

IN2090 – Databaser og datamodellering

01 – Introduksjon og oversikt over emnet

Leif Harald Karlsen
leifhka@ifi.uio.no



Universitetet i Oslo

Oversikt over kurset

- ◆ To hovedkategorier: Modellering og SQL
- ◆ Tirsdager: Forhåndsinnspilte videoer (dekker pensum)
- ◆ Torsdager: Digital forelesning med gjennomgang av eksempler og spørsmål
- ◆ Forventes at dere ser/deltar på begge

Oversikt over kurset

Uke	Tema	Dato	Foreleser	Innleveringer
1	Introduksjon og motivasjon	Tors 24.08	Leif Harald	
2	<i>Modellering</i> : Relasjonelle modellen	Tors 31.08	Leif Harald	
3	<i>Modellering</i> : ER-modellering 1	Tors 07.09	Leif Harald	
4	<i>Modellering</i> : ER-modellering 2	Tors 14.09	Leif Harald	
5	<i>DB</i> : Grunnleggende SQL	Tors 21.09	Leif Harald	
6	<i>DB</i> : Datamanipulering, skranker og views	Tors 28.09	Leif Harald	1. Modellering
7	<i>Modellering</i> : Normalformer	Tors 05.10	Leif Harald	
8	<i>Modellering</i> : Tapsfri dekomposisjon	Tors 12.10	Leif Harald	2. Enkel SQL
9	<i>DB</i> : Aggregering og sortering	Tors 19.10	Leif Harald	
10	<i>DB</i> : Ytre joins og mengde operatorer	Tors 26.10	Leif Harald	3. Normalformer
11	<i>DB</i> : Programmering med databaser	Tors 02.11	Leif Harald	
12	<i>DB</i> : Sikkerhet i databaser	Tors 09.11	Leif Harald	4. Avansert SQL
13	<i>DB</i> : Indekser og spørreprosessering	Tors 16.11	Leif Harald	5. Programmering med SQL
14	Repetisjon	Tors 23.11	Leif Harald	

Mer om timeplan og obliger:

<https://www.uio.no/studier/emner/matnat/ifi/IN2090/h23/index.html>



Leif Harald Karlsen

leifhka@ifi.uio.no

Gruppelærere og rettere

Gruppelærere:

- ◆ Camilla Dalby Borger
- ◆ Elise Marie Utermohlen
- ◆ Jamila Mehmandarova
- ◆ Julia Lundeby
- ◆ Peter Alexander Døvikén Løvold
- ◆ Sindre Bjørndal-Riis

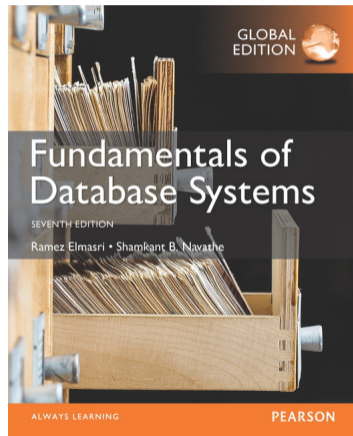
Rettere:

- ◆ Agnes Hunstad Birkemo
- ◆ Andrea Haugen
- ◆ Edis Kukuruzovic
- ◆ Ole-Sander Halla Johansen

Gruppetimer

- ◆ Én-til-én hjelp med ukesoppgaver og innleveringer
- ◆ 6 grupper
- ◆ Det er lov å bytte gruppe, men spør gruppelæreren først
- ◆ Gruppetimenes tema er alltid forelesningstemaet for uken før!
- ◆ Første gruppetime er neste uke

- ◆ Videoene og foilene til kurset
 - ◆ Publisert på timeplanen på semestersiden
 - ◆ Foiler/screencast fra torsdagsforelesningene blir også publisert
- ◆ Ukesoppgavene og innleveringsoppgavene
 - ◆ Gjennomgås på gruppetimene
- ◆ Boken *Fundamentals of Database Systems*
 - ◆ Tykk bok, men skal bare bruke deler av den
 - ◆ Se timeplanen for hva som er pensum
 - ◆ Gjenbrukes i andre kurs (IN3020)



Ukesoppgaver

- ◆ Publiseres i slutten av uken¹
- ◆ Gjennomgås på gruppetimen uken etter publisering
- ◆ Løsningsforslag publiseres i slutten av uken den gjennomgås

¹<https://www.uio.no/studier/emner/matnat/ifi/IN2090/h23/ukesoppgaver/>

Obliger og innleveringsoppgaver

- ◆ Fem innleveringsoppgaver med tilbakmelding
- ◆ Første av disse er obligatorisk (likt som i IN2010)
 - ◆ Fokus er på læring, ikke på løsning
 - ◆ Sterkt anbefalt å gjøre alle!
 - ◆ Essensiell del av undervisningen
 - ◆ Frihet under ansvar
- ◆ Publiseres torsdager, frist på torsdager²
- ◆ Leveres via Devilry
- ◆ To forsøk på obligen
- ◆ De andre kan man be om tilbakemelding på et andre forsøk
- ◆ Dersom du blir syk el., kan man be om utsatt frist

²<https://www.uio.no/studier/emner/matnat/ifi/IN2090/h23/innleveringer/index.html>

- ◆ Eksamen: 13. desember (4 timer, Silurveien)
- ◆ Bokstavkarakter (A-F)
- ◆ Alle skriftlige hjelpemidler er tillatt

- ◆ RDBMS: PostgreSQL 15²
 - ◆ Avansert relasjonell databasesystem
 - ◆ Åpen kildekode og støtter de fleste plattformer
 - ◆ Kan f.eks. bruke `psql`⁴ eller Edbit⁵
- ◆ Modelleringsverktøy: Dia³ eller `app.diagrams.net`
 - ◆ Både Dia og `app.diagrams.net` er generelle diagram-verktøy
 - ◆ Begge har åpen kildekode og støtter de fleste plattformer (Linux, Mac, Windows)
 - ◆ Trenger IKKE bruke dette om man ikke ønsker, kan tegne på papir eller bruke andre programmer!

²<https://www.postgresql.org/>

³<http://dia-installer.de/download/index.html>

⁴<https://www.postgresql.org/docs/14/app-psql.html>

⁵<https://gitlab.com/leifhka/edbit>

PostgreSQL



Databaser

- ◆ IMDB: Internet Movie DataBase
 - ◆ Filmdatabasen
 - ◆ Alle filmer, serier, episoder, skuespillere, roller, osv. frem til 2011
 - ◆ Stor database (2.4G)
 - ◆ >1 mil. filmer/serier/episoder, >10 mil. roller, >1.7 mil. personer, osv.
- ◆ Northwind
 - ◆ Eksempel-database laget av Microsoft Research for opplæring i databaser
 - ◆ Database over produkter, bestillinger, kunder, osv.
 - ◆ Liten database, men realistisk oppsett
- ◆ Personlig database
 - ◆ Alle har hver sin egne database
 - ◆ Du har full tilgang (kan lage, slette, oppdatere data, osv.)
- ◆ Satt opp av USiT for dere
- ◆ Mer info om tilkobling osv. kommer senere

- ◆ Nytt system utviklet av Institutt for Teoretisk Astrofysikk
- ◆ Brukes for diskusjon, spørsmål og kommentarer rundt pensum, oppgaver, obliger, osv.
- ◆ Bruker UiO-brukernavn og passord for innlogging
- ◆ Støtter:
 - ◆ Publisering med bruker og anonymt
 - ◆ Markering av riktig svar
 - ◆ Math-mode og code-blocks (SQL)
 - ◆ Kategorier og tags
 - ◆ osv.
- ◆ Guide ligger her: [https://www.mn.uio.no/english/services/it/discourse-uio/astro-discourse-instructions-\(students\).pdf](https://www.mn.uio.no/english/services/it/discourse-uio/astro-discourse-instructions-(students).pdf)

`https:
//astro-discourse.uio.no/c/in2090-23h`

AstroDiscourse: Råd og regler

- ◆ Gjerne hjelp hverandre, svar på hverandres spørsmål, osv.
- ◆ Vær hyggelig! :)
- ◆ Sjekk om spørsmålet du har allerede er postet
- ◆ Bruk alltid den mest spesifikke kategorien

- ◆ Mye ny og gammel teknologi som kan løse problemer
- ◆ Store språkmodeller (som ChatGPT, Bing, osv.) er flotte verktøy!
- ◆ Vil nok klare mange av oppgavene i dette kurset
- ◆ Derfor to viktige spørsmål:
 - ◆ Er det noen grunn til å lære seg pensum?
 - ◆ Hvordan bruke store språkmodeller i studiene?

Er det noen grunn til å lære seg pensum? (1)

JA! :D

- ◆ Språkmodeller er langt flinkere på problemene i dette kurset enn pensum generelt!
 - ◆ Problemene dere får er “enkle” og vanlige
 - ◆ Modellene er derfor godt trent på disse, men ikke “ekte” problemer
 - ◆ Modellene sliter med å sette sammen mindre program-snutter til større løsninger
- ◆ Ikke alltid mulig å bruke språkmodeller
 - ◆ Sensitive domener
 - ◆ For store eller komplekse problemer
 - ◆ Dette kan endre seg, men vanskelig å spå fremtiden

Er det noen grunn til å lære seg pensum? (2)

JA! :D

- ◆ Må gi riktig input/prompt
 - ◆ Krever ofte dyp forståelse av domenet
 - ◆ Må vite hva som er mulig og ikke
- ◆ Språkmodellene gjør feil
 - ◆ Feil i selve løsningen
 - ◆ Løser feil problem
 - ◆ Svarene må derfor sjekkes
- ◆ Språkmodellene kan kun det den kan
 - ◆ Nye teknologier kommer og går
 - ◆ Kan ikke forskning og utvikling
- ◆ Det er gøy!

Hvordan bruke store språkmodeller i studiene?

- ◆ Ikke bruk språkmodeller til å løse oppgaver!
 - ◆ Læring krever at man faktisk tenker
 - ◆ Lærer lite av å bare se en løsning
- ◆ Bruk dem heller for samtaler om pensum!
 - ◆ Still generelle spørsmål og oppfølgingsspørsmål
 - ◆ Husk: De gjør fortsatt feil! Test ut svaret den gir
- ◆ Læring er noe du må gjøre selv og kan ikke automatiseres!

Et par ord om læring

- ◆ Delta på undervisningen, gjør alle oppgaver, les tilbakemeldinger, osv.
- ◆ Husk: dere er er for å lære, ikke for å løse problemer!
- ◆ En av grunnene til endret oblig-opplegg
- ◆ Sjekk egen forståelse, og vær ærlige med dere selv
- ◆ Denne muligheten til å virkelig lære noe dypt får dere sjeldent igjen!

Eksempler på hva dere skal lære

Databaser og SQL

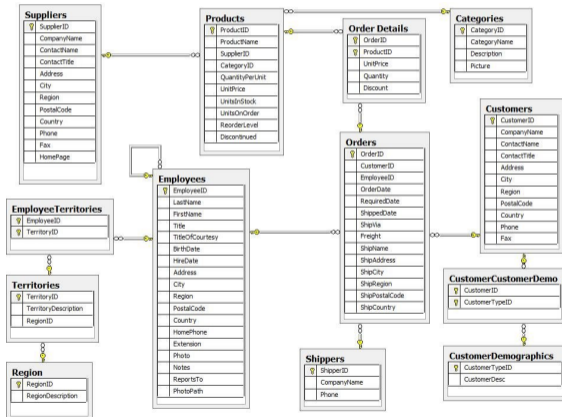
Finn topp 3 rangerte action-filmer med Harrison Ford som skuespiller.

```
SELECT f.title
FROM person AS p
  INNER JOIN filmparticipation AS fp USING (personid)
  INNER JOIN filmgenre AS fg USING (filmid)
  INNER JOIN filmrating AS fr USING (filmid)
  INNER JOIN film AS f USING (filmid)
WHERE p.firstname = 'Harrison' AND
      p.lastname = 'Ford' AND
      fp.parttype = 'cast' AND
      fg.genre = 'Action'
ORDER BY fr.rank DESC
LIMIT 3;
```

Eksempler på hva dere skal lære

Databaser og SQL

Northwind-databasen:



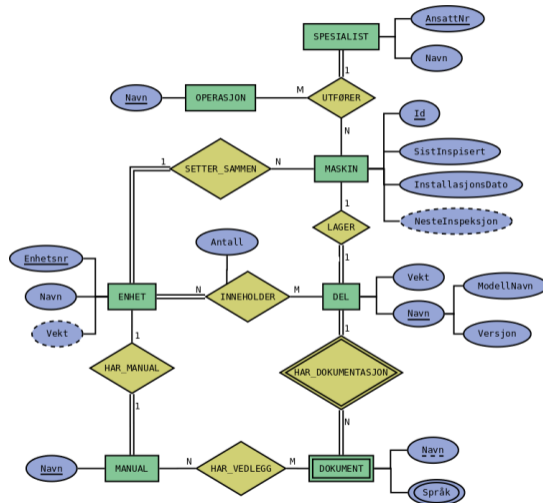
Eksempel-oppgave:

Finn navn og total regning for hver kunde som har kjøpt fler enn 1000 varer, sortert alfabetisk etter firmanavn.

```
SELECT c.company_name,
       sum(d.unit_price * d.quantity) AS bill
FROM customers AS c
     INNER JOIN orders AS o USING (customer_id)
     INNER JOIN order_details AS d USING (order_id)
GROUP BY c.customer_id, c.company_name
HAVING sum(d.quantity) > 1000
ORDER BY c.company_name;
```

Eksempler på hva dere skal lære

Modellering og ER



Eksempel ER-modell av informasjon om fabrikk som produserer enheter.

Nyttige lenker (Ikke pensum)

- ◆ Tidligere års forelesninger (opptak, slides, videoer, osv.)
 - ◆ <https://www.uio.no/studier/emner/matnat/ifi/IN2090/h19/timeplan/index.html>
 - ◆ <https://www.uio.no/studier/emner/matnat/ifi/IN2090/h20/timeplan/index.html>
 - ◆ <https://www.uio.no/studier/emner/matnat/ifi/IN2090/h21/timeplan/index.html>
 - ◆ <https://www.uio.no/studier/emner/matnat/ifi/IN2090/h22/timeplan/index.html>
- ◆ Mer om databaser:
 - ◆ Engelsk: <https://en.wikipedia.org/wiki/Database>
 - ◆ Norsk: <https://no.wikipedia.org/wiki/Database>
- ◆ Mer om relasjons-databaser:
 - ◆ Engelsk: https://en.wikipedia.org/wiki/Relational_database
 - ◆ Norsk: <https://no.wikipedia.org/wiki/Relasjonsdatabase>
- ◆ Mer om datamodellering:
 - ◆ Engelsk: https://en.wikipedia.org/wiki/Data_modeling
 - ◆ Engelsk: https://en.wikipedia.org/wiki/Data_model
- ◆ *What is the benefit of learning SQL?* (engelsk):
<https://www.quora.com/What-is-the-benefit-of-learning-SQL>