

Asignatura	Datos del alumno	Fecha
<b>Investigación en Inteligencia Artificial</b>	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	

## Actividad 2: Propuesta de desarrollo y despliegue de un proyecto de inteligencia artificial

### Sistema Inteligente de Visión por Computadora para el Reconocimiento de Expresiones Faciales

#### Requisitos

La finalidad es crear un sistema que pueda ser utilizado por entidades de gobierno, empresas privadas o partidos políticos que necesiten presentar algún producto nuevo, imagen, comercial o discurso y poder obtener las expresiones de la muestra de personas a quienes se lo presentan antes de salir al mercado para saber el impacto que puede llegar a tener y si es el deseado.

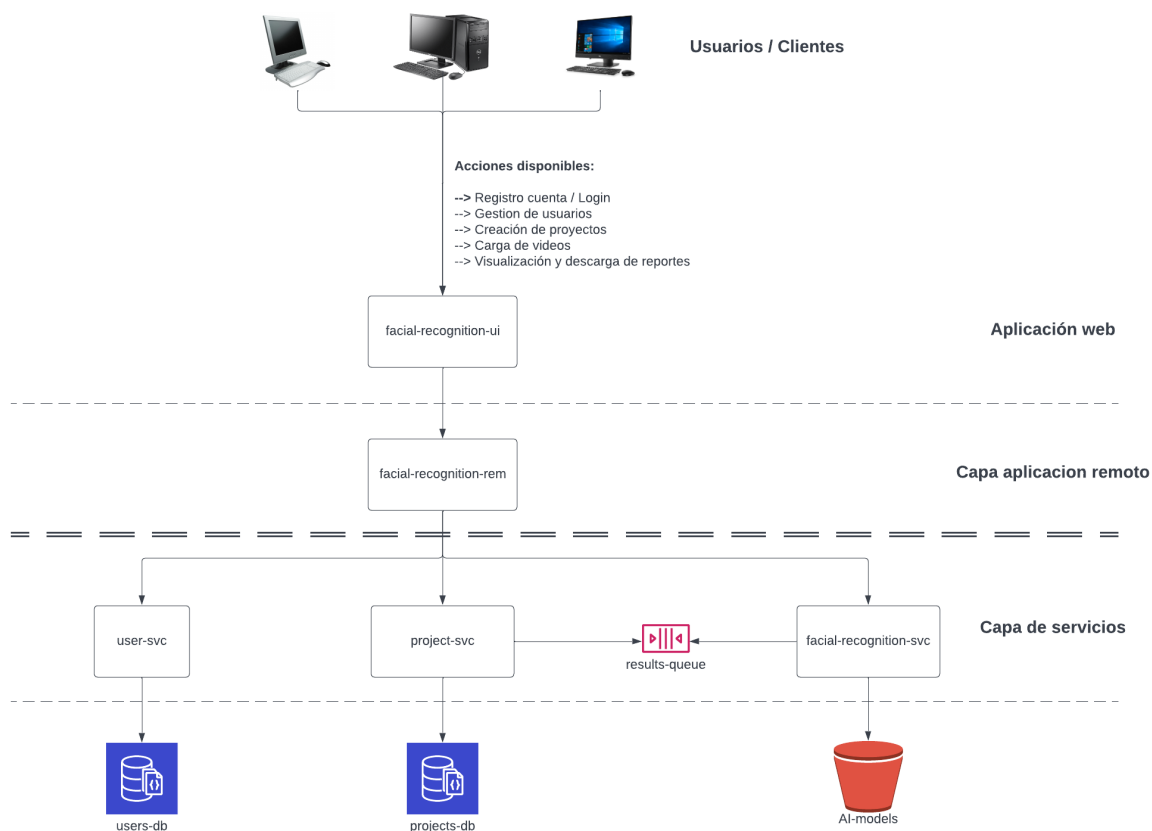
Actualmente esto se realiza haciendo cuestionarios a cada una de las personas pertenecientes al grupo muestra, con esta aplicación se pretende aumentar la cantidad de información recolectada para así ayudar a que las empresas puedan tomar una decisión final mas informada.

Ya que no se espera la misma reacción para diferentes situaciones, por ejemplo cuando se presenta el tráiler de una nueva película de terror, o el nuevo anuncio de la consola de video juegos de ultima generación, o del discurso presidencial para anunciar nuevos impuestos. Es por eso por lo que debemos recolectar información de las expresiones faciales del público muestra durante la presentación y contar con los resultados de impacto para cualquier momento/segmento de la presentación para así determinar si el impacto obtenido en las personas es el que se desea.

#### Diseño de la solución

El siguiente esquema es un diagrama general acerca de los módulos y aplicaciones que el sistema necesita para resolver el problema planteado.

Asignatura	Datos del alumno	Fecha
Investigación en Inteligencia Artificial	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	



### Tecnologías por utilizar:

- **Docker 19.x**

“Es una plataforma de software que permite crear, probar e implementar aplicaciones rápidamente. Docker empaqueta software en unidades estandarizadas llamadas contenedores que incluyen todo lo necesario para que el software se ejecute, incluidas bibliotecas, herramientas de sistema, código y tiempo de ejecución. Con Docker, puede implementar y ajustar la escala de aplicaciones rápidamente en cualquier entorno con la certeza de saber que su código se ejecutará.” [1]

- **Kubernetes 1.19**

“Kubernetes es un software de código abierto que le permite implementar y administrar aplicaciones en contenedores a escala. Kubernetes administra clústeres de instancias de computación de Amazon EC2 y ejecuta contenedores en las instancias con procesos destinados a implementación, mantenimiento y escalado.” [2]

- **Amazon Elastic Kubernetes Service (EKS)**

“Amazon Elastic Kubernetes Service (Amazon EKS) es un servicio de contenedores administrado para ejecutar y escalar aplicaciones Kubernetes en la nube o en las instalaciones.” [3]

Asignatura	Datos del alumno	Fecha
Investigación en Inteligencia Artificial	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	

- **Elastic Load Balancer**  
“Elastic Load Balancing (ELB) distribuye automáticamente el tráfico de aplicaciones entrantes entre varios destinos y dispositivos virtuales en una o varias zonas de disponibilidad (AZ).” [4]
- **MongoDB**  
“MongoDB es una base de datos NoSQL de uso general que proporciona compatibilidad con sistemas de almacenamiento de estilo JSON orientados a documentos. Gracias a su modelo de datos flexible, se pueden almacenar datos de cualquier estructura y proporciona compatibilidad plena con la indexación, además de partición y replicación.” [5]
- **Kafka**  
“Amazon MSK facilita la ingesta y el procesamiento de datos de transmisión en tiempo real con Apache Kafka completamente administrado.” [6]
- **S3**  
“Amazon Simple Storage Service (Amazon S3) es un servicio de almacenamiento de objetos que ofrece escalabilidad, disponibilidad de datos, seguridad y rendimiento líderes en el sector.” [7]
- **Java 11**  
Lenguaje de programación Java versión 11
- **Spring Framework 5**  
“Spring es un Framework Java diseñado para agilizar el desarrollo de aplicaciones empresariales, este implementa un contenedor de inversión de control (IoC) y es de código abierto, puede ser usado para la programación aplicaciones web o de escritorio estándar, cuenta con gran variedad de módulos que nos facilitan el trabajo, tenemos módulos para: acceso a datos con JDBC, ORM, JPA, etc., crear aplicaciones web MVC, entres otros.” [8]
  - Spring Security
  - Spring Data
  - Spring Kafka
- **ItextPDF**  
Librería para Java que nos facilita la creación de reportes PDF.
- **React**  
“React te ayuda a crear interfaces de usuario interactivas de forma sencilla. Diseña vistas simples para cada estado en tu aplicación, y React se encargará de actualizar y renderizar de manera eficiente los componentes correctos cuando los datos cambien.” [9]
- **Grunt**  
“Grunt es ejecutor de tareas de JavaScript que automatiza la minificación de scripts, la compilación de TypeScript, las herramientas de "lint" de calidad del código, los preprocesadores de CSS y las tareas repetitivas que se necesitan para admitir el desarrollo del cliente.” [10]
- **Python 3.10**  
Lenguaje de programación Python versión 3.10

Asignatura	Datos del alumno	Fecha
Investigación en Inteligencia Artificial	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	

- **Flask 2.1**  
“Flask es un “micro” Framework escrito en Python y concebido para facilitar el desarrollo de Aplicaciones Web bajo el patrón MVC.” [11]
- **Kafka-Python 2.0**  
Cliente de Python para conectarse a Kafka.
- **AffectNet**  
“Existing annotated databases of facial expressions in the wild are small and mostly cover discrete emotions (aka the categorical model). There are very limited annotated facial databases for affective computing in the continuous dimensional model (e.g., valence and arousal).” [12]
- **Numpy 1.22**  
Librería para Python para el manejo de arreglos
- **TensorFlow 2.8**  
“TensorFlow es una biblioteca de software de código abierto para computación numérica, que utiliza gráficos de flujo de datos. Los nodos en las gráficas representan operaciones matemáticas, mientras que los bordes de las gráficas representan las matrices de datos multidimensionales (tensores) comunicadas entre ellos.” [13]
- **Robot Framework**  
“Robot Framework is a generic open source automation framework. It can be used for test automation and robotic process automation (RPA).” [14]
- **Git**  
Sistema para el control de versiones.
- **Jenkins**  
“Jenkins es un servidor open source para la integración continua. Es una herramienta que se utiliza para compilar y probar proyectos de software de forma continua, lo que facilita a los desarrolladores integrar cambios en un proyecto y entregar nuevas versiones a los usuarios. Escrito en Java, es multiplataforma y accesible mediante interfaz web. Es el software más utilizado en la actualidad para este propósito.” [15]
- **Amazon SageMaker for ML Engineers**  
“Amazon SageMaker helps you streamline the machine learning (ML) lifecycle by automating and standardizing MLOps practices across your organization. You can easily build, train, deploy, and manage ML models” [16]
- **Jira Software**  
“Jira Software forma parte de una gama de productos diseñados para ayudar a equipos de todo tipo a gestionar el trabajo. En principio, Jira se diseñó como un gestor de incidencias y errores. Sin embargo, se ha convertido en una potente herramienta de gestión de trabajo para todo tipo de casos de uso, desde la gestión de requisitos y casos de prueba hasta el desarrollo de software ágil.” [17]

Asignatura	Datos del alumno	Fecha
<b>Investigación en Inteligencia Artificial</b>	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	

Para todas las aplicaciones se utilizarán las siguientes tecnologías:

- Git:  
Para el manejo del control de versiones.
- Docker:  
Se crearán contenedores Docker con la aplicación, de esta forma en caso de necesitar más capacidad de cómputo, se podrán crear más y así balancear la carga.
- Kubernetes:  
Nos proporciona herramientas para la creación y manejo de contenedores Docker.
- Elastic Kubernetes Service:  
Este servicio es la solución proporcionada por Amazon para Kubernetes, nos permite la creación de contenedores Docker onDemand, entonces si en algún momento se necesita aumentar la capacidad de cómputo para alguna de las aplicaciones, podremos crear más contenedores y balancearlos utilizando el load balancer.
- Elastic Load Balancer:  
Este servicio de Amazon nos ayuda a balancear la carga de trabajo entre todos los contenedores que tengamos disponibles para cada aplicación.
- Jenkins:  
Con esta aplicación podremos hacer CI/CD (Continuos Integration / Continuos Deployment) de todas nuestras aplicaciones. La aplicación compilada y empaquetada se subirá a un S3 de Amazon para su posterior despliegue.
- S3  
Este servicio de Amazon tendrá doble propósito:
  - Guardar las versiones de las aplicaciones compiladas listas para el despliegue.
  - Guardar los reportes generados en PDF.

### **facial-recognition-ui (webapp)**

Esta aplicación será la webapp, es la interfaz a la que el usuario final tendrá acceso para realizar todas las operaciones.

Esta aplicación utilizará las siguientes tecnologías:

- React  
Framework para programar en lenguaje de Javascript.
- Grunt  
Ejecutor de tareas que será útil para “minificar” archivos javascript y CSS entre otras.

### **facial-recognition-rem**

Esta aplicación es llamada aplicación remoto, su principal objetivo es recibir los “requests” de la “webapp”, validar la autenticación de usuario, registro de usuarios,

Asignatura	Datos del alumno	Fecha
<b>Investigación en Inteligencia Artificial</b>	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	

validar reglas de negocio y comunicarse a los servicios que tienen acceso a las bases de datos y al servicio de IA.

#### **user-svc**

Esta aplicación servicio, no es accesible desde el exterior de la empresa, solamente puede ser accedida desde dentro, ya sea desde otro servicio o de un remoto, pero nunca desde una webapp o “requests” manuales del exterior.

Su principal objetivo es el administrar los usuarios registrados en la aplicación, guardará, actualizará y leerá información de los usuarios.

#### **project-svc**

Esta aplicación servicio, no es accesible desde el exterior de la empresa, solamente puede ser accedida desde dentro, ya sea desde otro servicio o de un remoto, pero nunca desde una webapp o “requests” manuales del exterior.

Su principal objetivo es el administrar los proyectos de los usuarios, también escuchará una queue de Kafka para cuando existan resultados listos para ser guardados en un proyecto. También generará los reportes en PDF con los resultados.

Estas últimas 3 aplicaciones estarán desarrolladas utilizando:

- Java 11  
Esto es debido a que existen muchos programadores que programan en este lenguaje y encontrarlos no será difícil.
- Spring Framework  
Este framework nos ayudara a crear aplicaciones REST fácilmente y nos brindara la seguridad con Spring Security en la aplicación remoto.

#### **facial-recognition-svc**

Esta aplicación servicio, no es accesible desde el exterior de la empresa, solamente puede ser accedida desde dentro, ya sea desde otro servicio o de un remoto, pero nunca desde una webapp o “requests” manuales del exterior.

Su principal objetivo será el filtrar y preparar las imágenes para su posterior clasificación, hará uso de un dataset llamado AffectNet, el cual contiene imágenes de rostros de personas que nos será de mucha ayuda para el inicio del entrenamiento y despliegue de las primeras versiones.

Cada vez que obtenga un resultado de alguna imagen, mandará un mensaje a través de la queue de Kafka, el cual será leído por project-svc para posteriormente guardar el resultado en la base de datos. Este proceso lo hará por cada rostro detectado en dentro del video.

También notificara con un mensaje especial que el proceso ha terminado, de esta forma el servicio de Project-svc sabrá que ya puede generar el reporte final.

Para este servicio se utilizarán las siguientes tecnologías:

Asignatura	Datos del alumno	Fecha
<b>Investigación en Inteligencia Artificial</b>	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	

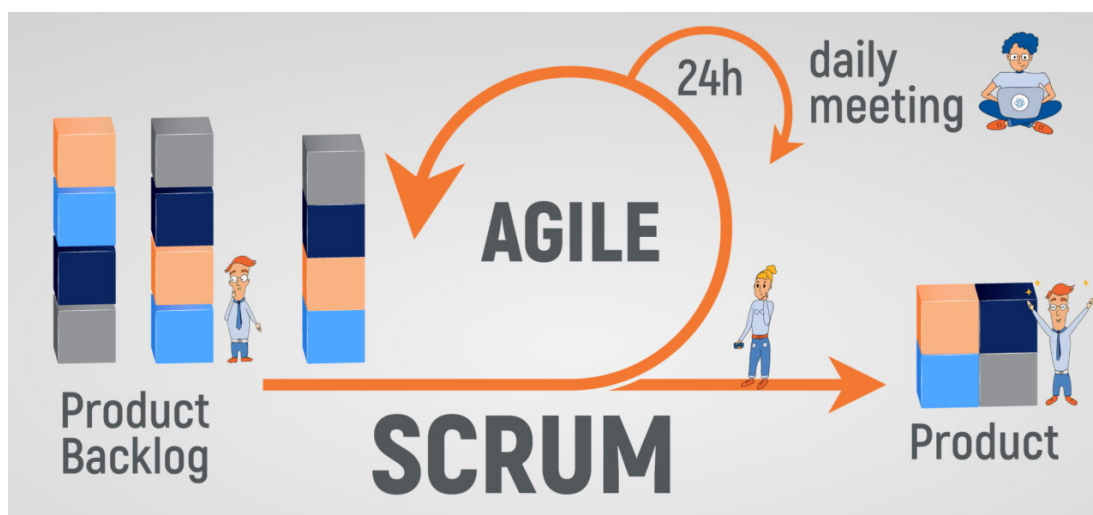
- Python  
Lenguaje de programación que es comúnmente utilizado para ámbitos de IA y ML.
- Numpy 1.22
- Tensorflow 2.8
- Amazon SageMaker for ML Engineers  
Amazon SageMaker for ML Engineers es un servicio de Amazon que nos ayuda automatizando y estandarizando el ciclo de vida para machine learning.

### Quality Assurance

Para verificar la calidad del sistema y hacer pruebas de integración entre todos los componentes, se realizarán pruebas manuales y automatizadas, para las pruebas manuales se utilizará la test suite de Jira y para las pruebas automatizadas se hará uso de Robot Framework, Robot es un framework de Python, diseñado para facilitar la creación de pruebas automatizadas.

### Metodología de desarrollo

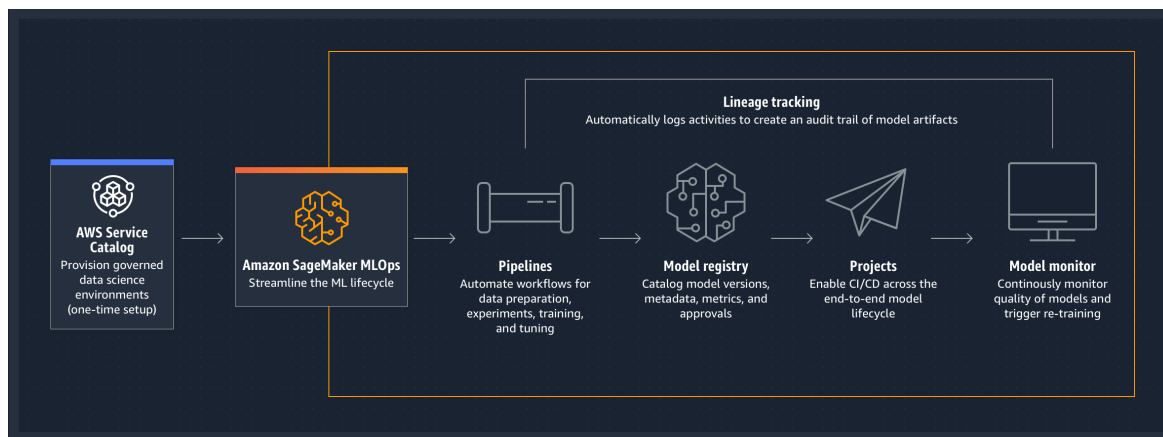
Para todas las aplicaciones a excepción de los módulos de la IA se utilizará Scrum, ya que de esta manera podremos tener un MVP rápidamente y podremos empezar a hacer pruebas con el mercado.



Fuente: <https://kruschecompany.com/agile-software-development-with-scrum-framework/>

Para el desarrollo de los módulos de IA se utilizará MLOps, ya que para re-entrenar la red, la filosofía es un poco distinta.

Asignatura	Datos del alumno	Fecha
Investigación en Inteligencia Artificial	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	



Fuente: <https://aws.amazon.com/sagemaker/mlops/>

En producción, se le pedirá a una muestra aleatoria de usuarios el permiso para utilizar su información (de manera anónima) para re-entrenar la red, entonces cada vez que se obtenga un incremento del 20% de datos, la red se re-entrenará y se probará para verificar su eficiencia.

Para asegurar el anonimato, las imágenes de los usuarios que se guarden, estarán localizadas en un bucket de S3 al cual solamente el servicio de entrenamiento tendrá acceso y no se guardará información personal de los usuarios.

### Equipos de trabajo:

Para poder iniciar el desarrollo del sistema se requerirán los siguientes perfiles:

- Product Owner
- Scrum Master
- Architect / Data Engineer
- Data Scientist
- DevOps
- UX
- 6 Developers
- 2 QA

Esto es para iniciar con 2 equipos de trabajo. Product Owner, Scrum Master, Architect/Data Engineer, Data Scientist y DevOps estarían trabajando en conjunto con los 2 equipos.

La persona de UX puede ser algún freelancer o alguien contratado bajo un contrato limitado.

3 Developers y 1 QA estarán asignados en un equipo y los demás en el otro.



Asignatura	Datos del alumno	Fecha
<b>Investigación en Inteligencia Artificial</b>	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	

## Bibliografía

- [1 Amazon, «Contenedores de Docker,» [En línea]. Available: ] <https://aws.amazon.com/es/docker/>. [Último acceso: 17 04 2022].
- [2 Amazon, «Kubernetes en AWS,» [En línea]. Available: ] <https://aws.amazon.com/es/kubernetes/>. [Último acceso: 17 04 2022].
- [3 Amazon, «Servicio de Kubernetes administrado,» [En línea]. Available: ] <https://aws.amazon.com/es/eks/>. [Último acceso: 18 04 2022].
- [4 Amazon, «AWS | Elastic Load Balancer,» [En línea]. Available: ] <https://aws.amazon.com/es/elasticloadbalancing/>. [Último acceso: 18 04 2022].
- [5 Amazon, «MongoDB en AWS,» [En línea]. Available: ] <https://aws.amazon.com/es/quickstart/architecture/mongodb/>. [Último acceso: 18 04 2022].
- [6 Amazon, «Apache Kafka completamente administrado,» [En línea]. Available: ] <https://aws.amazon.com/es/msk/>. [Último acceso: 18 04 2022].
- [7 Amazon, «Almacenamiento de datos,» [En línea]. Available: ] <https://aws.amazon.com/es/s3/>. [Último acceso: 18 04 2022].
- [8 Amazon, «Spring,» [En línea]. Available: ] <https://acodigo.blogspot.com/p/spring.html#:~:text=Spring%20es%20un%20Framework%20Java,m%C3%B3dulos%20que%20nos%20facilitan%20el>. [Último acceso: 18 04 2022].
- [9 Facebook Open Source, «React,» [En línea]. Available: <https://es.reactjs.org/>. [Último acceso: 18 04 2022].
- [1 Microsoft, «Uso de Grunt,» [En línea]. Available: <https://docs.microsoft.com/es-es/aspnet/core/client-side/using-grunt?view=aspnetcore-6.0>. [Último acceso: 18 04 2022].
- [1 Jose Domingo Muñoz, «Que es Flask?,» [En línea]. Available: ] <https://openwebinars.net/blog/que-es-flask/>. [Último acceso: 18 04 2022].
- [1 M. Mahoor, «AffectNet - Mohammad H. Mahoor, Ph.D.,» [En línea]. Available: ] [AffectNet - Mohammad H. Mahoor, Ph.D.,](#). [Último acceso: 18 04 2022].
- [1 J. Buhigas, «Todo lo que necesitas saber sobre Tensor Flow,» [En línea]. Available: ] <https://puentesdigitales.com/2018/02/14/todo-lo-que-necesitas-saber-sobre-tensorflow-la-plataforma-para-inteligencia-artificial-de-google/>. [Último acceso: 18 04 2022].
- [1 Robot Framework, «Robot Framework,» [En línea]. Available: ] <https://robotframework.org/>. [Último acceso: 18 04 2022].
- [1 Sentries, «Introduccion a Jenkins,» [En línea]. Available: ] <https://sentries.io/blog/que-es-jenkins/#:~:text=Jenkins%20es%20un%20servidor%20open,nuevas%20versiones%20a%20los%20usuarios..> [Último acceso: 18 04 2022].

Asignatura	Datos del alumno	Fecha
<b>Investigación en Inteligencia Artificial</b>	Apellidos: Anchondo Trejo	18/Abril/2022
	Nombre: Victor Eduardo	

[1 Amazon, «MLOps,» [En línea]. Available:  
6] <https://aws.amazon.com/sagemaker/mlops/>. [Último acceso: 18 04 2022].

[1 Atlassian, «Para que sirve Jira?,» [En línea]. Available:  
7] <https://www.atlassian.com/es/software/jira/guides/use-cases/what-is-jira-used-for>. [Último acceso: 18 04 2022].