabelle Personen anklicken
- ◆ Im Datenbankfenster auf ENTWURF klicken
- ◆ Die Zeile Geburtstag anklicken
- ◆ Im Feldeigenschaftsbereich in Zeile Gültigkeitsregel ≥ 0 Und ≤ 31 eingeben
- ◆ Analoge Regeln für Geburtsmonat, Geburtsjahr und Sterbedaten eingeben

Aufgabe 8.6: Legen Sie für die Datenfelder Geburtstag, Geburtsmonat, Geburtsjahr, Sterbetag, Sterbemonat und Sterbejahr der Tabelle Personen Gültigkeitsregeln fest. Geben Sie für fehlerhafte Eingaben eine Fehlermeldung aus.

8.6.6 Verändern eines bestehenden Formulars

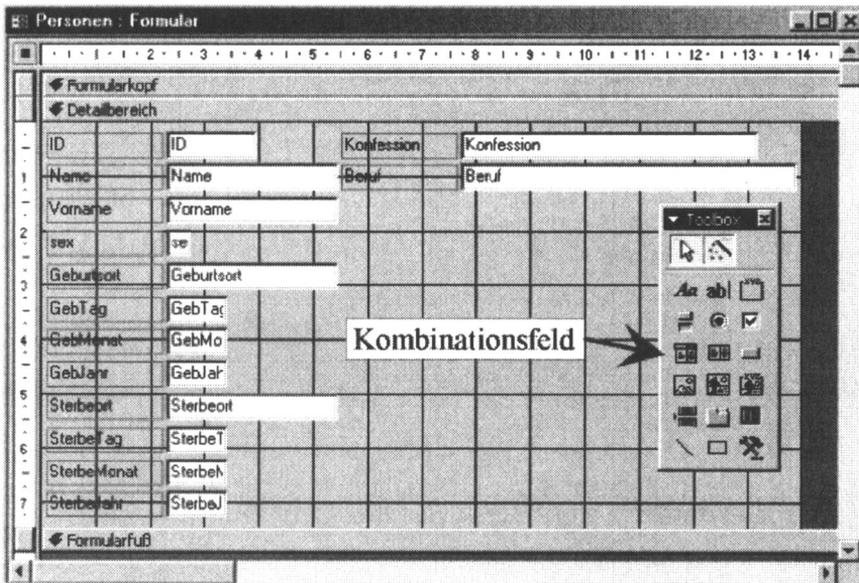
- ◆ Im Datenbankfenster FORMULARE aktivieren
- ◆ Klicken auf das Formular Personen
- ◆ Klicken auf ENTWURF in der Menüleiste des Datenbankfensters

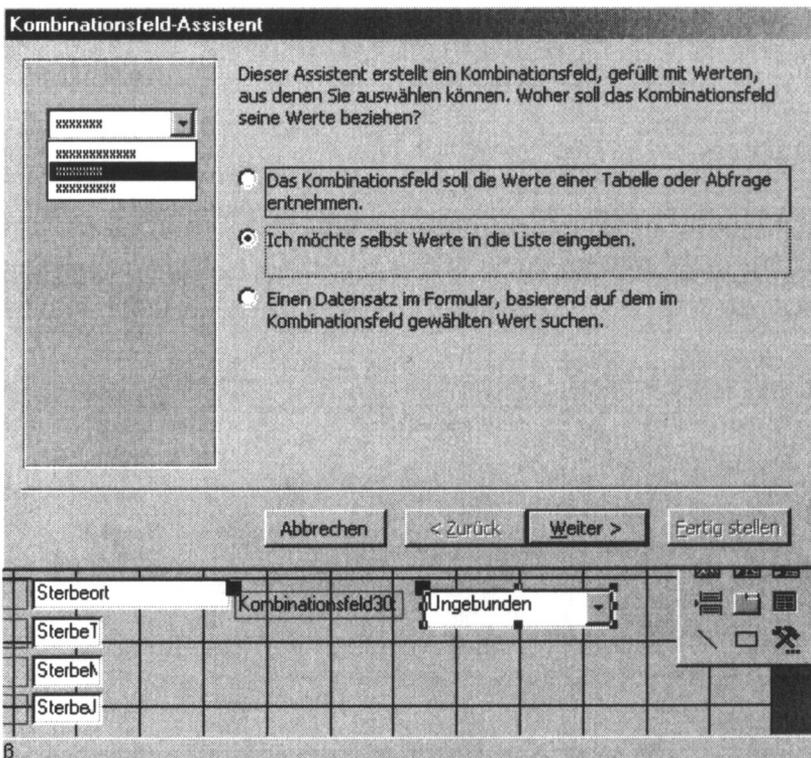
Sie erhalten das Formular Personen in der Entwurfsansicht. Das Formular besteht aus den drei Bereichen:

- Formulkopf,
- Detailbereich, in dem die mit Hilfe des Assistenten erzeugten Felder platziert sind,
- Formularfuß.

Sie können den Formulkopf- und den Formularfußbereich aufziehen und auch dort Elemente hinzufügen. Wir wollen mit Hilfe der Toolbox ein Kombinationsfeld hinzufügen:

- ◆ Klicken Sie in der Toolbox auf das Symbol für KOMBINATIONSFELD.
- ◆ Ziehen Sie das Feld auf.
ACCESS startet einen Assistenten, dessen Anweisungen Sie folgen.
- ◆ Sie wählen den mittleren Eintrag ICH MÖCHTE SELBST WERTE IN DIE LISTE EINGEBEN und klicken auf WEITER.





- ◆ Die Liste soll aus einer Spalte bestehen und die Werte
 - ev.
 - ev., später diss.
 - kath.
 - kath., später diss.
 - jüdisch
 - enthalten.

Kombinationsfeld-Assistent

Welche Werte sollen in Ihrem Kombinationsfeld erscheinen? Geben Sie bitte die Anzahl der Spalten der Liste und dann die Werte für jede Zelle ein.

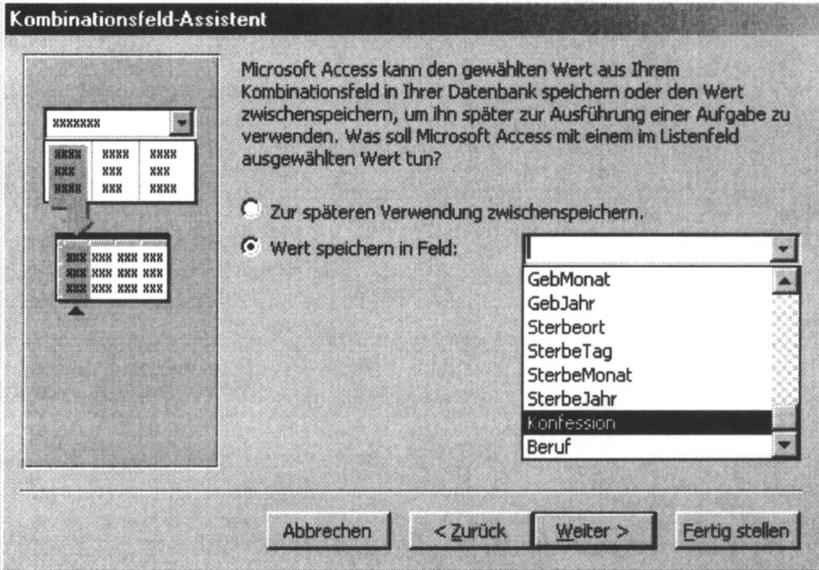
Um die Breite einer Spalte anzupassen, ziehen Sie entweder die rechte Begrenzung auf die gewünschte Breite, oder doppelklicken Sie auf die rechte Begrenzung, um die optimale Breite zu erhalten.

Spaltenanzahl:

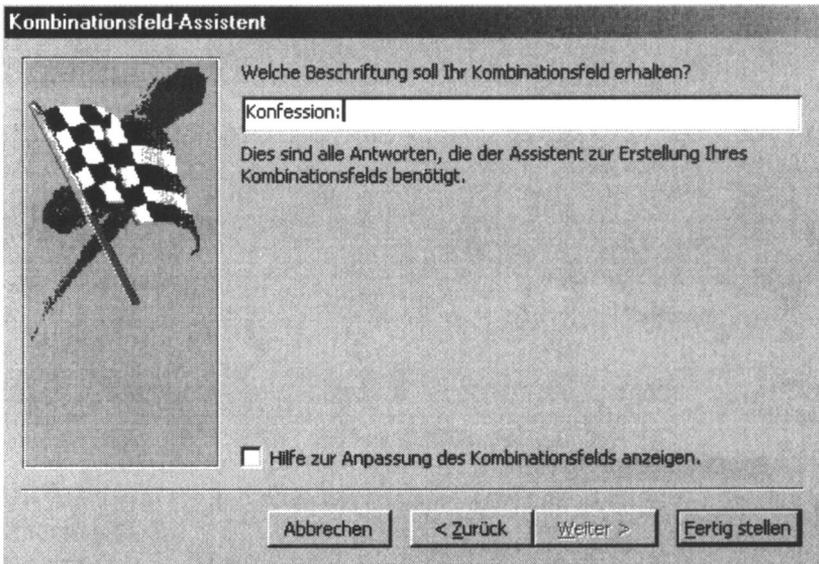
	Sp1
<input type="checkbox"/>	ev.
<input type="checkbox"/>	ev., später diss.
<input type="checkbox"/>	kath.
<input type="checkbox"/>	kath., später diss
<input type="checkbox"/>	jüdisch
<input checked="" type="checkbox"/>	keine Angabe
<input type="checkbox"/>	*

Abbrechen < Zurück Weiter > Fertig stellen

- ◆ Klicken Sie auf WEITER.
- ◆ Im nächsten Bild geben Sie an, wohin der Wert zu speichern ist. Wir möchten, dass die späteren Eingaben in das Datenfeld Konfession eingetragen werden.



- ◆ Mit dem nächsten Fenster geben Sie dem Kombinationsfeld einen Namen, hier *Konfession*:



◆ Klicken Sie auf FERTIG STELLEN und schließen Sie das Formular.

Wenn Sie das Formular öffnen und den Datenbestand mit Hilfe der Navigationsschaltflächen am unteren Rand durchblättern, werden Sie feststellen, dass im erstellten Kombinationsfeld die Konfessionszugehörigkeit angezeigt wird. Das ursprüngliche Textfeld Konfession, das vom Formularassistenten erzeugt wurde und das neue Kombinationsfeld Konfession: zeigen den gleichen Inhalt an. Das Textfeld kann vom Formular entfernt werden, indem Sie

- ◆ das Formular in der Entwurfsansicht öffnen,
- ◆ das Textfeld Konfession anklicken,
- ◆ die Entfernen-Taste drücken.

Schließen Sie anschließend das Formular.

Aufgabe 8.7: Verändern Sie das in Aufgabe 8.5 erstellte Formular, indem Sie für Konfession ein Kombinationsfeld einrichten.

Aufgabe 8.8: Geben Sie den folgenden Datensatz ein:

ID	Name	Vorname	sex	Geburtsort	Geburtsdatum			Sterbeort	Sterbedatum			Konf.	Beruf
20020	Baade	Fritz	m	Neuruppin	23	1	1893	Kiel	15	5	1974	ev.	Seminardirektor

Aufgabe 8.9: Das Kombinationsfeld soll über weitere Werte verfügen. Die Liste soll erweitert werden um die Werte „keine Angabe“ und „nicht bekannt“.
Hinweis: Rechte Maustaste → EIGENSCHAFTEN → DATENSATZ-HERKUNFT und Anfügen der neuen Werte.

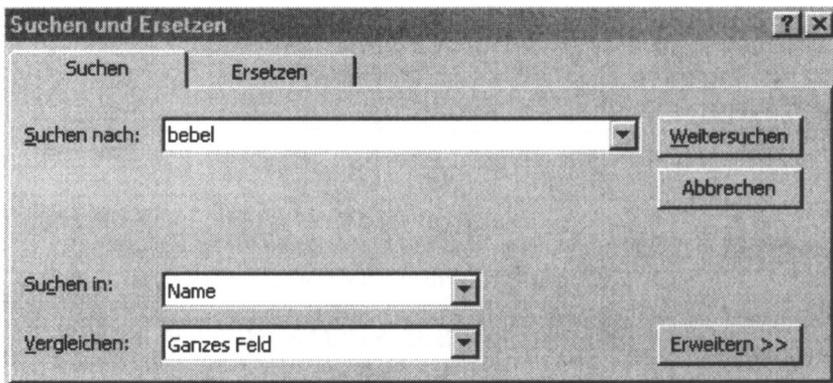
Aufgabe 8.10: Laden Sie die Übungsdatenbank SPDParl-AB.mdb von <http://www.hsr-trans.de>

8.7 Suchen, Sortieren, Filtern

8.7.1 Suchen und Ersetzen

Wir wollen in der Tabelle **Personen** der heruntergeladenen Datenbank **SPDParl.mdb** gezielt nach bestimmten Abgeordnete suchen. Dazu benutzen wir die 'SUCHEN UND ERSETZEN'-Funktion.

- ◆ Öffnen Sie die Tabelle **Personen**
- ◆ gehen Sie mit dem Cursor in das Feld **Name**
- ◆ Wählen Sie den Menüpunkt **BEARBEITEN → SUCHEN**
oder **[STRG] [E]**
oder **⌘**
- ◆ geben Sie den Suchbegriff ein, z.B. **'Bebel'**.



Probieren Sie selbständig weitere Varianten:

- Groß-, Kleinschreibung,
- Benutzen Sie die Platzhalter '?' und '*'.
? steht für ein beliebiges Zeichen,
* steht für eine beliebige Zeichenfolge.
Geben Sie beispielsweise **'M??er'**, **'M*er'** in das Suchfeld ein.
- Klicken Sie auf **WEITERSUCHEN**.

- Stellen Sie im Feld Vergleichen auf Anfang des Feldinhaltes und Teil des Feldinhaltes.

Diese Suchvorgänge können Sie auf die gleiche Weise in Formularen durchführen.

8.7.2 Sortieren

Mit der Sortierfunktion lässt sich die Tabelle nach einer oder mehreren Spalten sortieren. Die physische Reihenfolge der Zeilen bleibt dabei erhalten.

Sie können aufsteigend oder absteigend sortieren. Bei der deutschen Version von ACCESS gilt als Sortierreihenfolge:

- Sonderzeichen vor Ziffern,
- Ziffern vor Buchstaben,
- ß wie s.

Wenn nicht anders eingestellt, wird nicht zwischen Groß- und Kleinschreibung unterschieden.

◆ Öffnen Sie die Tabelle Personen

◆ Gehen Sie mit dem Cursor in das Feld Gebjahr

◆ Wählen Sie den Menüpunkt DATENSÄTZE → SORTIERUNG → ABSTEIGEND

oder 

	ID	Name	Vorname	sex	Geb	Geb	GebJahr
▶ +	50040	Ebert	Emil	m	0	0	0
+	201283	Schulze	Karl	m	0	0	0
+	210385	Stöckigt	Karl	m	0	0	0
+	120355	Leps	Karl	m	0	0	0
+	120320	Leistner	Kurt	m	0	0	0
+	40445	Drechsel	Auguste	w	0	0	0
+	40485	Drescher	Emil	m	0	0	0
+	40530	Drews	Friedrich	m	0	0	0
+	80554	Heilmann	Robert	m	0	0	0
+	250320	Wehle	Johann	m	0	0	0

Kontrollieren Sie die Geburtsjahre auf korrekte Werte. Zugelassen ist der Wert 0 für nicht eingegebene Jahre. Sinnvollerweise kann man annehmen, dass die Geburtsjahre größer als 1750 sind. Klicken sie auf das Symbol für absteigende

Sortierung und prüfen Sie die größten Jahreszahlen. Prüfen Sie in analoger Weise, ob die Monate zwischen 0 und 12 sowie die Tage zwischen 0 und 31 liegen.

Man kann auch nach mehreren Spalten sortieren. Gehen Sie folgendermaßen vor:

- ◆ Bringen Sie die Spalten in die Reihenfolge, in der sortiert werden soll; Sie können die Reihenfolge der Spalten verändern, indem Sie die Spalte markieren und per Drag & Drop an die gewünschte Position ziehen.
- ◆ Markieren Sie die Spalten, indem sie mit gedrückter Maustaste über die Spaltenüberschriften fahren, und
- ◆ klicken Sie anschließend das gewünschte Sortierungssymbol.

Beim Schließen einer Tabelle wird die aktuelle Sortierung beibehalten. Die Sortierung lässt sich aufheben durch den Menü-Befehl DATENSÄTZE → FILTER/SORTIERUNG ENTFERNEN.

8.7.3 Filter

8.7.3.1 Auswahlbasierter Filter

Filter bieten vielfältigere Möglichkeiten, als SUCHEN. Wir wollen aus der Tabelle Personen die weiblichen Abgeordneten des Reichstages herausfiltern. Sehr einfach ist diese Aufgabe mit einem auswahlbasierten Filter zu lösen.

- ◆ Setzen Sie den Cursor in eine Zelle der Spalte 'Sex', die den Wert 'w' beinhaltet.
- ◆ Wählen Sie den Menüpunkt DATENSÄTZE → FILTER → AUSWAHLBASIERTER FILTER
oder klicken Sie auf das Symbol .
- ◆ Den Filter entfernen Sie durch Klicken auf .

ID	Name	Vorname	sex	Geburtsort	geboren
10060	Agnes	Lore	w	Bochum	04.06.1876
10160	Ansorge	Marie	w	Löchau	15.12.1880
10220	Arendsee	Martha	w	Berlin	29.03.1885
10240	Arning	Marie	w	Bramsche	19.04.1887
20100	Bartels	Elise	w	Hildesheim	13.05.1880
20430	Behncke	Marie	w	Dangastermoor	21.02.1880
20810	Blos	Anna	w	Liegnitz	04.08.1866

Datensatz: 1 von 51 (Gefiltert)

Sie können bei dieser Art von Filter nur jeweils für ein Datenfeld eine Bedingung setzen.

8.7.3.2 Auswahlausschließender Filter

Es ist auch möglich, alle die Zeilen anzeigen zu lassen, die eine Bedingung nicht erfüllen:

- ◆ Setzen Sie den Cursor in die Spalte 'Geburtsort'.
- ◆ Wählen Sie den Menüpunkt DATENSÄTZE → FILTER → AUSWAHLAUSSCHLIESSENDER FILTER.

8.7.3.3 Formularbasierter Filter

Formularbasierte Filter können auf Tabellen oder Formulare angewendet werden. Filtern Sie aus dem Formular frmAbgeordnete alle diejenigen Reichstagsabgeordneten, die in Magdeburg geboren wurden und in einem Wahlkreis Preußens oder Anhalts ihr Erstmandat erhielten.

- ◆ Öffnen Sie das Formular FrmAbgeordnete.
- ◆ Wählen Sie den Menüpunkt DATENSÄTZE → FILTER → FORMULARBASIERTER FILTER
oder klicken Sie .
- ◆ Wählen Sie in der Dropdown-Liste für Geburtsort 'Magdeburg' und in der Liste für Land 'Preußen' aus.

Abgeordnete: Formularbasierter Filter

ID	<input type="text"/>	RTErsteintritt	<input type="text"/>
Name	<input type="text"/>	LP	<input type="text"/>
Vorname	<input type="text"/>	Bezeichnung	<input type="text"/>
sex	<input type="checkbox"/>	Region	<input type="text"/>
Geburtsort	"Magdeburg" ▾	Land	"Preußen"
geboren	<input type="text"/>	Provinz	<input type="text"/>
Sterbeort	<input type="text"/>	RGB	<input type="text"/>
gestorben	<input type="text"/>		
Beruf des Vaters	<input type="text"/>		
Ausbildung	<input type="text"/>		

Suchen nach Oder Oder

- ◆ Klicken Sie die Lasche ODER.
- ◆ Wählen Sie den wieder den Geburtsort 'Magdeburg' und diesmal für Land 'Anhalt' aus.

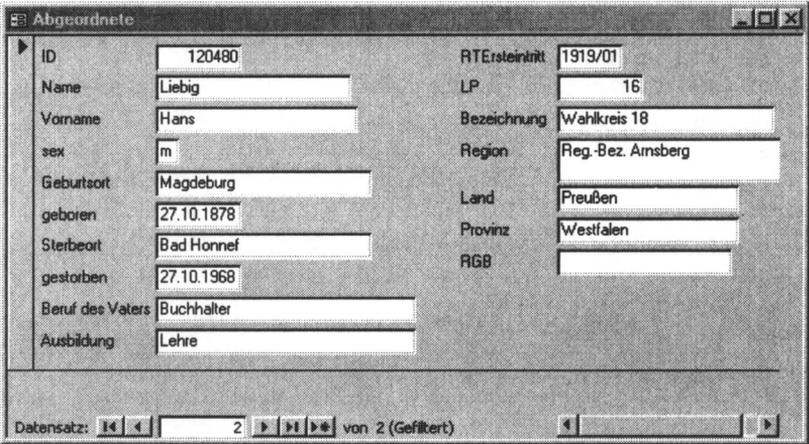
Abgeordnete: Formularbasierter Filter

ID	<input type="text"/>	RTErsteintritt	<input type="text"/>
Name	<input type="text"/>	LP	<input type="text"/>
Vorname	<input type="text"/>	Bezeichnung	<input type="text"/>
sex	<input type="checkbox"/>	Region	<input type="text"/>
Geburtsort	"Magdeburg" ▾	Land	"Anhalt"
geboren	<input type="text"/>	Provinz	<input type="text"/>
Sterbeort	<input type="text"/>	RGB	<input type="text"/>
gestorben	<input type="text"/>		
Beruf des Vaters	<input type="text"/>		
Ausbildung	<input type="text"/>		

Suchen nach **Oder** Oder

- ◆ Klicken Sie auf das Symbol , um den Filter zu aktivieren.
- ◆ Sie speichern den Filter, indem Sie auf das Symbol  klicken.

Das gefilterte Formular liefert zwei Datensätze. Der zweite Datensatz ist im folgenden Bild angezeigt:



The screenshot shows a data entry form titled "Abgeordnete" with the following fields and values:

ID	120480	RTErsteintritt	1919/01
Name	Liebig	LP	16
Vorname	Hans	Bezeichnung	Wahlkreis 18
sex	m	Region	Reg.-Bez. Amsberg
Geburtsort	Magdeburg	Land	Preußen
geboren	27.10.1878	Provinz	Westfalen
Sterbeort	Bad Honnef	RGB	
gestorben	27.10.1968		
Beruf des Vaters	Buchhalter		
Ausbildung	Lehre		

At the bottom of the form, it displays "Datensatz: 2 von 2 (Gefiltert)", indicating that 2 records are shown out of a total of 2 filtered records.

Der Filter wird als Abfrage gespeichert. In der Eigenschaftstabelle des Formulars ist unter der Eigenschaft 'FILTER' die Filterbedingung gespeichert.

8.8 Abfragen

Eine Abfrage kann man sich als eine virtuelle Tabelle vorstellen. Sie sieht aus wie eine Tabelle und kann (fast) überall dort stehen, wo Tabellen eingesetzt werden. Das Ergebnis einer Abfrage wird allerdings nicht gespeichert, sondern nur die Regeln, woher die Daten geholt und wie sie präsentiert werden.

In Kapitel 3 dieses Skripts wird die relationale Algebra erläutert. Eine 'Rechenoperation' mit einer oder mehreren Relationen (Tabellen) ergibt als Ergebnis wieder eine Relation. Damit kann die Verknüpfung von Relationen auch dort stehen, wo eine Relation vorgesehen ist.

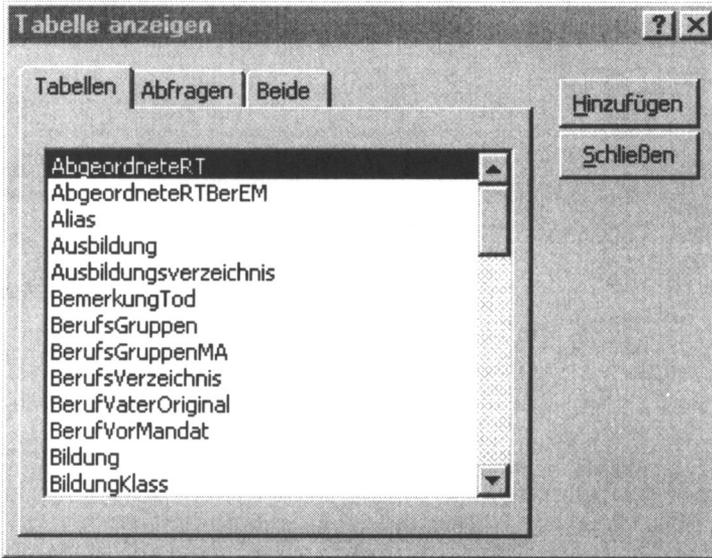
Die 'Rechenoperationen' der relationalen Algebra werden in ACCESS durch Abfragen realisiert. Die Grundlage eines Formulars muss nicht eine Tabelle sein. Das Formular kann ebenso gut auf einer Abfrage beruhen.

Mit Abfragen können Sie:

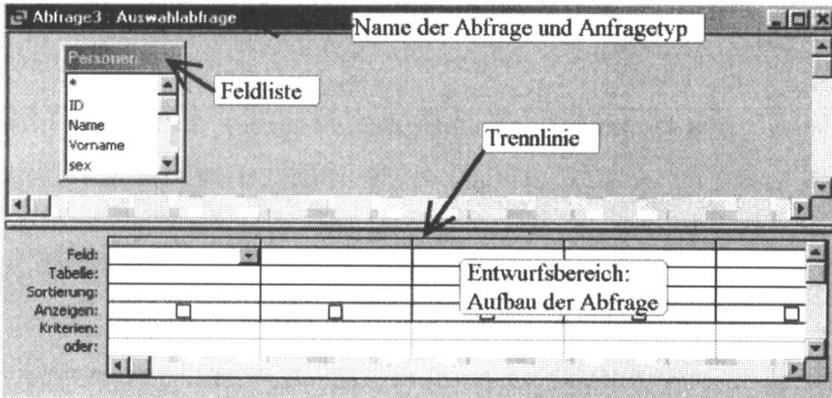
- eine Auswahl von Zeilen und Spalten aus einer Tabelle treffen (SELECTION, PROJECTION),
- die Daten aus mehreren Tabellen verbinden (JOIN), anzeigen und bearbeiten;
- Daten nach verschiedensten Kriterien sortieren;
- Berechnungen durchführen, z.B. aus Geburtsdatum und Sterbedatum das Alter ermitteln oder bestimmte Aggregatfunktionen (ANZAHL, MITTELWERT, STREUUNG) ausführen;
- Gruppen von Daten (Zeilen) aktualisieren, löschen oder korrigieren;
- doppelte Datensätze suchen, doppelte Zeilen unterdrücken (lokale Integrität);
- Datensätze suchen, die keinen über eine Spalte verbundenen Datensatz in einer zweiten Tabelle besitzen (globale Integrität), z.B. ein Reichstagsmandat ermitteln, für das kein Mandatsträger angegeben ist.

8.8.1 Auswahlabfrage manuell erstellen

- ◆ Aktivieren Sie im Datenbankfenster den Menüpunkt ABFRAGEN.
- ◆ Klicken Sie auf NEU.
- ◆ Klicken Sie doppelt auf den Eintrag ENTWURFSANSICHT.
- ◆ Wählen Sie die Tabelle aus, für die eine Abfrage erstellt werden soll, und klicken Sie auf HINZUFÜGEN.
- ◆ Blenden Sie das Dialogfenster durch Klick auf die Schaltfläche SCHLIEßEN aus.



Auf dem Bildschirm erscheint das Abfragefenster in der Entwurfsansicht.



Möchten Sie zu einem späteren Zeitpunkt weitere Tabelle oder Abfragen in die Abfrage einbinden, so klicken sie in der Abfrageentwurfsansicht auf das Icon  der Symbolleiste und das Fenster TABELLE ANZEIGEN öffnet sich erneut. Sie können die Laschen TABELLEN, ABFRAGEN oder BEIDE wählen und die gewünschten Objekte durch Doppelklick ins Abfragefenster ziehen.

8.8.1.1 Felder in die Abfrage aufnehmen

Im Abfragefenster sind die Felder zu bestimmen, die in die Abfrage aufgenommen werden sollen. Die Feldliste der Abfrage enthält alle Datenfelder der Tabelle. Zur Auswahl von Feldern stehen Ihnen mehrere Möglichkeiten zur Verfügung:

- ◆ Zur Auswahl eines Feldes klicken Sie dieses mit der Maus in der Feldliste an und ziehen es in eine freie Zelle der Zeile Feld des Entwurfsbereichs.
- ◆ Markieren Sie bei gedrückter [STRG]-Taste mehrere Felder in der Feldliste und ziehen Sie sie in den Entwurfsbereich.
- ◆ Klicken Sie im Entwurfsbereich in die Zeile FELD und wählen Sie im dann sichtbaren Listenfeld ein Feld aus.
- ◆ Doppelklicken Sie auf einen Feldnamen in der Feldliste.

8.8.1.2 Sortierreihenfolge der Abfrage festlegen

Nach welchem Feld die Daten im Abfrageergebnis sortiert werden, kann durch die Sortierung festgelegt werden:

- ◆ Klicken Sie in der Zeile SORTIERUNG das Feld an, für das Sie eine Sortierung festlegen möchten.
- ◆ Wählen Sie zwischen den Sortierreihenfolgen AUFSTEIGEND und ABSTEIGEND.

Sie können auch nach mehreren Feldern sortieren. Die Reihenfolge der Sortierung erfolgt von dem am weitesten links stehenden Feld zu dem am weitesten rechts stehenden Feld.

8.8.1.3 Felder aus der Abfrage entfernen

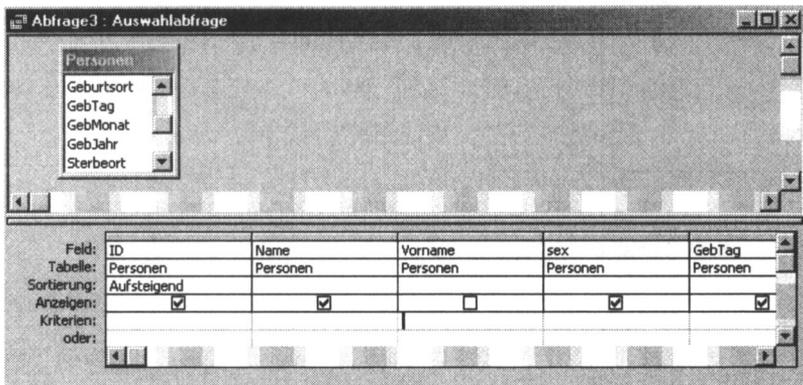
- ◆ Setzen Sie den Cursor in die betreffende Spalte und wählen Sie den Menüpunkt BEARBEITEN → SPALTE LÖSCHEN.
oder Markieren Sie die Spalte durch Klick auf den Spaltenkopf und betätigen Sie die [ENTF]-Taste.

Sie können auch die Anzeige unterdrücken, indem Sie auf das Kästchen in der Zeile ANZEIGEN klicken:

ID	Name	Vorname	sex
Personen	Personen	Personen	Personen
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

8.8.2 Abfrage ausführen

- ◆ Klicken Sie auf das Symbol !
oder rufen Sie den Menüpunkt ABFRAGE → AUSFÜHREN auf.



ID	Name	sex	GebTag	GebMonat	GebJahr
10050	Aderhold	m	30	6	1884
10060	Agnes	w	4	6	1876
10070	Agster	m	12	4	1858
10100	Albrecht	m	14	7	1855
10160	Ansorge	w	15	12	1880
10180	Antrick	m	24	11	1858
10220	Arendsee	w	29	3	1885
10240	Arning	w	19	4	1887
10280	Arzt	m	9	10	1880
10310	Auer	m	22	12	1874
10320	Auer	m	19	4	1846
10330	Aufhäuser	m	1	5	1884
20020	Baade	m	23	1	1893

Datensatz: 6 von 13

8.8.3 Abfrage speichern

- ◆ Rufen Sie den Menüpunkt DATEI → SPEICHERN auf oder klicken Sie das entsprechende Symbol .
- ◆ Geben Sie im folgenden Dialogfenster einen Namen für die Abfrage ein und bestätigen Sie mit Ok

Aufgabe 8.11: Erstellen Sie eine Abfrage auf der Grundlage der Tabelle Personen. Die Sortierung soll nach Geburtsjahr, Geburtsmonat, Geburtstag erfolgen.

8.8.4 Tabellen verknüpfen

Mit Hilfe von Abfragen lassen sich Tabellen verknüpfen. Sollen nicht die Abgeordneten im Mittelpunkt stehen, sondern die Reichstagsmandate, aber mit den Personendaten der Mandatsträger, so ist die Tabelle der Reichstagsmandate (MdR) mit der Personentabelle zu verknüpfen. Die Tabelle MdR enthält alle Mandate mit den dazu gehörenden Informationen, die Tabelle Personen alle direkt zur Person gehörenden Merkmale.

MDR Tabelle				
ID	LP	vonMonat	vonJahr	
10050	16	3	1919	
10050	17	6	1920	
10060	16	1	1919	
10060	17	6	1920	
10060	18	5	1924	
10060	19	12	1924	
10060	20	5	1928	
10060	21	9	1930	
10060	22	7	1932	
10060	23	11	1932	
10060	24	3	1933	
10070	12	6	1898	
10100	12	6	1898	
10100	13	1	1905	

Personen Tabelle					
ID	Name	Vorname	sex	GebTag	G
10050	Aderhold	Karl	m	30	
10060	Agnes	Lore	w	4	
10070	Agster	Alfred	m	12	
10100	Albrecht	Adolf	m	14	
10160	Ansoerge	Marie	w	15	
10180	Antrick	Otto	m	24	
10220	Arendsee	Martha	w	29	
10240	Arning	Marie	w	19	
10280	Arzt	Arthur	m	9	
10310	Auer	Erhard	m	22	
10320	Auer	Ignaz	m	19	
10330	Aufhäuser	Siegfried	m	1	
20020	Baade	Fritz	m	23	

Das Vorgehen ist analog zu dem mit einer einzelnen Tabelle.

8.8.4.1 Eine Verbindung zwischen den Tabellen herstellen

- ◆ Entwerfen Sie eine neue Abfrage. Fügen Sie die Tabellen MDR und Personen hinzu.
- ◆ Klicken Sie auf das Feld ID in der Feldliste von MDR und ziehen Sie es bei gedrückter Maustaste auf das Feld ID der Feldliste Personen.

Feld:	ID	LP	vonMonat	vonJahr	bisMonat	bisJahr	Name	Vorna-
Tabelle:	MDR	MDR	MDR	MDR	MDR	MDR	Personen	Personen
Sortierung:								
Anzeigen:	<input checked="" type="checkbox"/>	<input type="checkbox"/>						
Kriterien:								
oder:								

Damit sind die beiden Tabellen über die Felder ID verbunden. Es werden die Zeilen mit gleichen Dateninhalten, also gleicher ID, verbunden.

Es werden nur die Sätze angezeigt, für die in beiden Tabellen Werte für ID vorliegen. Enthält die Tabelle Personen Abgeordnete, die kein Mandat für eine Legislaturperiode des Reichstages hatten, so werden diese nicht angezeigt. Das gleiche gilt für Mandate, die auf keinen Mandatsträger verweisen (fehlerhafte oder noch nicht bekannte Personen-ID). Es handelt sich hier um einen INNER JOIN (s. Abschnitt 3.2.2).

Wenn man doppelt auf die Verbindungslinie zwischen den Tabellen klickt, öffnet sich ein Fenster, in dem man die Art des JOINS wählen kann, also auch OUTER JOINS aus der Sicht der einen oder anderen Tabelle. Ein OUTER JOIN aus Sicht der Tabelle Personen liefert alle Mandate und zusätzlich Zeilen für die Personen, die nicht im Reichstag vertreten waren mit NULL-Werten für die MdR-Merkmale.

Der OUTER JOIN aus Sicht der Tabelle MdR (Mandate) liefert alle Mandate verknüpft mit den Personendaten. Da jedes Mandat auf einen Mandatsträger in der Personen-Tabelle verweist, liefert der OUTER JOIN hier das gleiche Ergebnis wie der INNER JOIN. Wir müssen also in diesem Beispiel nicht den Typ der Verbindung verändern.

Welches Ergebnis wird uns die Abfrage liefern?

Die Tabelle MDR enthält alle Mandate mit den Informationen:

- Wer hatte das Mandat inne (ID des Parlamentariers)?
- In welcher Reichstagslegislaturperiode (Legislaturperioden-Nr.)?
- Von wann bis wann gehörte er in dieser Periode dem Parlament an?
- In welchem Wahlkreis wurde er gewählt (Nr. des Wahlkreises = WK)?
- Welche Besonderheiten traten auf (Bemerkung)?

Zu jedem Mandat wird über die ID der Name des Mandatsträgers aus der Tabelle Personen bereitgestellt.

ID	LP	vonMonat	vonJahr	bisMonat	bisJahr	Name	Vorname
10050	16	3	1919	6	1920	Aderhold	Karl
10050	17	6	1920	6	1921	Aderhold	Karl
10060	16	1	1919	6	1920	Agnes	Lore
10060	17	6	1920	5	1924	Agnes	Lore
10060	18	5	1924	12	1924	Agnes	Lore
10060	19	12	1924	5	1928	Agnes	Lore
10060	20	5	1928	9	1930	Agnes	Lore
10060	21	9	1930	7	1932	Agnes	Lore
10060	22	7	1932	11	1932	Agnes	Lore
10060	23	11	1932	3	1933	Agnes	Lore
10060	24	3	1933	6	1933	Agnes	Lore

10070	12	6	1898	6	1903	Agster	Alfred
10100	12	6	1898	6	1903	Albrecht	Adolf
10100	13	1	1905	1	1907	Albrecht	Adolf
10100	14	1	1907	1	1912	Albrecht	Adolf

Aufgabe 8.12: Erstellen Sie eine Abfrage über alle Mandate (Tabelle MDR), die neben den Feldern ID, LP, vonMonat, vonJahr noch den Namen, Vornamen und Geburtstag des Mandatsträgers enthält.

8.8.4.2 Vordefinierte Beziehungen

Die Beziehungen zwischen den Tabellen lassen sich vordefinieren. ACCESS greift beim Entwurf einer Abfrage auf diese Verknüpfungen zurück. Sie lassen sich aber in der Abfrageentwurfsansicht ändern. Diese Änderungen wirken dann aber nur lokal, d.h. nur für diese Abfrage. Die vordefinierten Beziehungen bleiben erhalten. So definieren Sie Beziehungen vor:

- ◆ Klicken Sie in der Menü-Leiste auf EXTRAS → BEZIEHUNGEN
- oder
- ◆ klicken Sie in der Symbol-Leiste auf das Icon 
- ◆ Klicken Sie in der Menü-Leiste auf BEZIEHUNGEN → TABELLE ANZEIGEN, um die Tabellen und Abfragen anzuzeigen,
- oder
- ◆ klicken Sie in der Symbol-Leiste auf das Icon 
- ◆ Sie fügen Tabellen oder Abfragen hinzu, indem Sie sie markieren und auf HINZUFÜGEN klicken.
- ◆ Die Verbindung realisieren Sie, indem Sie die Spalte einer der beiden Tabellen, über die die Verbindung vorgenommen werden soll, anklicken und per Drag & Drop auf die entsprechende Spalte der anderen Tabelle fallen lassen.
- ◆ Die Art der Verbindung wählen Sie im Verknüpfungsfenster.

8.8.5 UND- und ODER-Verknüpfungen

Wenn nicht alle Zeilen der Tabelle interessieren, kann man Kriterien vorgeben, die die Auswahl einschränken (SELECTION). Bedingungen für eine Abfrage werden im Entwurfsbereich in der Zeile KRITERIEN definiert. Sie können in verschiedenen Zeilen und Feldern eine oder auch mehrere Bedingungen eintragen.

UND-Verknüpfung

Die Bedingungen einer Zeile werden durch UND miteinander verknüpft. Nur wenn alle Bedingungen einer Zeile erfüllt sind, wird der Datensatz in das Ergebnis aufgenommen.

Beispiel: *Alle Abgeordneten, die vor 1900 geboren wurden und die nach 1945 noch gelebt haben.*

Geburtsjahr < 1900 UND Sterbejahr > 1945

Feld:	ID	Name	Vorname	Geb.Jahr	SterbeJahr
Tabelle:	Personen	Personen	Personen	Personen	Personen
Sortierung:					
Anzeigen:	<input checked="" type="checkbox"/>				
Kriterien:				<1900	>1945
oder:					

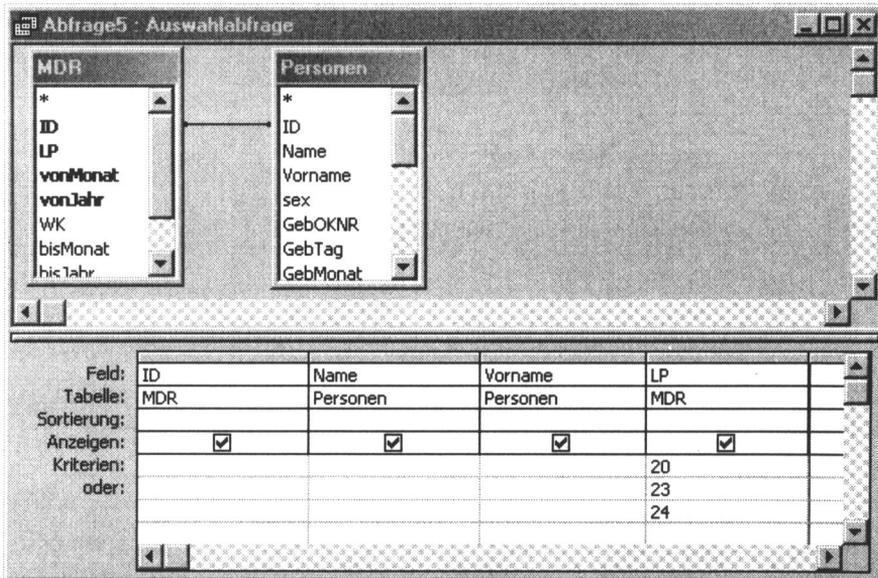
ODER-Verknüpfung

Wenn Sie in zwei Zeilen Bedingungen eintragen, werden die Zeilen durch ODER verbunden. Ist die Bedingung mindestens einer Zeile wahr, wird der Datensatz in die Ergebnisrelation aufgenommen.

Beispiel: *Alle sozialdemokratischen Reichstagsmandate der 20., 23. oder 24. Reichstagslegislaturperiode.*

LP = 20 ODER LP = 23 ODER LP = 24

Reichstagslegislaturperioden		
LP	von	bis
20	5 / 1928	9 / 1930
21	9 / 1930	7 / 1932
22	7 / 1932	11 / 1932
23	11 / 1932	3 / 1933
24	3 / 1933	6 / 1933



Beim Speichern der Abfrage versucht ACCESS zu vereinfachen. Die ODER-Bedingung wird intern umgewandelt in
20 ODER 23 ODER 24.

ID	Name	Vorname	LP
MDR	Personen	Personen	MDR
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			20 Oder 23 Oder 24

Das geht aber nur, wenn keine zusätzlichen UND-Verknüpfungen auftreten. Sucht man beispielsweise alle diejenigen, die nach März 1933 noch dem Reichstag angehörten und auch jene, welche im November 1932 noch vertreten waren, dann aber starben, so bekommt man folgenden logischen Ausdruck (LP=24) ODER (LP=23 UND Sterbemonat<3 UND Sterbejahr<=1933).

ID	Name	Vorname	LP	SterbeJahr	SterbeMonat
MDR	Personen	Personen	MDR	Personen	Personen
<input checked="" type="checkbox"/>					
			23	<=1933	<3
			24		

Die Abfrage liefert folgendes Ergebnis (Ausschnitt):

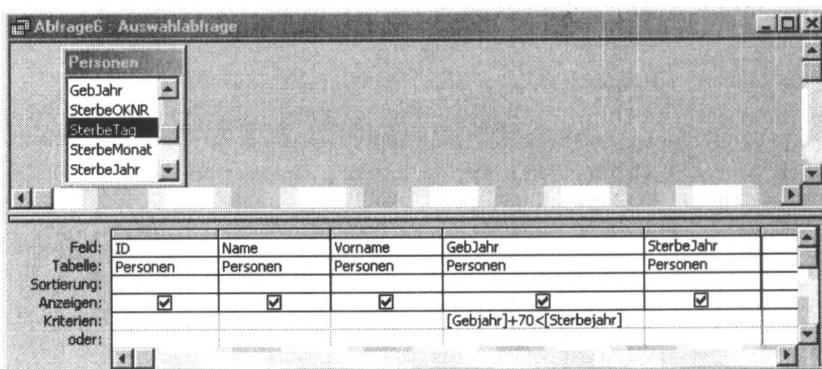
ID	Name	Vorname	LP	SterbeJahr	SterbeMonat
130370	Meier	Richard	23	1933	1
20680	Biedermann	Adolf	24	1933	5
50130	Eggerstedt	Otto	24	1933	10
160280	Pfölf	Antonie	24	1933	6
210290	Stelling	Johannes	24	1933	6
130210	Marum	Ludwig	24	1934	3
251110	Wurm	Mathilde	24	1934	4
81640	Husemann	Friedrich	24	1935	4
20937	Bohm-Schuch	Klara	24	1936	6
...

Aufgabe 8.13: Erstellen Sie eine Abfrage, die alle Reichstagsmandate der 20., 21. und 22. Legislaturperiode enthält.

Es ist auch möglich, komplexere logische Ausdrücke zu formulieren, in denen verschiedene Spalten miteinander verglichen werden. Die Spaltennamen müssen dabei in eckige Klammern gefasst werden.

Beispiel: *Alle Abgeordneten, die älter als 70 Jahre alt wurden.*

$[Geburtsjahr] + 70 > [Sterbejahr]$



	ID	Name	Vorname	GebJahr	SterbeJahr
▶	10020	Ackermann	Friedrich	1876	1949
	10060	Agnes	Lore	1876	1953
	10080	Ahrens	Heinrich	1876	1951
	10100	Albrecht	Adolf	1855	1930
	10110	Amann	Josef	1879	1971
	10120	Amlung	Georg	1896	1973
	10130	Ammon	Lina	1889	1968
	10160	Ansorge	Marie	1880	1955
	10190	Apel	Karl	1850	1929
	10200	Appelbaum	Karl	1879	1961
	10233	Arndt	Paul	1884	1964
	10250	Arnold	Anton	1872	1945

Datensatz: 1 von 1021

Von den 2424 in der Datenbank erfassten sozialdemokratischen Parlamentariern wurden 1021 (42 %) älter als 70 Jahre.

Aufgabe 8.14: Ermitteln Sie alle Abgeordneten, die zwischen dem 70. und 80. Lebensjahr starben.

8.8.6 Wichtige Regeln zur Formulierung von Abfragen

Es gelten folgende Syntaxregeln:

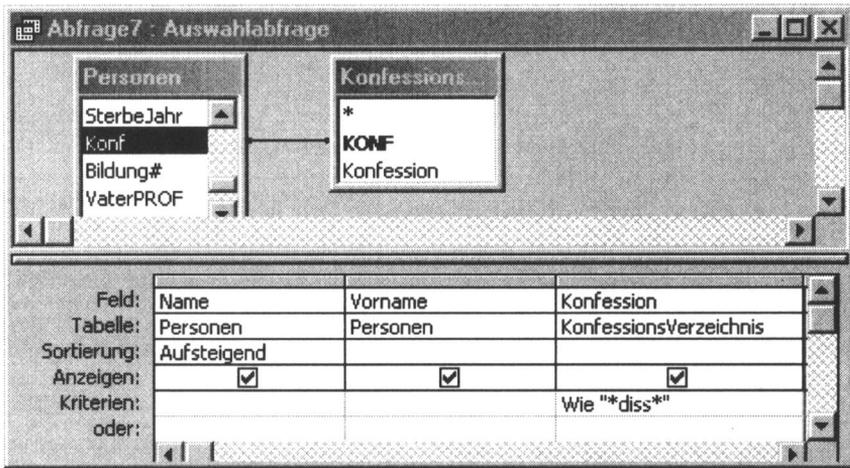
	Syntax	Beschreibung	Beispiel
Objekte	[Name]	Namen von Objekten, das sind Tabellen, Felder, Formulare, Formularelemente, Berichte usw., werden in eckige Klammern gesetzt.	[gebjahr]+70
Text	"..."	Texte (Dateninhalte) werden in Anführungszeichen gesetzt	"Bebel"

Datum	#tt.mm.jjjj#	Datumsangaben (nur vollständige mit Tag, Monat und Jahr werden akzeptiert) müssen durch das Zeichen # eingeschlossen sein. Das Jahr kann auch zweistellig eingegeben werden. Ein Wert unter 30 wird als 21. Jahrhundert interpretiert, 30 und darüber als 20. Jahrhundert	#4.6.1876# #12.07.10# wird interpretiert als 12.7.2010, #12.07.30# als 12.7.1930
-------	--------------	---	--

Der WIE-Operator hilft uns auf komfortable Weise in Textfeldern zu suchen. Bestimmte Zeichen stehen dabei als Platzhalter.

Zeichen	Bedeutung	Beispiel	
*	beliebige Anzahl von Zeichen	WIE "A*"	alle Abgeordneten, deren Name mit A beginnt
		WIE "[A-D]*"	alle Abgeordneten mit den Anfangsbuchstaben A bis D
		WIE "[AM]*"	alle Abgeordneten mit den Anfangsbuchstaben A oder M
		WIE "[!A]"	alle Buchstaben außer A
?	ein beliebiges Zeichen	WIE "???M"	alle Namen, die 4 Zeichen lang sind und deren viertes Zeichen ein M ist
		WIE "??M*"	der dritte Buchstabe eines beliebig langen Namens ist ein M
#	eine beliebige Ziffer	WIE "#*"	die Zeichenkette muss mit einer Ziffer beginnen
		WIE "####"	alle Zeichenketten, die 4 Ziffern lang sind

Beispiel: *Alle Parlamentarier, die im Feld Konfession an beliebiger Position den Eintrag 'diss' (dissidieren) haben.*



Die Tabelle Personen enthält nur den Code der Konfession. Die Ausprägungen der Konfession sind im Datenfeld Konfession der Tabelle KonfessionsVerzeichnis gespeichert. Die beiden Tabellen sind über den Schlüssel Konf miteinander verbunden. Das folgende Bild zeigt einen Ausschnitt der Ergebnismenge:

	Name	Vorname	Konfession
	Bömelburg	Theodor	kath., später diss.
	Borchardt	Julian	jüd., später diss.
	Borgmann	Hermann	ev., später diss.
	Bormann	Friedrich	ev., später diss.
	Borsch	Ludwig	kath., später diss.
	Böttger	Richard	ev., später diss., später erneut ev.
	Bracke	Wilhelm	diss.
	Brandenburg	Ernst	diss.
	Brandes	Alwin	ev., später diss.
	Braß	Otto	diss.
	Braun	Adolf	jüd., später diss.
	Braun	Otto	ev., später diss.
	Braun	Heinrich	jüd., später diss.

Aufgabe 8.15: Ermitteln Sie alle Katholiken, die dissidierten.

8.8.7 Logische Operatoren

Logische Operatoren liefern einen Wahrheitswert, von dessen Ausgang abhängt, ob die entsprechende Zeile zum Ergebnis dazugehört oder nicht. Die Operatoren UND und ODER haben wir schon angewendet. Hier soll noch der NICHT-Operator und der ZWISCHEN-Operator erwähnt werden.

	Syntax	Beschreibung
UND	Ausdruck1 UND Ausdruck2	Beide Ausdrücke müssen wahr sein
ODER	Ausdruck1 ODER Ausdruck2	Mindestens ein Ausdruck muss wahr sein
NICHT	NICHT Ausdruck	Der Wahrheitswert von Ausdruck wird umgekehrt
ZWISCHEN	ZWISCHEN Wert1 und Wert2	Der Wert des Feldes liegt zwischen den beiden Werten

Beispiel:

Alle Reichstagsmandate sozialdemokratischer Parlamentarier zwischen der 20. und 24. Legislaturperiode (einschließlich der 20. und 24.)

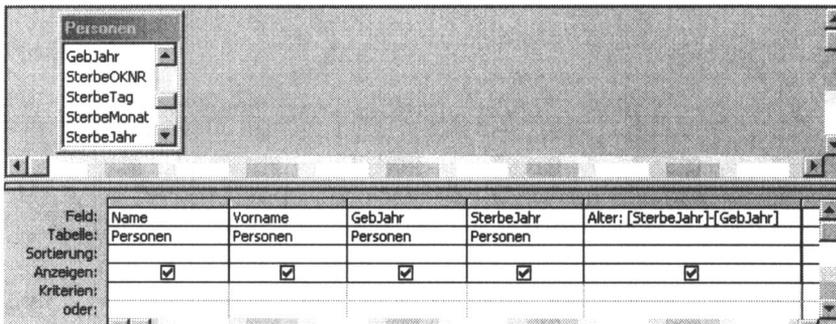
ID	LP	
MDR	MDR	
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	Zwischen 20 Und 24	

8.8.8 Berechnete Felder

Im Abfrageergebnis traten bisher Spalten der beteiligten Tabellen auf. Es ist auch möglich zusätzliche Spalten durch Rechenausdrücke zu erzeugen. Das können arithmetische Berechnungen, Zeichenkettenoperationen oder logische Verknüpfungen (die einen Wahrheitswert liefern) zwischen den Spaltenwerten und beliebigen Konstanten sein. Es können auch in ACCESS bzw. in Visual Basic for Applications (VBA) definierte Funktionen benutzt werden. Gehen Sie hierzu wie folgt vor:

- ◆ Klicken Sie in eine freie Spalte der Zeile FELD des Entwurfsbereiches.
- ◆ Geben Sie den Feldnamen ihrer Wahl ein und setzen Sie hinter den Namen einen Doppelpunkt.
- ◆ Tragen Sie die Berechnungsformel zur Ermittlung des Feldergebnisses ein. In der Berechnungsformel sind die unten aufgeführten Operatoren anwendbar. Alle benutzten Feldnamen werden in eckige Klammern gesetzt.

Beispiel: *Berechnung des Alters der Parlamentarier aus Sterbedatum und Geburtsdatum*
 Bei gegebenen Jahreszahlen gilt als angenäherter Wert für das Alter:
 Sterbejahr – Geburtsjahr



8.8.9 Mathematische Operatoren

Operator	Beschreibung	Beispiel
+	Addition von Zahlen	$[Geburtsjahr] + 70$
-	Subtraktion von Zahlen	$[Sterbejahr] - [Geburtsjahr]$
*	Multiplikation von Zahlen	$[Menge] * [Preis]$
/	Division von Zahlen	$[Erloes] / [Stundenzahl]$
^	Potenzieren einer Zahl mit dem angegebenen Exponenten	$[Laenge]^3$
MOD	Ganzzahliger Rest der Division zweier Zahlen	$51 \text{ MOD } 7$ (Ergebnis =2)

8.8.10 Anzeigeformat einstellen

Bei Divisionen entstehen häufig sehr viele Nachkommastellen, die das Ergebnis unübersichtlich machen. Daher ist es nützlich das gewünschte Ausgabeformat einzustellen:

- ◆ Klicken Sie in der Entwurfsansicht mit der rechten Maustaste in das berechnete Feld und rufen Sie den Kontextmenüpunkt EIGENSCHAFTEN auf.
- ◆ Wählen Sie im Listenfeld FORMAT ein gewünschtes Anzeigeformat aus

8.8.11 Textfelder aneinanderfügen

Mit dem Verknüpfungszeichen & können Sie Textfelder aneinanderfügen. Man kann beispielsweise den Vor- und Zunamen eines Abgeordneten in einer Spalte zusammengefasst darzustellen. Um die Lesbarkeit zu erhöhen, sollte man zwischen Name und Vorname ein Komma gefolgt von einem Leerzeichen einfügen:

[Name] & ", " & [Vorname]

ID	Fullname: [Name] & ", " & [Vorname]
Personen	
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Man erhält:

ID	Fullname
10010	Abelmann, Karl
10020	Ackermann, Friedrich
10030	Adams, Kurt
10040	Adelung, Bernhard
10050	Aderhold, Karl
10060	Agnes, Lore
10070	Agster, Alfred
10073	Ahrenholdt, Hans
...	...

8.9 Komplexe Beispiele

Beispiel: *Es sollen alle Abgeordneten des Deutschen Reichstages ermittelt werden, die vor 1900 geboren wurden und die nach 1945 noch gelebt haben.*

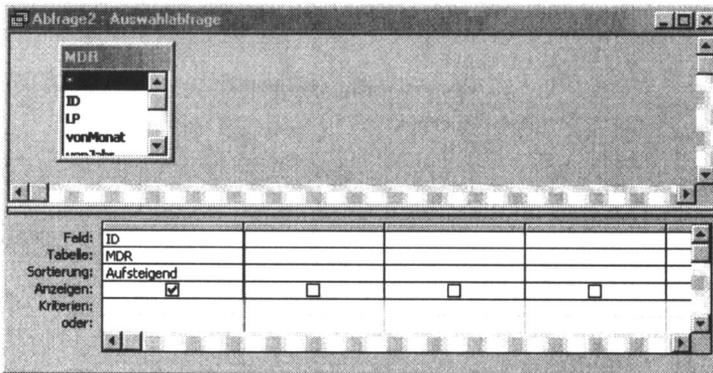
Als Auswahlkriterium gilt:

Geburtsjahr < 1900 UND Sterbejahr > 1945

Wenn für diese Abfrage nur die Personen-Tabelle benutzt wird, wie in Abschnitt 8.5.3, dann erhält man alle Abgeordneten, die dieser Bedingung entsprechen, also auch die Landtagsabgeordneten. Man muss daher die MDR-Tabelle (Reichstagsmandate) hinzunehmen. Die oben erstellte Abfrage der Mandate liefert aber die Abgeordneten mehrfach, entsprechend der Anzahl ihrer Mandate. Man muss ACCESS daher mitteilen, dass man jeden Abgeordneten des Reichstages nur einmal in der Ergebnistabelle haben möchte. Es soll von ACCESS die lokale Integrität – jede Zeile unterscheidet sich von jeder anderen – durchgesetzt werden.

Dazu werden wir im ersten Schritt eine Abfrage mit der Tabelle MDR erzeugen, die jeden Mandatsträger genau einmal enthält. Diese Abfrage kann genauso wie eine Tabelle in eine weitere Abfrage eingebunden werden. Sie wird im zweiten Schritt mit der Tabelle Personen verbunden, um den Namen, das Geburtsjahr und das Sterbejahr in die Abfrage zu integrieren.

Für den ersten Schritt darf in die Abfrage nur das Feld ID aufgenommen werden. Nimmt man weitere auf, z.B. die Nummer der Legislaturperiode, so unterscheiden sich die Zeilen für eine Person und werden durch die Forderung nach lokaler Integrität nicht unterdrückt.

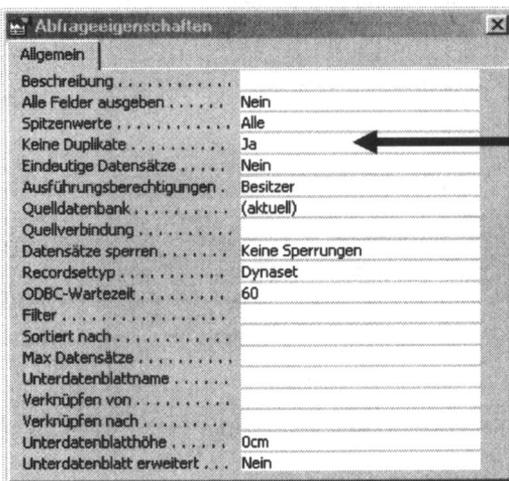


Als Ergebnis erhält man die nebenstehenden PROJECTION der Tabelle MDR.

ID
10050
10050
10060
10060
10060
10060
10060
10060
10070
.....

Im Abschnitt 8.4 wurde darauf hingewiesen, dass jedes Objekt von ACCESS Eigenschaften besitzt, die man auch verändern kann. Wir wollen jetzt für diese Abfrage die Eigenschaft einstellen, dass keine Zeile doppelt auftreten soll. Dabei gehen wir so vor:

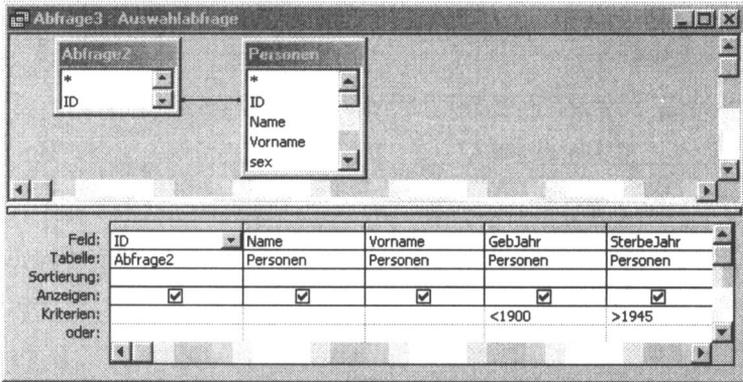
- ◆ Abfrage in der Entwurfsansicht öffnen
- ◆ Wenn das Eigenschaftsfenster bereits geöffnet ist, klicken Sie mit der linken Maustaste in den grauen Bereich des Abfragefensters. Wenn es nicht geöffnet ist, klicken Sie mit der rechten Maustaste in den grauen Bereich. Wählen Sie aus dem sich öffnenden Pulldown-Menü die Option EIGENSCHAFTEN.
- ◆ Stellen Sie im Fenster der Abfrageeigenschaften KEINE DUPLIKATE auf JA und schließen Sie das Fenster.



ID
10050
10060
10070
10100
10160
...

Jede Zeile der Abfrage tritt jetzt nur noch genau einmal auf:

Diese virtuelle Tabelle – als solche kann man eine Abfrage sehen – enthält alle Reichstagsmitglieder genau einmal. Sie stellt allerdings nur sehr wenige Informationen bereit. Daher verbinden wir sie mit der Tabelle Personen, um Namen, Vornamen, Geburtsjahr und Sterbejahr hinzuzufügen

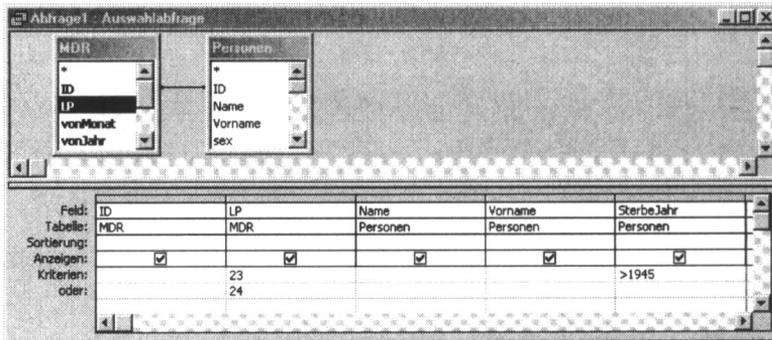


Die Bedingungen, dass die Abgeordneten vor 1900 geboren sein sollten und nach 1945 gestorben sind, wurden in der Zeile KRITERIEN formuliert. Als Ergebnistabelle erhält man:

ID	Name	Vorname	GebJahr	SterbeJahr
10060	Agnes	Lore	1876	1953
10160	Ansorge	Marie	1880	1955
10220	Arendsee	Martha	1885	1953
10240	Arning	Marie	1887	1957
10280	Arzt	Arthur	1880	1953
10330	Aufhäuser	Siegfried	1884	1969
20010	Baade	Fritz	1893	1974
20350	Becker	Heinrich	1877	1964
20370	Becker	Roman	1879	1949
20560	Bergholz	Albert	1892	1957
20570	Bergmann	Paul	1881	1951
20590	Bernhard	Nikolaus	1881	1957
20720	Biester	Louis	1882	1965
20730	Billian	Albert	1876	1954
20890	Böckler	Hans	1875	1951

Beispiel: *Alle Abgeordneten, die in den letzten beiden Legislaturperioden der Weimarer Republik im Reichstag vertreten waren (23. oder 24. Legislaturperiode) und nach 1945 starben, sollen ermittelt werden.*

Formal wird diese Bedingung so dargestellt:
 (LP = 23 ODER LP = 24) UND (SterbeJahr >1945)



Als Ergebnistabelle erhält man:

ID	LP	Name	Vorname	SterbeJahr
10060	23	Agnes	Lore	1953
10060	24	Agnes	Lore	1953
10160	23	Ansorge	Marie	1955
10160	24	Ansorge	Marie	1955
10280	23	Arzt	Arthur	1953
10280	24	Arzt	Arthur	1953
10330	23	Aufhäuser	Siegfried	1969
10330	24	Aufhäuser	Siegfried	1969
20010	23	Baade	Fritz	1974
20010	24	Baade	Fritz	1974
20350	23	Becker	Heinrich	1964
20350	24	Becker	Heinrich	1964
20560	23	Bergholz	Albert	1957
20590	24	Bernhard	Nikolaus	1957
20680	24	Biedermann	Adolf	1933
20720	23	Biester	Louis	1965

Das Ergebnis entspricht offensichtlich nicht unseren Erwartungen. Nicht die Mandatsträger mit den geforderten Merkmalen wurden aufgelistet, sondern die

Mandate! Für die Legislaturperiode 24 enthält die Ergebnistabelle Abgeordnete, die vor 1945 gestorben sind. Die Abfrage muss also anders formuliert werden, um das gewünschte Ergebnis zu erzielen.

Was haben wir falsch gemacht? Die Bedingungen werden zeilenweise geprüft, d.h., wir müssen uns eine Klammer um den logischen Ausdruck LP= 23 und SterbeJahr > 1945 vorstellen. Es wurde also ausgewertet:

(LP= 23 UND SterbeJahr > 1945) ODER LP= 24

Das sind alle Abgeordneten, die der Legislaturperiode 23 angehörten und nach 1945 starben oder diejenigen, die dem Reichstag in der 24. Legislaturperiode angehörten, also alle der 24. Periode!

Die 24. Legislaturperiode muss also auch noch bzgl. des Sterbejahres eingeschränkt werden:

(LP= 23 UND SterbeJahr > 1945) ODER (LP= 24 UND SterbeJahr > 1945)

Wer sich in formaler Logik auskennt, kann den im Beispiel formulierten Ausdruck auch sofort zum richtigen Ergebnis umformulieren.

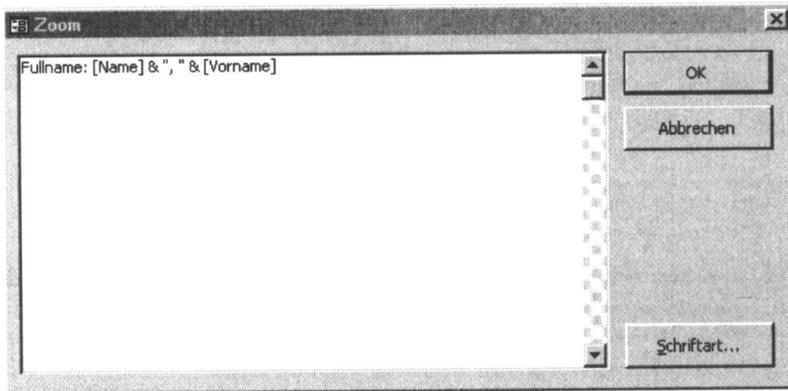
ID	LP	Name	Vorname	SterbeJahr
MDR	MDR	Personen	Personen	Personen
<input checked="" type="checkbox"/>				
	23			>1945
	24			>1945

Aufgabe 8.16: Ermitteln Sie diejenigen Abgeordneten, die für die 24. Legislaturperiode dem Reichstag angehörten und vor 1945 starben.

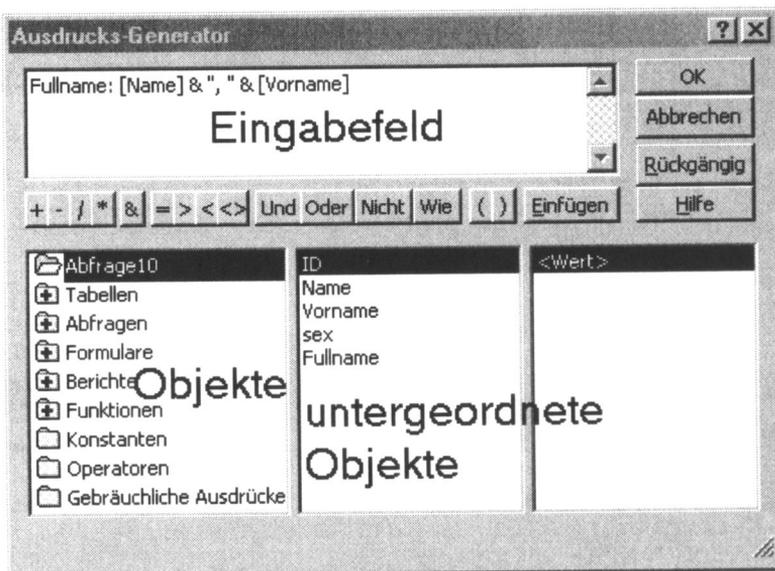
8.9.1 Der Ausdrucksgenerator

Die Eingabe von komplexen Rechenausdrücken für berechnete Felder wird in der Entwurfansicht der Abfrage schnell unübersichtlich.

ACCESS bietet zwei Möglichkeiten für eine verbesserte Eingabe: Durch die Tastenkombination [SHIFT] [F2] öffnet sich ein kleines Editorfenster, in das der Ausdruck eingegeben werden kann.



Mehr Hilfe bietet der etwas gewöhnungsbedürftige Ausdrucksgenerator. Im Ausdrucksgenerator werden Ihnen Objekte angeboten wie Tabellen, Abfragen, Formulare und in diesen definierte Elemente, die Sie per Drag & Drop in Ihren Berechnungsausdruck ziehen können. Ebenfalls bereitgestellt werden Operatoren (mathematische Operatoren, Vergleichsoperatoren, logische Operatoren) und alle in ACCESS definierten Funktionen und Konstanten.



So benutzen Sie den Ausdrucksgenerator:

- ◆ Klicken Sie mit der rechten Maustaste in die Zelle des Entwurfsbereichs, in der der Ausdruck stehen soll.
- ◆ Klicken Sie im sich öffnenden Kontextmenü auf **AUFBAUEN** .

Ein bereits in der Zelle vorhandener Ausdruck wird im Eingabefeld angezeigt. Man kann also den Ausdrucksgenerator schließen, das Ergebnis anzeigen lassen und erneut den Generator aufrufen. Damit ist der schrittweise Aufbau eines berechneten Feldes möglich.

Sie können den Ausdruck über Tasteneingabe bearbeiten oder durch Klicken mit der Maus auf die entsprechenden Buttons, bzw. Doppelklicken auf Elemente im Detailbereich (untergeordnete Objekte).

Die durch Pluszeichen gekennzeichneten Ordner enthalten Unterordner, die auf die Windows-übliche Weise durch Doppelklick geöffnet werden können.

Im Ausdrucksgenerator werden im Fenster **OBJEKTE** auch Funktionen zur Verfügung gestellt. Eine Funktion ist ein vordefinierter Rechenausdruck ('Formel'), der als Ergebnis einen Wert liefert, den man in anderen Ausdrücken, speziell auch in anderen Funktionen einsetzen kann. Einer Funktion müssen Eingangswerte, so genannte Funktionsargumente oder Parameter, mit denen sie die Berechnung ausführen soll, übergeben werden. Es gibt Funktionen, die keine Argumente, genau ein Argument oder mehrere Argumente verlangen.

Die folgende Übersicht enthält die festgelegte Schreibweise für Funktionen:

Syntax	Funktionsname(Argument1; Argument2; ...)
Funktionsname	Was berechnet werden soll, legt der Funktionsname fest. Z.B. schneidet die Funktion LGLÄTTEN(<Zeichenfolge>) die vorangestellten Leerzeichen einer Zeichenkette ab. Steht beispielsweise im Datenfeld [Vorname] die Zeichenfolge " Heinrich", so liefert die Funktion LGLÄTTEN ([Vorname]) das Ergebnis "Heinrich". Weitere Beispiele. DATUM() JAHR([Geburtsdatum]) LINKS([PLZ];3)
Argumente	Die in die Berechnung eingehenden Größen, z.B. für die Funktion LINKS (die die links stehenden Zeichen einer Zeichenfolge liefert) der Name der Zeichenfolge und wie viel Zeichen von links berücksichtigt werden sollen.

Syntaktische Regeln für die Benutzung von Funktionen:

- Dem Funktionsnamen folgt unmittelbar die öffnende Klammer.

- Die Anzahl der Argumente ist für jede Funktion genau vorgeschrieben.
- Die einzelnen Argumente werden durch Semikolon getrennt.
- Falls keine Argumente benötigt werden, muss trotzdem das Klammerpaar angegeben werden.
- Funktionen lassen sich in andere Funktionen verschachteln und in Rechenausdrücke einbinden. Das Ergebnis der Funktion wird dort eingesetzt, wo die Funktion aufgeschrieben wurde.
- Die Datentypen der Argumente müssen den Vorgaben der Funktion entsprechen. Dort, wo beispielsweise die Länge einer Zeichenkette angegeben werden soll, muss eine numerische Größe oder eine Variable oder ein Datenfeld das eine numerische Größe enthält, stehen.

Beispiel: *Bilden Sie für das Sterbealter der Abgeordneten Altersklassen:*
 Altersklasse 1: gestorben im Alter unter 50
 Altersklasse 2: gestorben im Alter von 50 bis 70
 Altersklasse 3: gestorben im Alter über 70
 Wir benutzen dazu die WENN-Funktion.
 Syntax: WENN(<Bedingung> ; <Wahr-Aktion> ; <Falsch-Aktion>)

Bedingung	wird von der Funktion ausgewertet und ergibt einen Wahrheitswert. Ist <Bedingung> wahr, so wird die 'Wahr-Aktion' ausgeführt, ist <Bedingung> falsch, wird die 'Falsch-Aktion' ausgeführt.
Wahr-Aktion	ein beliebiger Rechenausdruck oder eine Verknüpfung von
Falsch-Aktion	Zeichenketten.

Wir wollen die Altersklassen schrittweise aufbauen. Die benutzte Abfrage soll die berechnete Spalte Alter bereits enthalten, so dass wir uns auf sie beziehen können.

WENN([Alter] <50; 1 ; 2)

Dieser Ausdruck fragt, ob das Alter unter 50 liegt. Wenn ja, dann wird der Wert 1 zugewiesen, d.h. die Altersklasse soll 1 sein. Wenn das Alter 50 oder größer ist, wird 2 zugewiesen. Das stimmt noch nicht mit der Aufgabenstellung überein. An dieser Stelle muss noch differenziert werden zwischen größer als 70 Jahre und 70 oder kleiner. Das drückt die folgende Funktion aus:

WENN([Alter] >70; 3 ; 2)

Diese Funktion besagt, wenn das Alter größer als 70 ist, dann soll das Ergebnis 3 sein, bei kleiner oder gleich 70 soll der Wert 2 angenommen werden. Da Funktionen geschachtelt sein dürfen, setzen wir diesen Ausdruck in die erste Funktion an die Stelle, wo der Wert 2 steht, ein:

WENN([Alter] <50; 1 ; Wenn([Alter] >70; 3 ; 2))

Um alle die Abgeordneten, für die kein Alter ermittelt werden kann weil Geburts- oder Sterbedaten nicht bekannt sind von der Auswertung auszuschließen, wurde für Geburtsjahr und Sterbejahr gefordert, dass sie größer als Null sein sollen (Spalte Kriterium).

Field:	Name	Vorname	Geb.Jahr	Sterbe.Jahr	Alter: [Sterbejahr]-(Gebjahr)	Altersklasse: Wenn([alter]<50;1;Wenn([alter]>70;3;2))
Table:	Personen	Personen	Personen	Personen		
Sortierung:	Aufsteigend					
Anzeigen:	<input checked="" type="checkbox"/>					
Kriterium:			>0	>0		
oder:						

Die Abfrage liefert folgendes Ergebnis:

ID	Name	Vorname	GebJahr	SterbeJahr	Alter	Altersklasse
10010	Abelmann	Karl	1877	1928	51	2
10020	Ackermann	Friedrich	1876	1949	73	3
10030	Adams	Kurt	1889	1944	55	2
10040	Adelung	Bernhard	1876	1943	67	2
10050	Aderhold	Karl	1884	1921	37	1
10060	Agnes	Lore	1876	1953	77	3
10070	Agster	Alfred	1858	1904	46	1
10080	Ahrens	Heinrich	1876	1951	75	3
10085	Ahsendorf	Hermann	1881	1945	64	2
10090	Albert	Hermann	1887	1933	46	1
10095	Albertz	Hermann	1877	1945	68	2
10100	Albrecht	Adolf	1855	1930	75	3
10107	Allruth	August	1873	1922	49	1
10110	Amann	Josef	1879	1971	92	3
10113	Amelung	Friedrich	1889	1955	66	2
10120	Amlung	Georg	1896	1973	77	3
10130	Ammon	Lina	1889	1968	79	3
10140	Änderl	Franz	1883	1951	68	2
10160	Ansorge	Marie	1880	1955	75	3
10170	Anthes	Wilhelm	1877	1934	57	2
10180	Antrick	Otto	1858	1924	66	2
10190	Apel	Karl	1850	1929	79	3

Datensatz: 1 von 1951

Das Ergebnis entspricht der Aufgabenstellung.

Aufgabe 8.17: Folgende Altersklassen sind gegeben:

Altersklasse	Alter
1	unter 40
2	40 bis unter 50
3	50 bis unter 60
4	60 bis unter 70
5	70 und älter

Erstellen Sie eine Abfrage, die den Personen die entsprechende Altersklasse zuordnet.

9. Auswertung der Datenbank

9.1 Übung 1: Einfache Abfragen – Trennen von Informationen

Informationen liegen nicht immer in der gewünschten Struktur oder in einem geeigneten Format vor. Die Ursachen dafür können subjektiver oder auch objektiver Natur sein, beispielsweise mangelhafter Entwurf, quellenbedingte Informationslücken oder Übernahme fremder Datenbestände.

Datenkonvertierungen werden von ACCESS in vielen Fällen automatisch ausgeführt. Sind beispielsweise Datumsangaben im Textformat erfasst worden, die in Datumsfelder eingelesen werden oder verändert man das Format einer Spalte von Text nach Datum, so führt ACCESS automatisch eine Konvertierung durch. Daten, die kein korrektes Datum darstellen, werden nicht übernommen. Das System legt ein Fehlerprotokoll im Tabellenformat an.

ACCESS stellt eine Reihe von systemeigenen Funktionen zur Verfügung, mit denen Datenkonvertierungen durchgeführt werden können. Eine Funktion kann überall dort stehen, wo sonst ein Datenwert steht. Sie besteht aus einem Funktionsnamen, dem man gewisse Funktionsargumente oder Parameter übergibt. Die Funktion liefert als Ergebnis einen Wert zurück, der dann an der Stelle eingefügt wird, an der der Funktionsname stand. Die Anzahl der Funktionsargumente ist für jede Funktion festgelegt. Es gibt Funktionen, die kein Argument, ein oder mehrere Argumente verlangen. Die einzelnen Parameter stehen in einer festen Reihenfolge durch Semikolon getrennt.

Folgendes Beispiel soll die Syntax einer Funktion illustrieren:

Funktion	Parameter
LINKS(„nach 1945“; 4)	Die LINKS -Funktion liefert die linken Zeichen einer Zeichenkette als Ergebnis. Die Zeichenkette ist im Beispiel „nach 1945“. Wie viele Zeichen übergeben werden sollen, gibt der zweite Parameter an, hier 4. Das Ergebnis ist also 'nach'. Die Funktionsargumente müssen von runden Klammern eingeschlossen sein.

Datumsangaben treten in Quellen oft als unscharfe Informationen auf. Geburts- und Sterbedaten für Personen der jüngeren Zeitgeschichte lassen sich noch relativ genau recherchieren. Aber je tiefer man in die Vergangenheit zurück-

geht, desto häufiger findet man Jahresangaben mit Zusätzen wie ‘um’, ‘etwa’, ‘nach’ oder ‘vor’.

Solche Zeitangaben lassen sich besser im Textformat erfassen. Das hat einerseits den Vorteil, dass die gesamte Informationsmenge erfasst wird, andererseits bedeutet es aber eine ungünstige Informationsdarstellung. Will man beispielsweise das Alter berechnen, so geht das nicht direkt mit Textfeldern. Für Textfelder ist eine Rechenoperation wie etwa

$$\text{Alter} = \text{Sterbejahr} - \text{Geburtsjahr}$$

nicht definiert.

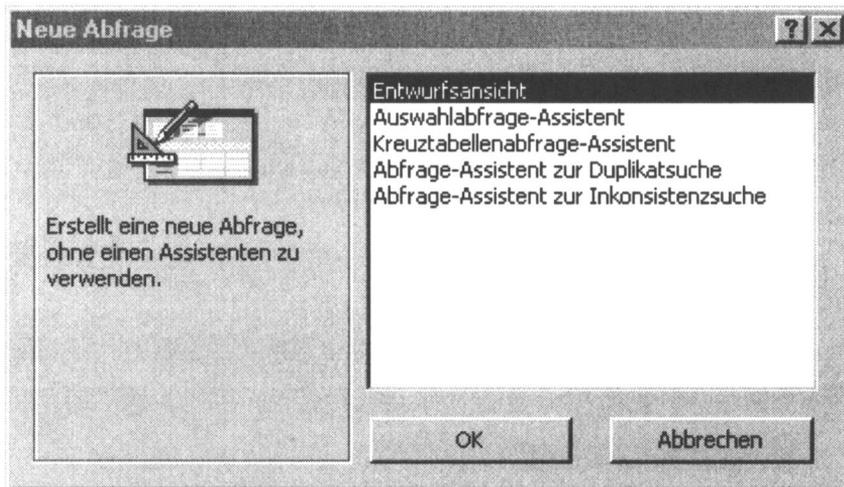
Für die Auswertung des Datenbestandes ist eine Speicherung des Datums in mehrere Felder vorteilhaft, z.B. in:

Datenfeld	Datentyp	Beispiel: „seit 26.4.1928“
Geburtstag	Integer	26
Geburtsmonat	Integer	4
Geburtsjahr	Integer	1928
Hinweis	Zeichenkette der Länge 5	seit

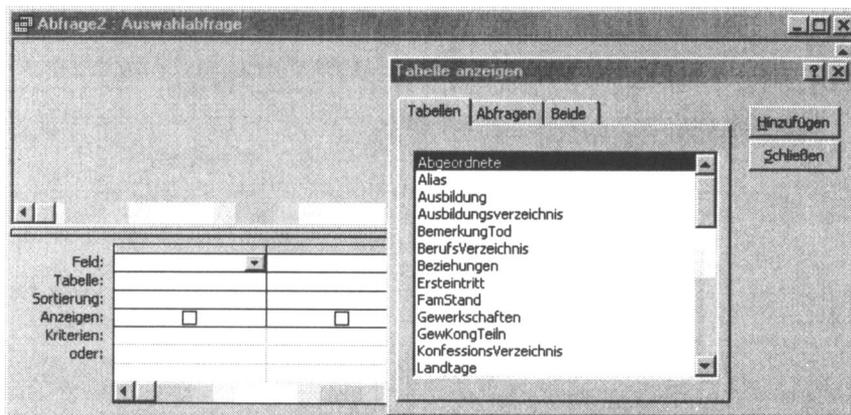
In der folgenden Übung gehen wir von der Tabelle **Abgeordnete** aus. Es handelt sich um eine Tabelle, die nur zu Übungszwecken angelegt wurde. Die Textfelder **geboren** und **gestorben** enthalten die Geburts- und Sterbedaten der Abgeordneten. Wir wollen, um besser nach Jahren abfragen zu können, die Informationen in den Feldern **geboren** und **gestorben** trennen in **Geburtstag**, **Geburtsmonat**, **Geburtsjahr** bzw. **Sterbetag**, **Sterbemonat**, **Sterbejahr**. Das soll durch berechnete Felder erreicht werden.

Wir erzeugen eine Abfrage, testen diese fehlerfrei und wandeln sie in eine **TABELLENERSTELLUNGSABFRAGE** um. Danach kann eine neue Tabelle erstellt werden, die die alte ersetzt.

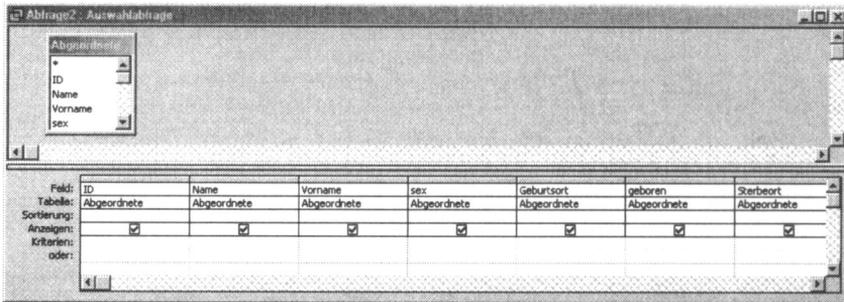
- ◆ Klicken Sie im Datenbankfenster auf die Lasche **ABFRAGE**.
- ◆ Wählen Sie **NEU**  **Neu**.
- ◆ Erstellen Sie die Abfrage in der **ENTWURFSANSICHT**.



- ◆ Wählen Sie im Fenster TABELLEN ANZEIGEN die Tabelle Abgeordnete.
- ◆ Klicken Sie auf HINZUFÜGEN.
- ◆ Klicken Sie auf SCHLIEßEN.

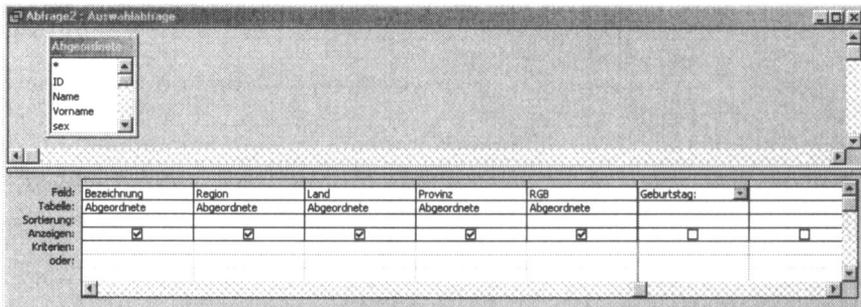


- ◆ Zur Auswahl eines Feldes Klicken Sie dieses in der FELDLISTE an und ziehen es per Drag & Drop in den Abfrageentwurfsbereich. Sie können auch alle Felder in der Liste markieren, indem Sie auf den Listenkopf doppelklicken und dann die markierten Felder in den Entwurfsbereich ziehen.



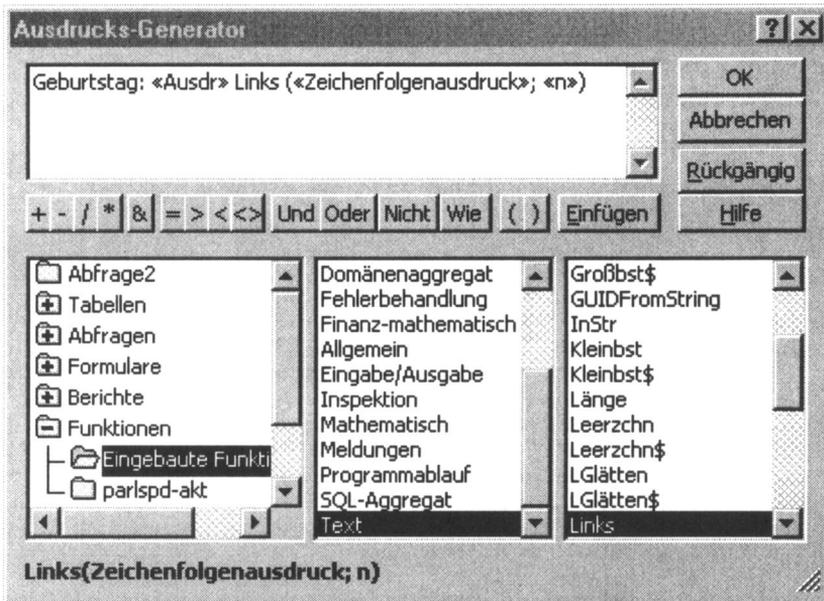
◆ Setzen Sie den Cursor in eine freie Spalte des Entwurfsbereiches in die Zeile FELD.

◆ Geben Sie einen Namen für das berechnete Feld ein: 'Geburtstag:'.

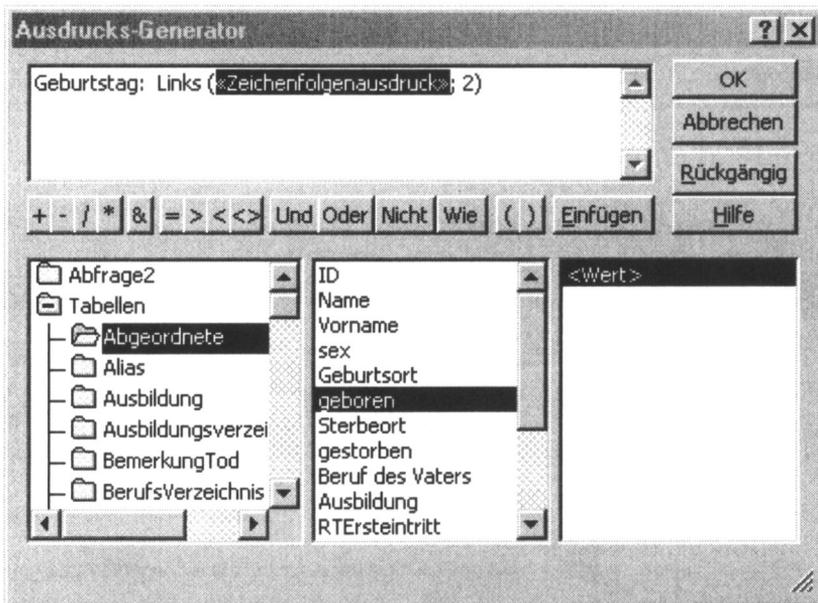


◆ Benutzen Sie den Assistenten um den Geburtstag zu ermitteln. Klicken Sie dazu mit der rechten Maustaste auf Geburtstag: und wählen Sie aus dem Kontextmenü

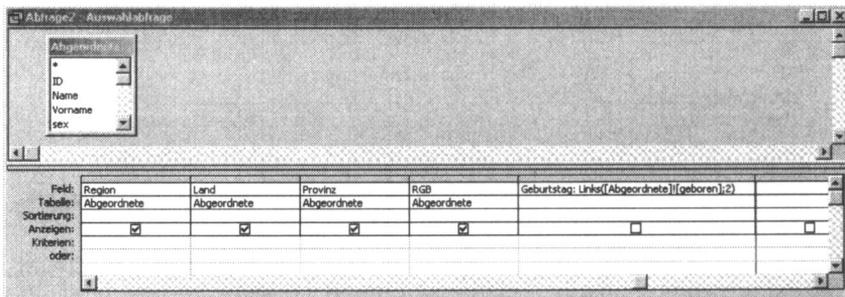
◆ Klicken Sie auf FUNKTIONEN → EINGebaUTE FUNKTIONEN → TEXT → LINKS und dann auf EINFÜGEN um die LINKS-Funktion in das Arbeitsfeld zu befördern.



- ◆ Löschen Sie «Zeichenfolgenausdruck» und lassen Sie den Cursor an dieser Stelle stehen.
- ◆ Klicken Sie auf TABELLEN → Abgeordnete → Geboren und auf EINFÜGEN, um das Tabellenfeld 'Abgeordnete' in der Funktion zu übernehmen.
- ◆ Löschen Sie «Ausdr» (dafür steht 'Geburtstag:') und ersetzen Sie «n» durch 2 (die linken zwei Zeichen des Datenfeldes Geboren enthalten den Tag).



- ◆ Klicken Sie auf OK und Sie gelangen zurück in die Entwurfsansicht der Abfrage.
- ◆ Ziehen Sie das Feld geboren neben das neu erstellte Feld Geburtstag, um das Ergebnis der Abfrage besser kontrollieren zu können.



- ◆ Gehen Sie analog vor für Geburtsmonat; die Zeichenkettenfunktion ist hier

TEIL(«Zeichenfolgenausdruck»; «Startwert»; «Länge»).
- Der Startwert ist 4 (der Monat beginnt an der 4. Position: tt.mm.jjjj), die Länge ist 2.
- Alternativ können Sie das berechnete Feld Geburtstag in die nächste freie Spalte kopieren und die entsprechenden Parameter überschreiben.
- ◆ Ziehen Sie das Feld geboren neben das neu erstellte Feld Geburtstag, um das Ergebnis der Abfrage besser kontrollieren zu können.
- ◆ Das Geburtsjahr ermitteln Sie auf die gleiche Art und Weise. Die entsprechende Funktion ist

RECHTS(«Zeichenfolgenausdruck»; «Länge»).

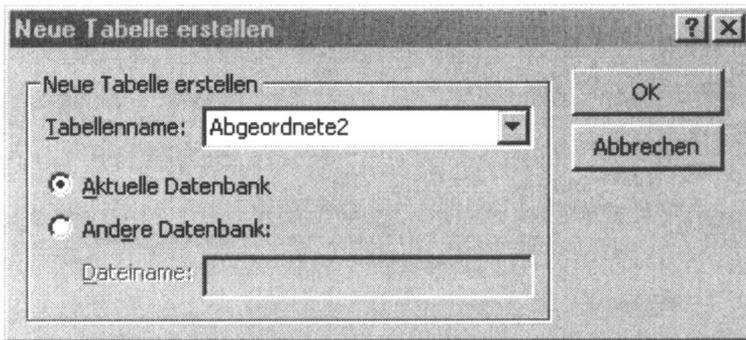
Das Ergebnis der Abfrage sieht folgendermaßen aus:

ID	Name	Vorname	sex	geboren	Geburtstag	Geburtsmonat	Geburtsjahr
10050	Aderhold	Karl	m	30.06.1884	30	06	1884
10060	Agnes	Lore	w	04.06.1876	04	06	1876
10070	Agster	Alfred	m	12.04.1858	12	04	1858
10100	Albrecht	Adolf	m	14.07.1855	14	07	1855
10160	Ansorge	Marie	w	15.12.1880	15	12	1880
10180	Antrick	Otto	m	24.11.1858	24	11	1858
▶ 10220	Arendsee	Martha	w	29.03.1885	29	03	1885
10240	Arning	Marie	w	19.04.1887	19	04	1887
10280	Arzt	Arthur	m	09.10.1880	09	10	1880
10310	Auer	Erhard	m	22.12.1874	22	12	1874
10320	Auer	Ignaz	m	19.04.1846	19	04	1846
10330	Aufhäuser	Siegfried	m	01.05.1884	01	05	1884
20010	Baade	Fritz	m	23.01.1893	23	01	1893
20040	Bader	Paul	m	13.02.1865	13	02	1865
20100	Bartels	Elise	w	13.05.1880	13	05	1880

Verfahren Sie analog mit dem Datenfeld gestorben.

Wenn Sie aus der Abfrage eine neue Tabelle erstellen wollen, die später die Tabelle Abgeordnete ersetzen soll, dann gehen Sie folgendermaßen vor:

- ◆ Rufen Sie den Menüpunkt ABFRAGE → TABELLENERSTELLUNGS-ABFRAGE auf
oder klicken Sie auf den Button  in der Symbolleiste und wählen TABELLENERSTELLUNGSABFRAGE.
- ◆ Geben Sie einen Namen für die neue Tabelle an: *Abgeordnete2* und Klicken Sie OK.



- ◆ Sie können das Ergebnis der Abfrage noch überprüfen, indem Sie sie öffnen. Die neue Tabelle wird erst erstellt, wenn Sie auf  klicken oder wenn Sie die Abfrage schließen und dann erneut öffnen.
- ◆ Beantworten Sie die Sicherheitsabfrage mit JA.
- ◆ Wenn Sie von Ihren Ergebnissen überzeugt sind, öffnen Sie die neue Tabelle *Abgeordnete2* in der Entwurfsansicht.
- ◆ Ändern Sie den Felddatentyp von Geburtstag, Geburtsmonat, Geburtsjahr, Sterbetag, Sterbemonat, Sterbejahr auf ZAHL und Feldgröße auf INTEGER

Abgeordnete2 : Tabelle

Feldname	Felddatentyp	Beschreibung
ID	Zahl	
Name	Text	
Vorname	Text	
sex	Text	
Geburtsort	Text	
geboren	Text	
Sterbeort	Text	
gestorben	Text	
Beruf des Vaters	Text	
Ausbildung	Text	
RTErsteintritt	Text	
LP	Zahl	
Bezeichnung	Text	
Region	Text	
Land	Text	
Provinz	Text	
RGB	Text	
Geburtsstag	Zahl	
Geburtsmonat	Zahl	
Geburtsjahr	Zahl	
Sterbetag	Zahl	
Sterbemonat	Zahl	
Sterbejahr	Zahl	

Feldeigenschaften

Allgemein	Nachschlagen
Feldgröße	Integer
Format	
Dezimalstellenanzeige	Automatisch
Eingabeformat	
Beschriftung	
Standardwert	0
Gültigkeitsregel	
Gültigkeitsmeldung	
Eingabe erforderlich	Nein
Indiziert	Nein

Die Feldbeschreibung ist optional. Sie hilft, den Feldinhalt zu erklären und wird auch in der Statusleiste angezeigt, wenn Sie dieses Feld in einem Formular markieren.

◆ Wenn Sie überzeugt sind, keine Fehler gemacht zu haben, löschen Sie die Datenfelder geboren und gestorben.

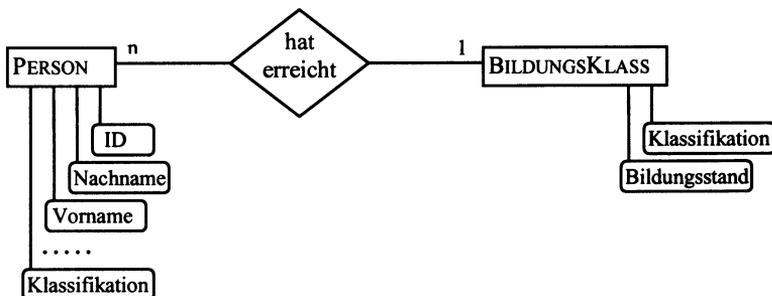
9.2 Übung 2: Verteilung des höchsten Bildungsabschlusses

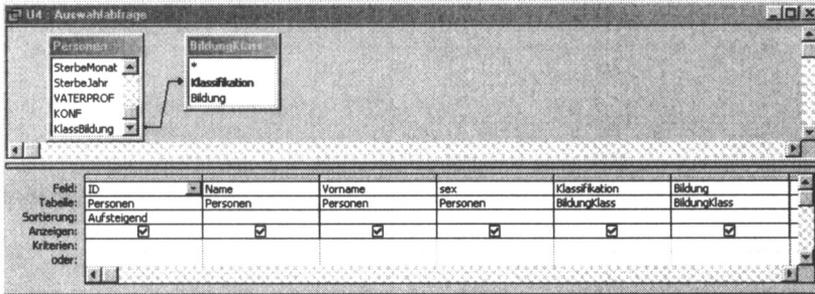
Die Tabelle Ausbildung enthält alle Bildungsmaßnahmen der Abgeordneten. Da nicht zu jeder Bildungsmaßnahme das Datum zu recherchieren war, wurde eine laufende Nummer der Bildungsmaßnahme erhoben, die die Reihenfolge der einzelnen Bildungsmaßnahmen für jeden Abgeordneten wiedergibt. Die Maßnahme mit der höchsten laufenden Nummer wäre damit jeweils die letzte und damit in der Regel die höchste Bildungsmaßnahme. Leider sind die angegebenen Bildungsmaßnahmen häufig nicht allgemein anerkannte Ausbildungsstufen. Daher wurde nach Prüfung aller Einzelmaßnahmen für jeden Abgeordneten ein höchster Bildungsabschluss nach der folgenden 9-stufigen Klassifikation vergeben.

Tabellenname: BildungKlass

Klassifikation	Bildung
0	k. A.
1	Volksschule
2	Mittelschule ohne Abschluss
3	Mittelschule mit Abschluss / "Tertia"-Abschluss
4	Höhere Schule ohne "Einjähriges"
5	Höhere Schule mit "Einjährigem" (Obersekundareife)
6	Höhere Schule mit Abitur
7	Lehrerseminar mit Abschluss
8	Universität ohne Abschluss
9	Universität mit Abschluss

Da für jeden Abgeordneten genau eine Angabe zum Bildungsabschluss vorliegt, wurde die Klassifikationsnummer in die Tabelle Personen übernommen. Folgendes ER-Diagramm liegt dem Sachverhalt zu Grunde:





Die Abfrage lieferte für jeden Abgeordneten folgende Informationen:

ID	Name	Vorname	sex	Klassifikation	Bildung
10010	Abelmann	Karl	m	1	Volksschule
10020	Ackermann	Friedrich	m	9	Universität mit Abschluss
10030	Adams	Kurt	m	9	Universität mit Abschluss
10040	Adelung	Bernhard	m	1	Volksschule
10050	Aderhold	Karl	m	1	Volksschule
10060	Agnes	Lore	w	1	Volksschule
10070	Agster	Alfred	m	8	Universität ohne Abschluss
10073	Ahrenholdt	Hans	m	0	k. A.
10075	Ahrens	August	m	0	k. A.
10080	Ahrens	Heinrich	m	1	Volksschule
10085	Ahsendorf	Hermann	m	0	k. A.
10090	Albert	Hermann	m	1	Volksschule
...

Uns interessieren jedoch nicht die Individualdaten, sondern die Verteilung des Bildungsniveaus: Wie viele sozialdemokratische Abgeordnete hatten welchen Schulabschluss erreicht? D.h. es muss nach den einzelnen Stufen des Bildungsabschlusses gruppiert werden.

Für solche Aufgaben stellt ACCESS eine Reihe von Aggregatfunktionen zur Verfügung. Bei einer Abfrage in Entwurfsansicht kann durch Klicken auf das Icon  der Symbolleiste die Zeile mit den Funktionen eingblendet werden.

Für alle Spalten wird GRUPPIEREN voreingestellt. GRUPPIEREN bedeutet, dass alle Werte der Zeilen, über die gruppiert wird, nur einmal in die Ergebnismenge übernommen werden.

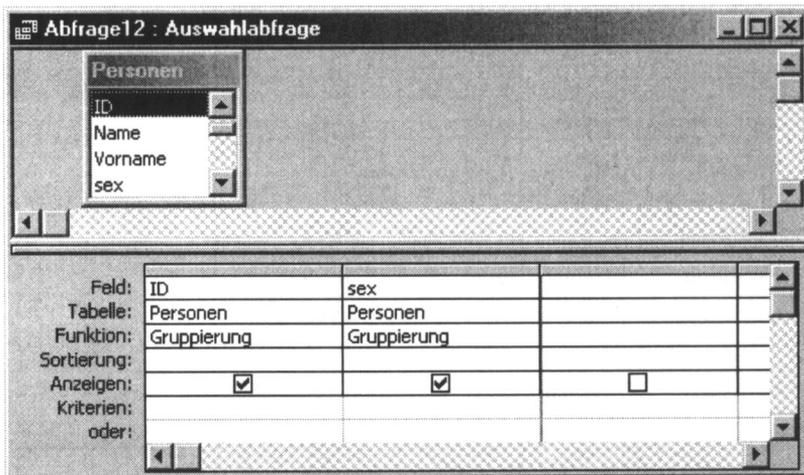
Es stehen die in der nebenstehende Abbildung Aggregat-Funktionen zur Verfügung:

Sie erreichen die Auswahl, indem Sie in einer Spalte in der Zeile FUNKTIONEN auf das Listenfeld klicken.

- Gruppierung
- Summe
- Mittelwert
- Min
- Max
- Anzahl
- Varianz
- ErsterWert
- LetzterWert
- Ausdruck
- Bedingung

Beispiel: *Gruppieren über mehrere Merkmale*

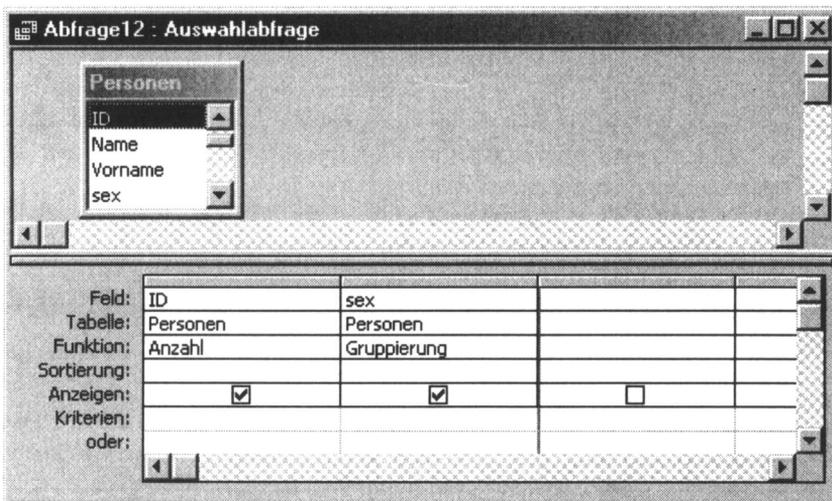
Zwei Datenfelder sind in die Abfrage übernommen wurden: ID und sex. Im ersten Falle wird über diese beiden gruppiert:



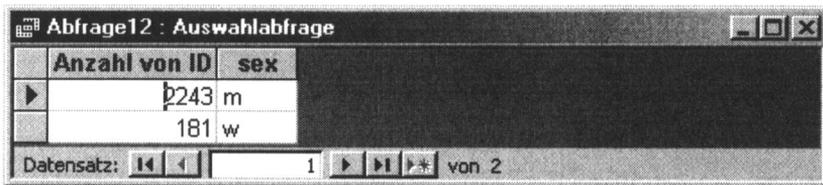
Die Ergebnistabelle enthält alle Abgeordneten, da ID als Schlüssel in der gesamten Spalte keine gleichen Werte enthält und damit kein Wertepaar ID, sex existiert, das gleiche Werte besitzt, um zu einer Gruppe zusammengefasst zu werden.

Wenn man nur nach sex gruppiert und für ID die Anzahl-Funktion wählt, wird innerhalb jeder Gruppe die Anzahl der zusammengefassten Zeilen ermittelt.

ID	sex
10010	m
10020	m
10030	m
10040	m
10050	m
10060	w
...	..



Von den 2424 Abgeordneten sind 2243 Männer und 181 Frauen.



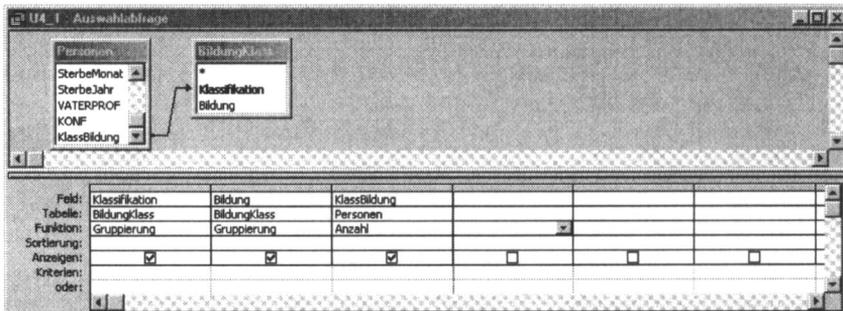
Sie müssen folgendes beachten: Wenn Sie über eine oder mehrere Spalten gruppieren, dann muss auch für die anderen Spalten eine Aggregatfunktion gewählt werden, die für die Gruppen jeweils genau einen Wert liefert.

Beispiel: *Fehlende Aggregatfunktion*

Feld:	ID	sex	
Tabelle:	Personen	Personen	
Funktion:		Gruppierung	
Sortierung:			
Anzeigen:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

In der Spalte ID würden für die beiden Gruppen 'm' und 'w' von sex mehrere Werte anfallen. ACCESS verlangt aber je Gruppe genau einen Wert. Es wird daher ein Fehlerhinweis eingeblendet.

Wir wollen Gruppen bilden nach dem Bildungsniveau der Abgeordneten. Also muss diese Spalte in die Abfrage integriert werden. Weiterhin interessiert uns die Anzahl der Fälle je Gruppe. Dazu müssen wir diese Spalte ein zweites Mal in die Abfrage einbeziehen, aber mit der Funktion ANZAHL. Das Ergebnis ist noch nicht ganz befriedigend. Wir haben nur den Code des Bildungsabschlusses, wir möchten den Abschluss natürlich auch in Textform angezeigt bekommen. Dazu nehmen wir die Spalte Bildung hinzu.



Das Ergebnis der Abfrage zeigt die folgende Tabelle:

Klassifikation	Bildung	Anzahl von KlassBildung
0	k. A.	601
1	Volksschule	1297
2	Mittelschule ohne Abschluss	97
3	Mittelschule mit Abschluss / "Tertia"-Abschluss	53
4	Höhere Schule ohne "Einjähriges"	41
5	Höhere Schule mit "Einjährigem" (Obersekundareife)	27
6	Höhere Schule mit Abitur	11
7	Lehrerseminar mit Abschluss	104
8	Universität ohne Abschluss	44
9	Universität mit Abschluss	149

Versuchen Sie bitte als Übung die Abfrage so zu verändern, dass Sie als Ergebnis die folgende Tabelle erhalten (Prozentangaben, Sortierung nach der Häufigkeit):

Klassifikation	Bildung	Anzahl	Prozent
1	Volksschule	1297	54
0	k. A.	601	25
9	Universität mit Abschluss	149	6
7	Lehrerseminar mit Abschluss	104	4
2	Mittelschule ohne Abschluss	97	4
3	Mittelschule mit Abschluss / "Tertia"-Abschluss	53	2
8	Universität ohne Abschluss	44	2
4	Höhere Schule ohne "Einjähriges"	41	2
5	Höhere Schule mit "Einjährigem" (Obersekundareife)	27	1
6	Höhere Schule mit Abitur	11	0

Lösung:

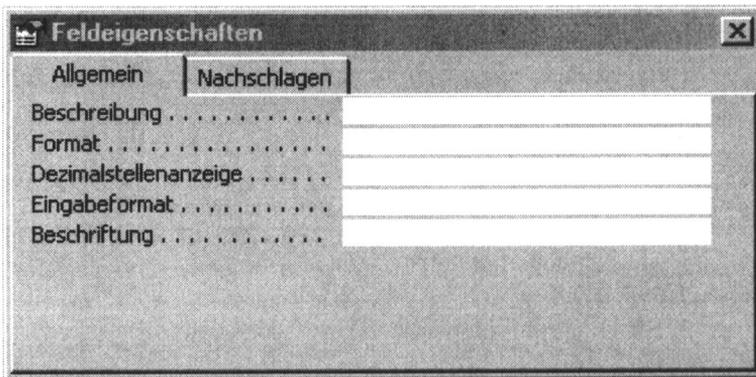
Wie man der Tabelle Personen entnehmen kann, beträgt die Anzahl der Parlamentarier insgesamt 2424. Es ist eine berechnete Spalte zu bilden, wir wollen sie Prozent nennen, die aus der Spalte Anzahl den Prozentsatz berechnet:

$$([\text{Anzahl}] * 100) / 2424$$

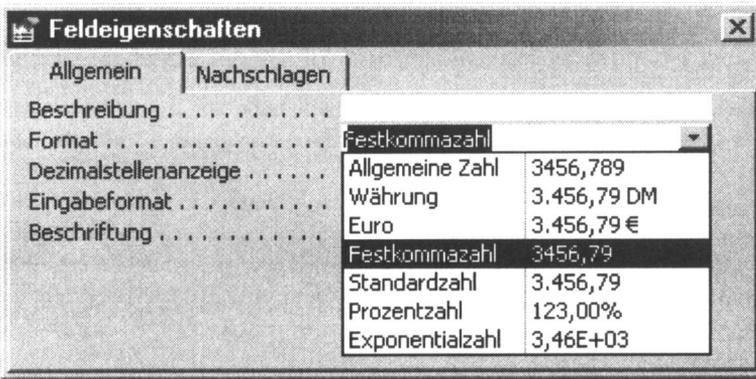
Die Nachkommastellen sollen gerundet werden. ACCESS stellt leider keine Funktion bereit, die eine Zahl rundet, sondern nur die INT-Funktion, die die Nachkommastellen abschneidet. Wenn man für weitere Berechnungen eine gerundete Zahl benötigt, muss man sich selbst einen Rechenausdruck herleiten, der unter Benutzung der INT-Funktion eine Zahl rundet.

Ein anderer Weg zu gerundeten Werten zu gelangen ist die Ausnutzung der Eigenschaften des berechneten Datenfeldes Prozent. Das Ausgabeformat gehört zu den einstellbaren Eigenschaften dieser Objektklasse. Die Anzahl der gewünschten Nachkommastellen kann angegeben werden. Bei dieser Formatierung werden die auszugebenden Werte gerundet.

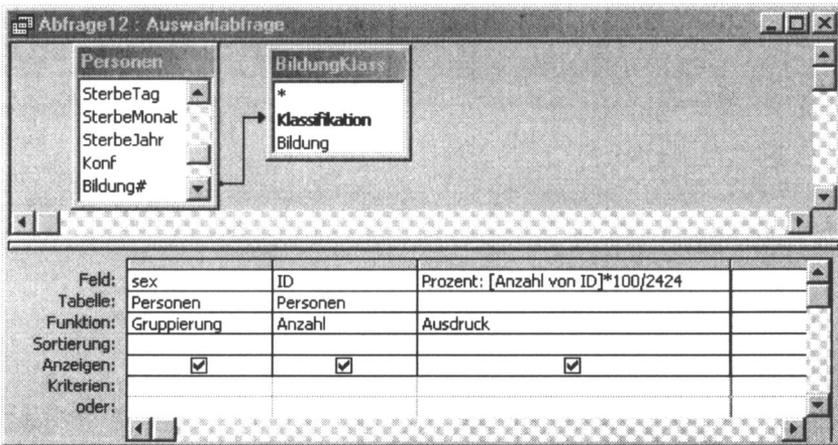
Wir klicken in der Entwurfsansicht mit der rechten Maustaste auf das Datenfeld Prozent. Es öffnet sich folgendes Kontextmenü: Wir wählen EIGENSCHAFTEN. Es öffnet sich das Fenster FELDEIGENSCHAFTEN.



Wir setzen den Cursor in das Feld FORMAT, klicken auf die Lasche und wählen FESTKOMMAZAHL und geben für DEZIMALSTELLENANZEIGE 0 ein (keine Nachkommastellen).



In der Zeile FUNKTION muss eine geeignete Funktion ausgewählt werden. In der zu erzeugenden Spalte Prozent soll nur gerechnet werden, wobei die in die Rechenoperation eingehende Zahl der Spalte Anzahl zu entnehmen ist. Dafür ist im Listenfeld AUSDRUCK zu wählen.



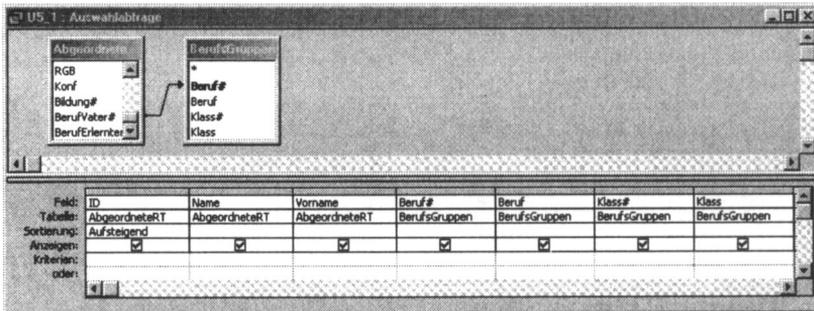
Die Aussagen lassen sich vertiefen, wenn Sie Ihre Untersuchungen nach weiblichen und männlichen Abgeordneten trennen oder zeitliche Schnitte vornehmen (Kaiserreich, Weimarer Republik oder Ersteintritt in den einzelnen Legislaturperioden).

9.3 Übung 3: Soziale Herkunft der Abgeordneten

Als Indikator für soziale Herkunft gilt im Allgemeinen der Beruf des Vaters. Die Berufe werden klassifiziert um die Probanden gruppieren zu können. Die Gruppierung erfolgt analog zu Übung 2.

Es sollen nur die Abgeordneten des Reichstages berücksichtigt werden. Deren personenbezogenen Daten sind bereits von der Tabelle Personen in die Tabelle AbgeordneteRT übertragen worden, so dass wir diese Tabelle benutzen können. Wie Sie aus der Tabelle aller Abgeordneten (Personen) speziell die Reichstagsabgeordneten ermitteln können, wurde im Kapitel 8, Beispiel 8, demonstriert.

Das folgende Bild zeigt eine Abfrage in der Entwurfsansicht, die zu jeder Person den Vaterberuf und die Klassifizierung liefert.



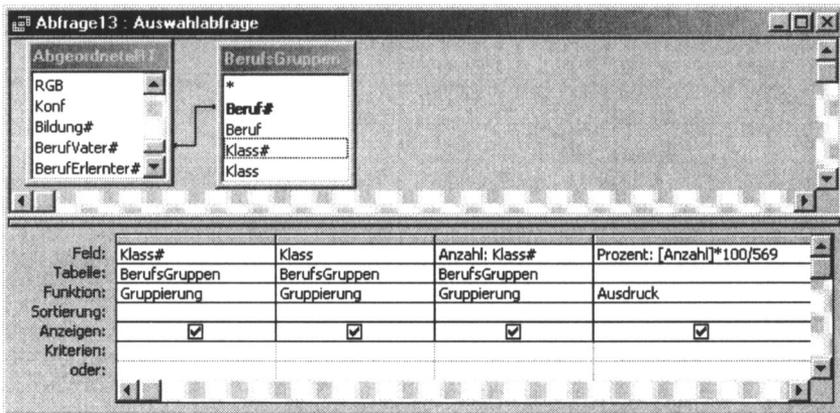
Diese Abfrage erzeugt die nach ID sortierte Tabelle (Ausschnitt):

ID	Name	Vorname	Beruf#	Beruf	Klas#	Klass
10050	Aderhold	Karl	514	Bergbauindustriearbeiter	1	ungelehrte Arbeiter (unselbständig)
10060	Agnes	Lore	514	Bergbauindustriearbeiter	1	ungelehrte Arbeiter (unselbständig)
10070	Agster	Alfred	9999	keine Angabe	99	k. A.
10100	Albrecht	Adolf	925	Schneidermeister	3	Selbständige
10160	Ansorge	Marie	508	Maurer, Stuckateur	2	gelernte Arbeiter (unselbständig)
10180	Antrick	Otto	90	freie Berufe	4	bürgerliche Berufe
...

Da auch hier nicht die Individualdaten interessieren, wollen wir die Abfrage so verändern, dass die Anzahl und die prozentuale Häufigkeit des Auftretens der einzelnen Klassifikationsstufen angezeigt werden.

Der Tabelle AbgeordneteRT entnehmen wir, dass im Datenbestand 569 sozialdemokratische Reichstagsabgeordnete erfasst wurden. Die Spalte Klass enthält die Klassifikation der Berufe, die Spalte Klass# die Kodierung. Über diese Spalten wird gruppiert und über eine dieser beiden muss die Anzahl je Gruppe ermittelt werden. Die Berechnung der Spalte Prozent gestattet folgender Ausdruck:

$$[\text{Anzahl}] * 100 / 569$$



Klass#	Klass	Anzahl	Prozent
1	ungelernte Arbeiter (unselbständig)	81	14
2	gelernte Arbeiter (unselbständig)	125	22
3	Selbständige	142	25
4	bürgerliche Berufe	44	8
5	politische Beamte	2	0
6	Arbeiterbeamte	3	1
8	untere/mittlere Angestellte/Beamte	36	6
99	k. A.	136	24

Über den Menüpunkt EIGENSCHAFTEN wurde für die Spalte ‚Prozent‘ FESTKOMMAZAHL mit 0 NACHKOMMASTELLEN eingetragen.

9.4 Übung 4: Ausgeübte Berufe vor Mandatsantritt

Wir beschränken uns in dieser Übung wieder auf die Reichstagsabgeordneten. Es werden die Tabellen `AbgeordneteRTBerEM` (Abgeordnete des Reichstages mit der Legislaturperiode des Erstmandats und dem Jahr des Eintritts in den Reichstag) und `BerufsGruppenMA` (Berufsgruppen bei Mandatsantritt) benutzt. Die beiden Tabellen wurden vorab aus anderen Tabellen durch Abfragen erstellt. Sie haben folgende Tabellenstruktur:

`AbgeordneteRTBerEM`

Feldname	Felddatentyp	Hinweis
ID	Zahl	ID des Abgeordneten
Periode	Zahl	Legislaturperiode des Erstmandats für den RT
Jahr	Zahl	Jahr des Ersteintritts in den Reichstag
Beruf	Zahl	Code des Berufs bei Ersteintritt

`BerufsGruppenMA`

Feldname	Felddatentyp	Hinweis
Beruf#	Zahl	Berufs-Code
Beruf	Text	Beruf
Klass#	Zahl	Klassifikations-Code
Klass	Text	Klassifikation

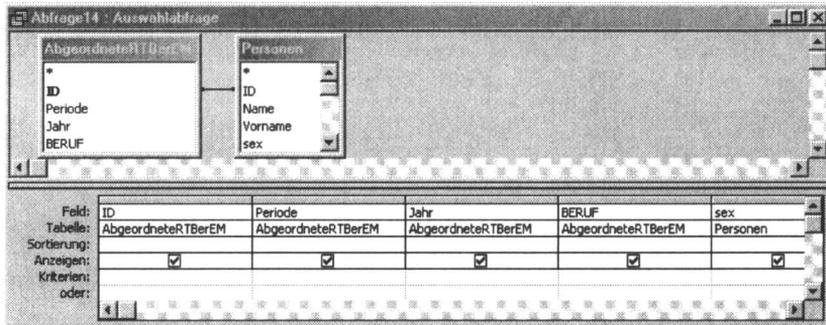
Wir wollen mit Hilfe einer Abfrage

1. für alle Reichstagsabgeordneten insgesamt,
2. für die männlichen und
3. für weiblichen Abgeordneten

Tabellen erstellen, die die Berufsklassen, die Anzahl und die prozentuale Verteilung enthalten.

Bestätigt sich unsere Erwartung, dass der Anteil an Arbeiterbeamten sehr hoch ist?

Da die Tabelle `AbgeordneteRTBerEM` das Geschlecht der Probanden nicht enthält, muss zuerst eine 'virtuelle Tabelle' – eine Abfrage – erzeugt werden, die die Variable `sex` bereitstellt. Wir verbinden die Tabelle `AbgeordneteRTBerEM` mit der Tabelle `Personen`.



Diese Abfrage, die wir qbUE4-1 nennen wollen, benutzen wir wie in den vorangegangenen Übungen, jedoch vorerst ohne die Spalte sex zu berücksichtigen. Bei der Prozentberechnung setzen wir für die Gesamtanzahl der Mdr 559 ein.

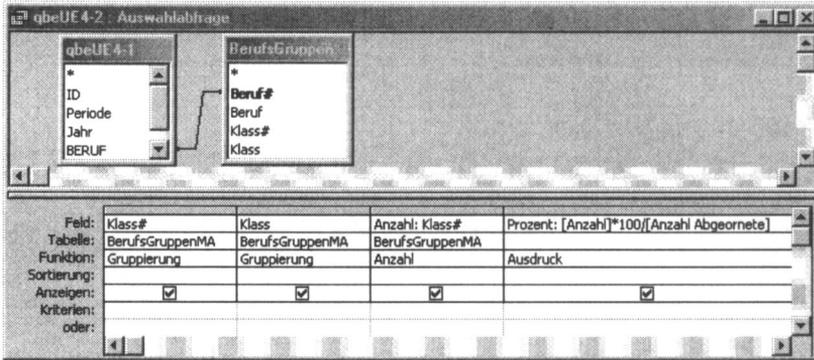
Die Anzahl der männlichen und weiblichen Abgeordneten kann man auf verschiedene Weise ermitteln. Erstellen Sie eine neue Abfrage unter Benutzung der Abfrage qbUE4-1, indem Sie nur die Spalte sex in die Abfrage einbinden und dann gruppieren.

Sie erhalten die nebenstehende Tabelle. Es sind 510 männliche und 49 weibliche Abgeordnete, insgesamt also 559 Reichstagsabgeordnete.

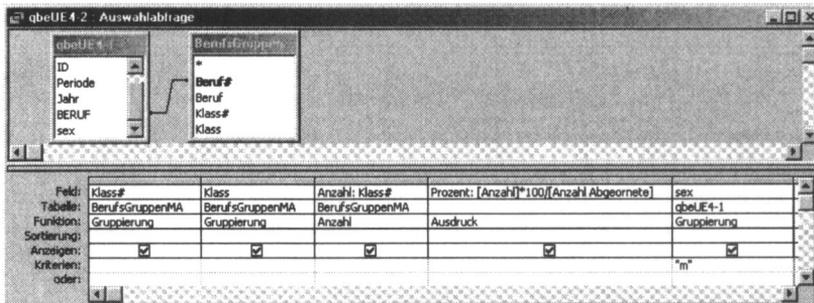
sex	Anzahl von sex
m	510
w	49

Wir entwerfen die Abfrage nach den prozentualen Anteilen wie im vorangegangenen Beispiel, geben aber bei der Prozentberechnung nicht die Anzahl der Abgeordneten ein, sondern den Text „Anzahl Abgeordnete“. ACCESS nimmt an, es handelt sich um einen Spaltennamen. Wird bei der Abarbeitung der Abfrage diese Spalte nicht gefunden, dann zeigt das Programm keinen Fehler an, sondern erfragt einen Wert, der dann in die Rechnung einbezogen wird. Wir geben 559 ein, die Anzahl der Abgeordneten insgesamt.

Man nennt diese Art der Abfrage Parameterabfrage, da man einen Wert – einen Parameter – für einen fehlenden Wert eingeben kann.



Die Ergebnistabellen sind am Ende dieser Übung zusammengefasst dargestellt. Wir verändern die Abfrage, indem wir das Geschlecht (Datenfeld sex) mit einbeziehen und für sex in die Zeile Kriterium 'm' und dann 'w' eintragen, jeweils mit den entsprechenden Werten für 'Anzahl Abgeordnete'.



Man kann zusätzlich eine zeitliche Differenzierung vornehmen, indem man einmal nur die Abgeordneten betrachtet, die ihr Erstmandat im Kaiserreich erhielten und diejenigen die in der Weimarer Republik erstmals in den Reichstag einzogen.

Nur die Abgeordneten des Kaiserreiches erhält man, wenn für das Jahr des Ersteintritts das Kriterium 'vor 1918' angegeben wird. Diese Abfrage soll qbeUE4-2 genannt werden.

ID	Periode	Jahr	BERUF	sex
AbgeordneteRTBerl	AbgeordneteRTBerl	AbgeordneteRTBerl	AbgeordneteRTBerl	Personen
Gruppierung	Gruppierung	Gruppierung	Gruppierung	Gruppierung
<input checked="" type="checkbox"/>				
		<1918		

Das sind 213 Abgeordnete.

In analoger Weise erzeugt man eine Abfrage, die alle Abgeordneten der Weimarer Republik liefert ('>1918'). Das sind 346 Abgeordnete.

9.4.1 Ergebnistabellen

Interpretieren Sie die Ergebnisse selbständig.

Alle Abgeordneten

Klass#	Klass	Anzahl	Prozent
1	ungelernte Arbeiter (unselbständig)	4	1
2	gelernte Arbeiter (unselbständig)	14	3
3	Selbständige	60	11
4	bürgerliche Berufe	45	8
5	politische Beamte	32	6
7	ohne Beruf	26	5
8	untere/mittlere Angestellte/Beamte	10	2
60	Partei	76	14
61	Publizistik	165	30
62	Gewerkschaft	103	18
63	Genossenschaft	4	1
64	Sonstige	2	0
65	Krankenkasse	11	2
99	k. A.	7	1

Alle männlichen Abgeordneten

Klass#	sex	Klass	Anzahl	Prozent
1	m	ungelernte Arbeiter (unselbständig)	3	1
2	m	gelernte Arbeiter (unselbständig)	12	2
3	m	Selbständige	59	11
4	m	bürgerliche Berufe	42	8
5	m	politische Beamte	31	6
7	m	ohne Beruf	3	1
8	m	untere/mittlere Angestellte/Beamte	9	2
60	m	Partei	71	13
61	m	Publizistik	157	28
62	m	Gewerkschaft	100	18
63	m	Genossenschaft	4	1
64	m	Sonstige	2	0
65	m	Krankenkasse	11	2
99	m	k. A.	6	1

Alle weiblichen Abgeordneten

Klass#	sex	Klass	Anzahl	Prozent
1	w	ungelernte Arbeiter (unselbständig)	1	2
2	w	gelernte Arbeiter (unselbständig)	2	4
3	w	Selbständige	1	2
4	w	bürgerliche Berufe	3	6
5	w	politische Beamte	1	2
7	w	ohne Beruf	23	47
8	w	untere/mittlere Angestellte/Beamte	1	2
60	w	Partei	5	10
61	w	Publizistik	8	16
62	w	Gewerkschaft	3	6
99	w	k. A.	1	2

Alle Abgeordnete des Kaiserreiches

Klass#	Klass	Anzahl	Prozent
1	ungelernte Arbeiter (unselbständig)	1	0
2	gelernte Arbeiter (unselbständig)	4	2
3	Selbständige	48	22
4	bürgerliche Berufe	15	7
5	politische Beamte	2	1
7	ohne Beruf	2	1
8	untere/mittlere Angestellte/Beamte	1	0
60	Partei	16	7
61	Publizistik	90	42
62	Gewerkschaft	29	13
63	Genossenschaft	2	1
65	Krankenkasse	6	3

Alle Abgeordneten der Weimarer Republik

Klass#	Klass	Anzahl	Prozent
1	ungelernte Arbeiter (unselbständig)	3	1
2	gelernte Arbeiter (unselbständig)	10	3
3	Selbständige	12	3
4	bürgerliche Berufe	30	9
5	politische Beamte	30	9
7	ohne Beruf	24	7
8	untere/mittlere Angestellte/Beamte	9	3
60	Partei	60	17
61	Publizistik	75	22
62	Gewerkschaft	74	22
63	Genossenschaft	2	1
64	Sonstige	2	1
65	Krankenkasse	5	1
99	k. A.	7	2

Männliche Abgeordnete des Kaiserreichs

Klass#	Klass	Anzahl	Prozent
1	Ungelernte Arbeiter (unselbständig)	1	0
2	gelernte Arbeiter (unselbständig)	4	2
3	Selbständige	48	22
4	Bürgerliche Berufe	15	7
5	politische Beamte	2	1
7	ohne Beruf	2	1
8	untere/mittlere Angestellte/Beamte	1	0
60	Partei	16	7
61	Publizistik	90	42
62	Gewerkschaft	29	13
63	Genossenschaft	2	1
65	Krankenkasse	6	3

Männliche Abgeordnete der Weimarer Republik

Klass#	Klass	Anzahl	Prozent
1	Ungelernte Arbeiter (unselbständig)	2	1
2	gelernte Arbeiter (unselbständig)	8	3
3	Selbständige	11	4
4	Bürgerliche Berufe	27	9
5	politische Beamte	29	10
7	ohne Beruf	1	0
8	untere/mittlere Angestellte/Beamte	8	3
60	Partei	55	19
61	Publizistik	67	23
62	Gewerkschaft	71	24
63	Genossenschaft	2	1
64	Sonstige	2	1
65	Krankenkasse	5	2
99	k. A.	6	2

Weibliche Abgeordnete der Weimarer Republik

Klass#	Klass	Anzahl	Prozent
1	Ungelernte Arbeiter (unselbständig)	1	2
2	gelernte Arbeiter (unselbständig)	2	4
3	Selbständige	1	2
4	Bürgerliche Berufe	3	6
5	politische Beamte	1	2
7	Ohne Beruf	23	47
8	Untere/mittlere Angestellte/Beamte	1	2
60	Partei	5	10
61	Publizistik	8	16
62	Gewerkschaft	3	6
99	k. A.	1	2

9.5 Übung 5: Mandatsdauer in Abhängigkeit vom Bildungsniveau und von der sozialen Herkunft

Zuerst ermitteln wir die Mandatsdauer für Abgeordnete des Reichstags. Dazu gehen wir von der Tabelle Reichstagsmandate (MdR) aus. Sie enthält alle Reichstagsmandate der Abgeordneten mit den folgenden Datenfeldern:

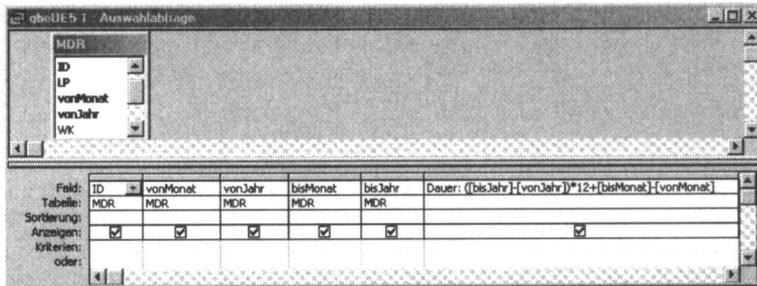
MdR

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	ID des Abgeordneten
	LP	Zahl	Reichstag-Legislaturperiode
	vonMonat	Zahl	Monat des Eintritts in den Reichstag
	vonJahr	Zahl	Jahr des Eintritts in den Reichstag
	WK	Zahl	Nummer des Wahlkreises
	bisMonat	Zahl	Monat der Mandatsniederlegung in dieser Legislaturperiode
	bisJahr	Zahl	Jahr der Mandatsniederlegung in dieser Legislaturperiode
	Bemerkung	Text	Bemerkung zum Mandat

Aus diesen Angaben lässt sich für jedes Mandat die Verweildauer im Reichstag feststellen. Folgender Rechenausdruck ermittelt aus zwei Datumsangaben die Dauer in Monaten:

$$\text{Dauer [in Monaten]} = 12 * (\text{Endjahr} - \text{Anfangsjahr}) + (\text{Endmonat} - \text{Anfangsmonat}) + 1$$

Wir erzeugen folgende Abfrage:



Die Abfrage qbeUe5-1 liefert die Mandatsdauer für jedes Mandat.

ID	vonMonat	vonJahr	bisMonat	bisJahr	Dauer
10050	3	1919	6	1920	16
10050	6	1920	6	1921	13
10060	1	1919	6	1920	18
10060	6	1920	5	1924	48
10060	5	1924	12	1924	8
10060	12	1924	5	1928	42
10060	5	1928	9	1930	29
10060	9	1930	7	1932	23
10060	7	1932	11	1932	5
10060	11	1932	3	1933	5
...

Diese Abfrage enthält noch Mandate. Um nicht Mandate, sondern Abgeordnete als Betrachtungsebene zu erhalten, muss man die Mandatsdauer je Abgeordneter summieren. Das ergibt für jeden Reichstagsabgeordneten genau einen Zeile.

ID	Dauer: (([bisJahr]-[vonJahr])*12+[bisMonat]-[vonMonat])
MDR	
Gruppierung	Summe
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Folgende Schritte sind durchzuführen:

1. Gruppieren nach ID liefert für jeden Abgeordneten eine Zeile,
2. Summieren nach Dauer innerhalb der Gruppe ergibt die Gesamtdauer je Abgeordneter.

Alle anderen Datenfelder wurden aus der Abfrage herausgenommen, da sie für die Gruppenbildung nicht von Relevanz sind. Die nebenstehende Tabelle enthält die Verweildauer je Abgeordneter in Monaten.

ID	Dauer
10050	29
10060	182
10070	61
10100	278
10160	156
....

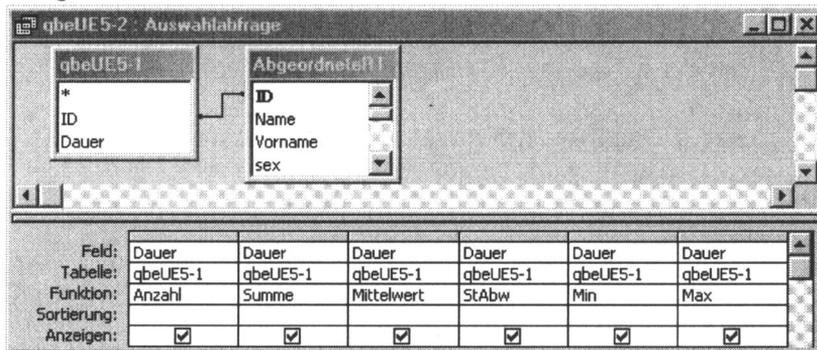
Für die Abgeordneten insgesamt und differenziert nach männlichen und weiblichen Abgeordneten wollen wir die statistischen Grundgrößen

- Summe,
- Anzahl,
- Mittelwert,
- Standardabweichung,
- kleinster Wert,
- größter Wert

ermitteln.

Die erzeugte Abfrage qbeUE5-1 muss in eine neue Abfrage eingebunden werden, um mit einer Tabelle oder Abfrage verbunden zu werden, die das Geschlecht enthält. Wir benutzen die Tabelle AbgeordneteRT.

Für die Auswertung 'Abgeordnete insgesamt' nehmen wir nur Dauer in die neue Abfrage auf, je einmal für Anzahl, Summe, Mittelwert, Standardabweichung, Minimalwert und Maximalwert.



Für die nach Geschlecht differenzierte Betrachtung gruppieren wir nach sex.

sex	Dauer	Dauer	Dauer	Dauer	Dauer	Dauer
AbgeordneteRT	qbeUE5-1	qbeUE5-1	qbeUE5-1	qbeUE5-1	qbeUE5-1	qbeUE5-1
Gruppierung	Anzahl	Summe	Mittelwert	StAbw	Min	Max
	<input checked="" type="checkbox"/>					

Die Abfragen liefern folgende Ergebnistabellen:

Sozialdemokratische Abgeordnete des Reichstages insgesamt.

Anzahl von Dauer	Summe von Dauer	Mittelwert von Dauer	StAbw von Dauer	Min von Dauer	Max von Dauer
559	52752	94,37	91,98	0	519

Sozialdemokratische Abgeordnete des Reichstages, differenziert nach Geschlecht:

sex	Anzahl von Dauer	Summe von Dauer	Mittelwert von Dauer	StAbw von Dauer	Min von Dauer	Max von Dauer
m	510	49520	97,10	94,21	0	519
w	49	3232	65,96	57,66	6	173

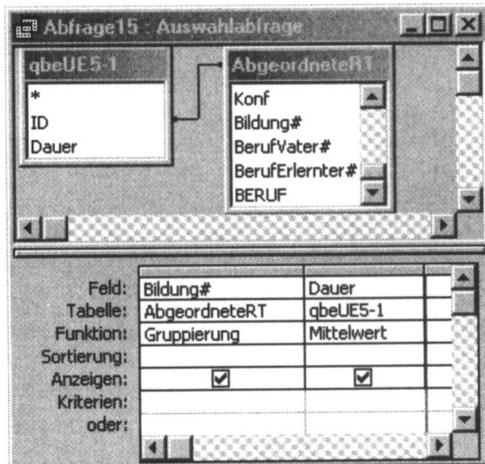
Die männlichen Abgeordneten waren im Durchschnitt 97 Monate, die weiblichen 66 Monate im Parlament.

Diese Abfrage lässt sich mit anderen Abfragen oder Tabellen, z.B. mit einer Abfrage zum Bildungsniveau oder zur sozialen Herkunft der Abgeordneten, durch einen JOIN verbinden.

9.5.1 Mandatsdauer und Bildungsniveau

In der folgenden Abfrage soll die durchschnittliche Verweildauer zum Bildungsniveau ins Verhältnis gesetzt werden.

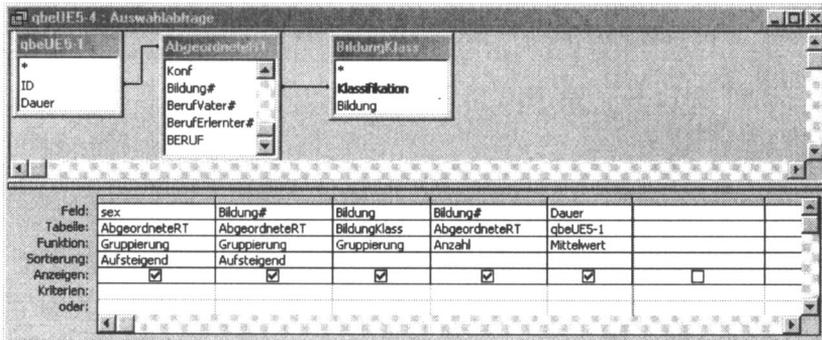
Die Tabelle AbgeordneteRT (sozialdemokratische Reichstagsabgeordnete) enthält den zehnstufigen Bildungsschlüssel. Wir verbinden die Abfrage qbeUE5-1 (Reichstagsabgeordnete mit Datenfeldern ID und Verweildauer im Reichstag) über die ID mit AbgeordneteRT. Relevant sind die Datenfelder Bildungsschlüssel (Bildung#) und Dauer (Verweildauer im Reichstag je sozialdemokratischer Abgeordneter). Über Bildungsschlüssel wird gruppiert und innerhalb der Gruppen für Dauer der Mittelwert errechnet.



Zur besseren Lesbarkeit nehmen wir das Datenfeld BildungKlass (Bildungsstufen im Textformat) mit in die Abfrage auf. Wir erhalten folgende Ergebnistabelle:

Bildung#	Bildung	Mittelwert von Dauer [Monate]
0	k. A.	26,00
1	Volksschule	90,56
2	Mittelschule ohne Abschluss	93,74
3	Mittelschule mit Abschluss / "Tertia"-Abschluss	161,88
4	Höhere Schule ohne "Einjähriges"	80,95
5	Höhere Schule mit "Einjährigem" (Obersekundareife)	105,38
6	Höhere Schule mit Abitur	176,50
7	Lehrerseminar mit Abschluss	92,44
8	Universität ohne Abschluss	121,54
9	Universität mit Abschluss	88,19

Bei einer Differenzierung zwischen männlichen und weiblichen Abgeordneten muss die Variable sex hinzugenommen werden. Um die Ergebnisse aussagekräftiger zu machen, wurde auch die Anzahl der Probanden je Bildungsstufe ausgegeben.

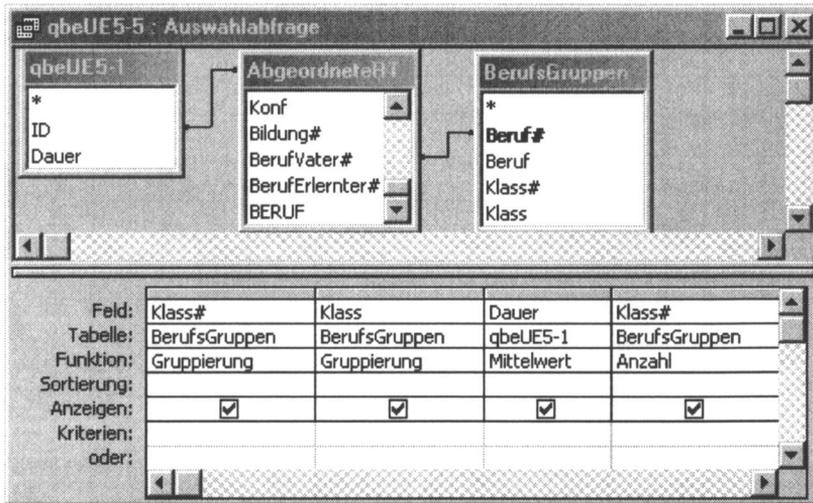


Wir erhalten die Tabelle:

sex	Bildung#	Bildung	Anzahl von Bildung#	Mittelwert von Dauer
m	0	k. A.	3	26,00
m	1	Volksschule	331	93,07
m	2	Mittelschule ohne Abschluss	37	100,14
m	3	Mittelschule mit Abschluss / "Tertia"-Abschluss	16	161,88
m	4	Höhere Schule ohne "Einjähriges"	16	74,38
m	5	Höhere Schule mit "Einjährigem" (Obersekundareife)	8	105,38
m	6	Höhere Schule mit Abitur	4	176,50
m	7	Lehrerseminar mit Abschluss	14	96,79
m	8	Universität ohne Abschluss	26	125,73
m	9	Universität mit Abschluss	55	90,51
w	1	Volksschule	31	63,81
w	2	Mittelschule ohne Abschluss	6	54,33
w	4	Höhere Schule ohne "Einjähriges"	3	116,00
w	7	Lehrerseminar mit Abschluss	4	77,25
w	8	Universität ohne Abschluss	2	67,00
w	9	Universität mit Abschluss	3	45,67

9.5.2 Mandatsdauer und soziale Herkunft

In der folgenden Abfrage wird die Verweildauer nach sozialer Herkunft untersucht. Das Vorgehen ist analog zu dem der vorangegangenen Abfrage. Statt der Kodierung der Bildung wird der Berufsschlüssel des Vaters eingebunden.



Verweildauer und soziale Herkunft (Beruf des Vaters) insgesamt

Klass#	Klass	Anzahl von Klass#	Mittelwert von Dauer	sex
1	ungelernte Arbeiter (unselbständig)	79	86,20	
2	gelernte Arbeiter (unselbständig)	123	92,73	
3	Selbständige	140	116,20	
4	bürgerliche Berufe	42	133,12	
5	politische Beamte	2	230,00	
6	Arbeiterbeamte	3	60,33	
8	untere/mittlere Angestellte/Beamte	36	110,64	
99	k. A.	134	60,10	

Verweildauer und soziale Herkunft (Beruf des Vaters) männliche Abgeordnete

Klass#	Klass	Anzahl von Klass#	Mittelwert von Dauer	sex
1	ungelernte Arbeiter (unselbständig)	72	86,74	m
2	gelernte Arbeiter (unselbständig)	112	95,02	m
3	Selbständige	129	119,04	m
4	bürgerliche Berufe	37	140,84	m
5	politische Beamte	2	230,00	m
6	Arbeiterbeamte	2	82,00	m
8	untere/mittlere Angestellte/Beamte	34	115,76	m
99	k. A.	122	61,52	m

Verweildauer und soziale Herkunft (Beruf des Vaters) weibliche Abgeordnete

Klass#	Klass	Anzahl von Klass#	Mittelwert von Dauer	sex
1	ungelernte Arbeiter (unselbständig)	7	80,71	w
2	gelernte Arbeiter (unselbständig)	11	69,45	w
3	Selbständige	11	82,91	w
4	bürgerliche Berufe	5	76,00	w
6	Arbeiterbeamte	1	17,00	w
8	untere/mittlere Angestellte/Beamte	2	23,50	w
99	k. A.	12	45,58	w

9.6 Übung 6: Kreuztabelle – Bildungsniveau und soziale Herkunft

Für statistische Analysen haben Kreuztabellen (auch Kontingenztabellen genannt) eine besondere Bedeutung. Man stellt zwei Merkmale einander gegenüber und zählt jeweils für eine Merkmalsausprägung des einen Merkmals, wie häufig die Werte des anderen Merkmals auftreten.² Kreuztabellen sind nur sinnvoll anwendbar auf Merkmale mit nichtmetrischem Messniveau, genauer, die Domäne (Wertebereich) besteht aus endlich vielen Elementen.

Im Folgenden sollen Kreuztabellen an einem Beispiel erläutert werden und daran anschließend wollen wir mit ACCESS eine Kreuztabelle erstellen

² Vgl. Sensch, (1995), S.144.

9.6.1 Alter bei Erstmandat und Geschlecht

Wir betrachten das Alter der Reichstagsabgeordneten bei ihrem Eintritt in den Reichstag. Gibt es hier Unterschiede zwischen männlichen und weiblichen Abgeordneten?

ID	Name	Vorname	sex	geboren	Ersteintritt	Alter	Klass
10050	Aderhold	Karl	m	30.06.1884	Mär 1919	34	2
10060	Agnes	Lore	w	04.06.1876	Jan 1919	42	3
10070	Agster	Alfred	m	12.04.1858	Jun 1898	40	3
10100	Albrecht	Adolf	m	14.07.1855	Jun 1898	42	3
10160	Ansorge	Marie	w	15.12.1880	Jun 1920	39	2
10180	Antrick	Otto	m	24.11.1858	Jun 1898	39	2
10240	Arning	Marie	w	19.04.1887	Dez 1924	37	2
10280	Arzt	Arthur	m	09.10.1880	Mai 1928	47	3
...

Der Übersichtlichkeit halber betrachten wir nicht das Alter, sondern klassifizieren nach Altersgruppen:

Klassifikation	Altergruppen
1	unter 30
2	30 bis unter 40
3	40 bis unter 50
4	50 bis unter 60
5	60 bis unter 70
6	70 und älter

Wir wollen jetzt tabellarisch das Attribut Geschlecht (sex) der Altergruppe (Klass) gegenüberstellen, d.h wir zählen aus, wie häufig Männer bzw. Frauen in den einzelnen Altersgruppen in den Reichstag gewählt wurden. Eine solche Tabelle ist eine Kreuztabelle.

		Altersgruppen					
		1	2	3	4	5	6
Geschlecht	m	12	171	220	76	8	2
	w	1	11	25	6	1	-

Im Kopf der Tabelle steht das Merkmal Altersgruppen mit den Ausprägungen {1, 2, 3, 4, 5, 6}, in den ersten beiden Spalten das Merkmal Geschlecht mit den Ausprägungen {'m', 'w'}. Statistik hat ein anderes Herangehen als Datenmo-

dellierung. Was hier Attribute mit ihren Wertebereichen oder Domänen waren, sind dort Merkmale und Ausprägungen.

In den Zellen der Tabelle steht die absolute Häufigkeit. Ob es darüber hinaus, dass es weniger Frauen sind, noch weitere Unterschiede gibt, ist schwer zu erkennen. Der Überblick wird wesentlich verbessert, wenn man nicht absolute Häufigkeiten betrachtet, sondern die Zeilen summiert und Prozentwerte berechnet. Solche Zeilenprozentwertberechnungen (oder auch Spaltenprozentwertberechnungen) werden von Statistikprogrammen unterstützt. Kreuztabellenabfragen in ACCESS ermitteln nur die absoluten Häufigkeiten. Prozentwerte müssen dann in weiteren Schritten relativ aufwändig berechnet werden. Da das Statistikprogramm SPSS (Statistical Package for Social Science) direkt auf ACCESS-Abfragen zugreifen kann, sollte man in ACCESS nur die gewünschte Sicht auf die Daten realisieren und anspruchsvollere statistische Auswertungen mit diesem Statistikprogrammpaket durchführen.

Kreuztabelle mit Zeilensumme

		Altersgruppen						Summe
		1	2	3	4	5	6	
Geschlecht	m	12	171	220	76	8	2	489
	w	1	11	25	6	1	-	44

Kreuztabelle mit Zeilensumme(Zeilensumme)

		Altersgruppen						Summe
		1	2	3	4	5	6	
Geschlecht	m	2,5	35,0	45,0	15,5	1,6	0,4	100,0
	w	2,3	25,0	56,8	13,6	2,3	-	100,0

Erst in der Tabelle mit Prozentberechnungen sieht man deutlich, dass die Frauen in der Altersgruppe der unter Vierzigjährigen gegenüber den Männern deutlich weniger vertreten sind. Dafür sind sie in der Altersgruppe der unter Fünfzigjährigen deutlich stärker vertreten. Es zeigt sich die Tendenz, dass Frauen in höherem Alter in den Reichstag einziehen als Männer.

9.6.2 Soziale Herkunft und Bildungsniveau

Wenden wir uns einem weiteren Beispiel zu. Wir haben die Vermutung, dass zwischen der sozialen Herkunft der Parlamentarier und ihrem Bildungsniveau ein statistischer Zusammenhang besteht. Wenn das so ist, dann müsste in einer Kreuztabelle für Probanden mit hohem sozialen Status auch ein hohes Bildungsniveau auftreten.

Für das Merkmal Bildungsniveau hatten wir bereits die folgende Klassifikation benutzt:

Berufsgruppen

Klass#	Klass
1	ungelernte Arbeiter (unselbständig)
2	gelernte Arbeiter (unselbständig)
3	Selbständige
4	bürgerliche Berufe
5	politische Beamte
6	Arbeiterbeamte
7	ohne Beruf
8	untere/mittlere Angestellte/Beamte
99	k. A.
8888	

Den Beruf des Vaters benutzen wir als Indikator für soziale Herkunft. Es wird die Berufsklassifikation der vorangegangenen Übung benutzt.

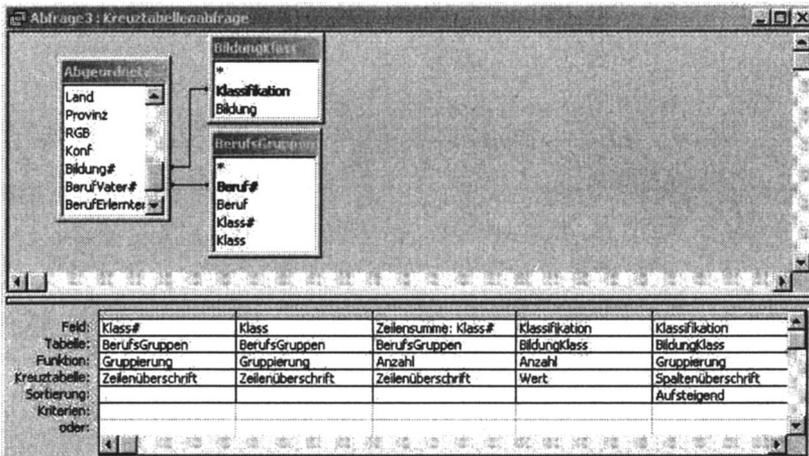
Bildungsklassifikation

Klassifikation	Bildung
0	k. A.
1	Volksschule
2	Mittelschule ohne Abschluss.
3	Mittelschule mit Abschluss / "Tertia"-Abschluss
4	Höhere Schule ohne "Einjähriges"
5	Höhere Schule mit "Einjährigem" (Obersekundareife)
6	Höhere Schule mit Abitur
7	Lehrerseminar mit Abschluss
8	Universität ohne Abschluss
9	Universität mit Abschluss

Sie gehen so vor:

- ◆ Klicken Sie im Datenbankfenster auf Abfrage → Neu und wählen Sie Entwurf.
- ◆ Ziehen Sie die Tabellen AbgeordneteRT und KonfessionsVerzeichnis in das Abfrageentwurf fenster.
- ◆ Klicken Sie auf die Lasche neben dem Symbol  in der Symbolleiste und wählen Sie den Abfragetyp Kreuztabelle.

- ◆ Ziehen Sie aus der Tabelle AbgeordneteRT das Feld LP und aus der Tabelle KonfessionsVerzeichnis die Felder Konf und Konfession in den Entwurfsbereich der Abfrage.
- ◆ Ändern Sie in der Zeile Funktionen den Eintrag unter Konf auf Anzahl (die anderen beiden bleiben auf Gruppieren).
- ◆ Wählen Sie in der Zeile Kreuztabelle für die Spalte LP 'Zeilenüberschrift', für die Spalte Konf 'Wert' und für Konfession 'Spaltenüberschrift'.
- ◆ Legen Sie als Sortierfolge für LP Aufsteigend fest.



Als Ergebnis erhalten wir folgende Kreuztabelle:

Kreuztabelle Bildungsniveau und soziale Herkunft – absolute Zahlen

Klass #	Klassifikation	Zeilen-summe	0	1	2	3	4	5	6	7	8	9
1	ungelernte Arbeiter (unselbständig)	79		68	6		2				1	2
2	gelernte Arbeiter (unselbständig)	123	1	94	13	5		2		3	3	2
3	Selbständige	140		79	15	2	7	3	1	3	9	21
4	bürgerliche Berufe	42		5		1	2	1		4	10	19
5	politische Beamte	2							2			
6	Arbeiterbeamte	3			1							2
8	untere/mittlere Angestellte/ Beamte	36		15	1	4	3		1	3	1	8
99	k. A.	134	2	101	7	4	5	2		5	4	4

Die Tabelle ist so zu deuten: Beispielsweise sagt das schattierte Feld aus, dass 10 Abgeordnete, deren Väter 'bürgerliche Berufe' ausübten, den Bildungsabschluss 'Universität ohne Abschluss' (=8) erreichten.

Zur Veranschaulichung der Ergebnisse wurden für die absoluten Werte Zeilenprozentage berechnet und die Spalte bzw. Zeile 'Keine Angabe' weggelassen. In ACCESS gestaltet sich die Berechnung einer solchen Tabelle als ausgesprochen umständlich. Das ist mit einem Statistikprogrammpaket wesentlich einfacher zu erledigen. SPSS kann direkt auf ACCESS-Tabellen und -Abfragen zugreifen. Statistische Berechnungen, die über statistische Grundmaße hinausgehen, sollten nicht mit ACCESS, sondern mit einem spezialisierten Programm durchgeführt werden

Kreuztabelle Bildungsniveau und soziale Herkunft – Zeilenprozentage

		Bildungsniveau									
		1	2	3	4	5	6	7	8	9	
soziale Herkunft	1	86,2	7,6		2,5				1,3	2,5	100
	2	76,4	10,6	4,1		1,6		2,4	2,4	1,6	100
	3	56,4	10,7	1,4	5,0	2,1	0,7	2,1	6,4	15,0	100
	4	11,9		2,4	4,8	2,4		9,5	23,8	45,2	100
	5						100,0				100
	6		33,3							66,7	100
	8	41,7	2,8	11,1	8,3		2,8	8,3	2,8	22,2	100

Der Zusammenhang ist, wie häufig bei sozialwissenschaftlichen Erscheinungen, schwach ausgeprägt. Abgeordnete, deren Väter Arbeiter waren, absolvierten nur zu einem geringen Teil ein Studium.

Abkürzungen

ANSI	American National Standards Institute
BIOSOP	Biographisches Handbuch der sozialdemokratischen Parlamentarier in den deutschen Reichs- und Landtagen 1867-1933
CAD	Computer Aided Design
CASE	Computer Aided Software Engineering
DB	Datenbank
DBMS	Datenbankmanagementsystem
DBS	Datenbanksystem
DM	Datenmodell
EERM	Erweitertes Entity Relationship Model
EDV	Elektronische Datenverarbeitung
ERM	Entity Relationship Model
HM	Hierarchisches Modell
ID	Identifikator
LP	Legislaturperiode
MDB	Mandat des Deutschen Bundestags
MDR	Mandat des Deutschen Reichstags
MDV	Mandat der Volkskammer der DDR
NF	Normalform
NF ²	Non-First-Normal-Form (Modell der geschachtelten Relationen)
NWM	Netzwerkmodell
OLE	object linking and embedding
OODM	Objektorientiertes Datenmodell
RDBMS	Relationales Datenbankmanagementsystem
RM	Relationales Modell
RT	Reichstag
SemDM	Semantisches Datenmodell
SPD	Sozialdemokratische Partei Deutschlands
SPSS	Statistical Package for Social Science
SQL	Structured Query Language
WK	Wahlkreis

Referenzliteratur

- Chen, P. P. S.; Knöll, H.-D.: Der Entity Relationship-Ansatz zum logischen Systementwurf. BI Wissenschaftsverlag, Mannheim, Wien, Zürich, 1991.
- Codd, E. F.: A Relational Model of Data for Large Shared Data Banks. Communications of the ACM, 13, Nr. 6, 1970, 377-387.
- Codd, E. F.: Extending the Database Relational Model to Capture More Meaning. ACM Transaction on Database Systems, Vol 4, No. 4, 397-434, 1979.
- Codd, E. F.: Relational database: A practical foundation for productivity. Communications of the ACM, 25, Nr. 2, 1982, 109-117.
- Elmasri, R.; Navathe, S.B.: Fundamentals of Database Systems. Benjamin Cummings, Redwood City, Ca, USA, 1994.
- Heuer, A.; Saake, G.: Datenbanken – Konzepte und Sprachen. International Thomson Publ., Bonn u.a., 1995.
- Howe, D. R.: Data Analysis for Data Base Design, Edward Arnold, Baltimor, Maryland, USA, 1983.
- Kemper, A.; Eickler, A.: Datenbanksysteme. Eine Einführung. Oldenburg, 2001.
- Lockemann, P. C., Schmidt, J. W. (Hrsg.): Datenbank-Handbuch. Springer Verlag, Berlin u.a., 1987.
- Scheffé, P.: Softwaretechnik und Erkenntnistheorie. Informatik-Spektrum 22, 1999, 122-135.
- Schröder, W. H.: Historische Sozialforschung: Identifikation, Organisation, Institution. HSR, Supplementheft No. 6, Köln, 1994.
- Schröder, W. H.: Historische Parlamentarismus-, Eliten- und Biographieforschung, Wilhelm Heinz Schröder, Wilhelm Weege, Martina Zech. HSR, Supplementheft No. 11, Köln, 2000.
- Schröder, W. H.: Sozialdemokratische Parlamentarier in den deutschen Reichs- und Landtagen 1867 - 1933, Droste Verlag Düsseldorf, 1995.
- Sensch, J.: Statistische Modelle in der Historischen Sozialforschung. HSR, Supplementheft No. 7, Köln, 1995.
- Trautloft, R.; Lindner, U.: Datenbanken – Entwurf und Anwendung, Verlag Technik, Berlin, 1991.
- Wedekind, H. et al.: Modellierung, Simulation, Visualisierung: Zu aktuellen Aufgaben der Informatik. Informatik-Spektrum 21, 1998, 265-272.
- Vetter, M.: Das Jahrhundertproblem der Informatik. Effektives Datendesign: Praxis-Erfahrungen, Gunter Müller-Ettrich (Hrsg.), Rudolf Müller Verlag, Köln, 1989.

Anhang A: Datenbankstruktur der Datenbank „Sozialdemokratische Parlamentarier in den deutschen Reichs- und Landtagen“ – Parl-SPD

1. Tabellen, die nur zu Übungszwecken erstellt wurden:

Abgeordnete (Abgeordnete des Deutschen Reichstags ohne Fremdschlüssel)

	Feldname	Felddatentyp	Hinweis
	ID	Zahl	ID des Abgeordneten
	Name	Text	Nachname des Abgeordneten
	Vorname	Text	Vorname(n) des Abgeordneten
	Sex	Text	Geschlecht: {'m', 'w'}
	Geburtsort	Text	Geburtsort
	Geboren	Text	Geburtsdatum im Format tt.mm.jjjj
	Sterbeort	Text	Sterbeort
	gestorben	Text	Sterbedatum im Format tt.mm.jjjj
	RTErsteintritt	Text	Jahr und Monat des Ersteintritts in den Deutschen Reichstag. Format jjjj/mm
	LP	Zahl	Legislaturperiode des Erstmandats für den Reichstag
	Bezeichnung	Text	Wahlkreisbezeichnung des Ersteintrittswahlkreises
	Region	Text	Region des Wahlkreises
	Land	Text	Land des Wahlkreises
	Provinz	Text	Provinz des Wahlkreises
	RGB	Text	Regierungsbezirk des Ersteintritts
	Konfession	Text	Konfession
	BerufVater	Text	Beruf des Vaters
	Ausbildung	Text	Ausbildungsabschluss

AbgeordneteRTBerEM (Abgeordnete des Deutschen Reichstags mit Beruf vor Erstmandat)

	Feldname	Felddatentyp	Hinweis
	ID	Zahl	ID des Abgeordneten
	Periode	Zahl	Legislaturperiode des Erstmandats für den RT
	Jahr	Zahl	Jahr des Ersteintritts in den Reichstag
	Beruf	Zahl	Code des Berufs bei Ersteintritt

AbgeordneteRT (Abgeordnete des Deutschen Reichstags mit Fremdschlüssel)

	Feldname	Felddatentyp	Hinweis
	ID	Zahl	ID des Abgeordneten
	Name	Zahl	Nachname des Abgeordneten
	Vorname	Text	Vorname(n) des Abgeordneten
	Sex	Text	Geschlecht: {'m', 'w'}
	Geburtsort	Text	Geburtsort
	Geboren	Text	Geburtsdatum im Format tt.mm.jjjj
	Sterbeort	Text	Sterbeort
	Gestorben	Text	Sterbedatum im Format tt.mm.jjjj
	RTErsteintritt	Text	Jahr und Monat des Ersteintritts in den Deutschen Reichstag. Format jjjj/mm
	LP	Zahl	Legislaturperiode des Erstmandats für den Reichstag
	Bezeichnung	Text	Wahlkreisbezeichnung des Ersteintrittswahlkreises
	Region	Text	Region des Wahlkreises
	Land	Text	Land des Wahlkreises
	Provinz	Text	Provinz des Wahlkreises
	RGB	Text	Regierungsbezirk des Ersteintritts
	Konf	Zahl	Konfession, kodiert
	Bildung#	Zahl	Höchster Bildungsabschluss, kodiert
	BerufVater#	Zahl	Beruf des Vaters
	BerufErlerner#	Zahl	Erlerner Beruf
	BERUF	Zahl	Code des Berufs bei Ersteintritt in der Deutschen Reichstag

2. Tabellen der Datenbank „Sozialdemokratische Parlamentarier in den Deutschen Reichs- und Landtagen“

Alias (weitere Namen der Abgeordneten)

	Feldname	Felddatentyp	Hinweis
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Name	Text	weiterer Name des Abgeordneten
	Bemerkung	Text	Hinweis zum weiteren Namen wie 'geb.', 'verh.', 'genannt', 'verw.'

Ausbildung (Ausbildungsmaßnahmen aller Abgeordneten)

	Feldname	Felddatentyp	Hinweis
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	LFDNR		Chronologische Reihenfolge; wegen fehlender Zeitangaben
	ABGNr		Fremdschlüssel Ausbildungsverzeichnis
	ANR		Fremdschlüssel OrteAusbildung
	Bemerkung		Hinweise zur Ausbildungsmaßnahme
	vonJahr		Beginn der Ausbildungsmaßnahme
	bisJahr		Ende der Ausbildungsmaßnahme

Ausbildungsverzeichnis (Ausbildungsmaßnahmen aller Abgeordneten)

	Feldname	Felddatentyp	Hinweis
	ABGNr	Zahl	
	Ausbildung	Text	Ausbildungsmaßnahme
	Code	Text	Vom Nutzer frei zu vergeben

BemerkungTod (Bemerkungen zu Todesumständen)

	Feldname	Felddatentyp	Hinweis
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Bemerkung		Hinweis zu Todesumständen

Berufsgruppen (standardisierte Berufsbezeichnung und Klassifikation)

	Feldname	Felddatentyp	Hinweis
	Beruf#	Zahl	Berufscode
	Beruf	Text	Berufe
	Klass#	Zahl	Code der Berufsklasse
	Klass	Text	Berufsklasse

BerufsgruppenMA (Berufsgruppen bei Mandatsantritt)

	Feldname	Felddatentyp	Hinweis
	Beruf#	Zahl	Berufscode
	Beruf	Text	Berufe
	Klass#	Zahl	Code der Berufsklasse
	Klass	Text	Berufsklasse

Berufsverzeichnis (Berufe wie im „Handbuch Sozialdemokratischer Parlamentarier“ angegeben)

	Feldname	Felddatentyp	Beschreibung
	PROF	Zahl	Code für den Beruf
	Beruf	Text	
	Code1	Zahl	Von Nutzer frei zu vergebender Code

BerufVaterOriginal (Berufe der Väter wie im „Handbuch Sozialdemokratischer Parlamentarier“ angegeben)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Beruf	Text	Berufsbezeichnung des Vaterberufs wie im „Handbuch Sozialdemokratischer Parlamentarier“

BerufVorMandat (Berufe vor Erstmandat; Mehrfachangaben je Abgeordneter)

	Feldname	Felddatentyp	Beschreibung
	id	Zahl	Fremdschlüssel Personen
	B#	Zahl	Berufs-Code
	Kenn	Text	

Bildung (Bildungsstand der Abgeordneten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen;
	Klassifikation	Zahl	Fremdschlüssel BildungKlass; Berufsgruppen-Code

BildungKlass (Liste der Klassifikation Bildungsstand)

	Feldname	Felddatentyp	Beschreibung
	Klassifikation	Zahl	Code für Bildungsniveau
	Bildung	Text	Folgende Bildungsniveaus sind möglich: 'k. A.', 'Volksschule', 'Mittelschule ohne Abschluss', 'Mittelschule mit Abschluss / "Tertia"-Abschluss', 'Höhere Schule ohne „Einjähriges“', 'Höhere Schule mit „Einjährigem“ (Obersekundareife)', 'Höhere Schule mit Abitur', 'Lehrerseminar mit Abschluss', 'Universität ohne Abschluss', 'Universität mit Abschluss'

Ersteintritt (Ersteintritt in eine sozialistische/sozialdemokratische Partei)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen
	Jahr	Zahl	Jahr des Ersteintritts
	ORGNr	Zahl	Fremdschlüssel PoOrganisation
	JahrUnsicher	Text	Hinweis, ob das Jahr eine gesicherte Information darstellt

Famstand (Familienstand)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen
	Familienstand	Text	'ledig', 'verh.', 'gesch.', 'verw.'
	HeiratsJahr	Zahl	Jahr der Eheschließung
	Bemerkung	Text	Bemerkungen, z.B. zur Zahl der Kinder

Gewerkschaften (Liste aller Einzelgewerkschaften)

	Feldname	Felddatentyp	Beschreibung
<input type="checkbox"/>	GEWNR	Zahl	Code für die Gewerkschaft
	Gewerk	Text	Bezeichnung der Gewerkschaft

GewkongTeiln (Teilnahme an Gewerkschaftskongressen)

	Feldname	Felddatentyp	Beschreibung
<input type="checkbox"/>	ID	Zahl	Fremdschlüssel Personen
<input type="checkbox"/>	Jahr	Zahl	
<input type="checkbox"/>	GEWNR	Zahl	Fremdschlüssel Gewerkschaften
	Bemerkung	Text	

Konfessionsverzeichnis (Liste aller Konfessionsausprägungen)

	Feldname	Felddatentyp	Beschreibung
<input type="checkbox"/>	KONF	Zahl	Kodierung für Konfession
	Konfession	Text	Es treten folgende Ausprägungen auf: 'altkath.' 'baptistisch' 'diss.' 'diss. (Mennonit)' 'diss., später Quäker' 'dt.-kath.' 'dt.-kath., später diss.' 'ev.' 'ev., diss.' 'ev., später diss.' 'ev., später diss., später erneut ev.' 'ev., später dt.-kath.' 'ev., später mennonitisch' 'jüd.' 'jüd., später diss.' 'jüd., später diss., später erneut jüd.'

Landtage (alle Landtage der deutschen Länderparlamente, s. „Handbuch Sozialdemokratischer Parlamentarier“)

	Feldname	Felddatentyp	Beschreibung
	LTG	Zahl	Nummer des Landtages
	LND	Zahl	Kennung für das Land
	LP	Zahl	Nr. der Legislaturperiode des Landtages
	Land	Text	Land (Redundante Information, da über LND auf ein Land verwiesen wird)
	vonMonat	Zahl	Beginnmonat des Landtages
	vonJahr	Zahl	Beginnjahr des Landtages
	bisMonat	Zahl	Endmonat des Landtages
	bisJahr	Zahl	Endjahr des Landtages

MdB (Mitglied des Bundestages)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen
	vonMonat	Zahl	Mitglied des Bundestages ab Monat
	vonJahr	Zahl	Mitglied des Bundestages ab Jahr
	bisMonat	Zahl	Mitglied des Bundestages bis Monat
	bisJahr	Zahl	Mitglied des Bundestages bis Jahr

MdL (Mandate der Landtage)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	LPG	Zahl	Fremdschlüssel Landtag
	von	Zahl	Jahr des Eintritts in den Landtag
	bis	Zahl	Jahr der Mandatsniederlegung in dieser Legislaturperiode
	Bemerkung	Text	Bemerkung zum Mandat; redundant ist hier das Land angegeben

MdR (Mandate der Reichstage)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	LP	Zahl	Fremdschlüssel RTL; Nr. der Reichstag-Legislaturperiode
	vonMonat	Zahl	Monat des Eintritts in den Reichstag
	vonJahr	Zahl	Jahr des Eintritts in den Reichstag
	WK	Zahl	Fremdschlüssel WK; Nummer des Wahlkreises
	bisMonat	Zahl	Monat der Mandatsniederlegung in dieser Legislaturperiode
	bisJahr	Zahl	Jahr der Mandatsniederlegung in dieser Legislaturperiode
	Bemerkung	Text	Bemerkung zum Mandat

MdV (Mitglieder der Volkskammer der DDR)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	von	Zahl	Jahr des Eintritts in die Volkskammer
	bis	Zahl	Jahr des Ausscheidens

Militärdienst (Militärdienste der Abgeordneten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	von	Zahl	Jahr des Beginns des Militärdienstes
	bis	Zahl	Jahr des Ausscheidens aus dem Militärdienst
	Dienst	Logisch	'ja' / 'nein': Militärdienst geleistet oder nicht
	Bemerkung	Text	Bemerkung zum Militärdienst

Nach45 (Verbleib nach 1945)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	lfdnr	Zahl	Fortlaufende Nummerierung (chronologisch) der Stationen je Abgeordneter
	OKNR	Zahl	Fremdschlüssel Ortsverzeichnis; Ortskennnummer
	vonJahr	Zahl	Beginn des Lebensabschnitts (Jahr)
	vonMonat	Zahl	Beginn des Lebensabschnitts (Monat)
	bisJahr	Zahl	Ende des Lebensabschnitts (Jahr)
	bisMonat	Zahl	Ende des Lebensabschnitts (Monat)
	Bemerkung	Text	Bemerkung zu Tätigkeit/Verbleib nach 1945

NSZ (Liste von Ereignissen während der NS-Zeit)

	Feldname	Felddatentyp	Beschreibung	
	Kennung	Zahl	Es sind folgende Labels zugelassen: 9 → - (keine der Folgenden) E → Emigration G → Gefangenschaft H → Haft I → Internierung K → Kriegsteilnehmer M → Militärdienst	
			Verbleib	Text

NSZeit (Verbleib der Abgeordneten während der NS-Zeit)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	LFDNR	Zahl	Lfd. Nr. des Ereignisses (in chronologischer Reihenfolge)
	vonJahr	Zahl	Anfangsjahr der Maßnahme / des Ereignisses
	vonMonat	Zahl	Anfangsmonat der Maßnahme / des Ereignisses
	bisJahr	Zahl	Endjahr der Maßnahme / des Ereignisses
	bisMonat	Zahl	Endmonat der Maßnahme / des Ereignisses
	Haft_Emig	Text	Kurzzeichen wie Tabelle NSZ
	Bemerkung	Text	Beschreibung der Maßnahme / des Ereignisses

NSZeitText

(Verbleib der Abgeordneten während der NS-Zeit als Übersicht für die Internet-Datenbank)

	Feldname	Felddatentyp	Beschreibung
<input checked="" type="checkbox"/>	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
<input checked="" type="checkbox"/>	Haft	Logisch	
	KZ	Logisch	
	Emigration	Logisch	
	Internierung	Logisch	
	Gefangenschaft	Logisch	
	Kriegsteilnahme	Logisch	
	Rueckkehr	Logisch	
	Verbleib	Logisch	
	Tod	Logisch	
	Bemerkung	Text	Beschreibung der Maßnahme / des Ereignisses

Orte (Tabelle noch in Bearbeitung durch DB-Administrator)

Feldname	Felddatentyp	Beschreibung
Ortsname	Text	Name des Ortes in amtlicher Schreibweise
Land	Text	Land
LND	Zahl	Statistische Kennziffer der amtlichen Statistik – 2-stellig, Land
RGB	Zahl	Statistische Kennziffer der amtlichen Statistik – 1-stellig, Regbz.
KRS	Zahl	Statistische Kennziffer der amtlichen Statistik – 1-stellig, Kreis
ORT	Zahl	Statistische Kennziffer der amtlichen Statistik – 1-stellig, Gemeinde
TGKZ	Text	Statistische Kennziffer der amtlichen Statistik – 3-stellig, Teilgemeindegennzeichen
Bemerkung	Text	Bezeichnung des Ortes wie im „Handbuch Sozialdemokratische Parlamentarier“
OKNR	Zahl	Ortskennung
TxtAmt	Text	Bezeichnung des Amtes
txtKreis	Text	Bezeichnung des Kreises
txtRegBez	Text	Bezeichnung des Regierungsbezirks
txtProv	Text	Bezeichnung der Provinz
txtLand	Text	Bezeichnung des Landes

Orteausbildung (Ausbildungsorte für mehrere Orte je Ausbildungsmaßnahme der Abgeordneten)

Feldname	Felddatentyp	Beschreibung
ID	Zahl	Fremdschlüssel Ausbildung; ID des Abgeordneten
ANR	Zahl	Fremdschlüssel Ausbildung; Lfd. Nr. der Ausbildungsmaßnahme
OKNR	Zahl	Fremdschlüssel Ortsverzeichnis; Ortskennziffer

OrteNSZeit (Orte zu einer Lebensstation in der NS-Zeit)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel NSZeit; ID des Abgeordneten
	LFDNR	Zahl	Fremdschlüssel NSZeit; Lfd. Nr. des Ortes zu einem Ereignis / Maßnahme
	OKNR	Zahl	Fremdschlüssel Ortsverzeichnis; Ortskennziffer

OrteWerdegang (Orte zu einer Lebensstation im Werdegang eines Abgeordneten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Werdegang; ID des Abgeordneten
	LFDNR	Zahl	Fremdschlüssel Werdegang; lfd. Nr. zu Ereignis / Maßnahme
	OKNR	Zahl	Fremdschlüssel Ortsverzeichnis; Ortskennziffer

Ortsverzeichnis (Liste aller in DB auftretenden Orte)

	Feldname	Felddatentyp	Beschreibung
	OKNR	Zahl	Ortskennung
	Ortsname	Text	Name des Ortes in amtlicher Schreibweise
	Land	Text	Land
	LND	Zahl	Statistische Kennziffer der amtlichen Statistik – 2-stellig, Land
	RGB	Zahl	Statistische Kennziffer der amtlichen Statistik – 1-stellig, Regbz.
	KRS	Zahl	Statistische Kennziffer der amtlichen Statistik – 1-stellig, Kreis
	ORT	Zahl	Statistische Kennziffer der amtlichen Statistik – 1-stellig, Gemeinde
	TGKZ	Text	Statistische Kennziffer der amtlichen Statistik – 3-stellig, Teilgemeindegkennzeichen
	Bemerkung	Text	Bezeichnung des Ortes wie im „Handbuch Sozialdemokratischer Parlamentarier“

Parteitage (Liste aller Parteitage soz. / sozialdemokratischer Parteien)

	Feldname	Felddatentyp	Beschreibung
	PNR	Zahl	Kennziffer zur Identifikation aller Parteitage
	OrgNr	Text	Kennziffer zur Identifikation aller Organisationen
	Jahr	Text	Jahr des Parteitages
	Monat	Zahl	Monat des Parteitages
	Ort	Zahl	Ortsname
	Partei	Zahl	Partei (Kürzel)

Parteitagsteilnahme (Alle Teilnahmen an Parteitag)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	PNR	Zahl	Fremdschlüssel Parteitage; Kennziffer zur Identifikation aller Parteitage
	Bemerkung	Text	Jahr des Parteitages

Personen (Alle sozialdemokratische Abgeordneten mit Grunddaten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	ID des Abgeordneten
	Name	Text	Nachname des Abgeordneten
	Vorname	Text	Vorname des Abgeordneten
	Sex	Text	Geschlecht; nur {'m' 'w'}
	GebOKNR	Zahl	Fremdschlüssel Ortsverzeichnis; Ortskennziffer für den Geburtsort
	GebTag	Zahl	Geburtstag
	GebMonat	Zahl	Geburtsmonat
	GebJahr	Zahl	Geburtsjahr
	SterbeOKNR	Zahl	Fremdschlüssel Ortsverzeichnis; Ortskennziffer für den Sterbeort
	SterbeTag	Zahl	Sterbetag
	SterbeMonat	Zahl	Sterbemonat
	SterbeJahr	Zahl	Sterbejahr
	Konf	Zahl	Fremdschlüssel Konfessionsverzeichnis
	Bildung#	Zahl	Fremdschlüssel BildungKlass; Kennziffer für den Bildungsstand
	VaterPROF	Zahl	Fremdschlüssel BerufVaterOriginal; Kennziffer für den Beruf des Vaters (Berufsbezeichnung wie im „Handbuch Sozialdemokratische Parlamentarier“)
	BerufVater#	Zahl	Fremdschlüssel Berufsgruppen; Kennziffer für die Berufsklassifikation des Vaters
	BerufErlerner#	Zahl	Fremdschlüssel Berufsgruppen; Kennziffer für die Berufsklassifikation des erlernten Berufs

PolOrganisation (Liste aller politischen Organisationen)

	Feldname	Felddatentyp	Beschreibung
	OrgNr	Zahl	Kennziffer der politischen Organisation
	Bezeichnung	Text	Kurzbezeichnung der pol. Organisation
	Parteiame	Text	Bezeichnung der politischen Organisation
	Zeitraum	Text	Zeitraum der Existenz der pol. Org.: von Jahr bis Jahr (nicht in erster Normalform)

PersonenTätig (Berufliche Tätigkeiten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	lfdnr	Zahl	Lfd. Nr. der Tätigkeit (chronologisch)
	TNR	Zahl	Fremdschlüssel Tätigkeiten; Kennziffer der Tätigkeit
	von	Zahl	Beginn der Tätigkeit (Jahr)
	bis	Zahl	Beendigung der Tätigkeit (Jahr)
	Bemerkung	Text	Hinweis auf die Tätigkeit

PersVerbandstag (Teilnahmen an Verbandstagen)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	VBTAGNR	Zahl	Fremdschlüssel Verbandstag; Kennziffer des Verbandstages

Pub (Publikationen der Abgeordneten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Titel	Text	Titel der Publikation
	Ort	Text	Erscheinungsort
	Jahr	Text	Erscheinungsjahr
	Bemerkung	Text	Hinweis zur Publikationstätigkeit

RTKand (Reichtagskandidaturen, (erfolgreich))

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	LP	Zahl	Fremdschlüssel RTL; Legislaturperiode der Reichstagskandidatur
	WK	Zahl	Fremdschlüssel WK; Wahlkreis der Kandidatur
	Tag	Zahl	Tag der Kandidatur
	Monat	Zahl	Monat der Kandidatur
	Jahr	Zahl	Jahr der Kandidatur
	Bemerkung	Text	Hinweis zur Kandidatur

RTL (Reichtagslegislaturperioden)

	Feldname	Felddatentyp	Beschreibung
	LP	Text	Legislaturperiode des Reichstags (1... ,24)
	vonMonat	Zahl	Anfangsmonat der Legislaturperiode
	vonJahr	Zahl	Anfangsjahr der Legislaturperiode
	bisMonat	Zahl	Endmonat der Legislaturperiode
	bisJahr	Zahl	Endjahr der Legislaturperiode

Tätig

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Lfdnr	Zahl	Laufende Nr. der Tätigkeit (chronologisch)
	von	Zahl	Beginn der Tätigkeit (Jahreszahl)
	bis	Zahl	Ende der Tätigkeit (Jahreszahl)
	Tätigkeit	Text	Bezeichnung der Tätigkeit
	Bemerkung	Text	Hinweis zur Tätigkeit (orig. aus „Handbuch Sozialdemokratische Parlamentarier“)

Tätigkeiten (Liste aller Tätigkeiten; Anbindung an Tabelle PersonTätig)

	Feldname	Felddatentyp	Beschreibung
	TNR	Zahl	Kennziffer der Tätigkeit
	Tätigkeit	Text	Bezeichnung der Tätigkeit

Titel (Titel der Abgeordneten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Titel	Text	Titel

Verbandstag (Liste aller Verbandstage der Berufsverbände)

	Feldname	Felddatentyp	Beschreibung
	VBTAGNR	Zahl	Kennziffer des Verbandstages
	Jahr	Zahl	Jahr des Verbandstages
	Verband	Text	Bezeichnung des Verbandes

Verwandt (Verwandtschaftsbeziehungen zu anderen Abgeordneten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Verwandtschaftsgrad	Text	Beschreibung des Verwandtschaftsgrades
	Name	Text	Nachname des Verwandten
	Vorname	Text	Vorname des Verwandten
	ID2	Zahl	Fremdschlüssel Personen; ID des Verwandten

Wanderschaft (Wanderschaften der Abgeordneten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Zielland	Text	Autokennzeichen des Ziellandes oder 'keine Ortsangabe'

Werdegang (Lebensstationen der Abgeordneten)

	Feldname	Felddatentyp	Beschreibung
	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Lfdnr	Zahl	Laufende Nr. der Tätigkeit (chronologisch)
	vonJahr	Zahl	Beginn der Tätigkeit (Jahreszahl)
	vonMonat	Zahl	
	bisJahr	Zahl	Ende der Tätigkeit (Jahreszahl)
	bisMonat	Zahl	
	Verlauf	Text	Hinweis auf den Beginnzeitpunkt, z.B. 'seit', 'nach' ...
	Inf_ges	Text	Hinweis auf Vertrauenswürdigkeit der 'von'-Angabe
	Vlf	Text	Hinweis auf den Endzeitpunkt, z.B. 'bis', 'um'
	Inf-ges	text	Hinweis auf Vertrauenswürdigkeit der 'bis'-Angabe
	Tätigkeit	Text	Bezeichnung der Tätigkeit

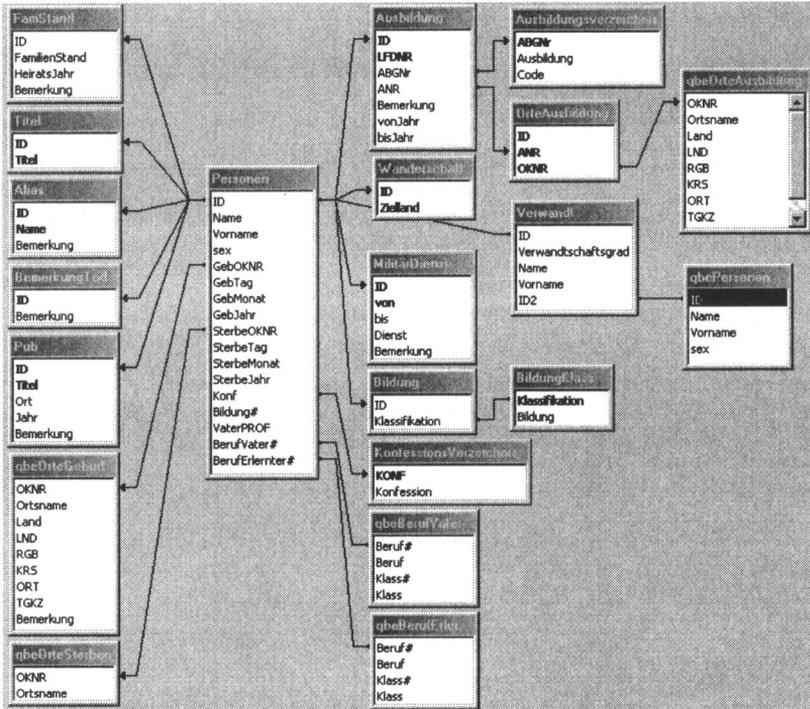
WerdegangTXT (Werdegang als Fließtext; 1:1- Beziehung zu Tabelle Personen)

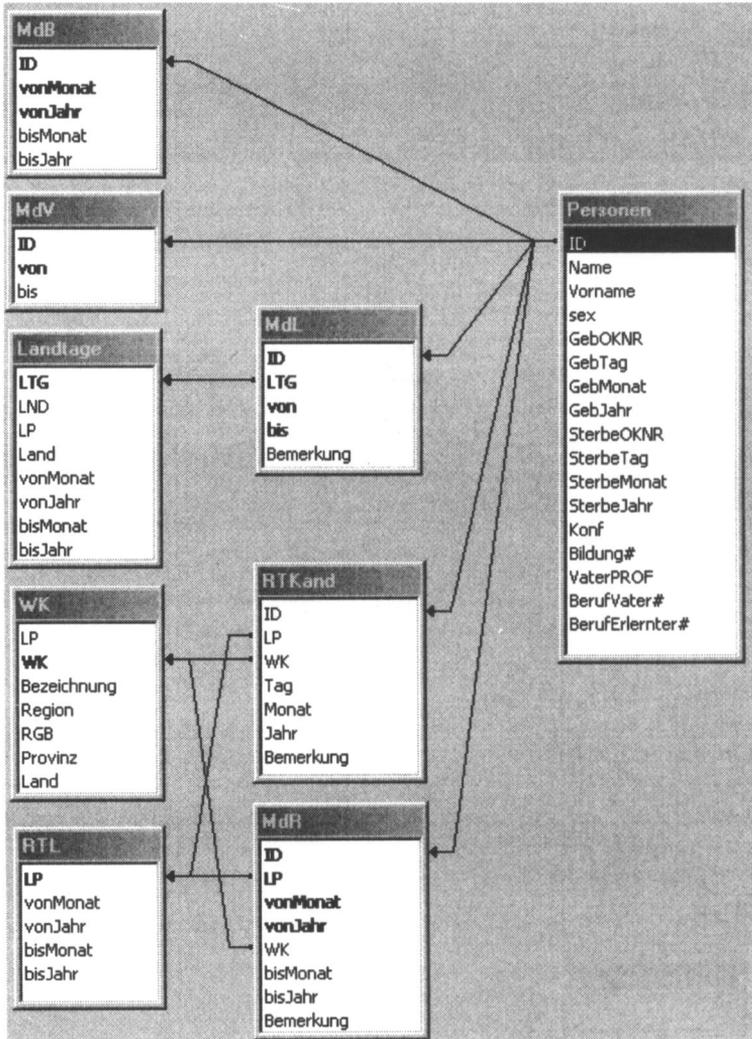
	Feldname	Felddatentyp	Beschreibung
<input type="checkbox"/>	ID	Zahl	Fremdschlüssel Personen; ID des Abgeordneten
	Werdegang	Text	Beschreibung der Tätigkeit wie im Handbuch

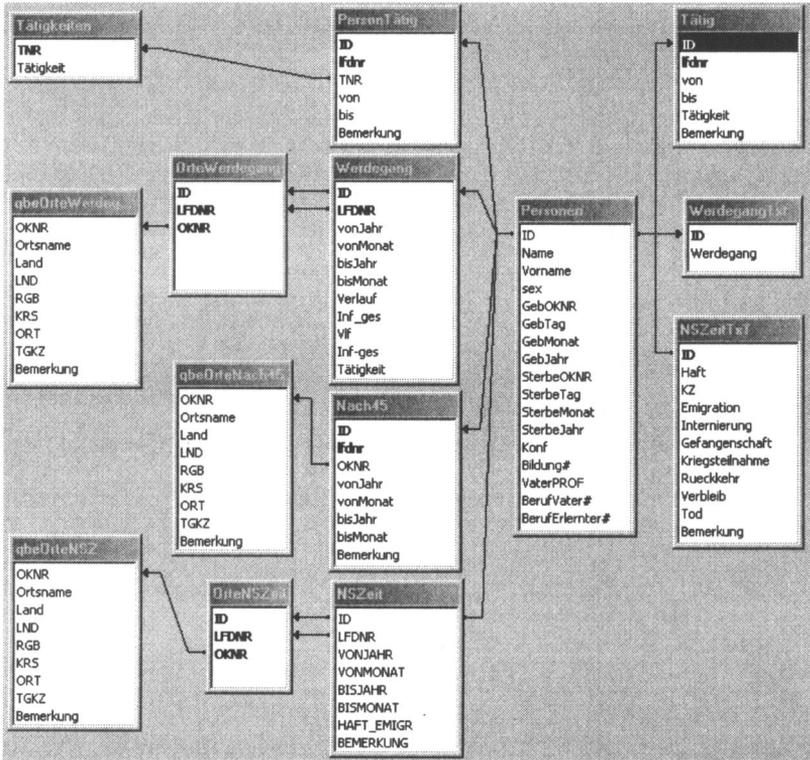
WK (Wahlkreise; veränderte Wahlkreise als verschiedene Wahlkreise betrachtet)

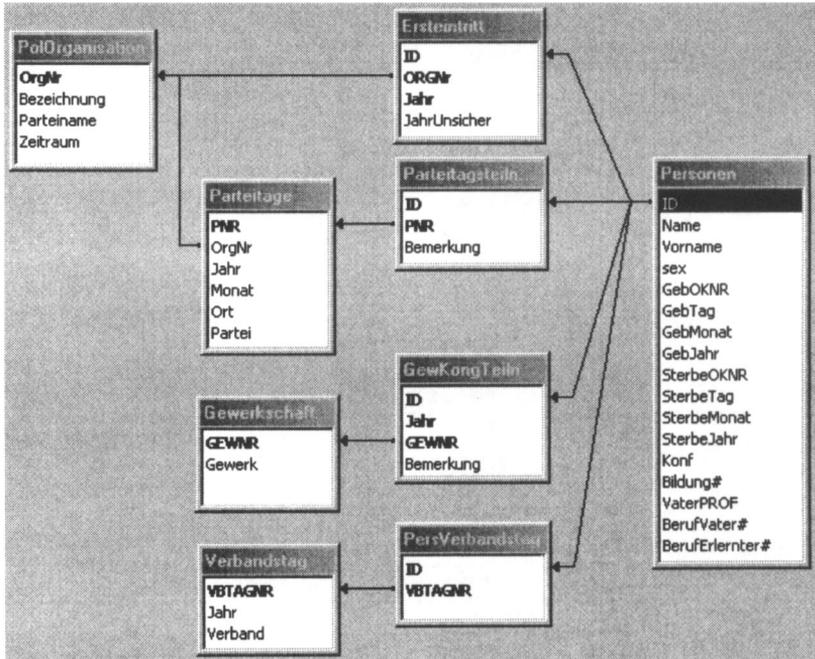
	Feldname	Felddatentyp	Beschreibung
<input type="checkbox"/>	WK	Zahl	Kennziffer des Wahlkreises
	LP	Text	Fremdschlüssel RTL; Nr. der Legislaturperiode
	Bezeichnung	Text	Wahlkreisname
	RGB	Zahl	Statistische Kennziffer Reg.bez.
	Land	Zahl	Statistische Kennziffer Land
	Provinz	Zahl	Statistische Kennziffer Provinz
	Region	Zahl	Auflistung von Kreisen / Teilregionen

3. Beziehungen zwischen den Tabellen









Anhang B: Übersicht: Übungsaufgaben im Text

Aufgabe 1.1: Versuchen Sie, sich den Unterschied zwischen den Begriffen 'Nachricht' und 'Information' anhand des folgenden Satzes klarzumachen:

„Diese Nachricht gibt mir keinerlei Information.“

Notieren Sie Ihre Gedanken.

Aufgabe 1.2: Was ist ein Datenbankmanagement?

- a. ein Datensatz
- b. ein Computerprogramm
- c. eine Datenbank mit dem dazugehörigen Datenbankverwaltungsprogramm

Wählen Sie die für Sie richtige Antwort aus und begründen Sie Ihre Entscheidung.

Aufgabe 1.3: Welche dieser Systeme klassifizieren Sie als Datenbankmanagementsystem?

- a. Oracle
- b. ACCESS
- c. dBase
- d. Excel

Wählen Sie die für Sie richtige Antwort aus und begründen Sie Ihre Entscheidung.

Aufgabe 1.4: Was sind Meta-Informationen?

- a. Spezifische Informationen
- b. Ein Maß für Informationsdichte
- c. Informationen über Informationen
- d. Fehlerhafte Daten

Wählen Sie die für Sie richtige Antwort aus und begründen Sie Ihre Entscheidung.

Aufgabe 3.1 Ist jede Tabelle eine Relation?

- a. Ja
- b. Nein

Wählen Sie die für Sie richtige Antwort aus und begründen Sie Ihre Entscheidung.

Aufgabe 3.2 Was sind NULL-Werte?

- a. die Zahl 0
- b. die Buchstabenfolge 'NULL'
- c. nicht bekannte oder nicht zutreffende Werte

Wählen Sie die für Sie richtige Antwort aus und begründen Sie Ihre Entscheidung.

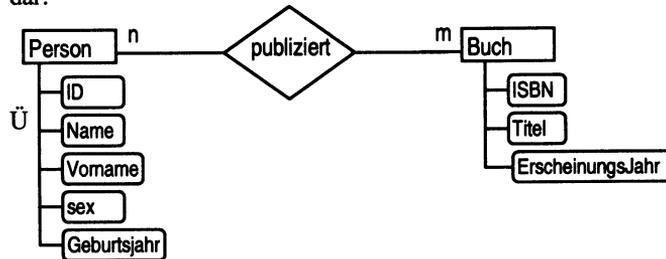
Aufgabe 3.3 Welchen Primärschlüssel für eine Tabelle mit den Grunddaten für Häftlinge halten Sie für geeignet?

- a. das Geburtsdatum
- b. Nachname und Vorname
- c. eine laufende Nummer
- d. keine von diesen

Wählen Sie die für Sie richtige Antwort aus und begründen Sie Ihre Entscheidung.

Aufgabe 5.1: Modellieren Sie die Publikationstätigkeit der Parlamentarier für den Fall, dass mehrere Parlamentarier gemeinsam als Buchautoren auftreten können.

Aufgabe 6.1: Das folgende ER-Diagramm stellt eine Lösung der Aufgabe 5.1 dar:



Überführen Sie das Diagramm in ein Relationenschema.

Aufgabe 8.1: Starten Sie ACCESS und legen Sie eine leere Datenbank mit dem Namen Parlamentarier an.

Aufgabe 8.2: Erstellen Sie eine Tabelle Personen mit den in 8.5.1 beschriebenen Datenfeldern.

Aufgabe 8.3: Geben Sie die Daten auf S. 106 in die Tabelle Personen ein.

Aufgabe 8.4: Definieren Sie einen geeigneten Primärschlüssel für die Tabelle Personen.

Aufgabe 8.5: Erstellen Sie das Formular mit dem Formularassistenten.

Aufgabe 8.6: Legen Sie für die Datenfelder Geburtstag, Geburtsmonat, Geburtsjahr, Sterbetag, Sterbemonat und Sterbejahr der Tabelle Personen Gültigkeitsregeln fest. Geben Sie für fehlerhafte Eingaben eine Fehlermeldung aus.

Aufgabe 8.7: Verändern Sie das in Aufgabe 8.5 erstellte Formular, indem Sie für Konfession ein Kombinationsfeld einrichten.

Aufgabe 8.8: Geben Sie den folgenden Datensatz ein:

ID	Name	Vorname	sex	Geburtsort	Geburtsdatum	Sterbeort	Sterbedatum	Konf.	Beruf
20020	Baade	Fritz	m	Neuruppin	23 1 1893	Kiel	15 5 1974	ev.	Seminardirektor

Aufgabe 8.9: Das Kombinationsfeld soll über weitere Werte verfügen. Die Liste soll erweitert werden um die Werte „keine Angabe“ und „nicht bekannt“.

Hinweis: Rechte Maustaste → EIGENSCHAFTEN → DATENSATZHERKUNFT und Anfügen der neuen Werte.

Aufgabe 8.10: Laden Sie die Übungsdatenbank SPDParl-AB.mdb von <http://www.hsr-trans.de>

Aufgabe 8.11: Erstellen Sie eine Abfrage auf der Grundlage der Tabelle Personen. Die Sortierung soll nach Geburtsjahr, Geburtsmonat, Geburtstag erfolgen.

Aufgabe 8.12: Erstellen Sie eine Abfrage über alle Mandate (Tabelle MDR), die neben den Feldern ID, LP, vonMonat, vonJahr noch den Namen, Vornamen und Geburtstag des Mandatsträgers enthält.

Aufgabe 8.13: Erstellen Sie eine Abfrage, die alle Reichstagsmandate der 20., 21. und 22. Legislaturperiode enthält.

Aufgabe 8.14: Ermitteln Sie alle Abgeordneten, die zwischen dem 70. und 80. Lebensjahr starben.

Aufgabe 8.15: Ermitteln Sie alle Katholiken, die dissidierten.

Aufgabe 8.16: Ermitteln Sie diejenigen Abgeordneten, die für die 24. Legislaturperiode dem Reichstag angehörten und vor 1945 starben.

Aufgabe 8.17: Folgende Altersklassen sind gegeben:

Altersklasse	Alter
1	unter 40
2	40 bis unter 50
3	50 bis unter 60
4	60 bis unter 70
5	70 und älter

Erstellen Sie eine Abfrage, die den Personen die entsprechende Altersklasse zuordnet.

Abbildungs- und Tabellenverzeichnis

Abb. 1.1: Datenbankbegriff	16
Abb. 1.2: Aufteilung in Software-Schichten nach Heuer/Saake (1995).....	17
Abb. 1.3: Arbeitsabläufe und ihre Widerspiegelung in einer Datenbank.....	19
Abb. 1.4: Phasenkonzept nach Lockemann/Schmidt (1987)	23
Abb. 1.5: Phasenkonzept nach Heuer/Saake (1995)	24
Abb. 2.1: Rekonstruktion von Verwandtschaftsbeziehungen	27
Abb. 3.1: Historische Einordnung der Datenbankmodelle nach Heuer/ Saake (1995).....	44
Abb. 3.2: Relationale Datenbank	48
Abb. 4.1: Darstellung eines Entity-Typs.....	57
Abb. 4.2: Darstellung von Relationships	58
Abb. 4.3: Landtagskandidaten	58
Abb. 4.4: Rollennamen	58
Abb. 4.5: Darstellung von Attributen.....	59
Abb. 4.6: Dreistellige Beziehung zwischen Professor, Vorlesung und Buch ..	59
Abb. 4.7: Dreistellige Beziehung als 2 zweistellige Beziehungen	60
Abb. 4.8: Mandatsträger als Ausschussmitglieder	61
Abb. 4.9: Optionale Beziehung.....	61
Abb. 4.10: Totale funktionale Beziehung	62
Abb. 5.1: Personen und Reichstage	64
Abb. 5.2: Personen, Reichstage und Wahlkreise	65
Abb. 5.3: Personen kandidieren in einem oder mehreren Wahlkreisen für Reichstage.....	65
Abb. 5.4: Personen nehmen Reichstagsmandat wahr.....	66
Abb. 5.5: Personen gehören Landtagen an.....	67
Abb. 5.6: Personen gehören nach 1945 Parlamenten an.	67
Abb. 5.7: Personen und Berufsausbildung	68
Abb. 5.8: Personen und Publikationstätigkeit	69
Abb. 5.9: Personen und Parteitage	69
Abb. 5.10: Werdegang von Abgeordneten.....	71
Abb. 5.11: Konfessionsbekenntnis.....	72
Abb. 5.12: Titel.....	72
Abb. 5.13: Familienstand.....	72
Abb. 5.14: Beruf des Vaters.....	72
Abb. 5.15: Weitere Namen	73
Abb. 5.16: Geburtsort	73
Abb. 5.17: Sterbeort.....	73
Abb. 5.18: Militärdienst.....	73
Abb. 5.19: Teilnahme an Gewerkschaftskongressen	74
Abb. 5.20: Teilnahme an Berufsverbandstagen	74

Abb. 5.21: Ersteintritt in eine sozialdemokratische Partei	74
Abb. 5.22: Erlebnis der NS-Zeit	74
Abb. 5.23: Entwicklung nach 1945.....	75
Abb. 6.1: Erhalt der Informationskapazität: $A = B$	76
Abb. 6.2: Transformation des ERM in Tabellen.....	77
Tabelle 6.1: Transformationsschema	77
Abb. 6.3: Personen mit ihren Attributen	78
Tabelle 6.2: Primärschlüssel der Tabellen von Beziehungstypen	79
Tabelle 6.3: Vereinfachungen bei zwingenden Beziehungen	79
Abb. 6.4: 1:1-Beziehung.....	80
Abb. 6.5: n:m-Beziehung.....	83
Abb. 6.6: 1:n-Beziehung.....	84