IBM AI Ethics(mod5)

Module 5

What is transparency?

Meet the team

One of the largest banking companies in the world is testing an AI system to automate their mortgage approval process. Abi, the Business Owner, is impressed with the system so far and has asked to see key details about the system to better understand it. Abi's request makes the team realize that the key details are not clearly documented and therefore not straightforward to understand. This story covers what the team did to increase transparency and create a standard for documenting details, from the sources of data to how the system makes predictions and recommendations.

Identify the issue

The company is in the phase of championing digital transformation using data and AI. Also, they've made significant investments in hiring AI experts to pioneer their AI innovation journey.

The company's AI team is excited to explore ideas for automating their mortgage approval process using AI. Within a month, the team developed a proof-of-concept system for automating the mortgage approval process. Stacey, the Head of AI, presents the proof of concept to the Business Owner, Abi.

Stacey concludes the presentation by saying that the team was determined to create an AI system that is trustworthy. To ensure this, they took the following actions:

- Evaluated the model's fairness

- Added explainability functions

- Tested for robustness against adversarial attacks

- Documented all data privacy measures

Stacey then turns towards Abi expecting him to approve the AI system.

Abi starts by saying, "I want to thank you and the team for the hard work in developing this proof of concept. I am so excited and looking forward to the new AI journey of our company.

"However, I am curious to know more about the data source, how we transformed the data, how all personally identifiable information was handled, the model performance, and other evaluation results. Can you provide more details on the entire process and on the decisions made?"

Stacey sends a message to Daniel and asks for the details requested.

Daniel sits down at his computer and starts to look at the files. While skimming through the details, he notices that each team member had their own way of recording their process and their findings. None of the details were logged consistently. Daniel forwards the link to the documentation hesitantly because he knows the information is not easily digestible.

As Stacey and Abi start reviewing the details, they realize the need to standardize the documentation process.

Abi concludes, "We can fix this! Let's connect with Patricia to discuss how we can increase transparency in the process and create a standard way for documenting details to enhance governance."

Explain the issue

Patricia, Stacey, and Abi gather in a meeting room. Patricia sets up a whiteboard session.

Patricia starts by writing down the high-level goals of transparency and governance.

Patricia points out that they need to ensure the process covers the organization's focus areas in building trust in AI systems. To refresh the team's memory, she mentions that the focus areas
are **fairness**, **robustness**, **explainability**, **transparency**, and **privacy**.

Patricia begins, "Let's try a simple exercise. Let's create a simple framework to help us understand what key details about the system are available and who can provide them."

Stacey and Abi are enthusiastic about the exercise and are ready to start right away.
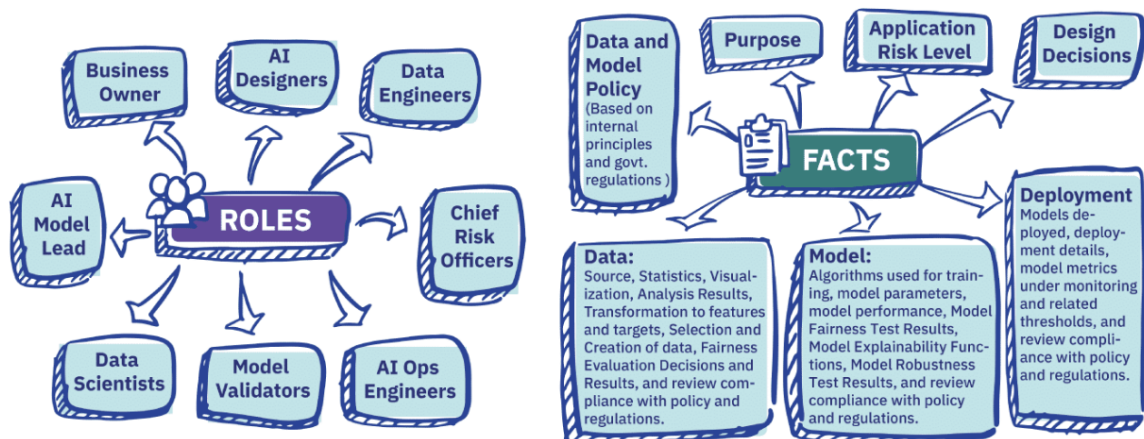
"Alright," says Patricia, "let's get started. Stacey, can you provide me with a list of roles that participated in creating our AI system? And Abi, can you provide me with a list of facts that the team needs to collect?"

Both Abi and Stacey start writing down the information. After a while, Patricia goes to the whiteboard, and divides the board into two sections. She asks Abi and Stacey to add the gathered details to get a holistic view.

Note: For the purposes of the story, things are done quickly, but in reality the process should be more deliberative.

The team looks at the board and feels motivated to see things finally coming together.
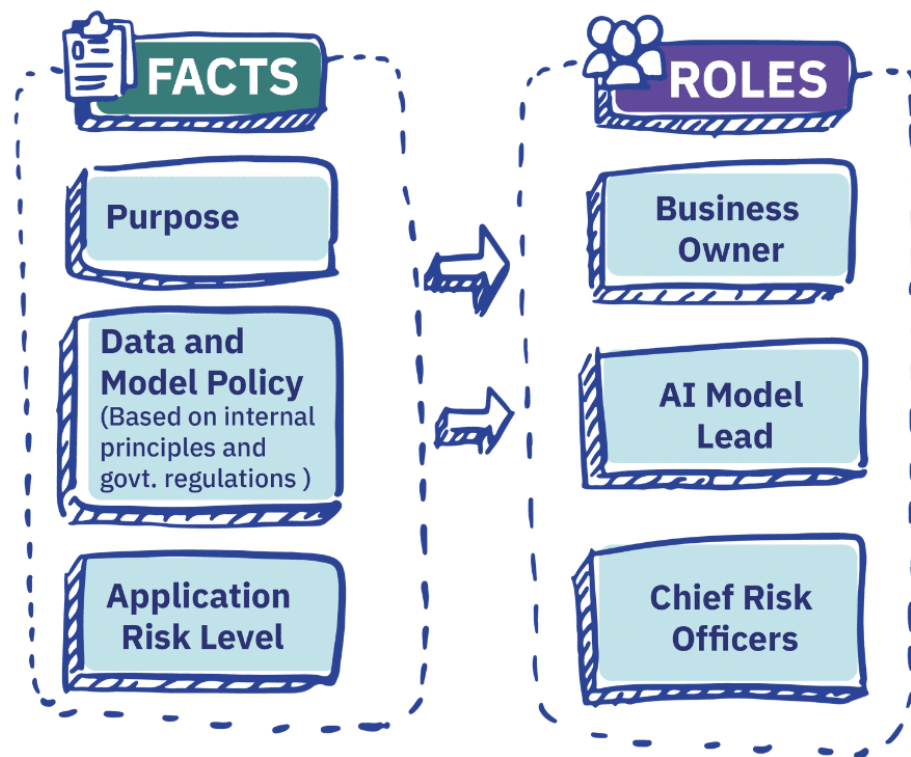
"This is great," enthuses Patricia. "Now, let's map the **roles** to the **facts**. By doing this mapping, we can clarify the responsibilities of the people involved in the process."

FACTS

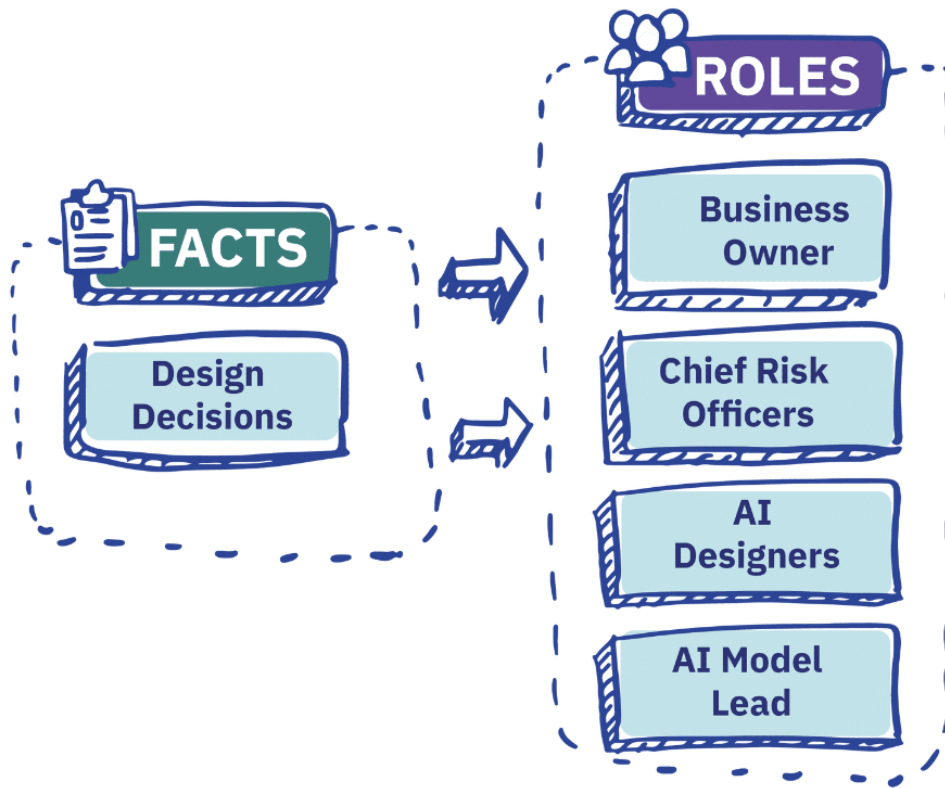- Purpose
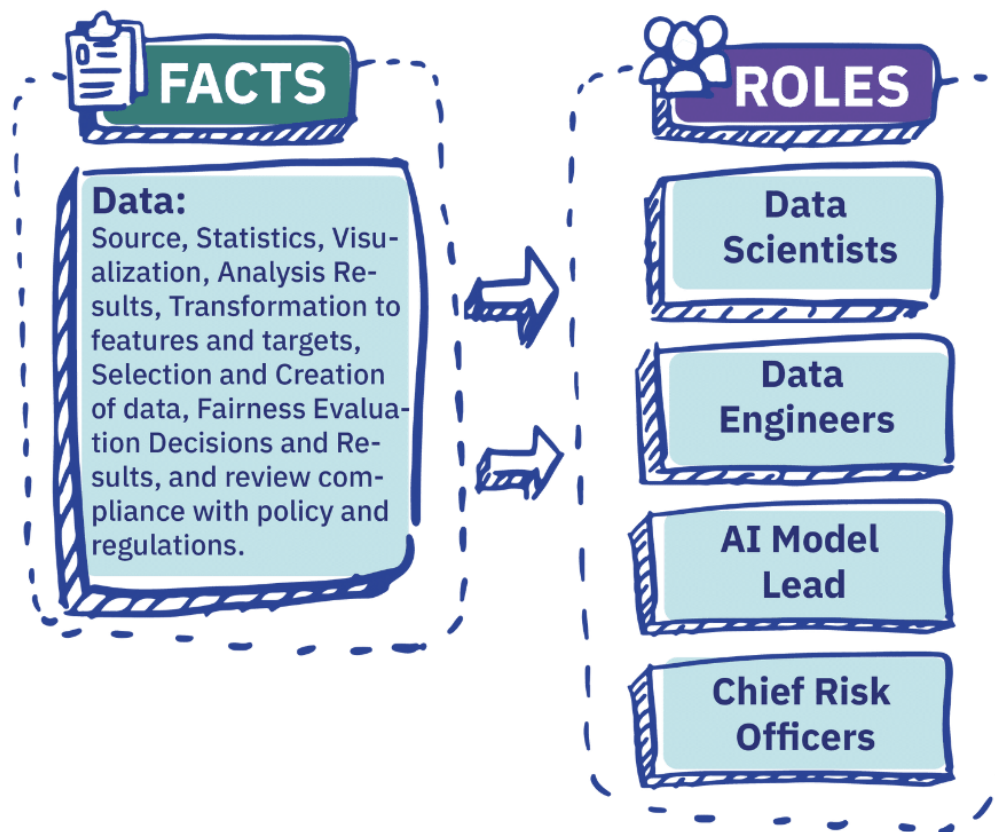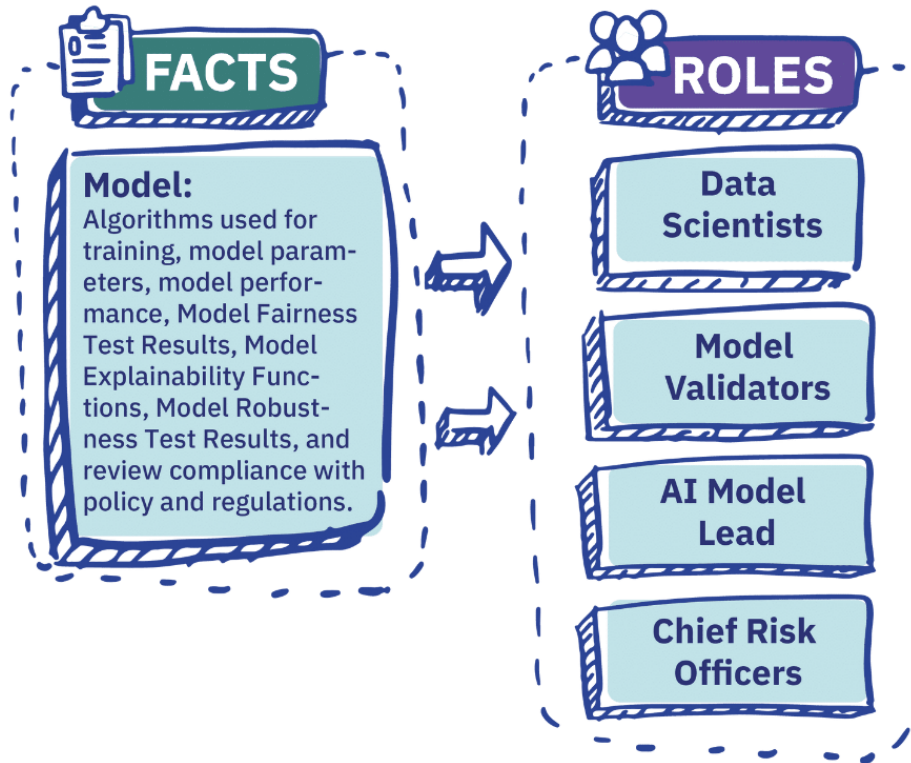- Data and Model Policy (Based on internal principles and govt. regulations)
- Application Risk Level

ROLES

- Business Owner
- AI Model Lead
- Chief Risk Officers

FACTS

Design Decisions

ROLES

Business Owner

Chief Risk Officers

AI Designers

AI Model Lead

FACTS

Data:
Source, Statistics, Visualization, Analysis Results, Transformation to features and targets, Selection and Creation of data, Fairness Evaluation Decisions and Results, and review compliance with policy and regulations.

ROLES

Data Scientists

Data Engineers

AI Model Lead

Chief Risk Officers

FACTS

Model:
Algorithms used for training, model parameters, model performance, Model Fairness Test Results, Model Explainability Functions, Model Robustness Test Results, and review compliance with policy and regulations.

ROLES

Data Scientists

Model Validators

AI Model Lead

Chief Risk Officers

After mapping the facts to the roles, Patricia and the team create a template for logging the facts at every pipeline stage. As a next step, they decide to try the overall flow and finalize what needs to be improved.

Stacey and Abi are happy and relieved that they were able to spot the standardization need at the right time! They walk out of the meeting with a clear plan for improving documentation and enhancing transparency.

Address the issue

Transparency is one of the keys to creating trustworthy AI. Going forward, the team will document what data is collected, how it is used, how it is stored, and who has access to the data. This transparency will allow questions about the AI system to be answered easily and quickly.

Reflection: Transparency in AI

Imagine that you are part of the team trying to maximize the transparency of an AI system. Think about the following questions, then take a few minutes and use the space provided to record your answers. (Writing an answer is a good way to process your thoughts. These answers are for your use only. You have the option to download your response and save it. It will not be saved in the text box when you move on in the course.)

For question 1:

How can you determine the roles involved in developing and deploying the AI system?

It is important to define roles before starting to work on the project or after the business discussion stage. Remember that building a diverse and inclusive team — including a diverse community of stakeholders — helps to build systems that are more trustworthy.

For question 2:

What types of facts can each role involved in the process provide?

Defining all of the different roles and all of the different facts needed, and then mapping the roles to facts they can provide, is a very important exercise in working toward transparency. For example, facts about the purpose of the system can be provided by the Business Owner, while facts about the data, such as the sources of the data or fairness evaluations, could be provided by Data Scientists, Data Engineers, the AI Model Lead, or even the Chief Risk Officer.

For question 3:

Who do you think should be responsible for ensuring compliance with the whole process?

Ensuring compliance is a big task! Often, a chief compliance officer or other leader oversees compliance and depends on a team or committee to enable and enforce compliance across the business. But compliance is a team sport, and everyone can take actions to help their organization maintain compliance.