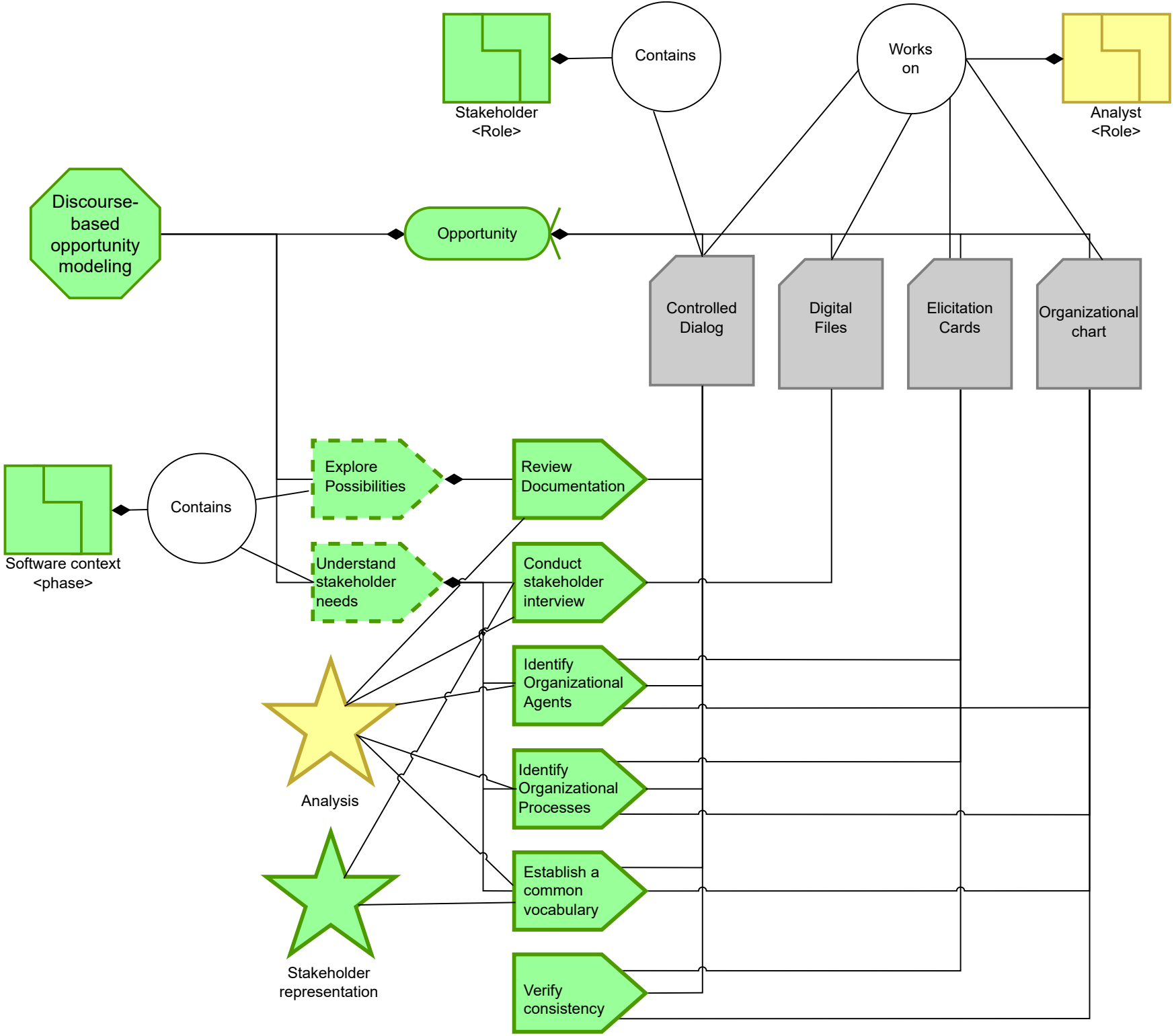
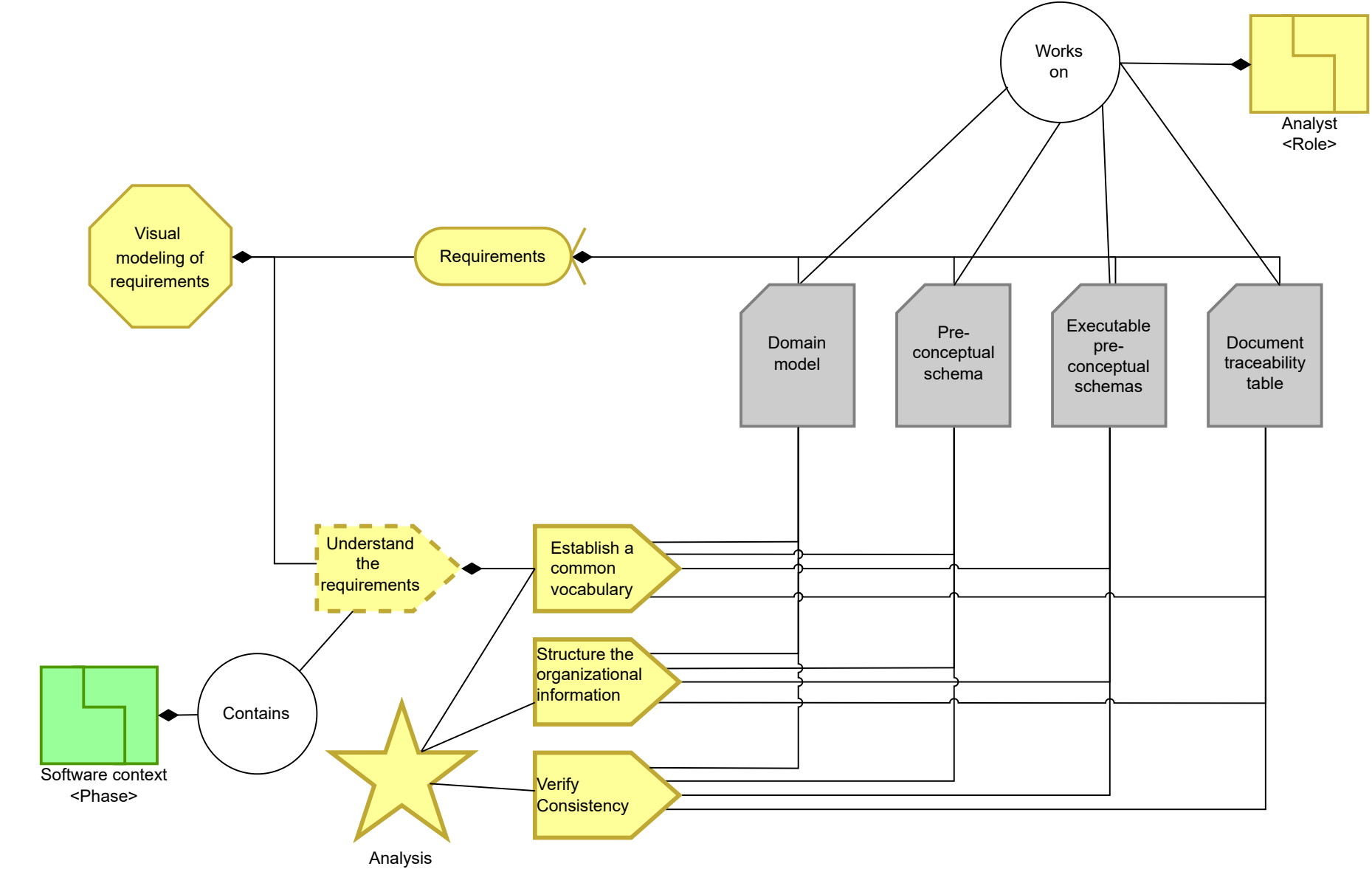
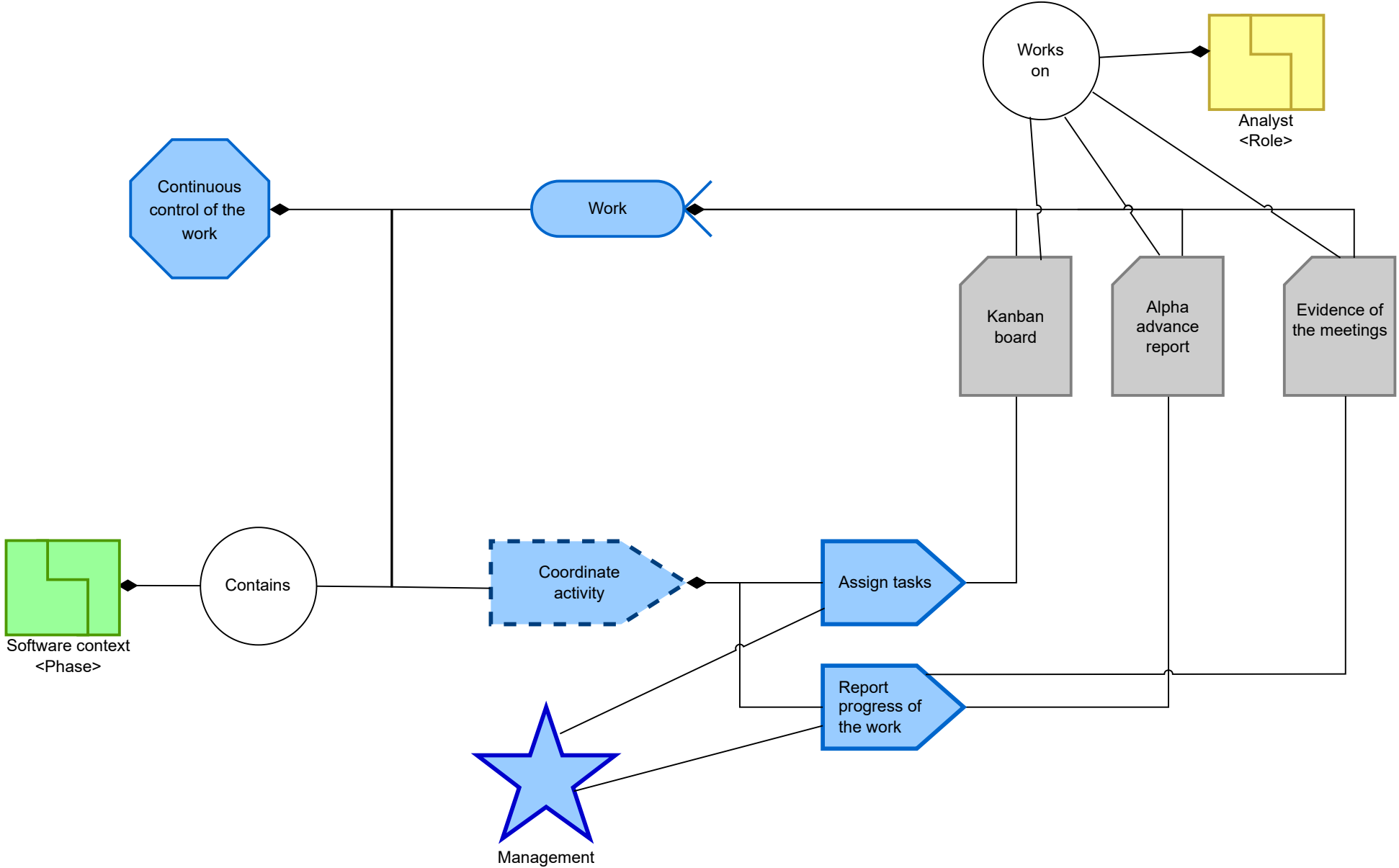
DELIVERABLE 1: SOFTWARE CONTEXT - DATA SCIENTIST

DIEGO VALENTÍN OSORIO MARÍN
JAIME ANDRÉS MONSALVE BALLESTEROS
SANTIAGO CASTRO TABARES
FREDY ALBERTO LOAIZA OROZCO

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


Stakeholder
<Role>
Contains
Works
on
Analyst
<Role>
Discourse-
based
opportunity
modeling
Opportunity
Controlled
Dialog
Digital
Files
Elicitation
Cards
Organizational
chart
Software context
<phase>
Contains
Explore
Possibilities
Understand
stakeholder
needs
Review
Documentation
Conduct
stakeholder
interview
Identify
Organizational
Agents
Identify
Organizational
Processes
Establish a
common
vocabulary
Verify
consistency
Analysis
Stakeholder
representation


Works on
Analyst
<Role>
Visual modeling of requirements
Requirements
Domain model
Pre-conceptual schema
Executable pre-conceptual schemas
Document traceability table
Understand the requirements
Establish a common vocabulary
Structure the organizational information
Verify Consistency
Contains
Software context
<Phase>
Analysis


Works on
Analyst <Role>
Continuous control of the work
Work
Kanban board
Alpha advance report
Evidence of the meetings
Software context <Phase>
Contains
Coordinate activity
Assign tasks
Report progress of the work
Management

CONTROLLED DIALOGUE

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 employees, clients, data scientists and business analysts.

Analyst: Who can play the role of an employee?
Stakeholder: A data scientist and business analyst.

Analyst: Would you please mention some characteristics of employees?
Stakeholder: They have name, id, job title, experience and e-mail.

Analyst: Would you please mention some characteristics of clients?
Stakeholder: They have name, id and an e-mail.

Analyst: Would you please mention some characteristics of business analysts?
Stakeholder: They have skill.

Analyst: Would you please mention some characteristics of data scientists?
Stakeholder: They have background, style and tool.

Analyst: Please list the main functions of the business analysts.
Stakeholder: He/she raises problems; collects data.

Analyst: Please list the main functions of the data scientists.
Stakeholder: The data scientist cleans data; extracts and interprets information; creates models; trains applications; performs solutions.

Analyst: Are these functions gathered in some sort of responsibility?
Stakeholder: Yes, cleans data, extracts, and interprets information are gathered as “analyzes data”; creates models and trains applications are gathered as “manages model”.

Analyst: Would you please mention some features of the data?
Stakeholder: Each data has value, type, format, volume, variety, velocity, quality and information.

Analyst: Would you please mention some features of company?
Stakeholder: It has description, name, data, and problem.

Analyst: Would you please mention some features of model?
Stakeholder: It has name, input, type, solution and application.

Analyst: Would you please mention some features of problem?
Stakeholder: It has description and solution.

Analyst: Would you please mention some features of tool?
Stakeholder: It has name and application.

Analyst: : Which of the mentioned features have features themselves?
Stakeholder: Quality has amount; information has description, validity and new value; solution has description, result and validity; application has performance.

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 amount; 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.

Analyst: What does the business analyst need in order to accomplish the raising of a problem?
Stakeholder: He/she only needs the appearance of a problem.

Analyst: What does the data scientist need in order to accomplish the creation of a model?
Stakeholder: He/she only needs that information validity = True.

Analyst: Would you please establish some sort of sequence in the functions and responsibilities you have just described?
Stakeholder: First the business analyst collects the data; then, in the data analysis, the data scientist cleans the data, then extracts the information and then interprets that information; In model managing, the data scientist creates the model and then he/she trains the application; then the data scientist performs a solution.

Analyst: What are the goals and problems associated with the function “creates model”?
Stakeholder: The only goal is “making the model”. This function has no problems.

Analyst: What are the goals and problems associated with the function “trains application”?
Stakeholder: The goals are “making the model” and “developing application”. This function has no problems.

Analyst: What are the goals and problems associated with the function “collects data”?
Stakeholder: The only goal is “fostering the data”. The problems are “there is low quality of data” and “lack of clear problems”.

Analyst: What are the goals and problems associated with the function “cleans data”?
Stakeholder: The only goal is “fostering the amount of quality of the data”. The problems are “data is cleaned slowly” and “there is low quality of data”.

Analyst: What are the goals and problems associated with the function “performs solution”?
Stakeholder: The only goal is “achieving that solution has validity”. The only problem is “solution does not have validity”.

Analyst: What are the goals and problems associated with the function “interprets information”?
Stakeholder: The only goal is “achieving that information has validity”. The only problem is “there is no validity from information”.

Analyst: What are the goals and problems associated with the function “extracts information”?
Stakeholder: The only goal is “fostering the information”. The only problem is “information extraction is done hardly”.

Analyst: What are the goals and problems associated with the function “raises problem”?
Stakeholder: The only goal is “achieving that problem has solution”. The problems are “problem does not have clarity” and “problem does not have solution”.

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.

BACK

ELICITATION CARDS

ACTOR	
EMPLOYEE	
FEATURES	NAME, ID, JOB TITLE, EXPERIENCE, E-MAIL

ANNOTATIONS

ACTOR	
CLIENT	
FEATURES	NAME, ID, E-MAIL

ANNOTATIONS

ACTOR	
DATA SCIENTIST	
FEATURES	BACKGROUND,STYLE

ANNOTATIONS
DATA SCIENTIST can be EMPLOYEE, also DATA SCIENTIST is related to TOOL

ACTOR	
BUSINESS ANALYST	
FEATURES	SKILL

ANNOTATIONS
BUSINESS ANALYST can be EMPLOYEE

OBJECT	
DATA	
FEATURES	VALUE, TYPE, FORMAT, VOLUME, VARIETY, VELOCITY

ANNOTATIONS
DATA is related to QUALITY and INFORMATION

OBJECT	
QUALITY	
FEATURES	AMOUNT

ANNOTATIONS

OBJECT	
INFORMATION	
FEATURES	DESCRIPTION, VALIDITY, NEW VALUE

ANNOTATIONS

OBJECT	
COMPANY	
FEATURES	NAME, DESCRIPTION

ANNOTATIONS
COMPANY is related to EMPLOYEE, CLIENT, DATA and PROBLEM

OBJECT	
MODEL	
FEATURES	NAME, INPUT,TYPE,

ANNOTATIONS
MODEL is related to SOLUTION and APPLICATION

OBJECT	
SOLUTION	
FEATURES	DESCRIPTION,RESULT, VALIDITY

ANNOTATIONS

OBJECT	
PROBLEM	
FEATURES	DESCRIPTION

ANNOTATIONS
PROBLEM is related to SOLUTION

OBJECT	
APLICATION	
FEATURES	PERFORMANCE

ANNOTATIONS

OBJECT	
TOOL	
FEATURES	NAME

ANNOTATIONS
TOOL is related to APPLICATION

FUNCTION	
CREATES	
ACTOR	DATA SCIENTIST
OBJECT	MODEL
CONSTRAINT	
VALIDITY ="TRUE"	

GOAL	PROBLEM
MAKING the MODEL	

FUNCTION	
TRAINS	
ACTOR	DATA SCIENTIST
OBJECT	APPLICATION
CONSTRAINT	
DATA SCIENTIST TRAINS the MODEL. after DATA SCIENTIST CREATES the MODEL	

GOAL	PROBLEM
- MAKING the MODEL - DEVELOPING APPLICATION	

FUNCTION	
COLLECTS	
ACTOR	BUSINESS ANALYST
OBJECT	DATA
CONSTRAINT	
BUSINESS ANALYST COLLECT the DATA. after BUSINESS ANALYST RAISE PROBLEM	

GOAL	PROBLEM
FOSTERING the DATA	- There is low QUALITY of DATA - Lack of clear PROBLEMS

FUNCTION	
CLEANS	
ACTOR	DATA SCIENTIST
OBJECT	DATA
CONSTRAINT	
DATA SCIENTIST CLEAN the DATA after BUSINESS ANALYST collec thet data	

GOAL	PROBLEM
FOSTERING the AMOUNT of QUALITY of the DATA	- DATA is CLEANED SLOWLY - There is low QUALITY of DATA

FUNCTION	
PERFORMS	
ACTOR	DATA SCIENTIST
OBJECT	SOLUTION
CONSTRAINT	
DATA SCIENTIST PERFORMS the DATA after DATA SCIENTIST TRAINS the model	

GOAL	PROBLEM
ACHIEVING that SOLUTION has VALIDITY	SOLUTION does not have VALIDITY

FUNCTION	
INTERPRETS	
ACTOR	DATA SCIENTIST
OBJECT	INFOMATION
CONSTRAINT	
DATA SCIENTIST INTERPRETS the DATA after DATA SCIENTIST EXTRACTS DATA	

GOAL	PROBLEM
ACHIEVING that INFORMATION has VALIDITY	There is no VALIDITY from INFORMATION

FUNCTION	
EXTRACTS	
ACTOR	DATA SCIENTIST
OBJECT	INFORMATION
CONSTRAINT	
DATA SCIENTIST EXTRACTS the DATA after DATA SCIENTIST CLEANS the DATA	

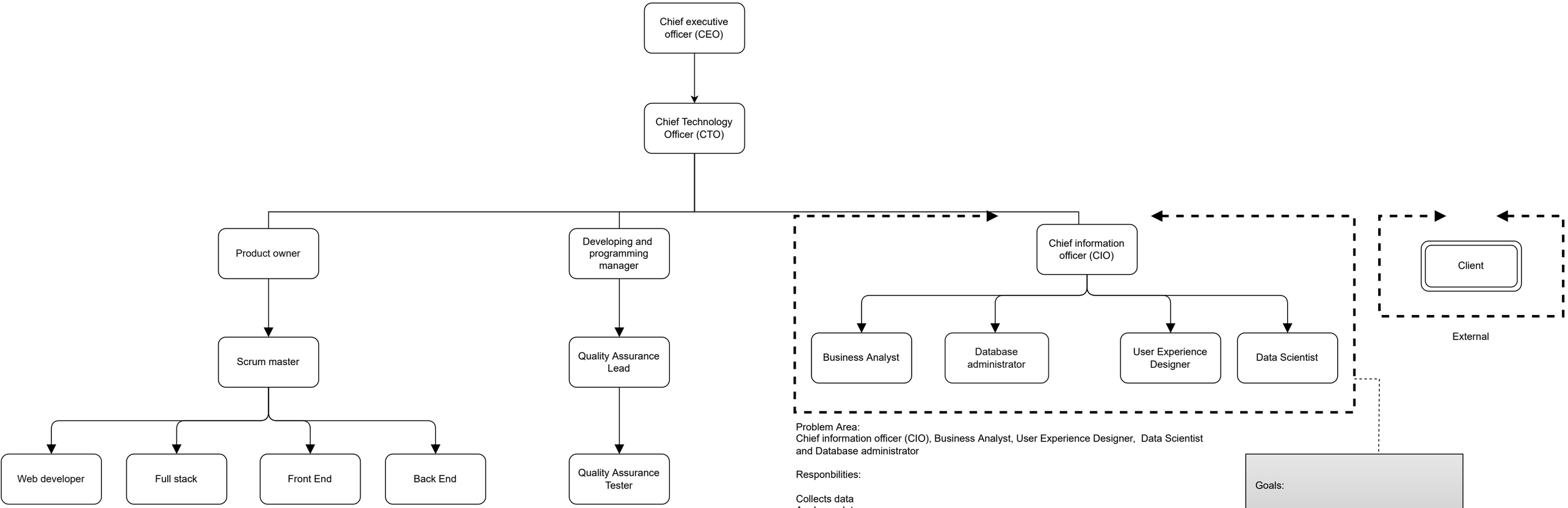
GOAL	PROBLEM
FOSTERING the INFORMATION	INFORMATION extraction is done hardly

FUNCTION	
RAISES	
ACTOR	BUSINESS ANALYST
OBJECT	PROBLEM
CONSTRAINT	
PROBLEM APPERS	

GOAL	PROBLEM
ACHIEVING THAT PROBLEM HAS SOLUTION	- PROBLEM does not have CLARITY - PROBLEM does not have SOLUTION

ORGANIZATIONAL CHART

BACK



Problem Area:
Chief information officer (CIO), Business Analyst, User Experience Designer, Data Scientist and Database administrator

Responbilities:

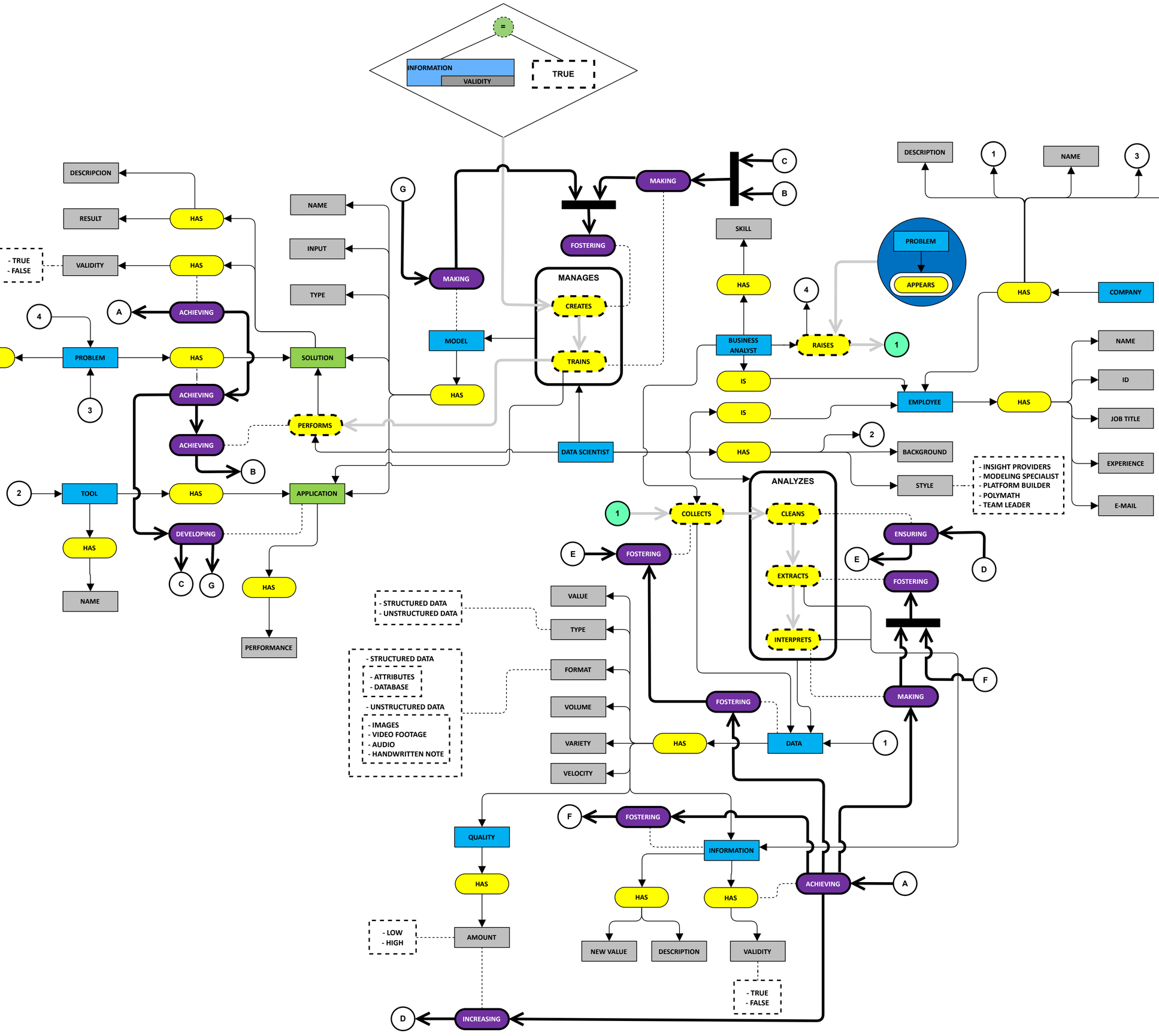
Collects data
Analyzes data
Manages model

Goals:

Making the model
Fostering the data
Fostering the amount of quality of the data
Achieving that solution has validity
Achieving that information has validity
Fostering the information
Developing application
Achieving that problem has solution

PRE-CONCEPTUAL SCHEMA

BACK

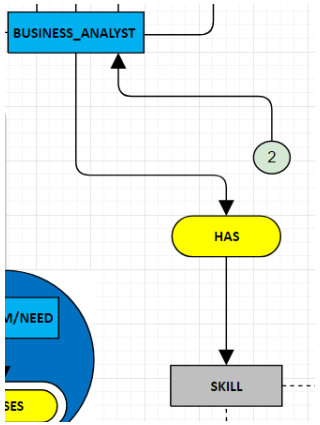


BACK

Document Traceability Table (Preconceptual Schema)

Original sound/Image/Text	Source	Location	Element	Kind of element	Observations
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
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 infered that the Data Scientist is employee of a company, in this case "The New York Times"
these positions require: experience in the use of data batch and streaming tools	Text	On Understanding Data Scientists, Page 1, Paragraph 4	Data Scientist has tool	Structural triad	- We interpreted "these position" as roles related to Data Scietist profession, so then he has tools
companies interface with their customers and clients	Text	Data Scientists at Work, Page XVI, Paragraph 3	Company has client	Structural triad	
The New York Times is also centuries old. It's a 163-year-old company	Text	Data Scientists at Work, Page 2, Paragraph 4	Company has name	Structural triad	- We interpret that "The New York Times" is related to the name of the company, son then company has name

The New York Times is also centuries old. It's a 163-year-old company, and I think it also stands for a set of values that I strongly believe in and is also very strongly associated with New York, which I like very much. When I think of The New York Times, I think of the sentiment expressed by Thomas Jefferson that if you could choose between a functioning democracy and a dysfunctional press, or a functioning press and a dysfunctional democracy, he would rather have the functioning press. You need a functioning press and a functioning journalistic culture to foster and ensure the survival of democracy	Text	Data Scientists at Work, Page 2, Paragraph 4	Company has description	Structural triad	- We interpret that what Chris Wiggins (Data Scientist) is taking about in this fragment is the description of "The New York Times" which is a company
as well as the early employees of the company	Text	Data Scientists at Work, Page 191, Paragraph 9	Company has employee	Structural triad	-We can interpret from the context of the article that the company can have employees like Data Scientist or Business Analyst
make sense of the abundant data that The New York Times gathers.	Text	Data Scientists at Work, Page 4, Paragraph 2	Company has data	Structural triad	- We inferred that the company, in this case "The New York Times", has data because it is said that there is abundant data that they gather.
Why is search relevance an important problem to tackle for LinkedIn?	Text	Data Scientists at Work, Page 86, Paragraph 7	Company has problem	Structural triad	- We interpret "problem to tackle for LInkedIn" as LinkedIn is a company that has a problem which has to be solved
but they are wrapped inside of a software product that clients interact with directly	Text	Data Scientists at Work, Page 242, Paragraph 2	Client has ID Client has email Client has Name	Structural triad	- As clients can interact with something we interpret them as actors, so it is implicit that they have a name, id, and email
as well as the early employees of the company	Text	Data Scientists at Work, Page 191, Paragraph 9	Employee has name Employee has ID Employee has experience	Structural triad	- We infered that name, ID, an experience are features that a company needs in an employee
Claudia Perlich is the Chief Scientist at Dstillery	Text	Data Scientists at Work, Page 150, Paragraph 1	Employee has job title		- We interpret that "Chief Scientist" is her job title at that company

	Image/Diagram	Partners class PS: Business Analyst https://drive.google.com/file/d/1REFHDtXkNuD_7WcOZBCgdtqi1YVMg-0O/view?usp=sharing	Business Analyst has skill	Structural triad	
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.
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
I believe that access to quality information and information relevant to our problems is the greatest challenge.	Text	On Understanding Data Scientists, Page 4, Paragraph 3	Data has information	Structural triad	We interpret from the context of the article that "information relevant to our problems" ca be interpreted as the Data carries with it the information that will be used
"Buscar los features o caracteristicas de los datos que sabemos que son importantes"	Video	https://youtu.be/Bl2sBiVdZHs, minute 1:37	Data has value	Structural triad	We can interpret that "features o caracteristicas" ca be refer to value that the Data carries with it
When dealing with data we have to take into consideration three important characteristics: volume, variety, and velocity	Text	On Understanding Data Scientists, page 1, paragraph 3	Data has Volume Data has Variety Data has Velocity	Structural triad	- We interpret from the context of the article that "we have to take into consideration three important characteristics" refers to Data have this three characteristics
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 1, paragraph 4	Quality has amount	Structural triad	- We infer that "quality of the data to be assessed" can be interpreted as data have different levels of quality

If we want to use information on whether this is a kid's or an adult's profile to help optimize people's experience, we'd rather ask people to tell us in their profile that it's a kid's profile.	Text	Data Scientists at Work, Page 26, Paragraph 6	Information has description	Structural triad	We ca infer from "information on whether this is a kid´s or an adult´s" that the information we are interested in includes a description
I believe that access to quality information and information relevant to our problems is the greatest challenge.	Text	On Understanding Data Scientists, Page 4, Paragraph 3	Information has validity	Structural triad	- We can understood from the context of the article that "quality information and information relevant" can be interpreted as the information must be validated for it´s analysis so the information has a validation
Data cleaning process In terms of data pre-processing in order to increase data quality.	Text	On Understanding Data Scientists, Page 3, Paragraph 3	Information has new value	Structural triad	- We can interpret from the context of the article that "data pre-processing" it is the information that will be use - We can interpret from the context of the article that the information that was extracted may be changed to further increase it´s quality so the information must be has new values
the predictive engines and data infrastructure required to effectively personalize recommendations	Text	Data Scientists at Work, Page 19, Paragraph 1	Model has name	Structural triad	-We interpret from the context of the interview that "predictive engines" can be interpret as model -We interpret from the context of the interview that "personalize recommendations" can be refer to a model is created to predict a specific content so a model must have a name to identify its purpose
"A estos modelos solo le vamos a alimentar una parte del total de nuestro dataset"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2: 25	Model has input	Structural triad	-We can interpret from the context of the video that "le vamos a alimentar" refers to the fact that the model is fed with data so the model has a input - We interpret that "alimentar" refers to there are a input
"Vamos a poder proceder a elegir un modelo matemático este modelo puede ser un modelo algebraico o estadístico"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2: 09	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 led the product data science team, a group of data scientists focused on creating innovative solutions to improve LinkedIn's products and create new ones.	Text	Data Scientists at Work, Page 85, Paragraph 4	Model has solution	Structural triad	- We can interpret from the context of the interview that "creating innovative solutions to improve" refers to a model has solution
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 to a model ca be applied to a different levels of prediction for a specific problem

We studied both versions of the model for both accuracy [AUC] and performance on A/B test results.	Text	Data Scientists at Work, Page 31, Paragraph 2	Aplication has performance	Structural triad	We interpret from the context of the article that "accuracy and performance" that a model can have different levels of performance
"Obtendremos una calificacion y esta nos va a indicar que tan acertadas son nuestras predicciones"	Text	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2: 38	Solution has result Solution has description	Structural triad	- We interpret from the context of the video that "predicciones" refers to the solution of the model - We interpret from the context of the video that "Obtendremos una calificación" refers to the fact that the solution brings with it a result so the solution has a result. - We interpret from the context of the video that "esta nos va indicar que tan acertadas son nuestras predicciones" refers to the solution has a description
we deliver an optimal solution	Text	Data Scientists at Work, Page 99, Paragraph 2	Solution has validity	Structural triad	- We interpret that "optimal solution" in thos context its related to the validity of the solution
because really the style of modeling of a physicist is usually about trying to identify a problem that is the key element, the key simplified description,which allows fundamental modeling.	Text	Data Scientists at Work, Page 5, Paragraph 1	Problem has description	Structural triad	- We interpret from the context of the interview that "the key simplified description" that the identified problem must have a description
because really the style of modeling of a physicist is usually about trying to identify a problem that is the key element, the key simplified description,which allows fundamental modeling.	Text	Data Scientists at Work, Page 5, Paragraph 2	Problem has solution	Structural triad	- We can interpret from the context of the interview that "which allows fundamental modeling" that the problem must have a model that will have its solution so the problem have a solution for different models
these positions require: experience in the use of data batch and streaming tools (e.g. Spark, AWS, Hadoop); understanding and experience in the use of high-performance machine learning algorithms and deep learning techniques; experience with programming languages such as Python, R, or Java	Text	On Understanding Data Scientists, page 1, paragraph 4	Tool has name	Structural triad	- We interpret from the context of the article that "e.g Spark, AWS, Hadoop" and "programming languages such as Python, R, o Java" can be interpreted by tool's name

My group here at The New York Times uses only open source statistical software, so everything is either in R or Python, leaning heavily on scikit-learn and occasionally IPython notebooks	Text	Data Scientists at Work, Page 6, Paragraph 8	Tool has application	Structural triad	- We interpret form the context of the article that "use of data batch and streaming tools" that tools ca be applis to wotk that Data Scientist make
Business analysts evaluate data looking for ways to improve organizational and strategic decision-making processes	Web Page	https://acortar.link/GTTg0H Lucidchart, What's the difference between a business analyst and a data analyst?, What is a Business Analyst?	Business Analyst is employee	Structural triad	- We interpret that "organizational" in this context means that the Bussines analyst is an employee at a company
"Con nuestro datos listos vamos a poder proceder a elegir modelo"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2:08	Data Scientist create model	Dynamic triad	We can infer that "proceder a elegir un modelo" refer to Data Scientist is the person who creates a model
"Que nosotros vamos a entrenar o ajustar para que sepa iterpretar los datos que le estamos dando"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2:14	Data Scientist trains application	Dynamic triad	We can interpret from the context of the video that the Data Scientist is training the application of the model.
"Podemos comenzar a generar predicciones"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2:21	Data Scientist performs solution	Dynamic triad	- We can interpret from the context of the video thtat "generar predicciones" ca be interpreted as Data Scientist performs solution - We interpret that prediction can be understood as a solution
Business analysts also collect data. However, they use the data to strategically find and propose ways to improve processes, procedures, products, software, and services.	Web Page	https://acortar.link/GTTg0H Lucidchart, What's the difference between a business analyst and a data analyst?, What is a Business Analyst?, Paragraph 1.	Business analyst collects data	Dynamic triad	- We can interpret that "Business analyst also collect data" refers to the Business analyst can collect data instead of a data scientist.
"Puedes comenzar a pensar en como limpiar estos datos"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 1:28	Data Scientist cleans data	Dynamic triad	- We can interpret from the context of the video that "Puede" ca be interptrede as Data Scientist
it is the information that is extracted from data	Text	On Understanding Data Scientists, Pagina 1, Parrafo 4	Data Scientist extracts information	Dynamic triad	- We can interpret from the context of the article that the one who extracts the information is the data scientist.
"Obtener fechas que indique eventos importantes de nuestros usuarios. Encontrar y relacionar los datos que nos indican obtener las respuestas a las preguntas planteadas anteriormente"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs , minute 1:47	Data Scientist interpret information	Dynamic triad	We can infer from the video that the person "asking the questions" is the data scientist and that this is information that comes out of the data.

By analyzing data, business analysts can identify problem areas	Web Page	https://acortar.link/GTTg0H Lucidchart, What's the difference between a business analyst and a data analyst?, What is a Business Analyst?	Business Analyst raises problem	Dynamic triad	- We interpret that "identify problem" in this context can be related to raise problem
"Vamos a poder proceder a elegir un modelo... que luego nosotros vamos a entrenar"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 1: 35	Data Scientis manages model	Responsibility	- We intrepret that the actions in this part of the video can be gathered in the respinsibility of managing model
the agent do a public information request to the state and get public university employee email addresses	Text	Data Scientists at Work, Page 114, Paragraph 6	Employee has email	Structural triad	We can infer that the employees of the company must be have email addresses
performing all the tasks in the analysis process	Text	On Understanding Data Scientists, Page 3, Paragraph 7	Data Scientist analizes data	Responsibility	- We infered that the one who does the "analyzes data" part is the 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	STYLE: - Insight providers - Modeling specialists - Plataform builder - Polymath - Team leader	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	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	

I believe that access to quality information and information relevant to our problems is the greatest challenge.	Text	On Understanding Data Scientists, Page 4, Paragraph 3	INFORMATION VALIDITY: -true -false	Note	- We can understood from the context of the article that "quality information and information relevant" can be interpreted as the information must be validated for it´s analysis so the information has a validation
we deliver an optimal solution	Text	Data Scientists at Work, Page 99, Paragraph 2	SOLUTION VALIDITY: -true -false	Note	- We can interpret from the context of the article that "optimal solution" can be no optimal solution so the solution should be validated
quality of the data	Text	On Understanding Data Scientists, Page 3, Paragraph 11	QUALITY AMOUNT: -High - Low	Note	- We inferred that qualita has amount because it can be measured in low or high
"Con nuestro datos listos"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2: 08	If information.validity = true	Conditional	- We can interpret of the context of the video that "datos listos" refers to information.validity = true
situations in which they've encountered problems	Text	Data Scientists at Work, Page 125, Paragraph 4	Problem appears	Event	- We can interpret that "encountered problems" its related to a problem that appeared
"Con nuestro datos listos vamos a poder proceder a elegir modelo"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2: 08	information.validity = true, then Data Scientist creates model	Implication	- We can interpret of the context of the video that "datos listos" refers to information.validity = true - We can infer that when "datos listos" then data scientist creates the model
"Este modelo puede ser un modelo algebraico o estadístico que nosotros vamos a entrenar"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2: 13	Data Scientist creates model, then Data Scientist trains application	Implication	- We can infer from the context or the video that when the data scientist creates the model then he/she can start training it
"Vamos a entrenar o ajustar para que sepa interpretar los datos que le estamos dando una vez hecho esto podemos comenzar a generar predicciones"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2: 15	Data Scientist trains application, then Data Scientist performs solution	Implication	- We interpret that prediction can be understood as a solution - We can infer that when the Data Scientist trains the model is when it can begin to yield solutions.
By analyzing data, business analysts can identify problem areas	Web Page	https://acortar.link/GTTg0H Lucidchart, What's the difference between a business analyst and a data analyst?, What is a Business Analyst?	problem appears, then Business Analyst raises problem	Implication	- We interpret that "identify problem" in this context can be related to raise problem

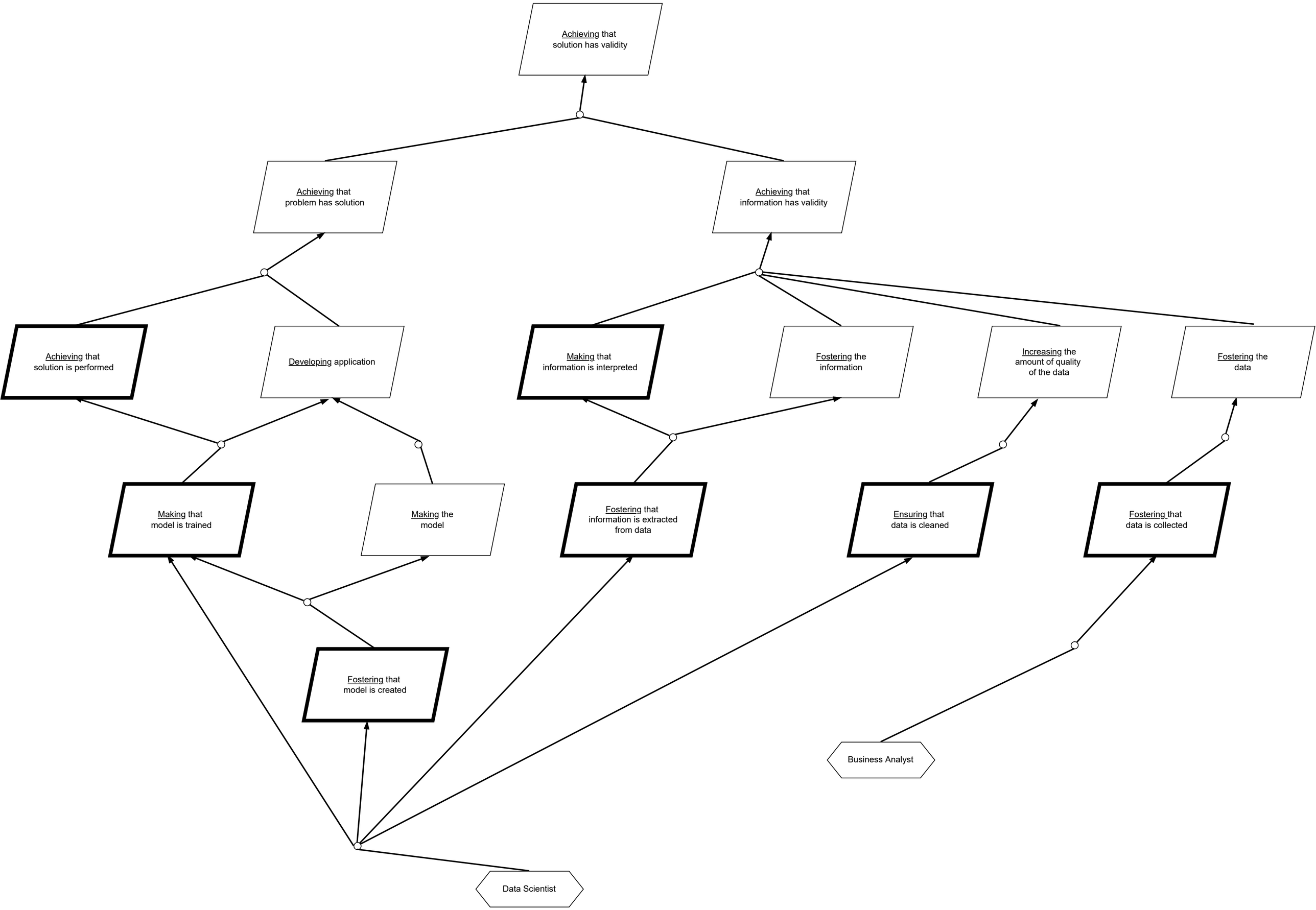
Business analysts also collect data. However, they use the data to strategically find and propose ways to improve processes, procedures, products, software, and services.	Web Page	https://acortar.link/GTTg0H Lucidchart, What's the difference between a business analyst and a data analyst?, What is a Business Analyst?, Paragraph 1.	Business Analyst raises problem, then Business Analyst collects data	Implication	- We can interpret that "they use the data to strategically find and propose ways to improve processes" refers to when the Bussines Analyst collects data start to find ways to improve procces raises problems that later will be resolve
Business analysts also collect data. However, they use the data to strategically find and propose ways to improve processes, procedures, products, software, and services.	Web Page	https://acortar.link/GTTg0H Lucidchart, What's the difference between a business analyst and a data analyst?, What is a Business Analyst?, Paragraph 1.	Business Analyst collects data, then Data Scientist cleans data	Implication	- From all the research, and work we have done with other teams, we infered that Business Analyst collect data, an then he passes this data to the data scientist who is the one in charge of cleaning it
People capable of gathering, cleaning and using data to extract knowledge	Text	On Understanding Data Scientists, Page 1, Paragraph 4	Data Scientist cleans data, then Data Scientist extracts information	Implication	- We interpreted that "people" referes to the data scientists interviewed in the article and "knowledge" its related to information
"Que patrones de datos obtener sobre tus usuarios... luego limpiar y organizar estos datos para poder trabajar con ellos"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 0: 45	Data Scientist extracts information, then Data Scientist interprets data	Implication	- We can interpret "obtener patrones" as Data Scientis extracting information, and, "datos para poder trabajr con ellos" as Data Scientist interpreting the data
we deliver an optimal solution	Text	Data Scientists at Work, Page 99, Paragraph 2	Achieving that solution has validity	Goal	- We interpret that "optimal solution" in thos context its related to the validity of the solution, so then they have to achieve that
access to quality information	Text	On Understanding Data Scientists, Page 4, Paragraph 2	Achieving that information has validity	Goal	- We interpret that they want the information to have quality, so then they ave to achieve its validity
prototype to make sure that my solution actually addresses the problem	Text	Data Scientists at Work, Page 287, Paragraph 2	Achieving that problem has solution	Goal	- We can interpret that "addresses" in this case is related to Achieving
Also helpful will be the extensibility of the language that is used for developing the models.	Text	Data Scientists at Work, Page 235, Paragraph 1	Developing application	Goal	- We can interpret that "model" in this case is related to application
in order to increase data quality	Text	On Understanding Data Scientists, Page 3, Paragraph 7	Increasing the amount of quality of the data	Goal	
access to quality information	Text	On Understanding Data Scientists, Page 4, Paragraph 2	Fostering the information	Goal	- We interpreted that they are fostering the data to have acces to quality information

"Con nuestro datos listos vamos a poder proceder a elegir modelo"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs, minute 2: 08	Making the model	Goal	- We can interpret that "elegir modelo" wants to emphasize that the are making the model
All these projects centered around data	Text	Data Scientists at Work, Page 152, Paragraph 8	Fostering the data	Goal	- We interpret the "centered around data" as they want to emphasize in data, so then they are fostering the data
they have to be implementers and show that they have both the ability and passion to build solutions	Text	Data Scientists at Work, Page 100, Paragraph 7	Achieving that solution is performed	Requirement	- We interpret "they have" as some kind of achievement by part of the data scientist, and that "build" in this case is related to perform
It was never about data for me. For me, data was and is a means to an end. For me, it's always been about the power of the model you can train, and so it's about learning algorithms.	Text	Data Scientists at Work, Page 50, Paragraph 4	Making that model is trained	Requirement	
It was clear to me by the end of my first day here that we should build a predictive model for looking at subscriber behavior.	Text	Data Scientists at Work, Page 11, Paragraph 6	Fostering that model is created	Requirement	We can inferred that one of the primary task for a Data Scientist is create or build a model
They should start by taking a class on information retrieval or learn from the vast array of resources available offline and online.	Text	Data Scientists at Work, Page 89, Paragraph 8	Fostering that information is extracted from data	Requirement	We can interpreted that they extracts the information from a vast array of resources availables offline and online, and we inferred that they was talking about data
I now lead the team that enables LinkedIn's 300M+ members to effectively and efficiently leverage LinkedIn's professional content to satisfy their information needs. The includes query understanding, but also scoring and ranking of results, and everything else we need to do to deliver the right results to the right searchers at the right time.	Text	Data Scientists at Work, Page 85, Paragraph 7	Making that information is interpreted	Requirement	When he mentions "Information needs" and "query understanding" we could inferred that one primary thing that Data Scientist should do is to make that the information is understandable
Most of the time you spent was doing feature design, data cleaning first of all, and then feature design	Text	Data Scientists at Work, Page 59, Paragraph 2	Ensuring that data is cleaned	Requirement	When the Data Scientist interviewed said "Data cleaning first of all", whe inferred that "Ensuring that data is cleaned" is mandatory

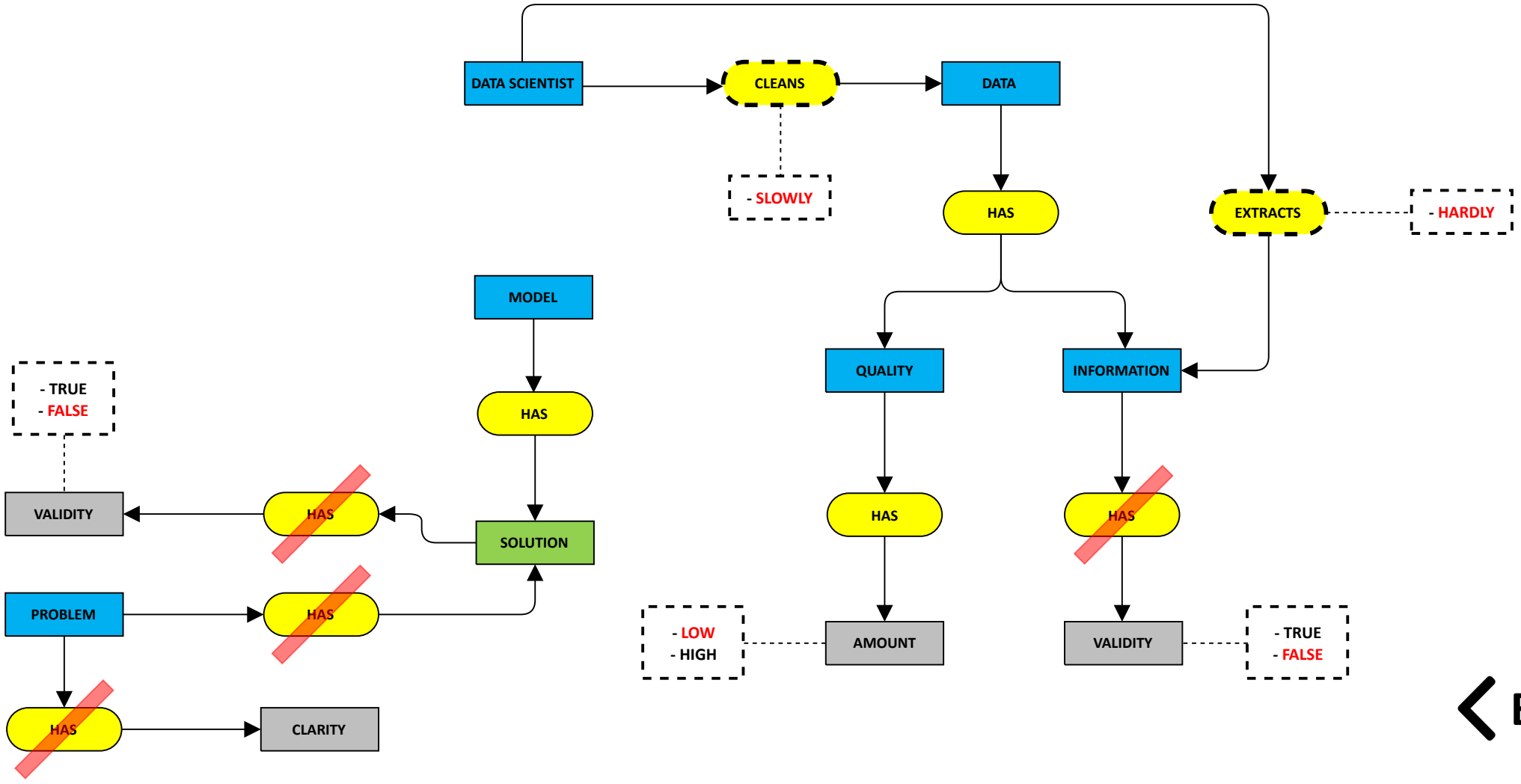
They should start by taking a class on information retrieval or learn from the vast array of resources available offline and online. Given the open source technology for search, they should learn by doing—for instance, implementing a basic search engine for a public data collection. It’s not hard to get started with search	Text	Data Scientists at Work, Page 89, Paragraph 7	Fostering that data is collected	Requirement	We inferred by the context of this interview that tha organization should foster the collection of the data, for that reason we deduced that "Fostering that data is collected" should be a requirement

GOAL DIAGRAM

BACK



PROBLEMS PRE-CONCEPTUAL SCHEMA



BACK

BACK

Document Traceability Table (Problems Preconceptual Schema)

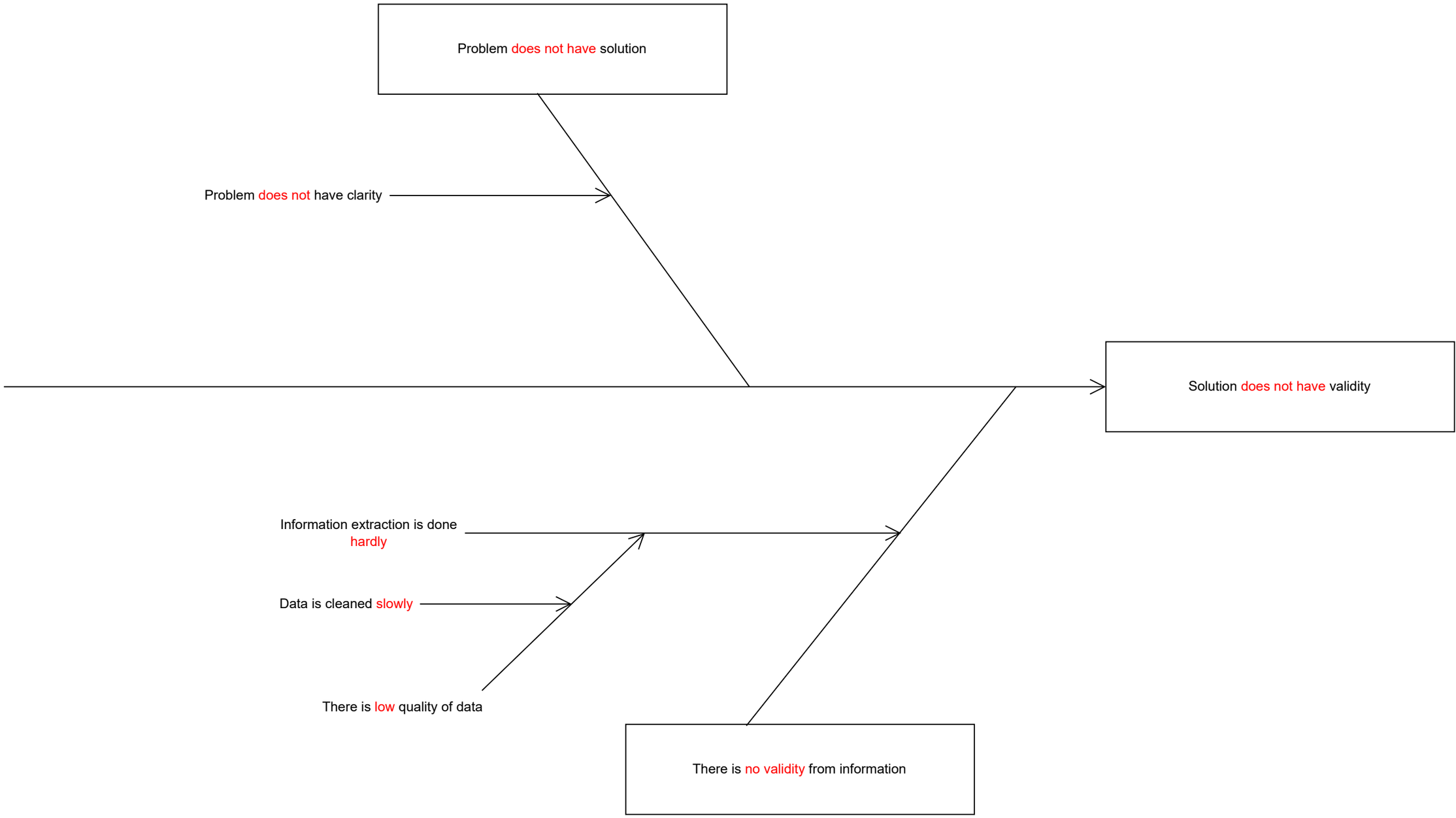
Original sound/Image/Text	Source	Location	Element	Kind of element	Observations
"Puedes comenzar a pensar en como limpiar estos datos"	Video	https://www.youtube.com/watch?v=Bl2sBiVdZHs , minute 1:28	Data scientist cleans data	Dynamic triad	- We can interpret from the context of the video that "Puede" ca be interptrede as Data Scientist
it is the information that is extracted from data	Text	On Understanding Data Scientists, Pagina 1, Parrafo 4	Data Scientist extracts information	Dynamic triad	- We can interpret from the context of the article that the one who extracts the information is the data scientist.
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
I believe that access to quality information and information relevant to our problems is the greatest challenge.	Text	On Understanding Data Scientists, Page 4, Paragraph 3	Data has information	Structural triad	We interpret from the context of the article that "information relevant to our problems" ca be interpreted as the Data carries with it the information that will be used
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 1, paragraph 4	Quality has amount	Structural triad	- We infer that "quality of the data to be assessed" can be interpreted as data have different levels of quality
The biggest difficulty is the access to quality information	Text	On Understanding Data Scintist, Page 4, Paragraph 2	Information does not have validity	Problem	- We interpret that the difficulty to get quality infomation refers to the infomation not having validity

I led the product data science team, a group of data scientists focused on creating innovative solutions to improve LinkedIn’s products and create new ones.	Text	Data Scientists at Work, Page 85, Paragraph 4	Model has solution	Structural triad	- We can interpret from the context of the interview that "creating innovative solutions to improve" refers to a model has solution
we deliver an optimal solution	Text	Data Scientist at Work, Page 99, Paragraph 2	Solution does not have validity	Problem	- We interpret that "optimal solution" in thos context its related to the validity of the solution but if the solution can be optimal then also can be no validity
because really the style of modeling of a physicist is usually about trying to identify a problem that is the key element, the key simplified description, which allows fundamental modeling.	Text	Data Scientist at Work, Page 5, Paragraph 2	Problem does not have solution	Problem	- We can interpret from the context of the interview that "which allows fundamental modeling" that the problem must have a model that will have its solution so the problem have a solution for different models - The different solutions of the model may not solve the proposed problem
the lack of clear problems	Text	On Understanding Data Scientists, Page 4, Paragraph 4	Problem does not have clarity	Problem	- We can interpret this as problems not having clarity
The biggest difficulty is the access to quality information	Text	On Understanding Data Scintist, Page 4, Paragraph 2	There is no validity from information	Problem	- We interpret that the difficulty to get quality infomation refers to the infomation not having validity
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	There is low quality of data	Problem	- We can interpret that "the anomalies that affect the quality of the data" relates to the low quality of it

The majority of the participants agree that this is still one of the most time-consuming and laborious tasks	Text	On Understanding, Page 3, Paragraph 7	Data is cleaned slowly	Problem	- We inferred that the task relates to cleaning data, and "most time_consuming" as it it's done slowly
the biggest difficulty is the access to quality information	Text	On Understanding Data Scientistss, Page 4, Paragraph 2	Information extraction is done hardly	Problem	We can interpreted that "hard" is a sinonym of "difficult".
I believe that access to quality information and information relevant to our problems is the greatest challenge.	Text	On Understanding Data Scientists, Page 4, Paragraph 3	INFORMATION VALIDITY: TRUE, FALSE	Note	- We can understood from the context of the article that "quality information and information relevant" can be interpreted as the information must be validated for it´s analysis so the information has a validation
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	AMOUNT: LOW	Problem	- We can interpret that "the anomalies that affect the quality of the data" relates to the low quality of it
we deliver an optimal solution	Text	Data Scientist at Work, Page 99, Paragraph 2	SOLUTION VALIDITY: TRUE, FALSE	Note	- We can interpret from the context of the article that "optimal solution" can be no optimal solution so the solution should be validated

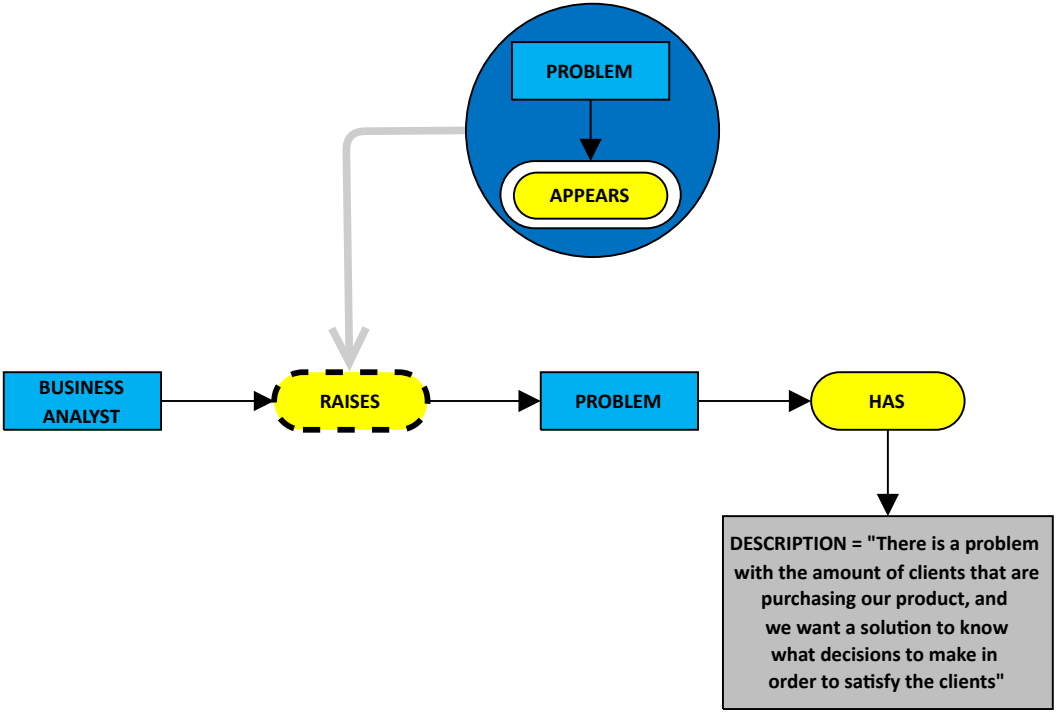
CAUSE-AND-EFFECT DIAGRAM

BACK



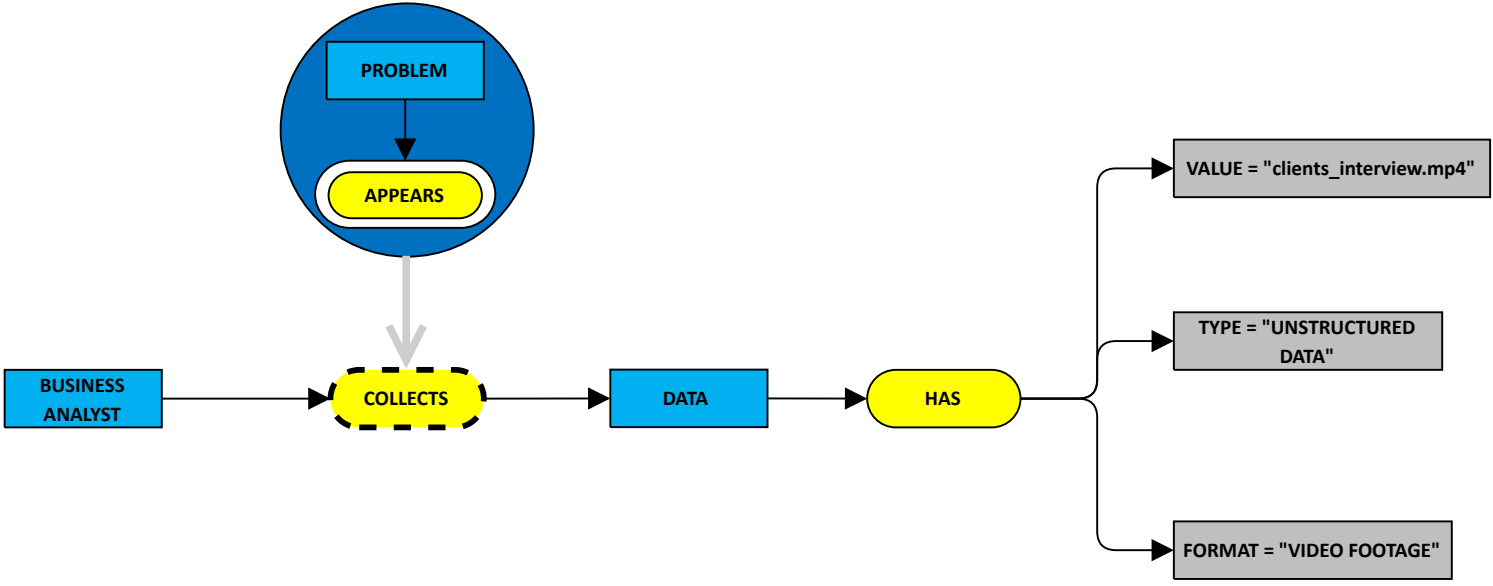
BACK

EXECUTABLE PS #1



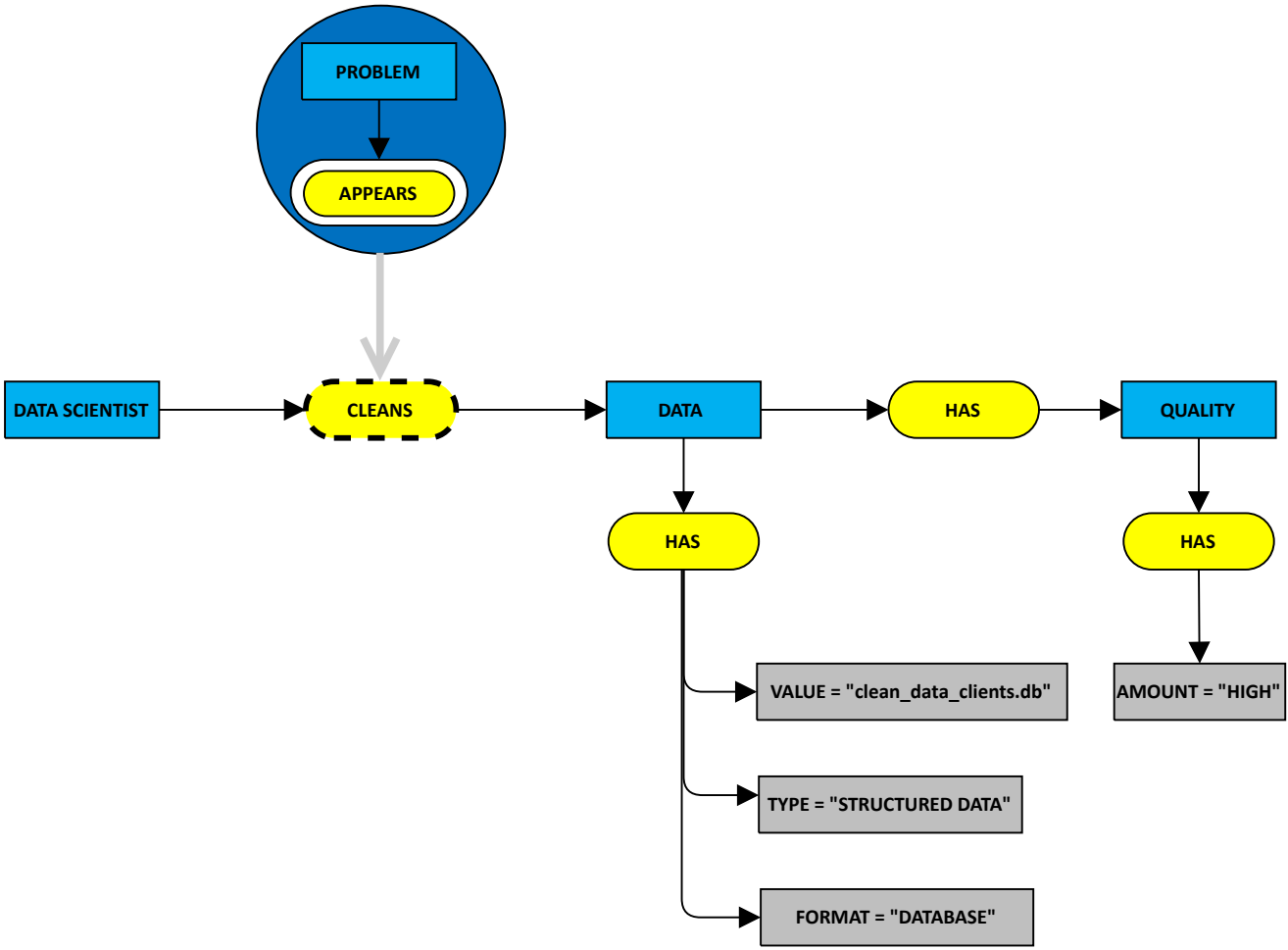
BACK

EXECUTABLE PS #2



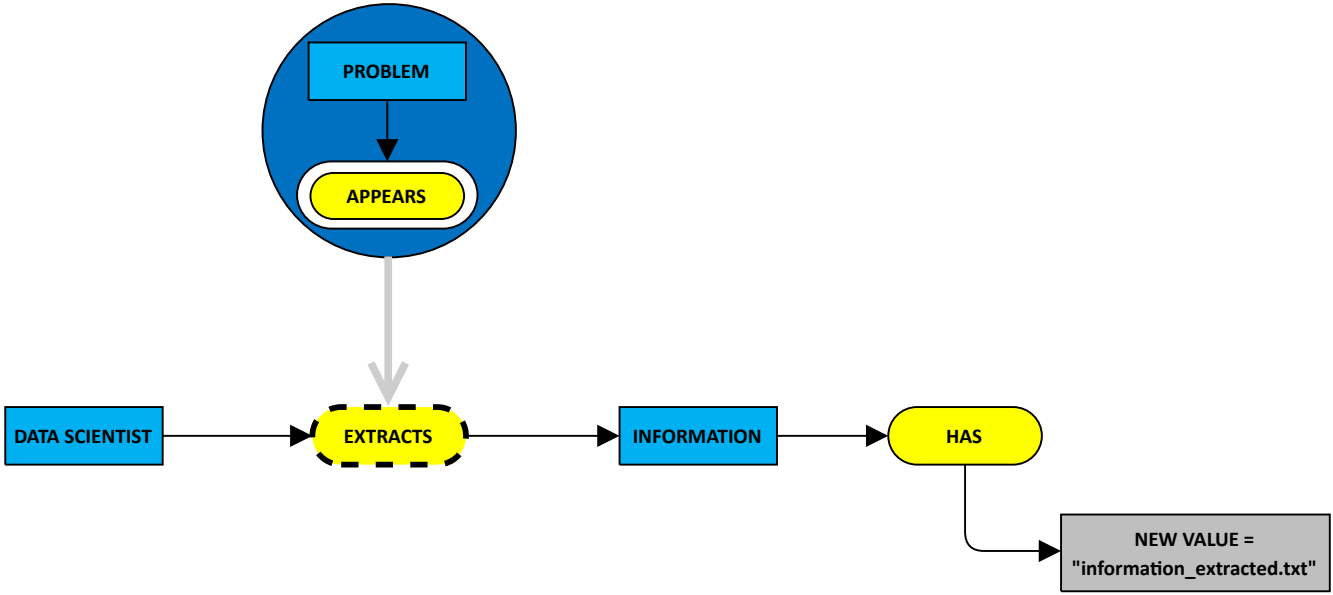
BACK

EXECUTABLE PS #3



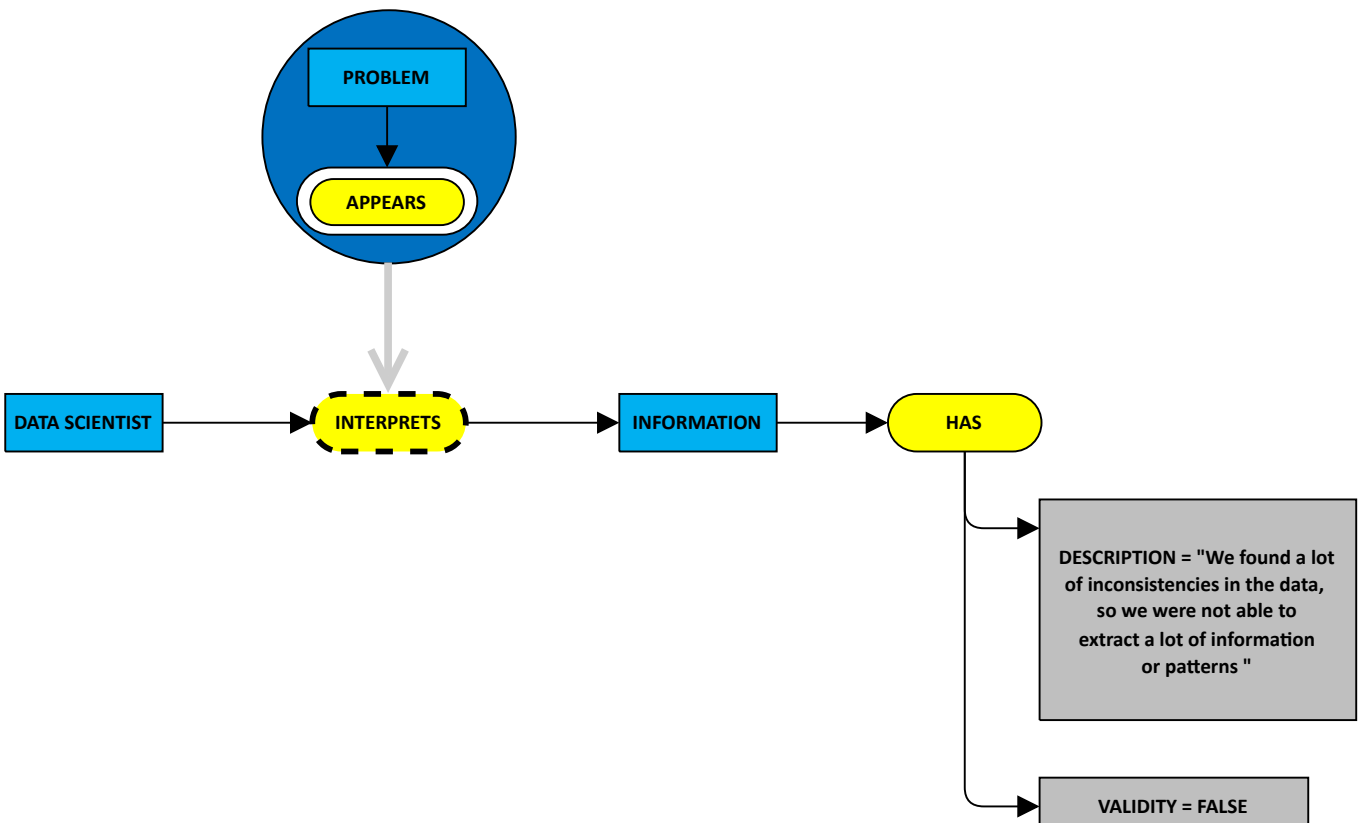
BACK

EXECUTABLE PS #4



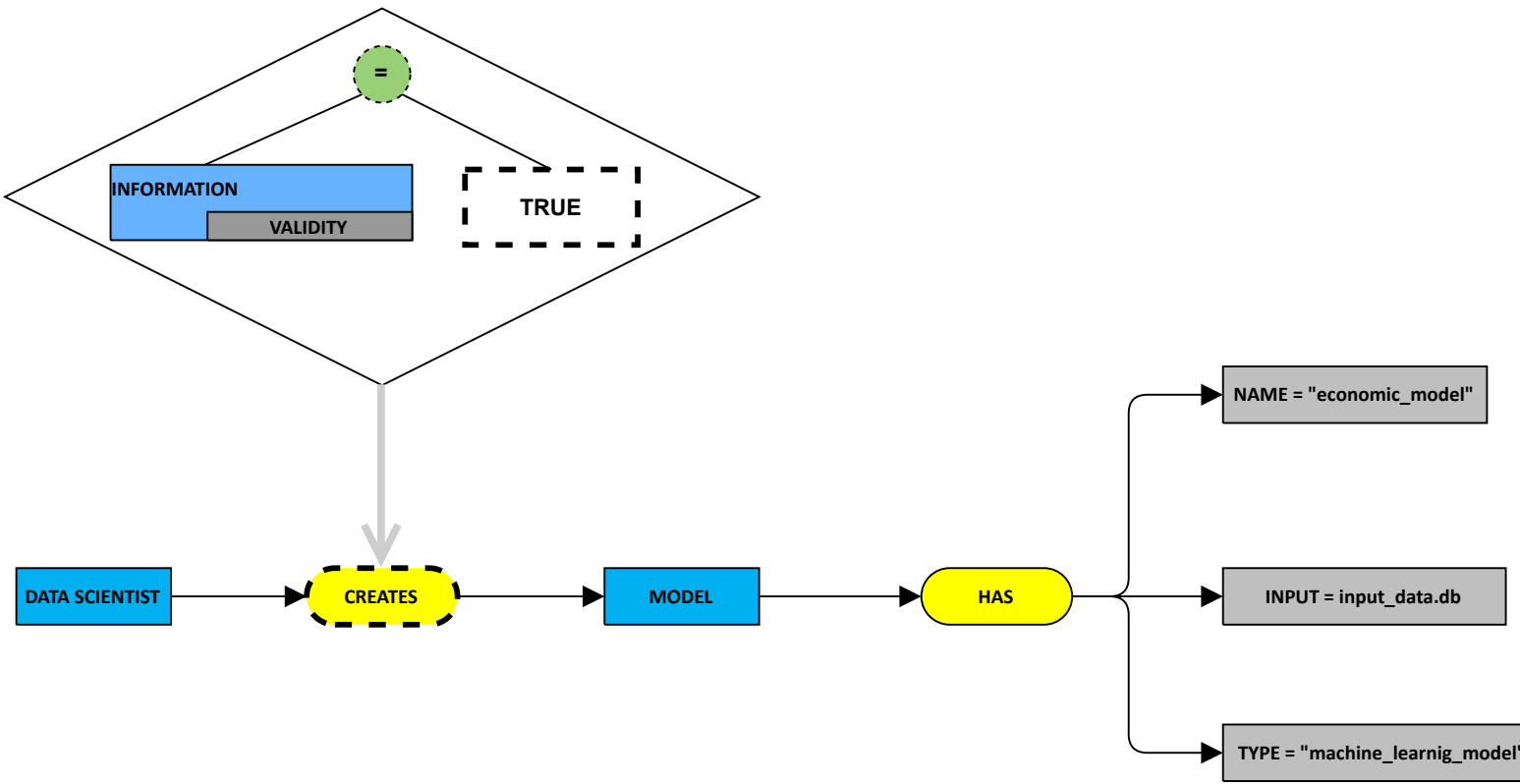
BACK

EXECUTABLE PS #5



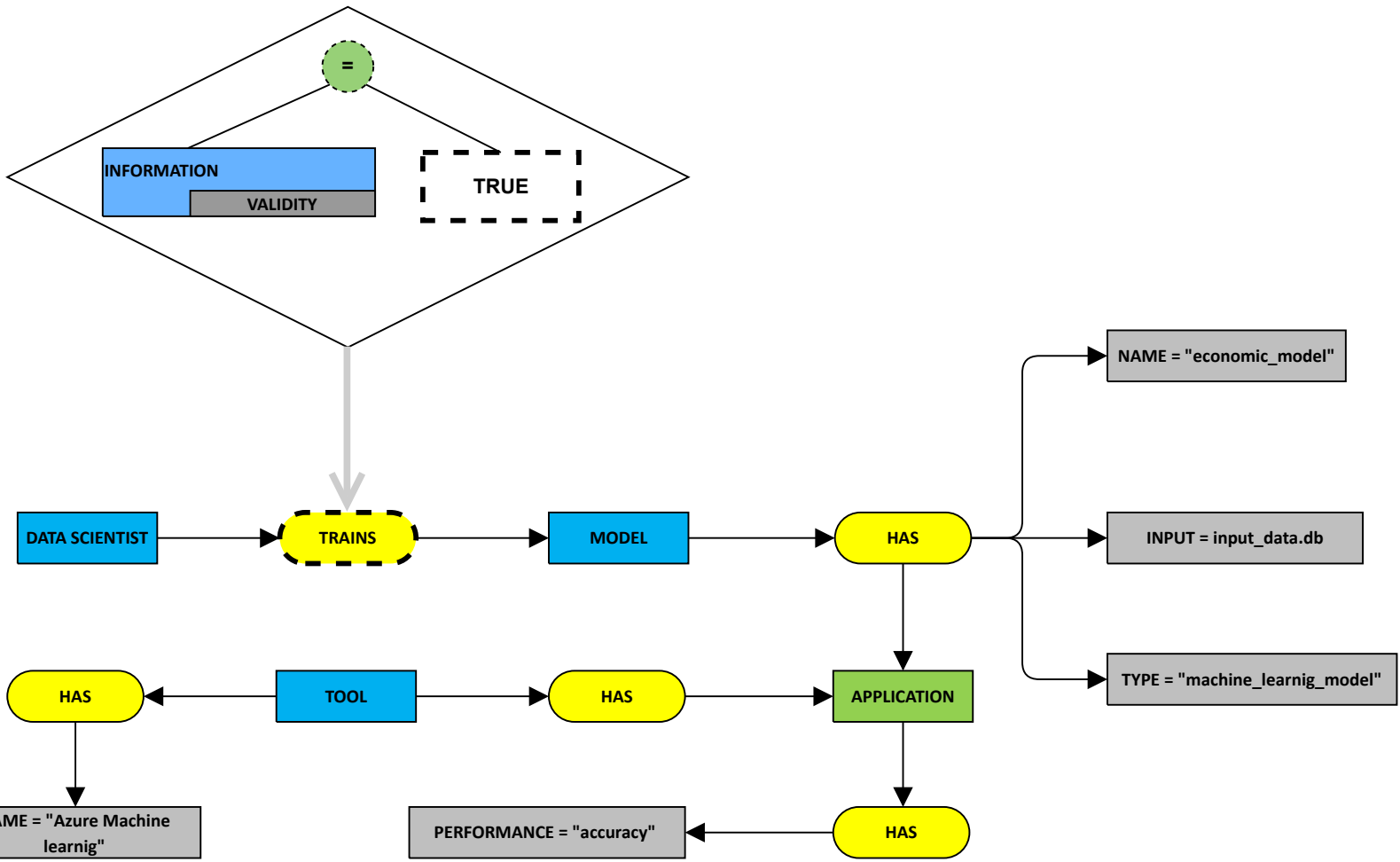
BACK

EXECUTABLE PS #6



BACK

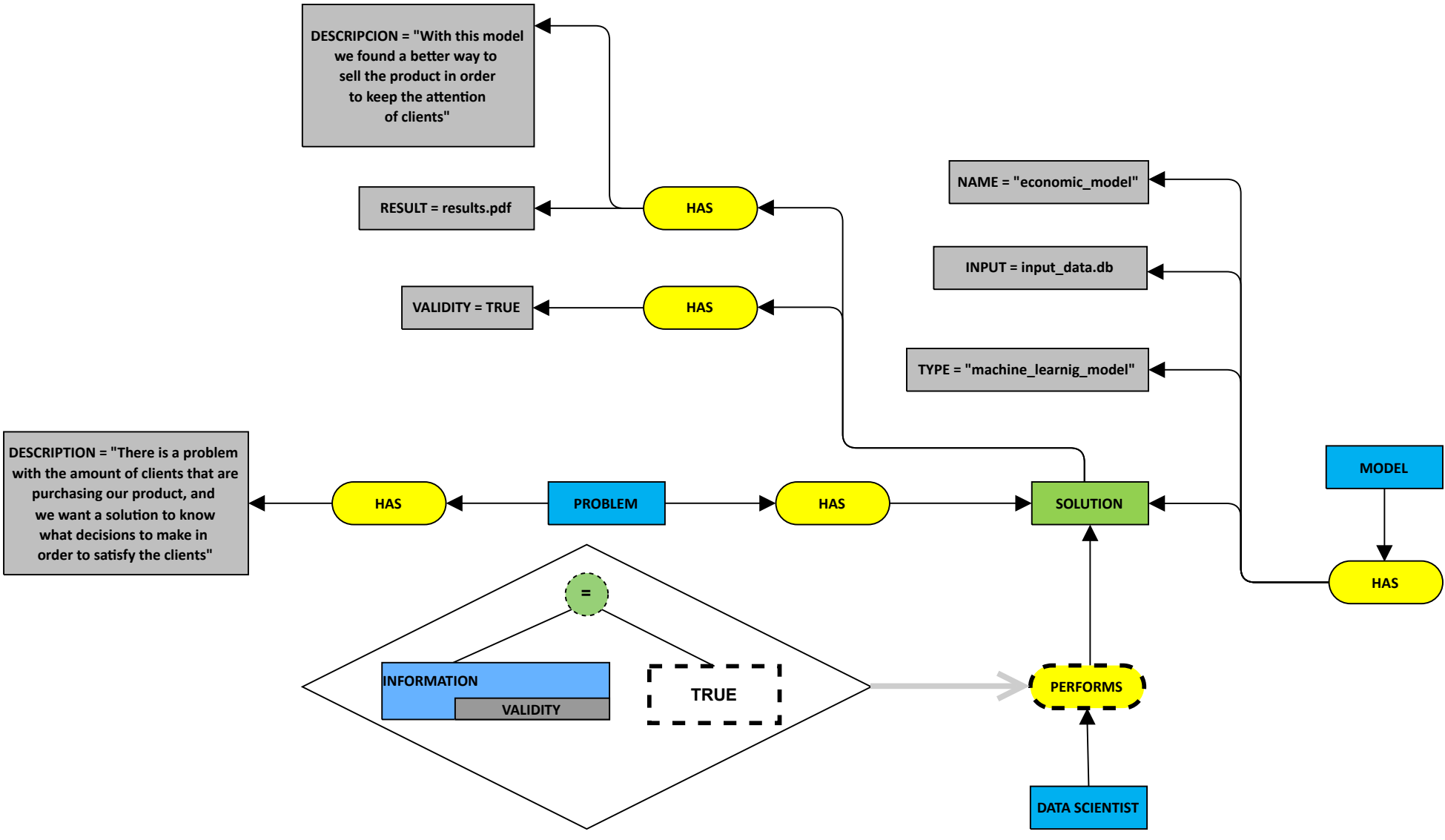
EXECUTABLE PS #7



APPLICATION				
TOOL.NAME	MODEL.NAME	MODEL.INPUT	MODEL.TYPE	PERFORMANCE
"Azure Machine Learning"	"economic_model"	input_data.db	"machine_learning_model"	"accuracy"

BACK

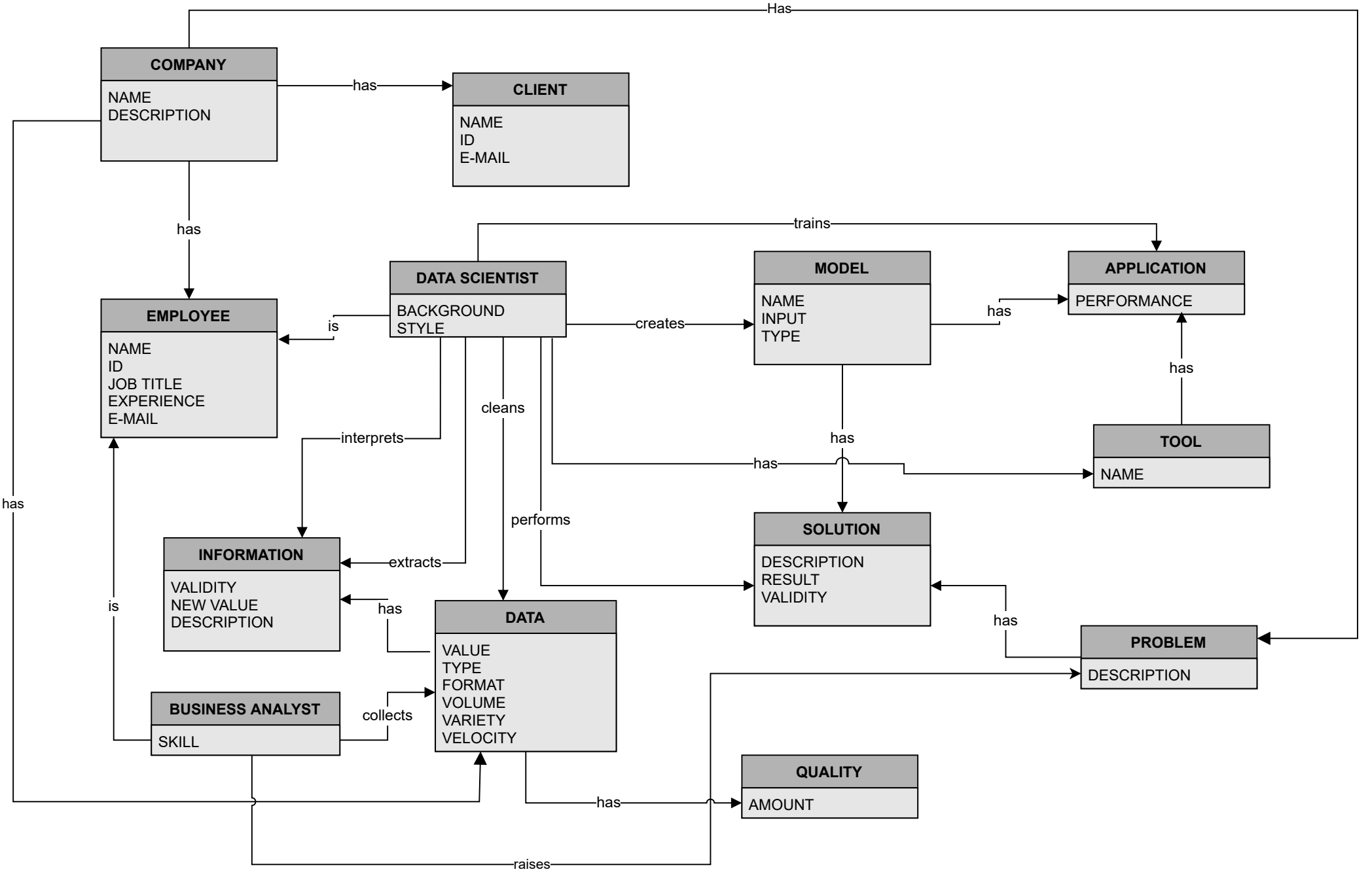
EXECUTABLE PS #8



SOLUTION						
MODEL.NAME	**MODEL.INPUT**	**MODEL.TYPE**	**PROBLEM.DESCRIPTION**	**RESULT**	**DESCRIPTION**	**VALIDITY**
"economic_model"	input_data.db	"machine_learning_model"	There is a problem with the amount of clients that are purchasing our product, and we want a solution to know what decisions to make in order to satisfy the clients	results.pdf	With this model we found a better way to sell the product in order to keep the attention of clients	TRUE

DOMAIN MODEL

BACK



KANBAN BOARD

BACK

GOAL	TO DO	DOING	DONE

Task 1 : Gathering information and interviews on data scientists

Responsible: Diego Valentin Osorio Marín, Fredy Alberto Orozco Loaiza

Task 2 : Making the controlled dialog

Responsible : Diego Valentín Osorio Marín

Task 3 : Collecting and organizing the digital files

Responsible : Team

Task 4 : Filling the elicitation cards

Responsible : Fredy Alberto Orozco Loaiza

Task 5 : Making and updating the organizational chart

Responsible : Team

Task 6 : Verify Consistency

Responsible : Team

Opportunity

Identified

- An opportunity was identified that a software-based solution could address.
- A stakeholder wants to make an investment in a better understanding potential value.
- Identified other stakeholders who want to share in the opportunity.

1/6

Stakeholders

Recognized

- Stakeholders identified.
- There is an agreement between the stakeholder groups to be represented.
- Responsibilities of stakeholder representatives defines

1/6

Stakeholders

Represented

- The representatives of the interested parties were summoned
- The representatives of the interested parties accept the responsibilities and authorized them
- The collaborative approach was agreed upon
- Representatives respect the way of working

2/6

Task 7 : Making the domain model

Responsible : Fredy Alberto Orozco Loaiza

Task 8 : Making, organizing and updating the pre-conceptual schema

Responsible : Team

Task 9 : Making, organizing and updating the executable pre-conceptual schemas

Responsible : Diego Valentín Osorio Marín

Task 10 : Making, organizing and updating the document of traceability table

Responsible : Team

Tarea 11 : Making Goal Diagram

Responsible : Diego Valentín Osorio Marín

Tarea 12 : Making Cause-and-Effect Diagram

Responsible : Jaime Andrés Monsalve Ballesteros, Diego Valentín Osorio Marín

Tarea 13 : Verify Consistency

Responsible : Santiago Castro Tabares

Task 14 : Making, organizing and updating the problems pre-conceptual schema

Responsible : Diego Valentín Osorio Marín

Requirements

Conceived

- The need for a new system is clear
- Users were identified
- Initial promoters were identified

1/6

Task 15 : Taking control of the kanban board

Responsible : Diego Valentín Osorio Marín, Jaime Andrés Monsalve Ballesteros

Task 16 : Taking the alpha advance report updated

Responsible : Santiago Castro Tabares

Task 17 : Taking some pictures to show evidence of the meetings

Responsible : Diego Valentín Osorio Marín

Task 18 : Attend classes with the teacher

Responsible : Team

Task 19 : Study class material

Responsible : Team

Tarea 20 : Conduct meetings

Responsible : Team

Task 21 : Elaboration of KANBAN Board

Responsible : Diego Valentín Osorio Marín

Work

Initiated

- The initiator of the work is known
- Work restrictions were clarified
- Sponsorship and funding model clarified
- Priority of work is clarified

1/6

Team

Seeded

- The team's mission is clear
- The team knows how to grow to achieve its mission
- The required competencies were identified
- The size of the team has been determined

1/5

Way of Working

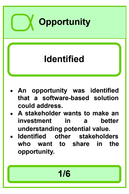
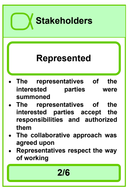
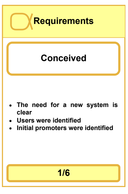
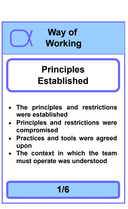
Principles Established

- The principles and restrictions were established
- Principles and restrictions were compromised
- Practices and tools were agreed upon
- The context in which the team must operate was understood

1/6

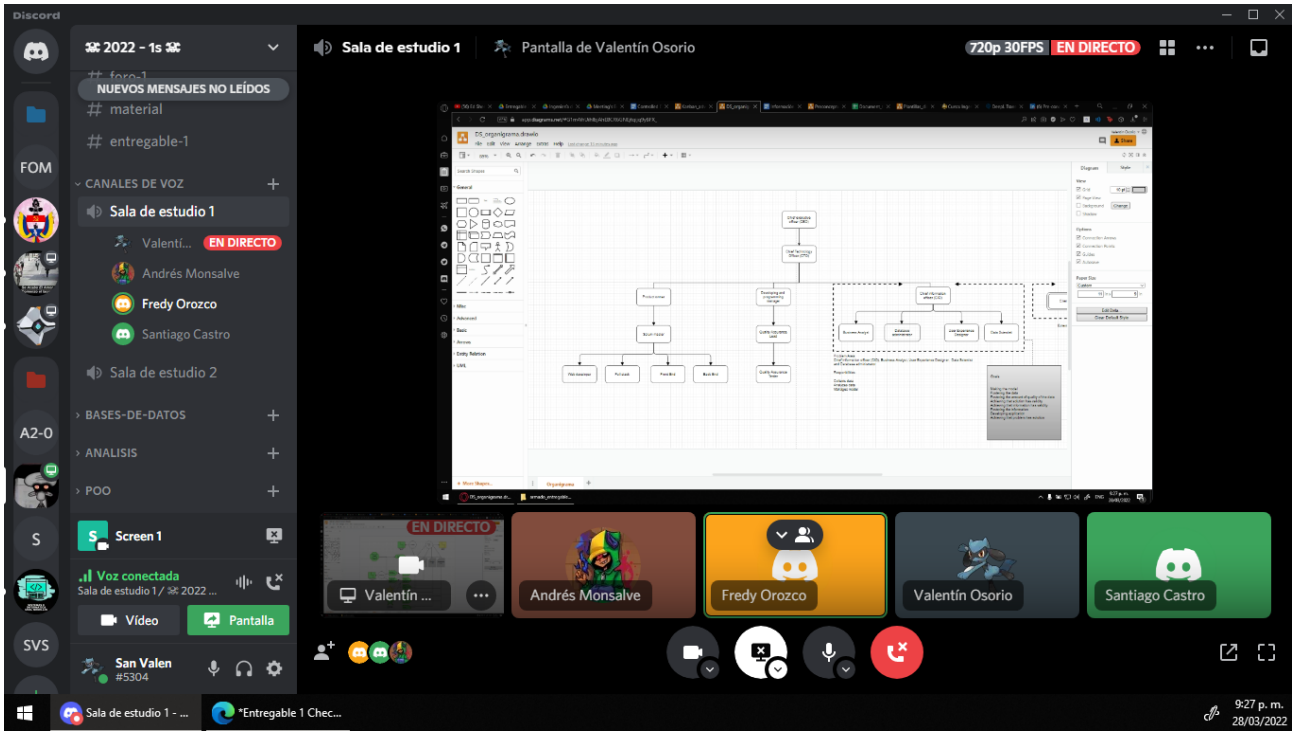
Alphas Advance Report

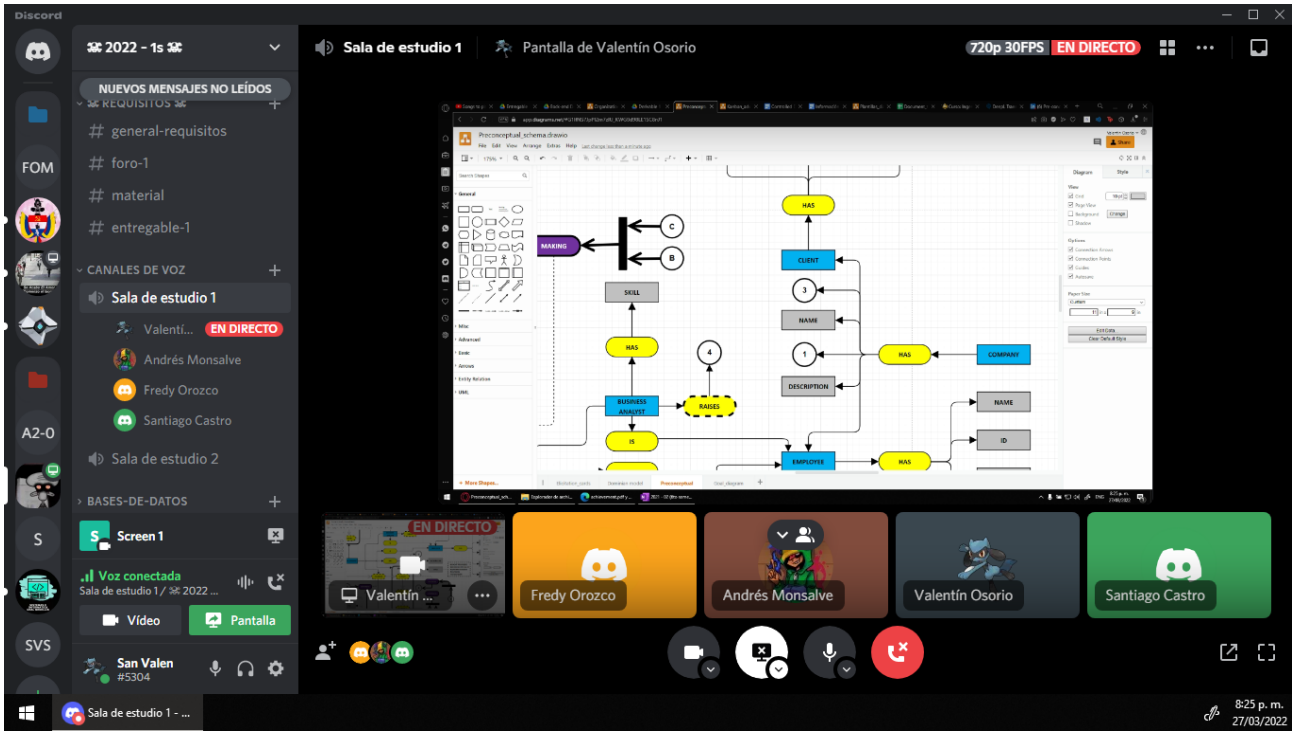
BACK

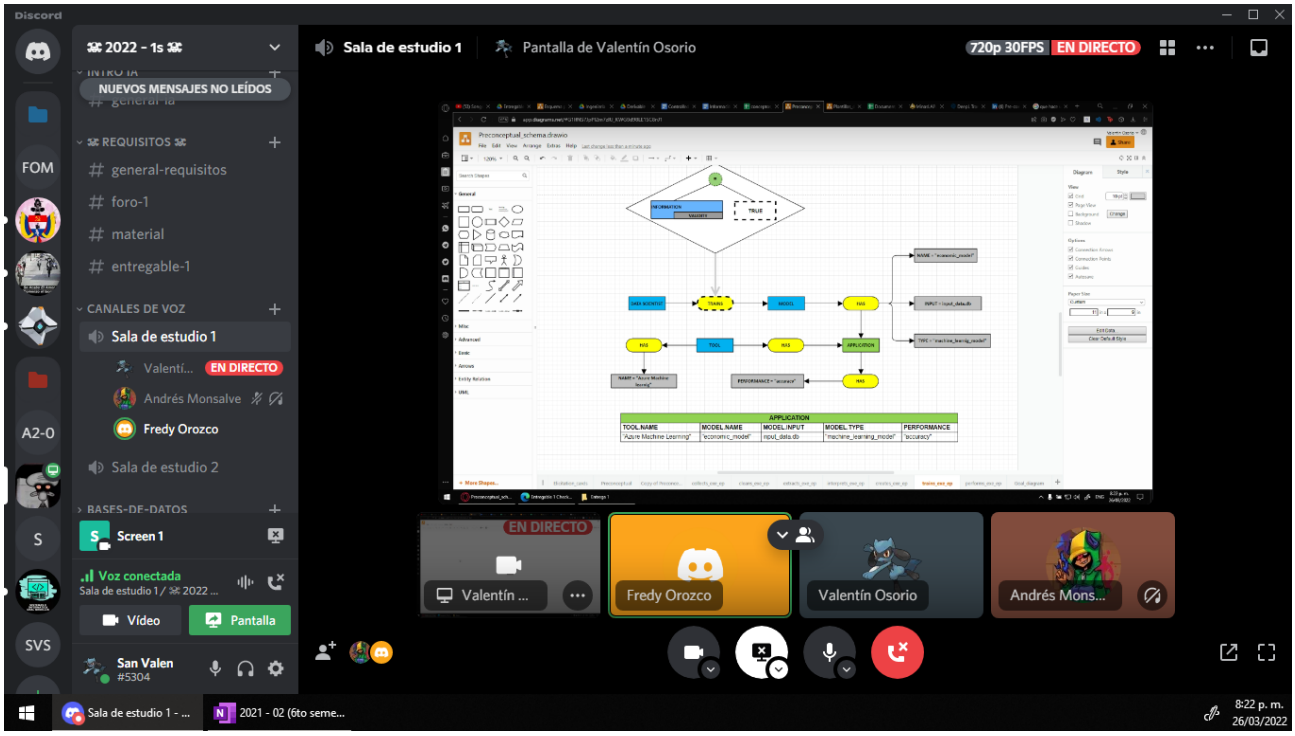
State	Summary of the achievement	Task	Date/Duration	Contents/Observations
		Ghatering information and interviews on data scientists	Start date: 19/03/2022 5 hours	The team got the information from an article, a book, some videos and web pages
		Collecting and organizing the digital files	Start date: 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.
		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 were 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.
		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.
		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
		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 pre-conceptual schema.
		Conduct meetings	Start date: 21/06/2022 10 hours	A meeting was held to work together and tasks were assigned.
		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
		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.

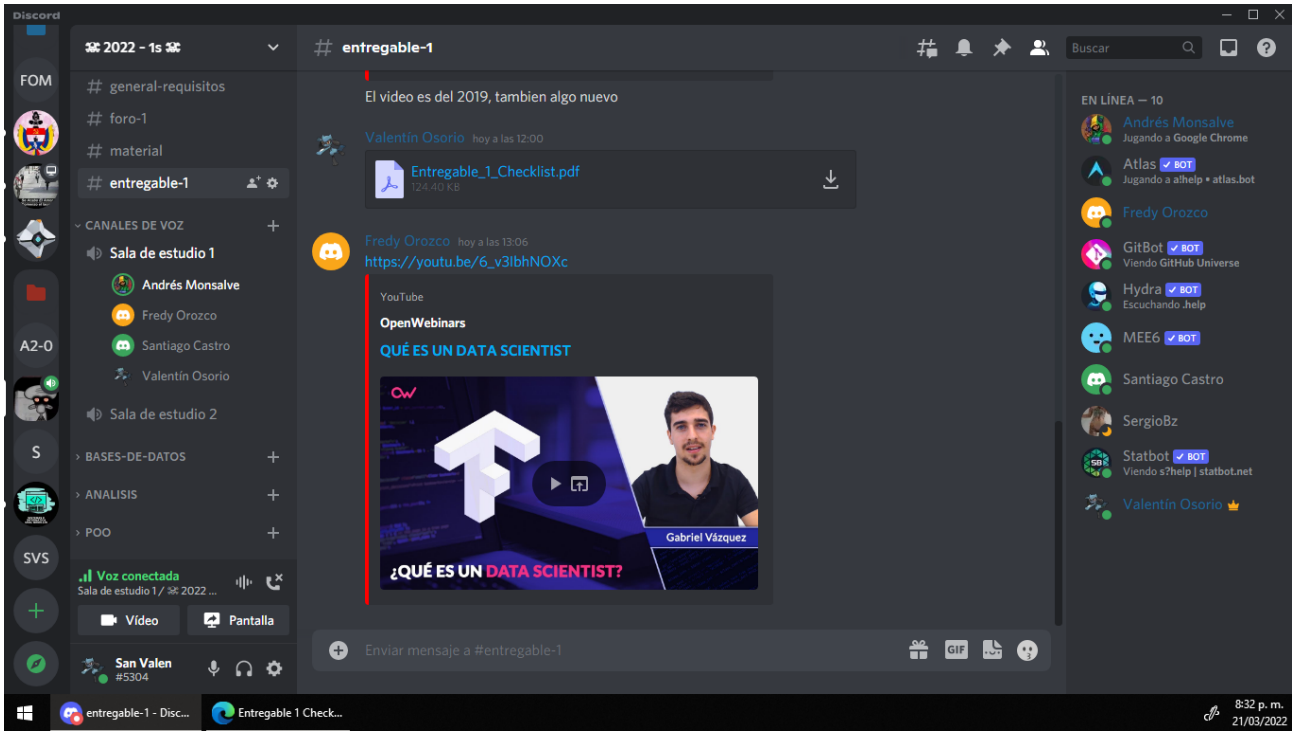
BACK

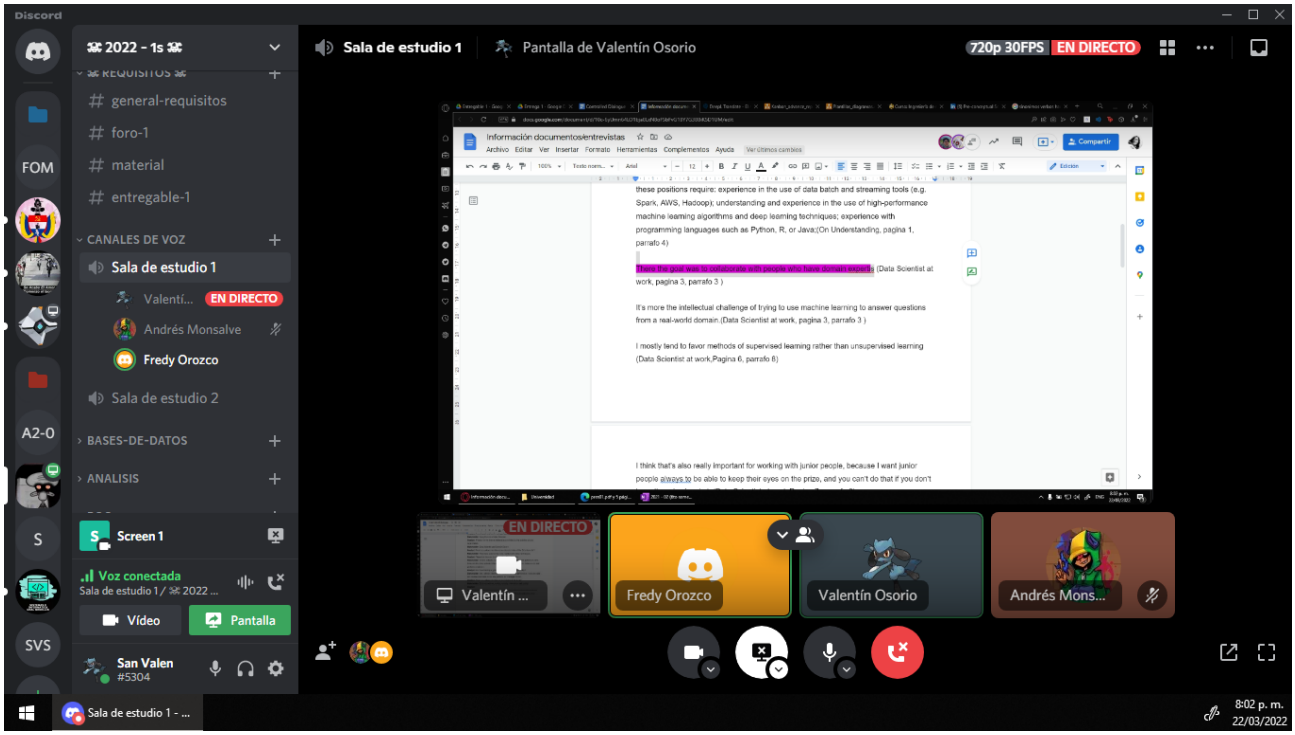
EVIDENCE OF THE MEETINGS

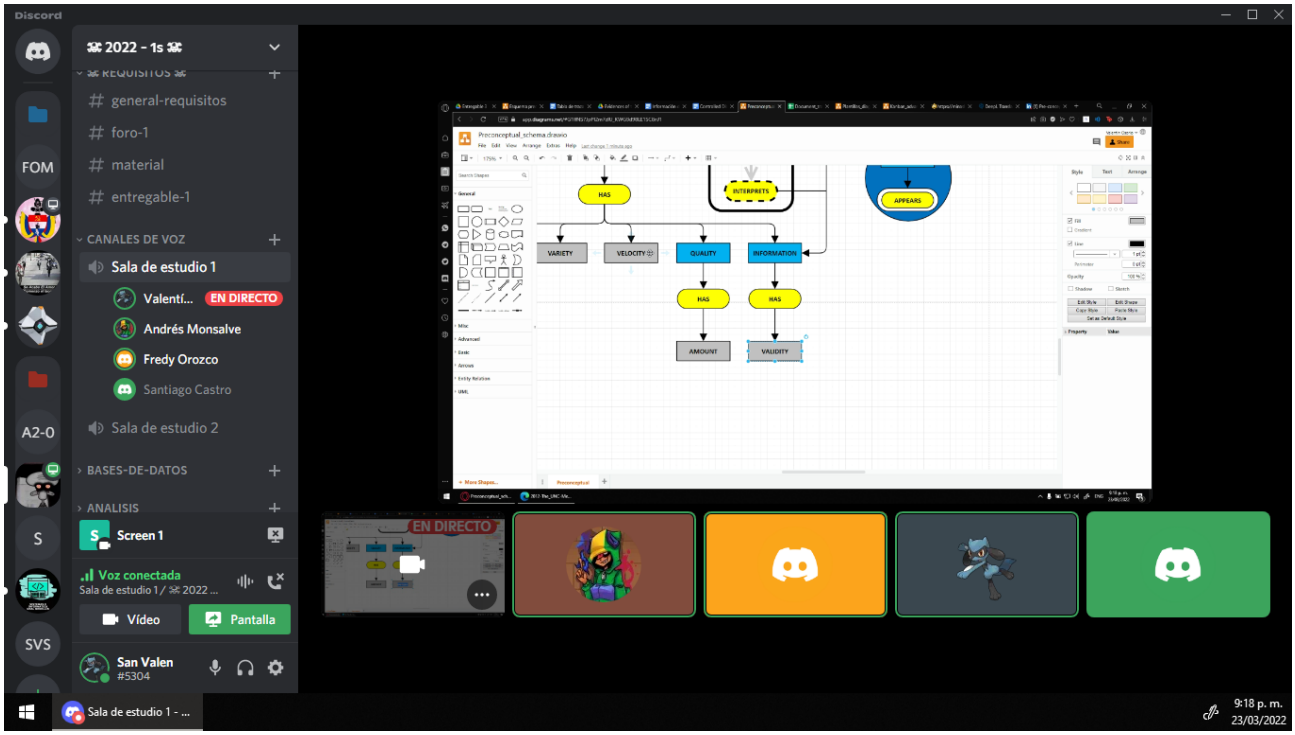





Discord
2022 - 1s
NUEVOS MENSAJES NO LEÍDOS
REQUISITOS
general-requisitos
foro-1
material
entregable-1
CANALES DE VOZ
Sala de estudio 1
Valentí... EN DIRECTO
Andrés Monsalve
Fredy Orozco
Sala de estudio 2
BASES-DE-DATOS
Screen 1
Voz conectada
Sala de estudio 1 / 2022 ...
Vídeo
Pantalla
San Valen
#5304
Sala de estudio 1
Pantalla de Valentín Osorio
720p 30FPS
EN DIRECTO
Valentín ...
Fredy Orozco
Valentín Osorio
Andrés Mons...
Sala de estudio 1 - ...
2021 - 02 (6to seme...
8:22 p. m.
26/03/2022


Discord
2022 - 1s
general-requisitos
foro-1
material
entregable-1
CANALES DE VOZ
Sala de estudio 1
Andrés Monsalve
Fredy Orozco
Santiago Castro
Valentín Osorio
Sala de estudio 2
BASES-DE-DATOS
ANALISIS
POO
Voz conectada
Sala de estudio 1 / 2022 ...
Vídeo
Pantalla
San Valen
#5304
entregable-1
Buscar
El video es del 2019, tambien algo nuevo
Valentín Osorio hoy a las 12:00
Entregable_1_Checklist.pdf
124.40 KB
Fredy Orozco hoy a las 13:06
https://youtu.be/6_v3IbhNOXc
YouTube
OpenWebinars
QUÉ ES UN DATA SCIENTIST
Gabriel Vázquez
¿QUÉ ES UN DATA SCIENTIST?
Enviar mensaje a #entregable-1
EN LÍNEA — 10
Andrés Monsalve
Jugando a Google Chrome
Atlas BOT
Jugando a a!help • atlas.bot
Fredy Orozco
GitBot BOT
Viendo GitHub Universe
Hydra BOT
Escuchando .help
MEE6 BOT
Santiago Castro
SergioBz
Statbot BOT
Viendo s?help | statbot.net
Valentín Osorio
entregable-1 - Disc...
Entregable 1 Check...
8:32 p. m.
21/03/2022


Discord
2022 - 1s
general-requisitos
foro-1
material
entregable-1
CANALES DE VOZ
Sala de estudio 1
Valentí... EN DIRECTO
Andrés Monsalve
Fredy Orozco
Sala de estudio 2
BASES-DE-DATOS
ANALISIS
Screen 1
Voz conectada
Sala de estudio 1 / 2022 ...
Vídeo
Pantalla
San Valen
#5304
Sala de estudio 1
Pantalla de Valentín Osorio
720p 30FPS
EN DIRECTO
these positions require: experience in the use of data batch and streaming tools (e.g. Spark, AWS, Hadoop); understanding and experience in the use of high-performance machine learning algorithms and deep learning techniques; experience with programming languages such as Python, R, or Java;(On Understanding, pagina 1, parrafo 4)
It feels like your work to understand and people who have domain expert (Data Scientist at work, pagina 3, parrafo 3)
It's more the intellectual challenge of trying to use machine learning to answer questions from a real-world domain.(Data Scientist at work, pagina 3, parrafo 3)
I mostly tend to favor methods of supervised learning rather than unsupervised learning (Data Scientist at work,Pagina 6, parrafo 8)
I think that's also really important for working with junior people, because I want junior people always to be able to keep their eyes on the prize, and you can't do that if you don't
EN DIRECTO
Valentín ...
Fredy Orozco
Valentín Osorio
Andrés Mons...
8:02 p. m.
22/03/2022


Discord
2022 - 1s
general-requisitos
foro-1
material
entregable-1
CANALES DE VOZ
Sala de estudio 1
Valentí... EN DIRECTO
Andrés Monsalve
Fredy Orozco
Santiago Castro
Sala de estudio 2
BASES-DE-DATOS
ANALISIS
Screen 1
Voz conectada
Sala de estudio 1 / 2022 ...
Vídeo
Pantalla
San Valen
#5304
Sala de estudio 1
HAS
INTERPRETS
APPEARS
VARIETY
VELOCITY
QUALITY
INFORMATION
HAS
HAS
AMOUNT
VALIDITY
EN DIRECTO
9:18 p. m.
23/03/2022

BACK

REFERENCES

#1. On Understanding Data Scientists, Sebastian Gutierrez

ISBN: 978-1-7281-6901-9.

#2. Data Scientists at Work, Paula Pereira, Jacome Cunha, Joao Paulo Fernandes

ISBN: 9781430265993.

#3. The Data Science Design Manual, Steven S. Skiena

ISBN: 978-3-319-55443-3.

#4. ¿Qué hace un Data Scientist?

https://www.youtube.com/watch?v=Bl2sBiVdZHs&ab_channel=Platzi

#5: QUÉ ES UN DATA SCIENTIST.

https://youtu.be/6_v3IbhNOXc