

## **DELIVERABLE 2: PROBLEM ANALYSIS - DATA SCIENTIST**

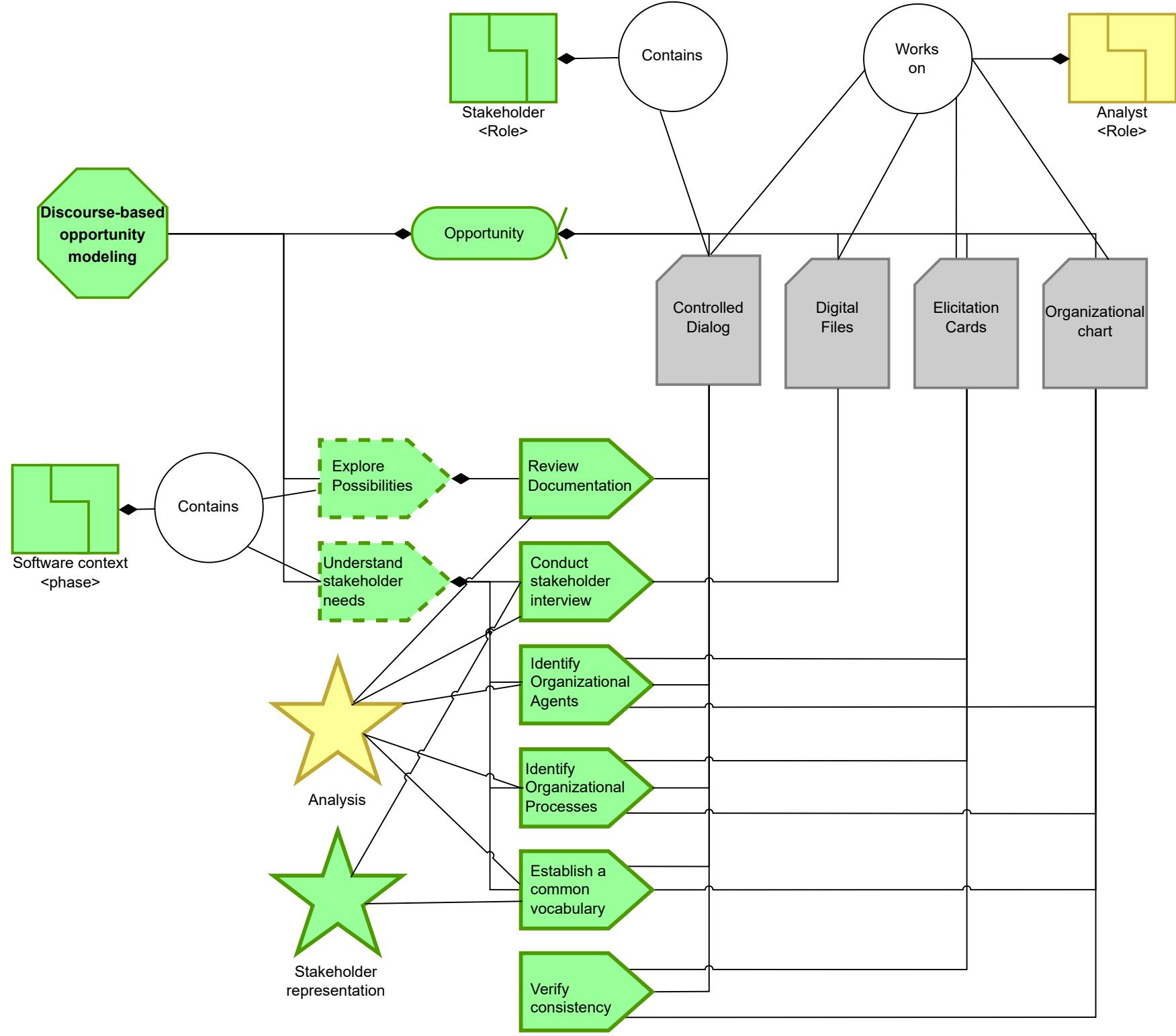
**DIEGO VALENTÍN OSORIO MARÍN**

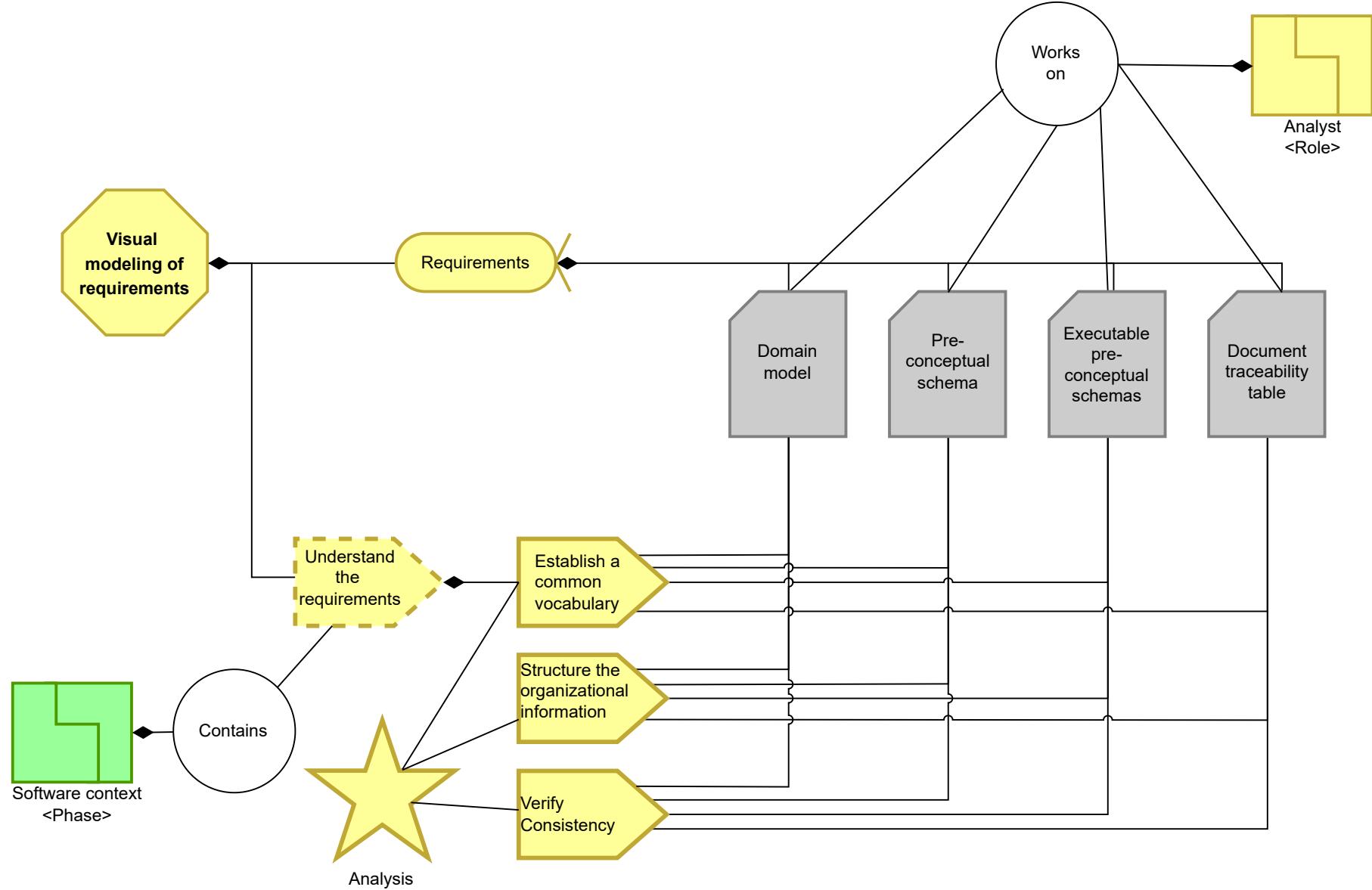
**JAIME ANDRÉS MONSALVE BALLESTEROS**

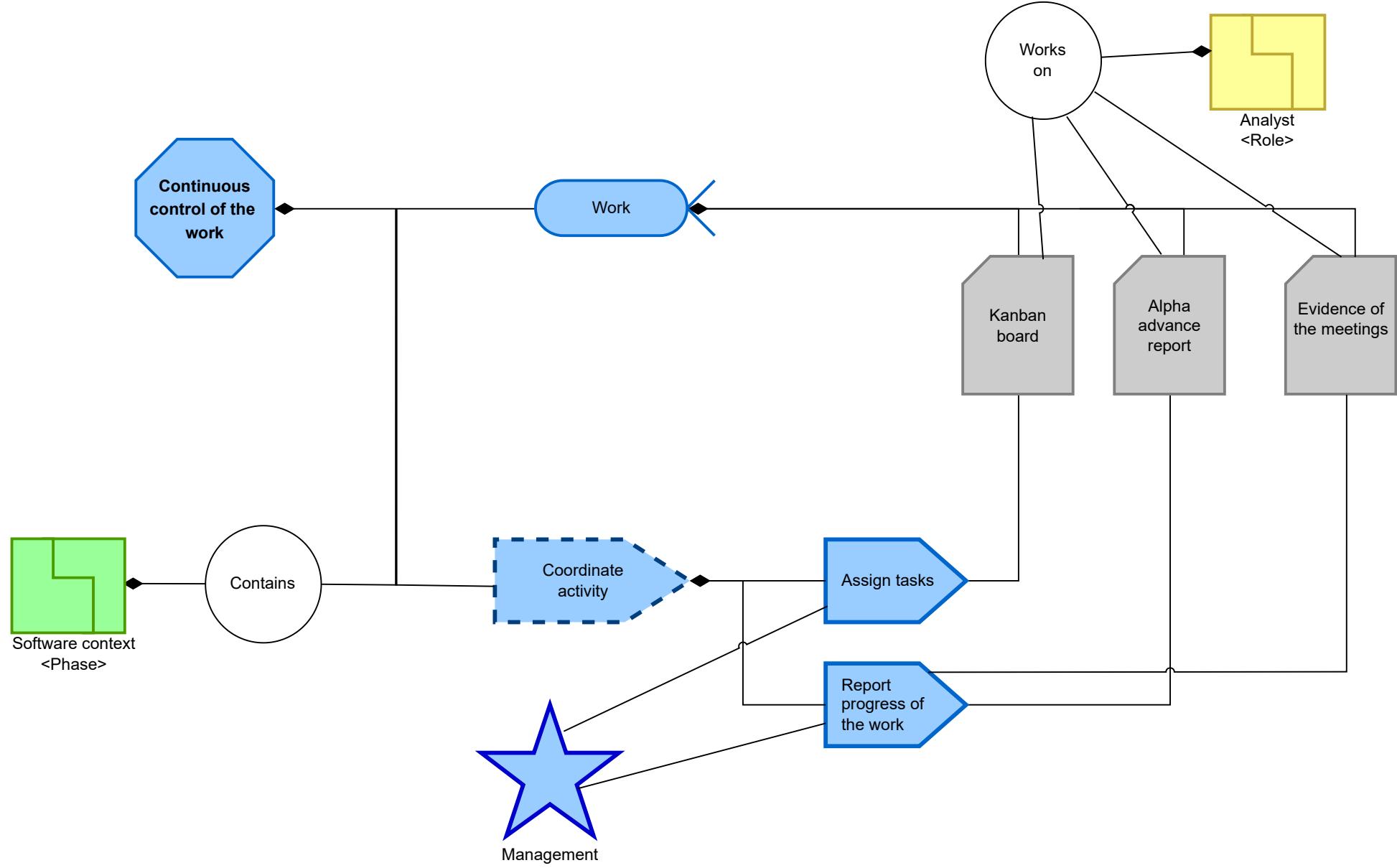
**SANTIAGO CASTRO TABARES**

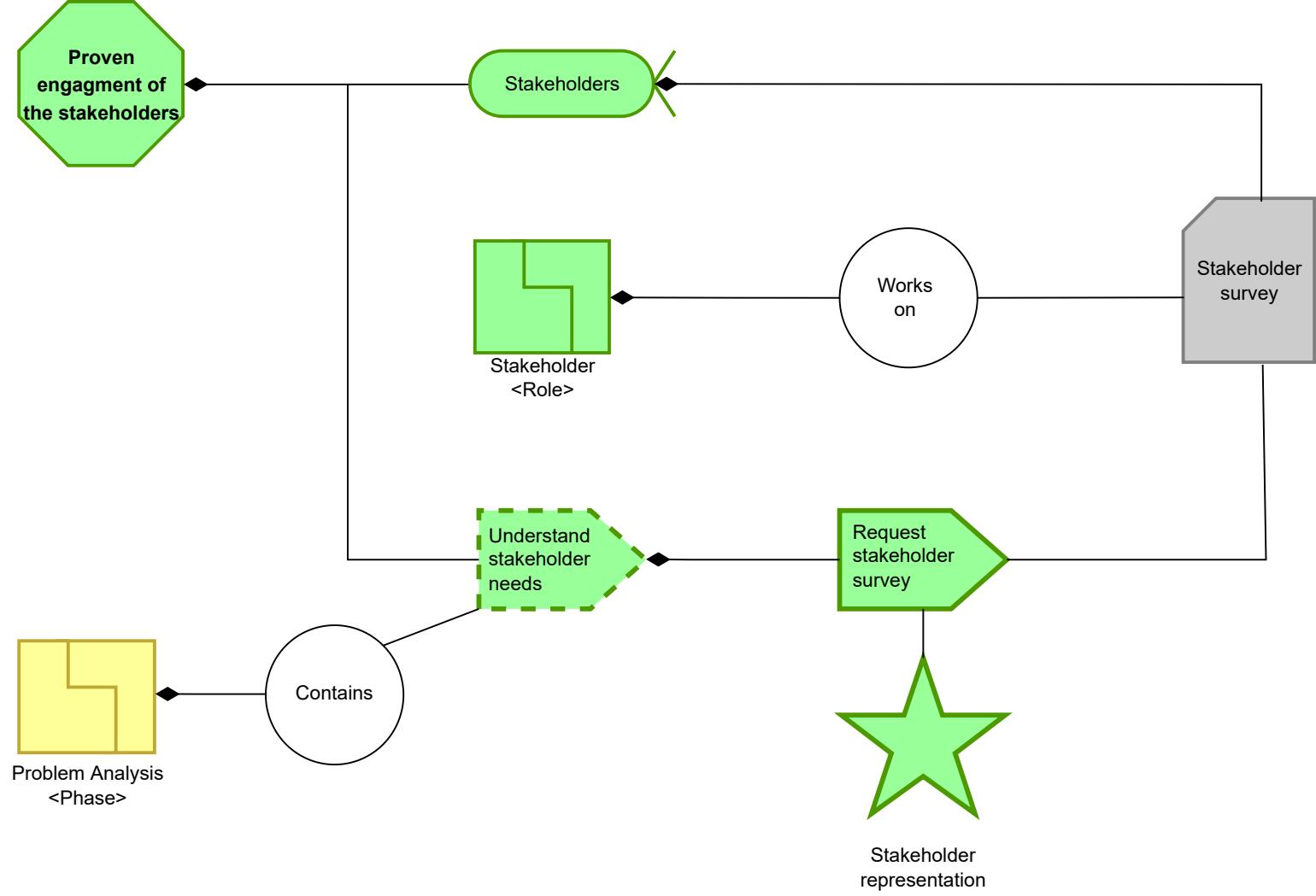
**FREDY ALBERTO OROZCO LOAIZA**

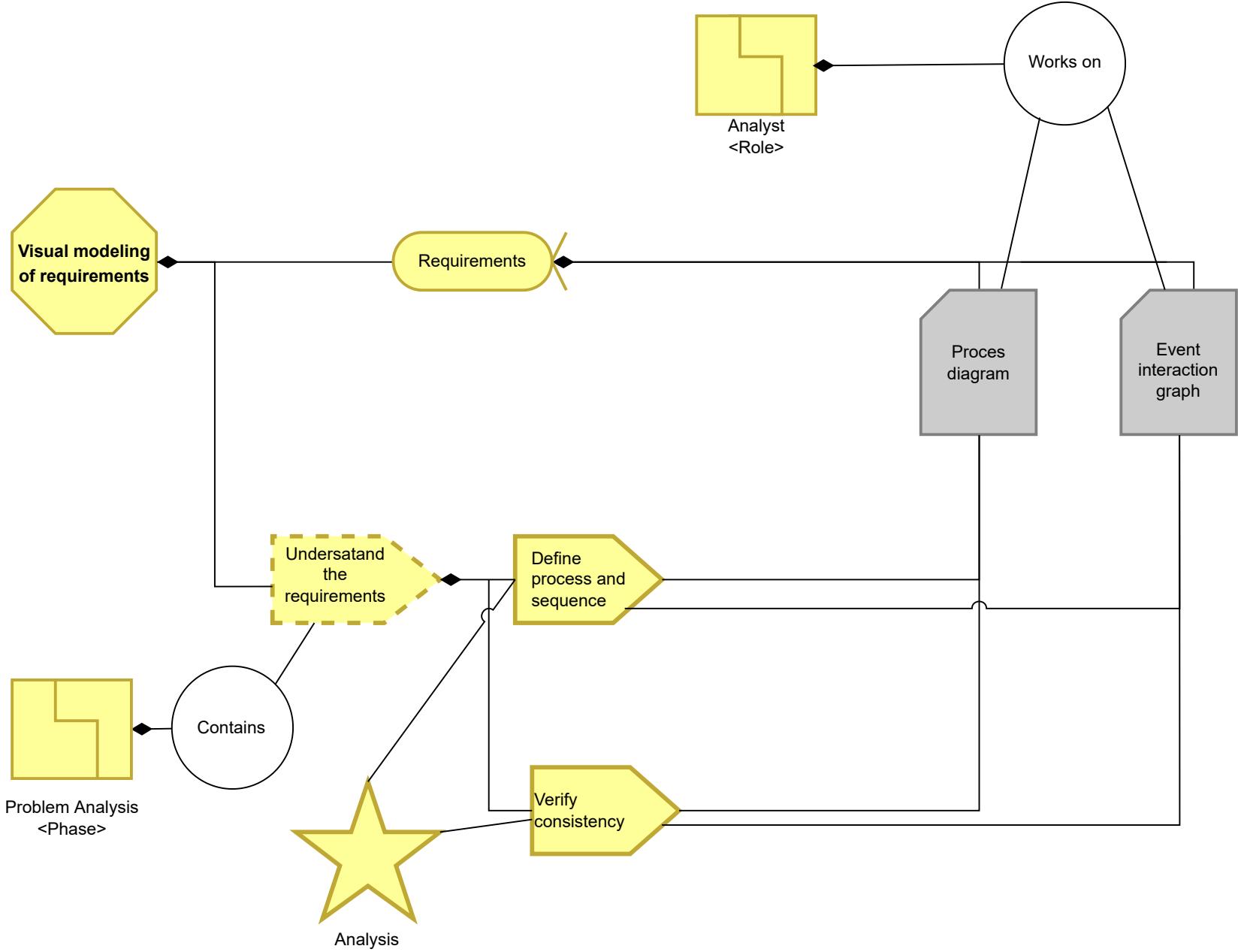
Computing and Decision Sciences Department. Faculty of Mines. Universidad Nacional de  
Colombia.

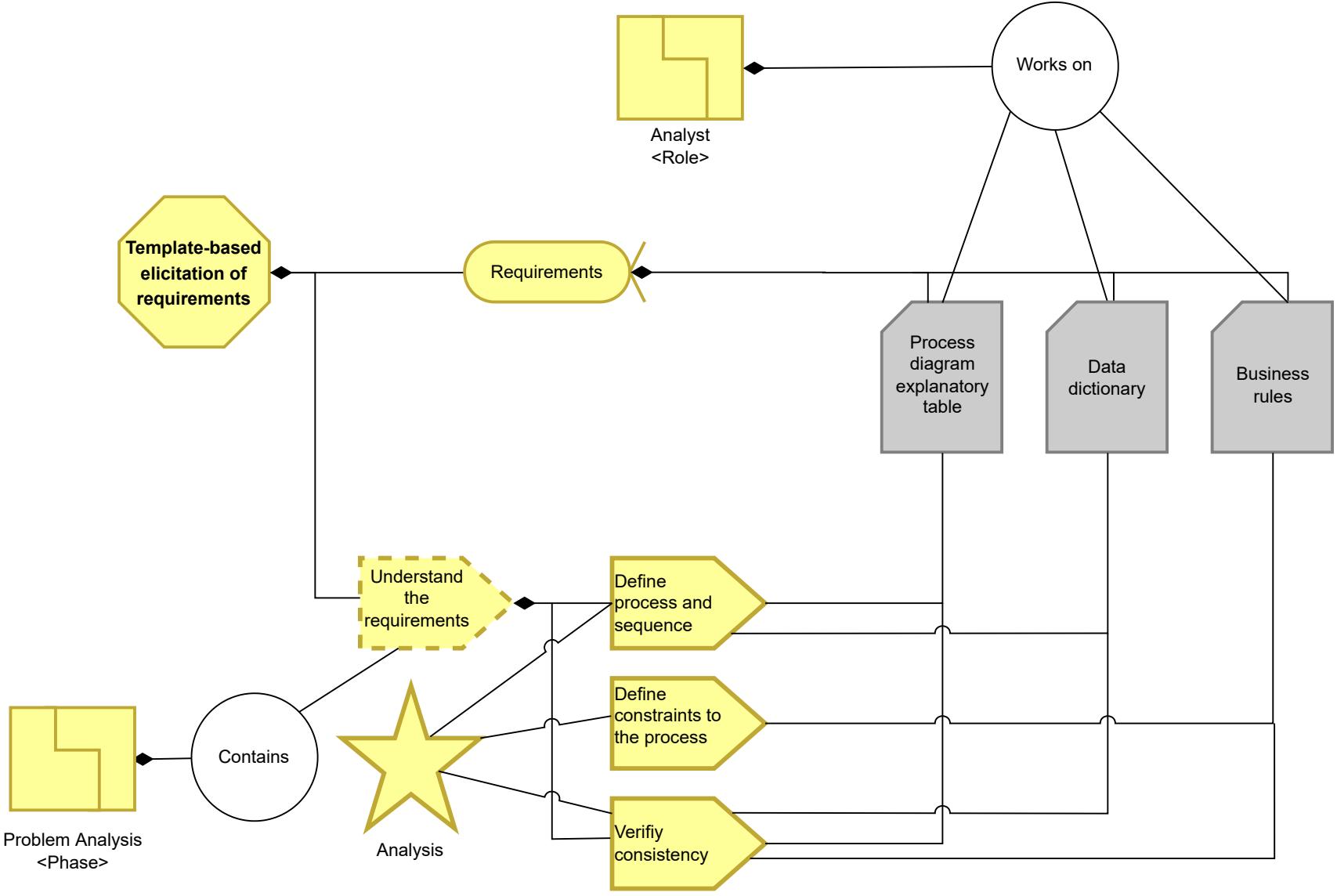


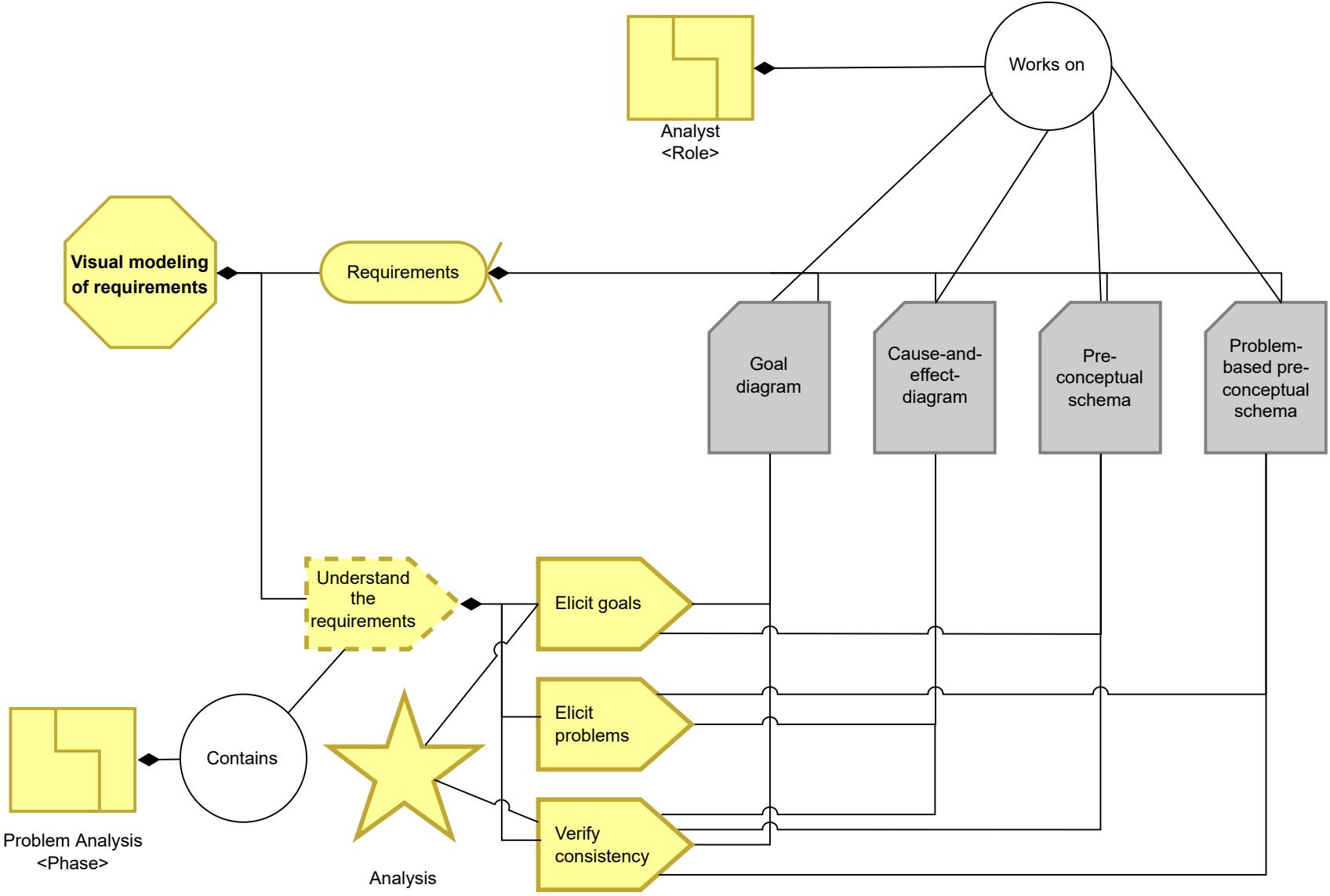


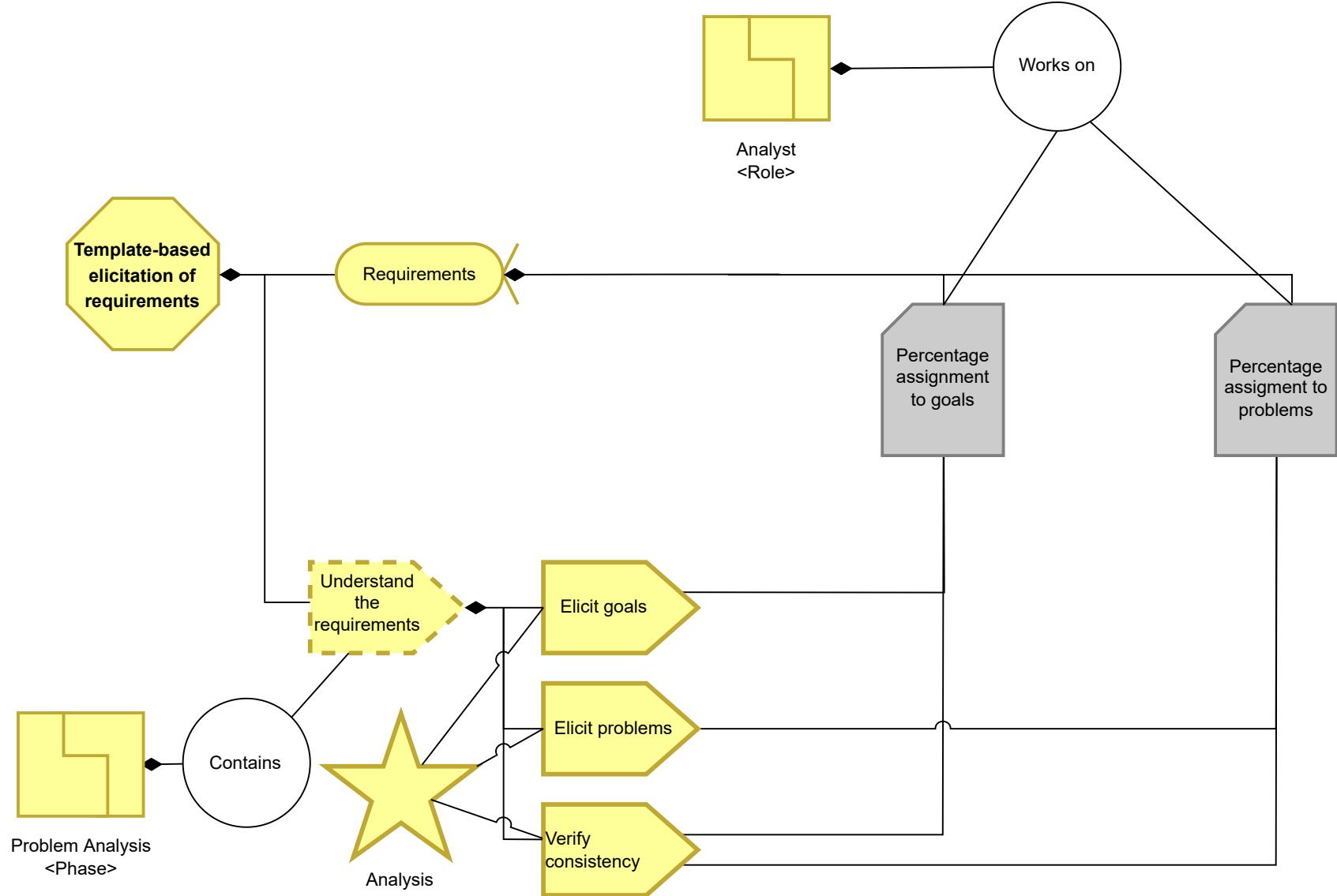


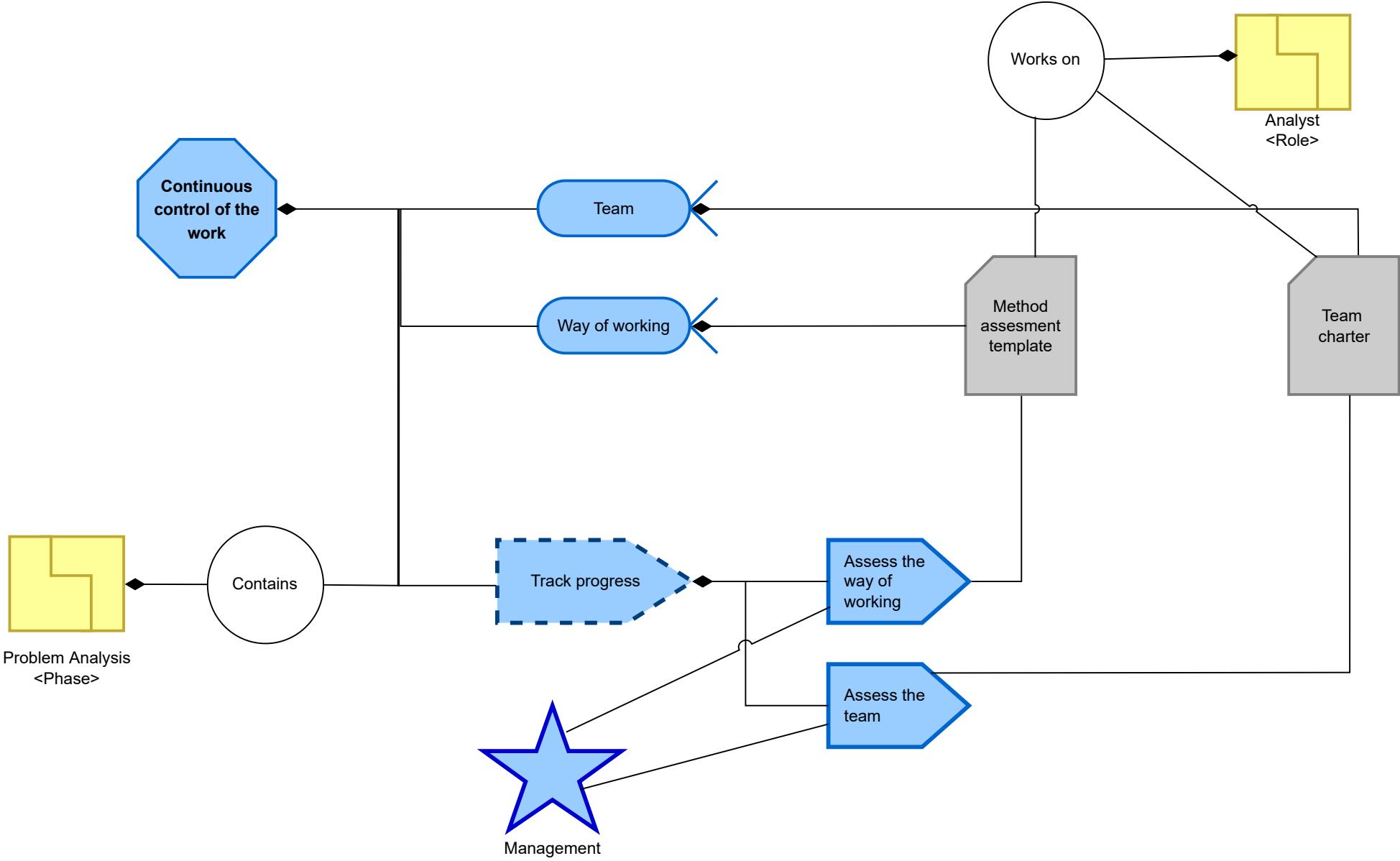


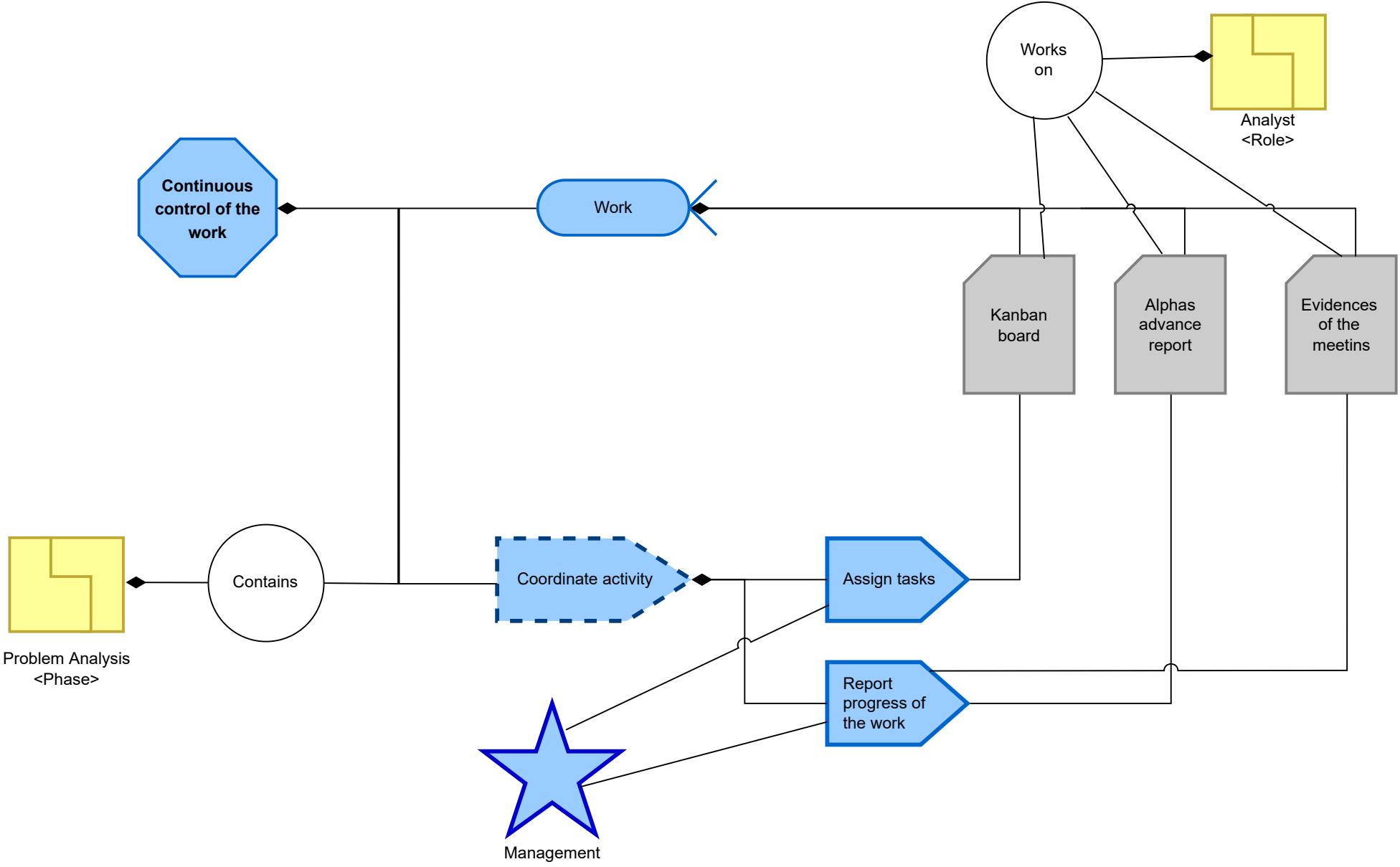












# CONTROLLED DIALOGUE

 BACK

**Analyst:** Good morning. With this interview, we aim to clarify the information concerning the problem domain in which we will work. Please answer the questions in the clearest way possible.

**Stakeholder:** O.k. Let's start.

**Analyst:** What is your role within the organization?

**Stakeholder:** I play the role of Data Scientist.

**Analyst:** Please, list the internal/external actors linked to the activities of your organization.

**Stakeholder:** We have users, employees, clients, data scientists, administrators and stakeholders.

**Analyst:** Who can play the role of a user?

**Stakeholder:** An employee, a stakeholder and an administrator.

**Analyst:** Who can play the role of an employee?

**Stakeholder:** A data scientist.

**Analyst:** Would you please mention some characteristics of users?

**Stakeholder:** They have name, e-mail, password, role and unique identification.

**Analyst:** Would you please mention some characteristics of employee?

**Stakeholder:** They have job title and experience.

**Analyst:** Would you please mention some characteristics of data scientists?

**Stakeholder:** They have background and style.

**Analyst:** Please list the main functions of a stakeholder.

**Stakeholder:** He/she provides data.

**Analyst:** Please list the main functions of the administrator.

**Stakeholder:** The administrator creates, edits and deletes users.

**Analyst:** Please list the main functions of the data scientists.

**Stakeholder:** The data scientist; evaluates user requirements, cleans data, extracts summary, determines model, builds applications, updates model, releases result and evaluates result.

**Analyst:** Are these functions gathered in some sort of responsibility?

**Stakeholder:** Yes, creates, edits and deletes users are gathered as “manage users”; cleans data and extracts summary are gathered as “pre-process data”; determines model, builds application and updates model are gathered as “process data”; releases and evaluates result are gathered as “post-process data”.

**Analyst:** Would you please mention some features of the data?

**Stakeholder:** Each data has version, file, type, classification, creation date, update date, description, size, format, quality, summary, application and unique name.

**Analyst:** Would you please mention some features of the model?

**Stakeholder:** It has type, application and unique name.

**Analyst:** Would you please mention some features of the application?

**Stakeholder:** It has version, creation date, tool, version control system, update date and solution.

**Analyst:** Would you please mention some features of the project?

**Stakeholder:** It has status, start date, end date, pre-processing tool, solution validation, users, user requirements, data and result.

**Analyst:** Would you please mention some features of the result?

**Stakeholder:** It has update date, file, creation date, version, type and feedback.

**Analyst:** Which of the mentioned features have features themselves?

**Stakeholder:** Summary has validity, creation date, description and file; Solution has update date, file, description, version, creation date and performance; user requirement has a user story, description, status and a unique name; feedback has comment and date.

**Analyst:** Which values or instances can be associated to which feature?

**Stakeholder:** True and False are instances of validity; insight providers, modeling specialist, platform builder, polymath and team leader are instances of style; low and high are instances of quality; structured data and unstructured data are instances of type of data; attributes, database, images, video footage, audio and handwritten note are instances of format of data; finished and in progress are instances of project status; administrator, data scientist and stakeholder are instances of user role; git, dvc, codecommit and sourcetree are instances of version control system; deterministic and random are instances of data classification.

**Analyst:** What does the administrator need in order to accomplish the managing of users?

**Stakeholder:** He/she only needs user.role = administrator.

**Analyst:** What does the administrator need in order to accomplish the creation of an user?

**Stakeholder:** He/she only needs user.role = administrator and the arising of an user.

**Analyst:** What does the data scientist need in order to accomplish the evaluation of user requirement?

**Stakeholder:** He/she only needs the arising of user requirements.

**Analyst:** What does the stakeholder need in order to accomplish the providing of data?

**Stakeholder:** He/she only needs: user.role = stakeholder and, the arising of an stakeholder or summary validity = false.

**Analyst:** What does the data scientist need in order to accomplish the pre-processing of the data?

**Stakeholder:** He/she only needs: user.role = data scientist and project status = in progress

**Analyst:** What does the data scientist need in order to accomplish the cleaning of data?

**Stakeholder:** He/she only needs: user.role = data scientist, project status = in progress, the data be provided before by the stakeholder and data quality = low.

**Analyst:** What does the data scientist need in order to accomplish the extraction of the summary?

**Stakeholder:** He/she only needs: user.role = data scientist, project status = in progress, the data be provided before by the stakeholder and, data quality = high or the data be cleaned before.

**Analyst:** What does the data scientist need in order to accomplish the processing of the data?

**Stakeholder:** He/she only needs: user.role = data scientist.

**Analyst:** What does the data scientist need in order to accomplish the determination of a model?

**Stakeholder:** He/she only needs: user.role = data scientist and summary validity = true.

**Analyst:** What does the data scientist need in order to accomplish the building of an application?

**Stakeholder:** He/she only needs: user.role = data scientist and the model be determined before.

**Analyst:** What does the data scientist need in order to accomplish the updating of a model?

**Stakeholder:** He/she only needs: user.role = data scientist, the model be determined before and solution performance < project solution validation.

**Analyst:** What does the data scientist need in order to accomplish the post-processing of the data?

**Stakeholder:** He/she only needs: user.role = data scientist.

**Analyst:** What does the data scientist need in order to accomplish the release of a result?

**Stakeholder:** He/she only needs: user.role = data scientist, and solution performance  $\geq$  project solution validation.

**Analyst:** What does the data scientist need in order to accomplish the evaluation of a result?

**Stakeholder:** He/she only needs: user.role = data scientist, and the appearance of feedback.

**Analyst:** Would you please establish some sort of sequence in the functions and responsibilities you have just described?

**Stakeholder:** First the data scientist evaluates the user requirements; then he waits for the stakeholder to show up, and then the stakeholder provides the data, then, in the pre-processing of the data, the data scientist cleans the data and then extracts a summary from the data; In the processing of the data the data scientist determines a model, then he can build and application or update the model; Last, in the post-processing of the data the data scientist releases the result and waits for feedback to be evaluated.

**Analyst:** What are the goals and problems associated with the function “extracts summary”?

**Stakeholder:** The goals are “Increasing the quality of data” and “Fostering data has summary”. The only problem is “Data scientist evaluates user requirements hardly”.

**Analyst:** What are the goals and problems associated with the function “cleans data”?

**Stakeholder:** The only goal is “Increasing the quality of data”. The only problems is “Stakeholder provides data wrongly”.

**Analyst:** What are the goals and problems associated with the function “provides data”?

**Stakeholder:** The only goal is “Ensuring project has data”. The only problems is “Stakeholder provides data wrongly”.

**Analyst:** What are the goals and problems associated with the function “determines model”?

**Stakeholder:** The only goal is “Obtaining the type of model”. The only problem is “data scientist does not have experience”.

**Analyst:** What are the goals and problems associated with the function “builds application”?

**Stakeholder:** The only goal is “Accomplishing application has solution”. The only problem is “Application does not have version control system”.

**Analyst:** What are the goals and problems associated with the function “updates model”?

**Stakeholder:** The goals are “Accomplishing application has solution” and “Enhancing the performance of the solution”. The only problem is “Application does not have version control system”.

**Analyst:** What are the goals and problems associated with the function “evaluates result”?

**Stakeholder:** The goals are “Conserving the end date of the project” and “Controlling the status of the project”. This function has no problems.

**Analyst:** What are the goals and problems associated with the function “releases result”?

**Stakeholder:** The only goal is “achieving the result”. This function has no problems.

**Analyst:** What are the goals and problems associated with the function “evaluates user requirements”?

**Stakeholder:** The only goal is “Obtaining the type of model”. The problems are “Data scientist does not have experience” and “Data scientist evaluates user requirements hardly”.

**Analyst:** What are the goals and problems associated with the function “creates user”?

**Stakeholder:** The only goal is “Maintaining the users”. This function has no problems.

**Analyst:** What are the goals and problems associated with the function “edits user”?

**Stakeholder:** The only goal is “Maintaining the users”. This function has no problems.

**Analyst:** What are the goals and problems associated with the function “deletes user”?

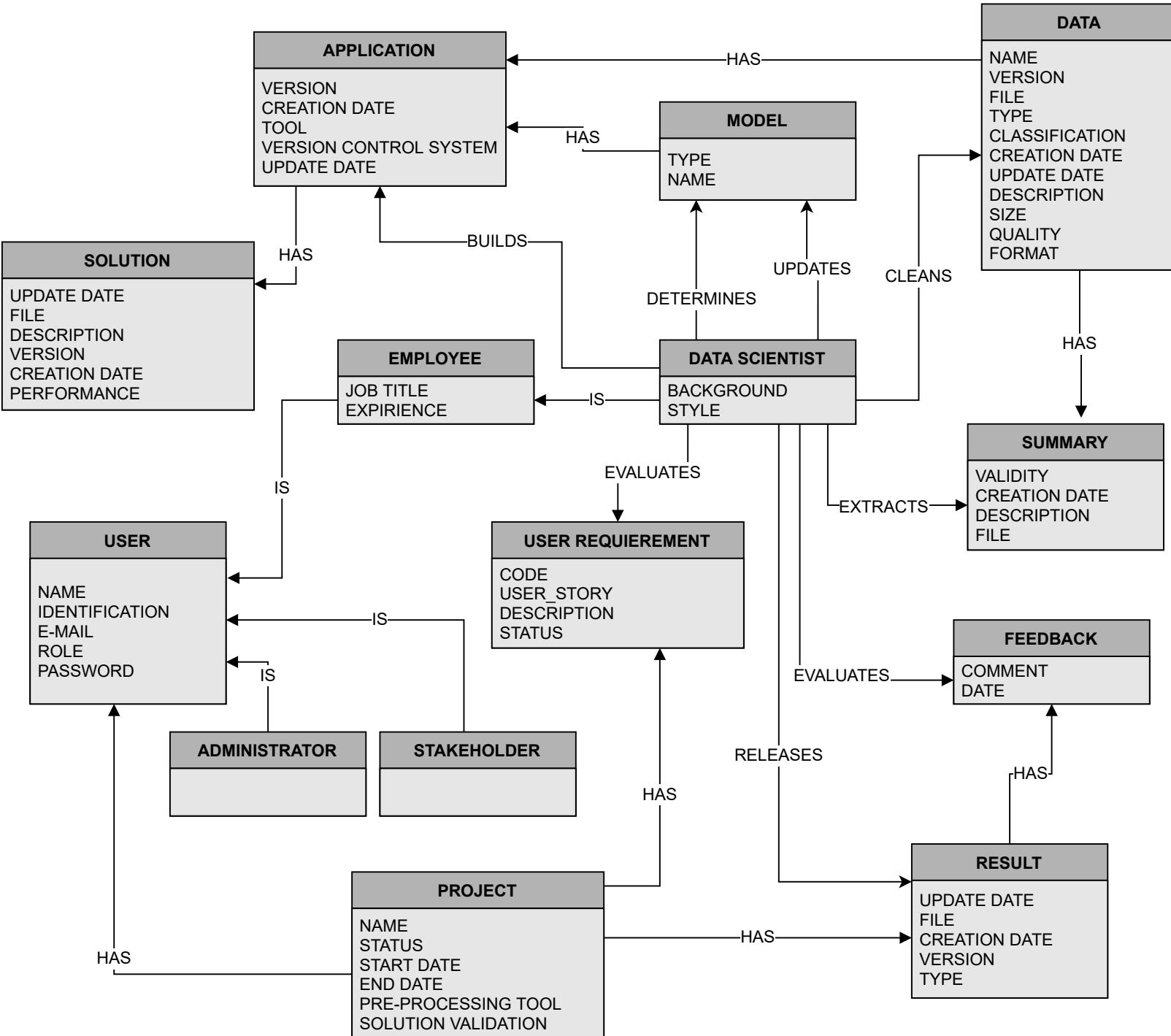
**Stakeholder:** The only goal is “Maintaining the users”. This function has no problems.

**Analyst:** Thank you for your valuable information. We will be in contact in order to clarify any doubts that may arise in this process.

**Stakeholder:** Thank you. I'll be in touch.

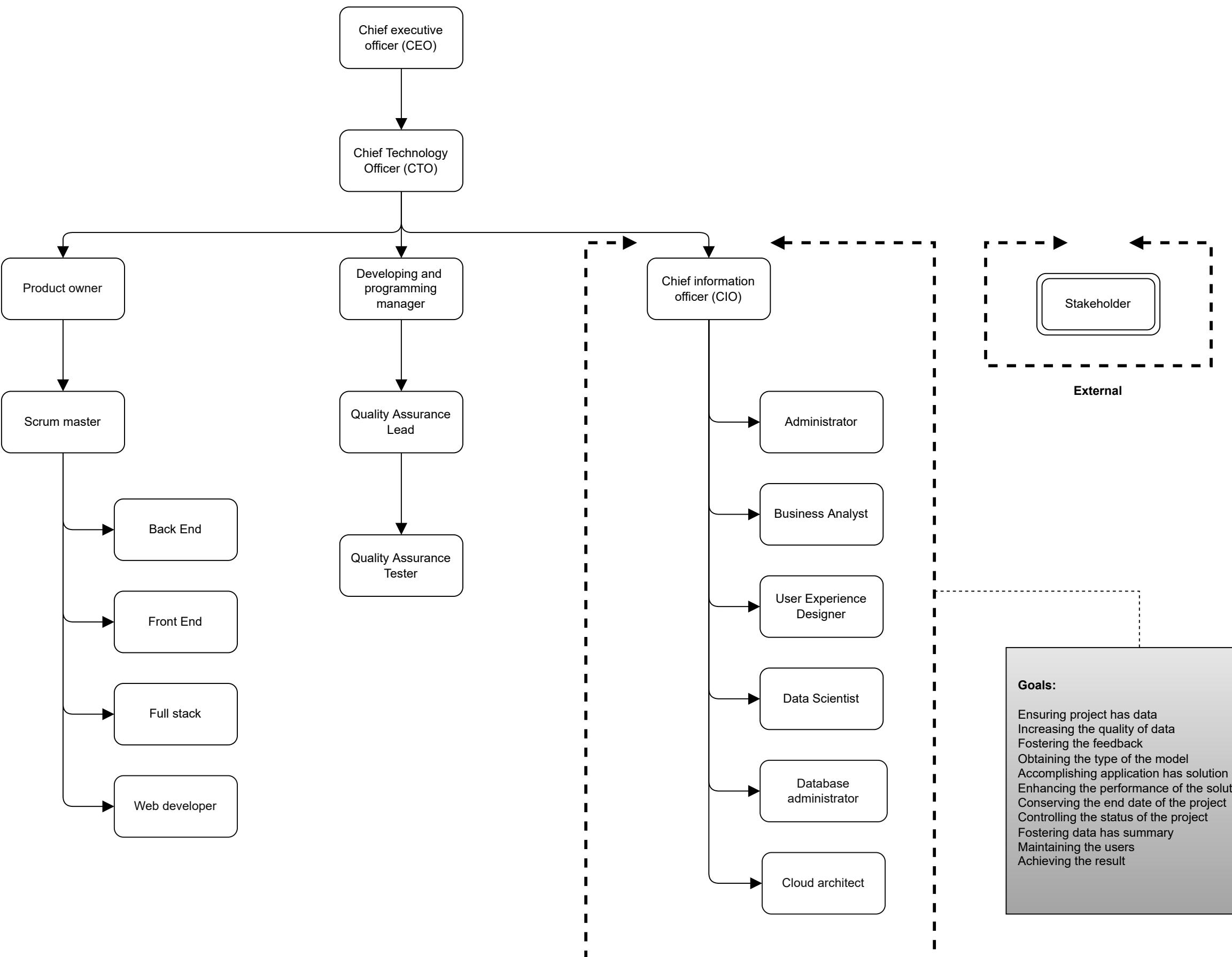
# DOMAIN MODEL

[BACK](#)



# ORGANIZATIONAL CHART

[BACK](#)



**Goals:**

- Ensuring project has data
- Increasing the quality of data
- Fostering the feedback
- Obtaining the type of the model
- Accomplishing application has solution
- Enhancing the performance of the solution
- Conserving the end date of the project
- Controlling the status of the project
- Fostering data has summary
- Maintaining the users
- Achieving the result

**Problem Area:**

Chief information officer (CIO), Business Analyst, User Experience Designer, Data Scientist, Database administrator, Cloud architect and Administrator

**Responsibilities:**

Manages users  
Pre-process data  
Process data  
Post-process data

# ELICITATION CARDS

◀ BACK

ACTOR	
USER	
FEATURES	NAME, E-MAIL, PASSWORD, ROLE, and IDENTIFICATION is unique
ANNOTATIONS	

ACTOR	
EMPLOYEE	
FEATURES	JOB TITLE, EXPERIENCE
ANNOTATIONS	

ACTOR	
ADMINISTRATOR	
FEATURES	ADMINISTRATOR can be USER
ANNOTATIONS	

ACTOR	
STAKEHOLDER	
FEATURES	STAKEHOLDER can be USER
ANNOTATIONS	

ACTOR	
DATA SCIENTIST	
FEATURES	DATA SCIENTIST can be EMPLOYEE
ANNOTATIONS	

	<b>OBJECT</b>	<b>ANNOTATIONS</b>
	<b>DATA</b>	
<b>FEATURES</b>	<b>NAME</b> is unique, <b>VERSION</b> , <b>FILE</b> , <b>TYPE</b> , <b>CLASSIFICATION</b> , <b>CREATION DATE</b> , <b>UPDATE DATE</b> , <b>DESCRIPTION</b> , <b>SIZE</b> , <b>FORMAT</b> , <b>QUALITY</b>	<b>DATA</b> is related to <b>SUMMARY</b> and <b>APPLICATION</b>
	<b>OBJECT</b>	<b>ANNOTATIONS</b>
	<b>SUMMARY</b>	
<b>FEATURES</b>	<b>VALIDITY</b> , <b>CREATION DATE</b> , <b>DESCRIPTION</b> , <b>FILE</b>	
	<b>OBJECT</b>	<b>ANNOTATIONS</b>
	<b>APPLICATION</b>	
<b>FEATURES</b>	<b>VERSION</b> , <b>CREATION DATE</b> , <b>TOOL</b> , <b>VERSION CONTROL SYSTEM</b> , <b>UPDATE DATE</b>	<b>APPLICATION</b> is related to <b>SOLUTION</b>
	<b>OBJECT</b>	<b>ANNOTATIONS</b>
	<b>SOLUTION</b>	
<b>FEATURES</b>	<b>UPDATE DATE</b> , <b>FILE</b> , <b>DESCRIPTION</b> , <b>VERSION</b> , <b>CREATION DATE</b> , <b>PERFORMANCE</b>	
	<b>OBJECT</b>	<b>ANNOTATIONS</b>
	<b>MODEL</b>	
<b>FEATURES</b>	<b>TYPE</b> , <b>NAME</b> is unique	<b>MODEL</b> is related to <b>APPLICATION</b>
	<b>OBJECT</b>	<b>ANNOTATIONS</b>
	<b>USER REQUIEREMENT</b>	
<b>FEATURES</b>	<b>CODE</b> is unique, <b>USER STORY</b> , <b>DESCRIPTION</b> , <b>STATUS</b>	
	<b>OBJECT</b>	<b>ANNOTATIONS</b>
	<b>PROJECT</b>	
<b>FEATURES</b>	<b>NAME</b> is unique, <b>STATUS</b> , <b>START DATE</b> , <b>END DATE</b> , <b>PRE-PROCESING TOOL</b> , <b>SOLUTION VALIDATION</b>	<b>PROJECT</b> is related to <b>USER</b> , <b>USER REQUIEREMENT</b> , <b>DATA</b> and <b>RESULT</b>
	<b>OBJECT</b>	<b>ANNOTATIONS</b>
	<b>RESULT</b>	
<b>FEATURES</b>	<b>UPDATE DATE</b> , <b>FILE</b> , <b>CREATION DATE</b> , <b>VERSION</b> , <b>TYPE</b>	<b>RESULT</b> is related to <b>FEEDBACK</b>
	<b>OBJECT</b>	<b>ANNOTATIONS</b>
	<b>FEEDBACK</b>	
<b>FEATURES</b>	<b>COMMENT</b> , <b>DATE</b>	

FUNCTION	GOAL	PROBLEM
<b>EXTRACTS</b> <b>ACTOR</b> DATA SCIENTIST <b>OBJECT</b> SUMMARY <b>CONSTRAINT</b>  USER.ROLE = DATA SCIENTIST and PROJECT.STATUS = IN PROGRESS and ((DATA.QUALITY = HIGH and STAKEHOLDER PROVIDES DATA) or DATA SCIENTIST CLEANS DATA)	<ul style="list-style-type: none"> <li>- <b>G10:</b> Increasing the quality of data</li> <li>- <b>G6:</b> Fostering data has summary</li> </ul>	<b>C1 :</b> DATA SCIENTIST evaluates USER REQUIREMENTS hardly

FUNCTION	GOAL	PROBLEM
<b>CLEANS</b> <b>ACTOR</b> DATA SCIENTIST <b>OBJECT</b> DATA <b>CONSTRAINT</b>  USER.ROLE = DATA SCIENTIST and PROJECT.STATUS = IN PROGRESS and (DATA.QUALITY = HIGH and STAKEHOLDER PROVIDES DATA)	<b>G10:</b> Increasing the quality of data	<b>C3 :</b> STAKEHOLDER provides DATA wrongly

FUNCTION	GOAL	PROBLEM
<b>PROVIDES</b> <b>ACTOR</b> STAKEHOLDER <b>OBJECT</b> DATA <b>CONSTRAINT</b>  (SUMMARY.VALIDITY = FALSE and USER.ROLE = STAKEHOLDER) or (STAKEHOLDER SHOWS UP and USER.ROLE = STAKEHOLDER)	<b>G11:</b> Ensuring project has data	<b>C3 :</b> STAKEHOLDER provides DATA wrongly

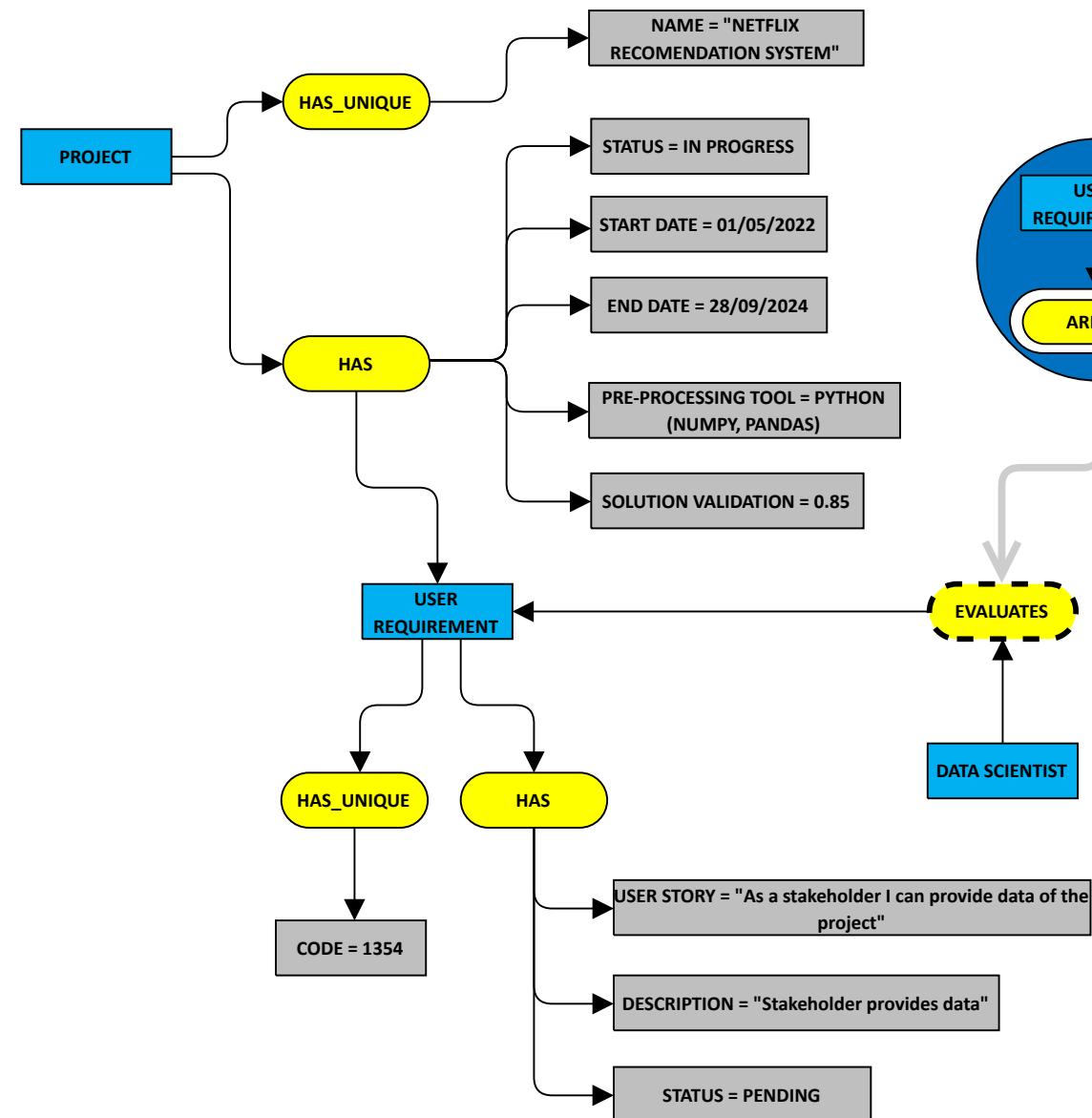
FUNCTION	GOAL	PROBLEM
<b>DETERMINES</b> <b>ACTOR</b> DATA SCIENTIST <b>OBJECT</b> MODEL <b>CONSTRAINT</b>  SUMMARY.VALIDITY = TRUE and USER.ROLE = DATA SCIENTIST	<b>G8 :</b> Obtaining the type of model.	<b>C5 :</b> DATA SCIENTIST does not have EXPERIENCE

FUNCTION	GOAL	PROBLEM
<b>BUILDS</b> <b>ACTOR</b> DATA SCIENTIST <b>OBJECT</b> APPLICATION <b>CONSTRAINT</b>  USER.ROLE = DATA SCIENTIST and DATA SCIENTIST DETERMINES MODEL	<b>G2 :</b> Accomplishing application has solution	<b>C4 :</b> APPLICATION does not have VERSION CONTROL SYSTEM

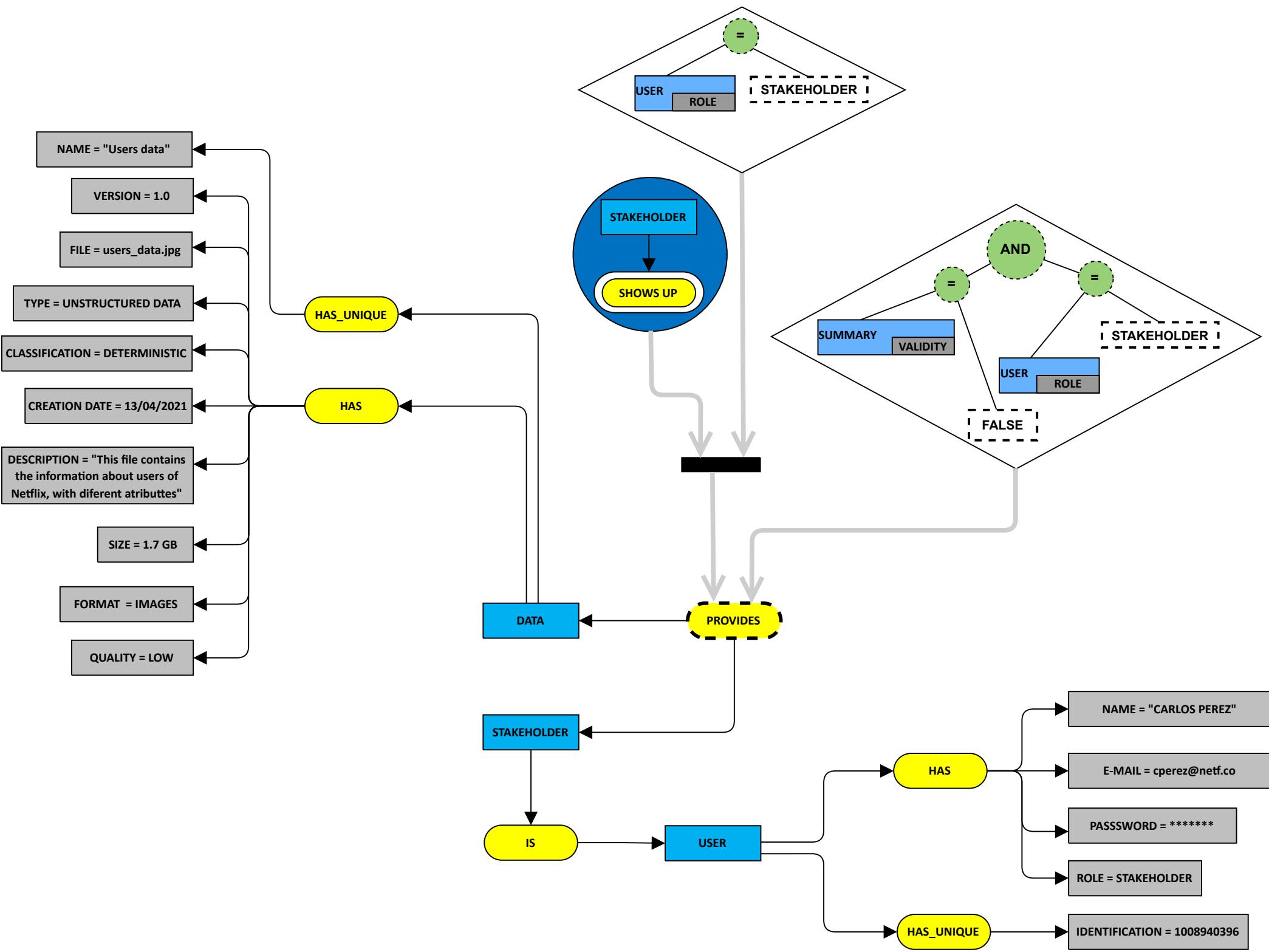
FUNCTION	GOAL	PROBLEM
<b>UPDATES</b> <b>ACTOR</b> DATA SCIENTIST <b>OBJECT</b> MODEL <b>CONSTRAINT</b>  USER.ROLE = DATA SCIENTIST and SOLUTION.PERFORMANCE < PROJECT.SOLUTION VALIDATION and DATA SCIENTIST DETERMINES MODEL	<ul style="list-style-type: none"> <li>- <b>G2 :</b> Accomplishing application has solution</li> <li>- <b>G3 :</b> Enhancing the performance of the solution</li> </ul>	<b>C4 :</b> APPLICATION does not have VERSION CONTROL SYSTEM

# EXECUTABLE PRE-CONCEPTUAL SCHEMAS

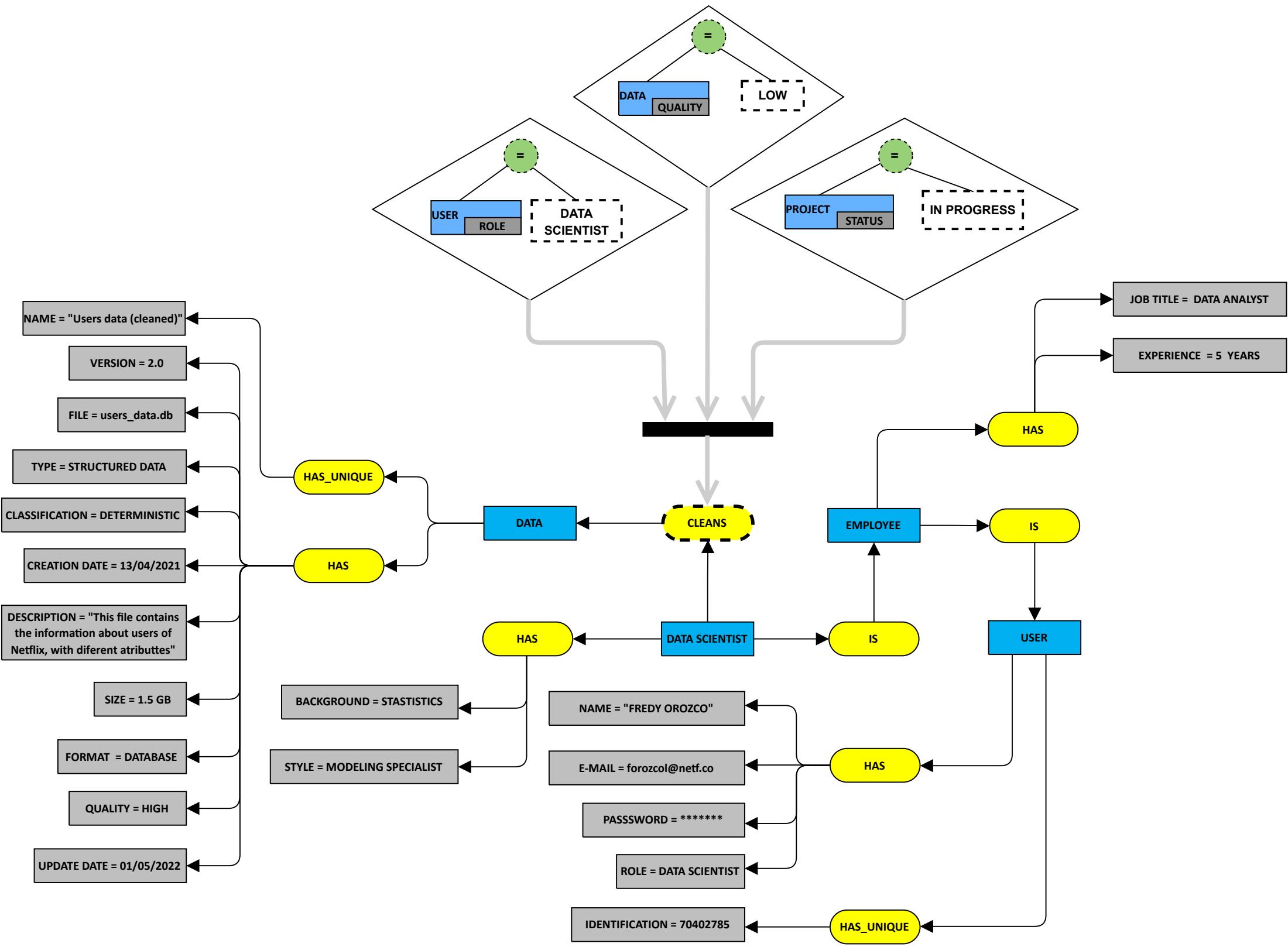
#1 < BACK



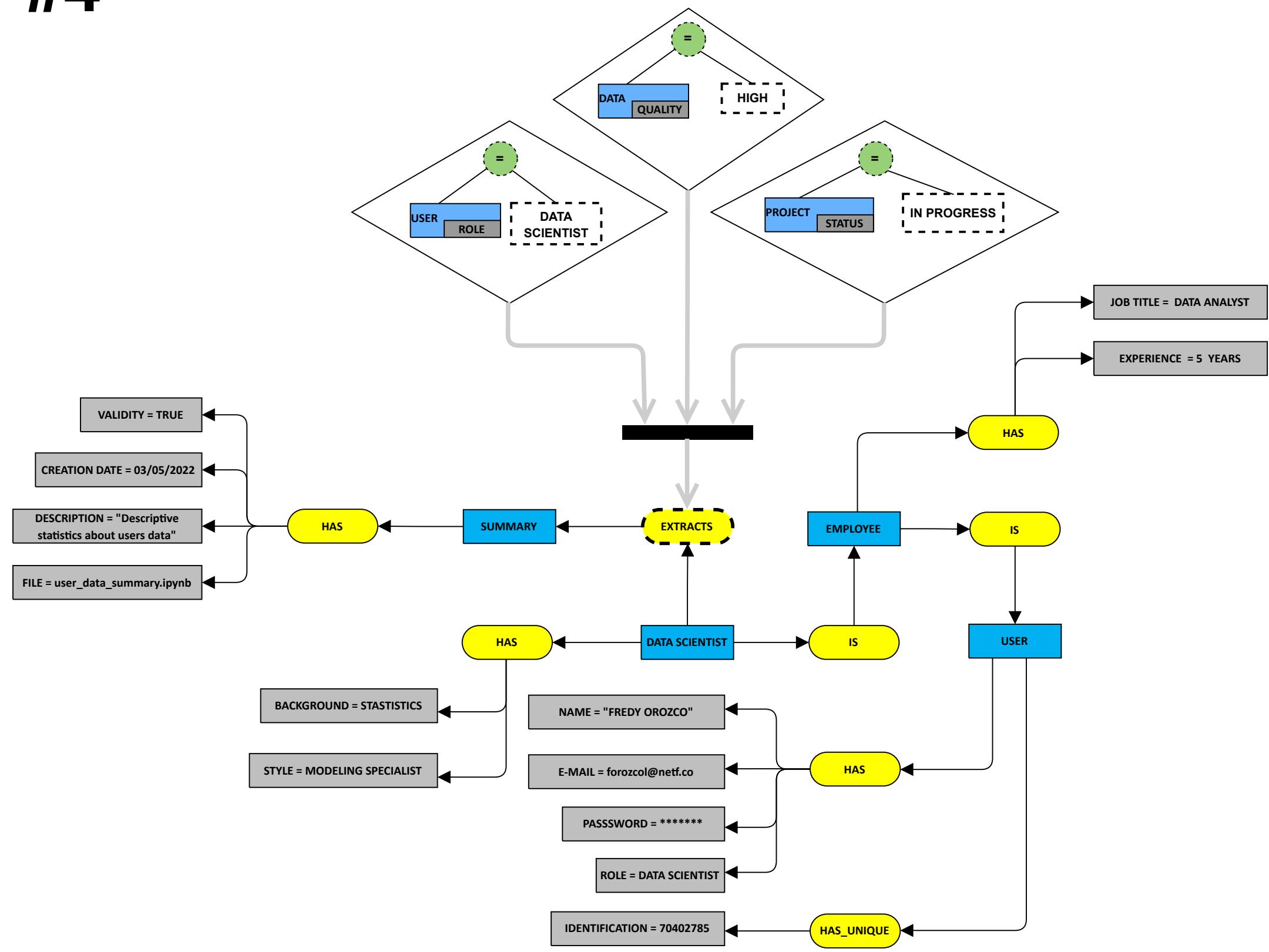
# #2



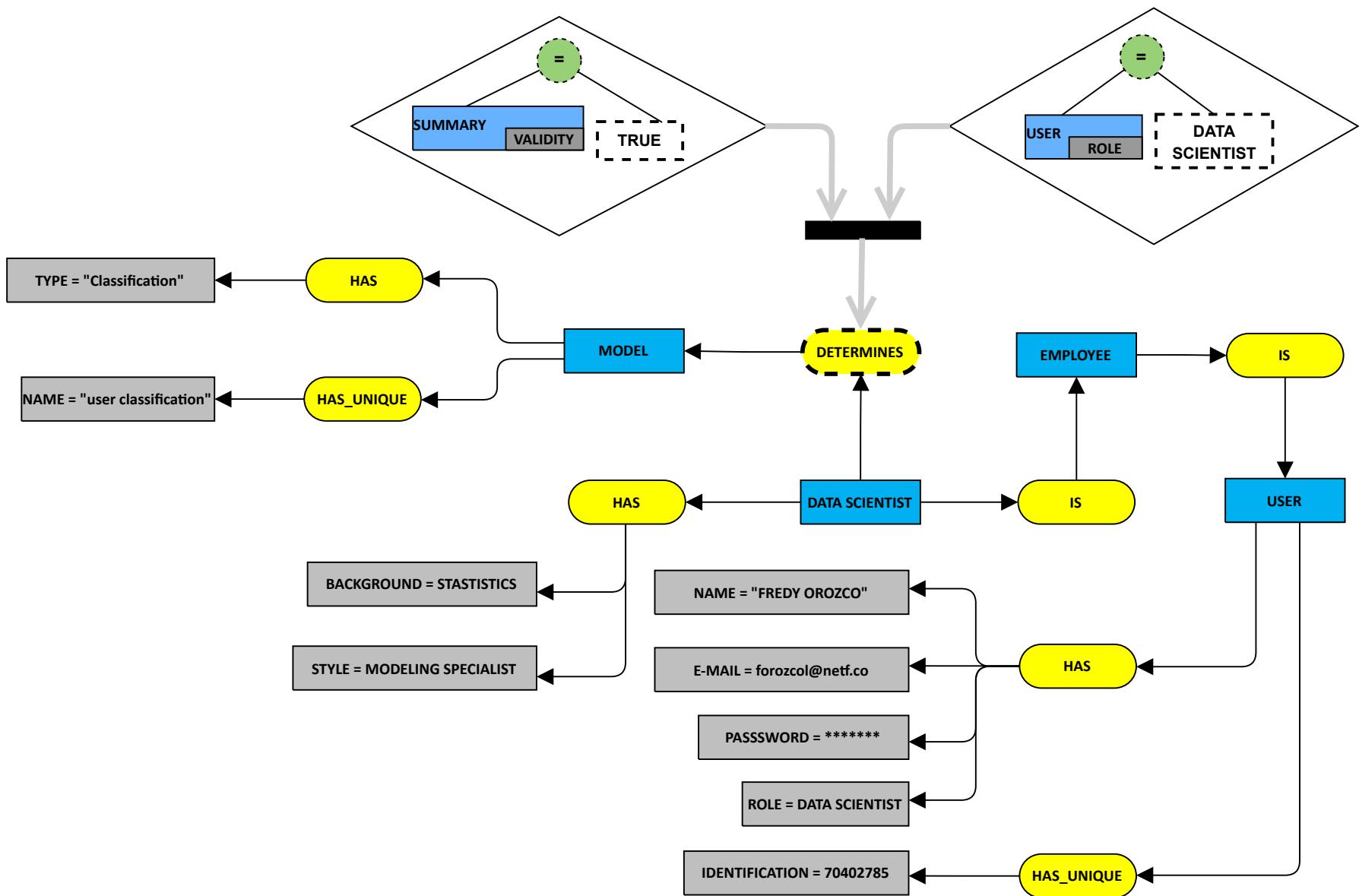
# #3



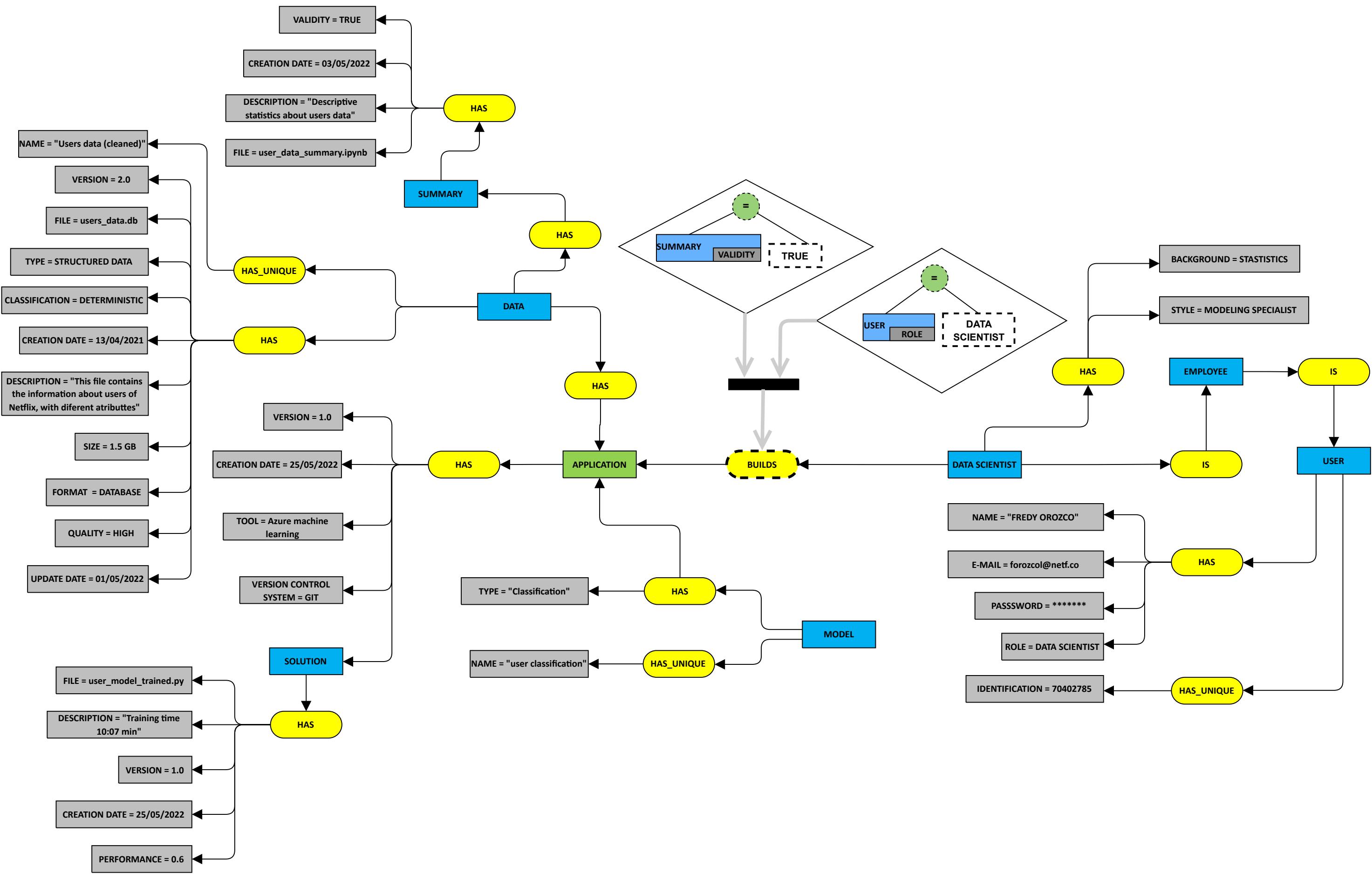
# #4



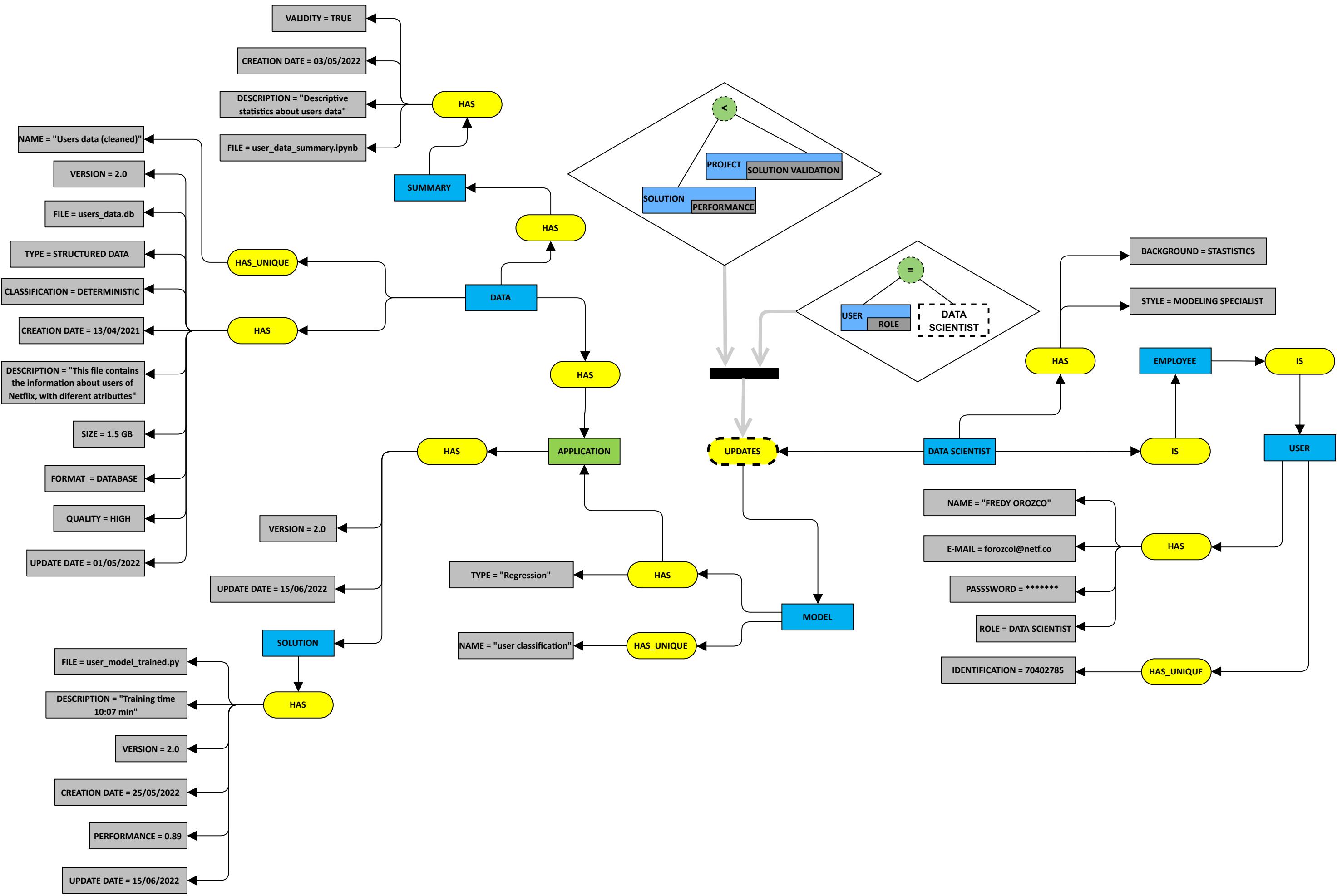
# #5



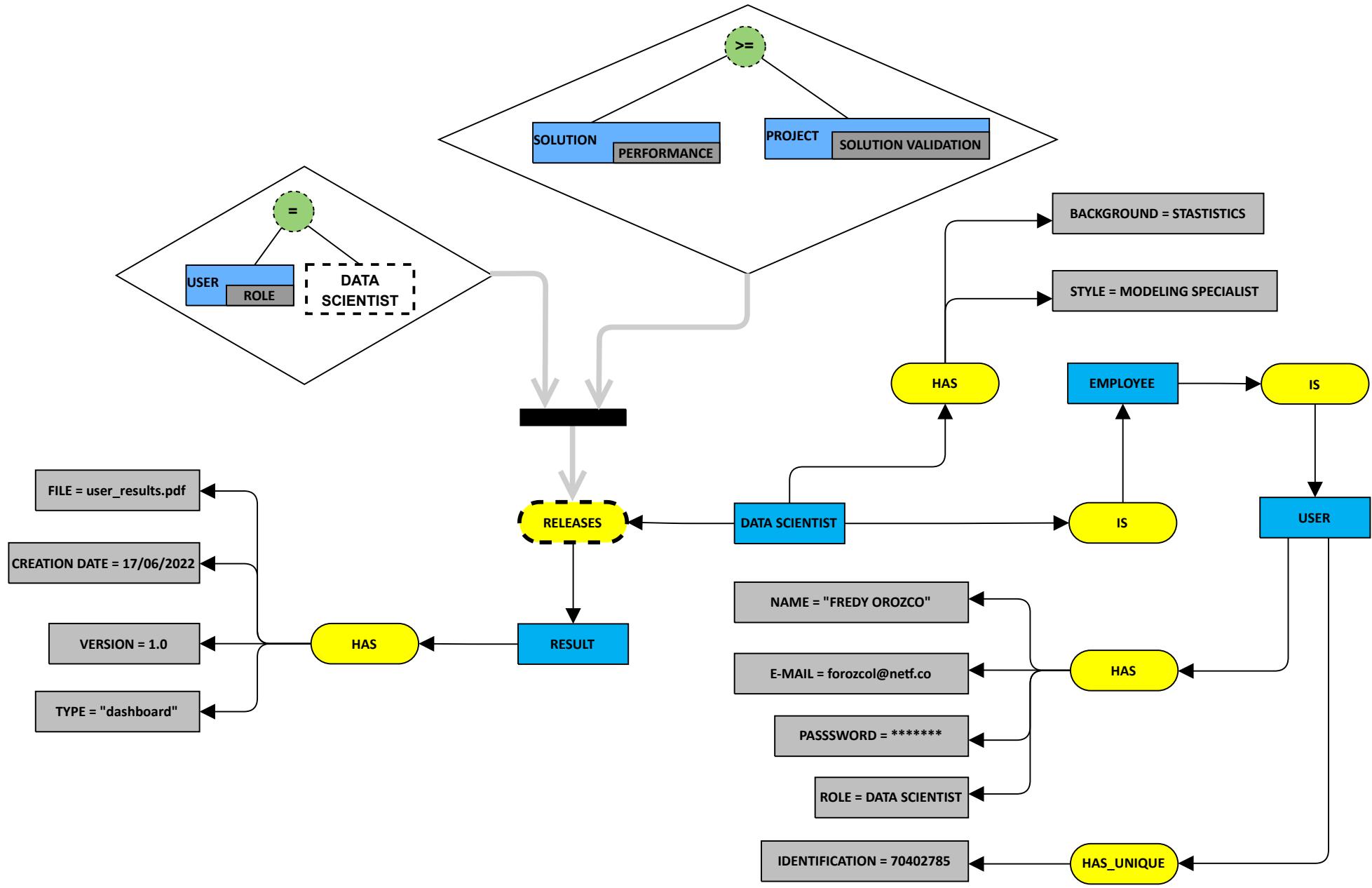
## #6



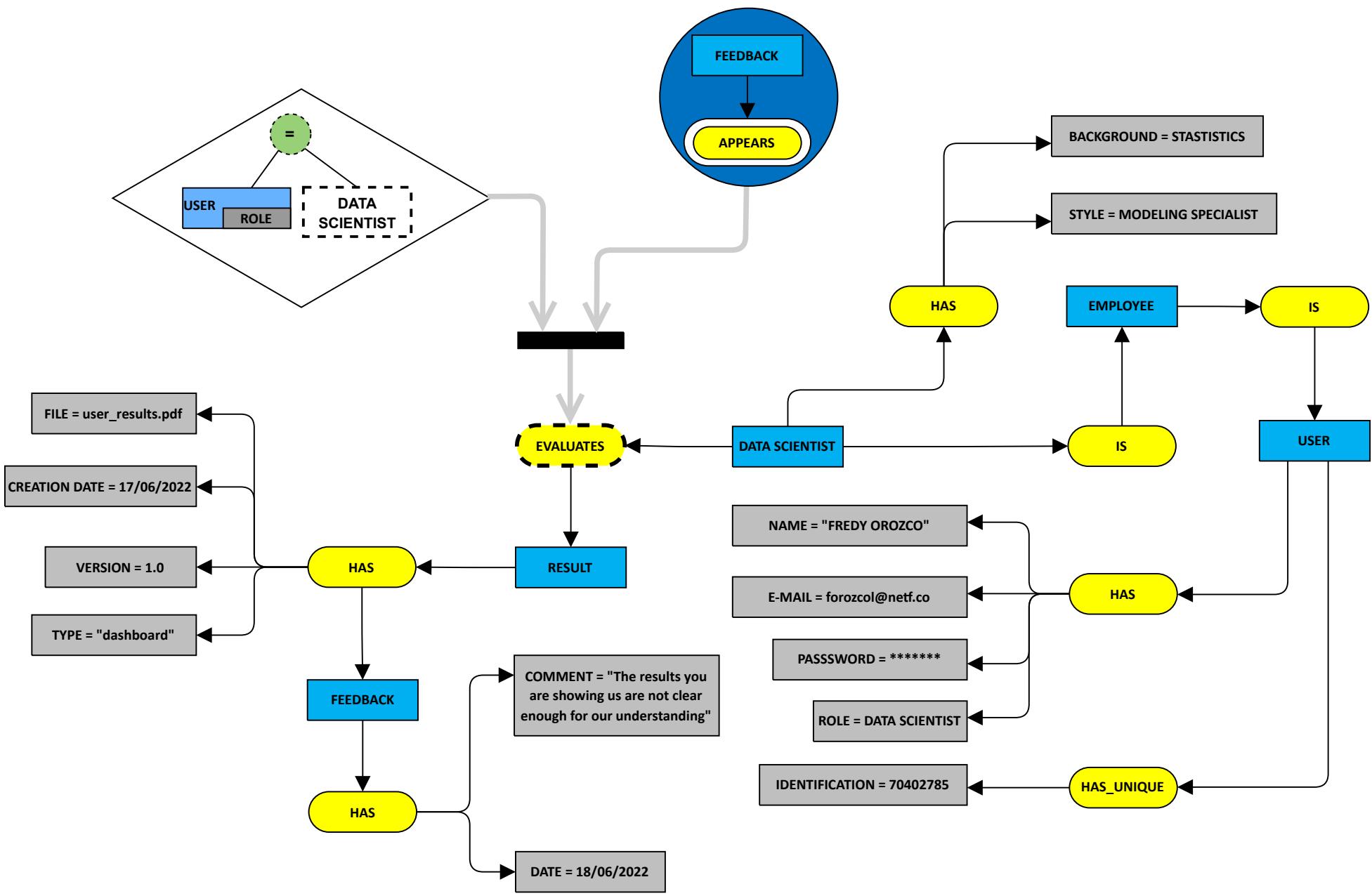
## #7



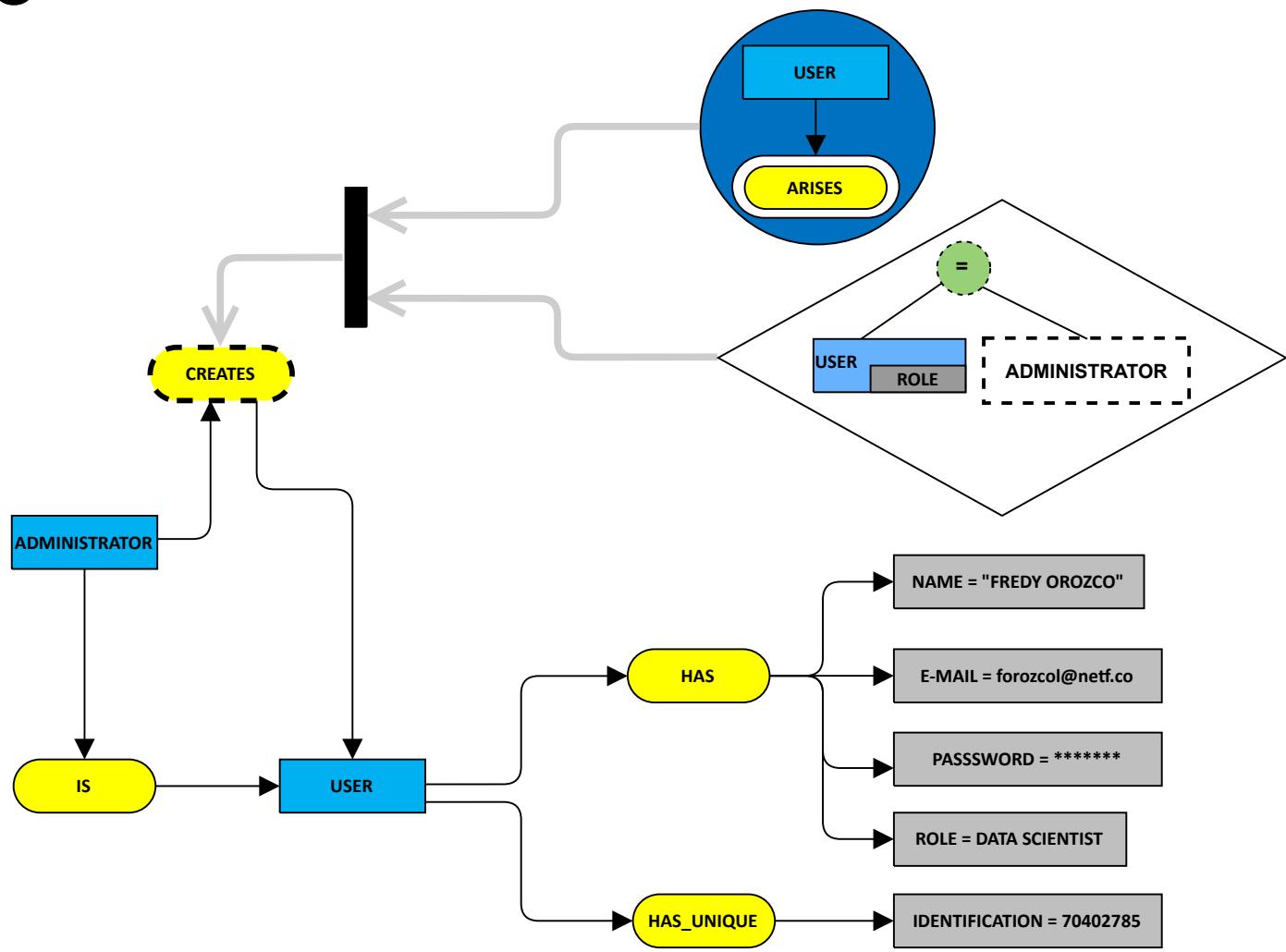
# #8



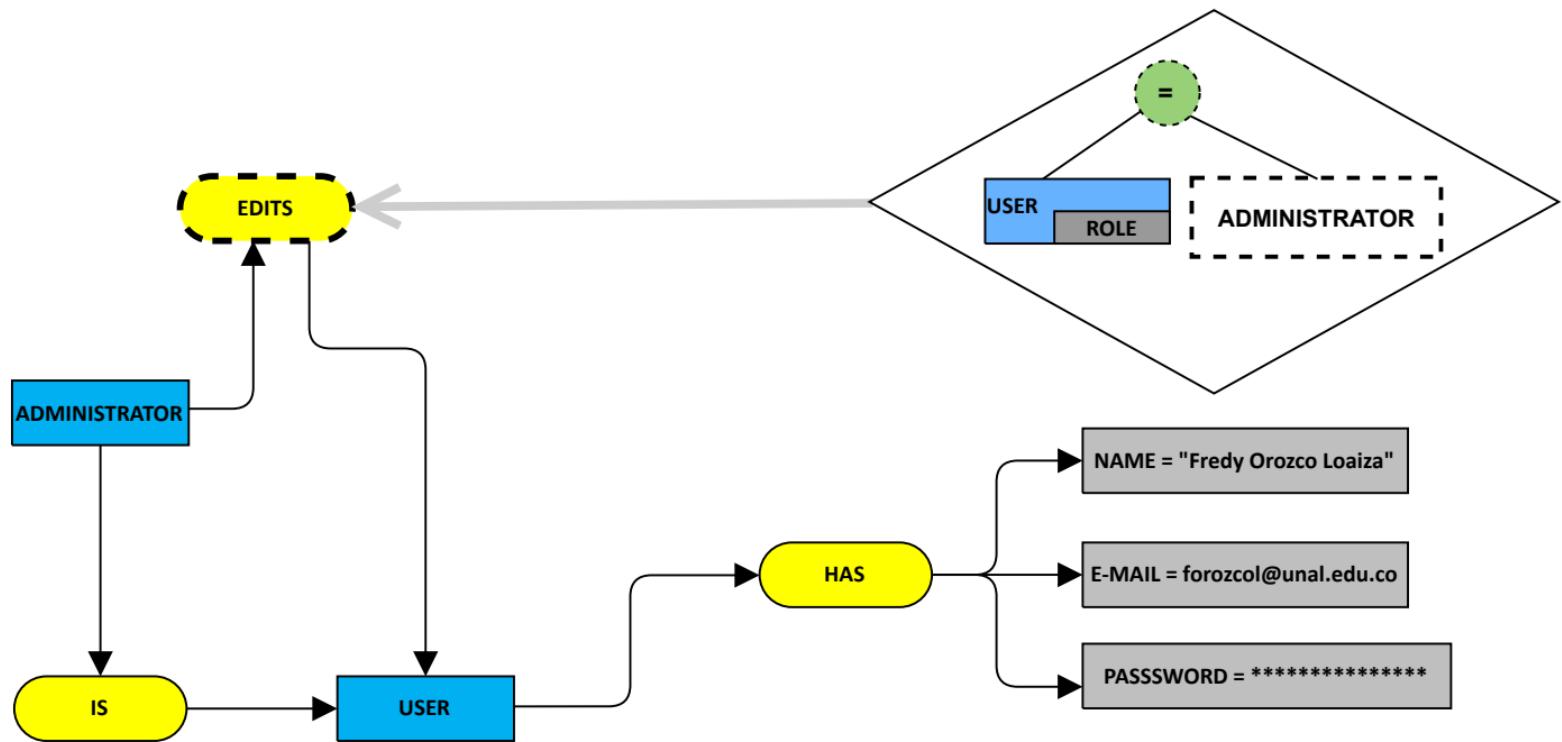
# #9



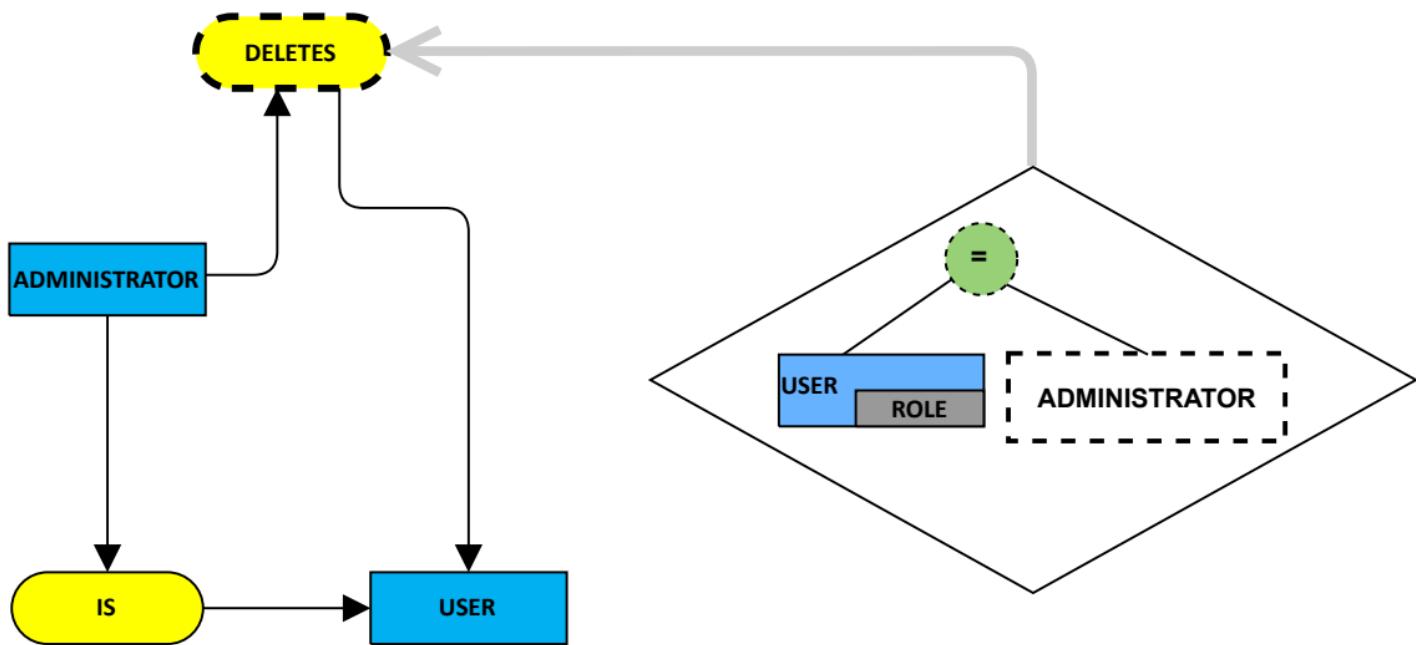
# #10



# #11

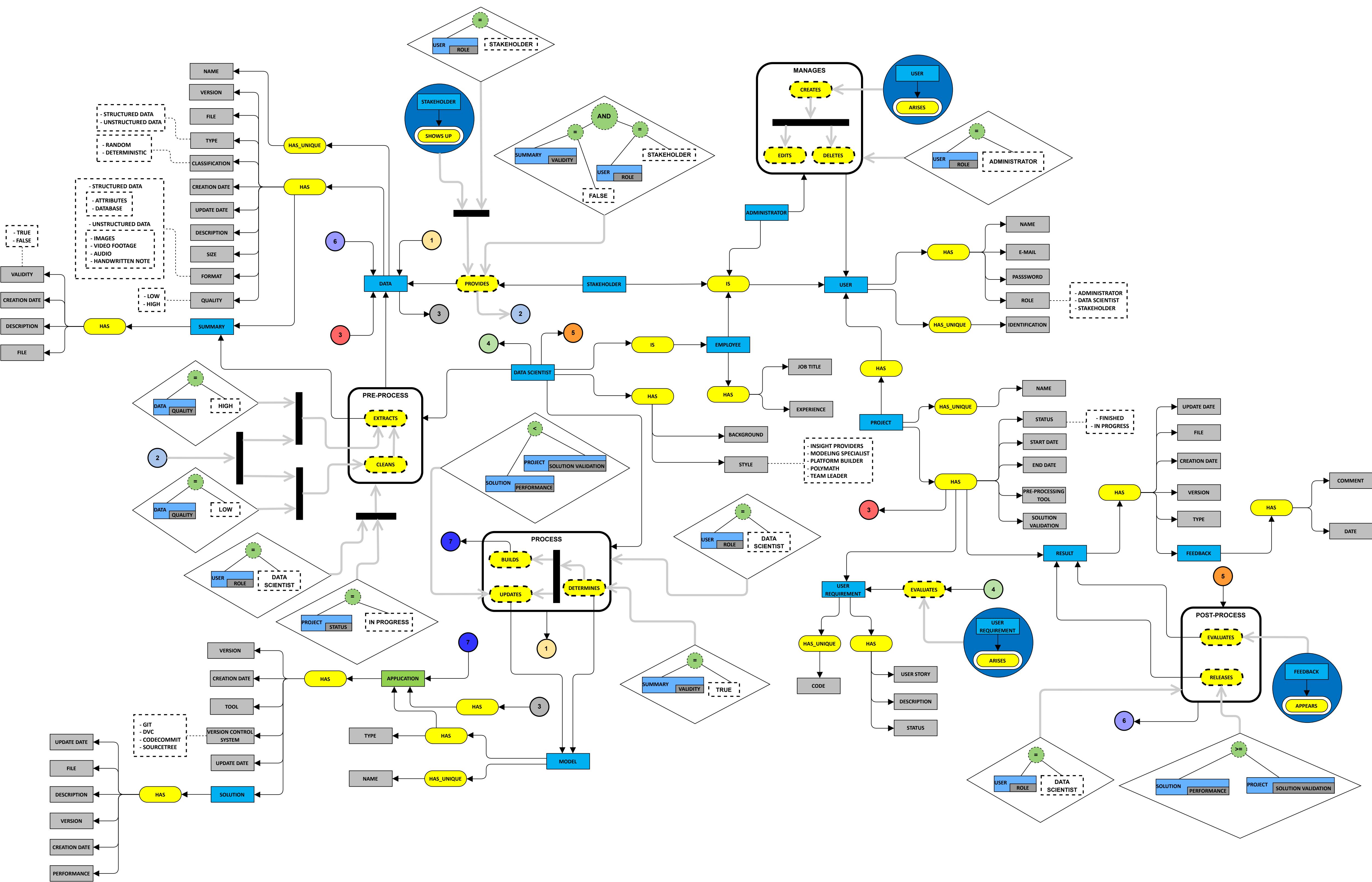


# #12

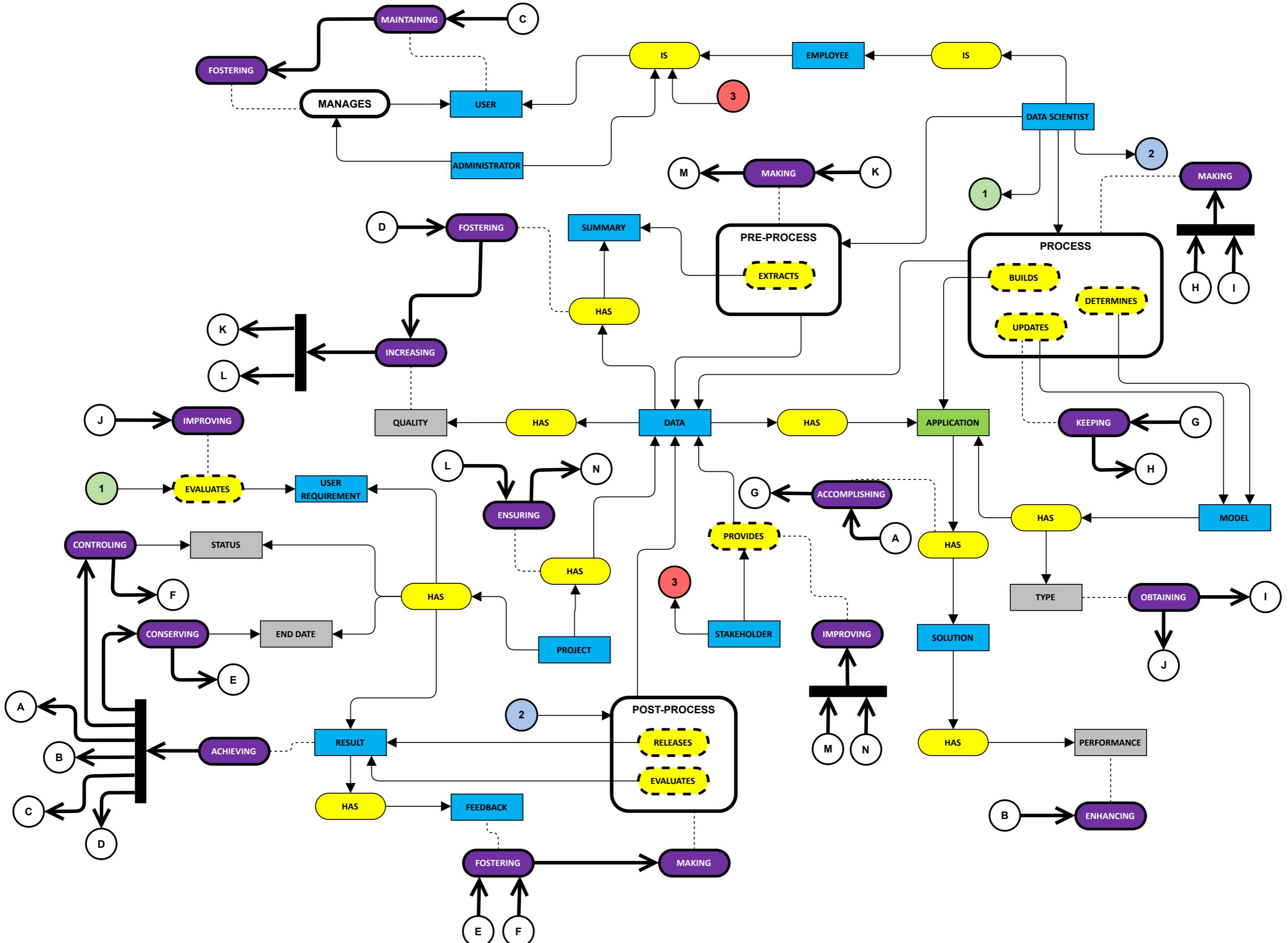


# PRE-CONCEPTUAL SCHEMA

[BACK](#)



# PRE-CONCEPTUAL SCHEMA WITH ACHIEVEMENT SETS



## Document Traceability Table (Preconceptual Schema) < BACK

Original sound/Image/Text	Source	Location	Element	Kind of element	Observations
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 12:32	Data has unique name	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 12:33	Data has version Data has file Data has creation date Data has update date Data has description Data has size	Structural triad	The stakeholder agreed
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 40:02	Data has classification	Structural triad	- When he says: "whether I am working with deterministic data or random data" we infer that Data can be have classification
An aspect which several mentioned was the lack of metrics that would enable the quality of the data to be assessed beforehand.	Text	On Understanding Data Scientists, Page 3, Paragraph 2	Data has quality	Structural triad	We can infer that "quality of the data" refers to Data has a quality
This data can be in a variety of formats, such as structured (attributes in a database) or unstructured data (images, video footage, audio, handwritten notes).	Text	On Understanding Data Scientists, Page 1, Paragraph 2	Data has type Data has format	Structural triad	-We can infer "structured" and "unstructured data" ca be two different type of Data. -We can infer that each of the formats has its own format.
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 17:34	Data has summary	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 17:38	Summary has validity Summary has creation date Summary has description Summary has file	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 10:50	Stakeholder is user	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 8:10	Employee is user	Structural triad	The stakeholder agreed
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 05:56	Administrator is user	Structural triad	The stakeholder agreed
Chris Wiggins is the Chief Data Scientist at The New York Times	Text	Data Scientists at Work, Page 1, Paragraph 1	Data Scientist is employee	Structural triad	- We inferred that the Data Scientist is employee of a company, in this case "The New York Times"
We see that the vast majority ideally pursue people with a background in Engineering, Computer Science, Mathematics, Statistics, Physics, and other related fields	Text	On Understanding Data Scientists, Page 1, Paragraph 5	Data Scientist has background	Structural triad	- We interpret from the context of the article that "people" refers to data scientist
The authors characterized the roles of data scientists in a large software company and explored various working styles of data scientists, having identified five different styles (insight providers, modeling specialists, platform builders, polymaths, and team leaders)	Text	On Understanding Data Scientists, Page 2, Paragraph 5	Data Scientist has style	Structural triad	- We interpret from the context of the article that "working style" refers to data scientist have a working style in you role in a company
Smallwood draws on her broad technical expertise in experimentation, analytics, and recommendations acquired over the course of more than twenty years of experience	Text	Data Scientists at Work, Page 19, Paragraph 2	Employee has experience	Structural triad	- We interpret form the context of the article that Smallwood is an employee in the companies in which it works - We interpret that "more than twenty years of experience" refers to an employee have experience
Claudia Perlich is the Chief Scientist at Distillery	Text	Data Scientists at Work, Page 151, Paragraph 1	Employee has job title	Structural triad	- We interpret that "Chief Scientist" is her job title at that company in which her works
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 4:49	User has unique identification	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 4:49	User has name User has E-mail User has password User has role	Structural triad	The stakeholder agreed
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 14:36	Project has user	Structural triad	- When he says: "depends on the type of project you have" we infer that there is a project in which work is being done - When he says: "depends on the type of project you have" we infer that there is a person that work on a project so we interpret that this persons are users
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 23:54	Project has unique name	Structural triad	The stakeholder agreed
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 23:55	Project has status Project has start date Project has end date Project has pre-processing tool Project has solution validation	Structural triad	The stakeholder agreed
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 14:36	Project has data	Structural triad	When he says: "I am working for an environmental company and I hope that the data you provide me with" we interprre from the context of the interview that "working for ar an enivroment company" refers to a project and this project hava data
spend time studying model results and iterating on models	Text	Data Scientists at Work, Page 7, Paragraph 1	Project has result	Structural triad	We interpret from the context of the article that "model results" can be interpreted as a project on which the Data Scientist is working has a solution
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:31	Project has user requirement	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 23:54	Result has update date Result has file Result has creation date Result has version Result has type	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 26:20	Result has feedback	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 26:46	Feedback has comment Feedback has date	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:44	User requirement has unique code	Structural triad	The stakeholder agreed
Structural triad	Diagram	<a href="https://drive.google.com/file/d/1REFHDxkNuD_7WcOZBCgdq1YVMg-00/view?usp=sharing">https://drive.google.com/file/d/1REFHDxkNuD_7WcOZBCgdq1YVMg-00/view?usp=sharing</a> , business analyst pre-conceptual diagram	User requirement has user story User requirement has description User requirement has status	Structural triad	This information was taken from the work of another team in the requirements engineering course.
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:41	Model has unique name	Structural triad	
"Vamos a poder proceder a elegir un modelo matemático este modelo puede ser un modelo algebraico o estadístico"	Video	<a href="https://www.youtube.com/watch?v=B12sBIVdZHs">https://www.youtube.com/watch?v=B12sBIVdZHs</a> , minute 2:09 Platzi - ¿Qué hace un Data Scientist?	Model has type	Structural triad	-We can interpret from the context of the video that "puede ser un modelo algebraico o estadístico" that a model can have various names
I usually can start by asking: How predictive is this model that we've built?	Text	Data Scientists at Work, Page 7, Paragraph 1	Model has application	Structural triad	- We can interpret from the context of the article that "How predictive is this model" refers to a model can be applied to a different levels of prediction for a specific problem
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 20:25	Data has application	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 20:48	Application has version Application has creation date Application has tool Application has version control system Application has update date	Structural triad	The stakeholder agreed

Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 21:52	Application has solution	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 21:56	Solution has update date Solution has file Solution has description Solution has version Solution has creation date Solution has performance	Structural triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:05	Stakeholder provides data	Dynamic triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 6:22	Administrator creates user	Dynamic triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 6:36	Administrator edits user Administrator deletes user	Dynamic triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 17:31	Data Scientist extracts summary	Dynamic triad	The stakeholder agreed
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 30:20	Data Scientist cleans data	Dynamic triad	When he says: "He cleans it only once" by the context of the recording he means that as a data scientist he cleans the data.
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 39:26	Data Scientist builds application	Dynamic triad	- When he says: "the construction of the model" by the context of the recording we interpret that a Data Scientist builds a model that is applied to a requirement - We infer from the context of the interview that the construction of the model can use a lot of data this refers to application
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 36:47	Data Scientist updates model	Dynamic triad	- When he says: "is the base model and then you add to it ..." by the context of the recording we infer that Data Scientist updates model
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 35:44	Data Scientist determines model	Dynamic triad	-When he says "In your model when you already get it "we infer from the context of the interview that Data Scientist Determinates a model!
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:42	Data Scientist evaluates user requirement	Dynamic triad	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 25:40	Data Scientist releases result	Dynamic triad	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 26:23	Data Scientist evaluates result		
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 06:10	Administrator manages user	Responsibility	The stakeholder agreed
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 45:25	Data Scientist pre-process data Data Scientist process data Data Scientist post-process data	Responsibility	- He says: "basically three stages one is the preprocessing which is everything that has to involve the data, the processing... and then there is the post processing." we interpret that Data Scientist have this three responsibilities
This data can be in a variety of formats, such as structured (attributes in a database) or unstructured data (images, video footage, audio, handwritten notes).	Text	On Understanding Data Scientists, Page 1, Paragraph 2	TYPE: -structured data - unstructured data	Note	
This data can be in a variety of formats, such as structured (attributes in a database) or unstructured data (images, video footage, audio, handwritten notes).	Text	On Understanding Data Scientists, Page 1, Paragraph 2	FORMAT: -STRUCTURED DATA: - attributes -database -UNSTRUCTURED DATA: - images -video footage -audio - handwritten note	Note	
quality of the data	Text	On Understanding Data Scientists, Page 3, Paragraph 11	QUALITY -High -Low	Note	- We inferred that qualita has amount because it can be measured in low or high
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 40:02	CLASSIFICATION: -Random - Deterministic	Note	- When he says: "whether I am working with deterministic data or random data" we infer that Data can be have classification
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 17:48	VALIDITY: -true -false	Note	
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 05:06	ROLE: -Administrator -Data Scientist -Stakeholder	Note	
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 24:00	STATUS: -Finished -In progress	Note	
The authors characterized the roles of data scientists in a large software company and explored various working styles of data scientists, having identified five different styles (insight providers, modeling specialists, platform builders, polymaths, and team leaders)	Text	On Understanding Data Scientists, Page 2, Paragraph 5	STYLE: -Insight providers - Modeling specialists -Platform builder -Polymath -Team leader	Note	
The adoption of such ML platform depends on the cloud strategy of the organization. In case, when an		<a href="https://ml-ops.org/content/state-of-mlops">https://ml-ops.org/content/state-of-mlops</a> , MLOps Informational Standard	VERSION CONTROL SYSTEM: -Git -DVC -Codecommit - Sourcetree	Note	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:21	Stakeholder shows up	Event	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 6:27	User arises	Event	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 26:31	Feedback appears	Event	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:50	User requirement arises	Event	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:17	If User.role = Stakeholder	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 18:02	If Summary.validity = false AND user.role = stakeholder	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 07:36	If User.role = Administrator	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 25:37	If Solution.performance >= project.solution_validation	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 18:22	If User.role = DataScientist	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 17:05	If Summary.validity = true	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 22:20	If Solution.performance < project.solution_validation	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 14:31	If Data.quality = low	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 18:19	If Project.status = in progress	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 14:31	If Data.quality = high	Conditional	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 14:32	If Stakeholder shows up AND User.role = Stakeholder, then Stakeholder provides Data	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 18:02	If Summary.validity = false AND user.role = stakeholder, then Stakeholder provides Data	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 6:29	If User arises, then Administrator creates User	Implication	The stakeholder agreed

Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 7:50	If User.role = Administrator, then Administrator manages User	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 6:37	If Administrator creates User, then Administrator edits User OR Administrator deletes User	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 16:50	If Stakeholder provides Data, then Data.quality = low OR Data.quality = high	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 14:32	If Stakeholder provides Data AND Data.quality = low, then Data Scientist cleans Data	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 17:32	If Stakeholder provides Data AND Data.quality = high, then Data Scientist extracts Summary	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 17:30	If Data Scientist cleans Data, then Data Scientist extracts Summary	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 18:16	If User.role = Data Scientist AND Projecto.status = In Progress then Data Scientist pre-process Data	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 22:15	If Solution.performance < project.solution_validation, then Data Scientist updates model	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 18:42	If User.role = Data Scientist, then Data Scientist process Data	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 18:57	If summary.validity = true, then Data Scientist determines Model	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 18:48	If Data Scientist determines Model, then Data Scientist build Application OR Data Scientist updates Model	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 11:46	If User requirement arises, then Data Scientist evaluates User requirement	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 25:37	If User.role = Data Scientist, then Data Scientist post-process data	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 26:35	If feedback appears, then Data Scientist evaluates result	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 25:39	If Solution.performance >= project.solution_validation, then Data Scientist releases result	Implication	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 37:31	Fostering the users be managed	Requirement	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 37:35	Maintaining the users	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 36:44	Making the data be pre-processed	Requirement	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 34:28	Making the data be processed	Requirement	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 34:03	Keeping the application be updated	Requirement	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 34:38	Obtaining the type of the model	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 34:17	Enhancing the performance of the solution	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 34:11	Accomplishing application has solution	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 36:35	Improving the data be provided	Requirement	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 36:48	Ensuring project has data	Goal	The stakeholder agreed
In order to increase data quality	Text	On Understanding Data Scientists, Page 3, Paragraph 7	Increasing the quality of data	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 37:08	Fostering data has summary	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 35:45	Controlling the status of the project	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 35:47	Conserving the end date of the project	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 37:44	Achieving the result	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 35:35	Fostering the feedback	Goal	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 34:58	Improving the user requirements be evaluated	Requirement	The stakeholder agreed
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 35:21	Making the data be post-processed	Requirement	The stakeholder agreed

# STAKEHOLDER SURVEY < BACK

PROJECT NAME	DATA SCIENTIST				
STAKEHOLDER NAME	CARLOS ALBERTO ALVAREZ HENAO				
DATE	06 / 04 / 2022				
CRITERIA	STRONGLY AGREE	AGREE	NEITHER	DISAGREE	STRONGLY DISAGREE
THE INFORMATION IS COMPLETE ACCORDING TO OUR MEETINGS	X				
THE WORK PRODUCTS I COULD REVIEW ARE CLEAR		X			
THE PROBLEM IS CLEARLY IDENTIFIED	X				
GOALS ARE RELATED TO THE ORGANIZATION	X				
THE SEQUENCE OF THE PROCESS IS ARTICULATED	X				
THE OVERAL QUALITY OF THE DELIVERABLES IS ADEQUATE		X			
SUGGESTIONS	<ul style="list-style-type: none"> <li>- Include in the project the pre-processing, processing and post-processing of the data, grouping together the actions that were already in place.</li> <li>- To be able to determine the time taken for each of the data stages in a project.</li> </ul>				
SIGNATURE					

# STAKEHOLDER SURVEY

PROJECT NAME	DATA SCIENTIST				
STAKEHOLDER NAME	JUAN CAMILO JARAMILLO TASCÓN				
DATE	03/05/2022				
CRITERIA	STRONGLY AGREE	AGREE	NEITHER	DISAGREE	STRONGLY DISAGREE
THE INFORMATION IS COMPLETE ACCORDING TO OUR MEETINGS	X				
THE WORK PRODUCTS I COULD REVIEW ARE CLEAR		X			
THE PROBLEM IS CLEARLY IDENTIFIED		X			
GOALS ARE RELATED TO THE ORGANIZATION	X				
THE SEQUENCE OF THE PROCESS IS ARTICULATED	X				
THE OVERAL QUALITY OF THE DELIVERABLES IS ADEQUATE		X			
SUGGESTIONS	Include more leaf concepts				
SIGNATURE					

# GOAL DIAGRAM < BACK

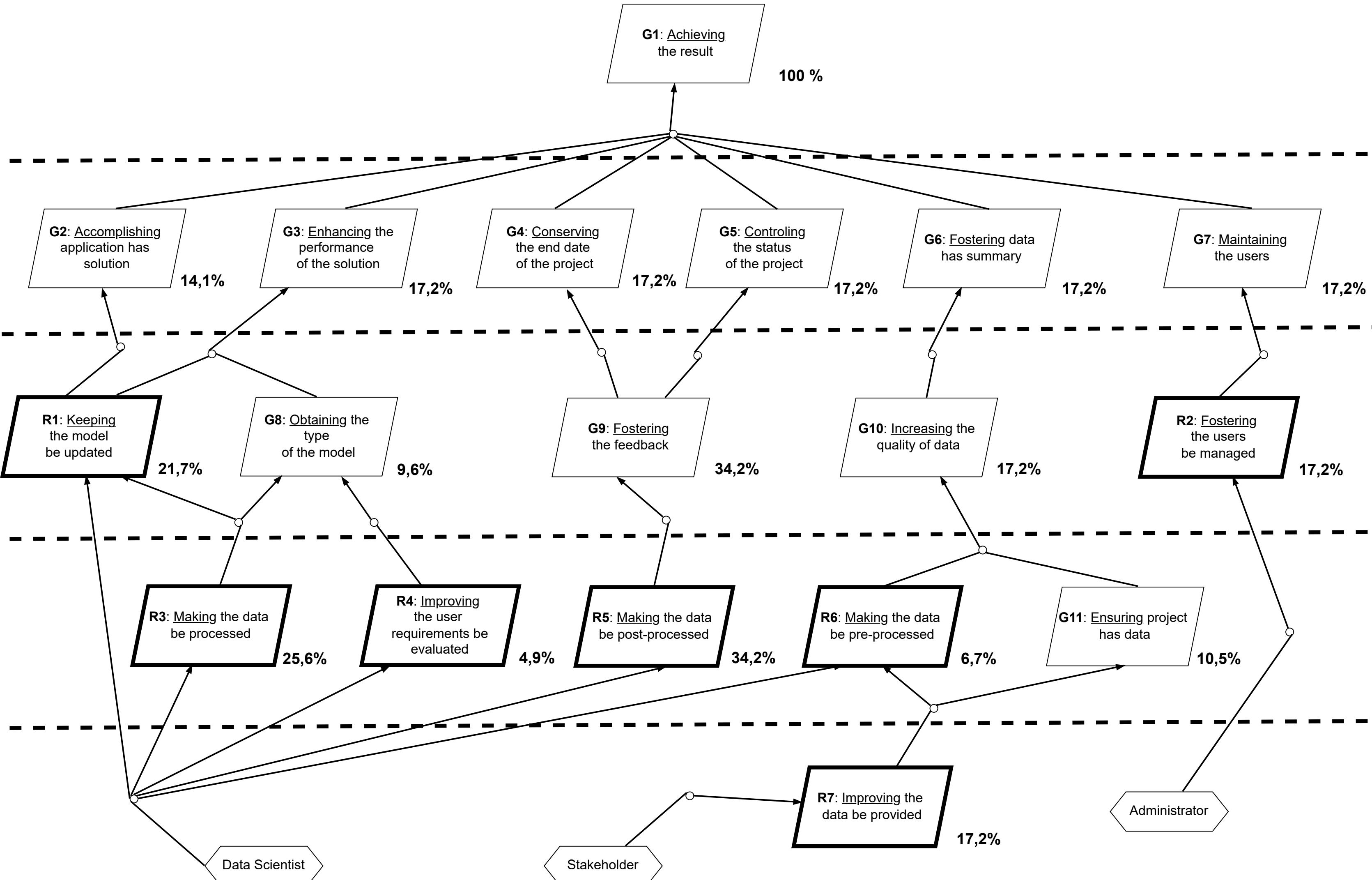
5

4

3

2

1

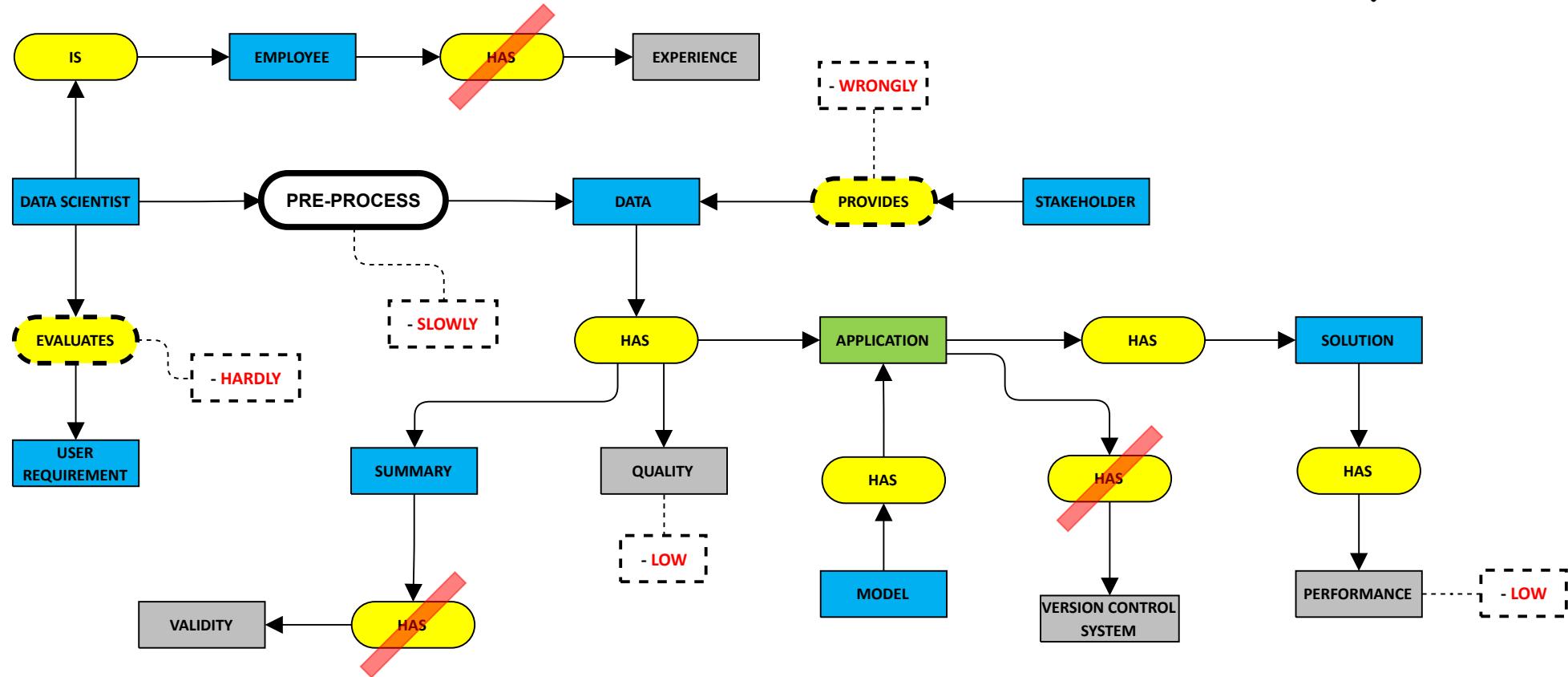


## Percentage Assignment to Goals ↵ BACK

Father	KAOS element	Level	%	Achievement verb	Verb type	Verb weight	Verb %	Element type	Element weight	Element %	Weighted %	Leaf %
ROOT	G1. Achieving the result	5	100,0	ACHIEVE	ACHIEVEMENT	1	100,0	GOAL	2	100,0	100,0	0,0
			100,0			1	100,0		2	100,0	100,0	
G1	G2. Accomplishing application has solution	4	16,67	ACCOMPLISH	ACHIEVEMENT	1	9,1	GOAL	2	16,7	14,1	0,0
G1	G3. Enhancing the performance of the solution	4	16,67	ENHANCE	IMPROVEMENT	2	18,2	GOAL	2	16,7	17,2	0,0
G1	G4. Conserving the end date of the project	4	16,67	CONSERVE	MAINTENANCE	2	18,2	GOAL	2	16,7	17,2	0,0
G1	G5. Controlling the status of the project	4	16,67	CONTROL	MAINTENANCE	2	18,2	GOAL	2	16,7	17,2	0,0
G1	G6. Fostering data has summary	4	16,67	FOSTER	IMPROVEMENT	2	18,2	GOAL	2	16,7	17,2	0,0
G1	G7. Maintaining the users	4	16,67	MAINTAIN	MAINTENANCE	2	18,2	GOAL	2	16,7	17,2	0,0
			100,0			11	100,0		12	100,0	100,0	
G2, G3	R1. Keeping the model be updated	3	22,7	KEEP	MAINTENANCE	2	22,7	REQUIREMENT	1	19,9	21,8	0,0
			22,7			2	22,7		1	19,9	21,8	
G3	G8. Obtaining the type of the model	3	8,6	OBTAIN	MAINTENANCE	2	8,6	GOAL	2	11,4	9,5	0,0
			8,6			2	8,6		2	11,4	9,5	
G4, G5	G9. Fostering the feedback	3	34,2	FOSTER	IMPROVEMENT	2	34,2	GOAL	2	34,2	34,2	0,0
			34,2			2	34,2		2	34,2	34,2	
G6	G10. Increasing the quality of data	3	17,1	INCREASE	IMPROVEMENT	2	17,1	GOAL	2	17,1	17,1	0,0
			17,1			2	17,1		2	17,1	17,1	
G7	R2. Fostering the users be managed	3	17,1	FOSTER	IMPROVEMENT	2	17,1	REQUIREMENT	1	17,1	17,1	17,1
			17,1			2	17,1		1	17,1	17,1	
G8, R1	R3. Making the data be processed	2	26,0	MAKE	ACHIEVEMENT	1	24,9	REQUIREMENT	1	26,0	25,6	25,6
			26,0			1	24,9		1	26,0	25,6	
G8	R4. Improving the user requirements be evaluated	2	4,3	IMPROVE	IMPROVEMENT	2	6,4	REQUIREMENT	1	4,3	5,0	5,0
			4,3			2	6,4		1	4,3	5,0	
G9	R5. Making the data be post-processed	2	34,2	MAKE	ACHIEVEMENT	1	34,2	REQUIREMENT	1	34,2	34,2	34,2
			34,2			1	34,2		1	34,2	34,2	
G10	R6. Making the data be pre-processed	2	8,6	MAKE	ACHIEVEMENT	1	5,7	REQUIREMENT	1	5,7	6,7	0,0
			8,6			1	5,7		1	5,7	6,7	
G10	G11. Ensuring project has data	2	8,6	ENSURE	MAINTENANCE	2	11,4	GOAL	2	11,4	10,5	0,0
			8,6			3	17,1		3	17,1	17,1	
G11, R6	R7. Improving the data be provided	1	17,2	IMPROVE	IMPROVEMENT	2	17,2	REQUIREMENT	2	17,2	17,2	17,2
			17,2			2	17,2		2	17,2	17,2	99,1

# PROBLEMS PRE-CONCEPTUAL SCHEMA

◀ BACK

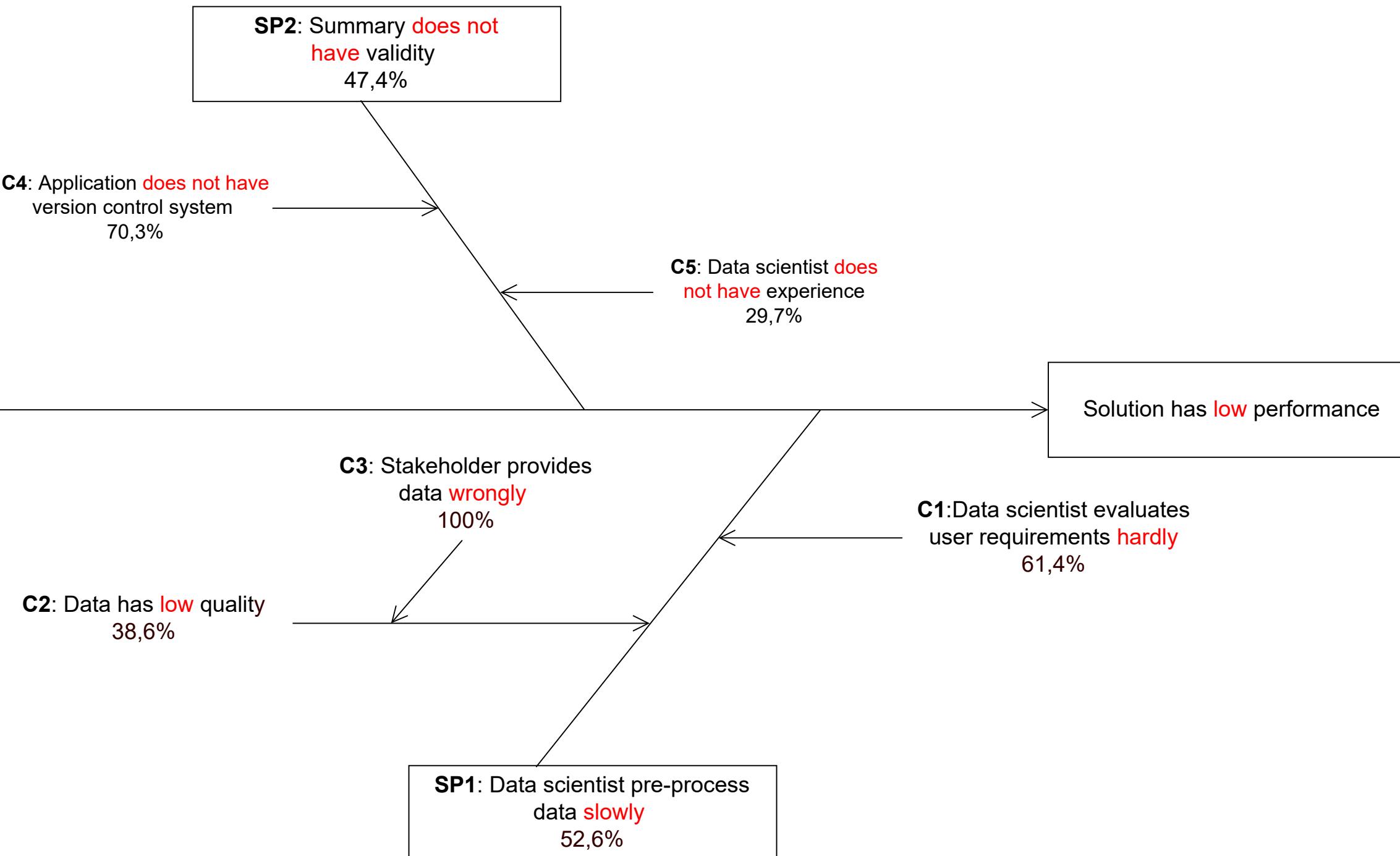


## Document Traceability Table (Problems Preconceptual Schema)

Original sound/Image/Text	Source	Location	Element	Kind of element	Observations
the difficulties she feels are related to the fact that she has no training in the field of data science, and that was the reason that led her to pursue a master's degree focused on data analysis.	Text	On Understanding Data Scientists, Page 4, Paragraph 2	Employee does not have experience	Problem	We can deduce from the context of the article that this person works in a place where she is an employee, from here we note that she expresses not having enough experience.
Chris Wiggins is the Chief Data Scientist at The New York Times	Text	Data Scientists at Work, Page 1, Paragraph 1	Data Scientist is Employee	Structural triad	- We inferred that the Data Scientist is employee of a company, in this case "The New York Times"
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:42	Data Scientist Evaluates User Requirement	Dynamic triad	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 29:23	Data Scientist Evaluates User Requirement hardly	Problem	
Meeting with stakeholder data scientist Carlos Alberto Alvarez Henao	Video	interview_data_scientist.mkv 59:20	Data Scientist Pre-Process Data	Responsibility	During the meeting he referred that these processes performed by the data scientist should be grouped into pre-processing, processing and post-processing.
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 28:30	Data Scientist Pre-Process Data slowly	Problem	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 11:05	Stakeholder provides data	Dynamic triad	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 29:43	Stakeholder provides data wrongly	Problem	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 17:34	Data has summary	Structural triad	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascon	Video	entrevista_2.mkv 28:20	Summary does not have validity	Problem	
An aspect which several mentioned was the lack of metrics that would enable the quality of the data to be assessed beforehand.	Text	On Understanding Data Scientists, Page 3, Paragraph 2	Data has quality	Structural triad	We can infer that "quality of the data" refers to Data has a quality

In the remaining cases, participants claim that most of the anomalies that affect the quality of the data concern inputs that have been wrongly introduced by humans	Text	On Understanding Data Scientists, Page 3, Paragraph 11	Data has low quality	Problem	- We can interpret that "the anomalies that affect the quality of the data" relates to the low quality of it
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 20:25	Data has application	Structural triad	
I usually can start by asking: How predictive is this model that we've built?	Text	Data Scientists at Work, Page 7, Paragraph 1	Model has application	Structural triad	- We can interpret from the context of the article that "How predictive is this model" refers to a model can be applied to different levels of prediction for a specific problem
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 31:47	Application does not have version control system	Problem	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 21:52	Application has solution	Structural triad	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 21:56	Solution has performance	Structural triad	
Meeting with Stakeholder Data Scientist Juan Camilo Jaramillo Tascón	Video	entrevista_2.mkv 28:13	Solution has low performance	Problem	

# CAUSE-AND-EFFECT DIAGRAM < BACK



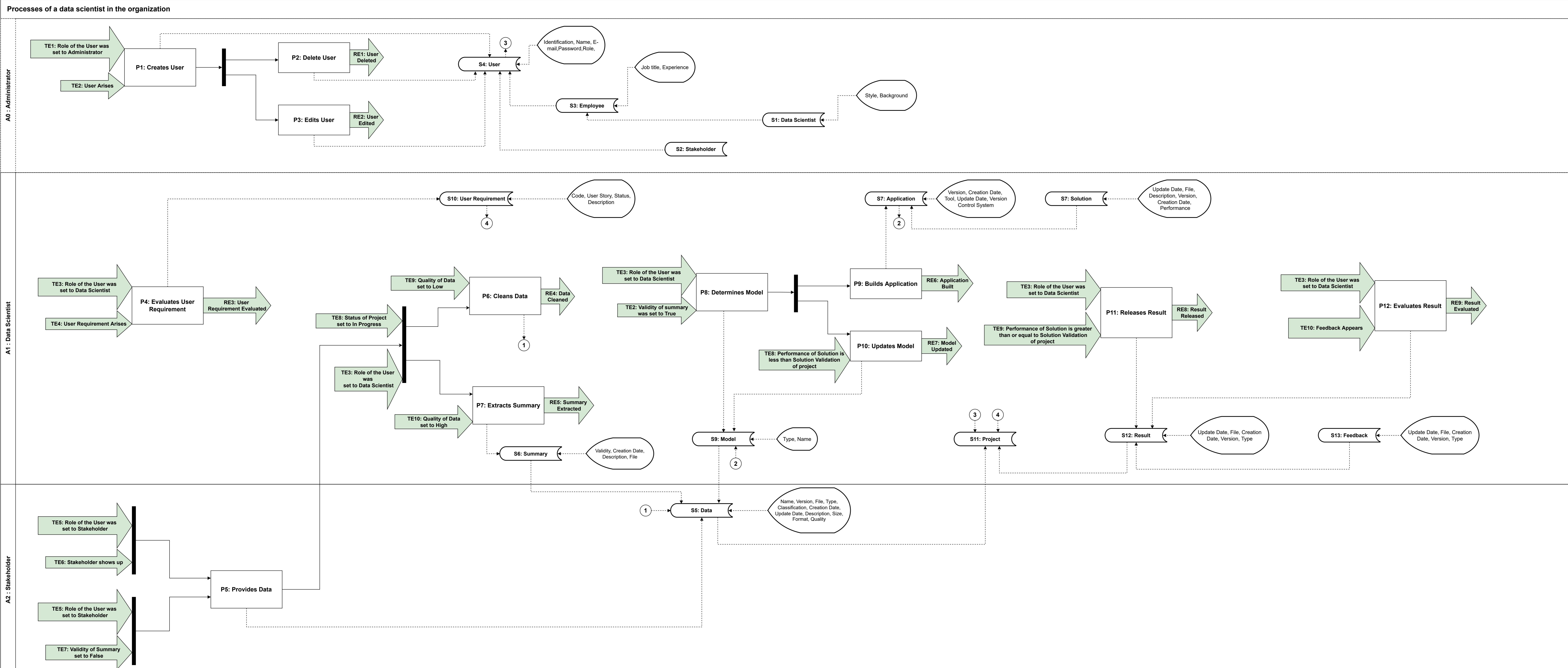
## Percentage Assigment to Problems < BACK

	C1	C3	C4	C5
P1				
P2				
P3				
P4	G8			G8
P5		G11		
P6		G10		
P7	G10, G6			
P8				G8
P9			G2	
P10			G2,G3	
P11				
P12				

	C1	C3	C4	C5
P1				
P2				
P3				
P4	9,6			9,6
P5		10,5		
P6		17,2		
P7	34,4			
P8				9,6
P9			14,1	
P10			31,3	
P11				
P12				
Total	44	27,7	45,4	19,2
		136,3		
	32,3%	20,3%	33,3%	14,1%

	FORMULATE	VALUE	PERCENTAGE
SP2	C4+C5	47,4%	
SP1	C2+C1	52,6%	
		100,0%	
C1	32,3%	32,3%	61,4%
C2	C3	20,3%	38,6%
C3	20,3%	20,3%	100,0%
C4	33,3%	33,3%	70,3%
C5	14,1%	14,1%	29,7%

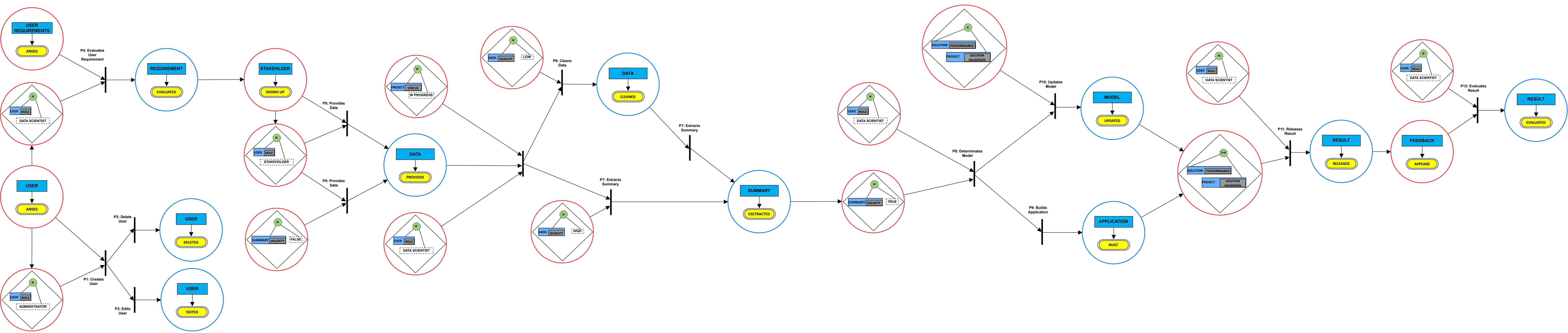
# PROCESS DIAGRAM < BACK



## Process Diagram Explanatory Table < BACK

Name	Goal	Duration/Frequency	How?/Where?	Problems	Business rules	Events
P1: Creates User	G7 : Maintaining the users	5 minutes / 4 times a month	For each user, the personal information is registered. This process takes place were the software is developed.		BR1: ADMINSTRATOR MANAGE THE USER	- TE1: Role of the User was set to Administrator - TE2: User Arises
P2: Deletes User	G7 : Maintaining the users	2 minutes / once a month	Users have sometimes the need to quit the system. In such a case, the user is deleted. This process takes place were the software is developed.		BR1: ADMINSTRATOR MANAGE THE USER	RE1: User Deleted
P3: Edits User	G7 : Maintaining the users	4 minutes / 9 times a month	This process takes place were the software is developed whenever the user information needs some adjustment		BR1: ADMINSTRATOR MANAGE THE USER	RE2: User Edited
P4: Evaluates user requirements	G8 : Obtaining the type of model.	24 hours / 2 times a month	When a project is developed, it comes with some user requirements that the data scientist must assess. This process takes place in the data scientist workspace	- C5 : DATA SCIENTIST does not have EXPERIENCE - C1 : DATA SCIENTIST evaluates USER REQUIREMENTS hardly	BR15:DATA SCIENTIST EVALUATES USER REQUIEREMENTS	- TE3: Role of the User was set to Data Scientist - TE4: User Requirement Arises - RE3: User Requirement Evaluated
P5: Provides Data	G11: Ensuring project has data	10 minutes / 2 times a month	It is very important that the stakeholder provides the data with which the data scientist will work. This process takes place in the data scientist workspace	C3 : STAKEHOLDER provides DATA wrongly	BR2:STAKEHOLDER PROVIDES DATA 1, BR3: STAKEHOLDER PROVIDES DATA 2	- TE5: Role of the User was set to Stakeholder - TE6: Stakeholder shows up - TE5: Role of the User was set to Stakeholder - TE7: Validity of Summary set to False
P6: Cleans Data	G10: Increasing the quality of data	36 hours / 2 times a month	In many cases, the data provided are not of the best quality, so it is necessary to clean and deputate them. This process takes place in the data scientist workspace	C3 : STAKEHOLDER provides DATA wrongly	BR9:DATA SCIENTIST CLEANS THE DATA	- TE9: Quality of Data set to Low - TE8: Status of Project set to In Progress - TE3: Role of the User was set to Data Scientist - RE4: Data Cleaned
P7: Extracts Summary	- G10: Increasing the quality of data - G6: Fostering data has summary	2 hours / 2 times a month	A summary of the data is necessary to see how the data behaves and to find relationships in it. This process takes place in the data scientist workspace	C1 : DATA SCIENTIST evaluates USER REQUIREMENTS hardly	BR10:DATA SCIENTIST EXTRACTS DATA 1 BR11: DATA SCIENTIST EXTRACTS DATA 2	- TE8: Status of Project set to In Progress - TE3: Role of the User was set to Data Scientist - TE10: Quality of Data set to High - RE5: Summary Extracted
P8: Determines Model	G8 : Obtaining the type of model.	1 hour / 2 times a month	This process takes place in the data scientist workspace. It is important to determine a model that best fits the behavior of the data.	C5 : DATA SCIENTIST does not have EXPERIENCE	BR7:DATA SCIENTIST DETERMINES THE MODEL	- TE3: Role of the User was set to Data Scientist - TE2: Validity of summary was set to True
P9: Builds Application	G2 : Accomplishing application has solution	6 hours / 2 times a month	This process takes place in the data scientist workspace. The application is necessary to be able to work with the model and the data, using a programming and development tool to work with them.	C4 : APPLICATION does not have VERSION CONTROL SYSTEM	BR5:DATA SCIENTIST BUILDS APPLICATION	RE6: Application Built
P10: Updates Model	- G2 : Accomplishing application has solution - G3 : Enhancing the performance of the solution	2 hours / 8 times a month	This process takes place in the data scientist workspace. Sometimes the model needs refinement if the desired performance is not obtained at first, so it is updated.	C4 : APPLICATION does not have VERSION CONTROL SYSTEM	BR6: DATA SCIENTIST UPDATES THE MODEL	- TE8: Performance of Solution is less than Solution Validation of project - RE7: Model Updated
P11: Releases Result	G1 : Achieving the result	4 hours / 2 times a month	This process takes place in the data scientist workspace. It is important to be able to show the conclusions of what was obtained in the data processing, therefore the result must be concise and with the key points that were found to show in a clear and simple way to the stakeholders.		BR13:DATA SCIENTIST RELASES RESULT	- TE3: Role of the User was set to Data Scientist - TE9: Performance of Solution is greater than or equal to Solution Validation of project - RE8: Result Released
P12: Evaluates Result	- G4 : Conserving the end date of the project - G5 : Controlling the status of the project	2 hours / 6 times a month	This process takes place in the data scientist workspace. Continuous evaluation of results is important as stakeholders give feedback that should be evaluated on results		BR14: DATA SCIENTIST EVALUATES RESULT	- TE3: Role of the User was set to Data Scientist - TE10: Feedback Appears - RE9: Result Evaluated

# EVENT INTERACTION GRAPH < BACK



## Data Dictionary < BACK

Name	Alias	Type	Process involved	Features
A0: Administrator		Agent	P1: Creates User, P2: Deletes User, P3: Edits User	name, e-mail, password, role, identification
A1: Data Scientist		Agent and I/O storage	P1: Creates User, P2: Deletes User, P3: Edits User, P4: Evaluates user requirements, P6: Cleans Data, P7: Extracts Summary, P8: Determines Model, P9: Builds Application, P10: Updates Model, P11: Releases Result, P12: Evaluates Result	name, e-mail, password, role, identification, job title, experience, style, background
A2: Stakeholder		External agent and I/O storage	P1: Creates User, P2: Deletes User, P3: Edits User, P5: Provides Data	name, e-mail, password, role, identification
A3: User		Agent and I/O storage	P1: Creates User, P2: Deletes User, P3: Edits User	name, e-mail, password, role, identification
A4: Employee		Agent and I/O storage	P1: Creates User, P2: Deletes User, P3: Edits User	name, e-mail, password, role, identification, job title, experience
S5: Data		I/O Storage	P5: Provides Data, P6: Cleans Data,	name, version, file, type, classification, creation date, update date, description, size, format, quality
S6: Summary		I/O Storage	P7: Extracts Summary	validity, creation date, description, file
S7: Application		I/O Storage	P9: Builds Application, P10: Updates Model	version, creation date, tool, update date, version control system
S8: Solution		I/O Storage	P9: Builds Application, P10: Updates Model	update date, file, description, version, creation date, performance
S9: Model		I/O Storage	P8: Determines Model, P10: Updates Model	type, name
S10: User requirement		I/O Storage	P4: Evaluates user requirements	code, user story, status, description
S11: Project		I/O Storage		name, status, start date, end date, pre-processing tool, solution validation
S12: Result		I/O Storage	P11: Releases Result	update date, file, creation date, version, type
S13: Feedback		I/O Storage	P12: Evaluates Result	comment, date

## Business Rules < BACK

Code	Name	Description	Formulae	Source	Related business rule
BR1	ADMINISTRATOR MANAGE THE USER	User can manage other users only if his role is administrator.	USER.ROLE="ADMINISTRATOR"	STAKEHOLDER	BR16
BR2	STAKEHOLDER PROVIDES DATA 1	The stakeholder can provides the data if, the summary validation is false and the user role is stakeholder	(SUMMARY.VALIDITY="FALSE" & USER.ROLE="STAKEHOLDER")	STAKEHOLDER	BR16, BR18
BR3	STAKEHOLDER PROVIDES DATA 2	THE STAKEHOLDER CAN PROVIDES DATA IF THE STAKEHOLDER IS SHOWS UP AND THE USER ROLE IS STAKEHOLDER	(STAKEHOLDER SHOWS UP & USER.ROLE="STAKEHOLDER")	STAKEHOLDER	BR16
BR4	DATA SCIENTISTS PROCESS DATA	data scientist can processes the data if the user's role is data scientist	USER.ROLE="DATA SCIENTIST"	STAKEHOLDER	BR16
BR5	DATA SCIENTIST BUILDS APPLICATION	Data scientist can builds the application only if data scientist determines the model		STAKEHOLDER	BR7
BR6	DATA SCIENTIST UPDATES THE MODEL	Data scientist can updates the application only if data scientist determines the model and the performance solution less than the project solution validation	DATA SCIENTIST DETERMINES THE MODEL & SOLUTION. PERFORMANCE < PROJECT. SOLUTION_VALIDATION	STAKEHOLDER	BR16, BR7
BR7	DATA SCIENTIST DETERMINES THE MODEL	The data scientist can determines the model only if summary validation is true	SUMMARY.VALIDITY="TRUE"	STAKEHOLDER	BR18
BR8	DATA SCIENTIST PRE-PROCESS DATA	the data scientist can pre-processes the data if the user's role is data scientist and the project's status is in progress	USER.ROLE="DATA SCIENTIST" & PROJECT.STATUS="IN PROGRESS"	STAKEHOLDER	BR16, BR17
BR9	DATA SCIENTIST CLEANS THE DATA	The data scientist can cleans the data if the stakeholder provides the data and the data quality is low.	STAKEHOLDER PROVIDES DATA & DATA.QUALITY="LOW"	STAKEHOLDER	BR19, BR2, BR3
BR10	DATA SCIENTIST EXTRACTS DATA 1	The data scientist can extracts the data only if the data scientist cleaned the data		STAKEHOLDER	BR9
BR11	DATA SCIENTIST EXTRACTS DATA 2	The data scientist can extracts the data if the data quality is high and the stakeholder provides the data.	STAKEHOLDER PROVIDES DATA & DATA.QUALITY="HIGH"	STAKEHOLDER	BR19, BR2, BR3
BR12	DATA SCIENTIST POST-PROCESS DATA	Data scientist can post-process data only if the user role is data scientist	USER.ROLE="DATA SCIENTIST"	STAKEHOLDER	BR16
BR13	DATA SCIENTIST RELEASES RESULT	The data scientist can releases the results only if the performance of the solution is greater than or equal to the solution validation of the project.	SOLUTION.PERFORMANCE >= PROJECT.SOLUTION_VALIDATION	STAKEHOLDER	
BR14	DATA SCIENTIST EVALUATES RESULT	The data scientist can evaluates the results only if the feedback appears.		STAKEHOLDER	
BR15	DATA SCIENTIST EVALUATES USER REQUIEREMENTS	The data scientist can evaluates the requirements only if the user requirements arises		STAKEHOLDER	
BR16	USER ROLE	the user role can only be ADMINISTRATOR, DATA SCIENTIST STAKEHOLDER		STAKEHOLDER	
BR17	PROJECT STATUS	the project status can only be "finish" or "in progress"		STAKEHOLDER	
BR18	SUMMARY VALIDITY	the validation of the summary can only be "TRUE" or "FALSE".		STAKEHOLDER	
BR19	DATA QUALITY	data quality can only be "LOW" or "HIGH".		STAKEHOLDER	

# Team Charter < BACK

<b>Team name</b>	Team #2: Data Scientist	<b>Date</b>	24/04/2022
<b>Team leader</b>	Diego Valentín Osorio Marín	<b>Sponsor</b>	UNAL
<b>Team purpose</b>	As the purpose of our team is to do all the analysis work on all the responsibilities of a data scientist in a software development company, using the UNC-METHOD.		
<b>Mission</b>	Our mission is to simplify the work of data scientists within a company, through a software product that increases productivity and optimizes workflow, taking into account certain objectives.		
<b>Objetives</b>	<ul style="list-style-type: none"> <li>- Identify and analyze data scientist problems and solutions within a company.</li> <li>- To come up with consistent models that correctly represent what our stakeholders express.</li> <li>- Make good use of time to generate good work products.</li> </ul>		
<b>Scope</b>	Analysis and development of what a data scientist does within a company.		
<b>Team members</b>			
Name	Roles	<b>Responsibilities</b>	
Diego Valentín Osorio Marín	Leader, Meeting Coordinator, Analyst, Moderator	<ul style="list-style-type: none"> <li>- Elaborate pre-conceptual schema.</li> <li>- Elaborate kanban board</li> <li>- Elaborate alphas advance report</li> <li>- Make corrections corresponding to deliverable 1</li> <li>- Elaborate goal diagram</li> <li>- Elaborate cause and effect diagram</li> <li>- Assess the team</li> <li>- Assign tasks</li> </ul>	
Jaime Andrés Monsalve Ballesteros	Documenter, Programmer, Tester	<ul style="list-style-type: none"> <li>- Elaborate document traceability table</li> <li>- Elaborate percentage assignment to goals and problems</li> <li>- Verify consistency</li> </ul>	
Santiago Castro Tabares	Analyst, Researcher, Programmer, Tester	<ul style="list-style-type: none"> <li>- Check consistency</li> <li>- Elaborate pre-conceptual schema.</li> <li>- Elaboration of process diagram and explanatory table</li> <li>- Elaborate Event interaction graph</li> </ul>	
Fredy Alberto Orozco Loaiza	Digital information manager, Analyst, Researcher	<ul style="list-style-type: none"> <li>- Contact the stakeholder</li> <li>- Elaborate pre-conceptual schema.</li> <li>- Elaborate of Elicitation cards</li> <li>- Elaborate of Business rules</li> <li>- Elaborate the Domain model</li> <li>- Elaborate Data Dictionary</li> </ul>	

## Method assesment template < BACK

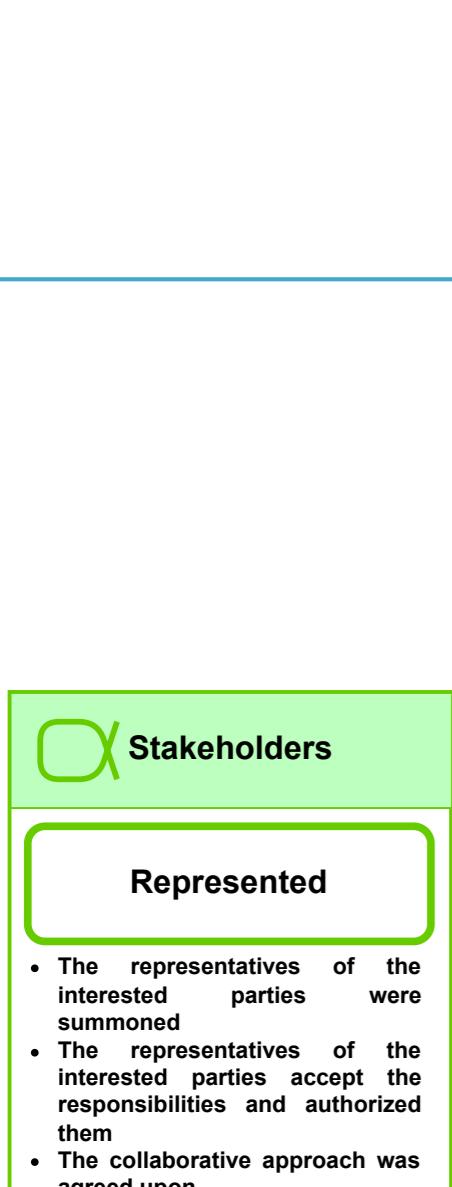
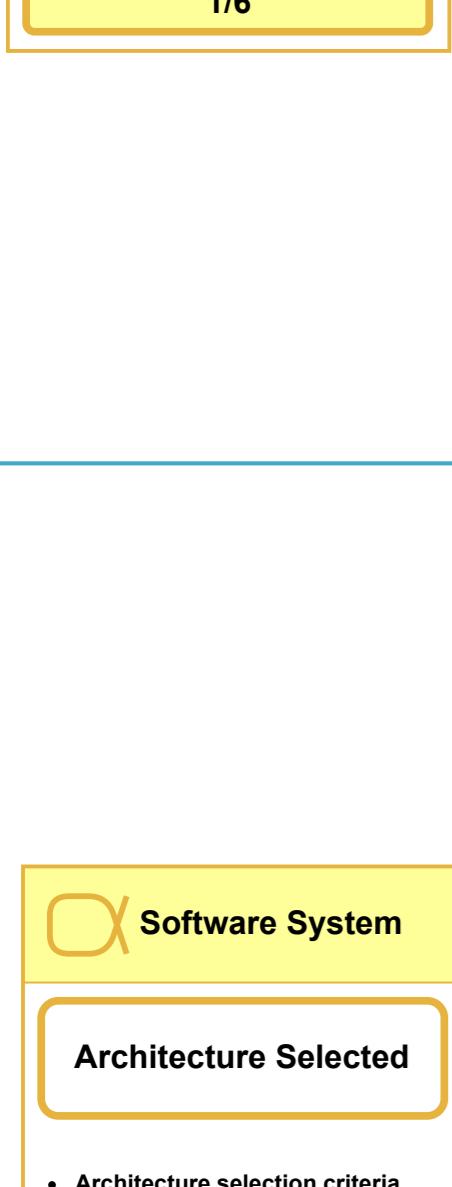
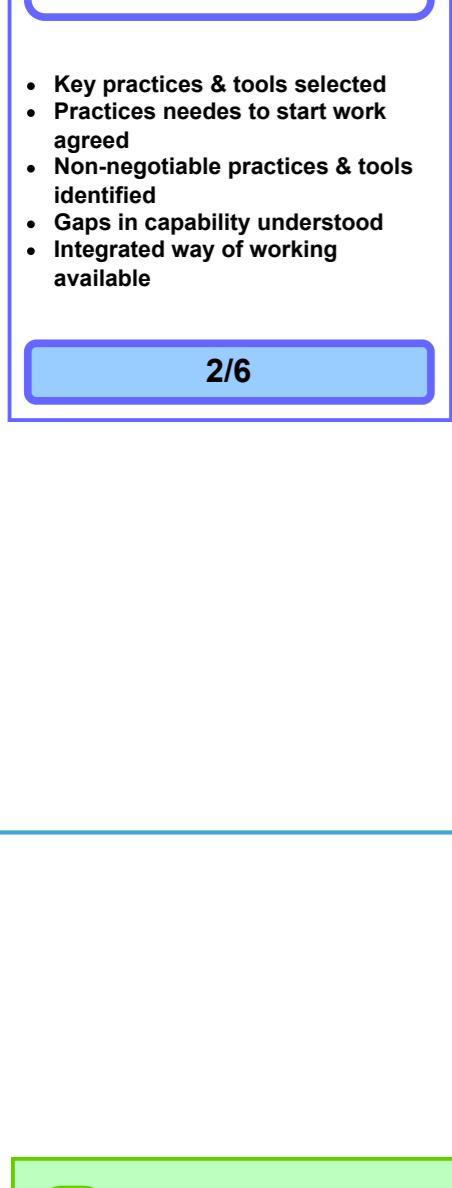
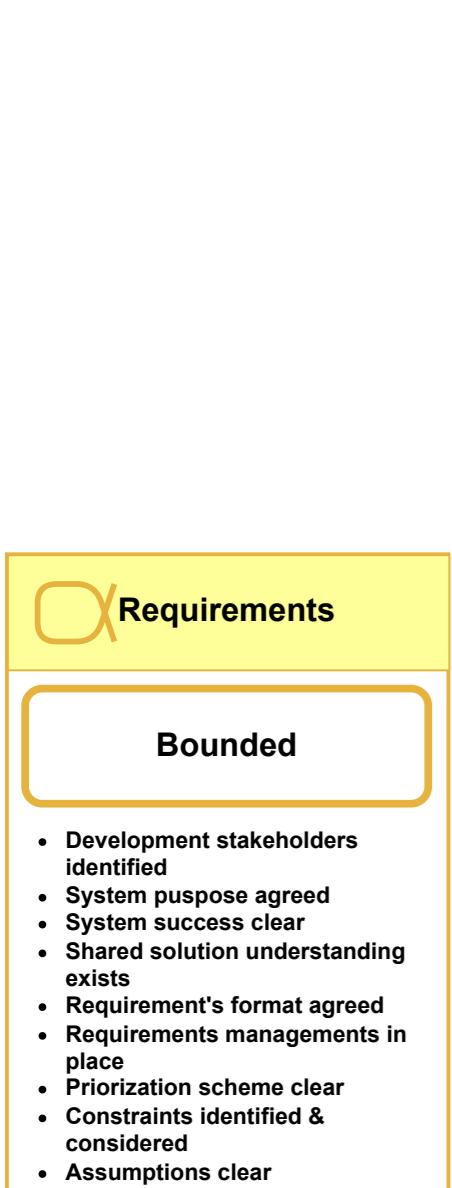
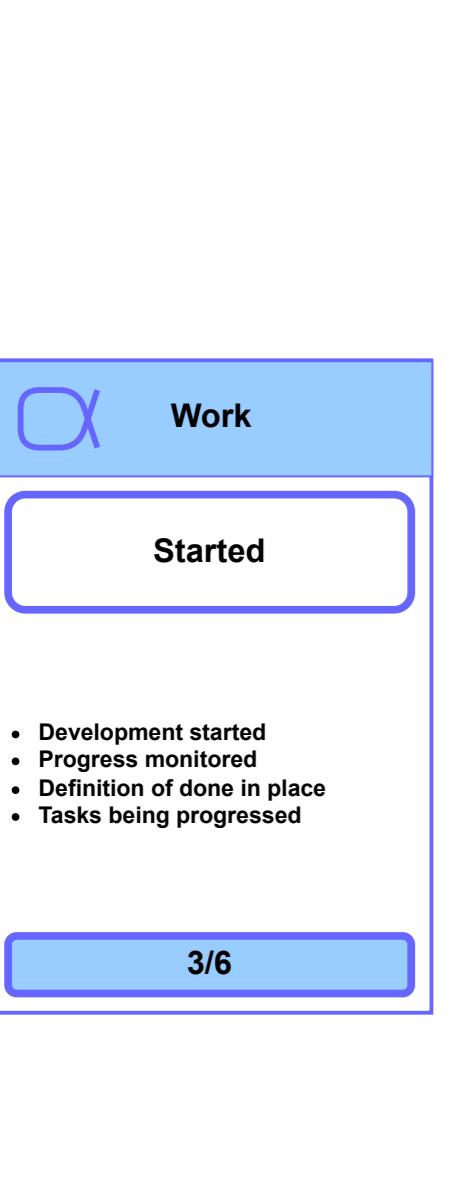
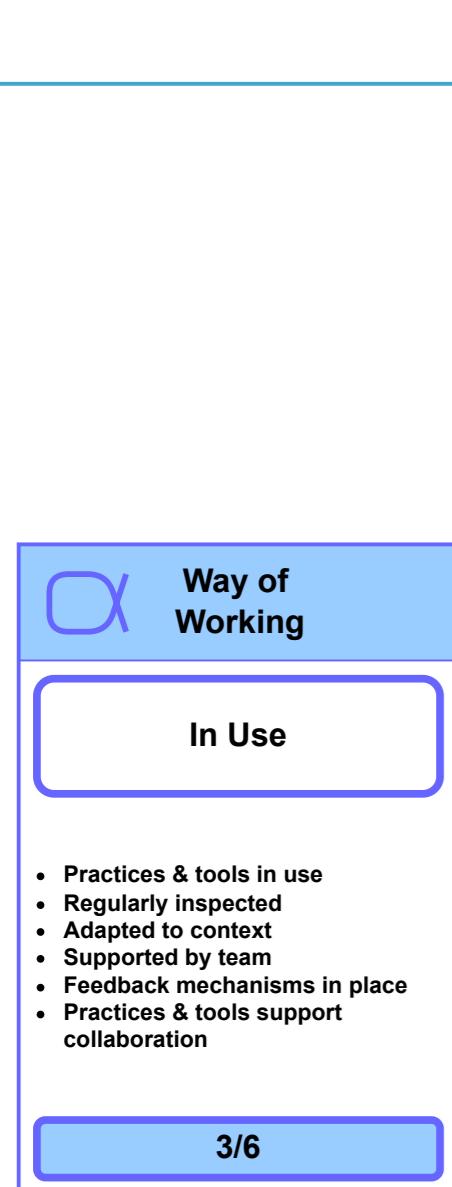
Team	Team #2 - Data Scientist		Date	27-4-2022
UNC-METHOD		Suggestions for improvement		
Phase	Practice	Consistency	Completeness	Correction
Software Context	Discourse-based opportunity modeling		Elicitation cards should be color-coded to better differentiate between object, actor and function cards.	
Problem Analysis	Template-based elicitation of requirements		The process diagram explanatory table should have an extra column for observations.	
Software Context	Discourse-based opportunity modeling		There should be a guide or advice on how to formulate the questions to the stakeholders, so that the best and maximum amount of information can be obtained.	
Problem Analysis	Visual modeling of requirements		There should be information on the syntax of the event interaction graph, as well as some examples to be consistent with the process diagram.	
Problem Analysis	Template-based elicitation of requirements		There should be examples of how to obtain the percentages of goals and problems, explaining step by step how to obtain them, since this process can be confusing and difficult to understand.	
Software Context - Problem Analysis	Visual modeling of requirements			For ease and speed of reading, references in the preconceptual schema should have different notation for dynamic relationship implications, achievement implications, and normal triads.
Problem Analysis	Visual modeling of requirements			In the process diagram there should be a way to better identify the actors, either by means of a color or a more characteristic symbol.
Problem Analysis	Visual modeling of requirements		There should be a better indication of what the process flow looks like.	
Problem Analysis	Visual modeling of requirements	In order to obtain the causes of the cause-effect diagram, from the problems pre-conceptual schema, there should be a way to hierarchize these causes from the same pre-conceptual.		
Problem Analysis	Template-based elicitation of requirements		Mathematical expressions should be allowed in the syntax.	

# KANBAN BOARD < BACK

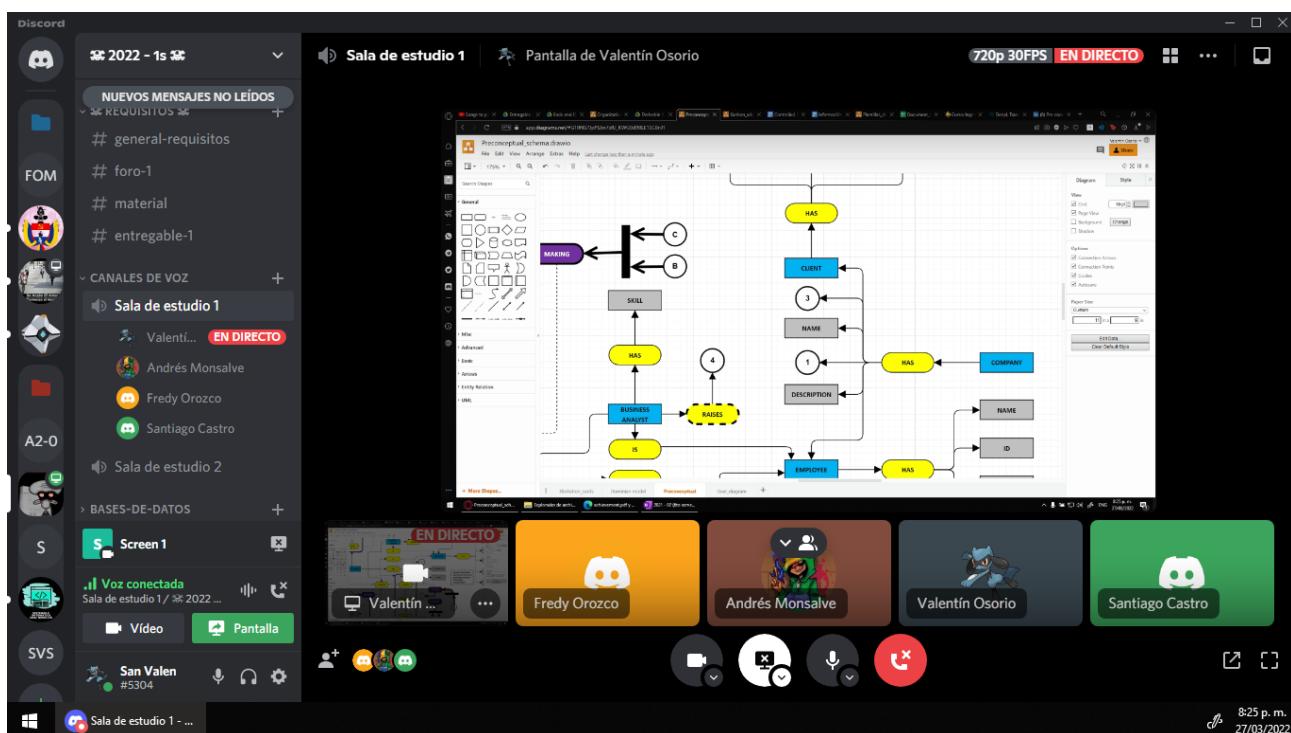
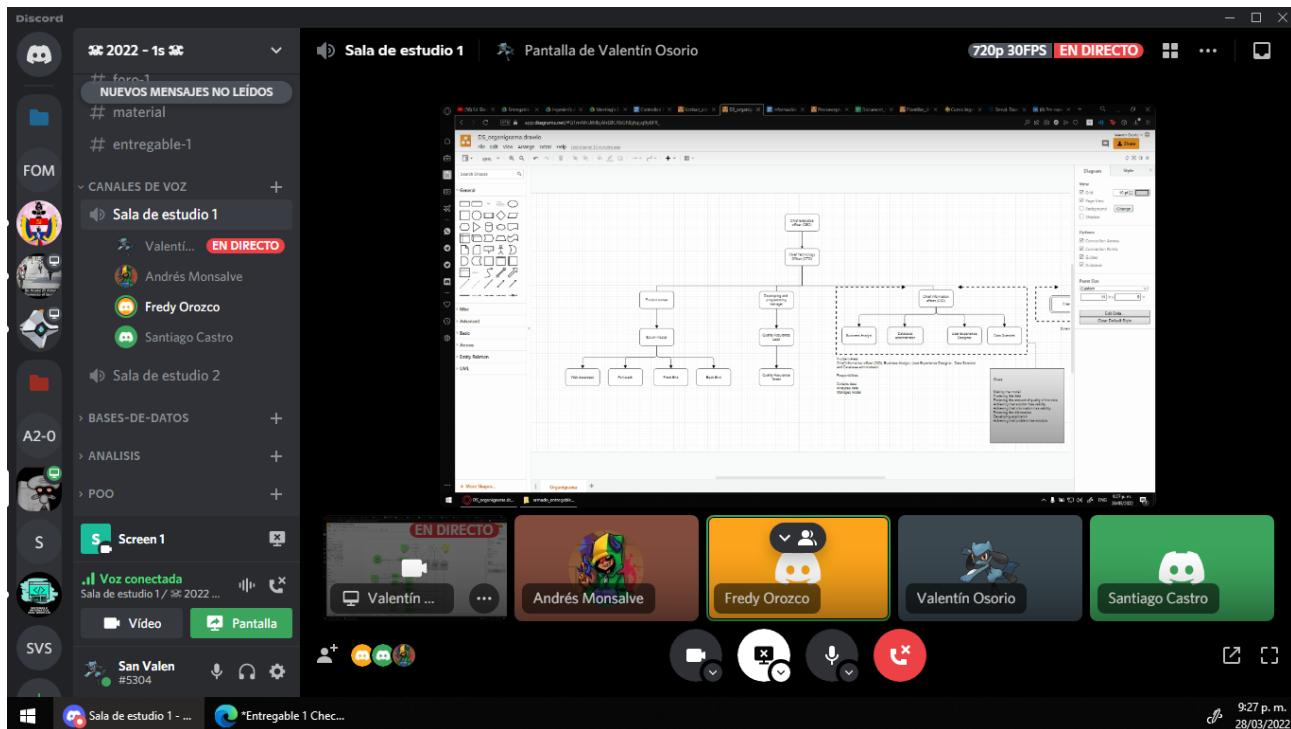
GOAL	TO DO	DOING	DONE
		<p>Task 1 : Gathering information and interviews on data scientists Responsible: Diego Valentin Osorio Marin, Fredy Alberto Orozco Loaiza</p> <p>Task 2 : Making the controlled dialog Responsible : Diego Valentin Osorio Marin</p> <p>Task 3 : Collecting and organizing the digital files Responsible : Team</p> <p>Task 4 : Filling the elicitation cards Responsible : Fredy Alberto Orozco Loaiza</p> <p>Task 5 : Making and updating the organizational chart Responsible : Team</p> <p>Task 6 : Verify Consistency Responsible : Team</p>	<p><b>Opportunity</b> <b>Identified</b></p> <ul style="list-style-type: none"> <li>An opportunity was identified</li> <li>A software-based solution could be used</li> <li>A stakeholder wants to make an investment in a better understanding potential users</li> <li>Identified other stakeholders who want to share in the opportunity.</li> </ul> <p>1/6</p> <p><b>Stakeholders</b> <b>Recognized</b></p> <ul style="list-style-type: none"> <li>Stakeholders identified.</li> <li>There is an agreement between the stakeholder groups to be represented.</li> <li>Responsibilities of stakeholder representatives defined</li> </ul> <p>1/6</p> <p><b>Stakeholders</b> <b>Represented</b></p> <ul style="list-style-type: none"> <li>The representatives of the interested parties were summoned.</li> <li>The representatives of the interested parties accept the responsibilities and authorized them.</li> <li>The collaborative approach was agreed upon.</li> <li>Responsibilities respect the way of working</li> </ul> <p>2/6</p> <p><b>Requirements</b> <b>Conceived</b></p> <ul style="list-style-type: none"> <li>The need for a new system is clear</li> <li>Users were identified</li> <li>Initial promoters were identified</li> </ul> <p>1/6</p> <p><b>Software System</b> <b>Architecture Selected</b></p> <ul style="list-style-type: none"> <li>Architecture selection criteria identified</li> <li>HW platforms selected</li> <li>Technologies selected</li> <li>System boundaries known</li> <li>Decisions on system organizations made</li> <li>Buy, build, reuse decisions made</li> <li>Key technical risks agreed to</li> </ul> <p>1/6</p> <p><b>Work</b> <b>Initiated</b></p> <ul style="list-style-type: none"> <li>The initiator of the work is known</li> <li>Work restrictions were clarified</li> <li>Sponsorship and funding model clarified</li> <li>Priority of work is clarified</li> </ul> <p>1/6</p> <p><b>Work</b> <b>Prepared</b></p> <ul style="list-style-type: none"> <li>The initiator of the work is known</li> <li>Work restrictions were clarified</li> <li>Sponsorship and funding model clarified</li> <li>Priority of work is clarified</li> </ul> <p>2/6</p> <p><b>Team</b> <b>Seeded</b></p> <ul style="list-style-type: none"> <li>The team's mission is clear</li> <li>The team knows how to grow to achieve its mission</li> <li>The required competencies were identified</li> <li>The size of the team has been determined</li> </ul> <p>1/5</p> <p><b>Team</b> <b>Formed</b></p> <ul style="list-style-type: none"> <li>Enough members recruited</li> <li>Roles understood</li> <li>How to work understood</li> <li>Members recruited</li> <li>Individual responsibilities accepted and aligned to competencies</li> <li>Members accepting work</li> <li>External collaborators identified</li> <li>Communication mechanisms defined</li> <li>Members commit to team</li> </ul> <p>2/5</p> <p><b>Way of Working</b> <b>Principles Established</b></p> <ul style="list-style-type: none"> <li>The principles and restrictions were established</li> <li>Principles and restrictions were compromised</li> <li>Practices and tools were agreed upon</li> <li>The context in which the team must operate was understood</li> </ul> <p>1/6</p> <p><b>Way of Working</b> <b>Foundation Established</b></p> <ul style="list-style-type: none"> <li>Key practices &amp; tools selected</li> <li>Practices needed to start work agreed</li> <li>Non-negotiable practices &amp; tools identified</li> <li>Gaps in capability understood</li> <li>Integrated way of working available</li> </ul> <p>2/6</p> <p><b>Stakeholders</b> <b>Involved</b></p> <ul style="list-style-type: none"> <li>Representatives assist the team</li> <li>Timely feedback and decisions provided</li> <li>Changes promptly communicated</li> </ul> <p>3/6</p> <p><b>Requirements</b> <b>Bounded</b></p> <ul style="list-style-type: none"> <li>Development stakeholders identified</li> <li>System purpose agreed</li> <li>System success clear</li> <li>Shared solution understanding exists</li> <li>Requirement's format agreed</li> <li>Requirements management in place</li> <li>Prioritization scheme clear</li> <li>Constraints identified &amp; considered</li> <li>Assumptions clear</li> </ul> <p>2/6</p> <p><b>Work</b> <b>Started</b></p> <ul style="list-style-type: none"> <li>Development started</li> <li>Progress monitored</li> <li>Definition of done in place</li> <li>Tasks being progressed</li> </ul> <p>3/6</p> <p><b>Team</b> <b>Collaborating</b></p> <ul style="list-style-type: none"> <li>Works as one unit</li> <li>Communication open and honest</li> <li>Focused on mission</li> <li>Members know each other</li> </ul> <p>3/5</p> <p><b>Way of Working</b> <b>In Use</b></p> <ul style="list-style-type: none"> <li>Practices &amp; tools in use</li> <li>Regularly inspected</li> <li>Feedback context</li> <li>Supported by team</li> <li>Feedback mechanisms in place</li> <li>Practices &amp; tools support collaboration</li> </ul> <p>3/6</p>

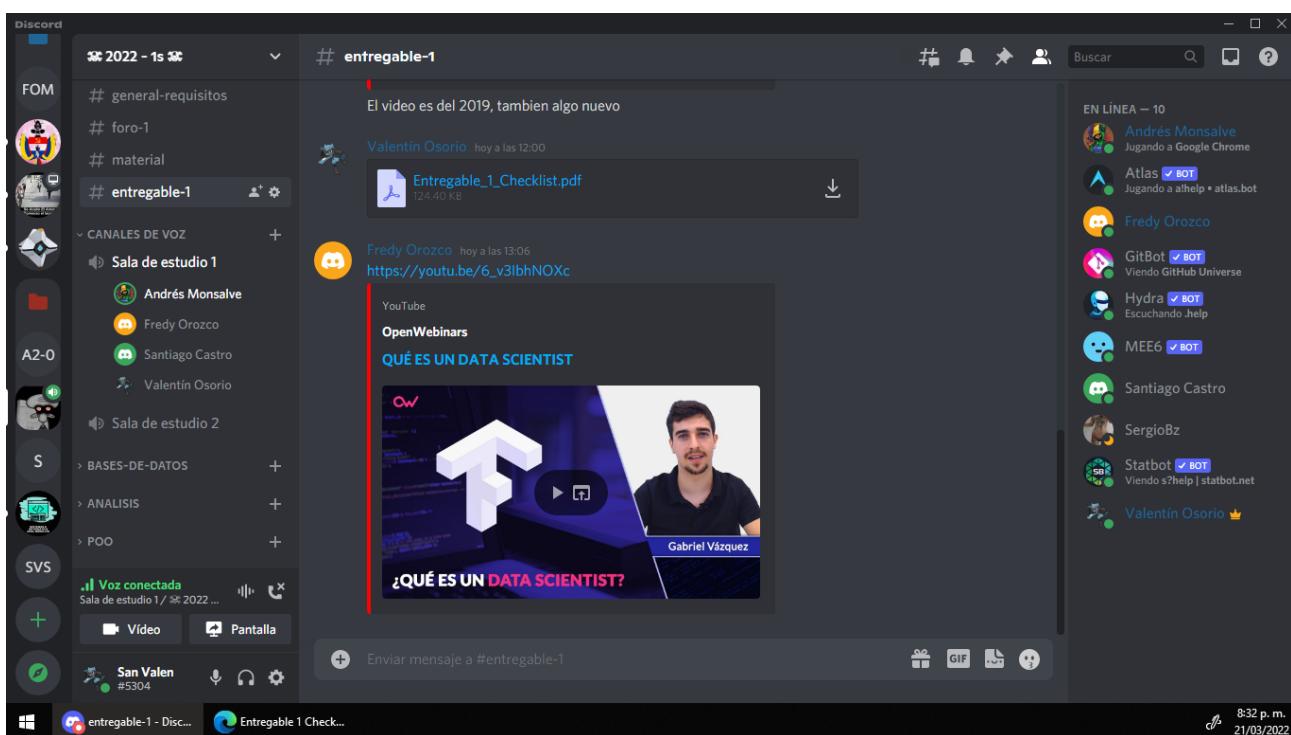
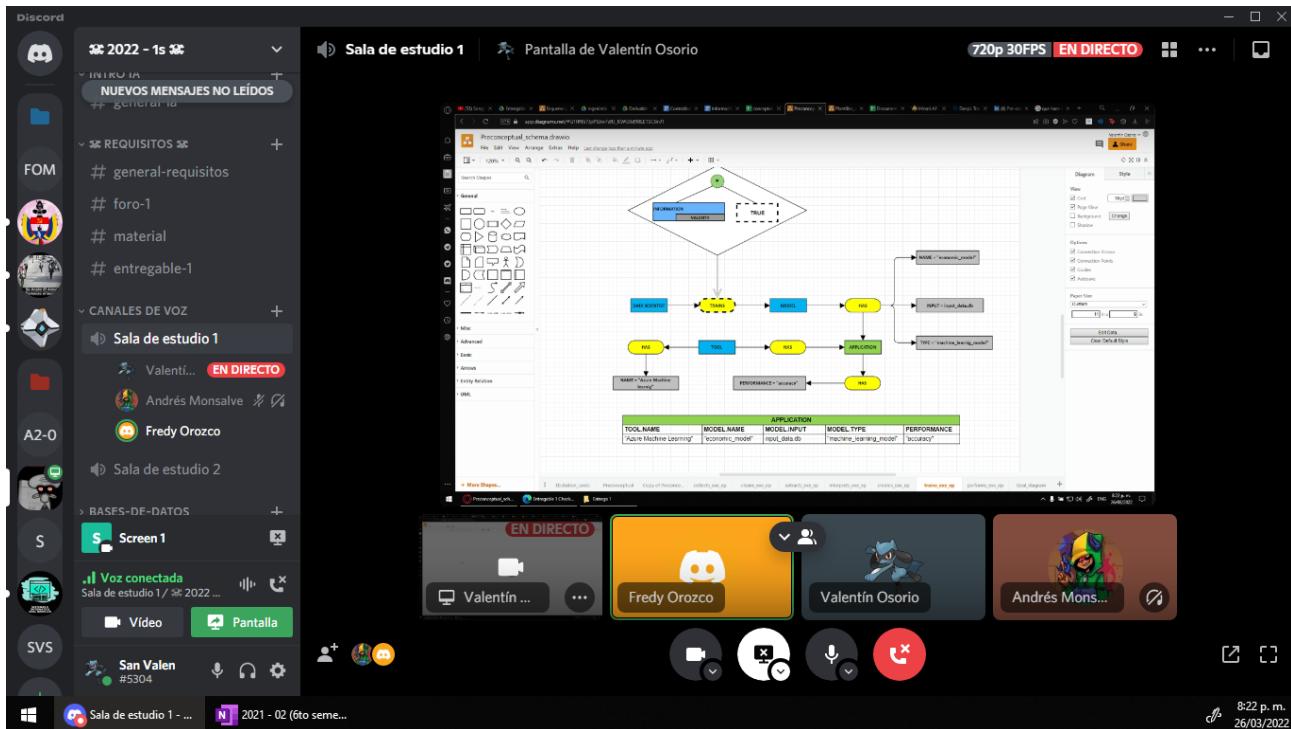
# Alphas Advance Report

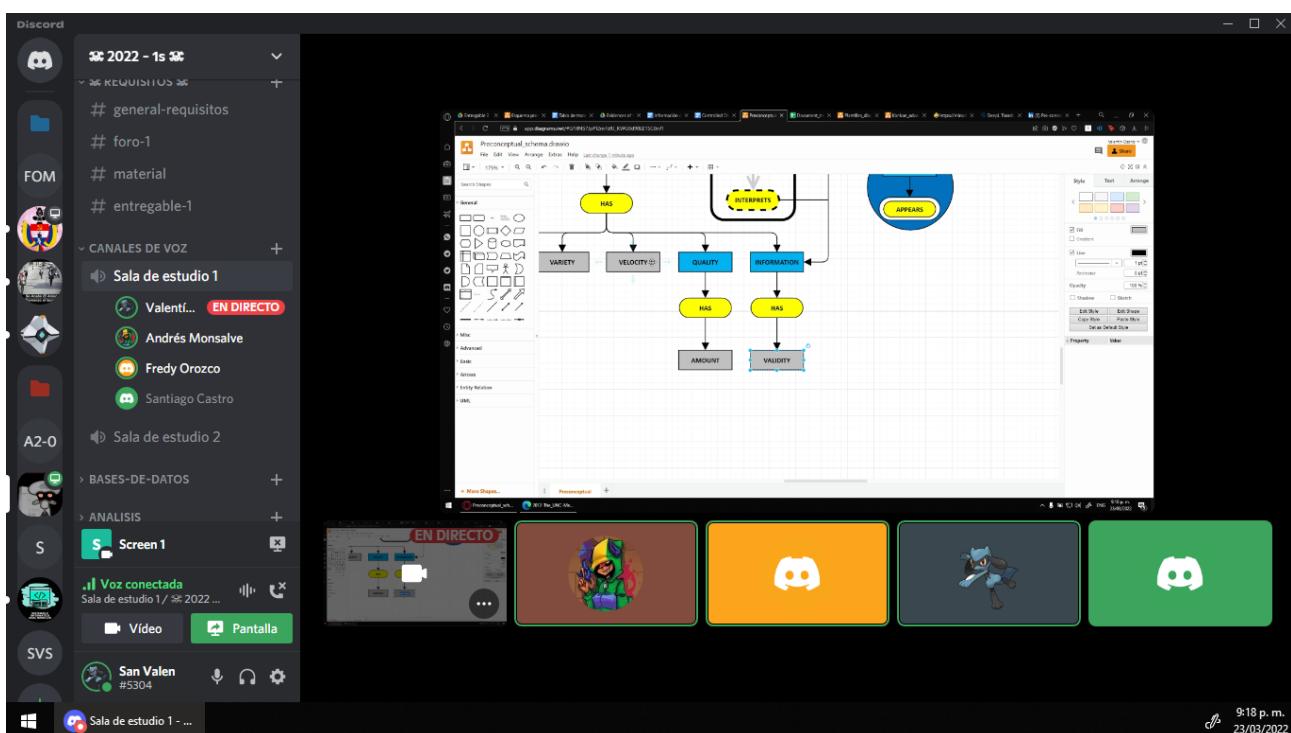
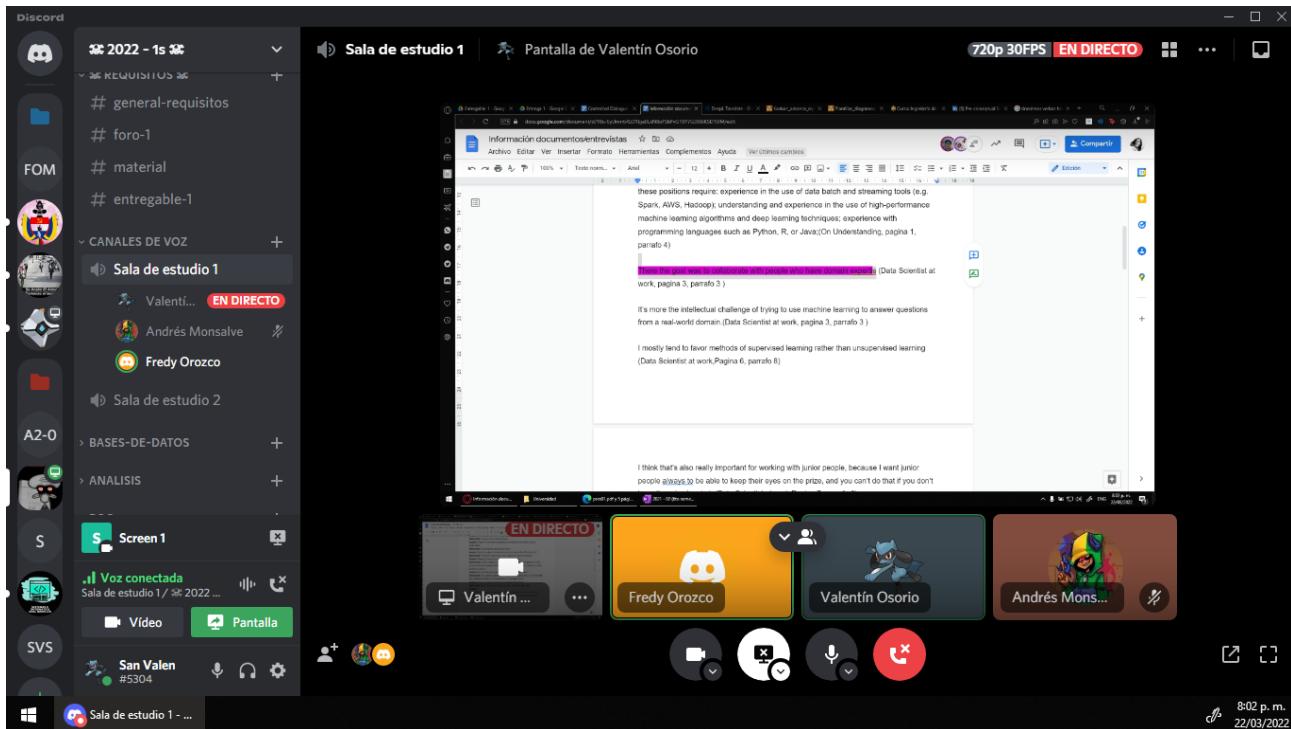
BACK

State	Summary of the achievement	Task	Date/Duration	Contents/Observations
 <b>Identified</b> <ul style="list-style-type: none"> <li>An opportunity was identified that a software-based solution could address.</li> <li>A team member wants to make an investment in better understanding of potential value.</li> <li>Identified other stakeholders who want to share in the opportunity.</li> </ul> <p>1/6</p>	The team got information about Data Scientists from several articles, book, and videos from YouTube	Gathering information and interviews on data scientists	Start date: 19/03/2022 5 hours	The team got the information from an article, a book, some videos and web pages
		Collecting and organizing the digital files	Start date: 20/03/2022 3 hours	The team organized the information they got in a document to have more summarized data
		Making the controlled dialog	Start date: 22/03/2022 15 hours and 20 minutes	Actors, attributes, functionalities, implications, events and conditionals are captured.
 <b>Represented</b> <ul style="list-style-type: none"> <li>The representatives of the interested parties were summoned.</li> <li>The representatives of the interested parties accept the responsibilities assigned to them.</li> <li>The collaborative approach was agreed upon.</li> <li>Representatives respect the way of working.</li> </ul> <p>2/6</p>	All the teams from the Requirements Engineer class created a shared drive where all the information of the roles is available for everyone	Filling the elicitation cards	Start Date: 26/03/2022 6 hours	4 actor cards, 9 object cards, 8 functionality cards
		Making a updating the organizational chart	Start Date: 15/03/2022 10 hours	There is a shared drive where all the teams organized their roles in the organizational charts
		Verify Consistency	Start Date: 26/03/2022 7 hours	Checked that the information obtained and what was being expressed by the team was congruent.
 <b>Conceived</b> <ul style="list-style-type: none"> <li>The need for a new system is clear.</li> <li>Users were identified.</li> </ul> <p>1/6</p>	The different artifacts from this deliverable were developed as intended.	Making, organizing and updating the document traceability table	Start date: 23/06/2022 10 hours	Document all the information contained in the preconceptual schema
		Verify consistency	Start date: 27/06/2022 3 hours	Checked that the information obtained and what was being expressed by the team was congruent.
		Making Goal Diagram	Start date: 26/06/2022 5 hours	8 goals, 7 requirements
		Making Cause-and-Effect Diagram	Start date: 27/06/2022 3 hours	Problems being consistent with elicitation cards and with preconceptual schema
		Making, organizing and updating the problems pre-conceptual schema	Start date: 28/06/2022 1 hours and 22 minutes	Problems being consistent with elicitation cards and cause-and-effect diagram
 <b>Architecture Selected</b> <ul style="list-style-type: none"> <li>Architecture selection criteria</li> <li>New platforms selected</li> <li>System boundary known</li> <li>Domain boundaries</li> <li>Organizations made</li> <li>Buy-build, reuse decisions made</li> <li>Key technical risks agreed to</li> </ul> <p>1/6</p>	The team decided about the technologies needed for the development of the software.	Making the domain model	Start date: 26/03/2022 1 Hours	13 classes
		Making, organizing and updating the pre-conceptual schema	Start date: 21/03/2022 22 Hours	8 dinamyc relationships, 8 implications, 1 event, 1 conditional
		Making, organizing and updating the executable pre-conceptual schemas	Start date: 26/06/2022 1 hour and 30 minutes	8 executable preconceptual diagrams were made.
 <b>Prepared</b> <ul style="list-style-type: none"> <li>The initiator of the work is known.</li> <li>Work restrictions were clarified.</li> <li>Sponsorship and funding model clarified.</li> <li>Priority of work is clarified.</li> </ul> <p>2/6</p>	In order to develop the deliverable we attended some classes to get feedback	Study class material	Start date: 08 /06/2022 4 hours	Studying the UNC Method and Documents related to the class
		Attend classes with the teacher	Start date: 22/06/2022 4 hours	The teacher was asked several questions about the preconceptual schema.
		Conduct meetings	Start date: 21/06/2022 10 hours	A meeting was held to work together and tasks were assigned.
 <b>Formed</b> <ul style="list-style-type: none"> <li>Enough members recruited.</li> <li>Roles understood.</li> <li>How to interact understood.</li> <li>Members introduced.</li> <li>Individual responsibilities accepted and aligned to competencies.</li> <li>Management work.</li> <li>External collaborators identified.</li> <li>Communication mechanisms defined.</li> <li>Members commit to team.</li> </ul> <p>2/5</p>	The team is composed of four members from the requirements engineer class: <ul style="list-style-type: none"> <li>- Diego Valentin Osorio Marín</li> <li>- Fredy Alberto Orozco Loaiza</li> <li>- Santiago Castro Tabares</li> <li>- Jaime Andrés Monsalve Ballesteros</li> </ul>	Conduct meetings	Start date: 21/06/2022 10 hours	A meeting was held to work together and tasks were assigned.
		Taking control of the kanban board	Start date: 27/06/2022 4 hours	Tasks were distributed at the beginning of each meeting.
		Taking the alpha advance report updated	Start date: 27/06/2022 3 hours and 30 minutes	The progress of each activity was reviewed at each meeting.
		Elaboration of KANBAN Board	Start date: 19/06/2022 3 hours	21 Tasks were made
 <b>Foundation Established</b> <ul style="list-style-type: none"> <li>Key practices &amp; tools selected.</li> <li>Practices needed to start work agreed.</li> <li>Non-negotiable practices &amp; tools identified.</li> <li>Key capability understood.</li> <li>Integrated way of working available.</li> </ul> <p>2/6</p>	We try to do a daily meeting in order to socialize the progress and to get feedback	Conduct meetings	Start date: 19/06/2022 40 hours	The team met at least once a day.
		Taking some pictures to show evidence of the meetings	Start date: 19/06/2022 10 minutes	There are 5 images of evidence of virtual meetings, but we also meet face-to-face at least 4 times.
 <b>Involved</b> <ul style="list-style-type: none"> <li>Representatives assist the team.</li> <li>Timely feedback and decisions provided.</li> <li>Changes promptly communicated.</li> </ul> <p>3/6</p>	The stakeholder, through an interview and a template, gave us an interview and suggestions for the project.	Interview with the stakeholder	Start date: 06/04/2022 2 hours	The team asked the stakeholder several questions about the takeaway from the project in order to obtain suggestions and feedback.
		Filling stakeholder survey	Start date: 06/04/2022 30 minutes	With the data, feedback, suggestions and signature of the interested party, a stakeholder survey was carried out.
		Collection of Extra Digital files	Start date: 06/04/2022 3 hours	We collected extra information from several sources and one interview with a data scientist
 <b>Bounded</b> <ul style="list-style-type: none"> <li>Development stakeholders identified.</li> <li>System purpose agreed.</li> <li>System success clear.</li> <li>Shared solution understanding exists.</li> <li>Requirements formally agreed.</li> <li>Design documents in place.</li> <li>Requirements scheme clear.</li> <li>Constraints identified.</li> <li>Assumptions defined.</li> <li>Key technical risks agreed to.</li> </ul> <p>2/6</p>	Several work products were elaborated in order to define the problem analysis phase, taking into account goals, problems, and the process.	Corrections to the Software Context Delivery	Start date: 06/04/2022 20 hours	Controlled Dialogue, domain model, organizational chart, elicitation cards and executable pre-conceptual schemas were corrected consistently to the corrections made to the domain.
		Define process and sequence	Start date: 06/04/2022 7 hours	With all the information provided by the stakeholder and several articles and sources the team were capable of defining the process and sequence.
		Elaboration of Process Diagram	Start date: 01/05/2022 10 hours	With all the information provided by the stakeholder and several articles and sources the team were capable of defining the process and sequence
		Elaboration of Event Interaction Graph	Start date: 02/05/2022 4 hours	Events were added
		Elaboration of Process Diagram: Explanatory Table	Start date: 02/05/2022 7 hours	Totally mapped to the Process diagram
		Define constraints to the process	Start date: 27/04/2022 8 hours	14 constraints
		Elaboration of Data Dictionary	Start date: 03/05/2022 3 hours	Totally mapped to pre-conceptual schema
		Elaboration of Business rules	Start date: 01/05/2022 3 hours	Totally mapped to pre-conceptual schema
		Elaboration of Goal Diagram	Start date: 30/04/2022 2 hours	11 goals, and 7 requirements
		Elaboration of Cause-and-effect Diagram	Start date: 01/05/2022 1 hour and 30 minutes	1 main problem, 2 sub problems and 5 causes
		Elicit goals	Start date: 06/04/2022 5 hours	11 goals
		Elicit problems	Start date: 06/04/2022 5 hours	8 problems
		Elaboration of Pre-conceptual Schema	Start date: 21/04/2022 30 hours	12 dinamyc relationships, 12 conditionals, 4 events
		Elaboration of Pre-conceptual Schema Traceability Table	Start date: 01/05/2022 20 hours	Document all the information contained in the preconceptual schema
		Elaboration of Problem-based Pre-conceptual Schema	Start date: 28/04/2022 2 hours	8 problems
		Elaboration of Problem-based Pre-conceptual Schema Traceability Table	Start date: 01/05/2022 10 hours	Document all the information contained in the problems preconceptual schema
		Elaboration of Percentage assignment to goals	Start date: 02/05/2022 4 hours	Percentages were assigned according to the instructions learned in class
		Elaboration of Percentage assignment to problems	Start date: 02/05/2022 3 hours	Percentages were assigned according to the instructions learned in class
		Verify Consistency	Start date: 04/05/2022 7 hours	All the members of the team are responsible of verifying consistency
 <b>Started</b> <ul style="list-style-type: none"> <li>Development started.</li> <li>Progress monitored.</li> <li>Delivery due dates in place.</li> <li>Tasks being progressed.</li> </ul> <p>3/6</p>	The entire team was constantly evaluated in order to achieve work objectives and produce work products.	Elaboration of Method assesment template	Start date: 27/04/2022 2 hours	Some suggestions were found that can be implemented in the UNC Method.
		Assign tasks	Start date: 21/04/2022 2 hours	After having the Preconceptual schema done, the team leader assigned different task in order to get progress
		Report progress of the work	Start date: 30/04/2022 2 hours	Each member of the team reported their progress each day
 <b>Collaborating</b> <ul style="list-style-type: none"> <li>Works as one unit.</li> <li>Communication open and honest.</li> <li>Members know each other.</li> </ul> <p>3/5</p>	All team members work proactively towards the same common goal.	Continuous Elaboration of Alphas Advance Report	Start date: 21/04/2022 3 hours	New states and activities were added
		Attend class in order to get feedback from the teacher.	Start date: 26/04/2022 5 hours	We got feedback from the professor several times in class
		Assess the team	Start date: 26/04/2022 1 hours	The team leader evaluated the team constantly
		Elaboration of Team Charter	Start date: 29/04/2022 2 hours	Each member of the team had different responsibilities
 <b>In Use</b> <ul style="list-style-type: none"> <li>Practices &amp; tools in use.</li> <li>Regularly inspected.</li> <li>Adaptations made.</li> <li>Supported by team.</li> <li>Feedback loops in place.</li> <li>Practices &amp; tools support collaboration.</li> </ul> <p>3/6</p>	Feedback was obtained from the professor to continue the work and achieve the objectives.	Continuous elaboration of Kanban Board	Start date: 21/04/2022 2 hours and 30 minutes	New activities and states were added. Some corrections were developed too.
		Assess the way of working	Start date: 23/04/2022 30 minutes	The team leader evaluated the way of working constantly
		Gathering evidence of the meetings	Start date: 21/04/2022 10 minutes	Screenshots of the virtual meeting and some presential meetings too.
		Attend class in order to get feedback from the professor	Start date: 26/04/2022 6 hours	We got feedback from the professor several times in class

# EVIDENCE OF THE MEETINGS < BACK







**Discord**

# 2022 - 1s

# general-requisitos

Fredy Orozco 02/04/2022 JFrameForn

Valentin Osorio 02/04/2022 <https://www.mtholyoke.edu/courses/quenell/s2003/ma139/js/count.html>

11 de abril de 2022

Valentin Osorio 11/04/2022 <https://bibliotecas.unal.edu.co/>

28 de abril de 2022

Andrés Monsalve 28/04/2022

Pasos y roles en el flujo de trabajo de data science (Design Patterns in Machine Learning).

EN LÍNEA – 10

- Andrés Monsalve Jugando a Google Chrome
- Atlas ✅ BOT Viendo /setup | atlas.bot
- Fredy Orozco
- GitBot ✅ BOT Viendo GitHub Universe
- Hydra ✅ BOT Escuchando help
- MEE6 ✅ BOT
- Santiago Castro
- SergioBz
- Statbot ✅ BOT
- Valentin Osorio

Voz conectada Sala de estudio 1 / 2022 ...

Video Pantalla

Enviar mensaje a #general-requisitos

9:27 p. m.  
1/05/2022

**Discord**

# 2022 - 1s

# entregarable-1

# foro-2

# entregarable-2

CANALES DE VOZ

Sala de estudio 1

Valenti... EN DIRECTO

Andrés Monsalve

Santiago Castro

Sala de estudio 2

BASES-DE-DATOS

ANALISIS

POO

Voz conectada Sala de estudio 1 / 2022 ...

Video Pantalla

Screen1

pres01.pdf y 6 pági... Sala de estudio 1 - ...

9:01 p. m.  
2/05/2022

**CAUSE-AND-EFFECT DIAGRAM**

BP2: Summary does not have validity

C4 Application does not have version control system

C5 Data scientist does not have experience

C3 Stakeholder provides data wrongly

C1 Data scientist evaluates user requirements hardly

Solution has low performance

**# entregable-2**

Aquí empieza el canal #entregable-2.

[Editar canal](#)

24 de abril de 2022

Fredy Orozco 24/04/2022

para sacar el porcentaje de los diagramas. Lo primero es calcular los porcentajes de diagramas de Objetivos, Una vez calculados. Procedemos a coger los procesos y las causas raíces. Le asociamos el porcentaje de los objetivos, Y en cada casilla ponemos la suma y luego sumamos la columna y por ultimo sacamos el porcentaje (Sumamos todo y dividimos las columnas por el) Este porcentaje va hacer el que pertenece a cada rama (editado)

**PERCENTAGE ASSIGNMENT TO PROBLEMS**

	C1	C3	C5	C6	C9	C10
P1		G4, R4	G1, G4, R4			
P2			G4, R4	G4, R4		
P3		G4, G6, R4		G4, G6, R4		
P4	G3, G4, R1					
P5				G2, G5, R2		
P6					G4, R7	
P7					G3, G5, R7	
P9					G2, G5, R7	G4, R7
P10		G4, R7, R9				
P12					G4, R6	
P13	G4, R10, R7, R6					
P14	G4, R8, R7, R6					

Enviar mensaje a #entregable-2

9:13 p.m.  
3/05/2022

**# entregable-1**

**NUEVOS MENSAJES NO LEÍDOS**

**# entregable-2**

**Sala de estudio 1**

Valenti... EN DIRECTO

Andrés Monsalve

Fredy Orozco

MEE6

Santiago Castro

**Sala de estudio 2**

**BASES-DE-DATOS**

**ANALISIS**

Screen1

Voz conectada

Envío de audio

Video Pantalla

San Valen #5304

Enviar mensaje a #entregable-2

Checklist Deliverable 2

Checklist Deliverable 2

**Deliverable**

- 1. Deliverable 1 corrections
  - 1.1. Evolution-cause (Domain model)
  - 1.2. (Organizational chart)
  - 1.3. (Executable PES)
  - 1.4. (Process flowchart)
- 2. Data alignment corrections
  - 2.1. SMART
  - 2.2. Stakeholder survey
  - 2.3. Team chart
  - 2.4. Process diagram (Santiago, Valente)
  - 2.5. Process diagram explanation table (Valente)
  - 2.6. Event interaction graph (Santiago, Valente)
  - 2.7. Cause-and-effect diagram
  - 2.8. Business rules (Fredy)
  - 2.9. Pre-conceptual schema with achievement rules and traceability table (Andrés)
  - 2.10. Goal Diagram
  - 2.11. Pre-conceptual pre-conceptual scheme and traceability table (Andrés)
  - 2.12. Cause-and-effect diagram
  - 2.13. Percentage assignment to goals (Andrés)
  - 2.14. Percentage assignment to problems (Andrés)
  - 2.15. Kanban board

**Consistency**

- All the processes have, at least, one goal (remember: requirements are different to goals). 6, 7, 8, 12, 13, 14.
- The suggestions for improving the former deliverable are developed. 1, 2, 6.
- The goals, requirements, and expectations from the goal diagram are consistently written according to the pre-conceptual schema added with achievement relationships. 11, 12.
- All the nouns and noun phrases in the cause-and-effect diagram are included as concepts in the pre-conceptual schema. 11, 13, 14.
- The pre-conceptual schema is completely mapped to the process diagram. 6, 7, 8, 11.
- All the conditions from the pre-conceptual schema are included in the business rules. 10, 11.
- All the concepts before "has" relationships are mapped as storage. 6, 7, 9, 11.
- All the causal relationships are mapped to events of the process diagram. 6, 7, 8, 11.
- All the dynamic relationships are mapped to processes of the process diagram. 6, 7, 11.
- The suggestions for improving the former deliverable are developed. 1, 2, 7, 8.

10:01 p.m.  
4/05/2022

## **DIGITAL FILES < BACK**

<https://drive.google.com/drive/u/0/folders/18KEbRElaHxSbiCZtfsCLBqiPhlrmScvo>