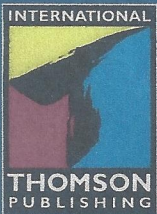
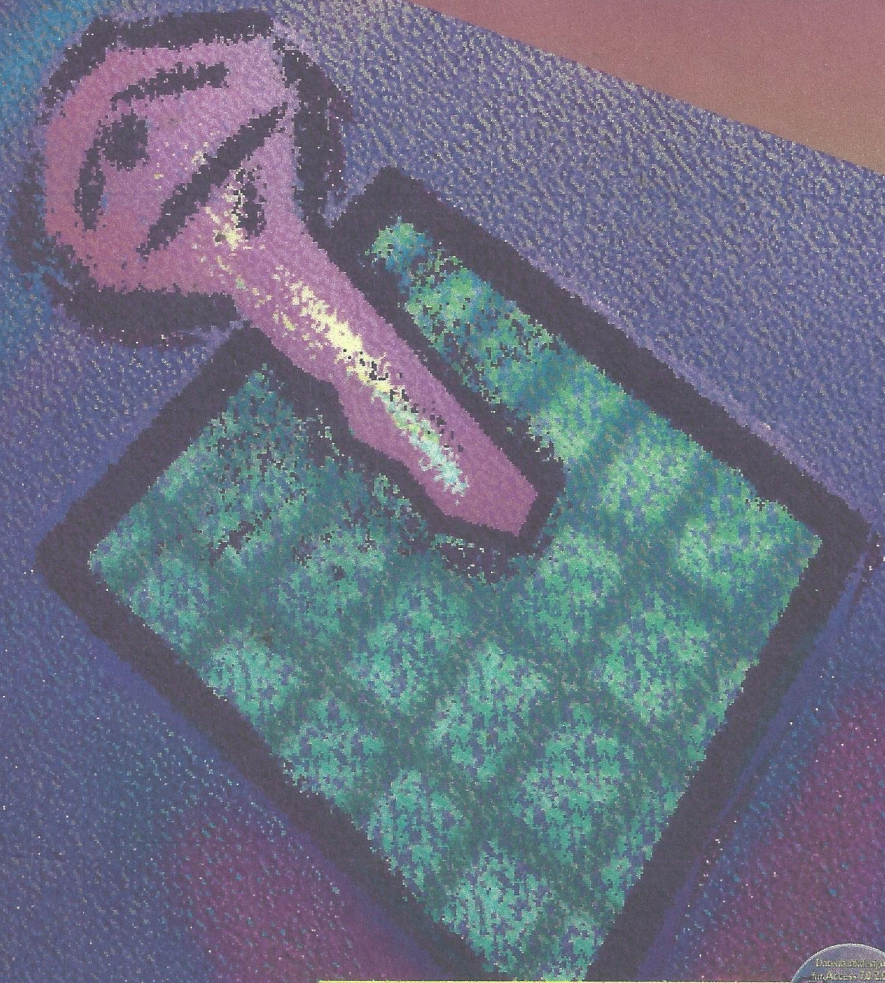


Manfred Sommer

Datenbankdesign für Access 7.0/2.0



**Inklusive Datenmodellierung
mit dem CASE-Tool ERwin**



Inhaltsverzeichnis



V	Warum man dieses Buch lesen sollte	11
	Wer dieses Buch lesen sollte	12
	Wie man dieses Buch nutzen sollte	13
	Was man auf der CD-ROM findet	14
	Danksagungen	15
	Über den Autor	16
1	Einleitung	17
1.1	Access als Windows-Datenbank	17
1.2	Access als relationale Datenbank	18
1.3	Access-Dateien und -Verzeichnisse	20
1.4	Access-Objekte und Namensgebung	24
1.5	Verschiedene Ansichten von Access-Objekten	30
2	Aufbau und Arbeitsweise von Access-Tabellen	33
2.1	Allgemeine Eigenschaften von Tabellen	33
2.2	Identifizierbarkeit und Auswahl des Primärschlüssels	37
2.3	Eindeutigkeit von Sekundärschlüsseln	54
2.4	Eingabepflicht	59
2.4.1	Eingabepflicht bei deskriptiven Attributen	59
2.4.2	Eingabepflicht bei Primärschlüsselattributen	64
2.4.3	Eingabepflicht bei Sekundärschlüsselattributen	65
2.4.4	Eingabepflicht bei Fremdschlüsselattributen	67
2.4.5	Nullwerte und leere Zeichenfolgen	67
2.5	Attributnamen	71
2.6	Domänen	81
2.6.1	Felddatentypen	81
2.6.2	Feldgröße	85

Inhalt		
2.6.3	Standardwerte	85
2.6.4	Gültigkeitsregeln und -meldungen	90
2.6.5	Nachschlagfelder	96
2.7	Beschreibung von Feldern und Tabellen	99
3	Nachteile von Redundanzen in Tabellen	101
3.1	Einfügeanomalien	103
3.2	Änderungsanomalien	106
3.3	Löschanomalien	107
3.4	Ein intuitiver Weg zur Redundanzfreiheit	108
4	Grundlagen der Entity-Relationship-Modellierung	115
4.1	Einordnung in die Drei-Ebenen-Architektur	115
4.2	Redundanzfreiheit versus Performance	118
4.3	Datenbankentwurf und Anwendungsentwicklung als Modellbildungsprozeß	121
4.4	Verhältnis der ER-Modellierung zur Normalisierung	122
4.5	ERM-Grundbegriffe: Entitäten, Attribute und Beziehungen	127
5	Zusammengesetzte Attribute und Mehrfachattribute	131
5.1	Zusammengesetzte Attribute	131
5.2	Mehrfachattribute	133
5.2.1	Denormalisierung	133
5.2.2	Zeilenweise Modellierung	134
5.2.3	Spaltenweise Modellierung	135
5.2.4	Vergleich der zeilenweisen und der spaltenweisen Modellierung	136
6	Abgeleitete (berechnete) Attribute	149
6.1	Berechnete Attribute mit konstanten Ausdrücken	149
6.2	Beispiel: Microsofts Probleme mit berechneten Attributen	153

	<u>Inhalt</u>	
6.3	Berechnete Attribute mit veränderlichen Ausdrücken	155
6.4	Performanceaspekte berechneter Attribute	158
7	Entitäts- und Beziehungsarten	161
7.1	Unabhängige und abhängige Entitäten; nicht-identifizierende und identifizierende Beziehungen	161
7.2	Charakterisierende und assoziative Entitäten	164
7.3	Zusammengesetzter Primärschlüssel versus Surrogatschlüssel	169
8	Hintergrundentitäten	177
8.1	Eingabekomfort	181
8.2	Eingabesicherheit	184
8.3	Performance	190
8.4	Zusammenfassung	201
8.5	Mehrspaltige Hintergrundtabellen für wechselseitig abhängige Domänen	204
9	Kardinalitäten und referentielle Integrität	207
9.1	Maximal- und Minimalkardinalitäten	207
9.1.1	Fünf Kardinalitätskonstellationen	208
9.1.2	Analyse der Ist-Kardinalitäten	209
9.1.3	Von Ist- zu Soll-Kardinalitäten	214
9.2	Referentielle Integritätsregeln	215
9.2.1	Referentielle Integrität ohne Aktualisierungs- und Löschweitergabe	217
9.2.2	Referentielle Integrität mit Aktualisierungswweitergabe	222
9.2.3	Referentielle Integrität mit Löschweitergabe	222
9.2.4	Referentielle Integrität mit Aktualisierungs- und Löschweitergabe	223
9.2.5	Referentielle Integrität mit nullifizierendem Aktualisieren und Löschen	224
9.2.6	Referentielle Integritätsregeln bei eingebundenen Tabellen	225
9.3	Zusammenhänge zwischen Kardinalitäten und RI-Regeln	227
9.4	Zusammenhänge zwischen Entitätsarten und RI-Regeln	228

10	Modellierung von Subtypen	231
10.1	Super- und Subtypen in IS-A-Beziehungen	231
10.2	Vollständigkeit und Disjunktheit	233
10.3	Supertyp-Lösung, Subtypen-Lösung und Supertyp-Subtypen-Lösung	234
10.4	Vor- und Nachteile der drei Lösungen	236
10.4.1	Transparenz und Flexibilität des Datenmodells	236
10.4.2	Abfragemöglichkeiten und -komfort	238
10.4.3	Datenintegrität beim Eingeben, Aktualisieren und Löschen	248
10.5	Empfehlungen zur Lösungswahl	259
10.6	Beispiel: Microsofts Probleme mit Subtypen	261
11	Grenzen der Datenmodellierungs-Assistenten	263
11.1	»Grundwissen über Datenbankmodellierung nicht erforderlich«	263
11.2	Der Irrweg: die Datenmodellierungs-Assistenten	265
11.2.1	Der Datenbank-Assistent	266
11.2.2	Der Tabellen-Assistent	270
11.2.3	Der Datenblatt-Assistent	272
11.2.4	Assistentenfreies Entfernen, Hinzufügen und Anpassen von Feldern	273
11.2.5	RI-Regeln des Datenbank-Assistenten	279
11.2.6	»Voll Normal«? – Der Tabellenanalyse-Assistent	280
11.3	Die Alternative: CASE-Tool-gestützte Datenmodellierung	285
12	Kurzeinführung in die ERwin-Bedienung	287
12.1	Voreinstellungen und Anzeigeoptionen der ERwin-Benutzeroberfläche	288
12.1.1	Voreinstellungen mit dem Menü Option	288
12.1.2	Anzeigeoptionen mit dem Menü Display	289
12.1.3	Mehrere Arbeitsbereiche und verschiedene gespeicherte Ansichten mit dem Menü Window	293
12.2	Die Editoren für das konzeptionelle Datenmodell	293
12.2.1	Der Entity Attribute Editor	294
12.2.2	Die Beziehungs-Editoren	295

	Inhalt	
12.3	Die zieldatenbankabhängigen Editoren	296
12.3.1	Der Referential Integrity Default Editor	298
12.3.2	Der Column Property Editor, der Validation Rule Editor und der Default Editor	299
12.4	Das Anlegen von Access-Tabellen mit ERwin	302
	Abbildungsverzeichnis	307
	Tabellenverzeichnis	315
	Abkürzungsverzeichnis	319
	Literaturverzeichnis	321
I	Index	325

Warum man dieses Buch lesen sollte



An Büchern zu Access einerseits und zu den theoretischen Grundlagen relationaler Datenbanken andererseits herrscht heute kein Mangel mehr. Nach meinem persönlichen Eindruck kommt aber in den meisten Access-Büchern, die ja größtenteils mehr oder weniger gelungene »Ersatzhandbücher« sind, die Frage, **wie eine Access-Datenbank zu strukturieren ist**, entschieden zu kurz. Das gilt auch für das zweite Kapitel »Entwerfen einer Datenbank« des von Microsoft mitgelieferten Benutzerhandbuchs. Mag sich der Leser auch noch soviel **Bedienungs-wissen** über Access angeeignet haben: der Entwurf der ersten eigenen Datenbank, deren Komplexität über die Verwaltung seiner Briefmarkensammlung minimal hinausgeht, wird ihn vor erhebliche Probleme stellen. Ohne **konzeptionelles Wissen** zum Datenbankentwurf sind diese kaum oder nur so unbefriedigend zu lösen sind, daß derart überstürzt und ad hoc entwickelte Datenbanken schon nach kurzer Nutzungszeit von Grund auf überarbeitet werden müssen.

Noch bedenklicher als die meist äußerst sparsame Thematisierung der Datenmodellierungsaspekte in Access-Büchern ist die Tatsache, daß die dabei zu beantwortenden Fragen leider häufig falsch abgehandelt werden. Mit fehlerhaft strukturierten Beispiel-Datenbanken und abwegigen Begründungen hierfür ließe sich ein ganzes Buch »Wie man Access-Datenbanken nicht entwerfen sollte« unschwer füllen. Als höflicher und positiv denkender Autor ziehe ich eine Darstellung, wie man es »richtig« machen sollte, vor – wohl wissend, daß es die eine, allgemein richtige Lösung häufig nicht gibt, sondern diese je nach den Anforderungen eines spezifischen Datenbankprojekts im Einzelfall zu suchen ist. Die dabei zu berücksichtigenden Aspekte sind ein Kernthema dieses Buches.

Ich greife bei der Darstellung zentraler Fragen des Datenbank-Designs einerseits auf didaktische Erfahrungen aus Lehrveranstaltungen zur Wirtschaftsinformatik und aus unternehmensorientierten Datenbank-Kursen (insb. zu Access) sowie andererseits auf Praxiserfahrungen aus kleinen und großen datenbankgestützten Softwareentwicklungsprojekten zurück. Natürlich geschieht dies unter Hinzuziehung einschlägiger theoretischer Literatur zu relationalen Datenbanken. Dies wirft die Frage auf, ob es denn nicht ausreiche, dem geneigten Leser eine Liste der Originalliteratur über das World Wide Web zur Verfügung zu stellen mit der Empfehlung, sich an der Quelle zu laben. Meine Erfahrung ist es aber, daß viele Access-Interessierte den Zugang zu dieser Literatur nicht finden, weil sie ihnen zu »abgehoben« vorkommt und sie die konkrete Umsetzbarkeit in

Access-Projekte vermissen. Der Clou dieses Buches ist es deshalb, genau diese **Verbindungsline von der Theorie des Datenbankentwurfs zur Software Access** zu ziehen.

Obwohl Access in diesem Buch die Hauptrolle besetzt, spielt ein weiteres Softwareprodukt eine anspruchsvolle Nebenrolle: das Datenmodellierungswerkzeug ERwin der Fa. Logic Works GmbH. Der Name signalisiert, daß es sich um ein CASE-Tool zur Entity-Relationship-Modellierung unter Windows handelt. Entity-Relationship (ER) ist heute der am weitesten verbreitete methodische Ansatz zur Datenmodellierung. Ich habe zuerst gezögert, ein solches CASE-Tool (CASE: computer aided software engineering = computergestützte Softwareentwicklung) mit in ein Access-Buch aufzunehmen. Drei Gründe waren hierfür ausschlaggebend:

- Datenmodellierungswerkzeuge wie z. B. ERwin werden zunehmend selbst in die semi-professionelle Datenbankentwicklung und mit weiter sinkenden Preisen dieser Tools vielleicht sogar in das home computing Eingang finden.
- Von ERwin gibt es eine Demo-Version, die es Ihnen erlaubt, mit den Beispieldateien dieses Buches zu experimentieren. Beides finden Sie auf der anliegenden CD-ROM.
- Datenmodellierungswerkzeuge wie ERwin sind den neuen Datenmodellierungs-Assistenten von Access 7.0, die mit völlig überzogenen Erwartungen vermarktet werden, turmhoch überlegen. Diesem Thema ist eigens das abschließende Kapitel 11 gewidmet.

Wer dieses Buch lesen sollte

Das Buch wendet sich an Studierende, Lernende in der Aus- und Weiterbildung sowie an Praktiker, die sich gründlicher mit Fragen des Entwurfs relationaler Datenbanken im Allgemeinen und der Realisierung in Access im Besonderen beschäftigen. Als Zielgruppe sind auch all jene willkommen, die Datenbankprojekte

- wegen eines Plattformwechsels (z. B. von einer xBASE-Datenbank zu Access oder zu einem SQL-Server),
- wegen eines problematischen Datenbankdesigns oder
- wegen eines unangebrachten Vertrauensvorschlusses in die Access-Datenmodellierungs-Assistenten (die jüngste, aber sicher schnell wachsende Klientel)

von Grund auf neu konzipieren möchten.

Deshalb sollte die Lektüre auch für den Leser und natürlich auch für die Leserin von Nutzen sein, die Datenbanken gar nicht mit Access, sondern mit einem

anderen relationalen Datenbankmanagementsystem zu entwickeln beabsichtigen. Für diesen Leserkreis verwende ich Access also weniger als prospektives Datenbanksystem, sondern quasi als Datenbank-Lernprogramm.

Wie man dieses Buch nutzen sollte

Wie in Vorworten häufig betont, gilt auch hier der Hinweis, daß dieses Buch nicht nur wie ein Lehrbuch linear vom Anfang bis zum Ende lesbar ist, sondern auch Quereinstiege zuläßt und als spezielles Vertiefungs- und Nachschlagewerk auch kapitelweise nutzbar sein sollte. Da es weder Access- noch Datenmodellierungskenntnisse voraussetzt, ist es auch für Newcomer geeignet. Einige Hinweise mögen den fortgeschrittenen Leserinnen und Lesern die selektive Nutzung dieses Buches erleichtern.

Die Einleitung wendet sich nur an absolute Access-Neulinge. Kapitel 2 (Aufbau und Arbeitsweise von Access-Tabellen) breitet alles Wissenswerte rund um Access-Tabellen aus. Erfahrene Access-Benutzer werden in diesem Kapitel sicher manche vertraute Passage überschlagen wollen. Nachschlagfelder als eine wesentliche Neuerung auf Tabellenebene werden im Abschnitt 2.6.5 nur kurz vorgestellt, dafür im Kap. 8 um so ausführlicher behandelt. Die Kapitel 3 (Nachteile von Redundanzen in Tabellen) und 4 (Grundlagen der Entity-Relationship-Modellierung) bereiten das Basiswissen für eine solide Datenmodellierung auf, von dem jedes konkrete Datenbankprojekt profitieren sollte. Wer mit diesen Grundlagen vertraut ist, kann sich gleich eines der vertiefenden Kapitel 5 bis 10 vornehmen, die sich mit Problemen befassen, die in vielen Datenbanken zu lösen sind:

- Bildet man Mehrfachattribute (z. B. die Monate eines Jahres) besser in den Spalten oder in den Zeilen einer Tabelle ab? (Kapitel 5)
- Soll man Attribute, die sich eigentlich jederzeit aus anderen Attributen berechnen lassen (z. B. den Bestellwert eines Artikels als Produkt aus Menge und Preis), in der Datenbank speichern oder nicht? (Kapitel 6)
- Was spricht für und was gegen zusammengesetzte Primärschlüssel? (Kapitel 7)
- Wie stellt man am besten sicher, daß bestimmte Informationen in Tabellen überall gleichartig und widerspruchsfrei auftreten? (Kapitel 8)
- Was ist referentielle Integrität und wie kann man sie in Access sinnvoll nutzen? (Kapitel 9)

- Wie geht man am besten mit der Tatsache um, daß sich Informationsobjekte manchmal einerseits sehr ähneln, aber andererseits Besonderheiten aufweisen, die man nicht einfach ignorieren kann (z. B. Privatkunden und Geschäftskunden)? (Kapitel 10)

Kapitel 11 (»Grundwissen über Datenbankmodellierung nicht erforderlich«) bringt die Grundintentionen dieses Buches mit einer Kritik an der um sich greifenden Assistentengläubigkeit abschließend noch einmal auf den Punkt.

Datenbankdesign steht insbesondere bei Praktikern gelegentlich in dem unverdient schlechten Ruf, zu theoretisch, deshalb eigentlich überflüssig und vor allem strohtrocken zu sein. Ich habe mich beim Schreiben deshalb von der Idee leiten lassen, daß ein gewisses Lesevergnügen dem Erkenntnisgewinn nicht schaden muß.

Wo immer sinnvoll wird im Text auf ergänzende Informationen der sehr ausführlichen und der regen Nutzung anempfohlenen Online-Hilfe von Access verwiesen. Die schlichte Wiedergabe von Hilfeinformationen auf Buchseiten halte ich für überholt, zumal Sie sich diese Informationen, wenn sie wirklich einmal in Papierform statt auf dem Bildschirm gewünscht werden, jederzeit ausdrucken können. Die Hinweise auf die Hilfethemen stehen kursiv in Klammern, werden jeweils durch *Hilfe:* eingeleitet und orientieren sich überwiegend an der Access 2.0-Hilfe, weil das Hilfesystem zu der mir bei der Endredaktion des Textes vorliegenden Beta-Version von Access 7.0 noch sehr unübersichtlich war.

Hinweise auf Access-Menüs sind in Kapitälchen formatiert, wobei Menüpunkte durch einen Punkt abgetrennt werden (Beispiel: DATEI.DATENBANK ÖFFNEN...).

Was man auf der CD-ROM findet

Daß sich dieses Buch nicht nur an Interessenten für Access 7.0 wendet, sondern genauso gut – und mit nicht weniger Gewinn – von Access 2.0-Anwendern gelesen werden kann, macht sich nicht nur an den entsprechenden Randsymbolen im Text bemerkbar, die auf Besonderheiten der älteren Versionen aufmerksam machen. Auf der anliegenden CD-ROM sind sämtliche Access-Beispieldateien (*.MDB) wegen des veränderten Datenbankformats sowohl in einer Version für Access 7.0 wie auch für Access 2.0 enthalten, wenn auch die 2.0-Dateien nicht in jedem Detail exakt mit der aktuellen Textfassung übereinstimmen. Die Datenbanken sind in Anlehnung an die Kapitelnumerierung mit ITP_01.MDB, ITP_02.MDB etc. benannt; werden in einem Kapitel mehrere Datenbanken verwendet, sind sie durch einen angehängten Buchstaben unterschieden (z. B. ITP_01A.MDB und ITP_01B.MDB). Diese Datenbanken sind keine Fertiggerichte zum Sofortverzehr. Man sollte sie als einen Steinbruch betrachten, aus dem man sich nach Bedarf gerade das an Know-how herausbricht, was man für das anste-