

INF115 Databases and Modelling

Adriaan Ludl

Department of Informatics

University of Bergen

Spring Semester 2021

20.01.2021



Course Responsible & Lecturer

Adriaan Ludl

adriaan.ludl@uib.no

2019 – ongoing: Postdoctoral fellow @ Department of Informatics UiB

- **Research on causal inference on genomics data**

2016 – 2019: Postdoctoral researcher @ Universitat de Barcelona, Spain

- **Research on network inference from neuronal activity**

2015 PhD, Université Pierre et Marie Curie, Paris, France

Introductory Lecture

Today's topics:

- Practical information about INF115
- Introduction (Part 1 / 2): What are databases ?

INF115 course content

An introduction to *methods for organizing, structuring, representing and storing large amounts of information.*

Emphasis on:

- techniques for **data modeling**,
- theory of **relational databases**.

Important topics: *relational algebra, query language, storage media and storage methods.*

Outline

- Course books
- Lecture schedule
- Group sessions
- Mandatory assignments
- Exam
- Software used in the course

INF115 Databases and Modelling

Course page updated regularly!

<https://mitt.uib.no/courses/27455>



Course Books

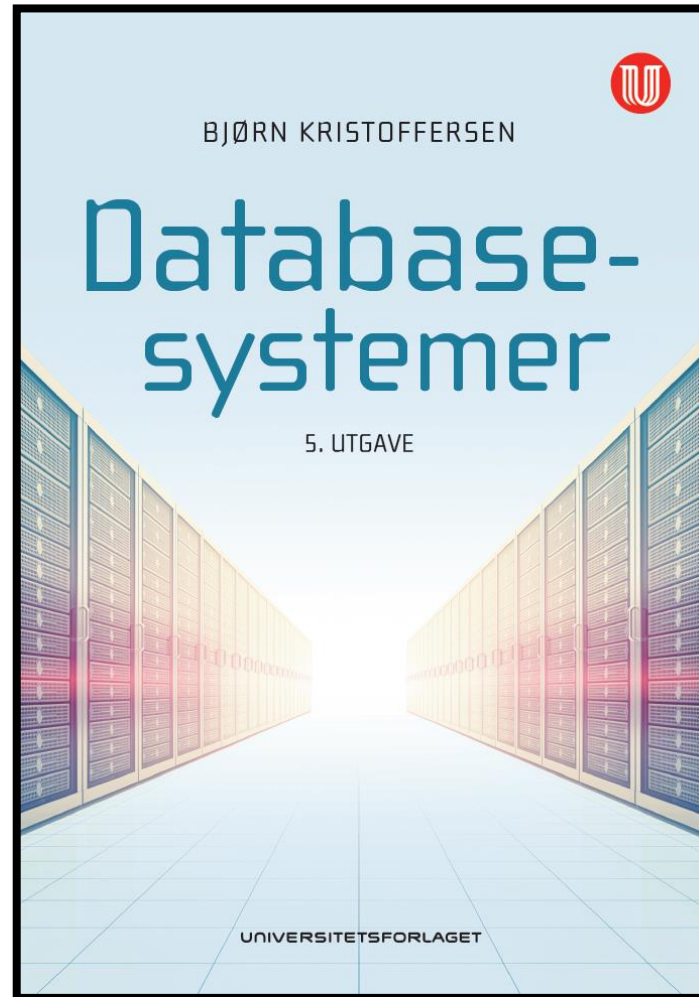
In Norwegian:

<https://www.dbsys.info/Databasesystemer/>

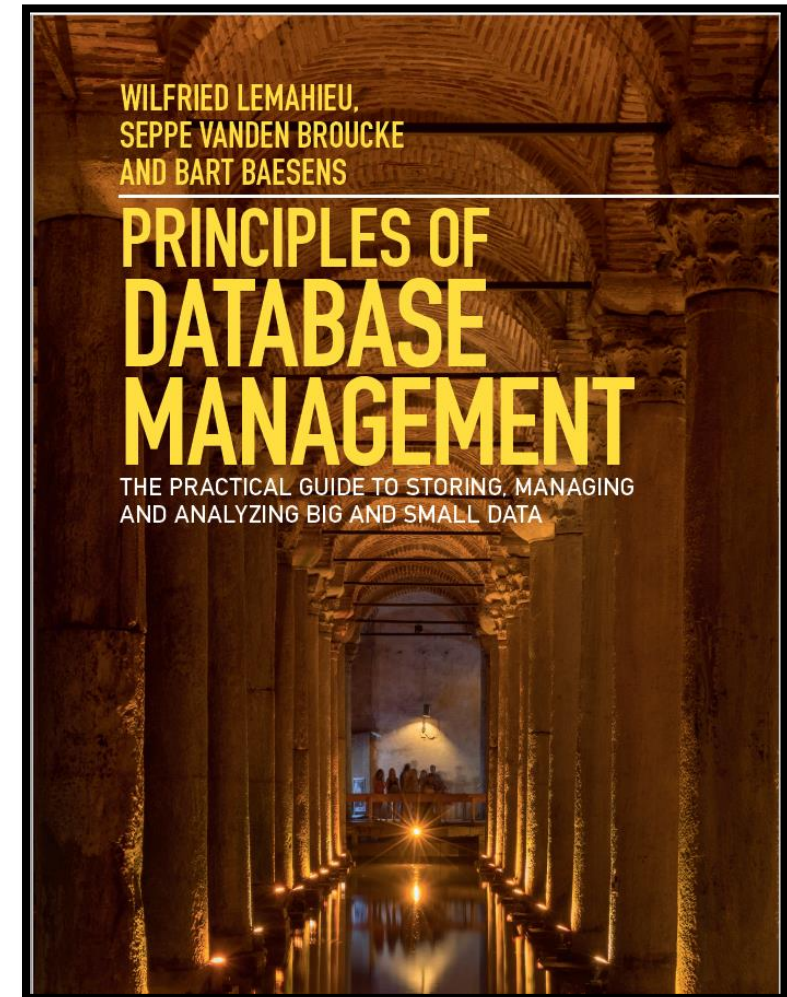
In English:

<https://www.pdbmbook.com/>

- Both cover the INF115 course material
- Students can choose the one they prefer



INF115 DB & Modelling



Lecture Schedule (preliminary)



**All lectures online,
will be recorded**

**on Tuesdays (10:15 to 12:00)
and Wednesdays (14:15 – 16:00)**

Week	Day	Nr.	Topic
3	20.01.2021	0	Introductory lecture
4	26.01.2021	1	Introduksjon (Part 2)
4	27.01.2021	2	Tabeller og enkel spørringer
5	02.02.2021	3	Tabelldefinisjon
5	03.02.2021	4	Datamanipulering
6	09.02.2021	5	Spørringer mot flere tabeller
6	10.02.2021	6	Avanserte spørreteknikker
7	16.02.2021	7	Relasjonsmodellen
7	17.02.2021	8	Relasjonsmodellen
8	23.02.2021	9	Datamodellering med E/R
8	24.02.2021	10	Datamodellering med E/R
9	02.03.2021		
9	03.03.2021	11	E/R-diagrammer til tabeller
10	09.03.2021		
10	10.03.2021	12	Normalisering

11	16.03.2021		
11	17.03.2021	13	Filer og indekser
12	23.03.2021		
12	24.03.2021	14	Transaksjoner
13	30.03.2021		(Paskeferie)
13	31.03.2021		(Paskeferie)
14	06.04.2021	15	Databaseadministrasjon
14	07.04.2021	16	Databaseadministrasjon
15	13.04.2021	17	Web-applikasjoner
15	14.04.2021	18	Web-applikasjoner
16	20.04.2021		(Guest lecture)
16	21.04.2021	19	SQL injection
17	27.04.2021	20	XML og JSON
17	28.04.2021	21	XML og JSON
18	04.05.2021	22	Bioinformatikk og databaser
18	05.05.2021		(Guest lecture)
19	11.05.2021	23	Via Objekter to NoSQL
19	12.05.2021	24	Via Objekter to NoSQL

Group sessions schedule

Sessions start on 1st of February:

- 6 groups to happen physically,
- 2 online groups.

**You can only attend the group
for which you are registered !
Self-registration from today
on mitt.uib.no**

- Teaching Assistant:
Håkon Tjeldnes
hakon.tjeldnes@uib.no



Group	Day	Time	Place	Capacity (People)
1	Monday	12:15 - 14:00	HiB, Room 105O1	18
2	Monday	12:15 - 14:00	HiB, Room 106O1	27
3	Wednesday	08:15 - 10:00	HiB 510N3 (5. etasje)	21
4	Monday	10:15 - 12:00	HiB 510N3 (5. etasje)	21
5	Monday	12:15 - 14:00	HiB 510N3 (5. etasje)	21
6	Monday	14:15 - 16:00	HiB 510N3 (5. etasje)	21
7	Thursday	14:15 - 16:00	online	30
8	Friday	10:15 - 12:00	online	30

Mandatory assignments



- **Three mandatory hand-in assignments,**
- **Due dates** will be announced on <https://mitt.uib.no/courses/27455/announcements>
- Each assignment counts for 10% of the final grade.
- Each assignment must be passed (50%) to be allowed to take the exam.
- You can discuss the assignments in groups,
- but **each student submits their solution individually.**



Exam INF115

A 3-hour written exam on 7th of June, 07.06.2021.

Note: digital exam on Inspira.

Exact location and time can be found [\(here\)](#) at a later date.

Be aware that location and time may change up to a few days before the exam!



Software used in the course



*Install it **now** before the first group sessions 😊*

- **MySQL Workbench**: visual database design tool :
<https://dev.mysql.com/downloads/workbench/>
- **PHP**: webpage script language
- **XAMPP**: web server stack, contains MariaDB and PHP:
<https://www.apachefriends.org/download.html>
- Recommended **E/R Diagram Editors**:
 - LucidChart
 - or MySQL Workbench Diagrams
- Recommended **SQL Stacks**:
 - myPhpAdmin
 - MySQL Workbench



For questions about the software please contact the TA or the student group leaders:

Teaching Assistant: Håkon Tjeldnes

hakon.tjeldnes@uib.no

How to study INF115 successfully 😊

- **Attend the lectures live online !**
 - *Ask questions*
 - ***Discuss with your colleagues***
 - *Read the **chapters** corresponding to the lectures*
- **Attend the study groups !**
 - *Try to answer each question by yourself,*
 - ***Be active during group sessions,***
 - *attend the consultancies for the mandatory assignments !*
- **Study all the topics and problems again during exam preparation**
- **Reach out to us if you have questions !**



Questions on course organization ?

INF115 Databases and Modelling

Course page updated regularly!

<https://mitt.uib.no/courses/27455>



Chapter 1 – Introduction (Part 1/2)

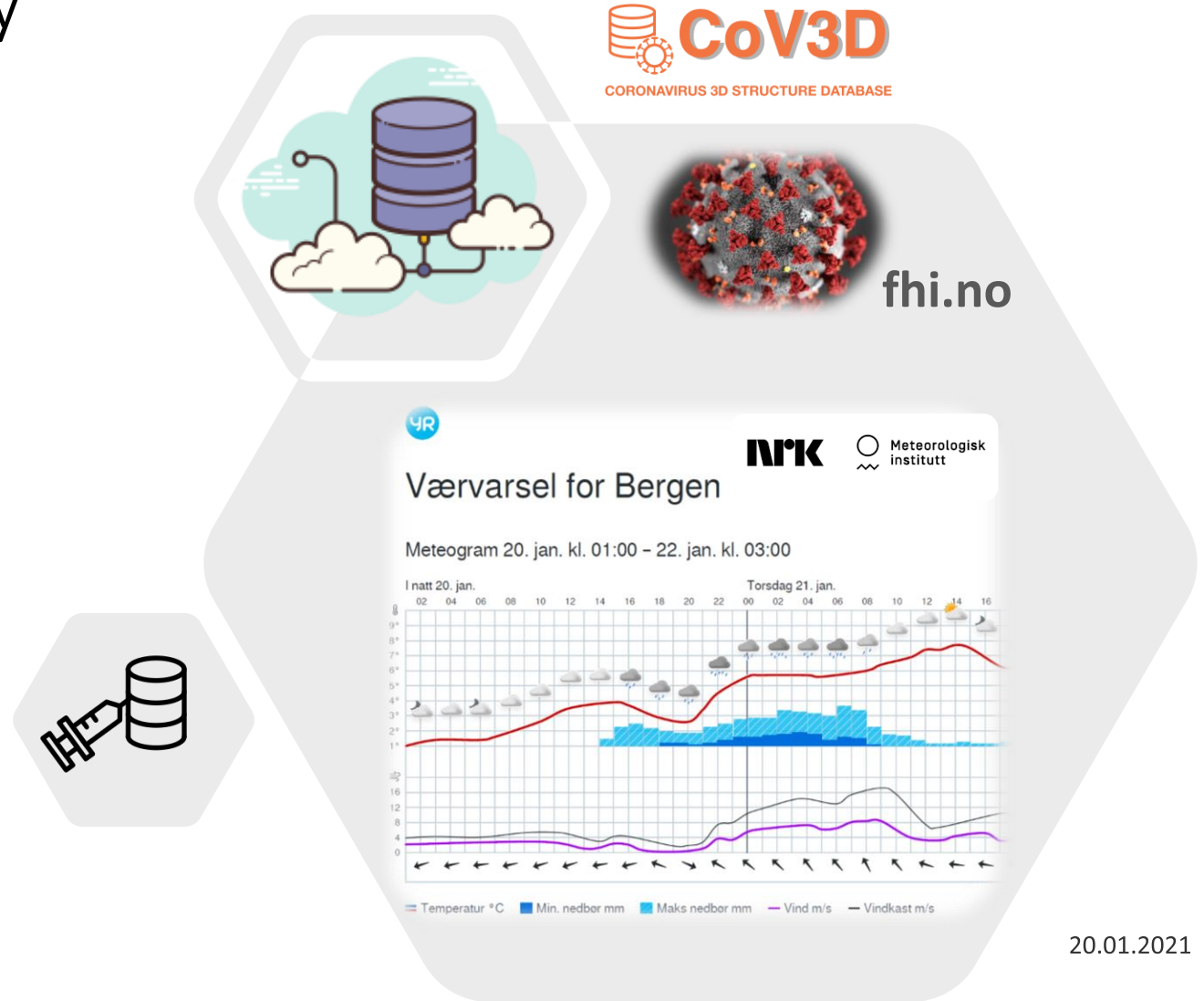
Learning objectives:

- Important applications of database (DB) systems
- **What** are database systems (DBS) ?
- **Why** are they used ?
- Common uses of DB systems



Databases are behind many software applications

- **Account management**
- **Cloud computing**, Email, Github ...
- **Health**: Statistics, *Coronavirus test results*, Hospitals, *Genetics* ...
- Banks, **Payment systems**
- **Universities**, Research, Weather forecast, **Machine learning** ...
- Media, **Wikipedia**
- Online services: **government** (tax), banks, stock market, shopping, travel, videos, games ...



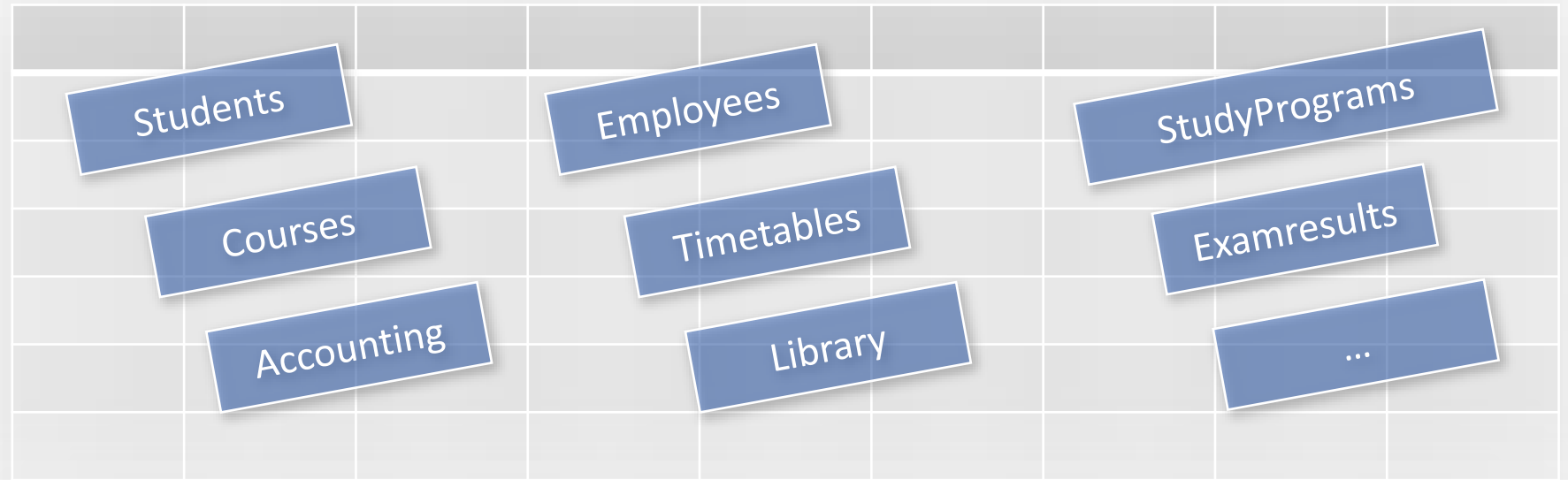
20.01.2021

What is a database ?

Database (DB) = a logically organized collection of data (*information*)

Examples: A university keeps many databases

➤ **every** database can contain **many tables**:

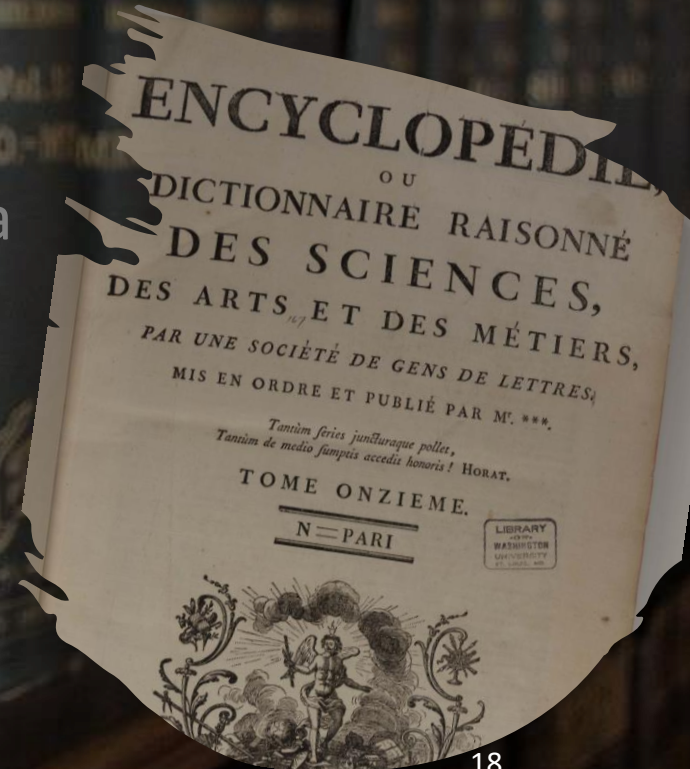


Why use databases ?



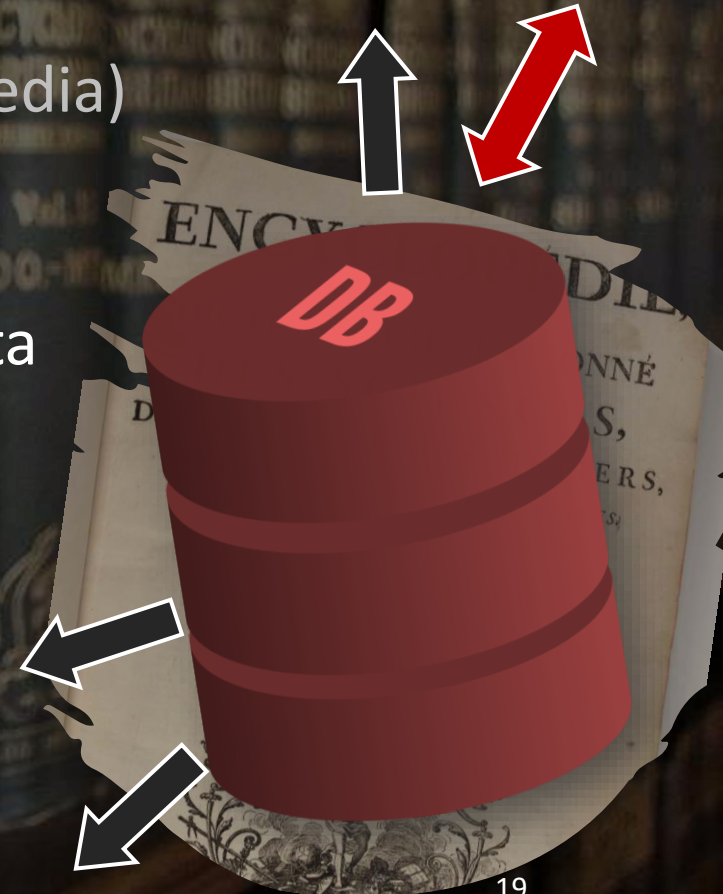
Why use databases ?

- Allows to collect & organize data from **many sources** in one system
 - **Create, update, delete** tables and entries
 - Streamline data **curation** and **administration**
- Store & access data **efficiently** (even big data: e.g. Wikipedia)
 - **Queries, indexing**, search, references ...
- Serve the *same data to many users simultaneously*
 - Allow users to **choose an appropriate view** of the data
 - **Security**: control access rights of users
 - Allow access over a **network** (internet or internal)
- Want to guarantee **consistency & integrity** of the data
 - Transaction management enables **reliable backups**



Why use databases ?

- Allows to collect & organize data from **many sources** in one system
 - **Create, update, delete** tables and entries
 - Streamline data **curation** and **administration**
- Store & access data **efficiently** (even big data: e.g. Wikipedia)
 - **Queries, indexing, search, references** ...
- Serve the **same** data to **many** users **simultaneously**
 - Allow users to **choose an appropriate view** of the data
 - **Security**: control access rights of users
 - Allow access over a **network** (internet or internal)
- Want to **guarantee consistency & integrity** of the data
 - Transaction management enables **reliable backups**



DIKW Pyramid (or hierarchy)

DIKW: Data, Information, Knowledge, Wisdom

“scientia potestas est” (= knowledge is power)

- Francis Bacon, philosopher (1561–1626)

Reference:

Henry, Nicholas L. (May–June 1974).

"Knowledge Management:

A New Concern for Public Administration".

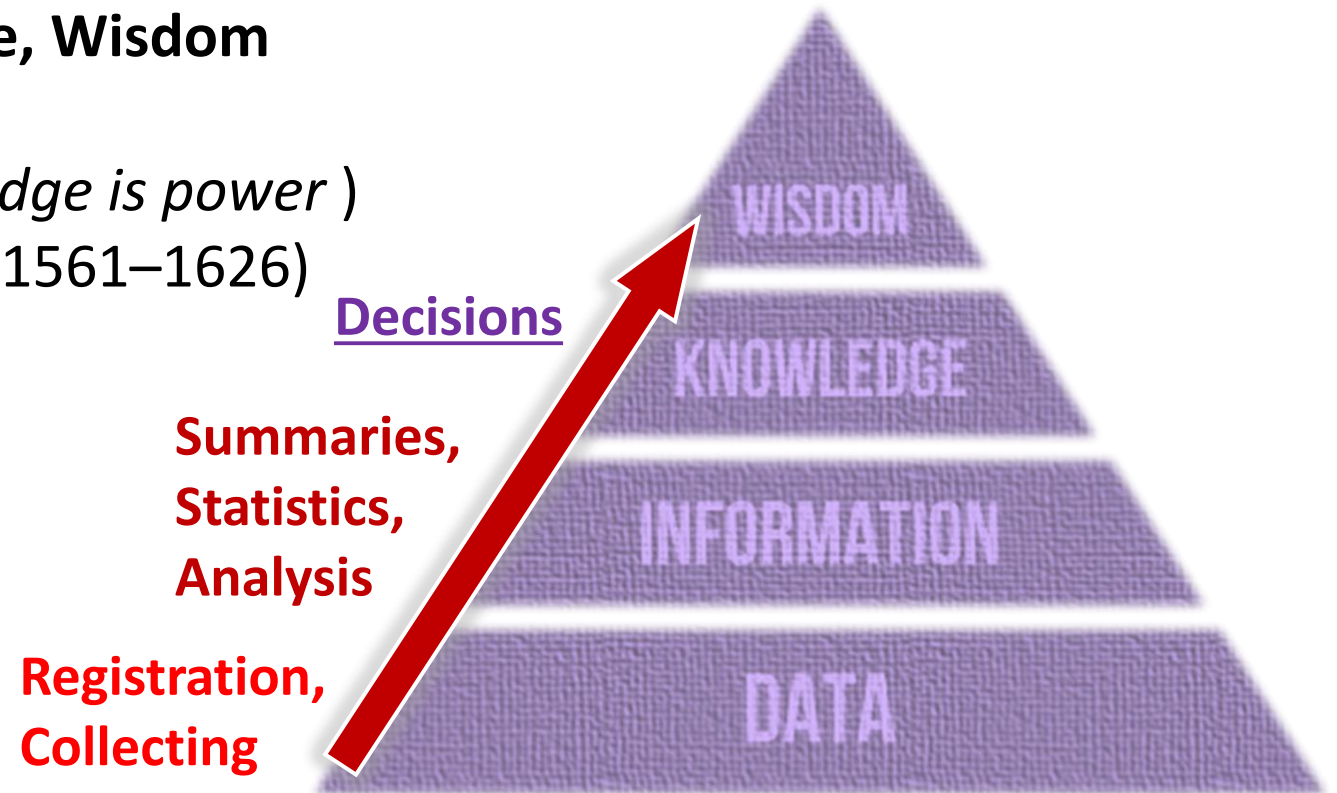
Public Administration Review. **34** (3): 189–196.

[doi:10.2307/974902](https://doi.org/10.2307/974902). [JSTOR 974902](https://www.jstor.org/stable/974902).

Rowley, Jennifer (2007). "The wisdom hierarchy: representations of the DIKW hierarchy".

Journal of Information and Communication Science. **33** (2): 163–180.

[doi:10.1177/0165551506070706](https://doi.org/10.1177/0165551506070706)



https://en.wikipedia.org/wiki/DIKW_pyramid

More Examples: Databases and Tables

- **Cloud computing**: UserAccounts, Passwords (encrypted), UserData, ...
 - **Email**: Contacts, Inbox, Sent, Labels, Spam ...
 - **Health**: Corona test results, Patients, Medication, Hospitals, Doctors, Appointments, ...
 - **Payment systems**: eID, Cardnumber, PaymentOperations, ...
 - **Wikipedia**: Articles, Pictures, Media, Users, ...
 - **Libraries**: Books, Journals, Articles, Loans, ...
-
- (Almost) All **businesses** use databases
 - Many systems must be **operational 24/7** (*critical services*)
 - Databases are part of larger **information processing systems**:
 - Examples: **Accounting** systems, Project **planning** systems, ...



Quizz on Introduction to Databases

Please answer this practice quizz on mitt.uib now 😊

(you can take it again now if you want)

- <https://mitt.uib.no/courses/27455/quizzes/16481>

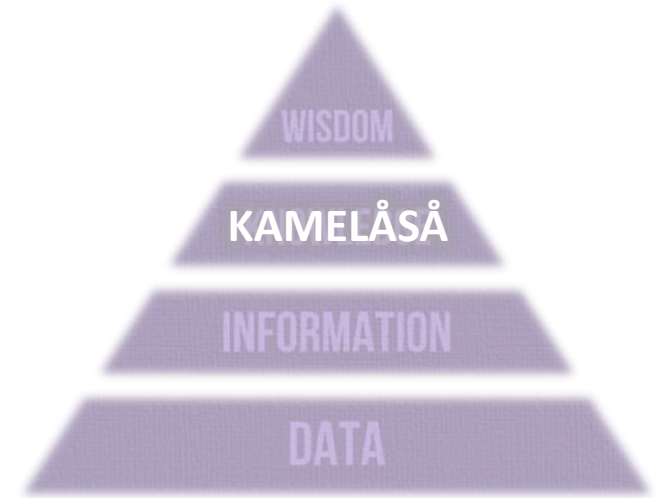
Summary: Chapter 1 – Introduction (Part 1/2)

- Databases are **behind many software applications**.
- A Database (DB) is a **logically organized collection of data** (*information*).
- DB systems (DBS) allow to **collect & organize data** from many sources.
- DBS enable us to store & access data **efficiently and reliably**.
- DBS are part of **information processing systems**.
- A DB may contain incomplete data, some values may be missing.
- A DB should not contain errors or contradictions.





Questions ?



- **Next lecture on Tuesday 26.01.2021 (10:15 – 12:00)**