

# Databassystem

## Inlämningsuppgift 2: SQL & XML-lab

### 1 Introduktion

Under varje år av längdskidstävlingarna i Tour de Ski har vallateamet för landslaget tröttnat på att alltid gissa vilka skidor, valla och struktur som lämpar sig bäst för alla åkare inför varje tävling. Idag används ibland papper och penna för att komma ihåg vilka skidor som fungerade bra och mindre bra under respektive tävling. Tyvärr håller inte detta måttet då konkurrenterna börjar bli duktigare på att erbjuda bättre skidor till sina åkare. Därför blir din uppgift att som konsult implementera en prototyp av ett tänkt databassystem som håller koll på all viktigt data som underlättar för landslagets vallateam. Det är högst viktigt att delar av systemet ska kunna möjliggöra vidare utveckling så att landslaget kan få förslag på vilka skidor som skulle fungera bäst inför en tävling. Det skulle underlätta mycket då utprovning av skidor är en tidskrävande och extremt viktig process. Prototypen behöver inte i ett första skede vara komplett utan fokus för ditt arbete är att realisera en databas från den givna kravspecifikation samt att formulera transaktioner och uppdateringsoperationer mot densamma. Dokumentation (i form av en implementationsrapport) av systemet är av högsta vikt eftersom en utvärdering av prototypen sker uteslutande från den skapade dokumentationen.

### 2 Verksamheten

Tour de Ski är längdskidstouren som hålls varje år och där landslagens åkare slåss om att komma högst på pallen. Varje landslag har både manliga och kvinnliga åkare som åker i två olika stilar: klassisk skidåkning och skate. Varje landslag har också vallateam som hjälper åkarna att erhålla så bra skidor som möjligt inför varje tävling. Eftersom väderlek, temperatur och snötyp spelar stor roll för hur snabbt man kan ta sig fram i skidspåren måste vallateamet ta fram rätt egenskaper på skida, valla och struktur (mönster under skidan). Dessa tre kombinationer måste vara perfekta för att skidan ska gå så snabbt som möjligt och därmed underlätta för åkaren att komma först. Därför vill landslaget ha din hjälp med att kunna lagra all data som anses viktig för varje åkare och tävling. Exempelvis, vilka skidor gick bäst under respektive tävling? Vilken valla och struktur som fungerar i ett visst läge osv.

**Skidåkare:** Identifieras av sitt namn och har en vikt. Landslaget tror att en lätt åkare är snabbare i lössnö jämfört med en tyngre åkare. Detta är avgörande för vilka åkare man tar ut inför varje tävling.

**Tävling:** Har ett namn och sker på ett förbestämt datum. I varje tävling kan många åkare från landslaget tävla. Det finns även åkare som ej deltar i alla tävlingar, pga. skador och liknande.

**Väder:** Eftersom vädret spelar stor roll vill man lagra vädrets typ (exempelvis soligt, mulet, regn osv) och vilken temperatur som råder vid respektive typ. Det kan vara så att vädret ändrar sig under en tävling. Då vill man lagra detta på ett effektivt sätt. Exempelvis, om det är både soligt och minus 5 grader och senare sedan gick det över till att bli mulet och minus 10 grader.

**Snö:** Egenskaper i snön spelar också stor roll inför varje tävling. Här vill man lagra snötyp (exempelvis, finkornig, grovkorning, slask, nysnö osv) och vilken luftfuktighet som råder i snön. För en tävling kan snön skifta. Vid start kan det vara grov snö medan på toppen av backen råder slask eftersom solen tagit mer där osv.

**Skida:** Varje skidåkare har väldigt många olika skidor att välja på. Eftersom många skidor ser likadana ut är det viktigt att veta vilken åkare som de tillhör. Eftersom varje åkare är sponsrade av ett visst fabrikat ska även detta lagras. Varje skida har även ett spann som beskriver om skidan fungerar bra i kallt eller varmt väder samt dess snötyp. Varje skidåkares skida identifieras av ett nummer tillsammans med skidåkaren.

**Valla:** Vallateamet lägger ner väldigt mycket tid på att hitta rätt valla som fungerar inför varje tävling. Det utförs mycket tester för att hitta rätt fästvalla och glidvalla. Därför vill man lagra vallans namn och dess typ (om det är burkvalla, klister, glidvalla osv).

**Struktur:** Om en skida ska gå kunna gå riktigt fort måste den även ha rätt struktur (mönster) under skidan. Normalt funkar en grovstruktur (kan se ut som ett julgransmönster) i varmt väder med slaskig snö medan en finare struktur lämpar sig bäst i kallt väder och hårdare snö. Strukturen identifieras av dess namn och måste finnas på varje skida.

**Rillverktyg:** Eftersom det finns olika fabrikat som tillverkar strukturer vill landslaget även lagra vilket fabrikat som används. Det kan vara så att fabrikket Swix fungerade bättre jämfört med Skigo trots att samma struktur användes under skidan. Därför vill man även lagra en kommentar för varje rillverktyg. Ett rillverktyg identifieras unikt tillsammans med en viss struktur.

## 2.1 Landslagets ER-modell

Nedan visas ER-modellen som landslaget tagit fram tillsammans med systemutvecklingskonsulterna som inledde arbetet:

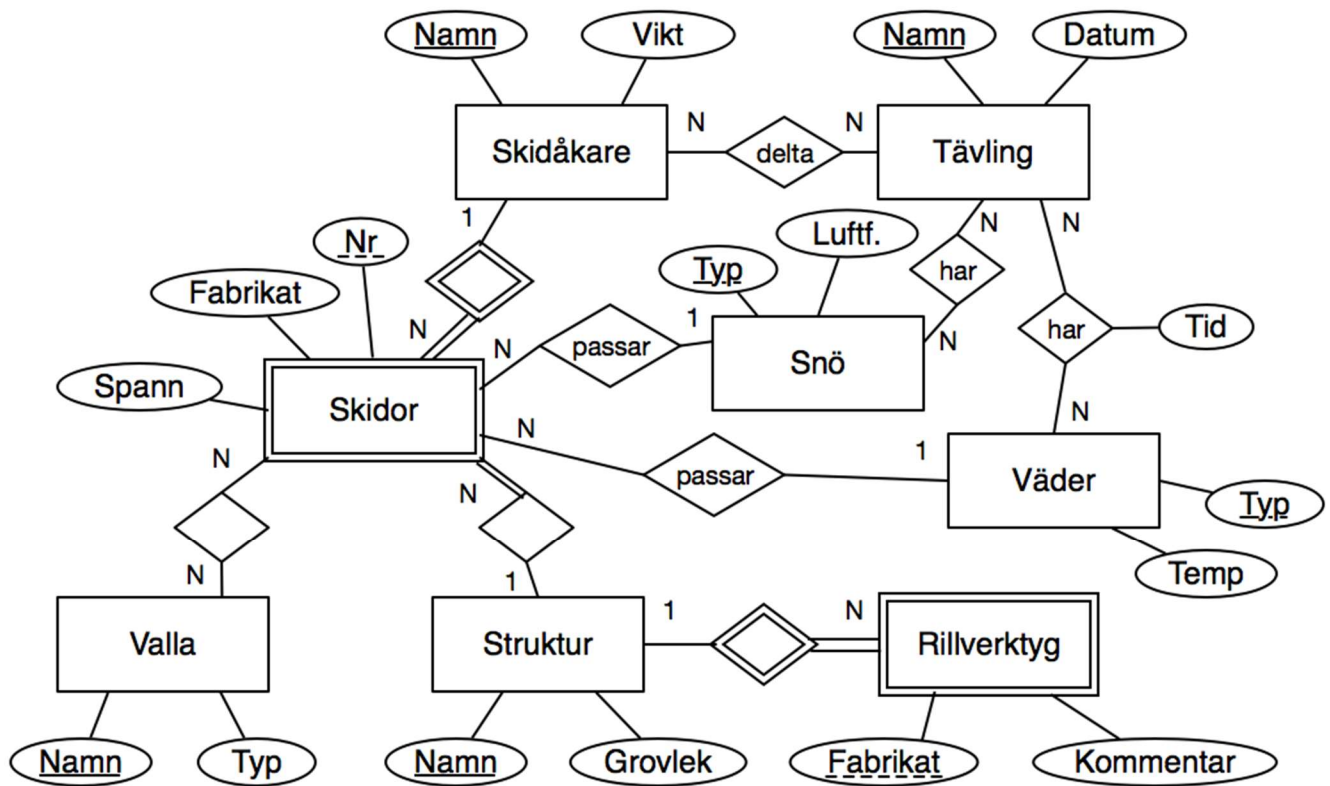


Fig. 1 ER-modell

### 3. Databaskrav och XML

Din uppgift är att utgå helt från den befintliga ER-modellen och realisera en databashanteraren MySQL. Prototypen ska dessutom fyllas med data så att det är möjligt att demonstrera lösningen för landslaget. Eftersom databasen inom en framtid också kommer kopplas till en hemsida ska du dessutom formulera frågeoperationer som kan användas mot databasen samt ange förslag på delar av databasen som kan skickas vidare med hjälp av XML. I kapitel 3.1 – 3.5 nedan finner du mer detaljerade krav som måste uppfyllas.

#### 3.1 Realisera databasen

Relationstabellerna enligt ER-modellen ska implementeras i databashanteraren MySQL. Det är viktigt att nyckelintegritet (primär och främmande nycklar) uppfylls på ett korrekt sätt enligt modellen. Attribut och dess passande datatyp ska implementeras på ett lämpligt sätt. Läs gärna "Laborationsinstruktioner.pdf" som finns tillgänglig i Canvas för att se vanliga designval för lämpliga datatyper.

#### 3.2 Transaktioner

Först när databastabellerna är implementerade med korrekt syntax och uppfyllt nyckelintegritet ska dessa fyllas med exempeldata. Det är viktigt att nedanstående data kan implementeras. **I de fall då data saknas för något attribut eller tidigare implementerad data faller det till din uppgift att hitta på lämplig data.** För att dessutom få en mer riktighet av lösningen ska du dessutom lägga in 3-4 egenkomponerade transaktioner för varje databastabell. Detta hjälper dig att testköra de frågeoperationer som listas i kapitel 3.3. Ett tips är att lägga in tillräcklig data så att det är möjligt att avgöra om frågeoperationen är korrekt formulerad, dvs. att rätt svar erhålls.

1. Skidåkaren Therese Johaug som väger 46kg.
2. Tävlingen "Mördarbacken" som hålls den 10/1-2016 och som skidåkaren Charlotte Kalla deltog i.
3. I tävlingen Oberstdorf som gick den 5/1-2016 bestod spåren mestadels av snötypen slask med luftfuktighet 100% trots att vädertypen var "Extremt kallt" med 20 minusgrader kl. 0800.
4. Skidorna som är tillverkade av Fischer med ett spann som är "Klisterskida" och som tillhör skidåkaren Markus Hellner. Skidan identifieras av nr 3. Denna skida består alltid av en grov struktur på 2mm som benämns "Grov Julgran". Skidan passar för snötypen slask och när vädret är regnigt.
5. Rillverktyget som är tillverkat av Swix har fått kommentaren: "Helt otroligt snabb i kallföre" skapar en snedskuren struktur som ofta finns på skidåkaren Anna Haags tredje par skidor som är tillverkade av "Fischer" där spannet är lågt.
6. Vallan Swix KX35 som är av typen klister finns på skidorna som tillhör både Emma Wikén och Stina Nilsson.
7. Tävlingen Lenzerheide där vädret skiftade från strålände solsken till spöregn under ett tidsintervall på 2h. Snön ändrade även karaktär från stenhårt till mjuk.
8. Skidåkaren Charlotte Kallas skidor som användes i tävlingen "Mördarbacken" och som vallades med glidvallan Swix HF8 och fästvallan Swix KX45.
9. Glidvallan vid namn Skigo HF-Gul som passar snötypen slask som användes under tävlingen Lenzerheide av Marit Björgen.

10. Rillverktyget från Skigo som gör en snedskuren struktur och som användes i tävlingen "Mördarbacken" av skidåkaren Anna Haag.

### 3.3 Frågeoperationer

Nedanstående operationer ska formuleras mot databasen. Det ska tydligt framgå vad koden gör och hur du har tänkt för att lösa operationerna. Om någon operation inte går att genomföra pga. din databasdesign, förklara då hur man skulle kunna designa databasen så att det är möjligt att genomföra operationen.

1. Hämta vikten för skidåkaren Therese Johaug.
2. Hämta namn och datum för tävlingen Mördarbacken.
3. Hämta vikten på alla åkare som deltog i tävlingen som hölls 5/1 – 2016.
4. Hämta namnet på skidåkarna som har en skida tillverkad av Madhus och som är markerad med nr 3 och som består av en snedskuren struktur.
5. Hämta vädertypen som gällde under tävlingen Lenzerheide kl 1200 och där Markus Hellners skidan nr 2 användes.
6. Hämta namnet på de personer som har samma vikt. (Tips. Skapa instanser av samma tabell och jämför deras vikt.).
7. Hämta namnet på de skidåkare som inte deltagit i någon tävling.
8. Hämta namnet på tävlingen som hade alla vädertyper som är lagrade i databasen.
9. Visa snötyperna som har förekommit någon gång under någon tävling.
10. Hämta namnet för de skidåkare som har deltagit i exakt två tävlingar. (Tips. Använd count).
11. Lista alla vallor och sortera namnet i omvänd ordning (Z först, A sist).
12. Hämta medelvikten för alla skidåkare.
13. Hämta medelvikten på skidåkare för respektive tävling (Tips. Använd Group by).
14. Hämta fabrikatet på alla rillverktyg som börjar med bokstaven "S".
15. Hämta fabrikatet på alla rillverktyg där kommentaren består av 5 bokstäver.
16. Hämta vädertypen som hade kallast temperatur.
17. Hämta den senaste tävlingen som hölls i Tour de Ski.
18. Lista den tävling som påbörjat den senaste dagen. (Tips. Använd curdate eller liknande).
19. Öka vikten med 10% för de skidåkare som väger mellan 50 och 60 kilo.
20. Ta bort skidåkaren som heter Anna Haag.
21. Ta bort Markus Hellners skida nr 2.

### 3.4 Grundläggande förståelse

En del av arbetet är att du ska visa din grundläggande förståelse inom databassystem genom att besvara några av landslagets frågor. Motivera dina svar väl.

1. Beskriv skillnaderna mellan datatyperna date och time.
2. Vilka konsekvenser erhålls om längden på en char överdrivs, dvs. ett högre värde sätts än vad som används?
3. Null-värden i en databas är något som i största mån bör undvikas. Men det finns även några fördelar med att lagra Null-värden. Ange ett exempel där det kan vara användbart att viss data lagrar ett Null-värde.

### 3.5 XML

I framtida arbete vill landslaget kunna presentera information på en intern hemsida. Ditt arbete inkluderar att visa förståelse av grunderna av XML i kontexten av relationsdatabaser samt ge förslag på en del av landslagets domän som kan presenteras i ett webbformulär med hjälp av XML. Besvara nedanstående delar.

1. Beskriv varför HTML inte är lämpligt för att strukturera data från en databas. Beskriv också skillnaderna mellan HTML och XML.
2. Beskriv varför trasiga XML-filer existerar (där start och sluttaggar inte matchar varandra) och hur detta kan undvikas.
3. Skapa en XML-fil som följer nedanstående DTD:

```
<!ELEMENT tävling(skidåkare*, snö*, väder*, datum)>
<!ELEMENT datum (#PCDATA)>
<!ELEMENT skidåkare (vikt)>
<!ELEMENT vikt (#PCDATA)>
<!ELEMENT snö(luftfuktighet)>
<!ELEMENT luftfuktighet (#PCDATA)>
<!ELEMENT väder (temperatur)>
<!ELEMENT temperatur (#PCDATA)>
<!ATTLIST tävling namn ID #REQUIRED>
<!ATTLIST skidåkare namn ID #REQUIRED>
<!ATTLIST snö typ ID #REQUIRED>
<!ATTLIST väder typ ID #REQUIRED>
```

4. Skapa en DTD som gör det möjligt att inkludera information om skidor, valla och struktur. Tanken är att landslaget vill använda informationen för att underlätta för vallapersonalens utprovning av skidor.

## 4 Redovisning

Din lösning ska lämnas in som en rapport och ska bestå av en laborationsrapport som beskriver lösningen och bilagor i form av källkod. Följande innehåll ska finnas i rapporten:

- Försättsblad med namn, program och personnummer.
- Tydligt framgå hur du har löst uppgiften.
  - Relationsdatamodell.
  - Konstruktion av databastabellerna.
  - Inmatning av transaktionerna.
  - Frågeoperationerna. Kort beskrivning av funktionaliteten av varje fråga.
  - Eventuell förlorad semantik. Finns det något i uppgiftsbeskrivningen som inte fångas i överföringen i ER-modellen?

- Besvara XML-frågorna, skapa XML-fil och DTD. Beskriv kort hur du kom fram till din lösning och varför den ser ut som den gör.
- Beskriv egna antaganden och krav som du har identifierat för att kunna lösa uppgiften.
- Källkod som löser uppgiften och som följer ER-modellen och dina relationstabeller.
- Kommentarer i källkoden som tydligt förklarar avsikten med koden.

## 5 Regler

Inlämningsuppgiften ligger till grund för examination i kursen och löses individuellt. Detta innebär att det är väldigt viktigt att materialet som skickas in till läraren är ditt eget och är enkelt att identifiera som ditt eget. Att kopiera text, bilder eller kod är inte accepterat. Däremot är det tillåtet att använda källhänvisningar för att visa vart information kommer ifrån samt diskutera lösningsförslag och strategier med lärare och studenter. Men uppgiften ska vara individuellt utförd efter egen förmåga. Grupparbete där källkod och skriftlig rapport skapas tillsammans tillåts inte.

## 6 Deadline

Sista inlämningsdatum för uppgiften står angivet i Canvas. Rapporten ska lämnas in elektroniskt via läroplattformen Canvas (instruktioner finns på kurshemsidan). Alla andra former av inlämning (ex e-post, utskrivna papper) är inte tillåtna.