

1-8 Veritabanı Uzmanlığı Kariyerine Hazırlık

Bu bölümde, veri, bilgi, veritabanları ve VTYS (Veritabanı Yönetim Sistemleri) kavramlarıyla tanıştınız. Ayrıca, hangi tür veritabanını kullandığınızdan (OLTP, OLAP veya NoSQL) ya da hangi veritabanı ortamında çalıştığınızdan (örneğin, Oracle, Microsoft, IBM veya Hadoop) bağımsız olarak, bir veritabanı sisteminin başarısının büyük ölçüde veritabanı yapısının ne kadar iyi tasarlandığına bağlı olduğunu öğrendiniz.

Bu kitap boyunca, veritabanı uzmanı olarak kariyerinizin temelini oluşturan yapı taşlarını öğreneceksiniz. Bu yapı taşlarını anlamak ve bunları etkili bir şekilde kullanma becerisi geliştirmek, organizasyon içinde farklı seviyelerde veritabanlarıyla çalışmaya sizi hazırlayacaktır. Böyle bir kariyer fırsatlarının küçük bir örneği Tablo 1.3'te gösterilmektedir.

Table 1.3 Database Career Opportunities

İş Unvanı	Açıklama	Gerekli Örnek Beceriler
Veritabanı Geliştiricisi	Veritabanı tabanlı uygulamalar oluşturma ve bakımı yapma	Programlama, veritabanı temelleri, SQL
Veritabanı Tasarımcısı	Veritabanılarını tasarlama ve bakımını yapma	Sistem tasarımı, veritabanı tasarımı, SQL
Veritabanı Yöneticisi	VTYS ve veritabanılarını yönetme ve bakımını yapma	Veritabanı temelleri, SQL, satıcı kursları
Veritabanı Analisti	Karar destek raporlaması için veritabanları geliştirme	SQL, sorgu optimizasyonu, veri ambarları
Veritabanı Mimarı	Veritabanı ortamlarının tasarlanması ve uygulanması	VTYS temelleri, veri modelleme, SQL, donanım bilgisi
Veritabanı Danışmanı	Şirketlerin iş süreçlerini iyileştirmek ve belirli hedeflere ulaşmak için veritabanı teknolojilerinden yararlanmalarına yardımcı olma	Veritabanı temelleri, veri modelleme, veritabanı tasarımı, SQL, VTYS, donanım, satıcıya özel teknolojiler, vb.
Veritabanı Güvenlik Görevlisi	Veri yönetimi için güvenlik politikalarını uygulama	VTYS temelleri, veritabanı yönetimi, SQL, veri güvenliği teknolojileri, vb.
Bulut Bilişim Veri Mimarı	Bulut tabanlı yeni nesil veritabanı sistemleri için altyapı tasarlama ve uygulama	İnternet teknolojileri, bulut depolama teknolojileri, veri güvenliği, performans ayarlama, büyük veritabanları, vb.
Veri Bilimci	Büyük ve çeşitli veri kümelerini analiz ederek içgörüler, ilişkiler ve öngörülebilir davranışlar oluşturma	Veri analizi, istatistik, ileri düzey matematik, SQL, programlama, veri madenciliği, makine öğrenimi, veri görselleştirme

Bu bölümde öğrendiğiniz gibi, veritabanı teknolojileri, veri miktarının artması (büyük veritabanları), giderek çeşitlenen veri türleri (yarı yapılandırılmış ve yapılandırılmamış veriler) ve artan işleme karmaşıklığı ve hızı gibi yeni zorluklarla başa çıkmak için sürekli olarak evrilmektedir. Veritabanı teknolojileri hızla değişebilse de, temel kavramlar ve beceriler değişmez. Bu kitapta veritabanı temellerini öğrendikten sonra, bilgilerinizi ve becerilerinizi geleneksel OLTP ve OLAP sistemlerinin yanı sıra aşağıdaki gibi ileri düzey, karmaşık veritabanı teknolojileriyle çalışmak için uygulamaya hazır olmanızı sağlamayı hedefliyoruz:

- Çok Büyük Veritabanları (VLDB'ler). Birçok satıcı, genellikle petabayt aralığındaki büyük veri miktarlarını destekleyen veritabanları ihtiyacını karşılamaktadır. (Bir petabayt, 1.000 terabayttan daha fazladır.) VLDB satıcıları arasında Oracle Exadata, IBM Netezza, HP Vertica ve Teradata bulunmaktadır. VLDB'ler şimdi, Büyük Veri veritabanları tarafından yerini almaktadır.
- Büyük Veri Veritabanları. Cassandra (Facebook) ve BigTable (Google) gibi ürünler, büyük miktarda "tablo dışı" veriyi yöneten veritabanı uygulamalarının ihtiyaçlarını karşılamak için "sütunlu-veritabanı" teknolojilerini kullanmaktadır. Bu konu hakkında daha fazla bilgiye 2. Bölümde ulaşabilirsiniz.
- Bellek içi veritabanları. Çoğu büyük veritabanı satıcısı, daha hızlı veritabanı işleme ihtiyacını karşılamak için bazı türde bellek içi veritabanı desteği de sunmaktadır. Bellek içi veritabanları, verilerini daha yavaş ikincil depolama (sabit diskler) yerine birincil bellekte (RAM) saklar. Bellek içi veritabanlarına örnek olarak IBM solidDB ve Oracle TimesTen verilebilir.

- Bulut Veritabanları. Şirketler artık bulut veritabanı hizmetlerini kullanarak veritabanı sistemlerini çevrelerine hızla ekleyebilirler, aynı zamanda yeni bir VTYS'nin toplam sahip olma maliyetini de düşürebilirler. Bir bulut veritabanı, yerel bir VTYS'nin tüm avantajlarını sunar, ancak organizasyonunuzun ağ altyapısında bulunmak yerine İnternet üzerinde yer alır. Bu konu hakkında daha fazla bilgiye 15. Bölümde ulaşabilirsiniz.

Bu kitap, bu konuların bazılarını değinmektedir, ancak hepsine—tek bir kitap, veritabanı teknolojilerinin tüm alanını kapsayamaz. Bu kitabın ana odak noktası, veritabanı temellerini öğrenmenize yardımcı olmak, veritabanı tasarımı becerilerinizi geliştirmenizi sağlamak ve SQL becerilerinizi ustalaştırmak, böylece başarılı bir veritabanı profesyoneli olma yolunda bir adım önde başlamanızı sağlamaktır. Ancak, önce kullanabileceğiniz araçları öğrenmeniz gerekmektedir. Bir sonraki bölümde, veri yönetimi için farklı yaklaşımları ve bu yaklaşımların tasarımlarınızı nasıl etkilediğini öğreneceksiniz.

Özet

- Veri, ham gerçeklerden oluşur. Bilgi, verinin işlenmesi sonucu anlamının ortaya çıkmasıdır. Doğru, ilgili ve zamanında bilgi, iyi karar verme için anahtar rol oynar ve iyi karar verme, küresel bir ortamda organizasyonların hayatta kalmasının anahtarıdır.

- Veri genellikle bir veritabanında saklanır. Bir veritabanı oluşturmak ve içeriğini yönetmek için bir veritabanı yönetim sistemine (VTYS) ihtiyacınız vardır. VTYS, kullanıcı ile veritabanı arasında aracılık yapan bir sistemdir. Veritabanı, topladığınız verileri ve "veri hakkında veri"yi, yani metadata'yı içerir.

- Veritabanı tasarımı, veritabanı yapısını tanımlar. İyi tasarlanmış bir veritabanı, veri yönetimini kolaylaştırır ve doğru ve değerli bilgiler üretir. Kötü tasarlanmış bir veritabanı, kötü kararlar alınmasına yol açabilir ve kötü kararlar, bir organizasyonun başarısız olmasına neden olabilir.

- Veritabanları, desteklenen kullanıcı sayısına, verilerin konumuna, depolanan veri türüne, verilerin kullanım amacına ve verilerin yapılandırılma derecesine göre sınıflandırılabilir.

- Veritabanları, manuel ve ardından bilgisayarla yapılmış dosya sistemlerinden evrilmiştir. Bir dosya sisteminde, veriler bağımsız dosyalarda saklanır.

Her biri kendi veri yönetim programlarına ihtiyaç duyan dosyalar. Bu veri yönetimi yöntemi büyük ölçüde modası geçmiş olsa da, özelliklerini anlamak, veritabanı tasarımını kavramayı kolaylaştırır.

- Dosya sistemi veri yönetiminin bazı sınırlamaları şunlardır: kapsamlı programlama gerektirir, sistem yönetimi karmaşık ve zor olabilir, mevcut yapıları değişiklik yapmak zordur ve güvenlik özellikleri muhtemelen yetersizdir. Ayrıca, bağımsız dosyalar genellikle gereksiz veri içerir, bu da yapı ve veri bağımlılığı sorunlarına yol açar.

- VTYS'ler, dosya sisteminin doğasında bulunan zayıflıkları gidermek için geliştirilmiştir. Veriyi bağımsız dosyalara koymak yerine, bir VTYS veritabanını son kullanıcıya tek bir veri deposu olarak sunar. Bu düzenleme, veri paylaşımını teşvik eder ve böylece bilgi adaları olasılığını ortadan kaldırır. Ayrıca, VTYS veri bütünlüğünü sağlar, gereksizliği ortadan kaldırır ve veri güvenliğini teşvik eder.

- Veritabanı teknolojileri bilgisi, hızla genişleyen BT endüstrisinde birçok kariyer fırsatına yol açar. Veritabanı alanında, geniş bir beceri ve uzmanlık yelpazesi için çeşitli uzmanlaşma alanları bulunmaktadır.

Anahtar Kelimeler

Ad hoc sorgu

Analitik veritabanı

Uygulama programlama

arayüzü (API)

İş zekası

Merkezi veritabanı

Bulut veritabanı

Veri

Veri anomalisi

Veri bağımlılığı

Veri sözlüğü

Veri tutarsızlığı

Veri bağımsızlığı

Veri bütünlüğü

Veri yönetimi

30 Part 1: Database Concepts

Veri işleme (DP) uzmanı	Dosya	Sorgu
Veri kalitesi	Genel amaçlı veritabanı	Sorgu dili
Veri çoğaltması	Bilgi	Sorgu sonuç kümesi
Veri tipi	Bilgi adaları	Kayıt
Veri ambarı	Bilgi	Yarı yapılandırılmış veri
Veritabanı	Mantıksal veri formatı	Tek kullanıcı veritabanı
Veritabanı tasarımı	Metadata	Sosyal medya
Veritabanı yönetim sistemi (VTYS)	Çoklu kullanıcı veritabanı	Yapısal bağımlılık
Veritabanı sistemi	NoSQL	Yapısal bağımsızlık
Masaüstü veritabanı	Çevrimiçi analitik işleme (OLAP)	Yapılandırılmış veri
Disiplin spesifik veritabanı	Çevrimiçi işlem işleme (OLTP)	Yapılandırılmış Sorgu Dili (SQL)
Dağıtık veritabanı	veritabanı	İşlemsel veritabanı
Kurumsal veritabanı	Operasyonel veritabanı	Yapılandırılmamış veri
Genişletilebilir İşaretleme Dili (XML)	Performans ayarları	Çalışma grubu veritabanı
Alan	Fiziksel veri formatı	XML veritabanı
	Üretim veritabanı	

Review Questions

- Aşağıdaki terimlerin her birini tanımlayın:
 - data
 - field
 - record
 - file
- Veri çoğaltması (data redundancy) nedir ve dosya sisteminin hangi özellikleri buna yol açabilir?
- Veri bağımsızlığı (data independence) nedir ve neden dosya sistemlerinde eksiktir?
- Bir VTYS (Veritabanı Yönetim Sistemi) nedir ve işlevleri nelerdir?
- Yapısal bağımsızlık nedir ve neden önemlidir?
- Veri, bilgi ve veritabanı arasındaki farkları açıklayın.
- Bir VTYS'nin rolü nedir ve avantajları nelerdir? Dezavantajları nelerdir?
- Farklı veritabanı türlerini listeleyin ve açıklayın.
- Bir veritabanı sisteminin ana bileşenleri nelerdir?
- Metadata nedir?
- Veritabanı tasarımı neden önemlidir?
- Bir veritabanı sistemi uygulamanın potansiyel maliyetleri nelerdir?
- Yapılandırılmamış ve yapılandırılmış veriyi örneklerle karşılaştırın. Tipik bir iş ortamında hangi tür daha yaygındır?
- Bir elektronik tablonun gerçekleştiremeyeceği bazı temel veritabanı işlevleri nelerdir?
- Son kullanıcılar tarafından oluşturulan elektronik tabloların, tipik bir dosya sistemiyle paylaştığı yaygın sorunlar nelerdir?
- Bilgisayarlı veri depolarının ortaya çıkmasıyla birlikte, son kullanıcıların iş verilerine doğrudan erişimini kaybetmesinin önemini açıklayın.
- Bir bulut veritabanının sahip olma maliyetinin, geleneksel bir şirket veritabanına göre neden daha düşük olabileceğini açıklayın.

Problemler

Lütfen Şekil P1.1'i paylaşın veya açıklayın, böylece sorularınızı yanıtlamanıza yardımcı olabilirim.

Figure P1.1 The File Structure for Problems 1–4

PROJECT_CODE	PROJECT_MANAGER	MANAGER_PHONE	MANAGER_ADDRESS	PROJECT_BID_PRICE
21-5Z	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	16833460.00
25-2D	Jane D. Grant	615-898-9909	218 Clark Blvd., Nashville, TN 36362	12500000.00
25-5A	George F. Dorts	615-227-1245	124 River Dr., Franklin, TN 29185	32512420.00
25-9T	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	21563234.00
27-4Q	George F. Dorts	615-227-1245	124 River Dr., Franklin, TN 29185	10314545.00
29-2D	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	25559999.00
31-7P	William K. Moor	904-445-2719	216 Morton Rd., Stetson, FL 30155	56850000.00

Çevrimiçi İçerik

Bu problem setinde gördüğünüz dosya yapıları, www.cengage.com adresinde bulunan Ch01_Problems adlı bir Microsoft Access veritabanında simüle edilmiştir.

1. Dosya kaç kayıt içeriyor? Her kayıta kaç alan var?
2. Şehre göre bir liste oluşturmak isterseniz hangi sorunla karşılaşsınız? Bu sorunu dosya yapısını değiştirerek nasıl çözersiniz?
3. Dosya içeriğini soyad, alan kodu, şehir, eyalet veya posta koduna göre sıralamak isterseniz, dosya yapısını nasıl değiştirirsiniz?
4. Hangi veri tekrarlarını (redundancy) tespit ediyorsunuz? Bu tekrarlar nasıl anomalilere yol açabilir?
5. Şekil P1.5'te gösterilen dosya yapısındaki ciddi veri tekrarlarını belirleyin ve tartışın.

Figure P1.5 The File Structure for Problems 5–8

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CODE	JOB_CHG_HOUR	PROJ_HOURS	EMP_PHONE
1	Hurricane	101	John D. Newson	EE	85.00	13.3	653-234-3245
1	Hurricane	105	David F. Schwann	CT	60.00	16.2	653-234-1123
1	Hurricane	110	Anne R. Ramoras	CT	60.00	14.3	615-233-5568
2	Coast	101	John D. Newson	EE	85.00	19.8	653-234-3254
2	Coast	108	June H. Sattlemir	EE	85.00	17.5	905-554-7812
3	Satellite	110	Anne R. Ramoras	CT	62.00	11.6	615-233-5568
3	Satellite	105	David F. Schwann	CT	26.00	23.4	653-234-1123
3	Satellite	123	Mary D. Chen	EE	85.00	19.1	615-233-5432
3	Satellite	112	Allecia R. Smith	BE	85.00	20.7	615-678-6879

6. Şekil P1.5'teki EMP_NAME ve EMP_PHONE içeriklerine bakarak hangi değişiklikleri önerirsiniz?
7. Problem 5'te incelediğiniz dosyadaki çeşitli veri kaynaklarını belirleyin.
8. Problem 7'deki cevabınıza göre, Şekil P1.5'te gösterilen dosyadaki veri tekrarlarını ortadan kaldırmak için hangi yeni dosyaları oluşturmalsınız?
9. Şekil P1.9'da gösterilen dosya yapısındaki ciddi veri tekrarlarını belirleyin ve tartışın. (Bu dosya, bir öğretmenin ders atama programı olarak kullanılması amaçlanmıştır. Veri tekrarlarının neden olduğu birçok sorundan biri, veri tutarsızlıklarının olası oluşumudur—Maria Cordoza adlı öğretmen için iki farklı baş harfin girilmesi gibi.)

Figure P1.9 The File Structure for Problems 9–10

BUILDING_CODE	ROOM_CODE	TEACHER_LNAME	TEACHER_FNAME	TEACHER_INITIAL	DAYS_TIME
KOM	204E	vWillston	Horace	G	MMF 8:00-8:50
KOM	123	Cordoza	Maria	L	MMF 8:00-8:50
LDB	504	Patroski	Donald	J	TTh 1:00-2:15
KOM	34	Hawkins	Anne	vV	MMF 10:00-10:50
JKP	225B	Risell	James		TTh 9:00-10:15
LDB	301	Robertson	Jeanette	P	TTh 9:00-10:15
KOM	204E	Cordoza	Maria	I	MMF 9:00-9:50
LDB	504	vWillston	Horace	G	TTh 1:00-2:15
KOM	34	Cordoza	Maria	L	MMF 11:00-11:50
LDB	504	Patroski	Donald	J	MMF 2:00-2:50

32 Part 1: Database Concepts

10.Şekil P1.9'da gösterilen dosya yapısına göre, KOM silindiğinde hangi sorunlarla karşılaşabilirsiniz?

11.Okulunuzun öğrenci bilgi sistemi kullanılarak sınıf programınızı yazdırın. Program, muhtemelen öğrenci kimlik numarası, öğrenci adı, ders kodu, ders adı, ders kredi saatleri, ders öğretmeni adı, dersin toplantı günleri ve saatleri, ve dersin sınıf numarasını içerecektir. Aşağıdaki işlemleri tamamlamak için Şekil P1.11'i bir şablon olarak kullanın.

Figure P1.11 Student Schedule Data Format

STU_ID	STU_NAME	CLASS_CODE	CLASS_NAME	CLASS_CREDHRS	INSTR_NAME	CLASS_DAYS	CLASS_TIMES	ROOM

- Şekil P1.11'de gösterilen şablonu kullanarak bir elektronik tablo oluşturun ve mevcut ders programınızı girin.
- Aynı elektronik tabloya, iki sınıf arkadaşınızın ders programını girin.
- Bu tasarımın neden olduğu veri tekrarları ve anomaliyi tartışın.