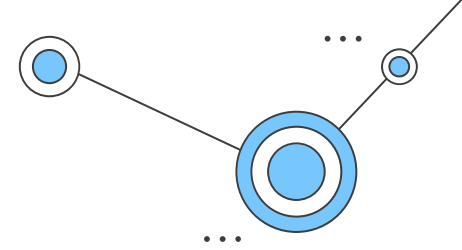


Ciclo de Vida de Datos y MLOPs

Herramientas, metodologías, puesta
en producción y mantenimiento de
proyectos de Data Science.

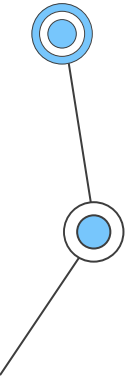
Autor: David Herrero Pascual
Tutor: Iván González Torre
Coordinador: Albert Solé Ribalta

Contenido



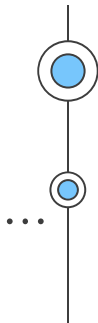
A lo largo de esta presentación:

1. Identificación de la problemática que supone la puesta en producción de modelos de Inteligencia Artificial.
2. Objetivos del Trabajo de Fin de Máster y la motivación.
3. Resumen del estado del arte y el contexto histórico de las técnicas MLOPs.
4. Herramientas que puedan ayudar a solventar los problemas expuestos.
5. Definición de un caso de uso para demostrar las capacidades MLOPs en un sistema demostrador.
6. Arquitectura y herramientas utilizadas para el caso de uso.
7. Demostración.
8. Conclusiones.

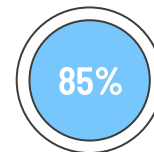
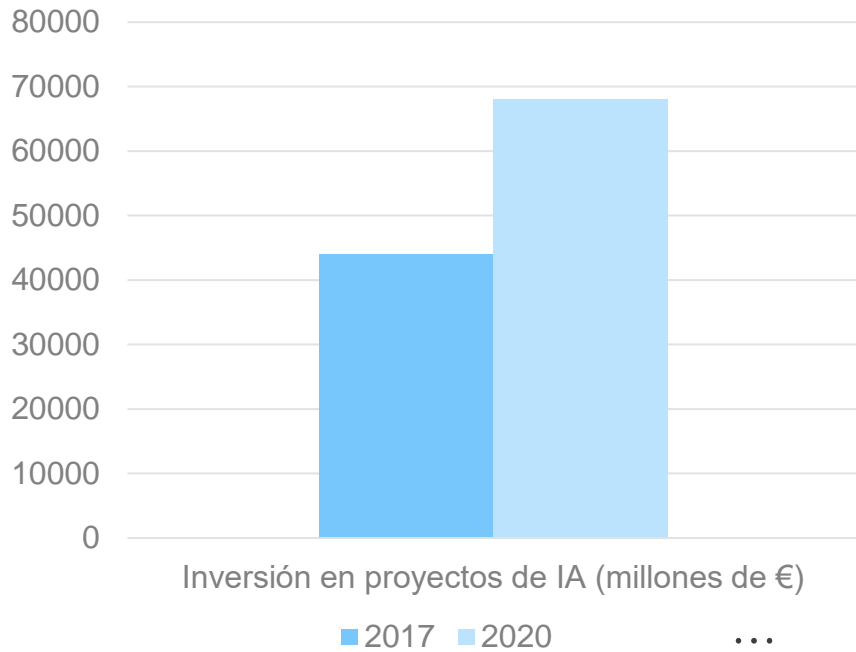


01

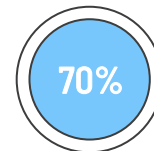
¿Por qué MLOps?



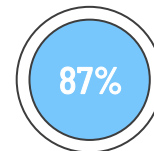
¿Por qué MLOPs?



Proyectos de IA que fallarán hasta 2022



Empresas que reportan impacto mínimo de su IA.



Proyectos de Data Science que nunca llegan a Producción

Objetivos y Motivación



Motivación

Guía de desarrollo
orientado a Data
Science.

...



Objetivo 1

Estado del Arte de
MLOps y contexto
histórico.

...



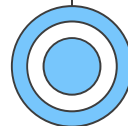
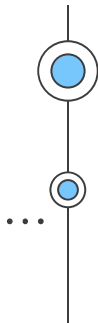
Objetivo 2

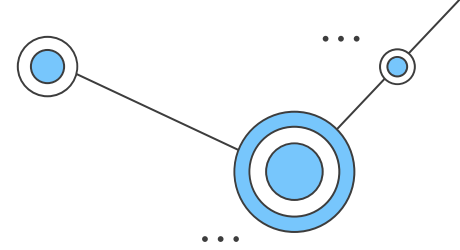
Comparativa de
herramientas y
demo.

...

02

Contexto

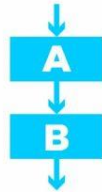




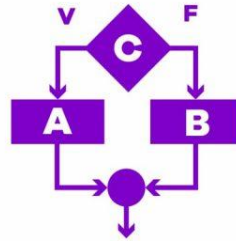
Programación Estructurada

- “La secuencia Goto, considerada perjudicial” – Edsger Dijkstra.
- Objetivo: Evitar el código spaghetti.
- Con el tiempo, quedó obsoleta por la complejidad de los desarrollos.

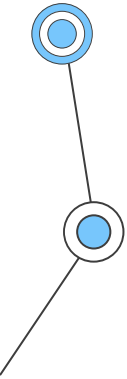
Secuencia



Selección o condicional

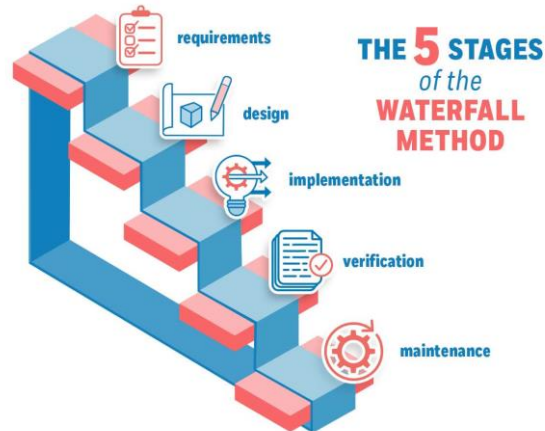


Iteración (ciclo o bucle)

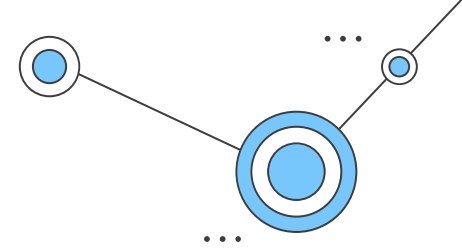


Metodología en Cascada

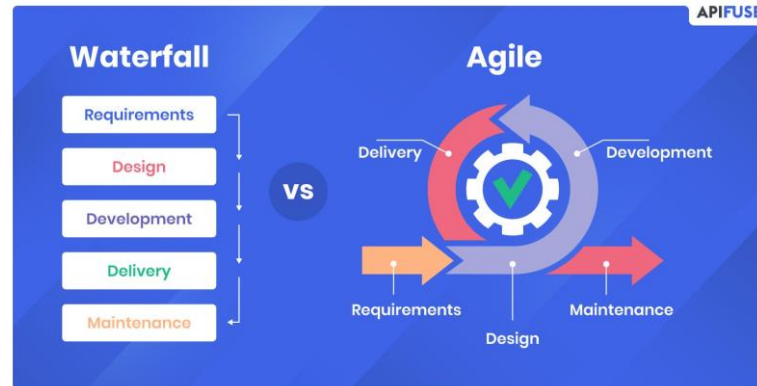
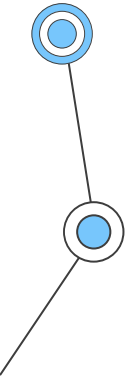
- Metodología de desarrollo software y gestión de proyectos.
- Rígida.
- Proyecto bien definido desde el principio.
- Pasos determinados.
- Correcta definición de requisitos, inamovibles de principio a fin.



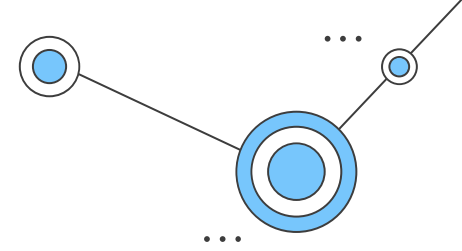
Metodologies Àgiles



- Manifesto Agile, 2001.
- 4 valors principals:
 - Individuos e interaccions VS processos y herramientas.
 - Producte funcional VS documentació exhaustiva
 - Col·laboració amb client VS negociació de un contracte.
 - Responder al canvi VS plan estricte.

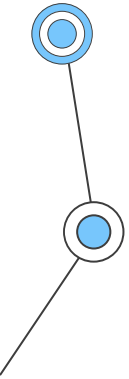
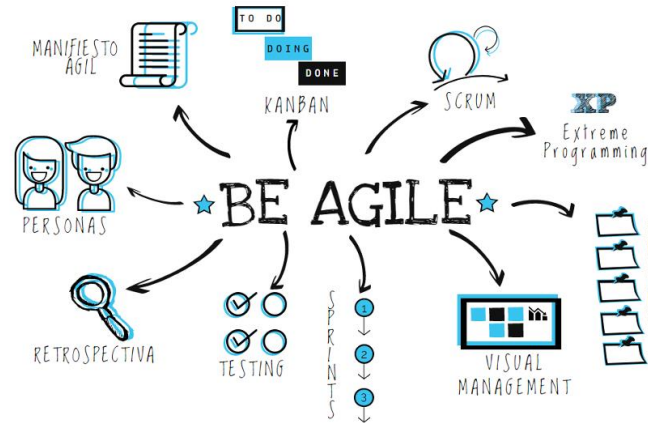


Metodologías Ágiles

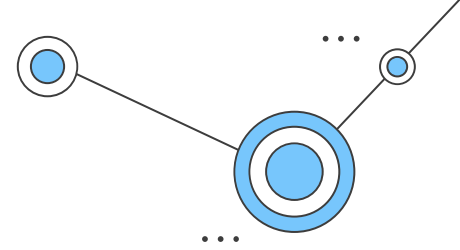


Beneficios:

- Simplificación de la sobrecarga de procesos → Iteraciones.
- Mejora de la calidad
- Mejora de la previsibilidad → Gestión del riesgo.
- Mejora de la productividad.

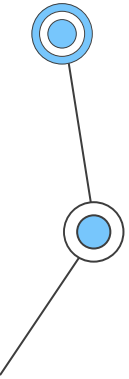
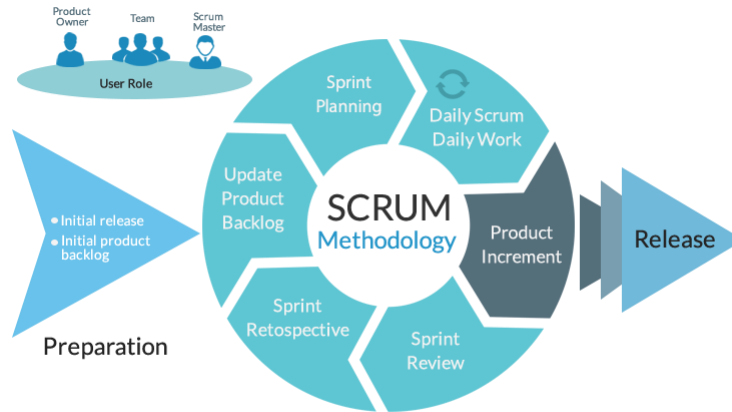


Metodologies Àgiles

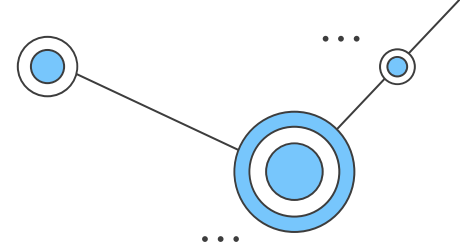


Beneficis:

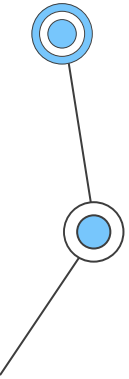
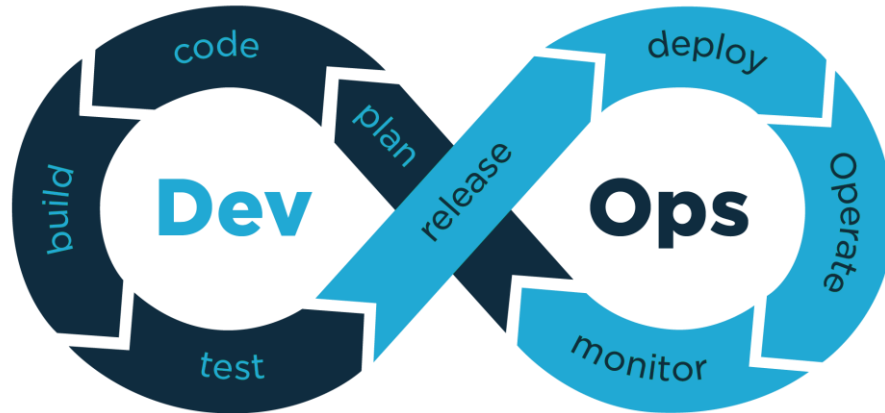
- Feedback continuu del client.
- Motivaci3 del equip → Objectius a curt pla.

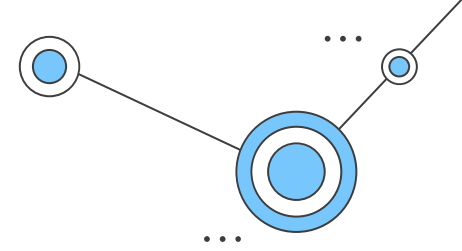


Técnicas DevOps



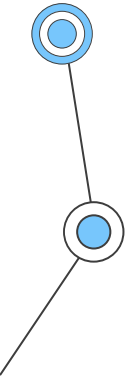
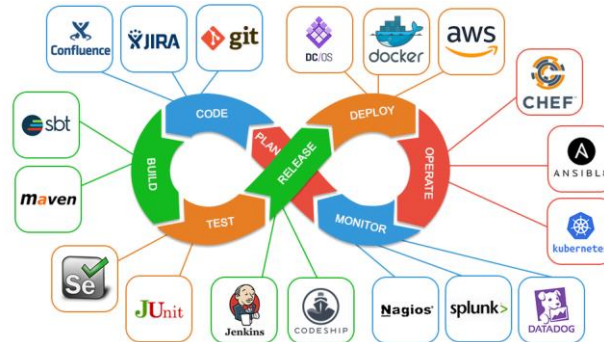
- Surgimiento → Problema de alineamiento entre desarrollo y operaciones.
- Ambos equipos existían independientemente el uno del otro.
- Despliegue y puesta en producción como fase crítica.





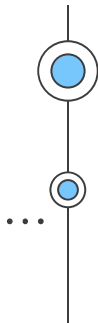
Implicaciones:

- Uso de las metodologías ágiles por definición.
- Integración y entrega continuas.
- Flujos de trabajo, repositorios Git.
- Gestión de servicios TI y gestión de incidentes y eventos no planificados.

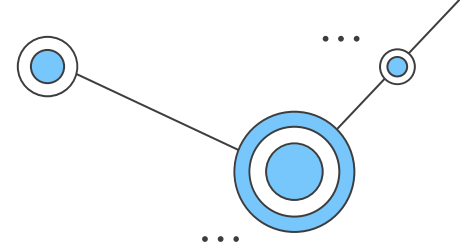


03

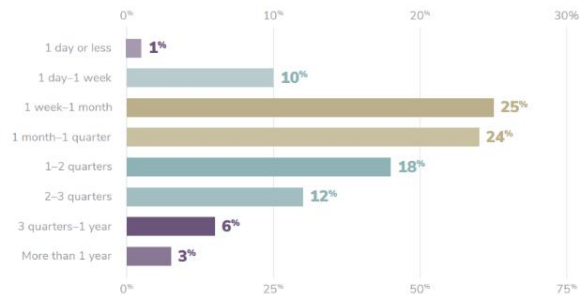
MLOps



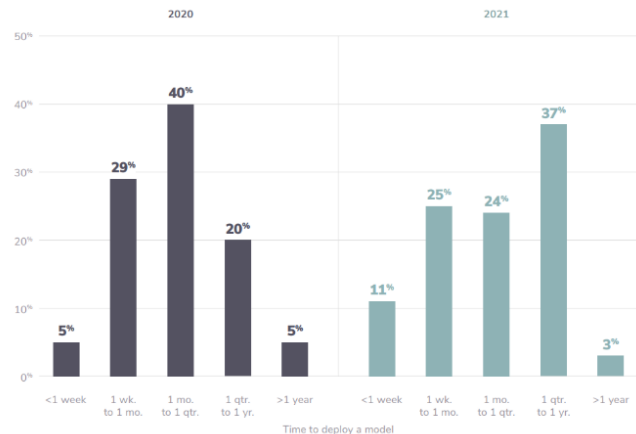
MLOps



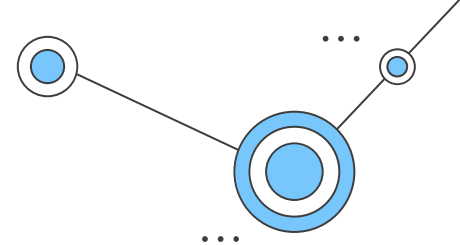
- Surgimiento → Problemas de puesta en producción de modelos de IA.
- DevOps aplicado a proyectos de Inteligencia Artificial.



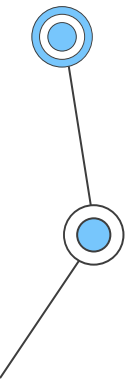
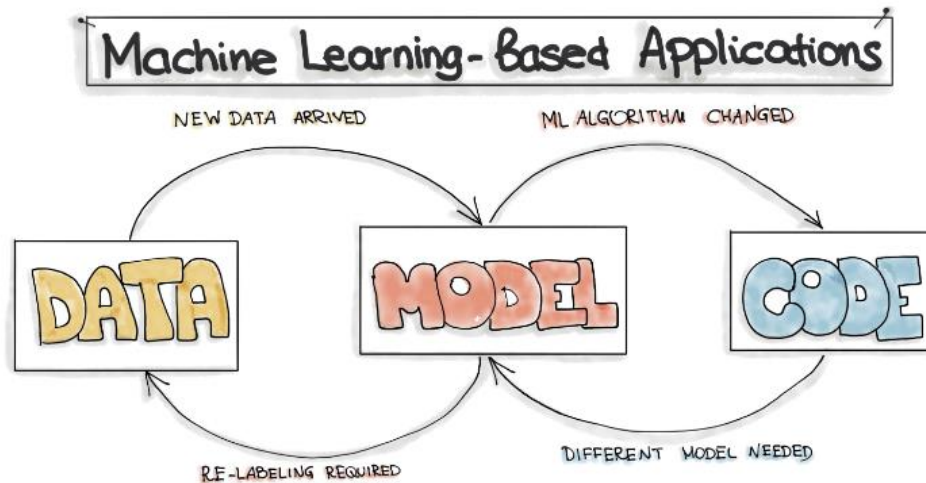
The total percentage of respondents who selected a month or longer was calculated with the underlying data before being rounded to the nearest percentage point. Categories do not add up to 100% because they have been rounded to the nearest percentage point.



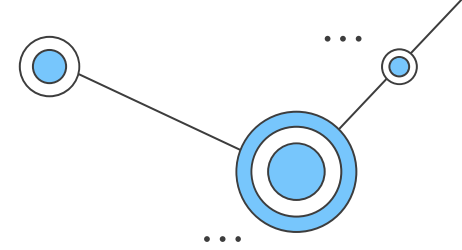
MLOps



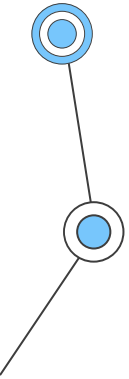
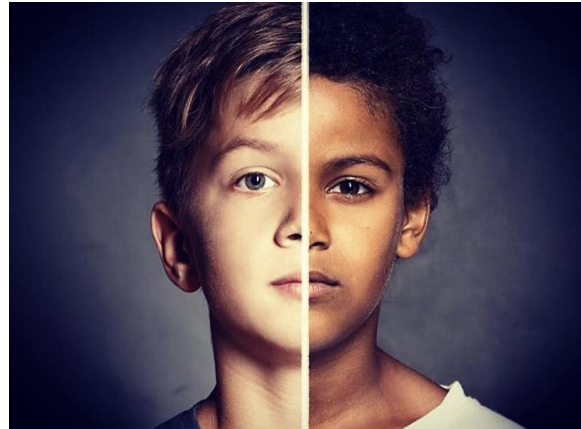
- 3 niveles de cambio: Datos, Modelo y Código.
- “Cambiar cualquier cosa cambia todo”.



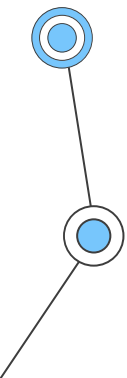
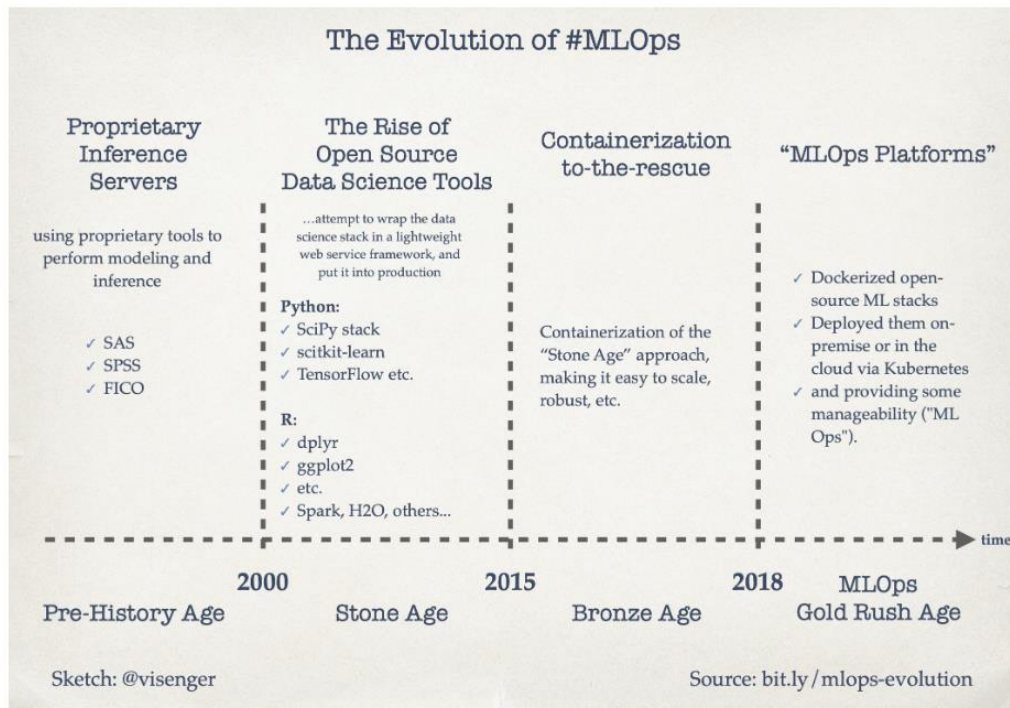
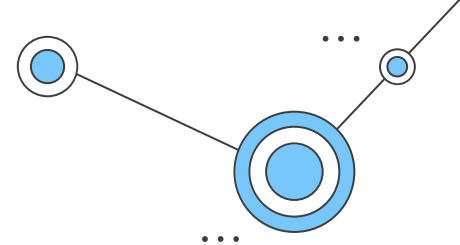
MLOps



- La calidad de los datos puede variar a lo largo del tiempo, o el modelo puede estar entrenado con datos que luego no se corresponden con la realidad.
- El rendimiento de los modelos en producción puede degenerar con el tiempo.
- Puede existir sesgo.

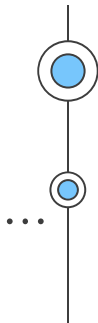


MLOps

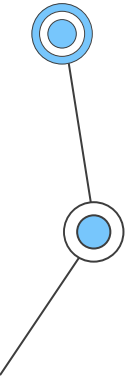
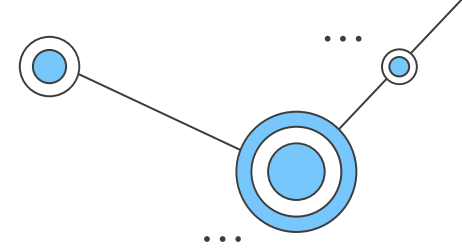
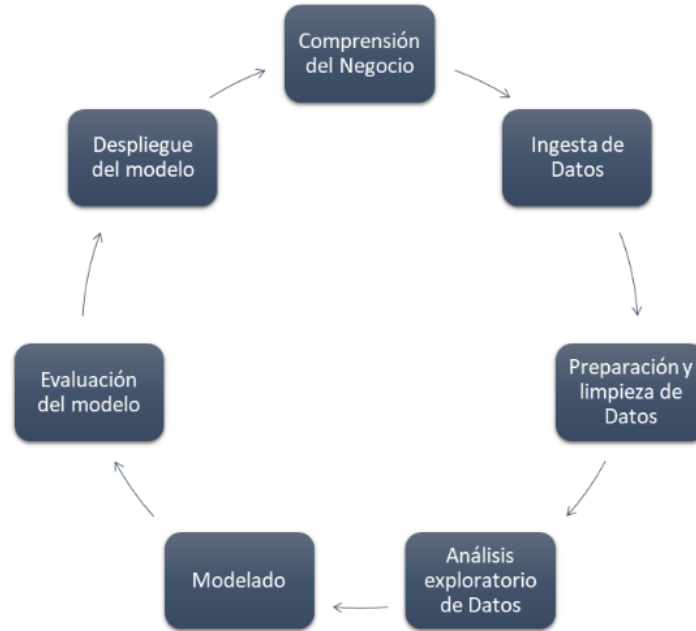


04

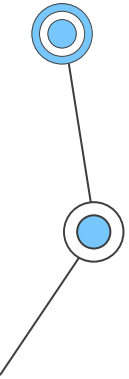
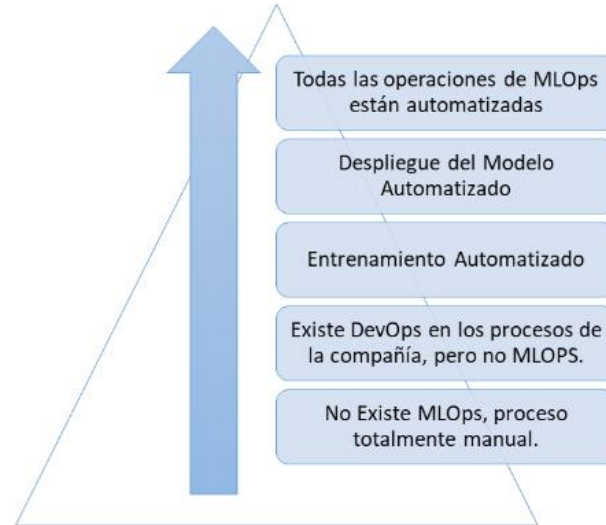
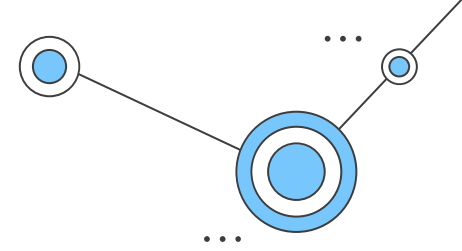
Implementación



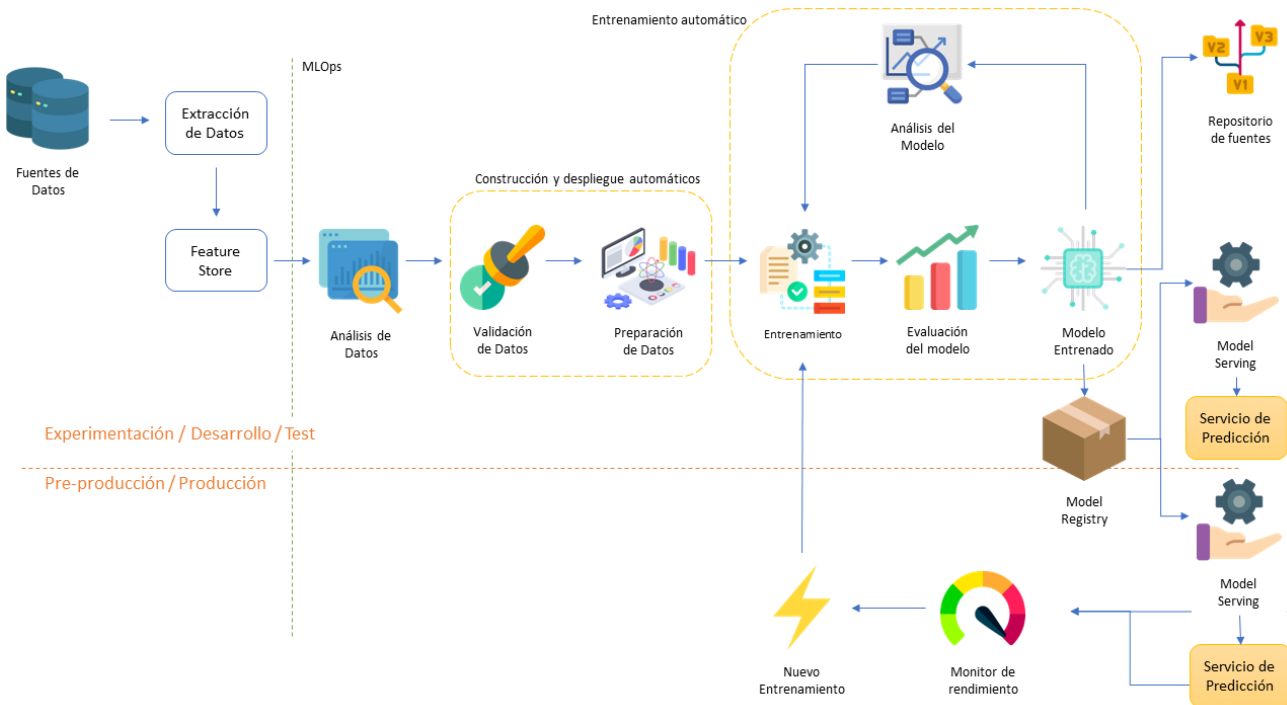
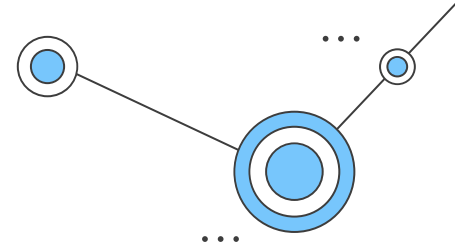
Ciclo de Vida en proyectos de IA

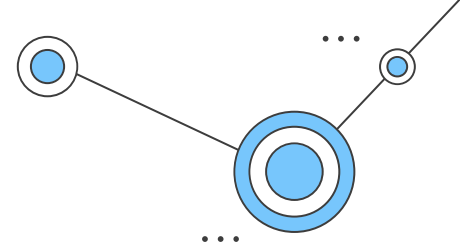


Grados de Madurez de MLOPs

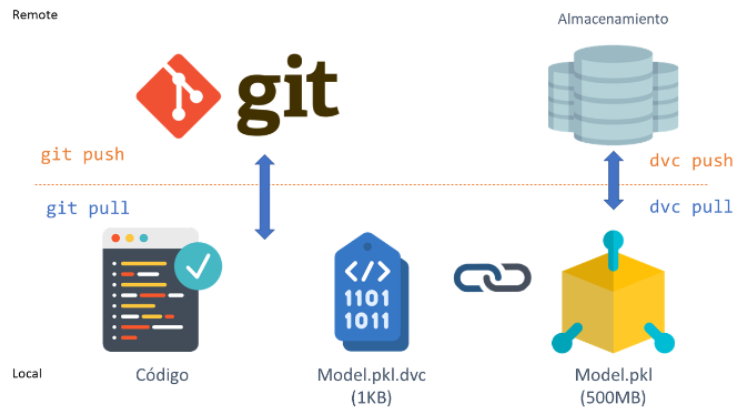


Arquitectura de referencia

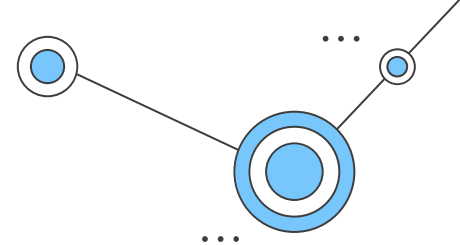




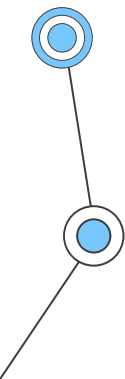
Control de versiones



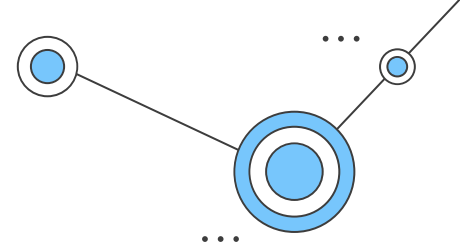
Herramientas



MLOps



Herramientas



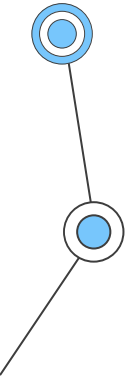
MLOps

mlflowTM



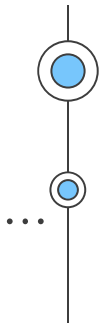
Kubeflow

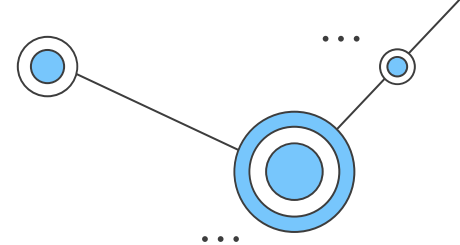
H2O.ai



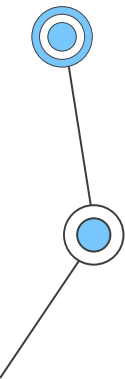
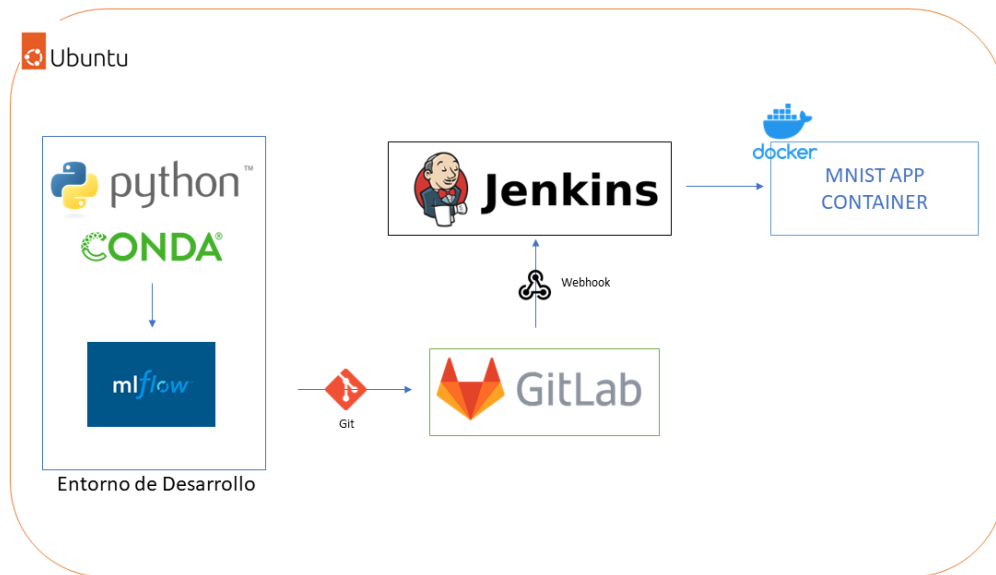
05

Demo



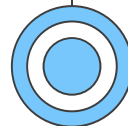
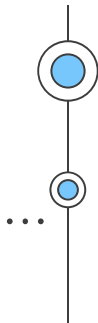


Arquitectura

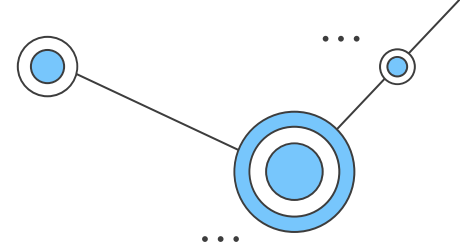


06

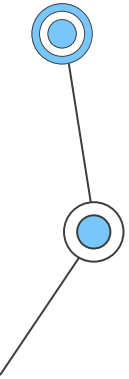
Conclusiones



Conclusiones



- Mejora de la velocidad de puesta en producción.
- Mejora de la trazabilidad de modelos.
- Mejora de la trazabilidad de los datasets.
- Mejora del mantenimiento y actualización de los modelos en producción.



Gracias!

CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#), infographics & images by [Freepik](#) and illustrations by [Stories](#)

