



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Lote 4: ADV5G-IMM-XR

Resumen del proyecto

Versión 1.0

Autores principales	YerbaBuena VR S.L., Universitat Politècnica de València
Entregado	31/07/2025

Avanzando-5G-Inmersivo-Telepresencia (TSI-063000-2021-110)

Resumen de proyecto, objetivos y casos de uso

El proyecto **AVANZANDO-5G-INMERSIVO** (*Advancing 5G for Immersive Holographic, Telepresence and Haptics Communications towards 6G*) ha logrado **crear un laboratorio inmersivo para aplicaciones 5G/6G en la Universidad Politécnica de Valencia** (UPV). Estas aplicaciones permiten la convergencia del mundo físico con el mundo virtual e implican el uso de tecnologías disruptivas como la robótica, la inteligencia artificial, los gemelos digitales, y las comunicaciones avanzadas 5G/6G. **El proyecto ha sido financiado a través de fondos europeos mediante el programa UNICO 5G I+D, dentro del marco del Plan de Recuperación, Transformación y Resiliencia.**

En el marco del proyecto AVANZANDO-5G-INMERSIVO, se encuentra el subproyecto **Avanzando-5G-Inmersivo-Telepresencia (TSI-063000-2021-109)**, el cual se focalizó en investigar e implementar **aplicaciones y/o casos de uso de las comunicaciones de telepresencia. Una de las actividades comprendidas en este subproyecto era AVANZANDO-5G-INMERSIVO-XR-A7: Comunicaciones de telepresencia en AR.** En concreto, esta actividad investigó e implementó aplicaciones inmersivas de video 180 y 360 grados para redes 5G avanzadas, así como la optimización de la captura y transmisión de dichas aplicaciones. Las aplicaciones fueron desarrolladas para verticales como educación, entrevistas y entretenimiento.

Los principales objetivos alcanzados se detallan a continuación:

- **Identificación de las tecnologías habilitadoras** para aplicaciones inmersivas con video inmersivo 180 y 360 grados, en el contexto de 5G avanzado y 6G.
- **Diseño e implementación** de un sistema de video multicámara inmersivo para el Laboratorio Inmersivo de la UPV.
- **Optimización en la transmisión de aplicaciones inmersivas** en tiempo real en redes 5G avanzadas, considerando distintos parámetros de vídeo, tales como formato, resolución espacial, resolución temporal, códec de vídeo, protocolo de *Streaming*, y protocolo de transporte.
- **Demostraciones y pilotos de casos de uso** utilizando redes 5G avanzadas.
- **Evaluar el valor de la tecnología** para los distintos casos de uso desplegados

Los casos de uso diseñados y desplegados durante el proyecto fueron:

1. **Entrevistas inmersivas.** En este caso de uso, se utilizó el sistema desarrollado para realizar entrevistas con la cámara 180 grados y de 2 a 4 participantes, con la inserción de in video con chroma key.
2. **Entretenimiento inmersivo.** En este caso de uso, se utilizó el sistema desarrollado para cubrir un pequeño concierto con músicos, utilizando la cámara 180 y 360 grados.
3. **Educación inmersiva.** En este caso de uso, se utilizó el sistema desarrollado para docencia, utilizando la cámara 180 y 360 grados con la inserción de imagen Picture-In-Picture con la presentación PowerPoint.



Ilustración 1. Imágenes del caso de uso de entrevistas inmersivas.

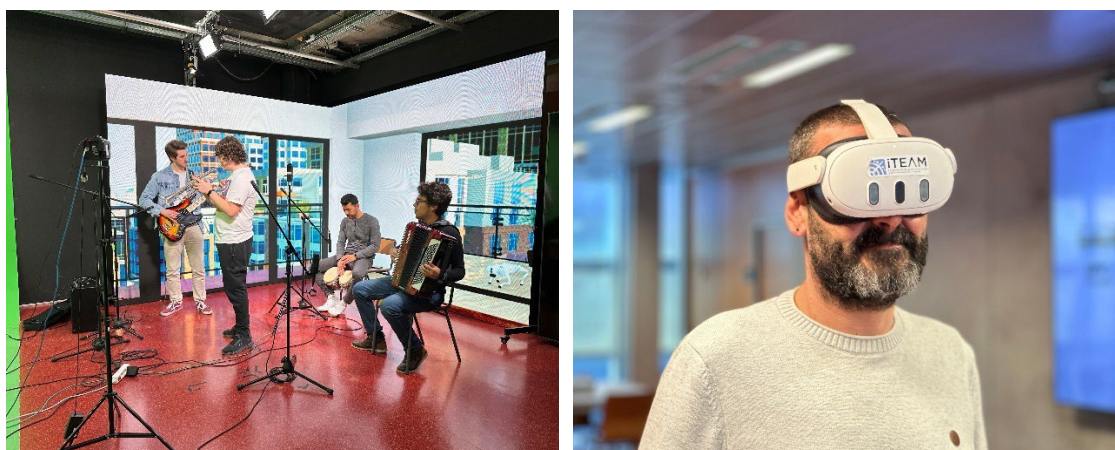


Ilustración 2. Imágenes del caso de uso de entretenimiento inmersivo.

Diseño del sistema de transmisión de video 180/360 inmersivo

El laboratorio inmersivo se dotó con un sistema de transmisión inmersiva en tiempo real, compuesto por:

- **Dispositivos de captura inmersivos.** En este caso, se dotó al laboratorio de dos cámaras 180 grados y una cámara 360 grados, todas con posibilidad de realizar transmisiones de audio y video en tiempo real a través de IP, a través de ethernet o Wi-Fi. Las cámaras tienen la posibilidad de ajustar configuraciones de resolución espacial y temporal, así como el protocolo de transmisión utilizado y tasa binaria. Además, las cámaras tienen la posibilidad de capturar audio con micrófonos internos o bien inyectar una señal de micrófono externo o mesa de mezclas.
- **Servidores de procesamiento de imagen y transmisión.** Se dotó al laboratorio de dos servidores, conectados directamente al núcleo de red 5G de UPV, formando así un pequeño *cluster* de *Edge Computing*. El primer servidor realiza las funciones de stitching y corrección de perspectiva de imagen. El segundo servidor realiza la transmisión de los contenidos hacia los dispositivos finales.

- **Dispositivos de usuario.** Se dotó al laboratorio de varios dispositivos de usuario para visualizar los contenidos inmersivos. Estos incluyen gafas Meta Quest 2, Meta Quest 3 y Meta Quest 3S, además de teléfonos inteligentes Xiaomi.
- **Dispositivos de conexión con la red 5G.** Se utilizaron varios dispositivos 5G (módems y CPE) para dar conectividad 5G a los dispositivos de captura y de usuario, disponibles en UPV.

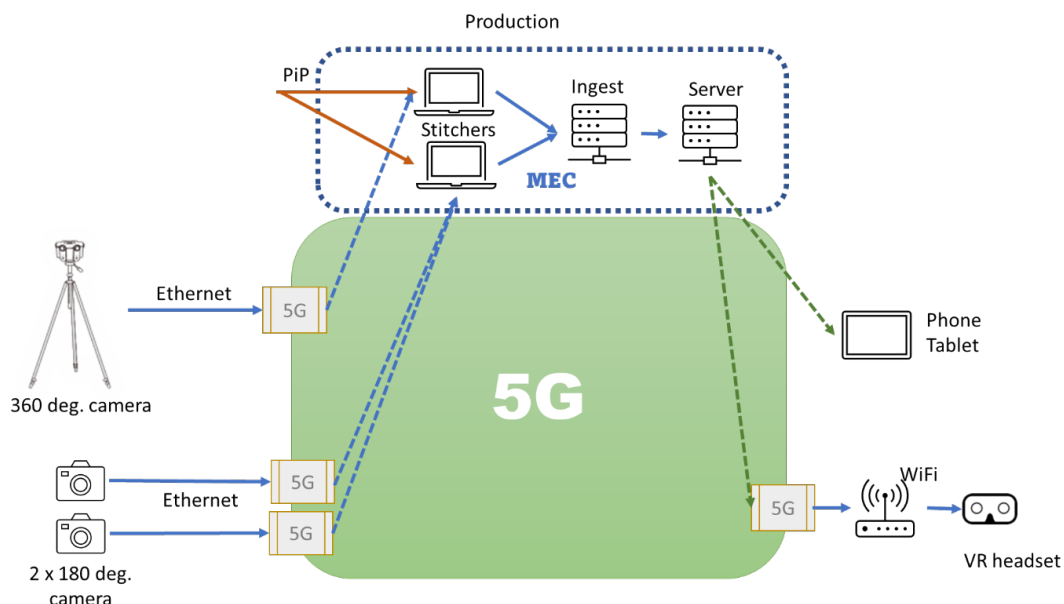


Ilustración 3. Arquitectura general del sistema inmersivo.

Respecto al **software**, para que los usuarios pudieran visualizar remotamente los contenidos, se desarrolló una **aplicación de usuario** específica para el proyecto. Esta aplicación está diseñada para dispositivos Quest, Android e iOS y permite entrar a un hall virtual, en el que se puede seleccionar el contenido a visualizar (las distintas cámaras desplegadas). Esta aplicación se desarrolló en dos fases: en la primera fase, solo era posible seleccionar una de las cámaras desplegadas, mientras que la segunda fase **permitiría al usuario seleccionar visualizar una de las múltiples cámaras desplegadas de forma fácil y rápida en tiempo real.**

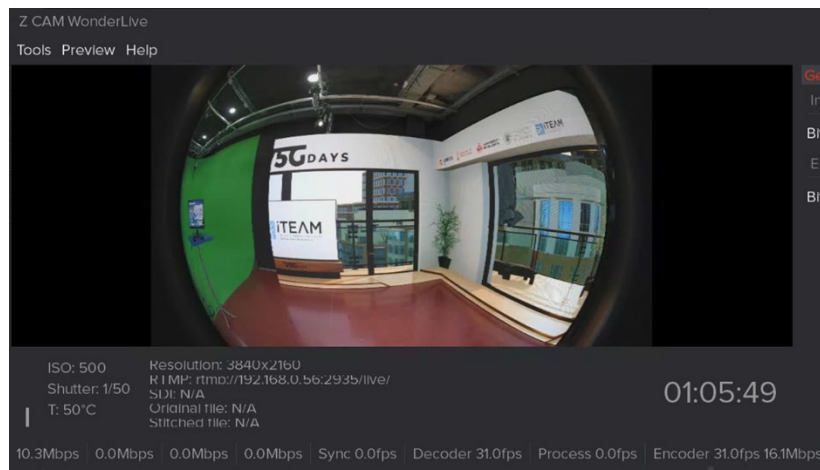


Ilustración 4. Transmisión recibida de cámara 180 sin corregir perspectiva.



Ilustración 5. Software de producción y stitching Voysys.

Experimentos y demostraciones finales

Las **demostraciones finales del caso de uso** a través de la red 5G han sido testeadas en **condiciones reales de operación** en el **evento V5GDAYS 2023 para el caso de uso de entrevistas**, así como en los experimentos en entorno real con participantes realizados para los casos de uso de entretenimiento y educación.

Se llevó a cabo **la difusión del proyecto** en diferentes medios de comunicación como prensa, YouTube, redes sociales – LinkedIn. A continuación, se muestran algunos de los enlaces de los medios de comunicación:

- https://www.linkedin.com/posts/ybvr_innovation-ieee-technology-activity-7236397212631498752-edOI?utm_source=share&utm_medium=member_desktop&rcm=ACoAAADoArsB1MxVilnXCeDqu-oKtpkPgpM15LM
- <https://x.com/YBVR360/status/1725196516589985970>

Para la preparación de las demostraciones, se llevaron a cabo numerosos experimentos previos en los que se determinó la configuración ideal de resolución espacial y temporal, además de la tasa binaria de cada cámara dependiendo del rendimiento de la red 5G.

Mediciones de red y aplicación

La red utilizada para los casos de uso de este proyecto es la **red de interior 5G** disponible en las instalaciones interiores del laboratorio de comunicaciones inmersivas de UPV. La red utilizada trabaja en la **banda n78, con 100 MHz de ancho de banda**. Por cada cámara desplegada en el laboratorio, se utiliza un módem 5G (Siemens, Advantech o ZTE) y por cada dispositivo de usuario (gafas o teléfono móvil) se utiliza un módem 5G o teléfono inteligente a modo de módem.

Para determinar la **configuración ideal de cámaras**, se realizó mediante prueba-error una optimización de la resolución espacial y temporal y tasa binaria. Durante estos experimentos, se estableció esta configuración base:

Tabla 1. Configuración de cámaras.

Ajuste	Valor
Resolución espacial cámara 180	3840 x 2160 pixels
Resolución temporal cámara 180	30fps inicialmente / 60 fps al final
Protocolo de transmisión cámara 180	RTSP inicialmente y SRT finalmente
Bitrate de transmisión cámara 180	12.5 Mbps
Resolución espacial cámara 360	3840 x 1920 pixels
Resolución temporal cámara 360	30 frames por segundo
Protocolo de transmisión cámara 360	RTMP
Bitrate de transmisión cámara 360	10 Mbps

Para determinar si una configuración era válida, se observó el comportamiento de la aplicación de usuario final y las estadísticas obtenidas en el servidor de stitching y transmisión. Se estableció el máximo de resolución temporal y espacial en 4Kp50 y se eligieron distintas tasas binarias, incrementándolas gradualmente, hasta que se observaban artefactos de audio o vídeo en la aplicación de usuario o se detectaba que la tasa binaria media de recepción en los servidores era mucho menor que la configurada en las cámaras.

Esta configuración base se utilizaría para **todos los casos de uso**, excepto en el caso de uso de entrevistas. Esto es debido a que este caso de uso se realizó al principio del proyecto, cuando la aplicación soportaba una sola cámara 180 grados. A partir de ese primer escenario base, se extendió la aplicación y se establecieron los ajustes descritos en la Tabla 1.

Durante los experimentos de configuración del proyecto, se realizaron **mediciones de latencia del sistema**. Para ello, se utilizó prueba de cronómetro, en la cual se posiciona un **cronómetro** de manera que sea visible para el operador. Con una de las cámaras 180 o 360, se graba este cronómetro y se observa en la aplicación de usuario

(teléfono móvil en este caso) el cronómetro. La diferencia de tiempo entre el cronómetro real y el grabado se puede considerar como la latencia extremo-a-extremo (Ilustración 6). La latencia se ajustó en el *software* de transmisión para adaptarse a las cámaras que proporcionasen mayor latencia, de tal manera que el video y audio de todas las cámaras estuviese sincronizado. **El valor de latencia extremo-a-extremo con el sistema actual es de aproximadamente 5 segundos.**

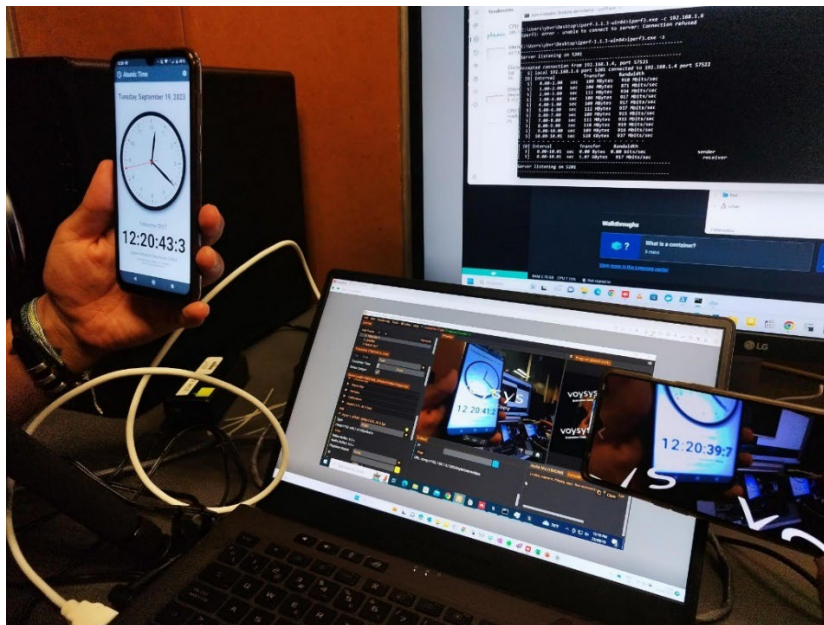


Ilustración 6. Prueba de latencia extremo-a-extremo.

Mediciones de calidad de experiencia

Durante el proyecto, se realizaron encuestas de **calidad de experiencia** por cada experimento de caso de uso. Esto es, una encuesta para el caso de uso de entrevistas, otro para el de entretenimiento y otro para el de educación. Estas encuestas de calidad de experiencia tenían como objetivo **recoger las opiniones de usuario y conocer el valor de las experiencias inmersivas** para distintos **casos de uso**. Estas encuestas de calidad de experiencia fueron completamente anónimas y los datos recogidos en ellas son solo accesibles al personal de investigación de UPV. No es posible trazar los datos recogidos con individuos concretos, por lo que se refuerza aún más la anonimidad de los resultados.

Las encuestas de calidad de experiencia tienen una **parte común** que determinan datos básicos de los usuarios, tales como:

1. Edad
2. Afiliación con UPV
3. Nivel de estudios
4. Tipo de contenido multimedia que el usuario consume y con qué frecuencia
5. Experiencia previa con contenido inmersivo

6. Calidad de audio y video de la aplicación

Además de estas preguntas, por cada caso de uso, se establecieron unas **preguntas específicas y adaptadas** para cada uno de ellos. Dependiendo de la naturaleza de las preguntas específicas, las respuestas posibles pueden ser texto, respuesta múltiple, múltiples opciones o en escala Likert (puntuación de 1 a 5).

Las encuestas de calidad de experiencia se realizaron en los 3 experimentos reales de cada caso de uso, a los que asistieron el siguiente **número de participantes**:

Experimento/Caso de uso	Número de participantes
Entrevistas inmersivas	19
Entretenimiento inmersivo	8
Educación inmersiva	7

A continuación, se resumen algunos de los resultados generales obtenidos durante los experimentos.

Entrevistas inmersivas

- La mayoría de los participantes tienen formación de máster universitario

Entretenimiento inmersivo

- La mayoría de los participantes son docentes o personal de investigación.
- La mayoría de los participantes tienen grado de máster.
- La mayoría de los participantes consumen contenido de video tradicional muy frecuentemente y la mayoría con fines de entretenimiento.
- La mayoría de los participantes alguna vez han probado tecnologías inmersivas, pero no las utilizan de manera habitual.
- La mayoría de los participantes experimentan música en directo de forma habitual.
- Todos los usuarios disfrutaron de la experiencia inmersiva.
- Los usuarios indican que les resultó fácil o muy fácil seguir la actuación de los músicos en forma inmersiva.
- La mayoría de los participantes indican que les resultó muy fácil usar la aplicación.
- La mayoría de los participantes prefieren la experiencia real a la inmersiva.
- La mayoría de los participantes prefieren la experiencia en gafas en contra del teléfono móvil.

- La mayoría de los participantes reportaron que no sufrieron malestar físico.
- La media de calidad visual es de 4, al igual que la calidad de audio.
- La mayoría de los participantes ven un buen futuro para esta clase de aplicaciones inmersivas.
- La media de satisfacción de la aplicación es de 4 de 5.
- Algunos de los comentarios de problemas experimentados son:
 - *La transición entre vistas 180° y 360°*
 - *En ocasiones se cortaba el audio. También, según mirase a un sitio u otro, no salía la escena completa, sino que se veía el hall.*
 - *Cuando se acaba lo que es el fondo y aparece el relleno del campus. Me ha sorprendido que se escuchara tan bien con las gafas a pesar de no llevar cascos.*
 - *Corte en la unión de imágenes en la cámara 360° que recaía justo en uno de los músicos y no se apreciaba bien su instrumento.*
 - *Calidad de video un poco mala.*
-

Educación inmersiva

- La mayoría de los participantes son docentes y/o personal de investigación.
- La mayoría de los participantes tienen nivel de educación de máster.
- La mayoría de los participantes consumen contenido de video muy frecuentemente, mayoritariamente de entretenimiento y educativo.
- La mayoría de los participantes utilizan de forma regular tecnologías inmersivas.
- La mayoría de los participantes han disfrutado de la aplicación inmersiva.
- La mayoría de los participantes han sido estudiantes hace menos de un año.
- La mayoría de los participantes consideran que ha sido fácil seguir el contenido en forma tradicional (71.4%).
- La mayoría de los participantes consideran que ha sido fácil seguir el contenido en forma inmersiva (71.4%).
- A la mayoría de los participantes les ha resultado muy fácil o fácil utilizar la aplicación inmersiva.
- A la mayoría de los participantes les pareció mejor la experiencia inmersiva que la tradicional.
- La mayoría de los participantes consideran que aprenden mejor con el método inmersivo, aunque no por mucha diferencia (57.1% vs 42.9% de los usuarios).
- La mayoría de los participantes consideran que la segunda parte del examen es más complicada que la primera.
- La mayoría de los participantes experimentaron Ligero malestar, pero totalmente aceptable al experimentar el contenido inmersivo.

- La mayoría de los usuarios puntuaron la calidad visual de la aplicación inmersiva en 3 puntos sobre 5.
- La mayoría de los usuarios puntuaron la calidad de audio de la aplicación inmersiva en 5 puntos sobre 5.
- La mayoría de los usuarios puntuaron la calidad visual de la videoconferencia en 3 y 5 puntos sobre 5.
- La mayoría de los usuarios puntuaron la calidad de audio de la videoconferencia en 3, 4 y 5 puntos sobre 5.
- La mayoría de los usuarios consideran que hay un futuro para este tipo de aplicaciones inmersivas para educación.
- Algunos usuarios señalaron que sienten que los métodos inmersivos mejoran la concentración, ya que la utilización de las gafas obliga al usuario a estar atento a la lección, sin distracciones, mientras que, con la videoconferencia normal, es mucho más fácil distraerse.
- Debido al reducido número de usuarios, no es posible realizar un análisis ANOVA de un factor y obtener resultados estadísticamente relevantes.

Mediciones de aprendizaje

Para el caso de uso de educación, además de realizar encuestas de calidad de experiencia, se planteó realizar un **estudio de aprendizaje**, mediante el cual se evaluaría si las tecnologías inmersivas **facilitan la retención de conocimientos** respecto a las videollamadas tradicionales. Para ello, se estableció una metodología basada en estudio de varianza (ANOVA) de un factor con 2 niveles, para comparar el **rendimiento** (mediante exámenes) entre un contenido impartido por **videoconferencia con Microsoft Teams** y uno impartido con la **aplicación inmersiva**. Los participantes fueron evaluados, pero debido al bajo número de participantes, no es posible extraer resultados estadísticamente significativos. Sin embargo, la metodología establecida **sienta las bases para realizar estudios sociológicos a gran escala en el futuro**.

Trabajo futuro

El proyecto ha presentado un gran potencial de evolución hacia distintos casos de uso aumentados, que pueden mejorar el aprendizaje, la inmersión, y el entretenimiento. Se han identificado las siguientes líneas de trabajo futuras:

- **Prueba de otros protocolos de transmisión.** Se han detectado problemas de fiabilidad de los protocolos tradicionales de streaming (RTMP, RTSP...) con canales sensibles al ruido, como los canales radio utilizados en las redes 5G y WiFi. Con el uso en este proyecto del protocolo SRT (Secure Reliable Transport) se ha mejorado sensiblemente la fiabilidad de la transmisión, pero aún queda lejos de las comunicaciones de fibra. Sería interesante estudiar más a fondo

estos problemas y explorar nuevos protocolos que están apareciendo en el mercado, como SST (Safe Streams Transport).

- **El streaming de video sobre 5G debe eliminar los actuales cuellos de botella** para aprovecharse de las enormes capacidades de baja latencia que ofrecen de los nuevos estándares de redes inalámbricas. La mejora de la fiabilidad debe ser independiente de la reducción de la latencia, que es imprescindible para aplicaciones interactivas.
- En los casos de uso de entrevistas y docencia se ha detectado la **carencia de soluciones de interacción en entornos inmersivos**. Abrir la posibilidad de que los alumnos pregunten al profesor en medio de una clase inmersiva, es algo que merece la pena investigar y poner los medios para desarrollar soluciones adecuadas.
- **El software de producción** empleado en este proyecto (Voysys) es una herramienta potente de la que solo se emplean unas pocas funcionalidades. La licencia necesaria para su uso sin marca de agua supone un gasto recurrente importante.
- **Utilización de la tecnología en otros casos de uso** que puedan beneficiarse del video inmersivo 180 y 360 grados.
- **Utilización de entornos con 6 grados de libertad (entornos virtuales) con video 180 y 360 integrado.**
- **Incorporación de captura de video estereoscópico.**
- Utilización de **funcionalidades avanzadas** de baja latencia y tasa binaria aumentada de las redes 5G/6G.
- Utilización de APIs 5G para la solicitud de calidad bajo demanda automatizada.
- Realización de **experimentos a gran escala (gran número de usuarios)** para obtener resultados estadísticos relevantes sobre el valor de las tecnologías inmersivas en la sociedad.