raising the knowledge flowers

FlowR

At FlowR, we put our knowledge in the service of progress.

How do we achieve this? By combining our skills with our imagination, we design databases that are well-documented and perfectly adapted to the needs of our clients.

We are on a permanent quest to enhance the solution to our task.

That is, to blossom.

Project Manager: Livia Toca

From: National College of Computer Science TUDOR VIANU Bucharest

Project Description

Over the past decades, medical researchers and doctors have found cures for a series of infectious diseases, but there are many that still represent a constant threat to the health and lives of millions of people worldwide. Moreover, new diseases have emerged.

Efforts Up to Now

There are many health professionals and organizations specialized in the tracking of infectious diseases around the world that have directed their efforts towards countering infectious diseases.

These efforts were first based on data which was gathered at a national level and only updated from time to time, which was not enough for controlling wide-spread epidemics.

Direct collaboration was a problem, as well.

Making a Difference

One of the fundamental needs of the members global medical community concerned with infectious disease is to be provided with a means of instantly accessing information and constantly being in touch with each other.

Health professionals involved in this project will form a global medical union, that will be useful to the entire medical community.

The beneficiaries will be the members of the medical union, and, indirectly, health organizations, insurance companies and state governments, through the reports that will be issued on a regular basis.

Our Vision and Objectives

When we were faced with the idea of modeling a global database that would contain information on infectious diseases, we first thought of a structure that would comprise logically all data and would be used for statistics and informing.

The medical union members will share opinions on a constantly updated collection of relevant data regarding the most recent cases of infectious diseases across the globe.

The database will only be concerned with sharing information on disease cases and with medical collaboration; we will not provide an in-depth documentation on the medical theory regarding infectious diseases.

The members of the medical union should be able to:

- introduce data about the disease cases they have encountered
- get the information they need regarding medical cases that other health professionals had dealt with
- communicate and collaborate

To sum up, the database needs to have the following characteristics:

- Flexibility: The data structures must be created in such a way that they would fit to all situations that might occur.
- Comprehensiveness: It must contain all information required to give a detailed view on the different aspects of tracking infectious diseases.
- Relevancy: The data modeled in the database must be chosen according to the possibility of it being used in statistical reports on a global scale and in tracking the trends of infectious diseases.

Internet Research

on the topic of infectious diseases:

- www.who.int

World Health Organization, the United Nations specialized agency for health.

- www.isid.org

International Society for Infectious Diseases, an organization which enhances research, prevention, and treatment of infectious diseases.

- www.cdc.gov/ncidod

The official site of the Centers for Disease Control and Prevention.

on the infectious diseases databases **already existent:**

- www.gideononline.com

The homepage of GIDEON, the world's premier global infectious diseases database.

- www.diseasesdatabase.com

A database with an index of human diseases, medications, symptoms.

raising the knowledge flowers

FlowR

References

Books and Articles on the topic of infectious diseases:

- International travel and health. Geneva: World Health Organization, 2005.
- Rajnik, Michael. “Rhinoviruses”. E-Medicine. October 18, 2006.
- Lemonick, Michael. Park, Alice. “The truth about SARS”. Time. May 5, 2003.

Open Meeting with:

Mrs. Cristina Popescu
Infectious Disease Specialist from: “Matei Bals” Institute of Infectious Diseases, Bucharest

Mr. Gabriel Popescu
Infectious Disease Specialist from: “Matei Bals” Institute of Infectious Diseases, Bucharest

Interviews with:

Mrs. Gabriela Cozmanciuc
Doctor from Euroclinic, Bucharest

Mrs. Cristiana Oprea, M.D., Ph.D.
Researcher from “Dr.Victor Babes” Clinical Hospital of Infectious and Tropical Diseases, Bucharest

The Interviews

Previous to the first interview, we had conducted some research over the Internet regarding the infectious diseases and the existing databases concerned with the information regarding infectious diseases.

We understood:

- what a health professional is
- the topic of infectious diseases

We attended an open meeting, where we talked with **Mrs. Cristina Popescu** and **Mr. Gabriel Popescu**, two ID specialists from the Matei Bals Institute of Infectious Diseases who told us some generalities about the infectious diseases and the way in which they deal with patients who have this type of illnesses.

We were familiarized with:

- the following concepts: disease, symptom, sign and pathogen
- the procedure of establishing a disease diagnosis

After this meeting:

We had the idea of tracking the symptoms and signs that a case manifests, as well as the doctor's observations if the diagnosis he establishes for a case is not certain.

The Interviews

On the first interview, **Mrs. Gabriela Cozmanciuc** gave us a positive feedback about our understanding of the topic up to this point and about the early draft of our project. She confirmed the basic structure of our project and advised us to seek further development.

She also clarified some questions we had about:

- the notions of disease, pathogen and strain
- the evolution and phases of a disease
- the way in which a patient's location is determined

After this interview we included strains and phases in our conceptual model and improved our diagram.

On the second interview, **Mrs. Cristiana Oprea** offered us insight into research and into the communication methods the members of the medical community use.

She also clarified some questions we had about:

- the research work in the infectious diseases field
- the qualifications and licenses of health professionals
- the procedures regarding medical conferences

After the second interview we modeled qualifications and we developed forms of collaboration in our project.

About Users

- In order to have access to the database, health professionals must first register for a database account and thus become users. A user can only have one account.
- User data is entered by the applicant when he registers for an account. Specifically, the data includes the username, password, first and last name of the applicant, work place, medical licenses and contact information.
- The identity and medical licenses of every user need to be checked before he or she is granted the right to use the database. This is done using the first name, last name and work place provided by the user.
- The registration can be either approved or rejected. A rejected account is unusable, but the data provided continues to exist in the database until it is manually deleted.
- If the account is approved, the registration process is complete and the user can enter new data into the database, query the data that has been previously entered and communicate with other users.

Collaboration

- Users can collaborate by organizing themselves in discussion groups, virtual conferences or research projects, on different topics concerning infectious diseases.
- Any user can become the manager of a group, conference or project. He or she is automatically registered as the first member, sets the name, chooses the topic or creates a new one that better describes the collaboration and also has the right to designate another person to be the manager instead of himself or herself. The manager can dissolve a group, but the group's messages are kept into the database until they are deleted manually.
- Other users may then register to become members of projects and groups or to attend conferences.
- Registrations are approved or rejected by the manager. E-mail notifications are sent to users for each of these actions. The manager may also delete users' registrations, if needed.
- Users can cancel their registrations at any time, but they have the right to re-register following the same procedure.
- The members of a discussion group can write messages that are e-mailed to everyone in that group. Messages can be deleted by the manager. When a registration is deleted, all its messages are removed as well.
- Some of the discussion groups and conferences may be associated with a certain research project. Users who work on the research project may register to these groups and participate in the conferences.

Qualifications and Medical Facilities

- Users may have more than one qualification and must provide the license numbers of the certificates that give them the right to profess that qualification. License numbers are specific to each country.
- The data that users provide about the medical facility where they work is used to track the current location of disease cases treated by these users, in order to obtain regional statistics.

Dealing with Space and Climate

- Climate and geographical information, including the average temperature and humidity of each season, are needed into the database in order to locate disease cases, track epidemics and issue regional and climate-specific reports.

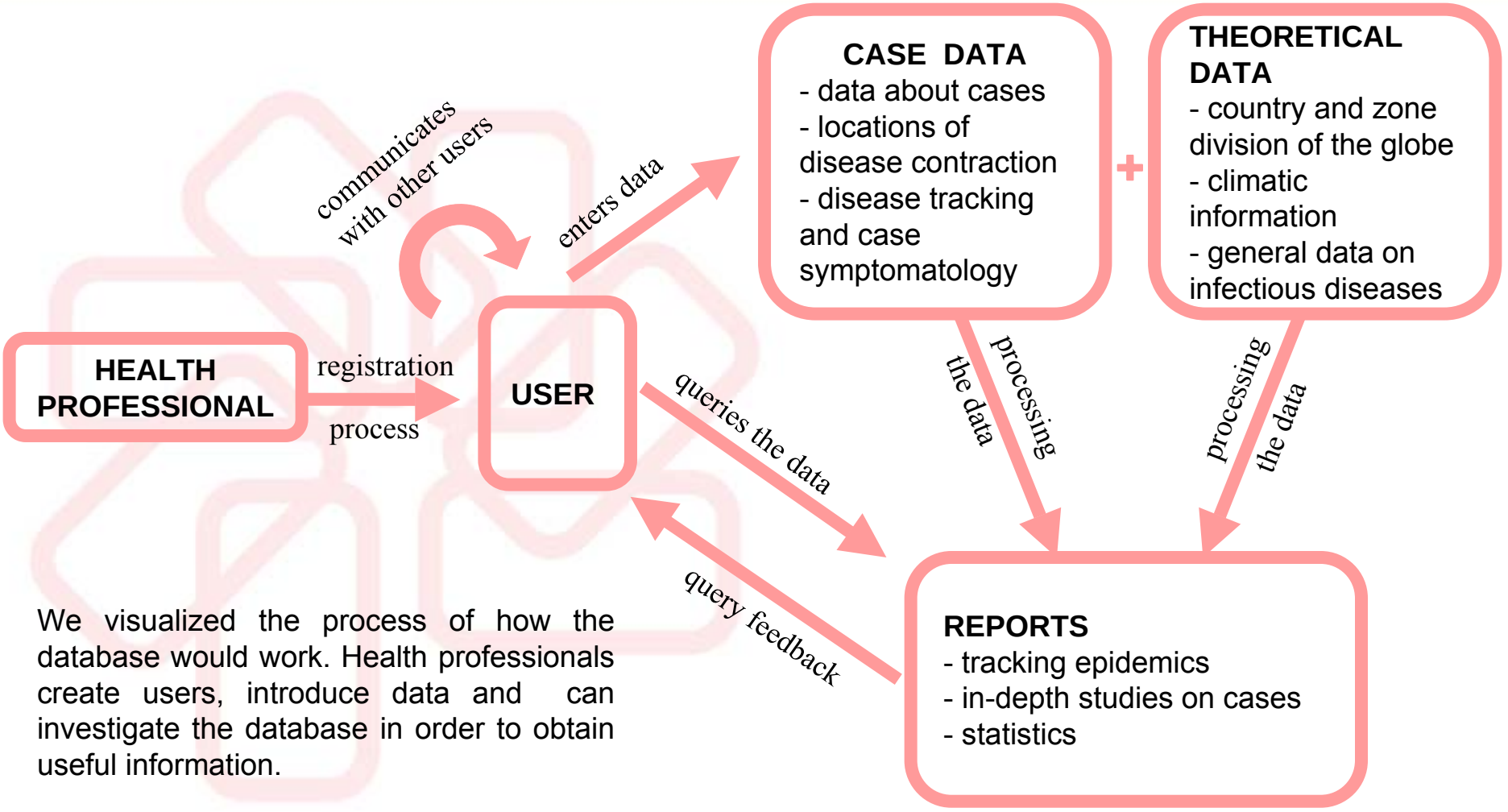
Diseases and Pathogens

- A pathogen may cause more than one disease: for example, the adenovirus causes both respiratory tract disease and conjunctivitis.
- A pathogen has several strains. Each strain locates in one organ or tissue. For example, the herpetic virus strain HSV1 affects the area around the mouth and the HSV2 strain locates itself in the genital area. Each of these strains also causes a different set of signs and symptoms.

Data about Cases

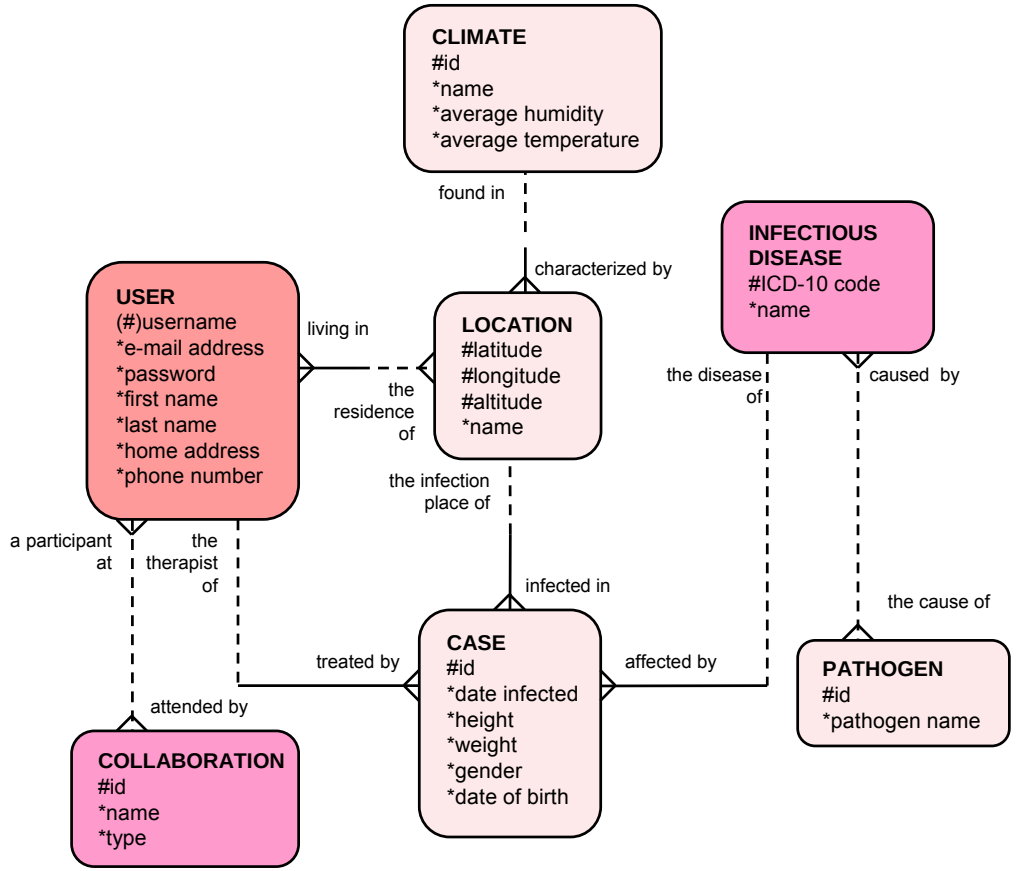
- Users enter data about disease cases encountered among their patients.
- Doctors always assign a preliminary diagnosis for a case, based on the symptoms and physical signs the patient has.
- Afterwards, they perform further investigations in order to identify the pathogen's strain responsible for the respective disease, which gives the firm diagnosis.
- When the doctor changes the diagnosis previously entered, he will document the case in the observations and comments sections.
- Periodically, doctors enter additional data about their cases by observing their statuses and making comments on the patients' condition. The data includes the signs and symptoms that the patient manifests.
- The evolution of a disease can be divided into several phases such as incubation, invasive phase, convalescence etc. For each case status, the phase of the disease that the patient is currently experiencing is also mentioned.
- The work domain of a patient is useful to establish connections between certain work environments and predispositions to different types of infectious diseases.
- The outcome of a case is entered after the pathogen has been eradicated from the organism, after the patient's death or if the doctor loses contact with the patient. After the patient is cured, the doctor still tracks the case by observing its aftereffects.

Data Flow Diagram



We visualized the process of how the database would work. Health professionals create users, introduce data and can investigate the database in order to obtain useful information.

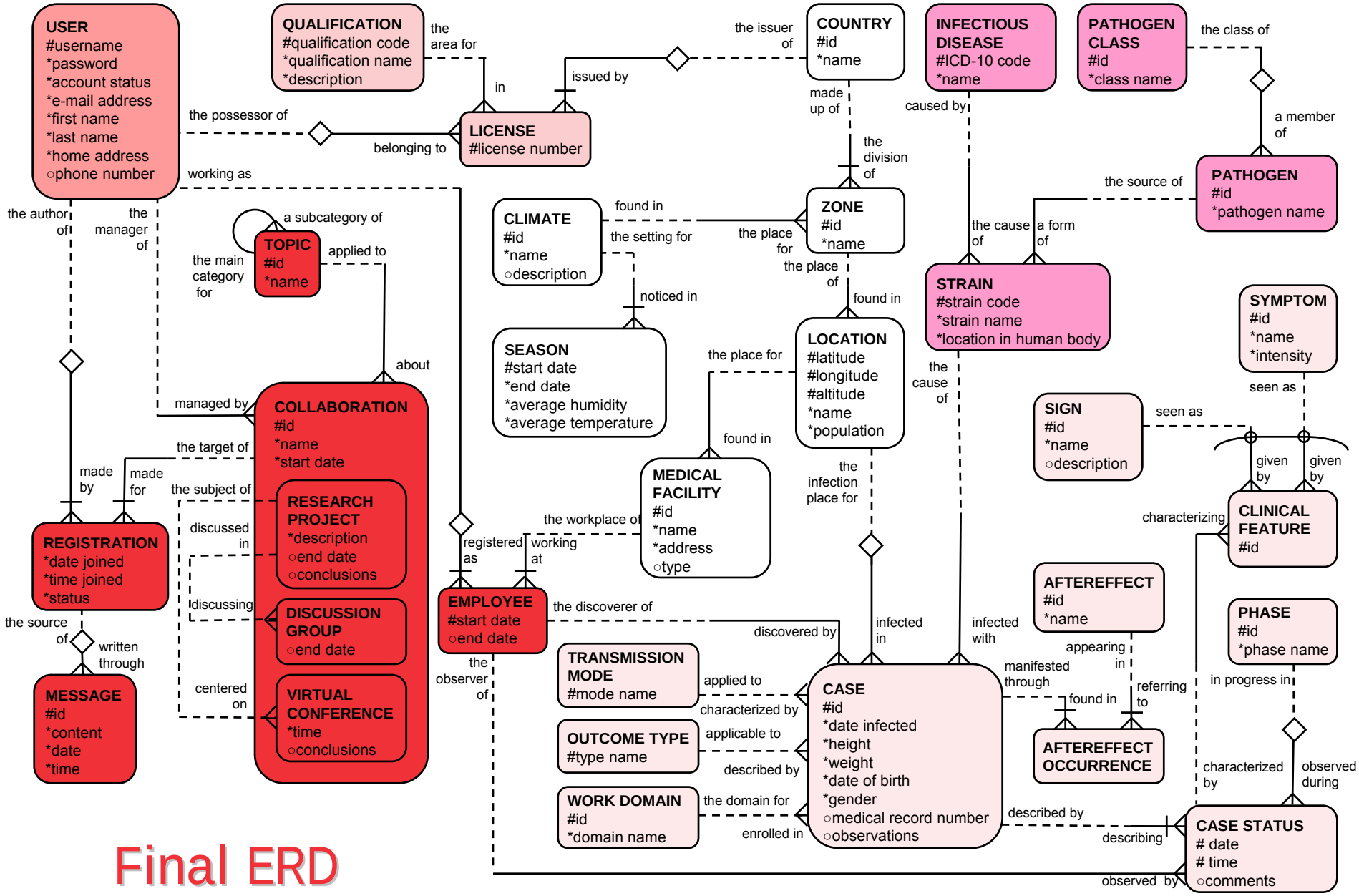
Preliminary ERD



The preliminary ERD contains the most important entities, that generate the fundamental relationships for the final ERD structure.

We converted this diagram to accurately represent our concept of how the final database would be structured by:

- normalizing the diagram
- adding other significant entities
- finding the attributes hidden in the intersection entities and the relationships which they generated
- identifying the errors
- solving many-to-many relationships


USER
#username
*password
*account status
*e-mail address
*first name
*last name
*home address
○phone number
QUALIFICATION
#qualification code
*qualification name
*description
the area for
in
issued by
the issuer of
COUNTRY
#id
*name
INFECTIOUS DISEASE
#ICD-10 code
*name
PATHOGEN CLASS
#id
*class name
the class of
a member of
PATHOGEN
#id
*pathogen name
the source of
the possessor of
belonging to
LICENSE
#license number
made up of
the division of
working as
the author of
the manager of
a subcategory of
TOPIC
#id
*name
the main category for
applied to
about
CLIMATE
#id
*name
○description
found in
the setting for
noticed in
SEASON
#start date
*end date
*average humidity
*average temperature
ZONE
#id
*name
the place for
the place of
found in
LOCATION
#latitude
#longitude
#altitude
*name
*population
caused by
the cause of
a form of
STRAIN
#strain code
*strain name
*location in human body
SYMPTOM
#id
*name
*intensity
seen as
SIGN
#id
*name
○description
given by
CLINICAL FEATURE
#id
characterizing
managed by
COLLABORATION
#id
*name
*start date
the target of
made by
made for
the subject of
RESEARCH PROJECT
*description
○end date
○conclusions
discussed in
DISCUSSION GROUP
○end date
discussing
VIRTUAL CONFERENCE
*time
○conclusions
centered on
REGISTRATION
*date joined
*time joined
*status
the source of
written through
MESSAGE
#id
*content
*date
*time
registered as
working at
the workplace of
MEDICAL FACILITY
#id
*name
*address
○type
found in
the place for
the infection place for
EMPLOYEE
#start date
○end date
the discoverer of
the observer of
discovered by
infected in
infected with
TRANSMISSION MODE
#mode name
applied to
characterized by
OUTCOME TYPE
#type name
applicable to
described by
WORK DOMAIN
#id
*domain name
the domain for
enrolled in
CASE
#id
*date infected
*height
*weight
*date of birth
*gender
○medical record number
○observations
AFTEREFFECT
#id
*name
appearing in
manifested through
found in
referring to
AFTEREFFECT OCCURRENCE
PHASE
#id
*phase name
in progress in
characterized by
observed during
described by
describing
CASE STATUS
date
time
○comments
observed by

Final ERD

ERD Explanations

Users

USER

When a health professional registers into the database, an instance of the USER entity is created. If a username is taken, even if the account is still in the trial period, it cannot be used by another person. However, this is possible after an account has been manually deleted by the database administrator.

After the registration has been made and before the account is approved, the account status is set to pending.

If the account has been approved, its status becomes active; otherwise, it is set to rejected.

EMPLOYEE

An intersection entity between USER and MEDICAL FACILITY. It is used to track a health professional's workplace since registration. If the end date field is blank, the user still works at the corresponding facility.

ERD Explanations

Collaboration

COLLABORATION	The entity details are entered by the manager when he or she chooses to initiate the collaboration, except for the conclusions of the project and conference.
REGISTRATION	When an user makes a registration, the status is “pending”. After confirmation by the manager, it becomes “approved”. If the registration is canceled, the status is “inactive”.
RESEARCH PROJECT	If the end date is not filled in, the project is in progress. The description states the purpose of the project and a short plan of action. Conclusions are filled by the manager after the project has ended.
VIRTUAL CONFERENCE	For virtual conferences, the manager posts the conclusion, not the full conference text; we also remember the date and time when the conference began.
MESSAGE	Messages that are e-mailed to group members are also kept in the database.

Qualifications and Medical Facilities

QUALIFICATION Instances of this entity are areas related to medicine in which a person can obtain licenses.

MEDICAL FACILITY The type of the medical facility refers to its profile (infectious diseases, AIDS, pediatric profile etc). If it has no profile, the field is left blank.

Space and Climate

ZONE Zones are subdivisions of a country (states in the US, departments in France, provinces in India, cantons in Switzerland, etc). Metropolises whose districts or sectors are recorded as separate locations are also considered zones.

LOCATION Represents a village, town, city, or in the case of large metropolises, a district or sector. The population attribute holds the number of inhabitants of that location at the last count. In the case of infections, we remember the location that is closest to the place of infection.

SEASON Each climate has a different set of seasons. Even if a season has the same name in two or more climates, it is recorded once for each climate where it appears, because it has different characteristics.

Diseases and Pathogens

INFECTIOUS DISEASE

The ICD-10 code is a unique identifier for diseases according to the International Classification of Diseases published by the WHO (World Health Organization) and it is used worldwide. The name attribute contains the scientific name of the disease.

STRAIN

The strain code is its short name, for example HSV1 or H5N3. The strain name is the official name according to the International Committee on Taxonomy of Viruses.

PATHOGEN CLASS

This entity is used to categorize pathogens according to their biological class: bacterial, viral, fungal etc.

SYMPTOM

Aspect of a disease that is described by the patient: pain, lack of appetite, insomnia, etc. Its intensity may be high, medium or low.

SIGN

Visible aspect of a disease's clinical manifestation: stains, rashes or changes in the skin color. The description is filled in by the health professional.

Data about Cases

CASE	The CASE entity holds data about a person that contracted an infectious disease. The height and weight are kept for the moment of the infection. The date infected is the approximate date when the patient contracted the disease.
WORK DOMAIN	The domain in which the patient was working when he contracted the infection. This includes the school environment in case of students.
OUTCOME	Cases for which the outcome has not been filled in are still in progress.
TRANSMISSION MODE	The way through which the case contracted the infectious disease: through blood, water, air, food and so on.
CASE STATUS	The status represents a stage in the evolution of a case when the doctor observes the signs and symptoms of the patient and makes comments. This data is useful for anyone who is interested in that particular case.
AFTEREFFECT	Condition of the patient that is caused by a disease and appears or persists when the pathogen is no longer present in the patient's organism.
AFTEREFFECT OCCURRENCE	Intersection entity between AFTEREFFECT and CASE. Provides a list of aftereffects that can be noticed after evolution of the case is complete.

ERD Explanations

Relationships

USER- MEDICAL FACILITY	At the same medical facility, there can be many health professionals who own database accounts. A health professional can have several work places at a time: for example, he or she can work both at a private clinic and at a public hospital. A health professional can also interrupt his or her work in a medical facility and then come back.
EMPLOYEE-CASE	Connects a case with the doctor that discovered it and with the medical facility where it has first been registered.
EMPLOYEE-CASE STATUS	Shows what health professional is currently monitoring the case and entering observations about it into the database. When the employment period of a user at a medical facility ends, the case either continues to be monitored by the same doctor at another facility or is transferred to another doctor.
CASE-LOCATION	The relationship indicates the place where the person contracted the disease. It can be different from the location of the hospital where the case was or is being treated.
CASE-WORK DOMAIN	The relationship is optional at both ends because the person might have been unemployed when the infection occurred.
INFECTIOUS DISEASE-STRAIN	Disease conditions caused by different strains of the same virus share the same ICD-10 code. For example, the J-09 code of Avian Influenza is applicable in the case of all H5N strains (H5N1, H5N2 etc).

Assumptions

We have assumed that:

- A conference, discussion group or research project has only one topic.
- Conferences, discussion groups and research projects must be initiated by a user in order to be created.
- Only one GPS position is recorded for a location.
- Only one climate exists in a specific zone.
- At a certain time, the case cannot be in more than one phase.
- We do not keep information about a health professional's cases that commenced prior to registration.
- Measurement units are recorded in metric:
  - case height: meters
  - case weight: kilograms
  - season temperature: degrees Celsius
  - season humidity: percents
  - location altitude: meters

Time Constraints

- For the RESEARCH PROJECT and EMPLOYEE, the start date must be before the end date.
- For the CASE, the date infected must be after the date of birth.
- For the CASE STATUS, the date observed must be after the date infected and between the start date and end date of the EMPLOYEE that observed it.
- For the REGISTRATION to a collaboration, the date joined and time joined attributes must be before the start date of the collaboration and the start time of the virtual conference.

Space Constraints

- Latitude and longitude are memorized in degrees, minutes, seconds and direction (east or west for longitude, north or south for latitude) and must be comprised within the maximum values: 180E to 180W degrees for longitude and 90N to 90N degrees for latitude.

Fixed Values for Attributes

- Account status: pending, active, rejected.
- Registration status: pending, active, inactive.
- Patient gender: male, female.
- Outcome type: cured, died, abandoned.

Sample Report 1:

Statistics for several countries regarding the number of new AIDS cases as a percent of the total population, for one year.

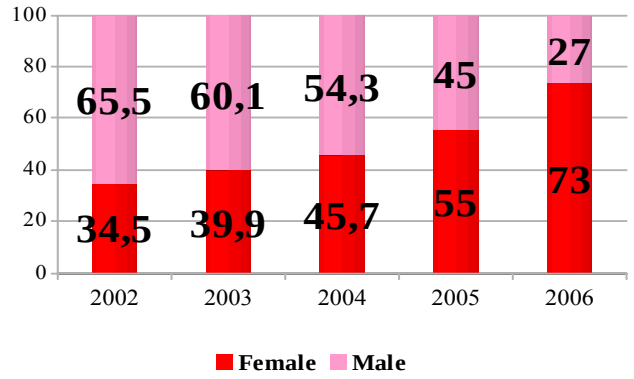
The data needed for this report concerns the location and infection date of the cases and the disease they are infected with.

Country ID	Country name	Percent infected
1	Angola	8.6%
2	Cambodia	1.6%
3	Ethiopia	4.4%
4	Thailand	1.5%
5	Zambia	16.5%
…	…	…

Sample Report 2:

Structure of the HIV infected population by gender – statistics in the period 2002-2006 for one country. Trends, such as (here) a gradual shift of the major virus bearers from men to women, can be noticed this way.

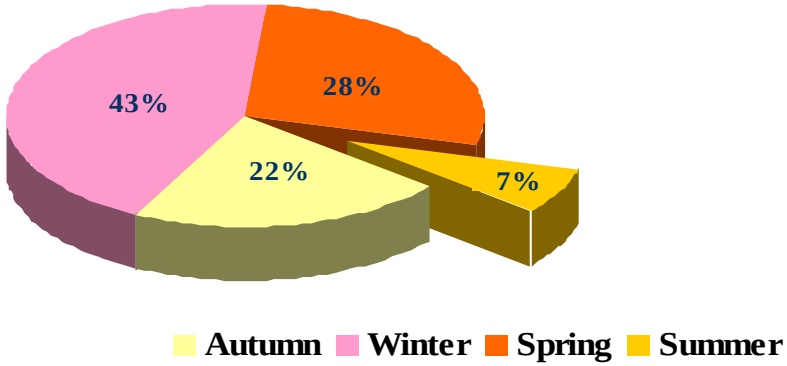
The report is based on the data about case gender, location, disease and infection date.



Sample Report 3:

The number of influenza cases in the course of an year – average values for the last decade, recorded in the temperate climate.

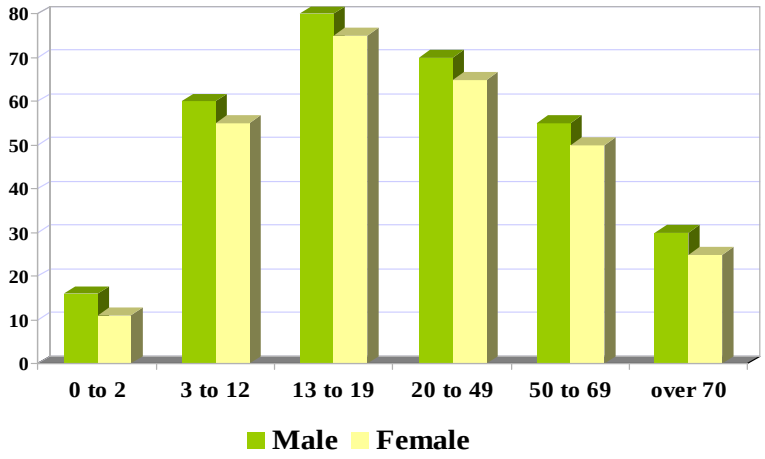
This report requires data about case infection date and the start and end dates of seasons.



Sample Report 4:

The graph represents the survivability rates of people diagnosed with tuberculosis, according to age and gender.

The report can be generated using the data about the outcome of infections and the patient's age and gender information.



Conclusions

- Our database aims to be a reliable resource not only for health professionals worldwide, but also for non-profit health organizations that are willing to provide funding for infectious diseases prevention and treatment.
- We consider that our database will be really helpful for doctors who encounter new cases of infectious diseases, as they can easily contact other health professionals that have experience in working with patients suffering from that particular disease.
- The database would also be important for research, as it allows health professionals to collaborate through discussion groups, projects and virtual conferences and access important information.
- The infrastructure of our database ensures trustworthy data storage and effective archive queries. It can also be easily used to generate reports, statistics at a global scale and track the trends of infectious diseases.

The personal accomplishments our team has gained are very rewarding:

- Strengthening our team-work abilities
- Learning the needs of a complex field, infectious diseases
- Earning a real-world knowledge