Extra Course Project 2024/2025 Databases

Antonio Walter De Fusco

Alessio Del Sorbo

Index

Detailed description

The goal set is to create an application for the management of professional tennis tournaments and the relationship between professionals in this field.

What is meant to be managed is the combination of matches between the various tennis players and the ranking in the respective circuits by the professionals, whether male or female. The relationships of the professionals with the sponsors and with the coaches will also be managed.

A player is identified by an ID and has the following attributes:

- Personal data composed of:
  - Name;
  - Surname;
  - Age;
- I earn
- Points obtained

A player can be female or male. Each player has a ranking established by an organizing body. Male players have a ranking in the ATP, female players have a ranking in the WTA circuit, both male and female players have a ranking in the ITF circuit.

An organizing body is identified by an ID and has the following attributes:

- President
- Name
- Website
- Year
- First place

Each organizing body organizes different tournaments.

A tournament is identified by ID and has the following attributes:

- Points
- Typology
- Name

- Year
- Location
- Stadium
- Surface (e.g. dirt, concrete or grass)
- Cash prize

Each tournament consists of several matches where a set number of players meet.

Each match is identified by the tournament in which it was played, the date and time. It also has the following attributes:

- Player Set 1
- Player Set 2

Each player represents several sponsors throughout his career and is likewise coached by several managers.

Each sponsor is identified by name and has the following attributes:

- Type (the type of sponsor's products)
- Motto
- Color

Each coach is identified by an ID and has the following attributes:

- Name
- Surname
- Former player (indicates whether or not he was a former player)

Description of operations

The following operations are present within the platform:

- Sponsor and Coach Management:
  - Inserting a new coach;
  - Insertion of a new representation between a player and a sponsor;
  - Printout of all sponsors represented by a given player;
  - Print all players trained by a particular coach;
  - Printout of all sponsors that are represented by players trained by a particular coach.

- Player, tournament and match management:
  - Creating a new tournament;
  - Annual printout of name, surname, points obtained and earnings of each player;
  - Print the winner of a tournament given the name of the tournament and the year it took place;
  - Printout of points obtained by a player in the last year;
  - Insert the points obtained by each player in his career into the points_obtained attribute of Player.

- Ranking and entity management:
  - Update of the first place for all institutions;
  - Print the ranking (position and points) of the players in the ATP and update the position in the ranking;
  - Print the ranking (position and points) of the WTA players and update the position in the ranking;
  - Print the ranking (position and points) of the ITF players and update the position in the ranking;
  - Printout of all tournaments a player has participated in.

Frequent operations

Among which the 3 most frequent operations are:

- Insertion of a new representation between a player and a sponsor;
- Annual printout of name, surname, points obtained and earnings of each player;
- Printout of all tournaments a player has participated in.

Specification Analysis

Nouns:

- Player(ID, Personal Data<Name, Surname, Age>, Earnings, Points_obtained)
    - Female
    - Male
- Sponsor(Name, Type, Motto, Color)
- Coach(ID , Name, Surname, Ex_player)
- Organizing Body (ID , Name, President, Website, Year, First Place)
- Tournament(ID , Name, Type, Year, Location, Stadium, Surface, Prize_money, Points)
- Match(Tournament , Date , Time , Player_1 Set , Player_2 Set)

Verbs:

- Representing a Sponsor
    - Representation [Player, Sponsor]
        - (Contract_Duration, Compensation)
- Preparing a player
    - Preparation [Coach, Player]
        - (Start_date, End_date)

- Position yourself in a ranking
    - WTA Ranking [Female Player, Organizing Body]

        - (Position)
    - ATP Ranking [Male Player, Organizing Body]
        - (Position)
    - ITF Ranking [Player, Organizing Body]
        - (Position)
- Organize a tournament
    - Organize [Organizing Body, Tournament]
- Composing a tournament
    - Composition [Match, Tournament]
- Take part in a match
    - Encounter [Player, Match]
        - (Role, Winner)
- Participate
    - Participation [Player, Tournament]
        - (Position)

Glossary of terms

Term	*Description*	*Links*
Player	The professional player who can represent one or more sponsors, be trained by multiple coaches, play several matches and participate in tournaments. Has a position in the ITF ranking	Representation, Female Player, Male Player Match, Participation, Preparation, ITF Ranking
Female	The player who has a position in the WTA women's ranking	Player, WTA Ranking
Male	The player who has a position in the ATP men's ranking	Player, ATP Ranking
Sponsor	The company or brand that chooses to be represented by a player	Representation
Trainer	The professional figure who has the aim of preparing players from a physical and technical point of view	Preparation
Organizing body	The company or circuit that organizes professional tournaments and has its own ranking	Organization, WTA Ranking, ATP Ranking, ITF Ranking
Tournament	The tennis competition in which athletes take part and to which several matches belong	Organization, Participation, Composition
Match	The tennis match in which two players compete against each other	Composition, Meeting

Data Dictionary

Entity

Entity`	Description	Attributes	Identifier
Player	A tennis player who can participate in a tournament, meet another player in a match, represent a sponsor, be trained by a coach and have a ranking with an organizing body.	● Personal data: Attribute composed of: ○ Domain Name: String of up to 20 characters; ○ Surname: String of up to 20 characters; ○ Age: Whole positive ● Gain: Redundant attribute Domain: Positive integer; ● Points earned: Redundant attribute Domain: Positive integer;	ID: A unique auto-incremental code

Tournament	A tournament that is organized by an organizing body consists of matches in which players can participate.	• Points Domain: positive integer • Typology Domain: String of up to 20 characters; • Name Domain: String of up to 40 characters; • Year Domain: Positive integer (YEAR in sql); • Location Domain: String of up to 20 characters; • Stage: Multi-valued attribute Domain: String of up to 20 characters for each element; • Surface Domain: String of up to 20 characters; • Cash Prize Domain: Positive integer;	ID: A unique auto-incremental code

Organizing Body	An organization that organizes tournaments and assigns a ranking to players who participate in the tournaments	• First place Domain: Positive integer • Year Domain: Positive integer (YEAR in sql); • Name Domain: String of up to 20 characters; • Website Domain: String of up to 40 characters; • President Domain: String of up to 20 characters	ID: Auto-incremental identification code
Match	A game that is part of a tournament and allows players to meet each other. Weak entity	• Set_Player_1 Domain: Positive integer; • Set_Player_2 Domain: Entire positive; • Now Domain: Positive integer; • Date Domain: DATE	Date, Time and ID(Tournament) The date and time a match is played are uniquely associated with a tournament.
Sponsor	A sponsor who has a player represent his brand	• Typology Domain: String of up to 50 characters; • Motto Domain: String of up to 100 characters; • Color Domain: String of up to 20 characters;	Name: Unique sponsor string

Trainer	A coach who can train one or more players	• Surname Domain: String of up to 20 characters; • Name Domain: String of up to 20 characters; • ex_player Domain: integer value(0 - 1)	ID: Auto-incremental unique code
Female	A female player who has a ranking with the organizing body WTA. One of the child entities that have Player as their father		ID (Player): By the inheritance property of generalizations, the daughters of this one have the identifier of the father.
Male	A male player who has a ranking with the organizing body ATP. One of the child entities that have Player as their father		ID (Player): By the inheritance property of generalizations, the daughters of this one have the identifier of the father.

Associations

Associations	Description	Entities Involved	Attributes
Participation	A player's participation in a tournament	Player, Tournament	Position: Player Placement in a Tournament Domain: Positive integer
Composition	The composition of a tournament given by the union of several matches	Tournament, Match	

Encounter	The meeting of two players in a game	Player, Match	• Winner: If the player won the match Domain: Positive integer • Role: If it's player 1 or 2 Domain: Positive integer
Representation	The representation of a sponsor by a player	Player, Sponsor	• Compensation: value of sponsor representation Domain: positive integer; • Contract_Duration: duration in years of brand representation Domain: Positive integer
Preparation	The physical preparation of one or more players by a coach	Player, Coach	• Start_date Preparation start date Domain: DATE • End_date Preparation end date Domain: DATE
Organization	Organization of tournaments by an Organizing Body	Organizing Body, Tournament	
WTA Rankings	A player's ranking at an Organizing Body	Female, Organizing Body	Position: Position in the corresponding ranking by a player Domain: positive integer
ATP Rankings	A player's ranking at an Organizing Body	Male, Organizing Body	Position: Position in the corresponding ranking by a player Domain: positive integer
ITF Rankings	The ranking of a player, male or female, at an Organizing Body	Player, Organizing Body	Position: Position in the corresponding ranking by a player Domain: positive integer

Associations	Cardinality
Participation	Player (0,N) - Tournament(1,N) Maximum: M,N A player can participate in multiple tournaments, one or more players participate in a tournament
Composition	Tournament(1,N) - Match(1,1) Maximum: 1,N A tournament is made up of several games, a game makes up a single tournament
Encounter	Player(0,N) - Match(1,N) Maximum: M,N A player can participate in multiple games, multiple players can participate in a game
Representation	Player(1,N) - Sponsor(0,N) Maximum: M,N A player represents one or more sponsors, a sponsor can be represented by multiple players
Preparation	Player(1,N) - Coach(1,N) Maximum: M,N A player can have multiple coaches, a coach trains multiple players
Organization	Organizing Body(0,N) - Tournament(1,1) Maximum: 1,N An organizing body can organize multiple tournaments, a tournament is organized by only one organizing body
WTA Rankings	Female(1,1) - Organizing Body(1,N) Maximum: N,1 A female player is classified by only one organization, one organization ranks multiple female players
ATP Rankings	Male(1,1) - Organizing Body(1,N) Maximum: N,1 A male player is ranked by only one organization, one organization ranks multiple male players

ITF Rankings	Player(1,1) - Organizing Body(1,N) Maximum: N,1 A player is classified by only one organization, an organization classifies multiple players

Binding rules

Constraints expressed through constructs:
● A game consists of a maximum of two players; ● A male player cannot have a WTA ranking; ● A female player cannot have an ATP ranking ● A generic player can have an ITF ranking; ● A tournament must be organized by an organizing body; ● A tournament entry must have a winner; ● A tournament must have at least one match; ● A player may not play matches and consequently not participate in tournaments; ● A player, whether male, female or generic, must have a ranking with the respective organization; ● A player must have a sponsor; ● A player must be trained by at least one coach.
Non-expressible constraints:
● A player who has not played tournaments organized by his organization and consequently has no position in the ranking but his score is null and the position among the players with null scores is given by alphabetical order; ● If a player participates in a tournament he will play at least one match; ● A player wins or loses a match considering the difference in sets won against his opponent; ● Two matches cannot be played on the same date and at the same time.

Derivation rules

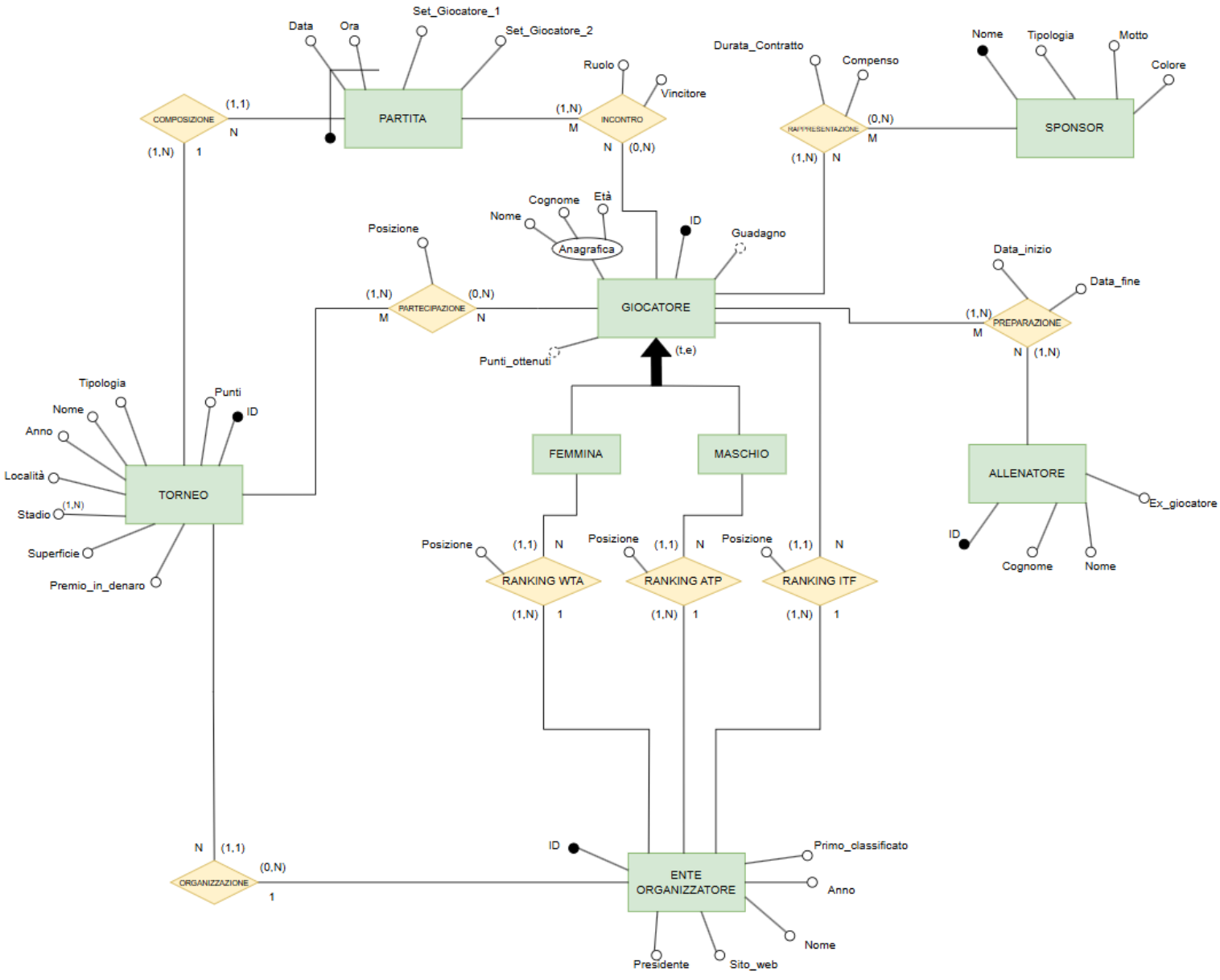
- The points obtained by a player are obtained:
  - Adding 100 for each participation in a tournament (The player gets points for his qualification);
  - Adding 50 for each match won;
  - Adding tournament points if your position in participation in that particular tournament is equal to 1.

- A player's earnings are calculated by adding the compensation from all sponsor representations to the prize money from tournaments in which the position is equal to 1;

Comments on the design choices made

- It was decided to generalise the player entity into male and female to allow players of all genders to participate in the corresponding ranking at the corresponding tournaments;
- The organizing bodies are mainly 3 characterized by the relationships that are attributed to the entity, when a calendar year ends a new body is established characterized by the year of activity;
- It was decided to make match a weak entity since the existence of a match is strictly associated with the existence of a reference tournament;
- It was decided that a coach will be able to train several players at the same time, giving players the opportunity to join together in preparation teams;
- It was decided to give sponsors the possibility to be represented by multiple players and the possibility for players to represent multiple sponsors;
- It was decided to make the organization of tournaments by the institutions mandatory in such a way that each match is necessary for the calculation of the ranking scores.

ER scheme


Data
Ora
Set_Giocatore_1
Set_Giocatore_2
PARTITA
COMPOSIZIONE
(1,1)
N
(1,N)
1
Ruolo
Vincitore
INCONTRO
(1,N)
M
N
(0,N)
Durata_Contratto
Compenso
RAPPRESENTAZIONE
(0,N)
M
(1,N)
N
Nome
Tipologia
Motto
Colore
SPONSOR
Cognome
Età
Nome
Anagrafica
ID
Guadagno
Posizione
PARTECIPAZIONE
(1,N)
M
(0,N)
N
GIOCATORE
Punti_ottenuti
(t,e)
Data_inizio
Data_fine
PREPARAZIONE
(1,N)
M
N
(1,N)
Tipologia
Nome
Anno
Località
Stadio
(1,N)
Superficie
Premio_in_denaro
Punti
ID
TORNEO
FEMMINA
MASCHIO
ALLENATORE
Ex_giocatore
ID
Cognome
Nome
Posizione
(1,1)
N
RANKING WTA
(1,N)
1
Posizione
(1,1)
N
RANKING ATP
(1,N)
1
Posizione
(1,1)
N
RANKING ITF
(1,N)
1
N
(1,1)
ORGANIZZAZIONE
(0,N)
1
ID
ENTE ORGANIZZATORE
Primo_classificato
Anno
Nome
Presidente
Sito_web

Application load specification

The database will have to manage all the tennis players who practice the sport in a professional context and who participate in the tournaments organized by the most important organizations in the world of tennis.

- About 7000 players are stored within the database

Of these, 60% are men and 40% are women.
Additionally each player:

- It represents approximately 10 sponsors per year out of a total of approximately 150 sponsors most present in the sector

- He participates in an average of 12 tournaments and plays approximately 5 matches in each tournament.

There are approximately 2000 active coaches and each player usually changes 1 during the course of the year.

The database stores the 3 most important organizations in the world of professional tennis, namely ATP, WTA and ITF. Specifically:

- The ATP organizes 60 tournaments a year
- The WTA organizes 50 tournaments a year
- The ITF organizes 500 tournaments a year

At least 50 matches are played for each tournament.

The main operations to be performed on this database are 15, of which the 3 most frequent are:

- Op. 1 - Inserting a new representation between a player and a sponsor (deductible frequency)
- Op. 2 - Annual printout of name, surname, points obtained and earnings of each player
- Op. 3 - Print all tournaments a player has participated in. (1/month)

Performance Analysis on ER Schema

Table of volumes

Name	*Type*	*Volume*
Player	E	7,000
Representation	R	70,000
Sponsor	E	150
Preparation	R	14,000
Trainer	E	2.000
Encounter	R	420,000
Match	E	30,500
Participation	R	84,000
Tournament	E	610
Female	SE	2.800
Male	SE	4.200
WTA Rankings	R	2.800
ATP Rankings	R	4.200
ITF Rankings	R	7,000
Organizing body	E	3
Organization	R	610
Composition	R	30,500

Table of operations

Operation	Type	Frequency
Op. 1	I	70,000/year
Op. 2	B	1/year
Op. 3	B	12/year

Access tables

Redundant attributes:

- #Points earned (Player)
- Earnings (Player)

Development of tables based on cases

Case 1 (Both attributes redundant)

Op. 1

Name	*Construct*	*Accesses*	*Type*
Representation	R	1	W
Player	E	1	R
Player	E	1	W

Op. 2

Name	*Construct*	*Accesses*	*Type*
Player	E	7,000	R

Op. 3

Name	*Construct*	*Accesses*	*Type*
Participation	R	84,000	R
Tournament	E	610	R

Total application load for case 1

- Op. 1 = 70,000 * (2S + 1L) = 70,000 * 5 = 350,000 hits/year
- Op. 2 = 1 * (7.000L) = 7000 accesses/year
- Op. 3 = 12 * (84,000L + 610L) = 12 * 84,610 = 1,015,320 accesses/year
- Redundant attributes = 2 * 4 bytes * 7,000 = 56,000 bytes

Total = (350,000 + 7,000 + 1,015,320) hits/year + 56,000 bytes = 1,372,320 accesses/year + 56,000 bytes on disk.

Case 2 (Only #Points obtained redundant)

Op. 1

Name	*Construct*	*Accesses*	*Type*
Representation	R	1	W

Op. 2

Name	*Construct*	*Accesses*	*Type*
Player	E	7,000	R
Representation	R	70000	R
Participation	R	84,000	R
Tournament	E	610	R

Op. 3

Name	*Construct*	*Accesses*	*Type*
Participation	R	84,000	R
Tournament	E	610	R

Total application load for case 2

- Op. 1 = 70,000 * (1S) = 70,000 * 2 = 140,000 hits/year

- Op. 2 = 1 * (7,000L + 70,000L + 84,000L + 610L) = 154,617 accesses/year

- Op. 3 = 12 * (84,000L + 610L) = 12 * 84,610 = 1,015,320 accesses/year

- Redundant attributes = 1 * 4 bytes * 7,000 = 28,000 bytes

Total = (140,000 + 154,617 + 1,015,320) hits/year + 28,000 bytes = 1,309,937 hits/year + 28,000 bytes on disk

Case 3 (Redundant Gain Only)

Op. 1

Name	*Construct*	*Accesses*	*Type*
Representation	R	1	W
Player	E	1	R
Player	E	1	W

Op. 2

Name	*Construct*	*Accesses*	*Type*
Player	E	7,000	R
Participation	R	84,000	R
Tournament	E	610	R
Encounter	R	420,000	R

Op. 3

Name	*Construct*	*Accesses*	*Type*
Participation	R	84,000	R
Tournament	E	610	R

Total application load for case 3

- Op. 1 = 70,000 * (2S + 1L) = 70,000 * 5 = 350,000 hits/year

- Op. 2 = 1 * (7,000L + 84,000L + 610L + 420,000L) = 511,610 accesses/year

- Op. 3 = 12 * (84,000L + 610L) = 12 * 84,610 = 1,015,320 accesses/year

- Redundant attributes = 1 * 4 bytes * 7,000 = 28,000 bytes

Total = (350,000 + 511,610 + 1,015,320) hits/year + 28,000 bytes = 1,876,930 hits/year + 28,000 bytes on disk

Case 4 (No redundant attributes)

Op. 1

Name	*Construct*	*Accesses*	*Type*
Representation	R	1	W
Player	E	1	R
Player	E	1	W

Op. 2

Name	*Construct*	*Accesses*	*Type*
Player	E	7,000	R
Participation	R	84,000	R
Representation	R	70,000	R
Tournament	E	610	R
Encounter	R	420,000	R

Op. 3

Name	*Construct*	*Accesses*	*Type*
Participation	R	84,000	R
Tournament	E	610	R

Total application load for case 4

- Op. 1 = 70,000 * (2S + 1L) = 70,000 * 5 = 350,000 hits/year
- Op. 2 = 1 * (7,000L + 84,000L + 70,000 + 610L + 420,000L) = 581,110 accesses/year
- Op. 3 = 12 * (84,000L + 610L) = 12 * 84,610 = 1,015,320 accesses/year

Total = (350,000 + 581,610 + 1,015,320) visits/year = 1,946,930 visits/year

Schema Restructuring

Changes made

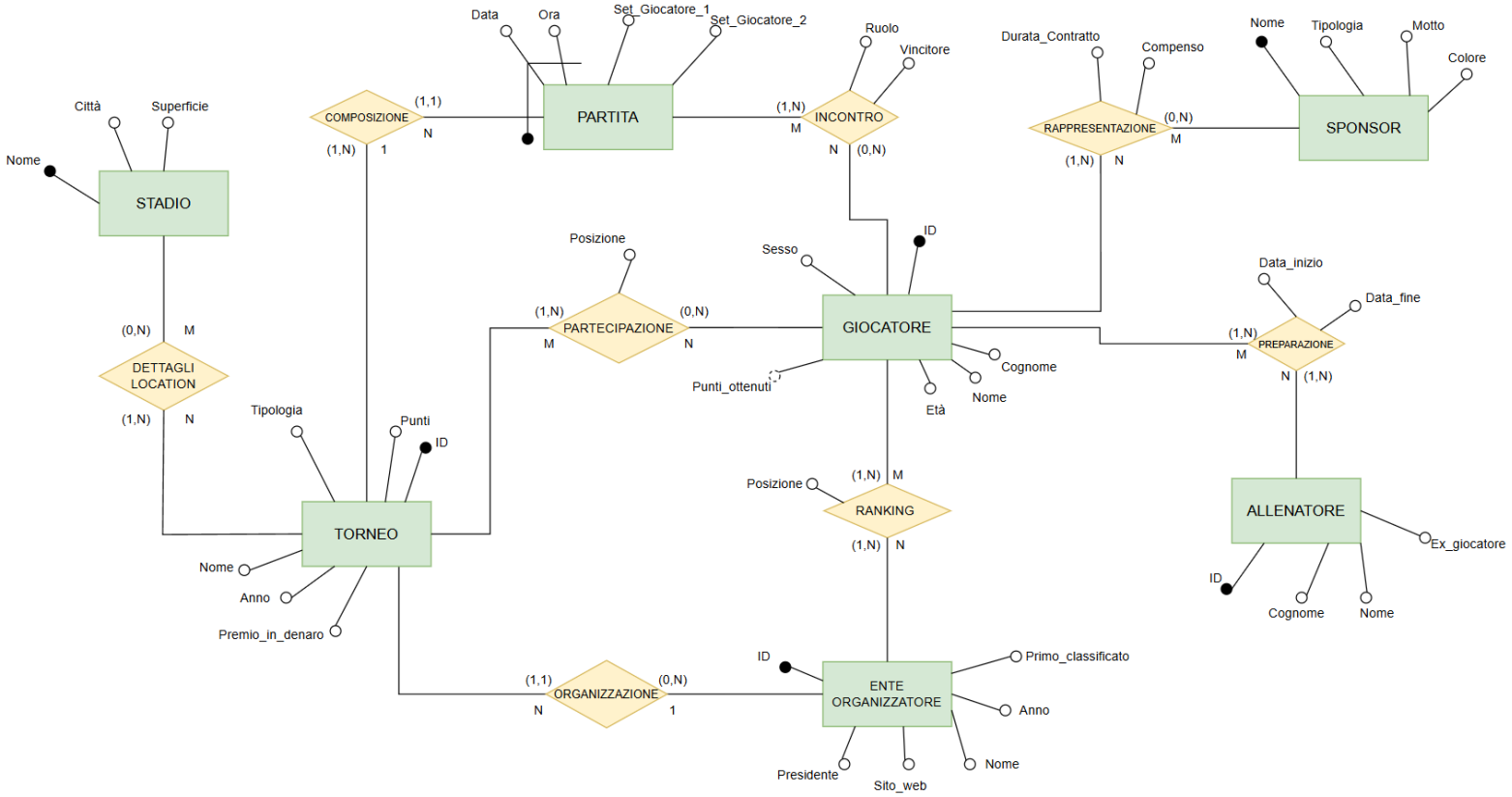
- Elimination of the hierarchy with a technique of merging the daughters into the father. An attribute "sex" is then introduced that will indicate whether the player is male or female. The 3 relations are merged into one called "Ranking" that preserves the attribute "Position".

- The composite attribute “Personal Data” of the player entity is eliminated by associating all sub-attributes to the entity.

- The multi-valued attribute “Stadium” of the Tournament entity is separated by creating a separate entity called “Stadium” that also includes the Tournament attributes regarding Location and Surface. The “Location Details” relationship is then introduced.

- The analysis of the application load showed that the best solution is to keep the redundant attribute “#Points Obtained” and eliminate “Gain” instead.

Refactored ER Scheme



Logical Mapping