# Data utilization

• What is the context? Why is the project relevant? What is it about? • What purpose (use case) does the database serve in your project? • Which decision should be supported and how exactly?

Wir nutzen die Daten, um eine Prognose zu machen, welcher Tennis Spieler ein gewisses ATP Turnier, egal ob real oder nicht, gewinnen würde.

Der Nutzen daraus kann verschieden sein, man kann es für Wetten verwenden, für TV-Programme, Spiel Simulationen, Video Spiele und weiteres.

Die Datenbank wird verwendet, um die vielen Daten zu persistieren und zu verwalten. Ohne die Datenbank könnten wir nicht effizient die Informationen verarbeiten für ein sinnvolles Resultat.

# Data architecture

• Which data (classes, attributes) do you need for your use case? • What software architecture is needed for this?

Siehe ER Diagramm.

Die Daten basieren auf vergangene ATP Matches.

# Data administration

• Which target group is the data analysis for? • Who gets access and how? • Are the security requirements?

Zielgruppe: Webseiten, Wettbüros, Videospiel Entwickler, Fernsehsender, ATP.

Zugriff: Die Endnutzer bekommen die Daten über eine REST-Schnittstelle. Die Administratoren bekommen zugriffen über MySql Workbench.

Sicherheit:

* Da es nicht um klassifizierte Daten handelt, müssen wir uns nur auf die Integrität, Konsistenz und Verfügbarkeit fokussieren.
* Um die Integrität zu gewährleisten, dürfen nur Administratoren Datensätze ändern. Über die REST-Schnittstelle können die Daten nicht manipuliert werden.
* Um sicherzustellen das wir die Daten nicht verlieren, führen wir Backups der Daten durch.

# Data technology

• Which database technology will you use and why? • How will the data be migrated? • What output will your query or analysis provide based on the input data?

Technologie: Wir verwenden MYSQL, da dieses Modul auf MYSQL basiert und wir gerne unser Wissen über MYSQL vertiefen möchten.

Daten Migration: Die öffentlich zur Verfügung stehenden Daten werden mittels einem CSV-Import importiert. Das CSV kommt von <https://data.world/tylerudite/atp-match-data>.

Funktionalitäten: Unser Projekt hat folgende 2 Funktionalitäten:

* Resultate von vergangen Turniere auszugeben.
* Prognosen über zukünftige oder „fantasy“ Turniere zu geben.

Dies gibt folgenden Input und Output:

Userinput => Vergangenes Turnier oder Turnier mit allen Teilnehmner und anderen Turnier Merkmalen.

Output => Turnierresultat in form eines JSON.

CREATE Script

CREATE TABLE Series (

sId int NOT NULL AUTO\_INCREMENT,

name varchar(255),

PRIMARY KEY (sId)

);

CREATE TABLE Tournament (

tId int NOT NULL AUTO\_INCREMENT,

sId int NOT NULL,

name varchar(255),

location varchar(255),

PRIMARY KEY (tId),

FOREIGN KEY (sId) references Series(sId)

);

CREATE TABLE Round (

rId int NOT NULL AUTO\_INCREMENT,

name varchar(255),

PRIMARY KEY (rId)

);

CREATE TABLE Player (

pId int NOT NULL AUTO\_INCREMENT,

firstName varchar(255),

lastName varchar(255),

countryCode varchar(5),

weight int,

height int,

birthYear int,

handed ENUM('right','left'),

backhand ENUM('one','two'),

PRIMARY KEY (pId)

);

CREATE TABLE Pitch (

piId int NOT NULL AUTO\_INCREMENT,

court ENUM('indoor','outdoor'),

surface ENUM('hard','grass', 'clay'),

PRIMARY KEY (piId)

);

CREATE TABLE Matches (

mId int NOT NULL AUTO\_INCREMENT,

date Date,

l1 int,

l2 int,

l3 int,

l4 int,

l5 int,

w1 int,

w2 int,

w3 int,

w4 int,

w5 int,

wrank int,

lrank int,

lpts int,

wpts int,

bestOf int,

tId int,

rId int,

wpId int,

lpId int,

piId int,

PRIMARY KEY (mId),

FOREIGN KEY (tId) references Tournament(tId),

FOREIGN KEY (rId) references Round(rId),

FOREIGN KEY (wpId) references Player(pId),

FOREIGN KEY (lpId) references Player(pId),

FOREIGN KEY (piId) references Pitch(piId)

);