

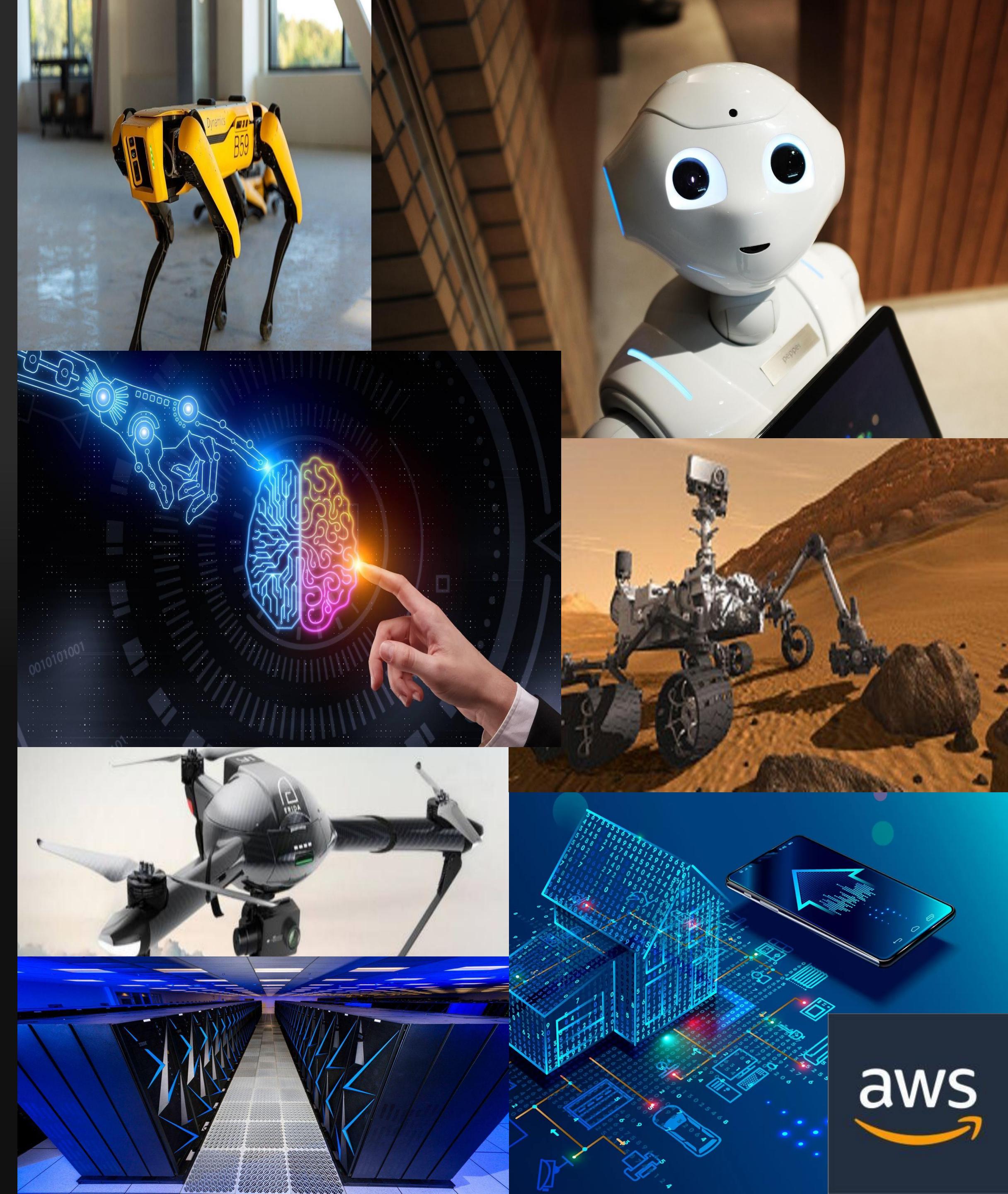
Programación distribuida

Con un enfoque a emprendimiento

Sesión V: Serverless e IA en Amazon web services

Leonardo Camargo Forero, Ph.D

2021





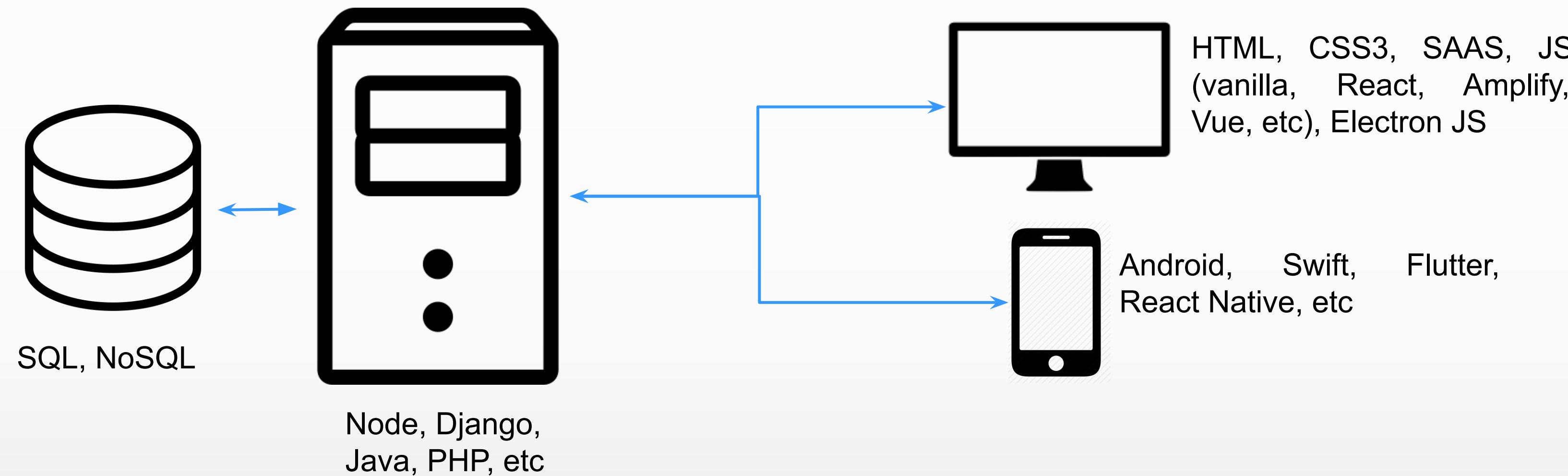
Agenda

- Serverless con AWS
- IA con AWS
- Conclusiones
- Nuevo plan de la clase
- Para la próxima clase

Serverless con AWS

El enfoque tradicional - Server

Un sistema web o una app



Escalabilidad?

- Incremento de usuarios
- Nuevas características

Administración?

- Instalación de servicios
- Resolución de errores
- Actualizaciones
- Soporte

Costo?

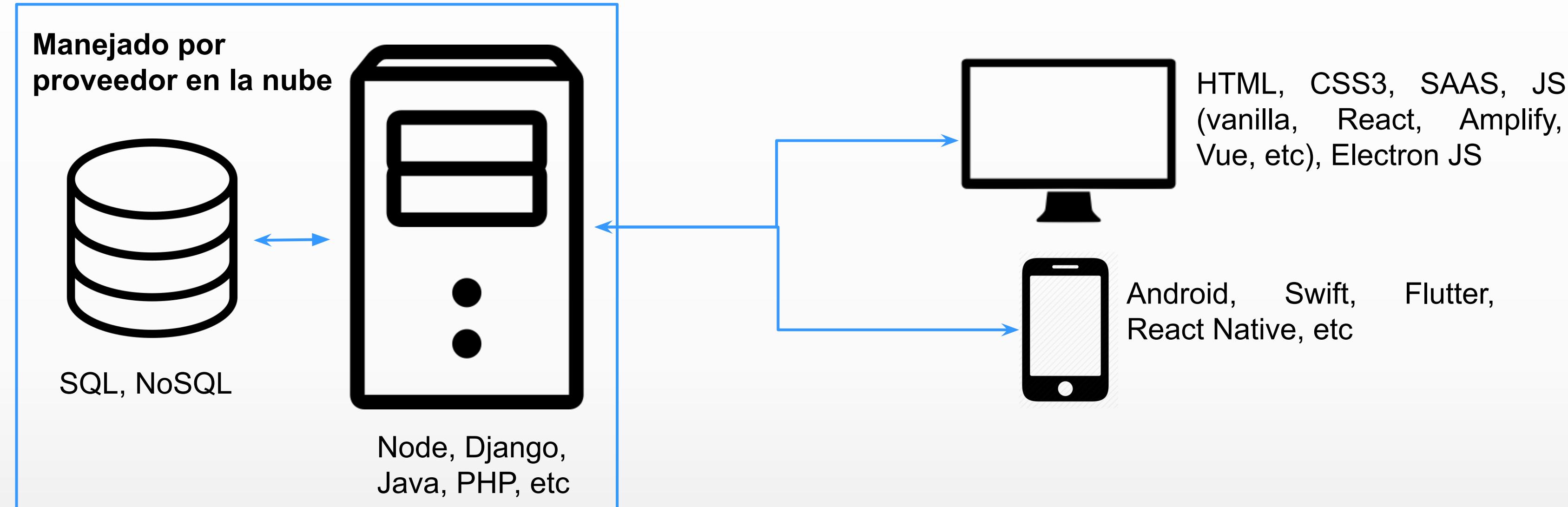
- Hardware
- Personal de administración
- Arquitecto
- Personal de backend
- Personal de frontend

Tiempo de lanzamiento?

- Adquisición e instalación de Hardware
- Arquitectura software
- Backend + frontend

Serverless

Un sistema web o una app



Escalabilidad?

- Auto scale
(contenedores, etc)

Administración?

- Solo si se desea

Costo?

- Pay as you go
- Arquitecto
- Personal de frontend

Tiempo de lanzamiento?

- Arquitectura software
- Frontend

Permite a los desarrolladores enfocarse en la lógica de negocio

Amazon web services



Amazon Web Services

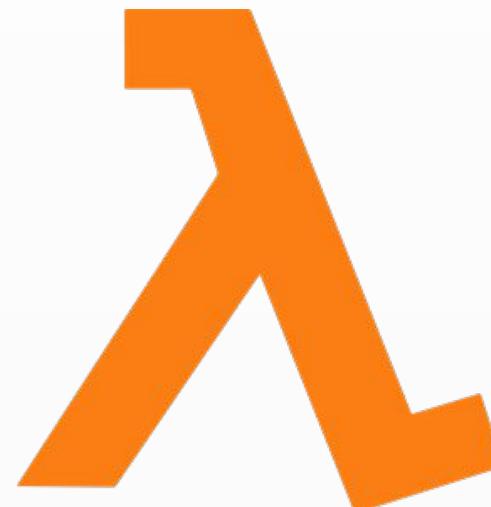
- Costos (servidores, administración, personal)
- Escalabilidad
- Seguridad
- Resiliencia
- Multitud de servicios: Más de 200 en 2021

AWS Certification Salaries

Certification	2019 Average Salary
AWS Certified Solutions Architect – Professional	\$148,456
AWS Certified DevOps Engineer	\$137,724
AWS Certified Solutions Architect - Associate	\$130,883
AWS Certified SysOps Administrator – Associate	\$130,610
AWS Certified Developer – Associate	\$130,272

AWS Serverless - algunos servicios

↑ Previous



AWS Lambda

Servicio de cómputo (funciones) sin aprovisionamiento.

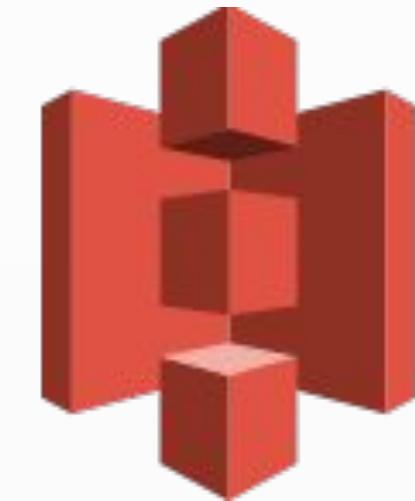
- Soporta: Node.js, Python, Go, Java, entre otros
- 1 millón de solicitudes gratis y 400000 GB/segundos(free tier)
- Tiempo máximo de ejecución: 15 minutos
- RAM máximo: 10 GB
- Hand in hand with Amazon API gateway



AWS DynamoDB

Key-value/document NoSQL database

- Más de 10 trillones de solicitudes por día
- Más de 20 millones de solicitudes por segundo
- Clientes: Lyft, Airbnb, Netflix, Disney, Samsung, Nike, etc
- Web, gaming, IoT, ad tech, retail, banking
- Junto con AppSync - real time



AWS S3

Object-based storage

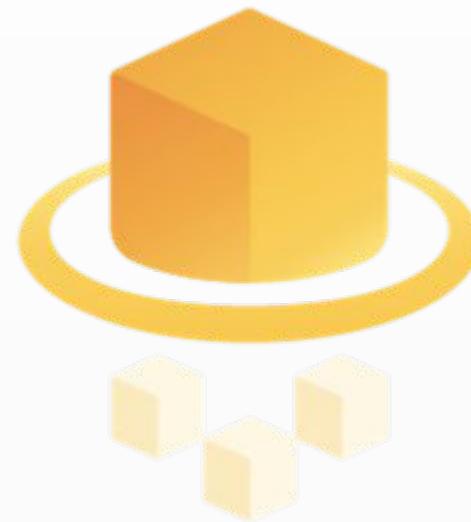
- Buckets / folders
- Data lakes
- AI core service
- 99.99999999% (11 9's) durabilidad

7

Next ↓

AWS Serverless - más servicios

↑ Previous



AWS Fargate

Elastic Container Service (ECS) / Elastic Kubernetes Service (EKS)



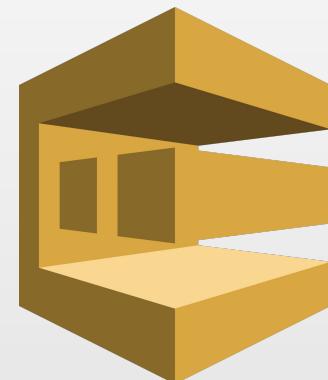
AWS EventBridge

Event bus, event-driven apps. Real-time data to AWS Lambda, etc



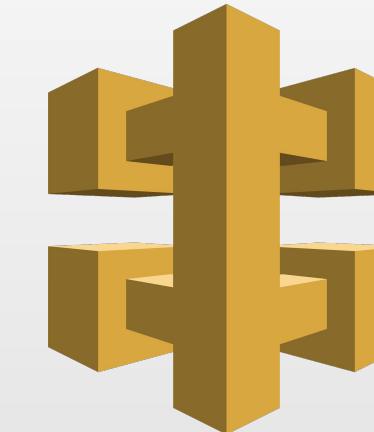
AWS Step functions

Visual workflow para orquestar AWS services, p.jem. workflow de funciones Lambda



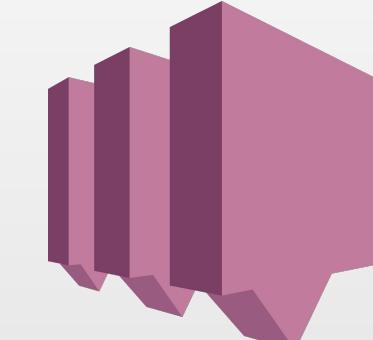
AWS Simple Queue Service (SQS)

Servicio de colas de mensajes (middleware). Microservicios, sistemas distribuidos, etc.



AWS API Gateway

RESTful y WebSocket APIs



AWS Simple Notification Service (SNS)

A2A (application-to-application) pub/sub + A2P (application-to-person) SMS, push, emails

8

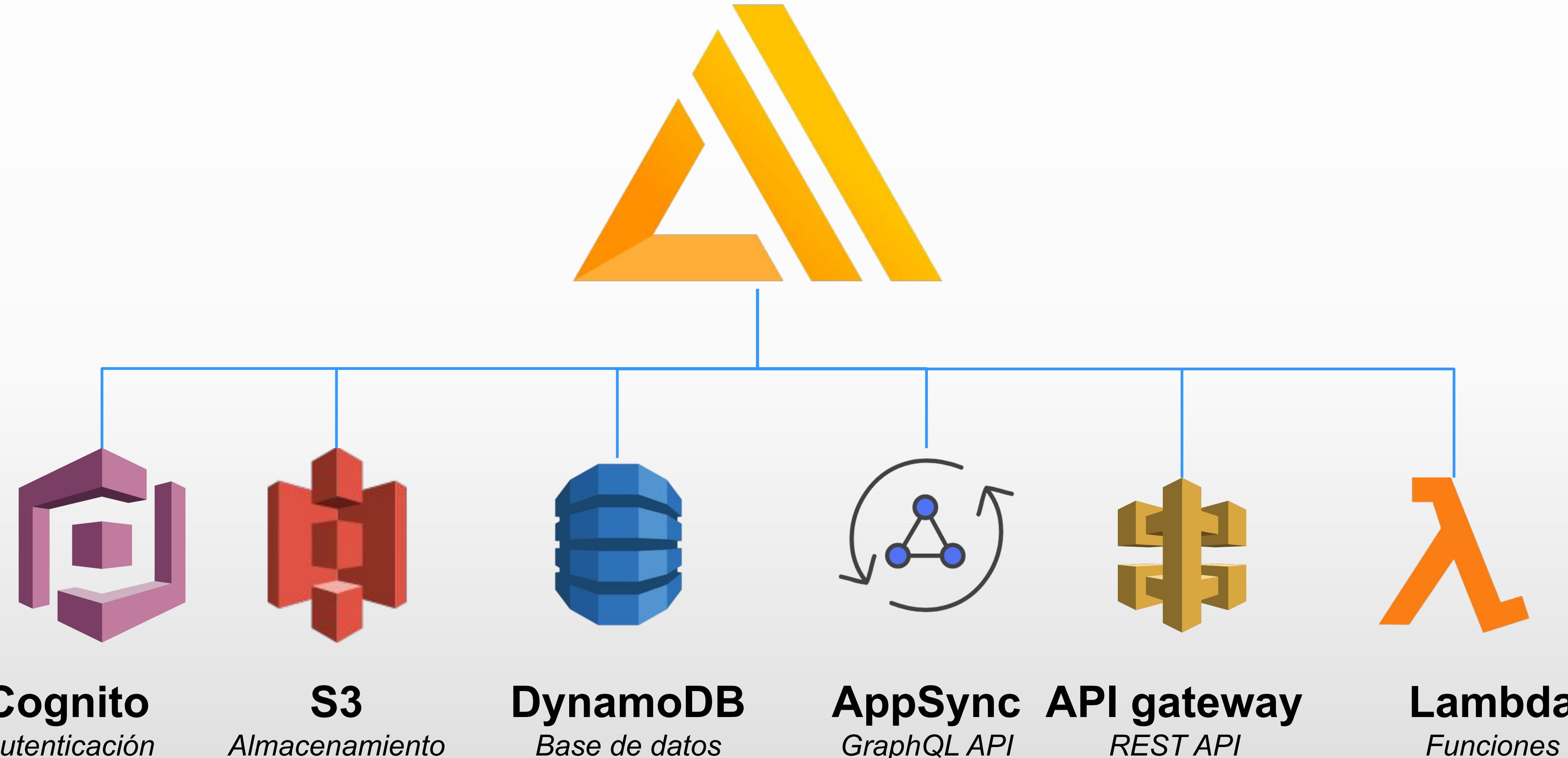
Next ↓

AWS Serverless - Amplify

↑ Previous

9

Next ↓

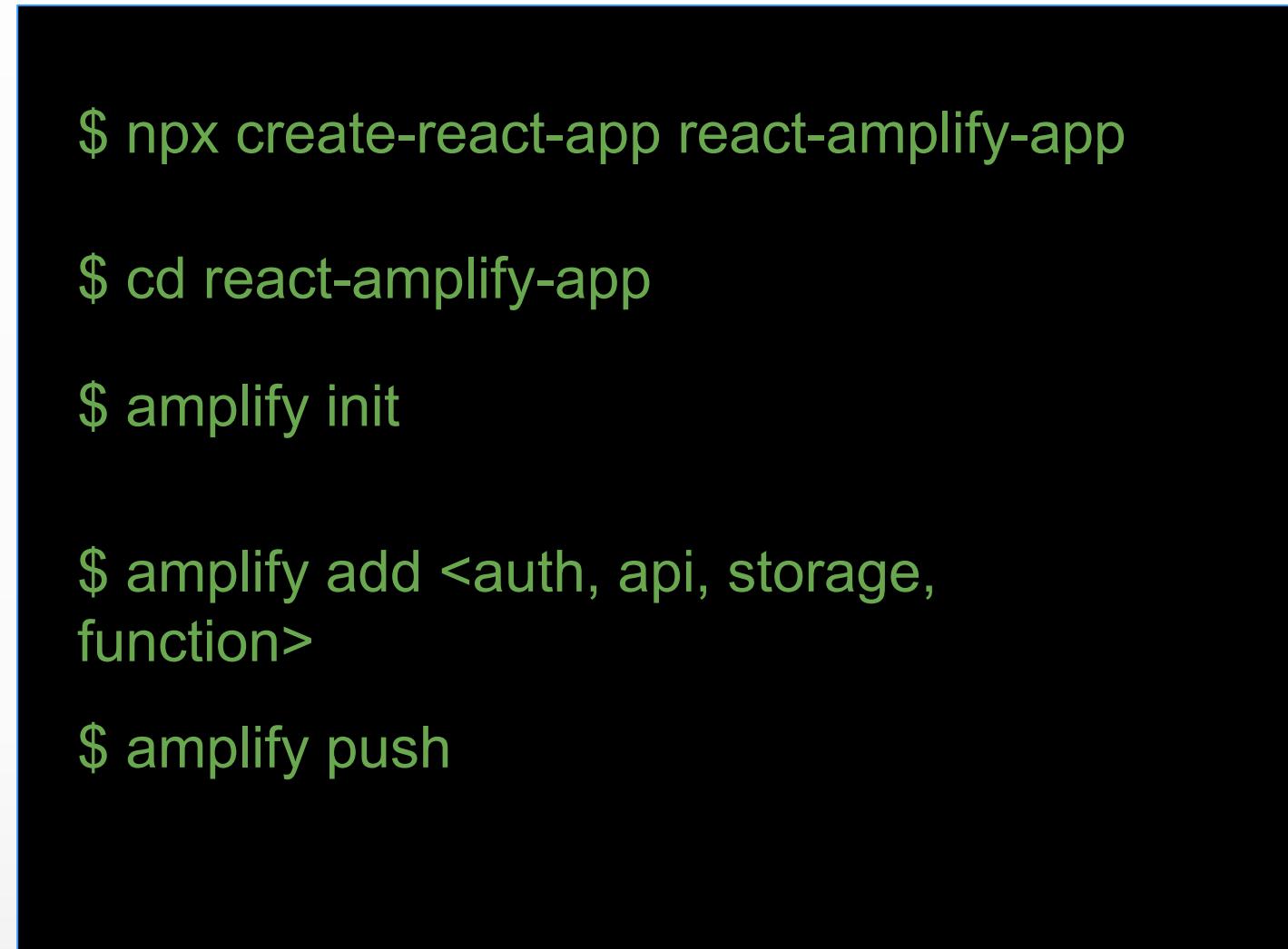


<https://aws.amazon.com/amplify/>

AWS Serverless - Amplify

1. CLI

```
$ npx create-react-app react-amplify-app
$ cd react-amplify-app
$ amplify init
$ amplify add <auth, api, storage, function>
$ amplify push
```



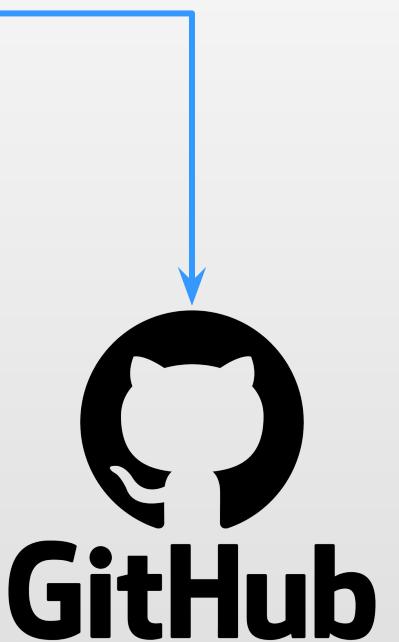
The screenshot shows the AWS Amplify Console interface for the 'react-amplify-app'. The left sidebar lists 'App settings' like General, Domain management, and Build settings. The main panel shows a preview of the React app with the URL <https://master.react-amplify-app.com>. Below the preview, the deployment status is shown as a sequence of four green checkmarks: Provision, Build, Deploy, and Verify. At the bottom, it indicates the last deployment was on 12/12/2018 at 7:32:46 AM and the last commit was an autogenerated message from Auto-build | GitHub - master.

2. Consola

3. Librerías

JavaScript, React, Angular, Vue, Next.js, and mobile platforms including Android, iOS, React Native, Ionic, Flutter.

10



AWS Serverless - Amplify

```

import React from "react";
import logo from "./logo.svg";
import "./App.css";
|
import Amplify from "aws-amplify";
import awsExports from "./aws-exports";
Amplify.configure(awsExports);

function App() {
  return (
    <div className="App">
      <header className="App-header">
        <img src={logo} className="App-logo" alt="logo" />
        <p>
          Edit <code>src/App.js</code> and save to reload.
        </p>
        <a
          className="App-link"
          href="https://reactjs.org"
          target="_blank"
          rel="noopener noreferrer"
        >
          Learn React
        </a>
      </header>
    </div>
  );
}

export default App;

```

app.js

```

type User @model @auth(rules: [{ allow: owner }]) {
  id: ID!
  name: String!
  lastname: String!
  posts: [Post] @connection(name: "UserPosts")
}

type Post @model @auth(rules: [{ allow: owner }]) {
  id: ID!
  title: String!
  content: String!
  comments: [Comment]
  @auth(
    rules: [
      { allow: owner }
      { allow: private, operations: [create, read, update, delete] }
    ]
  )
  user: User! @connection(name: "UserPosts")
}

type Comment {
  author: String!
  text: String!
}

```

schema.graphql



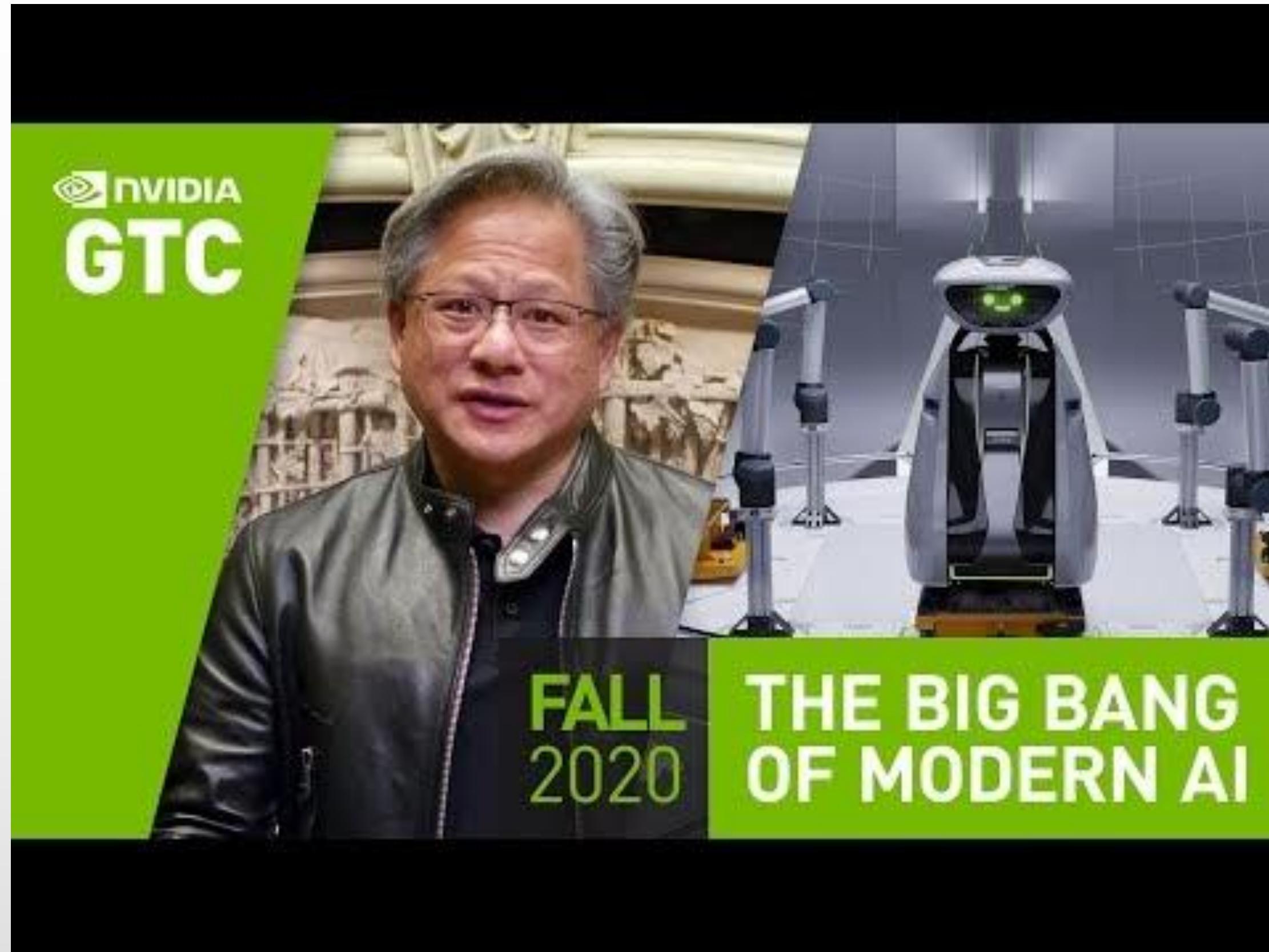
IA con AWS

Inteligencia artificial

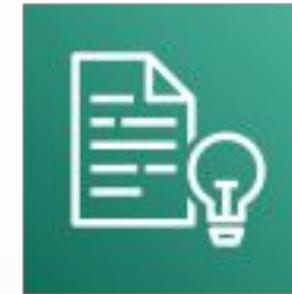
Artificial Intelligence (AI)

Artificial intelligence leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind.

<https://www.ibm.com/cloud/learn/what-is-artificial-intelligence>



AWS - servicios IA



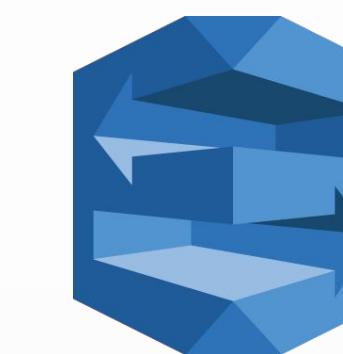
AWS Comprehend

NLP - Relaciones en texto no estructurado, referencias al lenguaje, gente, sentimientos, frases clave, etc



AWS CodeGuru

Asistente en el desarrollo de software - Revisión de código. Recomendaciones inteligentes, detección de líneas de código más costosas computacionalmente



AWS Lex

Chatbots (voz y texto) - lifelike - Alexa. Conectividad con lambda y otros servicios de AWS



AWS Forecast

Time series combinado con otros datos (seasonal behavior, etc) forecasting. La misma tecnología usada por Amazon. 50% más preciso que solo time series



AWS Textract

Extracción de texto, escritura a mano, tablas, etc



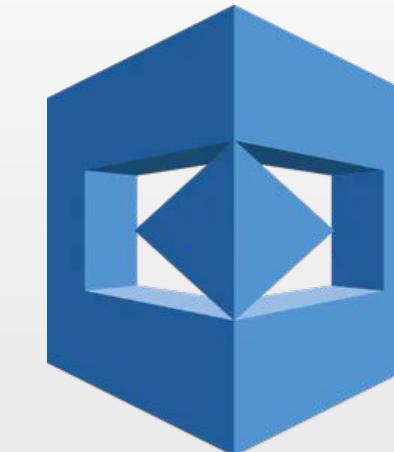
AWS Kendra

“Preguntas NLP” en vez de solo palabras para búsquedas. Reducción de resultados de búsqueda



AWS Fraud Detector

Detección de outliers en data sets. Ofrece real-time detection via API



AWS Rekognition

Imagen y video análisis. Objetos, personas, texto, escenas, actividades, contenido inapropiado, análisis facial, custom labels



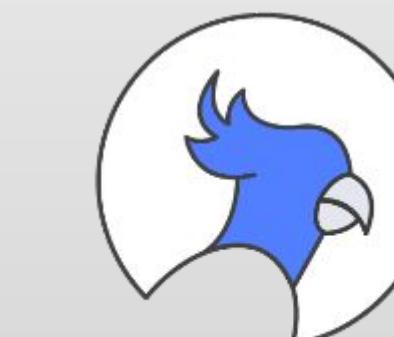
AWS Personalize

Recomendaciones en tiempo real. Recomendación de productos, direct marketing, etc.



AWS Translate

Traducción en tiempo real de documentos, sitios web. Comunicación entre usuarios multi-lenguaje



AWS Polly

Texto a voz. Múltiples voces, masculinas, femeninas y en múltiples lenguajes



AWS Transcribe

Voz a texto

Sin ningún conocimiento de IA

AWS IA - EC2



EC2



Amazon Machine Image (AMI)



Deep Learning AMI (Ubuntu 18.04) Version 48.0 - ami-0b70285e5215b80eb

MXNet-1.8.0 & 1.7.0, TensorFlow-2.4.2, 2.3.3 & 1.15.5, PyTorch-1.4.0 & 1.8.1, Neuron, & others. NVIDIA CUDA, cuDNN, NCCL, Intel MKL-DNN, Docker, NVIDIA-Docker & EFA support. For fully managed experience, check: <https://aws.amazon.com/sagemaker>

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

Select

64-bit (x86)

15

Instancias con GPU

Instancia	Cantidad de GPUs	Cantidad de vCPUs	RAM (GiB)	Precio (USD / hora)
g4dn.xlarge	1	4	16	\$0.526
g4dn.2xlarge	1	8	32	\$0.752
g4dn.4xlarge	1	16	64	\$1.204
g4dn.8xlarge	1	32	128	\$2.176
p3.2xlarge	1	8	61	\$3.06
g4dn.12xlarge	4	48	192	\$3.912
g4dn.16xlarge	1	64	256	\$4.352
p3.8xlarge	4	32	244	\$12.24
p3.16xlarge	8	64	488	\$24.48

Next

Previous

Ciencia de datos

Obtención de datos

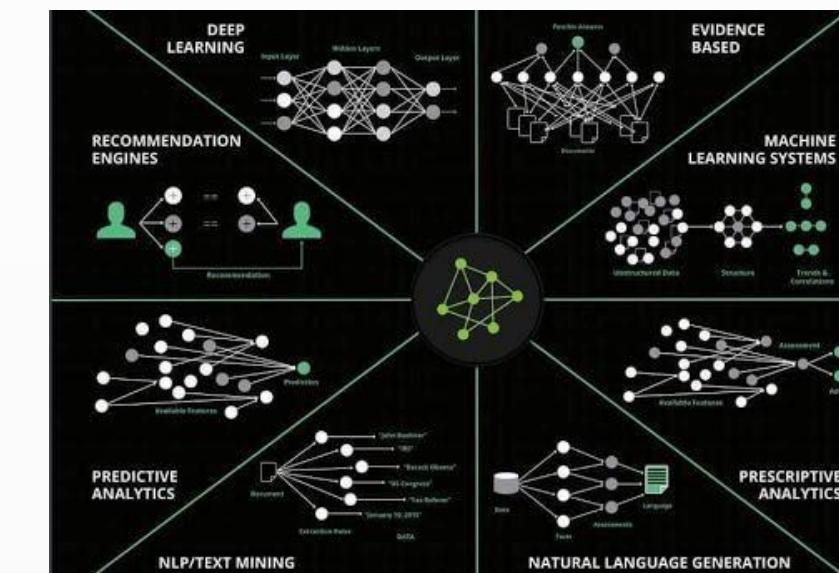
Obtención y revisión de datos necesarios para entrenamiento

S3, SageMaker, Kinesis, Glue, Batch, DMS, Step functions, etc



1

2



Entrenamiento

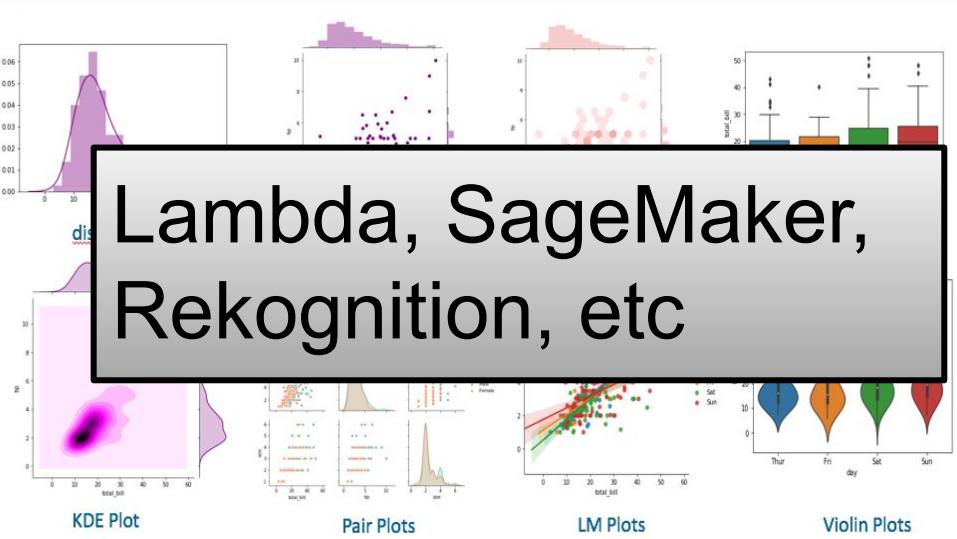
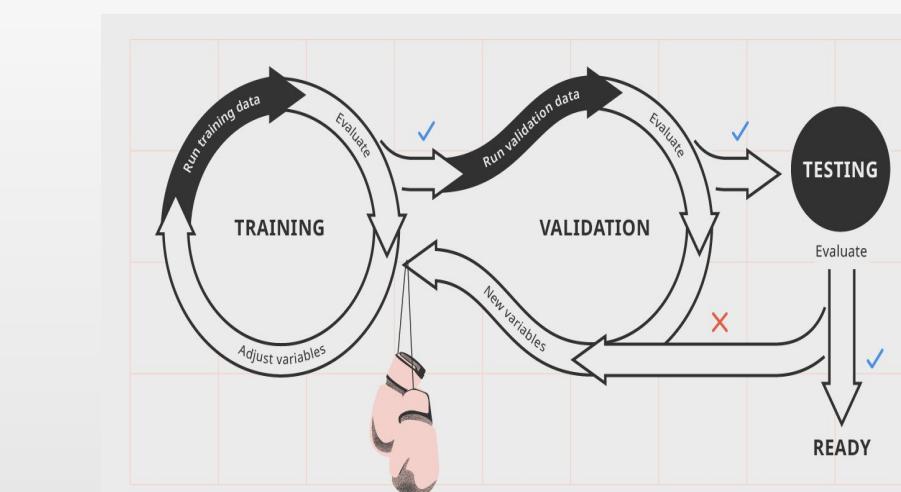
Generar el modelo más apropiado para representar los datos

3

Selección de técnica(s) de Inteligencia Artificial

- Aprendizaje supervisado o no supervisado
- Machine learning o Deep learning?

EC2, SageMaker, Rekognition, etc

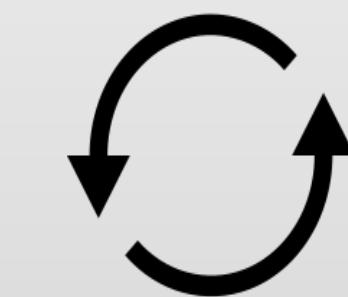


Lambda, SageMaker, Rekognition, etc

4

Despliegue de modelo de inferencia

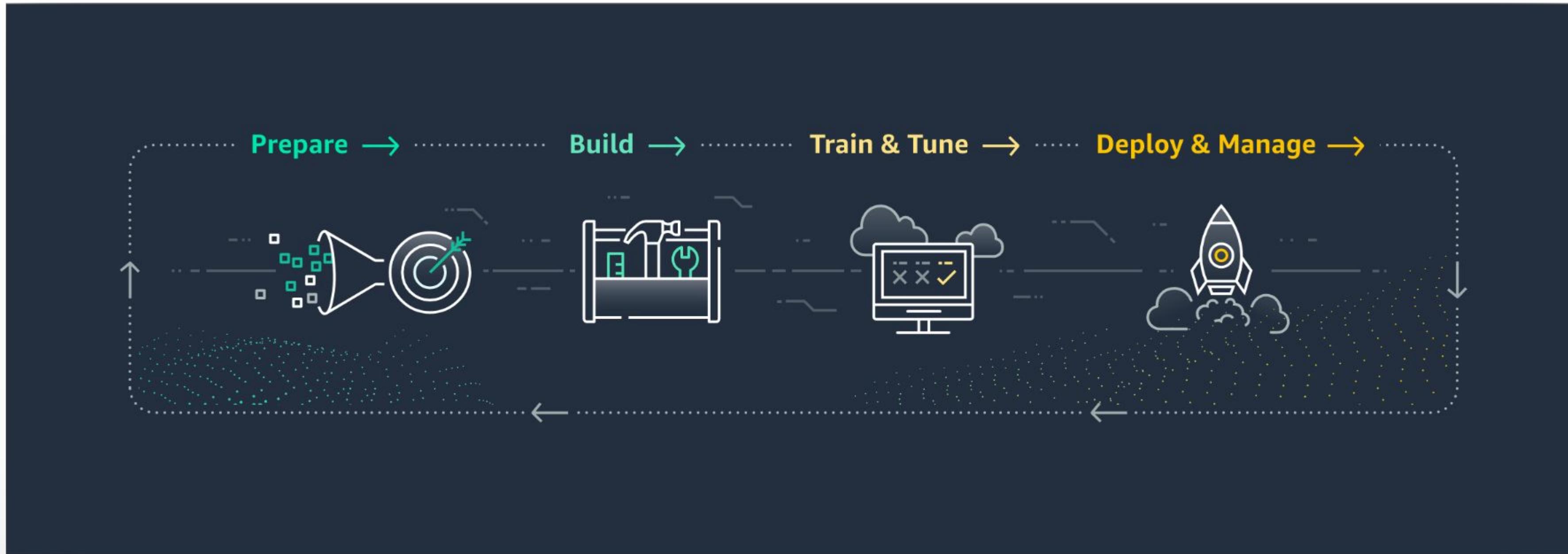
Predicción sobre nuevos datos



AWS Machine learning specialist (150K USD / año)

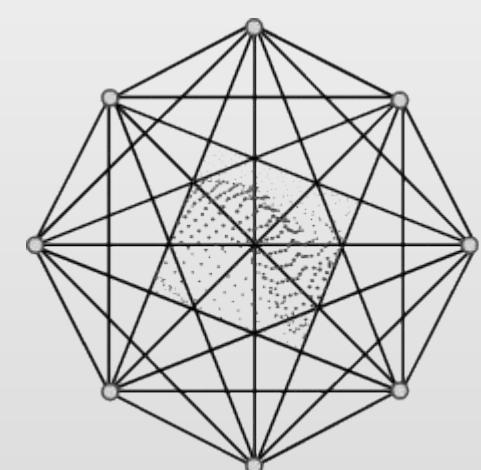
→ <https://www.udemy.com/course/aws-machine-learning/>

AWS IA - SageMaker -> científicos de datos



17

<https://aws.amazon.com/es/sagemaker/>

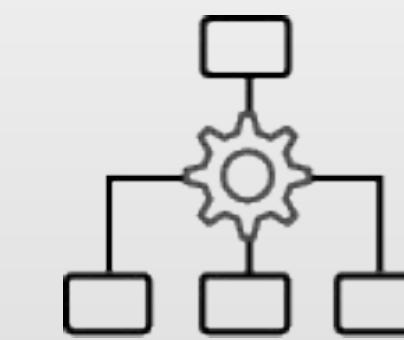


SageMaker studio: Carga de datos desde S3, etc, IDE, Notebooks, etc

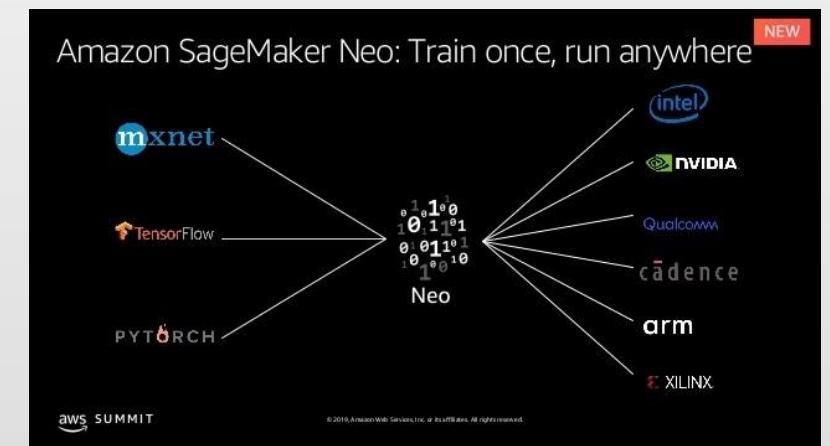
SageMaker data wrangler: Feature engineering



SageMaker clarify: Chequea el balance entre features y data y las predicciones asociadas - evitar bias - inspeccionar predicciones individuales



SageMaker pipelines: CI/CD - actualización con nuevos datos, re-entrenamientos, etc.



SageMaker NEO: Deep learning models compiler , EC2, AWS Greengrass ML, FPGA, Raspberry Pi 3, Texas Instruments' Sitara, Jetson TX1

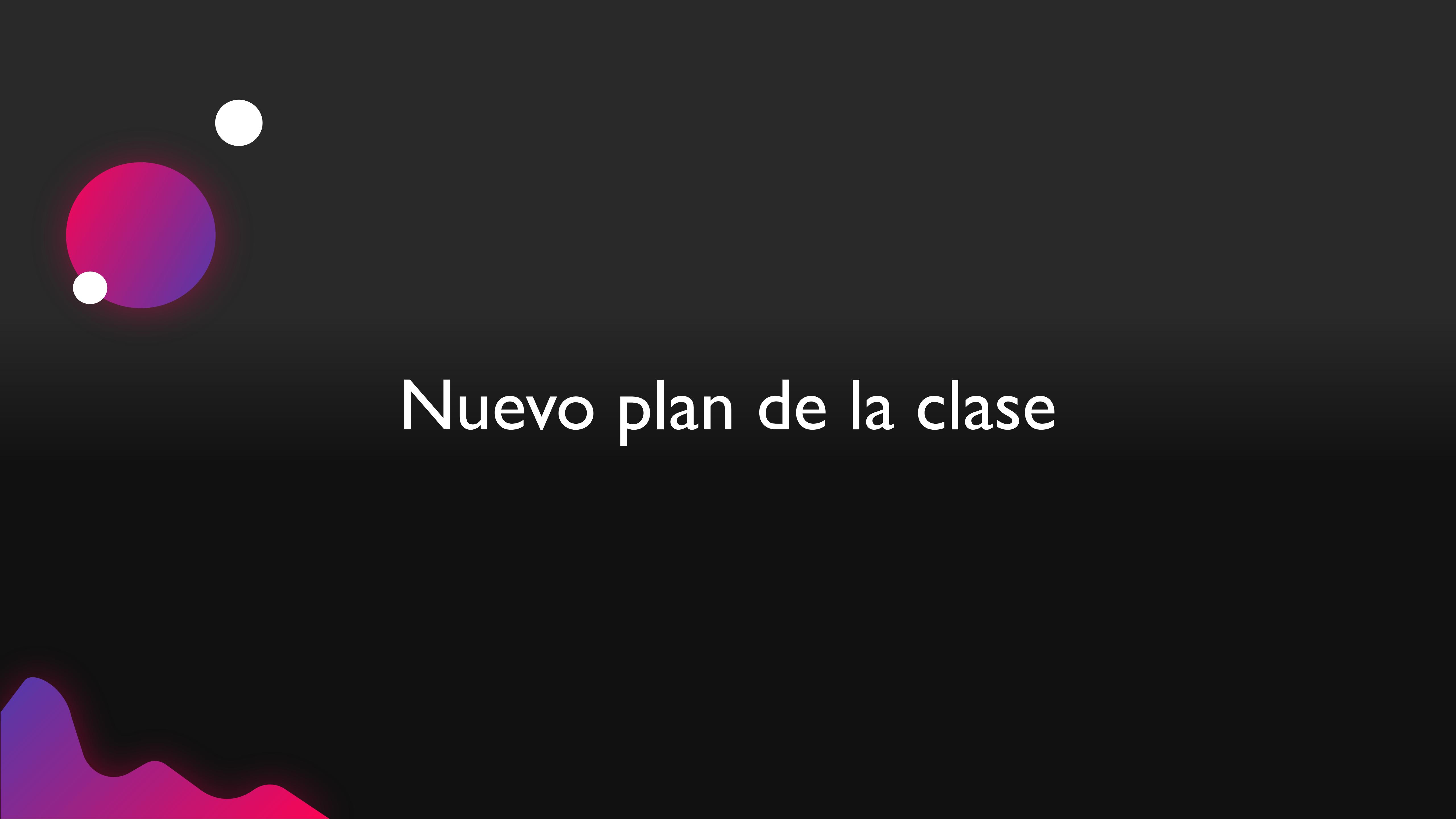
Conclusiones

Conclusiones

El uso de un enfoque serverless puede acelerar el desarrollo de una aplicación, pero reduce el control que se tiene sobre la infraestructura subyacente

El uso de servicios AI puede ofrecer todo tipo de características interesantes a cualquier producto

Es muy importante revisar los costos de estos servicios cuando se planea su utilización - cuando sea posible



Nuevo plan de la clase

Acerca de la clase



Seminarios

Supercomputación, IA, AWS, Robótica/IoT, manejo del tiempo, computación cuántica, etc.

Resumen de notas

- Canvas - 10 %
- Pitch deck V.1.0 - 10 %
- Pitch deck V.1.1 - 30 %
- MVP - 40 %
- Presentación libro: 10%

Otras cosas

- Todo en inglés
- Metodología de desarrollo de software ágil (Jira)
- Slack

Planeación

Agosto

16 - 22 Agosto



Actividades

- Martes: Seminario AWS
- Jueves: Modelo de datos preliminar

23 - 29 Agosto



Actividades

- Martes: Seminario visión (práctico) + <manejo del tiempo>
- Jueves: Revisión avances

Septiembre

22



Pitch deck v. 1.0
(10 %)

30 Agosto - 05 Septiembre



Actividades

- Martes: Pitch deck v.1.0
- Jueves: Pitch deck v.1.0



Presentación
Libro (10 %)

06 - 12 Septiembre



Actividades

- Martes: Seminario lenguaje corporal
- Jueves: Revisión avances

13 - 19 Septiembre



Actividades

- Martes: Presentación libro
- Jueves: Presentación libro

20 - 26 Septiembre



Actividades

- Martes: Práctica elevator pitch
- Jueves: Práctica elevator pitch

27 Septiembre - 03 Octubre



Actividades

- Martes: Revisión propuesta de valor, competencia y segmento de clientes
- Jueves: Revisión propuesta de valor, competencia y segmento de clientes

Next

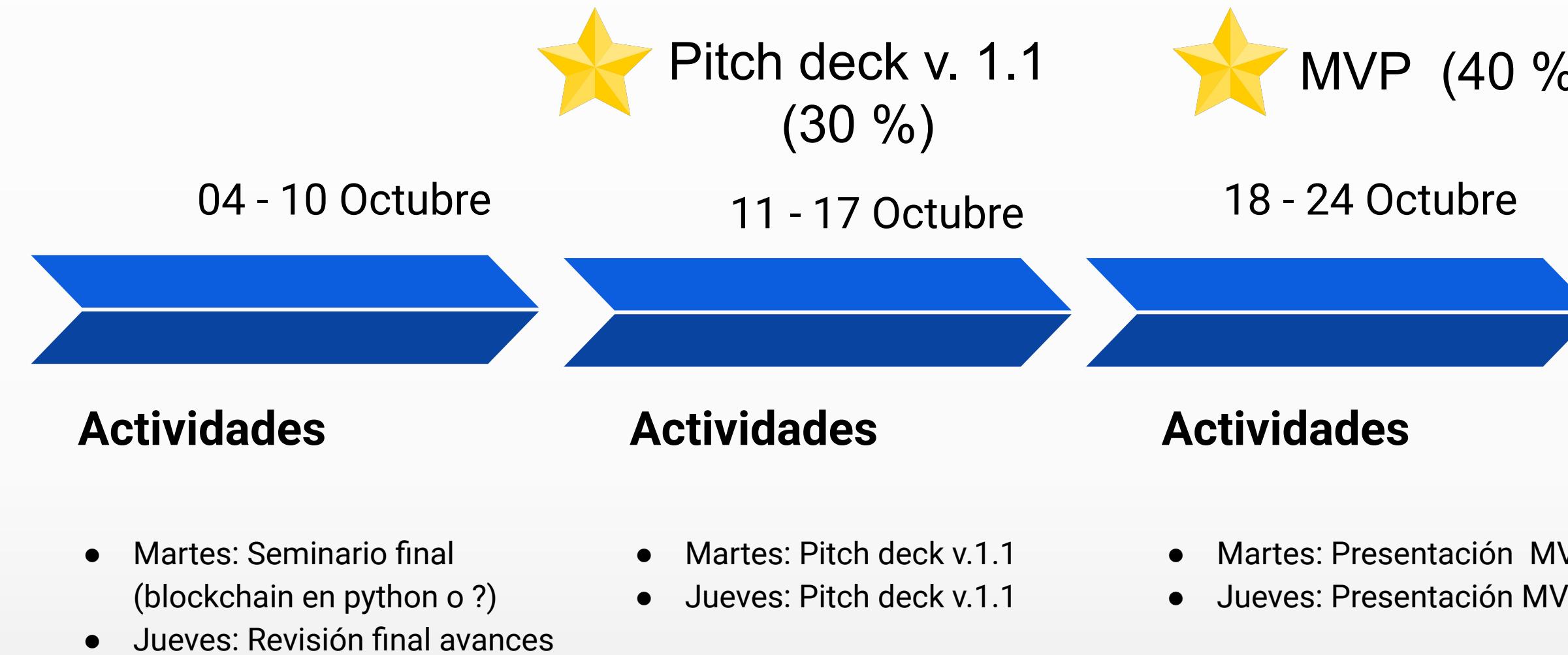
Previous

Octubre 22: Finalización clases
Octubre 29: Último día para las evaluaciones finales

Planeación

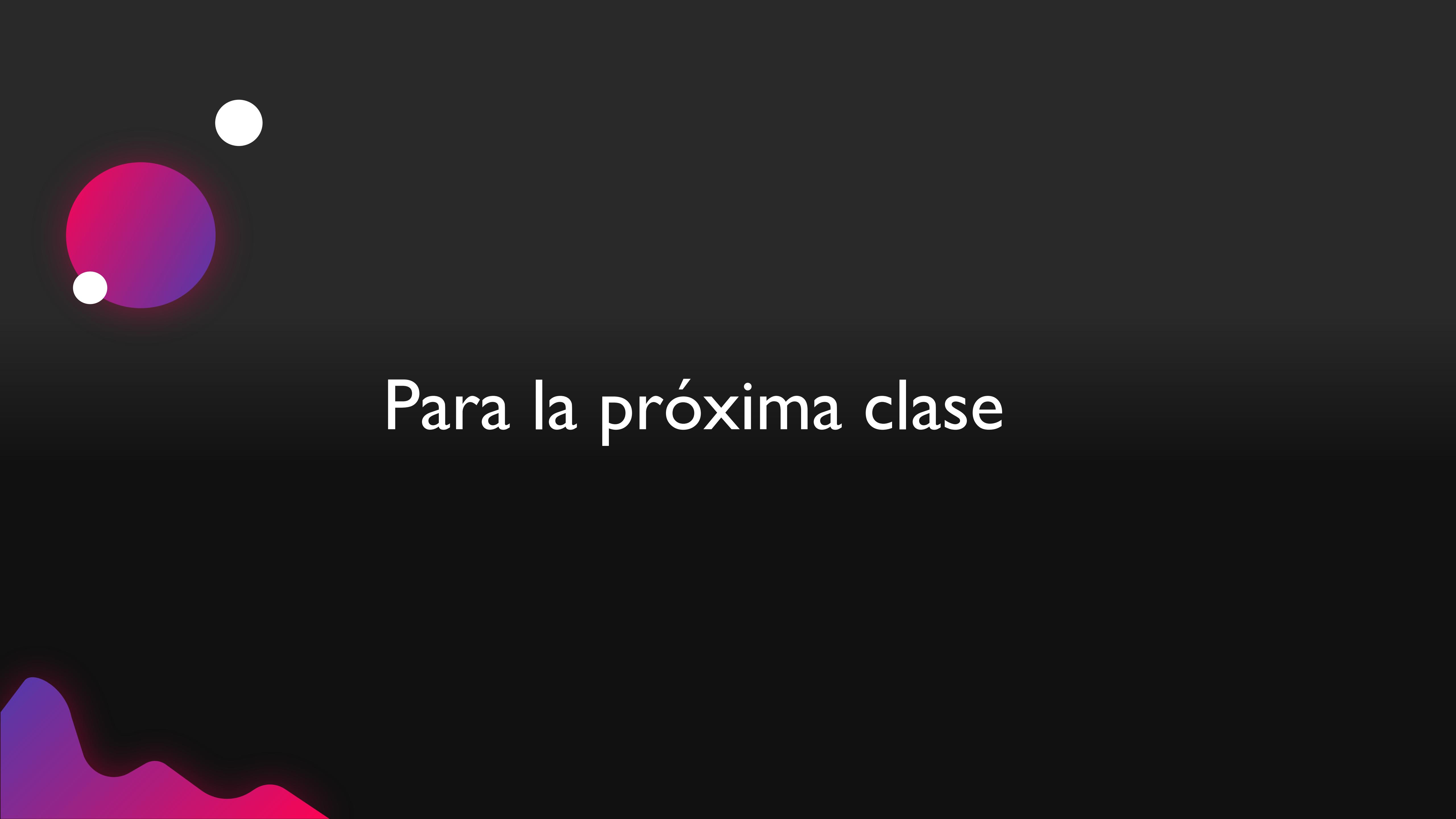
↑ Previous

Octubre



23

Next ↓



Para la próxima clase

Para la próxima clase

19 Agosto

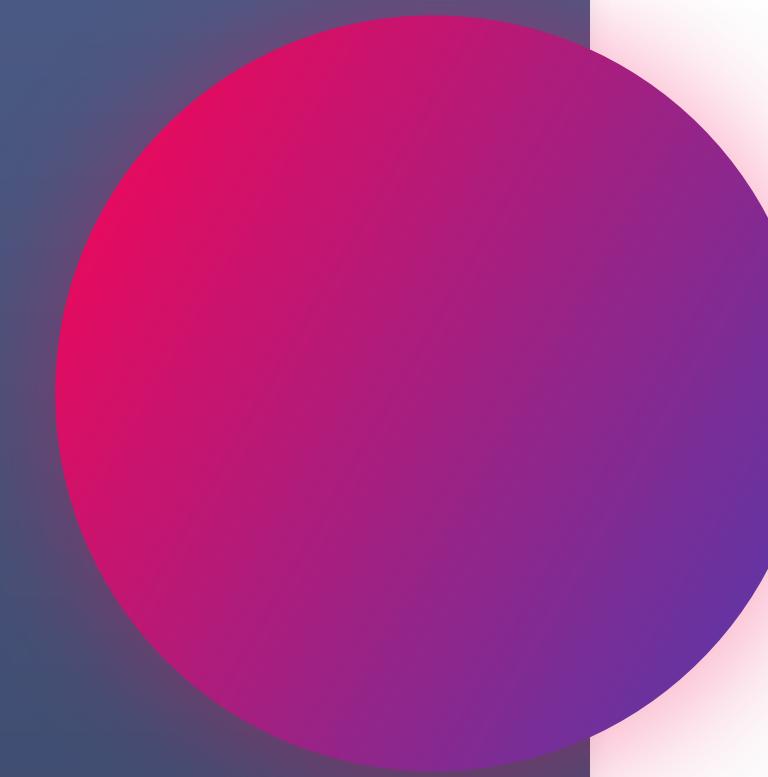
- Modelo preliminar base de datos
- Selección de metodología de desarrollo de software

25

24 Agosto:

- Instalar jupyter notebook
- Instalar numpy y OpenCV

Gracias



leonardocamargoforero



leonardocamargoforero1@gmail.com