



**CSCI 5408 – Data Management, Warehousing, Analytics**

**Assignment 1**

**Work done by,**

**Name: Guturu Rama Mohan Vishnu**

**Banner ID: B00871849**

**Email: [rm286720@dal.ca](mailto:rm286720@dal.ca)**

## **DECLARATION**

**I, Guturu Rama Mohan Vishnu, declare that in assignment 1 of CSCI 5408 course, data scrapping is not done programmatically or using any online or offline tools. However, the webpages or the domain mentioned in this document are visited manually, and some useful information is gathered for education purpose only. Information, such as email, personal contact numbers, or names of people are not extracted. The course instructor or the Faculty of Computer Science cannot be held responsible for any misuse of the extracted data.**

## Problem #2: World Health Organization

### Solution:

After exploring the WHO (World Health Organization) website, the list of 25 entity sets that I believe would be good enough to make a good data model are listed below.

Entity	Reasons for considering	Source	Possible attributes
Questions and Answers	This can be considered as an entity since this can solve doubts of people. This is a strong entity since it doesn't depend on any other entity.	Information has been taken from <a href="https://www.who.int/news-room/questions-and-answers">https://www.who.int/news-room/questions-and-answers</a>	Questions, Answers, Topic
COVID-19 vaccines	There are numerous types of vaccines for covid and there is so much of information about each vaccine and who to use and who shouldn't. This is a strong entity.	Information has been taken from <a href="https://www.who.int/health-topics/coronavirus#tab=tab_1">https://www.who.int/health-topics/coronavirus#tab=tab_1</a>	Vaccine names, Type, Invented by, Approved in region
Regions	This can make a super entity and there are a few sub entities under this since there are multiple regions based on	Information has been taken from <a href="https://www.who.int/about/who-we-are/regional-offices">https://www.who.int/about/who-we-are/regional-offices</a>	Office ID, Country, Address, Email id, Phone number

	continents. Every region has its own office and own events.		
Multimedia	This is a strong entity since it is present to represent the portfolio of WHO in a very unique way, in form of videos, graphics and audios.	Information has been taken from <a href="https://www.who.int/multi-media">https://www.who.int/multi-media</a>	Topic, Countries/Areas, Year, Publishing offices, Multimedia type
Publications	All the major publications done by WHO team would be found here. They can be of type books, journals or series. This is a strong entity.	Information has been taken from <a href="https://www.who.int/publications">https://www.who.int/publications</a>	Name of the publication, Published date, Type of the publication, Editors, Reference number
WHO	To know about WHO or to know who they are, to get a knowledge of what they do, where they work, who they work with, this page is needed. This is a strong entity.	Information has been taken from <a href="https://www.who.int/about">https://www.who.int/about</a>	Year of their invention, History of WHO
COVID – 19 disease	COVID – 19 disease has become such an affecting and harming disease in the world with people getting affected every single minute. There are	Information has been taken from <a href="https://www.who.int/health-">https://www.who.int/health-</a>	Overview, Prevention, Symptoms, Myth busters, Advice for people

	<p>numerous precautions to take, measures to follow, advice for people from WHO team and myth busters. All this information can be found in this page.</p>	<a href="https://www.who.int/topics/coronavirus#tab=tab_1">topics/corona virus#tab=tab_1</a>	
Statements	<p>Any update or messages from WHO team to the public would be given in the form of statements. This is a strong entity as it has nothing to do with any other entity.</p>	<p>Information has been taken from <a href="https://www.who.int/news-room/statements">https://www.who.int/news-room/statements</a></p>	<p>Statement given, Given by whom, When was it given, Topic regarding the statement</p>
Campaigns	<p>Campaigns are a series of organized actions which are done for one purpose. Campaigns attract a lot of people and the information required to know all about the campaigns can be found here. It is a strong entity.</p>	<p>Information has been taken from <a href="https://www.who.int/campaigns">https://www.who.int/campaigns</a></p>	<p>Topic, Date of happening, Place of happening, Number of people attended</p>
Events	<p>Events are organized by the WHO team to spread awareness to the public regarding particular things and</p>	<p>Information has been taken from <a href="https://www.who.int/news-room/events">https://www.who.int/news-room/events</a></p>	<p>Purpose of event, Place, Date, Number of people attended</p>

	also to help people. It is a strong entity.		
Photo library	Photo library is the place where all the photos of events, campaigns and all the health topics are posted.	Information has been taken from <a href="https://photos.hq.who.int/">https://photos.hq.who.int/</a>	Topic, Photo taken by, Where the photo is taken, When was it taken
Speeches	Speeches are given on a particular topic by a presenter to announce something or to declare something or to spread awareness of something.	Information has been taken from <a href="https://www.who.int/director-general/speeches">https://www.who.int/director-general/speeches</a>	Presenter, Topic, Countries, Year, Speech type
Research	An important part of WHO's work with global expert networks is the R & D Blueprint, a global strategy and preparedness plan that triggers the rapid activation of research and development activities during outbreaks.	Information has been taken from <a href="https://www.who.int/emergencies/research">https://www.who.int/emergencies/research</a>	Research start date, Research end date, Research topic, Regarding which disease
Funding	All the information about contributors who provide funding or in-kind contributions to support the work WHO does, is in here.	Information has been taken from <a href="https://www.who.int/emergencies/funding">https://www.who.int/emergencies/funding</a>	Amount, Received from, Through, Date

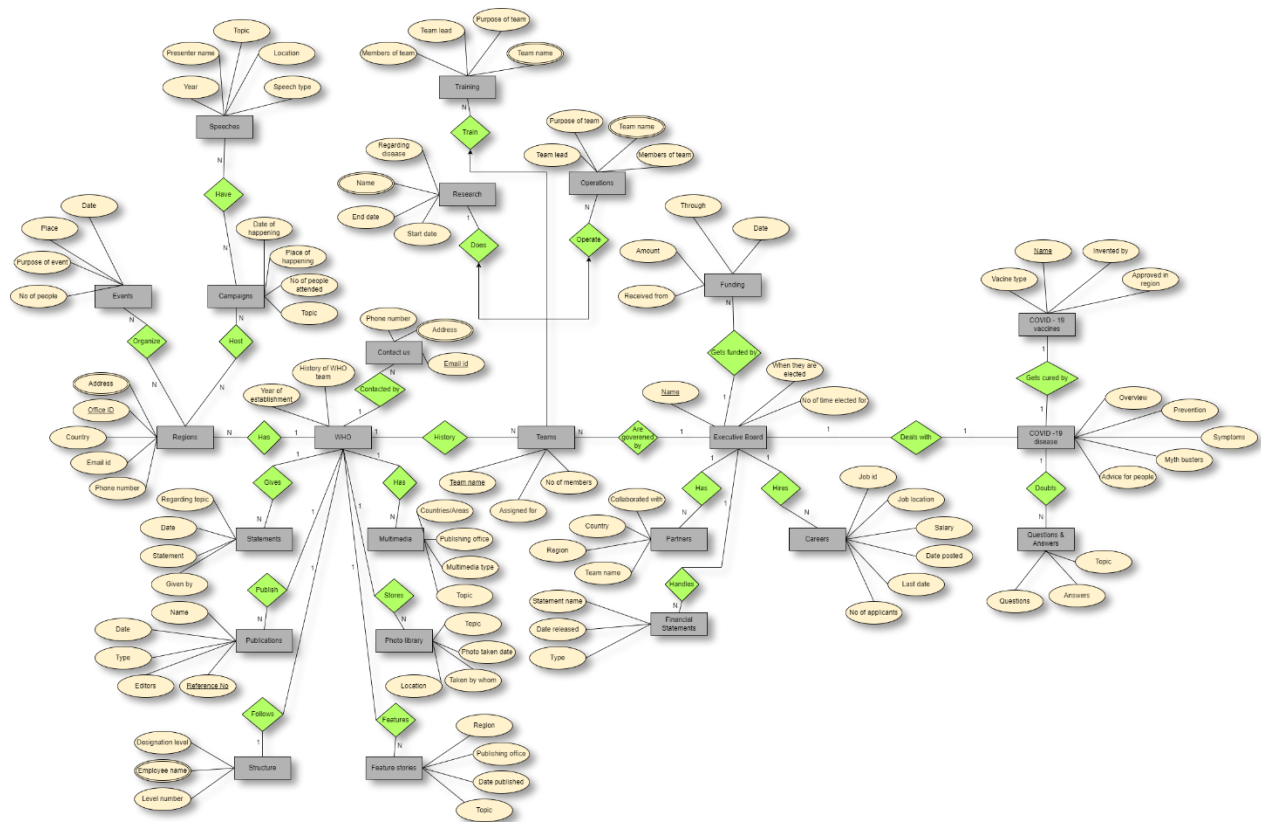
	Hence this is a perfect fit for an entity.		
Operations	The WHO's team supports best practices and standards for emergency operations centres and helps member states respond and detect public health emergencies more quickly.	Information has been taken from <a href="https://www.who.int/emergencies/operations">https://www.who.int/emergencies/operations</a>	Team name, Purpose of their team, Members of the team, Lead
Training	The training team at WHO targets at ensuring one billion people that are better protected from health emergencies.	Information has been taken from <a href="https://www.who.int/emergencies/training">https://www.who.int/emergencies/training</a>	Team name, Purpose of their team, Members of the team, Lead
Insights and Visualizations	Visual stories about health, wellbeing and healthcare driven by the data of WHO can be seen and known here. The visuals can be in the form of graphs, bar graphs, pie charts etc.	Information has been taken from <a href="https://www.who.int/data/stories">https://www.who.int/data/stories</a>	Visual type, Topic
Executive Board	The main members of the WHO team, the team which has the power to control WHO	Information has been taken from <a href="https://www.who.int">https://www.who.int</a>	Team member name, No of times

	is the executive board. The main functions of the board are to implement the decisions and policies of the Health Assembly, and advise and generally to facilitate it's work.	<a href="#">int/about/governance/executive-board</a>	elected for, When they are elected
Teams	Numerous teams are present in the World Health Organization to take care of individual purposes, assigned to their respective teams.	Information has been taken from <a href="https://www.who.int/teams">https://www.who.int/teams</a>	Team name, Team assigned for, Number of members in the team
Structure	Every company or organization will have a structure of their own and it is designed to represent a particular way of operating in the organization.	Information has been taken from <a href="https://www.who.int/about/structure">https://www.who.int/about/structure</a>	Employee name, Designation level, Level number
Contact us	This is very important since this is the way people will have the way to connect to the WHO team by different means.	Information has been taken from <a href="https://www.who.int/about/contact-us">https://www.who.int/about/contact-us</a>	Email id, Phone number, Address
Feature stories	Every feature story published by WHO team has a purpose, a story to be known to the world, Each office in each	Information has been taken from <a href="https://www.who.int/news-">https://www.who.int/news-</a>	Region, Publishing office, Date published, Topic



	country in each region has it's own stories to be told.	<a href="#">room/feature-stories</a>	
Careers	All the job postings and offerings would be found in this website.	Information has been taken from <a href="https://www.who.int/careers">https://www.who.int/careers</a>	Job posting id, Job location, Salary, Date posted, Application last date, No of applicants
Partners	All the collaborations details with other organizations, individuals or teams for the betterment of society are stored in this table.	Information has been taken from <a href="https://www.who.int/about/collaboration/partnerships">https://www.who.int/about/collaboration/partnerships</a>	Collaborated with, Region, Country, Team name
Financial statements	All the money flow in the organization, where the funding is being used, where it is being distributed are a very important factor of an organization.	Information has been taken from <a href="https://www.who.int/about/accountability/financial-statements">https://www.who.int/about/accountability/financial-statements</a>	Statement type, Date released, Statement name

### INITIAL ENTITY RELATIONSHIP DIAGRAM:



The issues I found in this ERD with solution:

- The first design issue I found in this Extended ERD is that in real life world, the Executive Board doesn't hire candidates to work for them, team does. So, careers should be related to teams.
- There can be more than one vaccine for covid – 19 disease. So the cardinality between COVID – 19 disease entity and vaccine entity should be (1 : N) and not (1 : 1).
- There is a cardinality issue between Regions and Campaigns. The cardinality present now is (N : N) because one region can have N campaigns but one campaign can't be at N regions. So the cardinality should be (1 : N).
- The same issue is present with Regions entity and Events entity. One region can have many events but one event cannot be present at many places at the same time. So the cardinality between them should be (1 : N).
- I have found another design trap in the Extended ERD. The correct design is that Teams should operate Operations and then Operations should manage Research and Training.

[illegible]