



# Building a HealthAI Chatbot to Expand Patient Support

Adding storytelling to healthcare

Luca Bianchi (He/Him)

AWS Serverless Hero  
CTO @ Neosperience Health

# Storytelling in healthcare

# What is storytelling?

**“Storytelling is everything that remains when you pull the facts out of reality”**

**Alessandro Baricco**

writer, director and performer

# Why does it matter in healthcare?

# Storytelling is fundamental to understanding what we don't know

- A medical test result is just a fact.
- Without a proper explanation from a doctor (the storytelling), It does not exist (do I need to be worried about this test?).
- A doctor explains to the rest of us facts related to our health.
- Making them real.

# Healthcare is broken

## In the pandemics aftermath

- Healthcare systems have been stressed beyond their limits.
- Hospitals are overwhelmed by delays in diagnosis and treatments.
- MDs are understaffed, needing more doctors to recover from the delays.





## Nevertheless...

- A good understanding of clinical reports is crucial for patients.
- Understanding and continuous follow up means more adherence to therapy
- Clinical protocols for diagnosis require a second opinion in double blind, which means doubling clinical reports for the same patient.
- Poor adherence to therapy, delays in diagnosis, and lack of follow-up are a recipe for healthcare failure.

# Using AI to support doctors

# AI in medicine

- A conversational interface can offer a 1:1 experience to patients
- The tone of voice could be tailored to the patient's profile
- Clinical data can be analyzed and sent to doctors with insights
- The human in the loop is fundamental make a diagnosis
- AI can help to tailor the response to patients

# Chatbots “could be” the perfect interface to patients

However, traditional chatbots are:

- too naive
- and constrained to a conversation flow
- cannot adapt the response to the patient

Large Language Models (LLM) to complement chatbots to create powerful conversational interfaces

# AlexaTM 20B

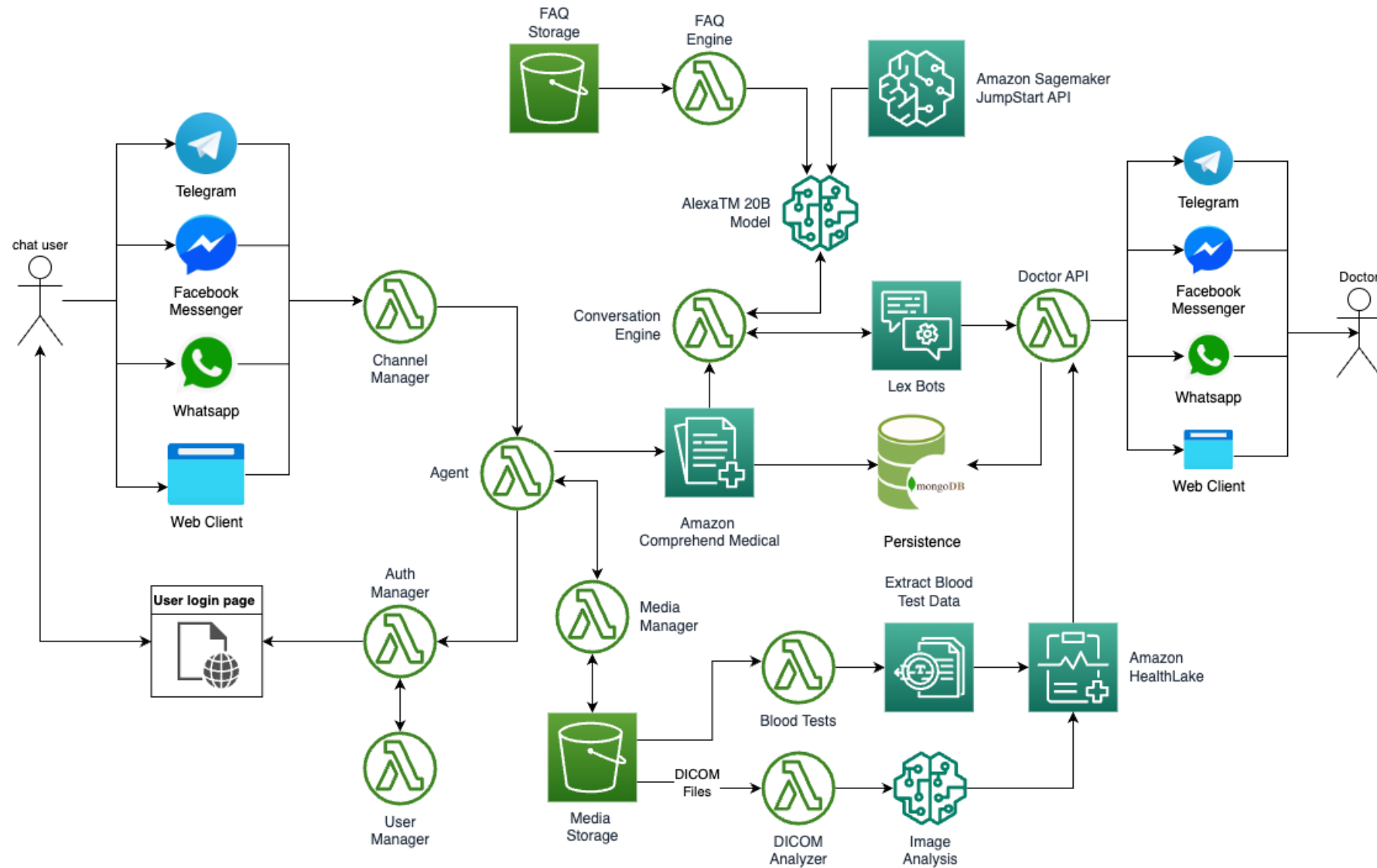


# AlexaTM 20B Model

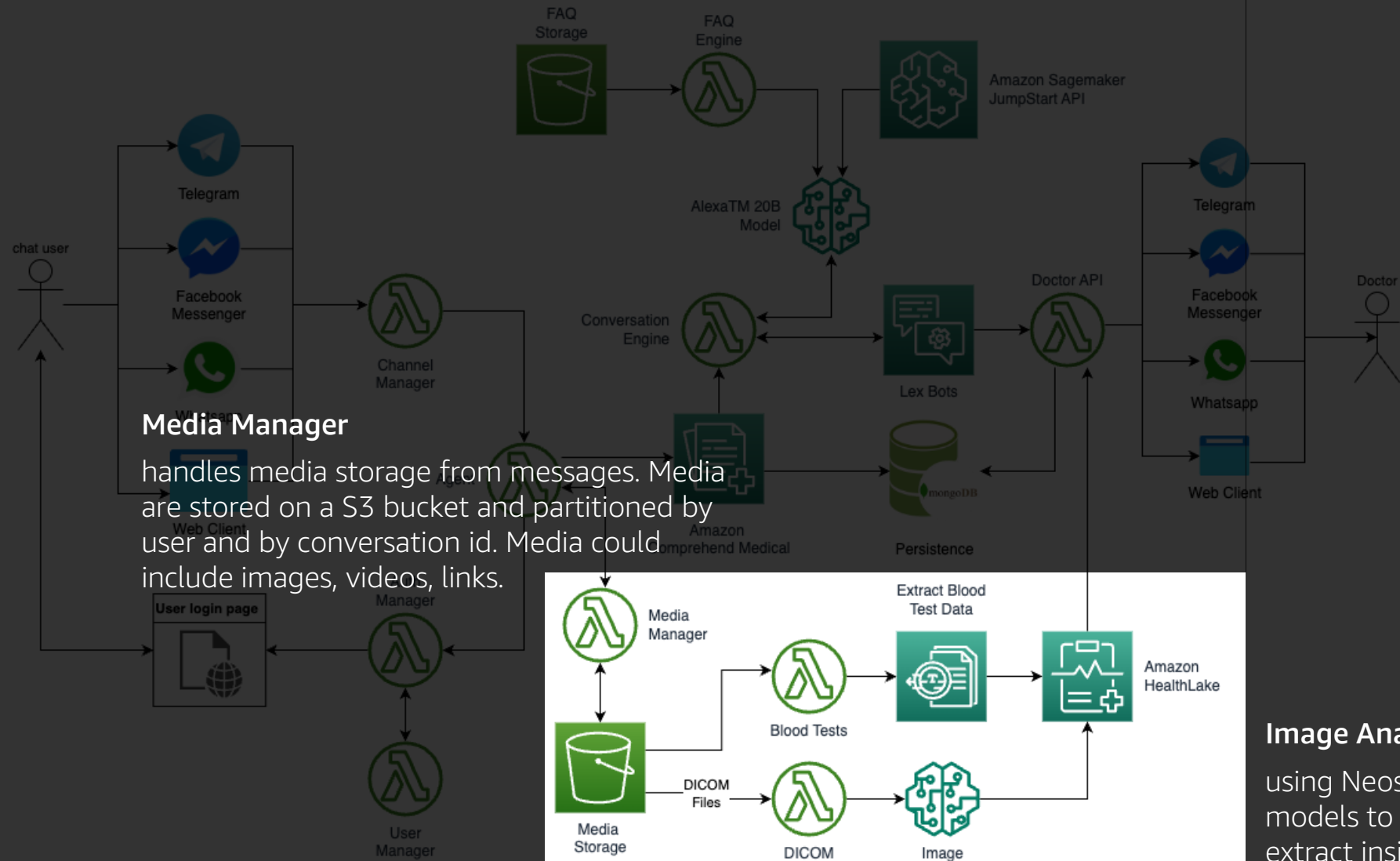
- Alexa Teacher Model with 20 Billion is an LLM with performances comparable to bigger models such as GPT-3 or LaMbDa.
- Encoder-decoder architecture, rather than decoder only.
- Excels other large language models on few-shot tasks.
- Suitable for summarization and machine translation.
- Learns new concepts and transfers knowledge with minimal human input

# Smart Omnichannel Functional Interactive Agent (SOFIA)

# HEALTH AI CHATBOT TO EXPAND PATIENT SUPPORT







**Media Manager**

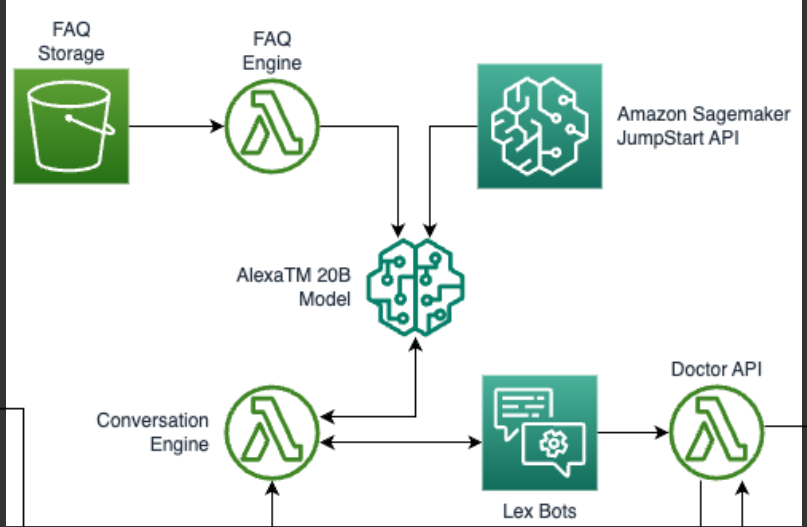
handles media storage from messages. Media are stored on a S3 bucket and partitioned by user and by conversation id. Media could include images, videos, links.

**Image Analysis Model**

using Neospirence Health models to analyze images and extract insights.

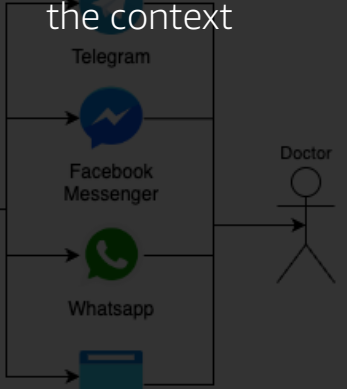
### FAQ Engine

receives formatted JSON lines files from user, stores it into a cloud local storage (S3) then uploads the file (properly formatted to AlexaTM to be usable by the chatbot



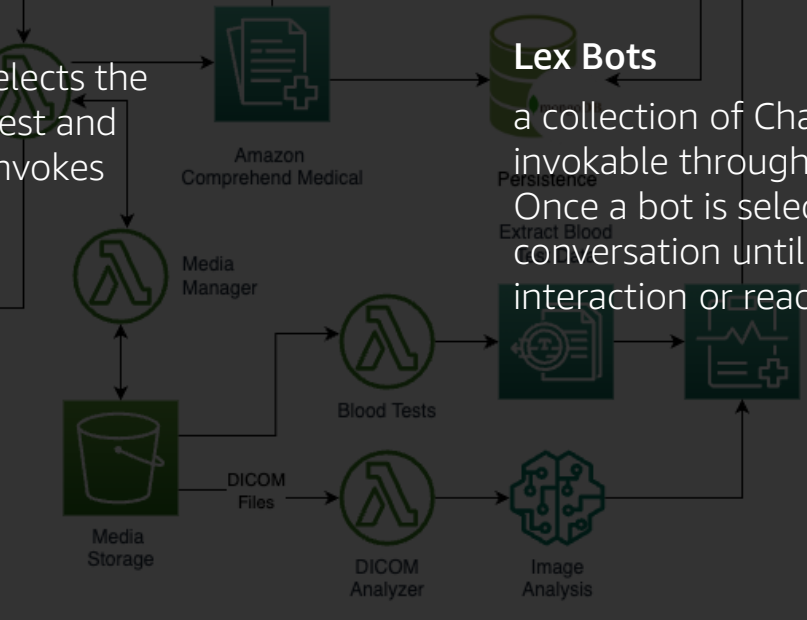
### AlexaTM 20B

implements FAQ capabilities and free text search. The engine is used also to answer to the user while maintaining the context



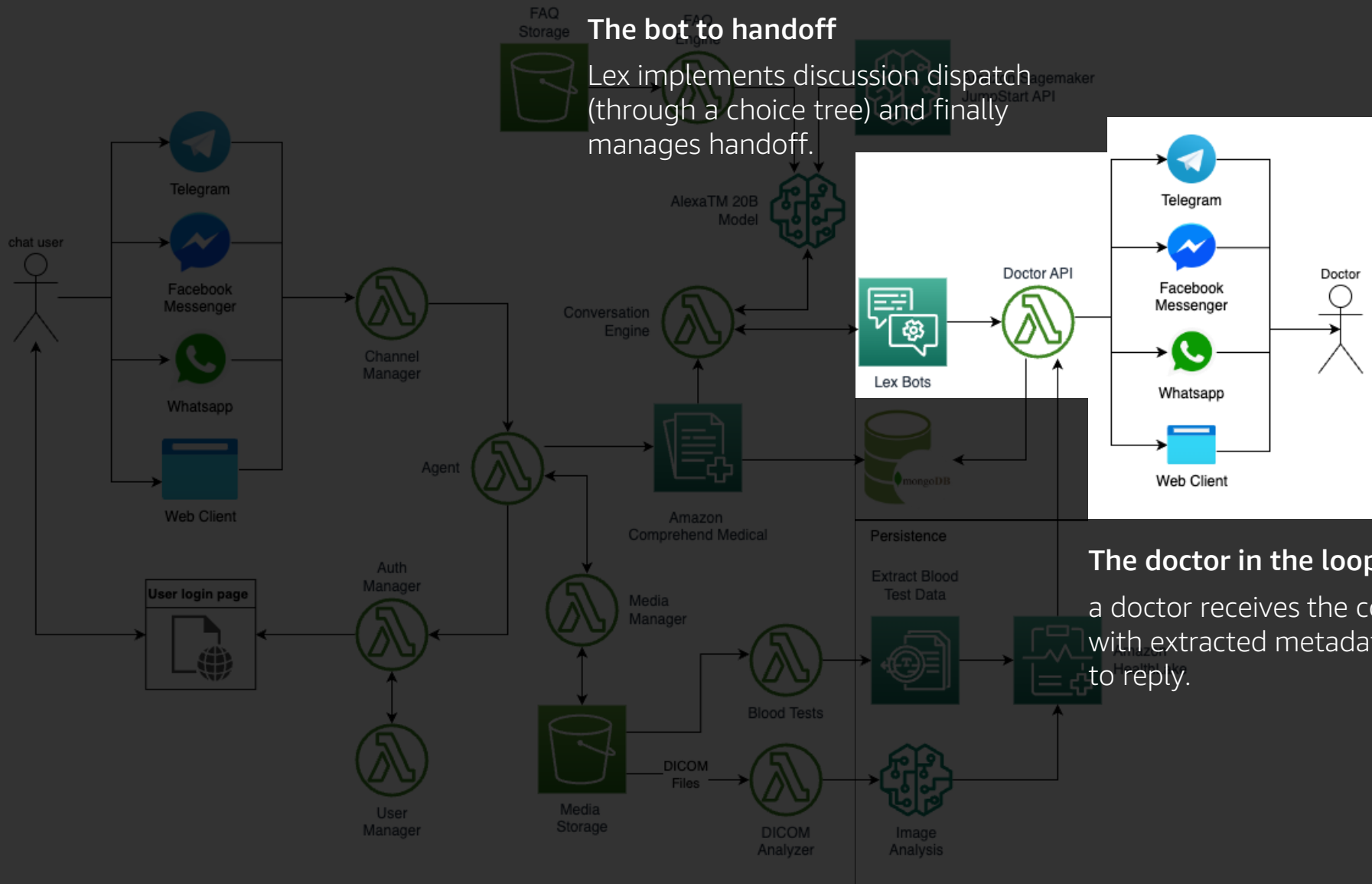
### Conversation Engine

handles bot conversation, Selects the right bot to handle the request and selects the conversation. It invokes Lex Bot



### Lex Bots

a collection of Chatbots deployed to AWS and invocable through an API Call to be selected. Once a bot is selected, it is attached to a given conversation until the user terminates the interaction or reaches the TTL.



**The bot to handoff**

Lex implements discussion dispatch (through a choice tree) and finally manages handoff.

**The doctor in the loop**

a doctor receives the conversation with extracted metadata and is able to reply.

## Deploying an AlexaTM 20B

- Easy to be deployed with Amazon Sagemaker JumpStart.
- JumpStart APIs offer support to automated deployment.
- AlexaTM 20B inference runs on g4dn.12xlarge instances.
- Average costs for inference is as low as 4\$/h

## Wrap up

- AI can help MDs and hospitals to overcome practitioners shortage.
- Lex can engage patients in 1:1 conversations.
- Amazon Comprehend Medical can detect relevant topics and sentiments from messages.
- AlexaTM 20B can tailor doctors' responses to patients' behavioral profiles.
- AlexaTM can be leveraged even in restricted or regulated contexts because data remains in your account.



# Thank you!

Luca Bianchi

Twitter: @bianchiluca

LinkedIn: <https://www.linkedin.com/in/lucabianchipavia/>