Informe de progreso

MLEDGE - Aprendizaje automático en la nube y en el borde

(Cloud and Edge Machine Learning)

Juniio de 2024

Información sobre el entregable

Nombre del documento:

Informe de Progreso

Versión actual: 1.0

Proyecto: MLEDGE - Aprendizaje automático en la nube y en el borde (Cloud and Edge Machine Learning)

Paquete de trabajo: P0 - Gestión del Proyecto

Tarea: A0.1: Gestión del proyecto

Entregable: E0.2 - M18 - Informe de Progreso

Autores: Santiago Andrés (IMDEA)

Revisores: Nikolaos Laoutaris (IMDEA)

Historial de Versiones

Versión	Fecha	Resumen de modificaciones
Version 1.0	24-06-2024	Versión inicial del documento

Índice

1. Introducción

Este informe de progreso tiene como objetivo presentar el progreso de los diferentes paquetes de trabajo y actividades del proyecto tras 18 meses de ejecución. Se trata además de un instrumento operativo destinado a coordinar a los socios de MLEDGE en la ejecución de las actividades del proyecto, y está dirigido principalmente a puntos de contacto administrativo. Este documento se estructura en las siguientes secciones:

- La sección 2 incluye una memoria de actividades por paquete de trabajo que abarca desde el mes 12, en el que se entregó la anterior versión de este informe, hasta el mes 18.
- La sección 3 incluye los siguientes pasos del proyecto.

2. Memoria de actividades

2.1. Objetivo del proyecto

El objetivo del proyecto MLEDGE es impulsar la implementación de FL como una capa intersectorial independiente pero optimizada sobre CloudEdge, utilizando aplicaciones y datos del mundo real para demostrar que esta sinergia puede producir grandes beneficios para todos. Con ello se podrá habilitar un ecosistema próspero de servicios FL en el borde seguros y eficientes capaces de facilitar el uso de datos personales y B2B confidenciales para entrenar modelos de ML para consumidores mientras se protege la privacidad de los datos y de sus propietarios.

Para allanar el camino a la adopción del FL en el borde de la red para un creciente número de aplicaciones que empleen modelos de ML, MLEDGE persigue el desarrollo de técnicas, librerías y componentes que permitan poner en marcha más ágilmente estos servicios. La Figura 1 resume la arquitectura de MLEDGE y los bloques del proyecto.

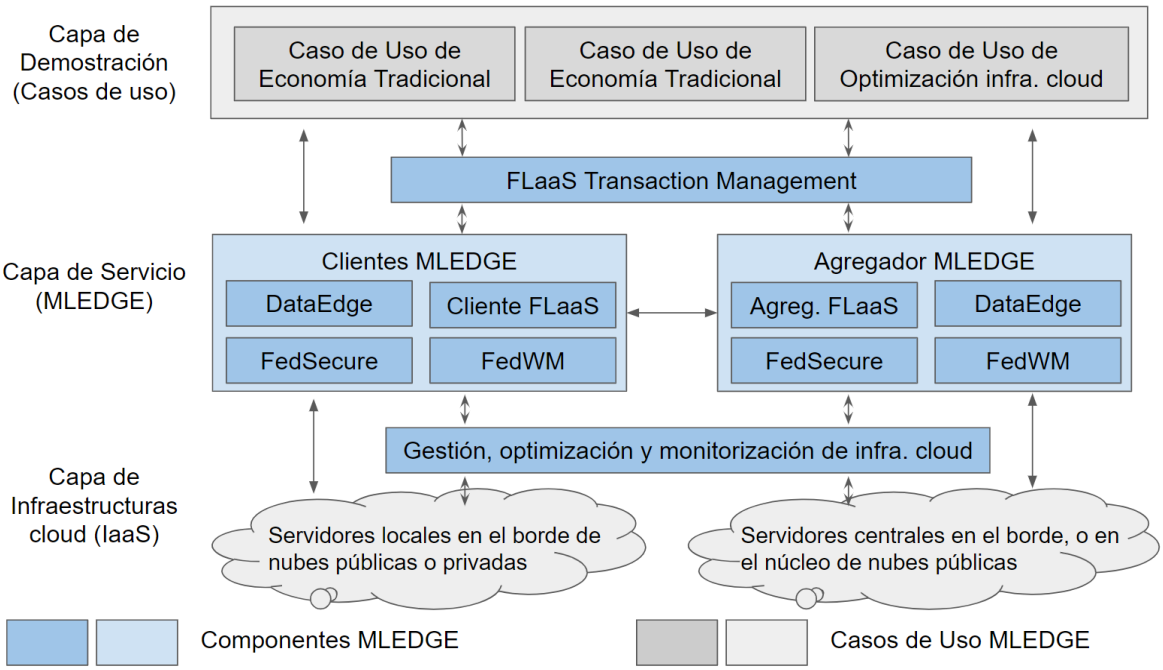

Figura 1. Diagrama de bloques de MLEDGE

2.2. Estructura de ejecución

. En la Figura 2 se proporciona un esquema de los paquetes de trabajo del proyecto y las relaciones entre ellos. El proyecto se estructura en 7 paquetes de actividades (P0-P6). P0 cubre la gestión del proyecto, y P1 tiene como objetivo definir los requisitos de los casos de uso y el diseño de la arquitectura del proyecto. Adicionalmente, se prevén 4 paquetes de trabajo técnicos (P2-P5), y un último paquete (P6) orientado a mostrar las pruebas de

concepto, y a la diseminación, explotación y comunicación de los resultados del proyecto.

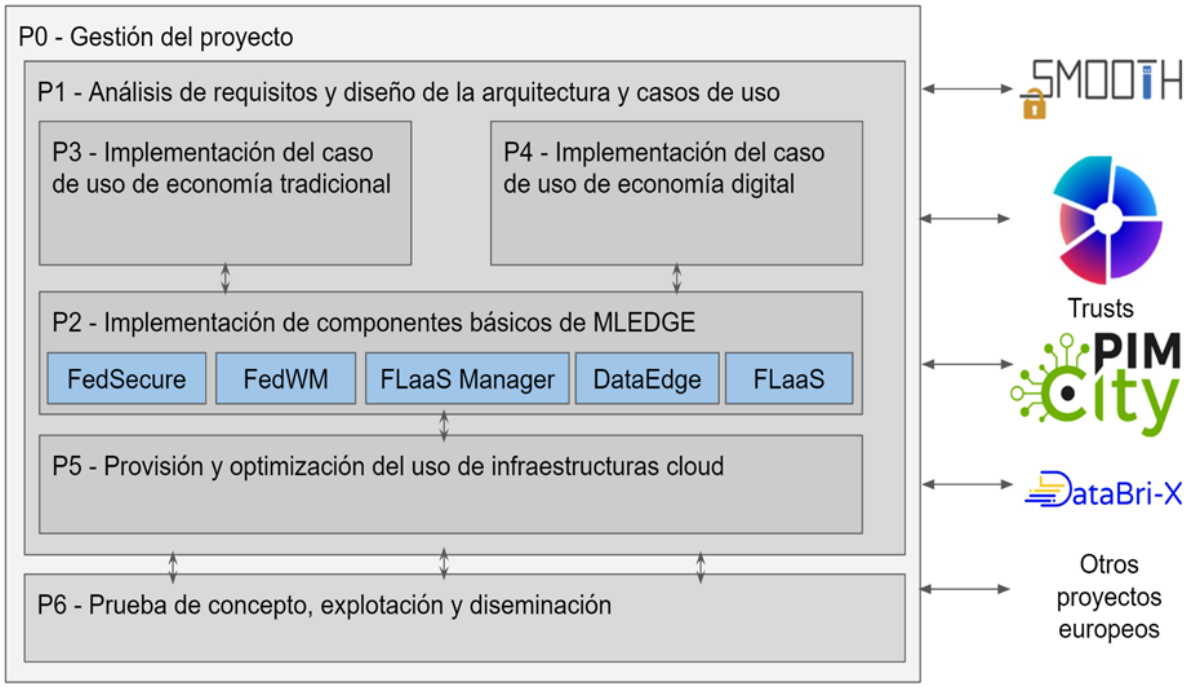

Figura 2. Paquetes de trabajo de MLEDGE

A continuación se ofrece un resumen de los avances y actividades por paquete de trabajo.

2.3. Actividad por paquete de trabajo

2.3.1. P0 - Gestión del Proyecto

Como parte de este paquete de trabajo se han realizado las siguientes actividades:

- Configurar y mantener la infraestructura tecnológica, incluidos sitio web y listas de correo.
- Organizar reuniones periódicas de seguimiento.
- Gestionar la coordinación estratégica mediante revisión periódica de la visión del proyecto, análisis y solución de problemas en la implementación del plan de trabajo.
- Supervisar el proceso de licitación de los componentes técnicos del proyecto
- Elaborar una presentación oficial del proyecto y presentarla a los socios adjudicatarios de los diferentes componentes técnicos
- Coordinar los diferentes componentes técnicos
- Controlar la calidad de la ejecución y los entregables del proyecto
- Participación en reuniones de revisión de los entregables del proyecto previstos para el mes 18.
- Supervisión de los equipos de investigación
- Elaborar el presente informe de actividad

2.3.2. P1 - Análisis de requisitos y diseño de la arquitectura y casos de uso

Como parte de este paquete de trabajo se han realizado las siguientes actividades:

- Supervisar desde un punto de vista técnico y administrativo el proceso de subcontratación, que se desarrolló en dos partes debido a que la primera licitación se declaró desierta.
- Coordinación inicial de los diferentes socios adjudicatarios del proyecto
- Reporte a la gestión del proyecto

2.3.3. P2 - Implementación de componentes básicos de MLEDGE

Como parte de este paquete de trabajo se han realizado las siguientes actividades:

- Actividades propias de investigación en los correspondientes campos
- Desarrollos vinculados a los diferentes componentes
- Documentación de la actividad investigadora
- Elaboración de artículos científicos - ver detalle en página web del proyecto, particularmente en este enlace.
- Elaboración de material para charlas y diseminación
- Selección de trabajos científicos para incorporar en los paquetes de trabajo técnicos P3-P5 (ver detalle en los correspondientes entregables)

2.3.4. P3 - Implementación del caso de uso de economía tradicional

Tras su adjudicación con fecha 3/2/2024 al consorcio formado por InmaRepro y Acuratio, y su posterior formalización con fecha 8/04/2024, la actividad del paquete de trabajo se ha concentrado en la definición de los requisitos y el diseño de la arquitectura de los casos de uso. Para ello se han realizado las siguientes actividades:

- Entendimiento del proyecto
- Participación en reunión de kick-off con el resto de los socios del proyecto
- Elaboración del entregable E3.1 - Estado del arte y diseño de los componentes del caso de uso de economía tradicional.
- Participación en reuniones de seguimiento con el equipo de proyecto de IMDEA Networks
- Revisión de los comentarios al entregable realizados por el equipo de proyecto

2.3.5. P4 - Implementación del caso de uso de economía digital

Tras su adjudicación con fecha 3/2/2024 al consorcio formado por Orange España y Acuratio, y su posterior formalización con fecha 8/04/2024, la actividad del paquete de trabajo se ha concentrado en la definición de los requisitos y el diseño de la arquitectura de los casos de uso. Para ello se han realizado las siguientes actividades:

- Entendimiento del proyecto
- Participación en reunión de kick-off con el resto de los socios del proyecto

- Elaboración del entregable E4.1 - Estado del arte y diseño de los componentes del caso de uso de economía digital.
- Participación en reuniones de seguimiento con el equipo de proyecto de IMDEA Networks
- Revisión de los comentarios al entregable realizados por el equipo de proyecto

2.3.6. P5 - Provisión y optimización de infraestructuras cloud

Tras su adjudicación con fecha 3/2/2024 al consorcio formado por Acuratio S.L., y su posterior formalización con fecha 8/04/2024, la actividad del paquete de trabajo se ha concentrado en la definición de los requisitos y el diseño de la arquitectura de los casos de uso. Para ello se han realizado las siguientes actividades:

- Entendimiento del proyecto
- Participación en reunión de kick-off con el resto de los socios del proyecto
- Elaboración del entregable E5.1 - Estado del artey diseño de los componentes de infraestructuras cloud
- Selección de los trabajos de investigación a ser integrados y testados en la plataforma FLaaS
- Participación en reuniones de seguimiento con el equipo de proyecto de IMDEA Networks
- Revisión de los comentarios al entregable realizados por el equipo de proyecto

2.3.7. P6 - Prueba de concepto, explotación y diseminación

Como parte de este bloque de trabajo se han realizado las siguientes actividades:

- Establecimiento de un feed do noticias en la página web del proyecto (https://mledge.networks.imdea.org/en/noticias/)
- Publicación periódica de noticias en la página de LinkedIn del proyecto (https://www.linkedin.com/company/mledge-project/) y en la página web del proyecto
- Elaboración de material y diseminación de trabajos en página web y redes sociales
- Presentaciones y charlas de diseminación del trabajo en el proyecto.

A continuación se resumen las publicaciones científicas que se han dado como resultado de los desarrollos del proyecto hasta la fecha:

- Santiago Andres Azcoitia, Costas Iordanou, y Nikolaos Laoutaris. (2023) Understanding the Price of Data in Commercial Data Marketplaces. IEEE International Conference on Data Engineering ICDE. Los Angeles, California, USA. April 2023.

- Tianyue Chu, Alvaro Garcia-Recuero, Costas Iordanou, Georgios Smaragdakis, y Nikolaos Laoutaris Securing Federated Sensitive Topic Classification against Poisoning Attacks. Network and Distributed System Security (NDSS) Symposium. 2023.

- Devriş Işler, Elisa Cabana, Álvaro Garcia-Recuero, Gieorgia Koutrika, Nikolaos Laoutaris, FreqyWM: Frequency Watermarking for the New Data Economy, aceptado para publicación en IEEE International Conference on Data Engineering ICDE'24.

- T. Chu, N. Laoutaris, "FedQV: Leveraging Quadratic Voting in Federated Learning," ACM SIGMETRICS'24

Adicionalmente, se han realizado las siguientes publicaciones y charlas:

- Santiago Andres Azcoitia. Presentación del paper Understanding the Price of Data in Commercial Data Marketplaces*.* en el IEEE International Conference on Data Engineering ICDE. Los Angeles, California, USA. April 2023. (enlace al Video).

- Tianyue Chu. Presentación del paper Securing Federated Sensitive Topic Classification against Poisoning Attacks. Network and Distributed System Security (NDSS) Symposium. 2023.

- Santiago Andrés Azcoitia y Alba Ribera Martínez Data Marketplaces and the Data Governance Act: A Business Model Perspective. September 2023. Kluwer Competition Law Blog.

- Santiago Andrés Azcoitia y Alba Ribera Martínez Data Marketplaces and the Data Governance Act: A Business Model Perspective. Noviembre 2023. Charla PLAMADISO (Platforms, Markets, and the Digital Society) en el Weizenbaum Institute for the Networked Society.

- Santiago Andrés Azcotiia, Charla: Towards a Human-centric data economy. Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad Politécnica de Madrid.

- Tianyue Chu. Presentación del paper: FedQV: Leveraging Quadratic Voting in Federated Learning. SIGMETRICS'24. Venecia. Italia

3. Siguientes pasos

En la segunda parte de 2024 se iniciarán las actividades de implementación de los diseños con las empresas adjudicatarias de los paquetes de trabajo 3, 4 y 5. En paralelo, se seguirá avanzando con las actividades de investigación de acuerdo con el plan de trabajo, que se irán progresivamente integrando en los casos de uso según se trabaje con estas compañías. El objetivo de final del año 2024 consiste en disponer de un prototipo intermedio de los componentes técnicos del proyecto que permita demostrar parcialmente los requerimientos y los componentes de diseño establecidos en los entregables 3.1, 4.1, y 5.1.

Respecto a la actividad de publicaciones, se estarán desarrollando los siguientes artículos científicos:

- Un modelo de aprendizaje federado de precios de los datos en mercados de datos comerciales.
- Un modelo de mercado de datos federado y algoritmos de optimización de adquisición de datos por parte de los compradores.
- Un modelo de aprendizaje federado utilizando BitTorrent para la comunicación con garantías de privacidad entre los diferentes participantes.

Adicionalmente, se seguirá trabajado para dotar de contenido a la página web y las redes sociales sobre el proyecto. También se buscarán oportunidades de diseminación para este año y, para el año 2025, de demostración de los prototipos intermedios de los componentes técnicos, que deberían estar listos a finales de año.