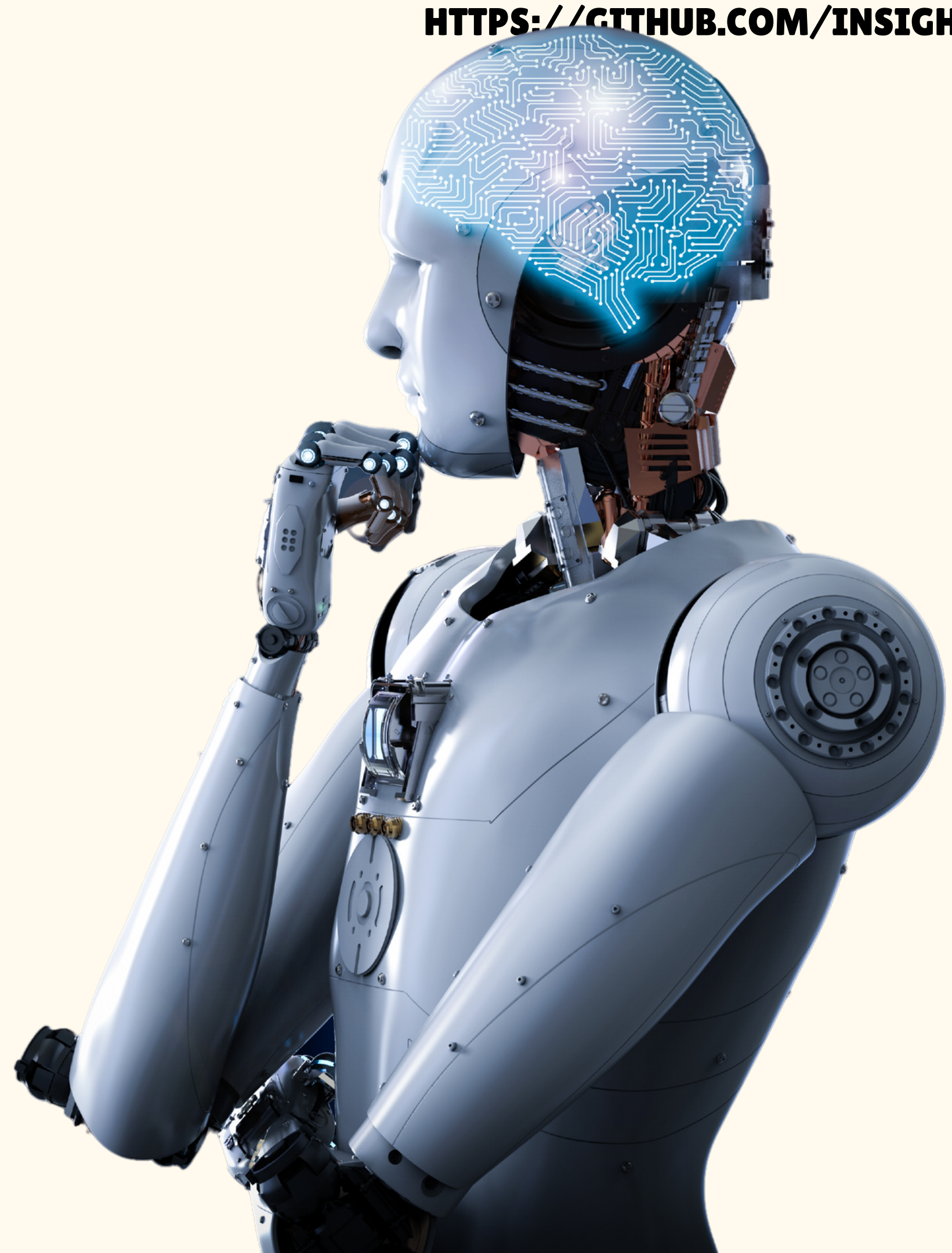


# TEST-DRIVING HUGGING FACE AGENTS & TOOLS



**HOW AGENTS &  
TOOLS  
SUPERCHARGE HF  
TRANSFORMERS**



# WHAT ARE TRANSFORMERS AGENT + TOOL



**AGENT :** A LARGE LANGUAGE MODEL SERVED OVER THE INFERENCE END POINT. THE AGENT IS PROVIDED WITH A PROMPT AND SET OF TOOLS TO EXECUTE THE CODE. AGENT OUTPUTS PYTHON CODE THAT IS EXECUTED ARBITRARILY

**TOOLS:** SINGLE PYTHON FUNCTION THAT IS ALREADY WRITTEN, OR CUSTOM WRITTEN BY THE USER

**SUPPORTING MODELS :** OPEN SOURCE MODELS THAT SUPPORT VARIOUS ACTIVITIES RANGING FROM IMAGE GENERATION, AUDIO GENERATION, TRANSLATION, AND EVEN DOCUMENTQA

IMPLEMENTATION OF CHAT BASED AND INDEPENDENT SESSION AGENTS

[HTTPS://GITHUB.COM/INSIGHTBUILDER](https://github.com/insightbuilder)

# CHALLENGE SOLVED

- **USING OPENSOURCE LLMS (OPEN ASSISTANT / STAR CODER) TO CONTROL THE TOOLS**
- **USE YOUR OWN LLMS TO WORK WITH AGENTS AND TOOLS**
- **USING OTHER OPEN SOURCE MODELS FOR COMPLETING THE USER TASKS**
- **FLEXIBILITY TO WRITE CUSTOM TOOLS AND PROMPTS & ASSIGN IT AS PART OF THE LLM TOOLBOX**
- **KEEPING YOUR DATA SAFE AS YOU CAN IMPLEMENT YOUR INFERENCE ENDPOINTS WITH OPEN SOURCE LLMS**
- **INTEGRATION WITH LANGCHAIN IS EXPECTED IN FUTURE VERSIONS**
- **APPS CREATED WITH OPEN SOURCE MODELS CAN BE UPDATED IN A CONTROLLED MANNER**



# TOOL BOX OF AGENTS : OPEN SOURCE MODELS

- **DOCUMENT QUESTION ANSWERING:** GIVEN A DOCUMENT (SUCH AS A PDF) IN IMAGE FORMAT, ANSWER A QUESTION ON THIS DOCUMENT (DONUT)
- **TEXT QUESTION ANSWERING:** GIVEN A LONG TEXT AND A QUESTION, ANSWER THE QUESTION IN THE TEXT (FLAN-T5)
- **UNCONDITIONAL IMAGE CAPTIONING:** CAPTION THE IMAGE! (BLIP)
- **IMAGE QUESTION ANSWERING:** GIVEN AN IMAGE, ANSWER A QUESTION ON THIS IMAGE (VILT)
- **IMAGE SEGMENTATION:** GIVEN AN IMAGE AND A PROMPT, OUTPUT THE SEGMENTATION MASK OF THAT PROMPT (CLIPSEG)
- **SPEECH TO TEXT:** GIVEN AN AUDIO RECORDING OF A PERSON TALKING, TRANSCRIBE THE SPEECH INTO TEXT (WHISPER)
- **TEXT TO SPEECH:** CONVERT TEXT TO SPEECH (SPEECHT5)
- **ZERO-SHOT TEXT CLASSIFICATION:** GIVEN A TEXT AND A LIST OF LABELS, IDENTIFY TO WHICH LABEL THE TEXT CORRESPONDS THE MOST (BART)
- **TEXT SUMMARIZATION:** SUMMARIZE A LONG TEXT IN ONE OR A FEW SENTENCES (BART)
- **TRANSLATION:** TRANSLATE THE TEXT INTO A GIVEN LANGUAGE (NLLB)

# AGENT PROMPT STRUCTURE

AGENT PROMPT IS STRUCTURED FOR CUSTOMISATION. THE PROMPT IS STRUCTURED BROADLY INTO FOUR PARTS.

A. **INTRO**: HOW THE AGENT SHOULD BEHAVE, EXPLANATION OF THE CONCEPT OF TOOLS.

B. **TOOL BOX**. THIS IS DEFINED BY A <<ALL\_TOOLS>> TOKEN THAT IS DYNAMICALLY REPLACED AT RUNTIME WITH THE TOOLS DEFINED/CHOSEN BY THE USER.

C. A SET OF **EXAMPLES** OF TASKS AND THEIR SOLUTION

D. **CURRENT EXAMPLE**, AND REQUEST FOR SOLUTION.

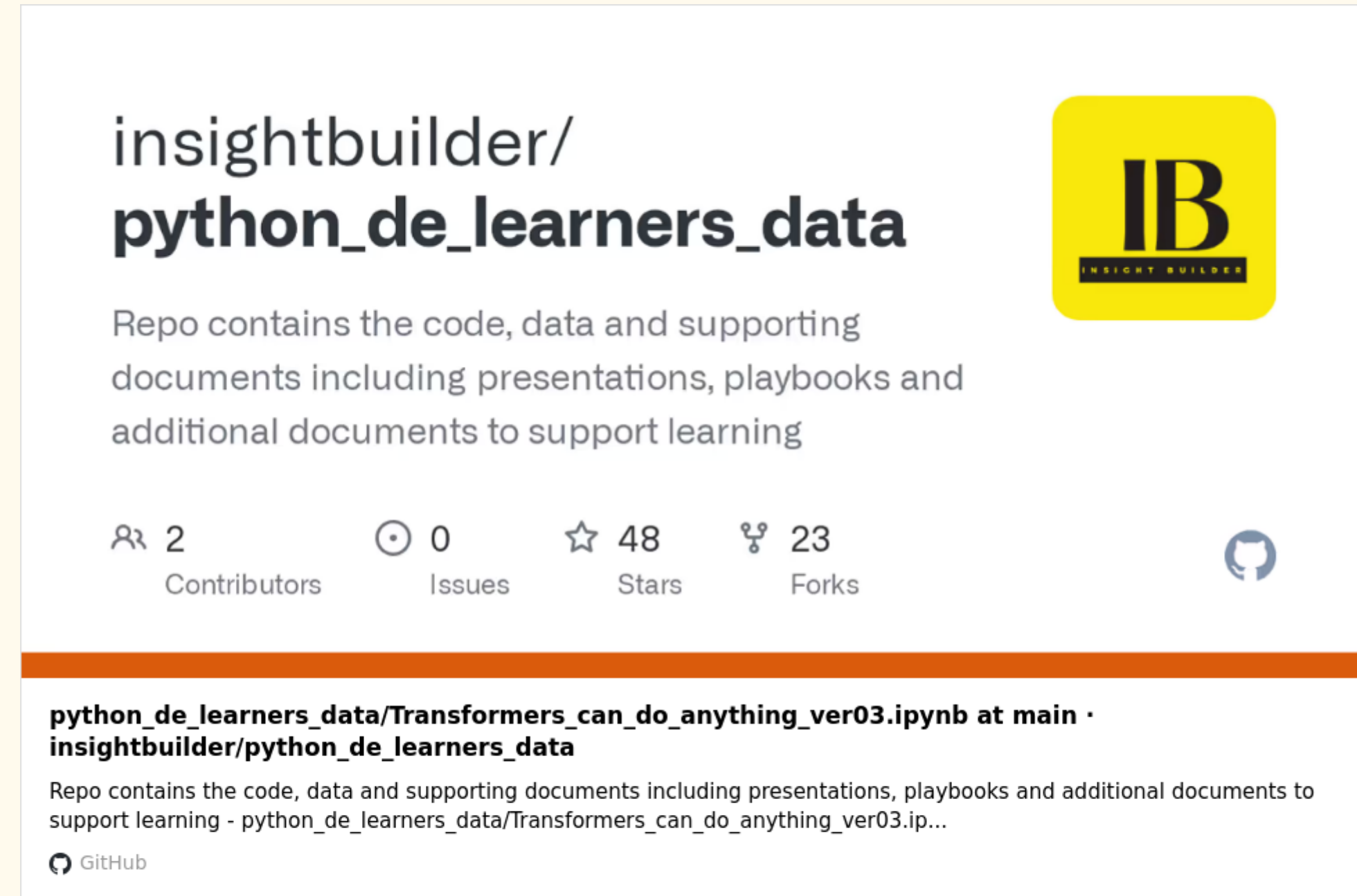
## TOOL PROMPT STRUCTURE

A. **TOOL NAME**: PRECISE NAME OF THE TOOL SO AGENT CAN IDENTIFY

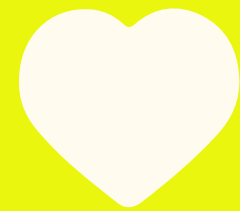
B. **TOOL DESCRIPTION**. PRECISELY DESCRIBE THE TASK THE TOOL CAN PERFORM. THE TASK IS EXPLAINED FIRST, FOLLOWED BY THE INPUT ARGUMENTS & RETURN VALUES

# LETS WORK WITH AGENTS

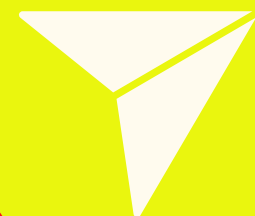
1. INITIALIZE THE AGENTS ON COLAB
2. UNDERSTAND WHAT AGENTS ARE DOING
3. EXECUTE SIMPLE TASKS LIKE IMAGE GENERATION, DOCUMENT QA
4. DEVELOP CUSTOM TOOLS AND INTEGRATE WITH AGENTS
5. FURTHER IDEAS ON DEVELOPING WITH OPEN SOURCE MODELS



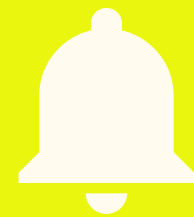
# THANKS FOR WATCHING



**LIKE**



**SHARE**



**SUBSCRIBE**