

JORNADAS

GIPUZKOA TECNOLÓGICA

Gemelos Digitales e IA Responsable

Tecnologías Habilitadoras para la Empresa del Futuro

Samara Ruiz

samara.ruiz@immersia.eu



Short bio

8º E.G.B

De mayor, informática



Ingeniería Informática

Análisis de datos, IA

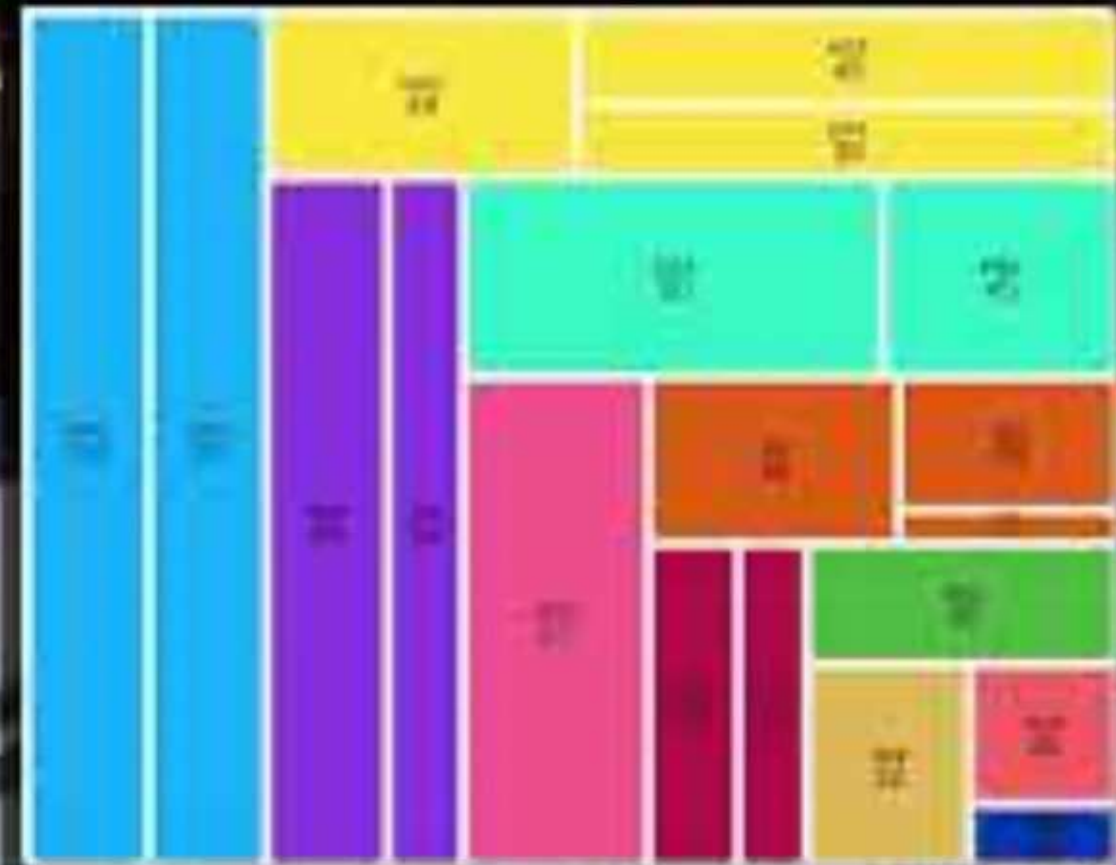
PhD

Learning Analytics



Immersia

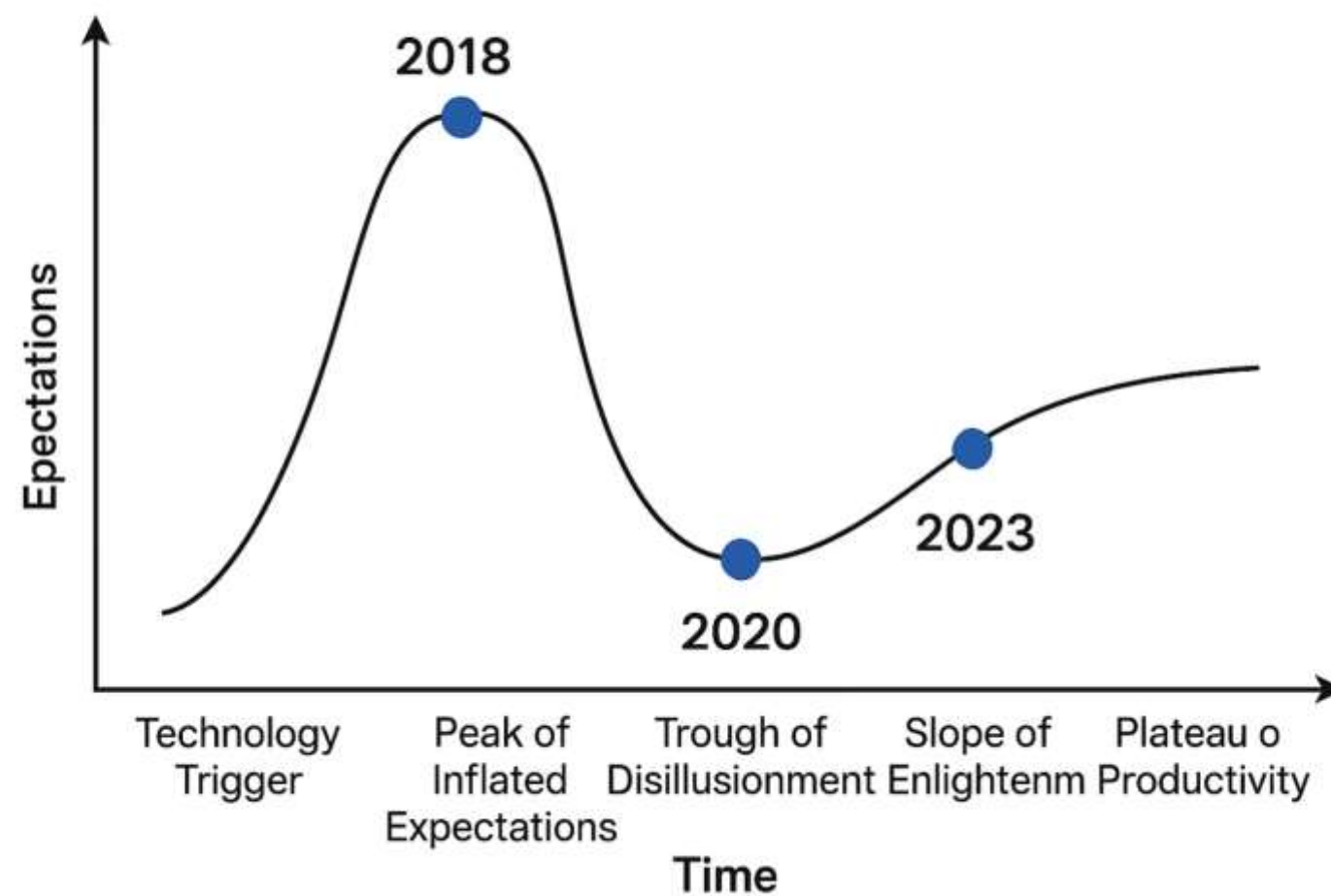
CTO & Co-founder



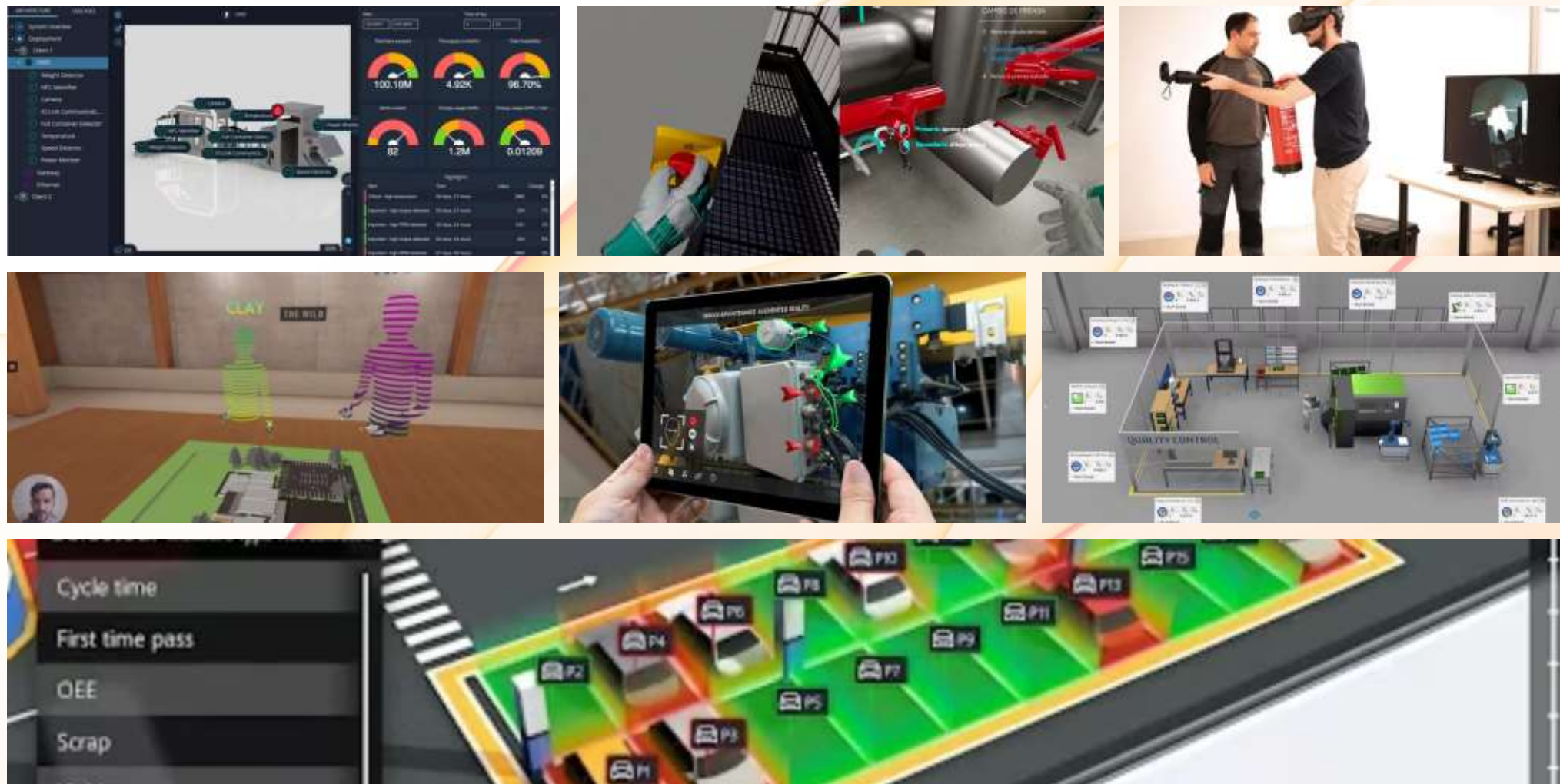
#digitaltwin



Gartner Hype Cycle for Digital Twin



Qué es y qué no es un gemelo digital





La NASA

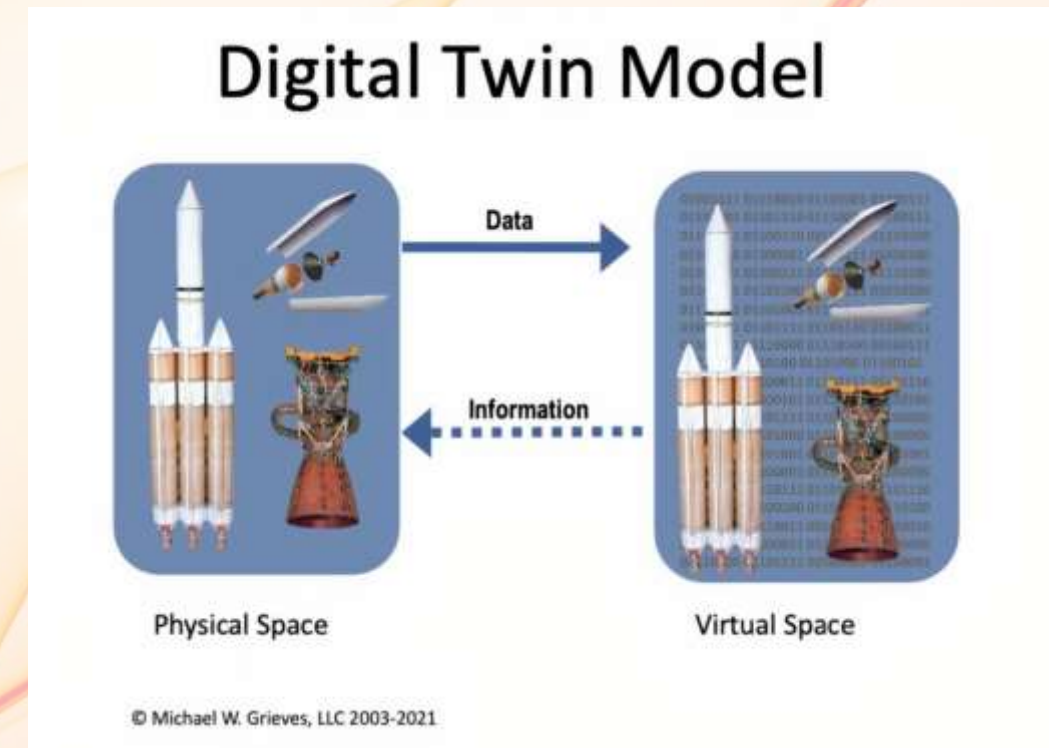
"La simulación probabilística multi-física y multi-escala de un vehículo o sistema que utiliza los mejores modelos físicos disponibles, actualizaciones los sensores, historia de vuelo, etc, para simular la vida de su gemelo volador. Es ultra realista y puede considerar uno o más sistemas de vehículos importantes e interdependientes."

Grieves, 2002

GESTIÓN DE VIDA DEL PRODUCTO

Representación virtual de algo físico. Una réplica digital que coexiste junto a su contraparte física para mejorar el diseño, monitoreo y gestión de productos a lo largo de su vida útil.

- Gemelo físico
- Gemelo digital
- Hilo digital



Tecnologías habilitadoras



IoT

Boschert & Rosen, 2016

"Enlace entre el mundo físico y el virtual, donde los **datos obtenidos a través de sensores** permiten representación precisa y actualizada del sistema."



IA

Tao & Qi, 2019

"Una réplica virtual que utiliza **algoritmos de IA para predecir** el rendimiento, diagnosticar fallas y optimizar el comportamiento de un sistema a partir de datos en tiempo real e históricos."



Visualización de datos

Tao et al., 2019

"Una plataforma digital centralizada que integra datos de diversas fuentes y dispositivos en una única interfaz conectada, ofreciendo una **vista unificada y en tiempo real** del estado de un sistema físico."



Modelado y simulación

Fuller et al. 2020

"Una réplica digital que emplea **técnicas de simulación para predecir** cómo un sistema responderá a diferentes condiciones ambientales o ajustes operacionales, permitiendo una planificación de contingencias y una mejora en el diseño del sistema."



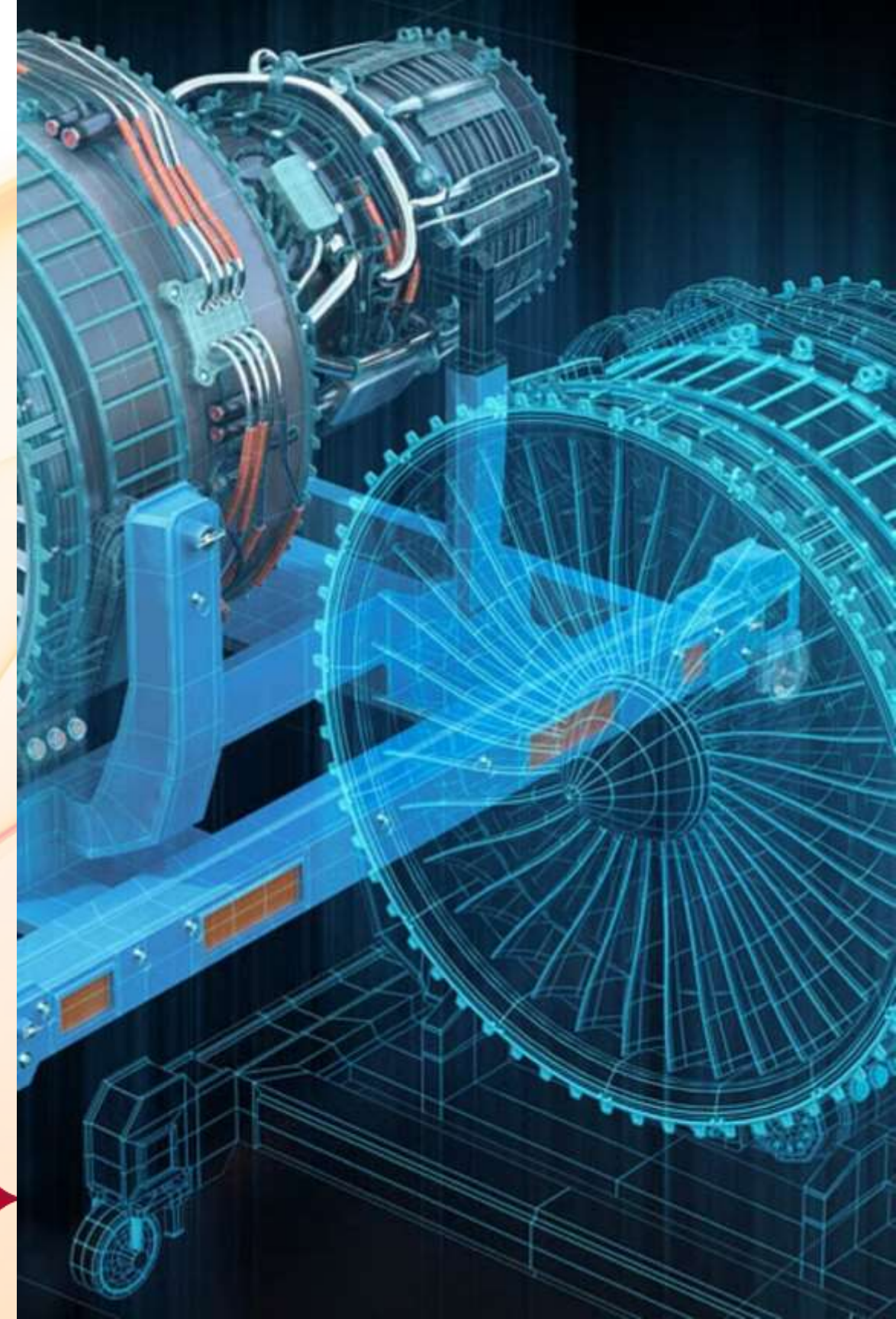
Big data y análisis de datos

Kaur et al., 2020

"Una entidad que se beneficia de **datos masivos de múltiples fuentes** para generar modelos precisos de predicción, optimización y mantenimiento preventivo de sistemas complejos."

ISO 23247-1, 2021

"Una representación digital de una entidad del mundo real, que está sincronizada con ella en una frecuencia y fidelidad adecuadas para el propósito de la simulación, análisis o toma de decisiones basada en datos."





ChatGPT

"Un gemelo digital es una representación virtual precisa y dinámica de un objeto, proceso o sistema físico, que se utiliza para simular, analizar y optimizar su rendimiento a lo largo de su ciclo de vida."

Marzo de 2025

Historia y evolución del concepto

1960s - Origen

La NASA crea modelos físicos de sus naves espaciales y sistemas que funcionan como primeros ejemplos de gemelos digitales, utilizados para monitorear su rendimiento y prever el comportamiento en situaciones críticas.

Era Industria 4.0 - Crecimiento

El crecimiento del concepto de gemelo digital se debe a los avances en IoT, IA, big data y computación en la nube, lo que ha llevado a una adopción generalizada en la manufactura, la ingeniería y la gestión de infraestructuras.

Futuro - Alcance

La convergencia de los gemelos digitales con tecnologías emergentes, como el metaverso, aprendizaje automático (machine learning) y la automatización avanzada, ampliará aún su alcance y capacidades.

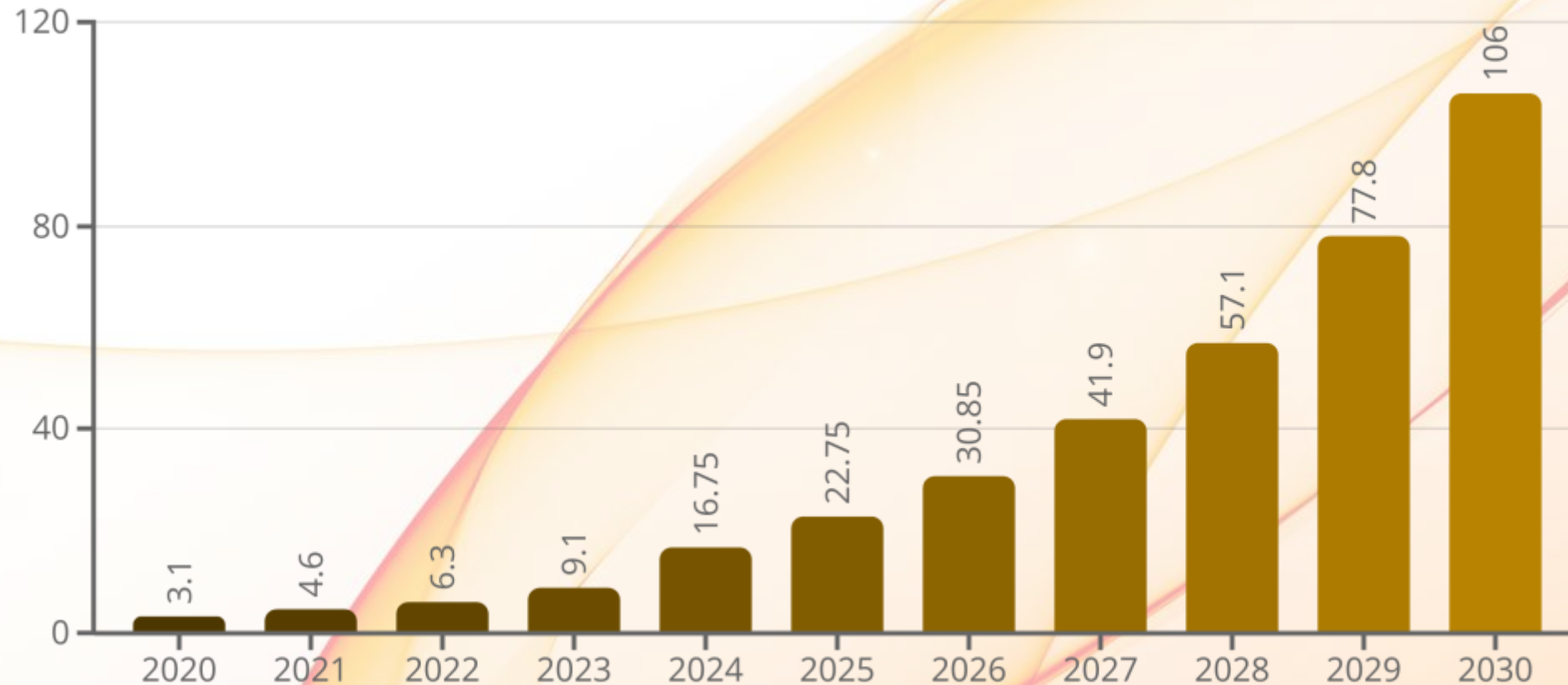
2002 - Oficialidad

El investigador Michael Grieves, del Instituto de Tecnología de de Florida, acuñó oficialmente el término "gemelo digital" con el objetivo de mejorar el diseño y la gestión de productos.

Hoy - Adopción masiva

Los gemelos digitales abarcan diversas aplicaciones, como procesos industriales, ciudades inteligentes, vehículos autónomos y medicina personalizada, transformando la toma de decisiones y la eficiencia operativa.

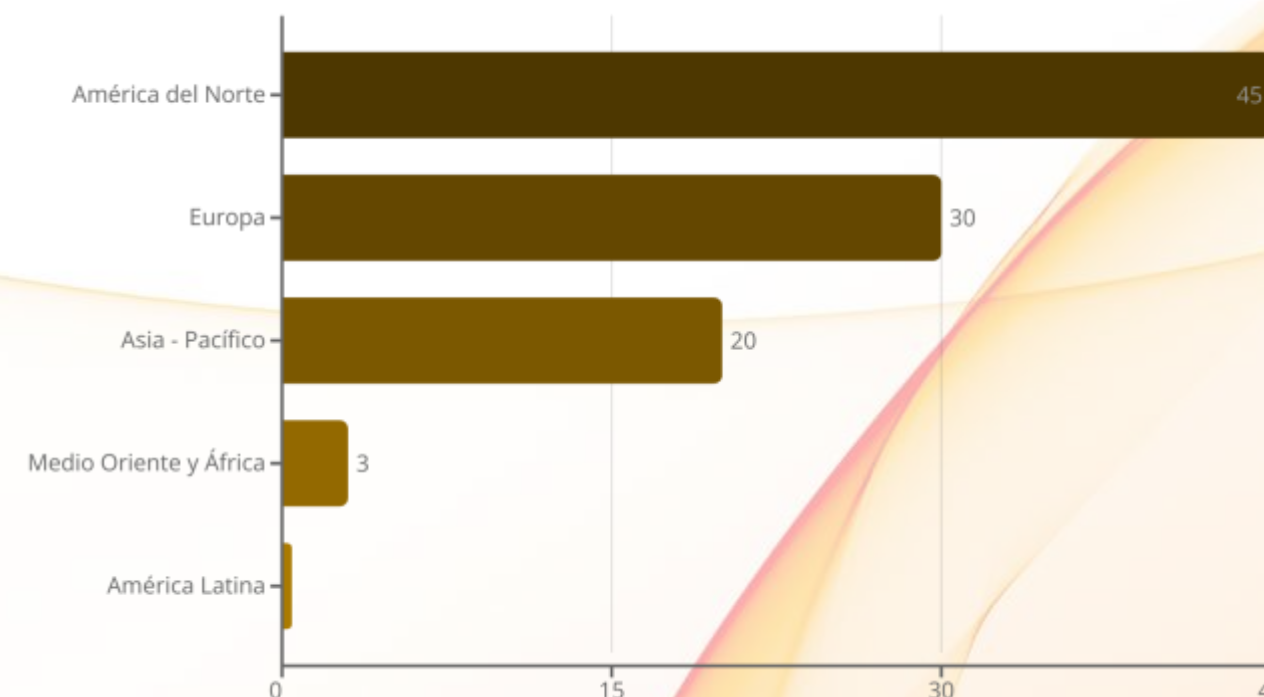
Crecimiento del mercado



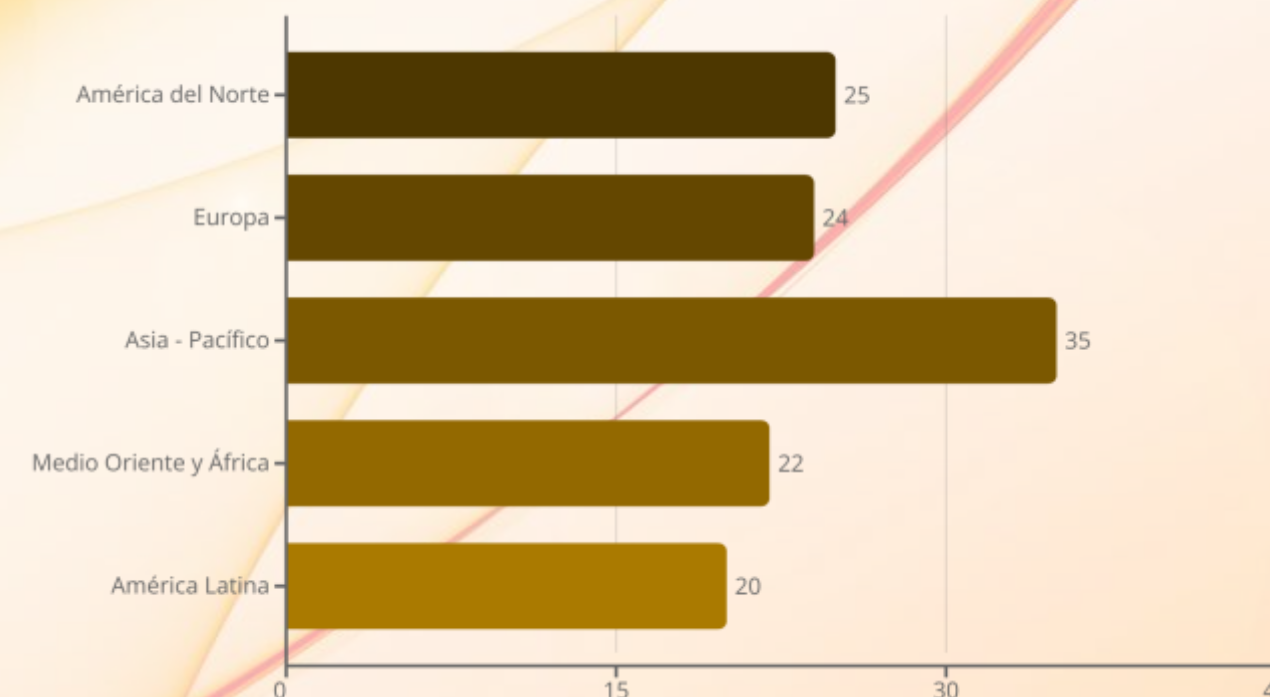
CAGR del 35.7%, alcanzando los 106 mil millones de dólares para 2030.

Tendencias del mercado

Cuota de mercado 2023



CAGR



América del Norte y Europa dominan actualmente, Asia-Pacífico podría liderar el crecimiento futuro en este mercado debido a sus ambiciosos proyectos de digitalización en varias industrias clave.

El gemelo digital por regiones

Norte América

Automotriz, Energía, Defensa

Alta inversión en infraestructura digital, adopción temprana de tecnologías gemelas digitales en manufactura y energía, apoyo de políticas para la digitalización.

Europa

Manufactura, Construcción, Salud

Avance en manufactura inteligente, políticas de sostenibilidad que impulsan la adopción de gemelos digitales para eficiencia energética y reducción de residuos.

Asia-Pacífico

Electrónica, Construcción, Smart Cities

Crecimiento en la construcción de ciudades inteligentes, aumento de la inversión en IoT y AI, y expansión de sectores tecnológicos en China, Japón, India y Corea.

Este Medio y África

Petróleo y Gas, Infraestructura

Implementación de gemelos digitales para mantenimiento predictivo en petróleo y gas, aumento de la digitalización en infraestructura pública.

América Latina

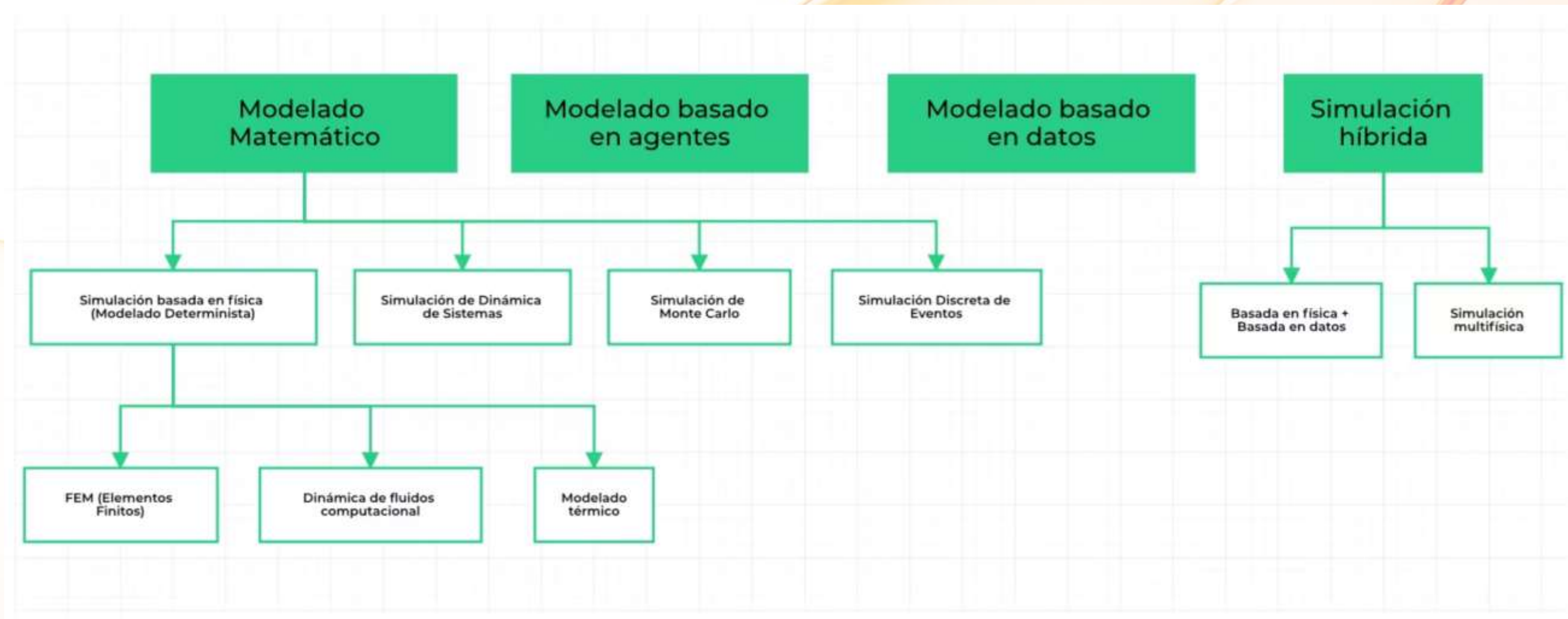
Agricultura, Energía

Mayor adopción en agricultura digital y monitoreo de energías renovables, crecimiento en la adopción de IoT en la industria.

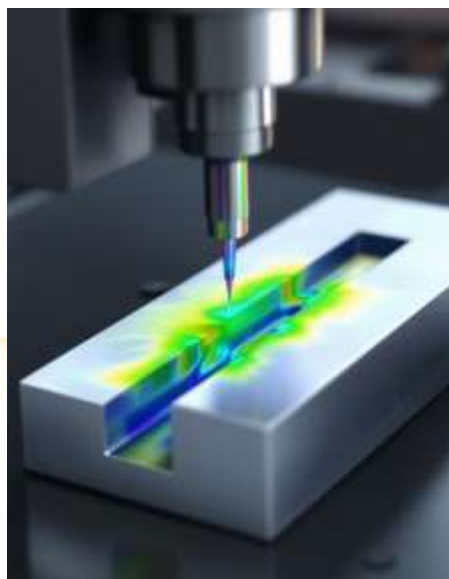
Arquitectura



Modelado y simulación



Modelado y simulación



$$\sigma_{xM} = \sqrt{\frac{1}{2} [(\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2] + 3(\tau_{xy}^2 + \tau_{yz}^2 + \tau_{xz}^2)}$$

FEM

Simular el proceso de estampado ya que las piezas de metal están sometidas a fuerzas extremas, lo que puede causar deformaciones o, en algunos casos, fracturas si el proceso no se gestiona correctamente.



$$\frac{dP(t)}{dt} = r P(t) \left(1 - \frac{P(t)}{K} \right)$$

DINÁMICA DE SISTEMAS

Simular cómo los niveles de producción y el inventario cambian en respuesta a la demanda fluctuante del mercado.



$$X = \frac{\text{Cantidad de productos procesados}}{\text{Tiempo total}}$$

SIMULACIÓN DE MONTECARLO

Simular la variabilidad en el tiempo de producción de diferentes etapas del proceso. considerando factores como el tiempo de ciclo de las máquinas y la disponibilidad de materiales, la simulación de monte carlo permite prever el rango de tiempo de producción total, ayudando a los gerentes a planificar mejor la producción y reducir el riesgo de incumplimiento de plazos.



BASADO EN AGENTES

```
if temperatura > 80 and vibracion > 10:
    detener_maquina = True
```

Optimizar la logística interna en un entorno de producción:

Cada robot o vehículo autónomo se trata como un agente individual con reglas propias de comportamiento, que interactúa con otros agentes y con el entorno. Esto permite replicar y optimizar el flujo de materiales y productos entre las distintas estaciones de la fábrica.

Modelado y simulación

Basado en datos



No requiere un modelo físico

Es útil cuando no se dispone de un modelo físico exacto (por ejemplo, por procesos muy complejos o no completamente conocidos)



Aprende de datos reales (Históricos y en Tiempo Real)

Modelos entrenados con grandes conjuntos de datos de SCADAS, MES, etc.



Permite predicción y optimización

Puede predecir el fallo de una máquina o la calidad de un lote antes de que finalice su producción..



Optimización Continua

Modelos actualizados con datos nuevos para mayor precisión.

Posibles problemas y riesgos



Selección incorrecta de variables

Modelos poco útiles o con alto error en condiciones reales.



Modelos sobreajustados (overfitting)

Sin validación práctica, los modelos pueden estar ajustados al dataset pero no generalizan bien.



Interpretabilidad limitada

Soluciones de IA "caja negra" dificultan la adopción por parte de equipos técnicos operativos.



Falta de comprensión del proceso productivo

No conocer cómo fluye el proceso, las restricciones operativas o condiciones físicas puede llevar a modelos inviables.



Optimización de objetivos irrelevantes



Ignorar restricciones de planta o mantenimiento



Resistencia del personal técnico o de planta

Si los modelos no son comprensibles ni útiles para los operativos, serán descartados o ignorados.



Costes ocultos y baja escalabilidad

Desarrollo a medida complejo, sin capacidad de ser mantenido por equipos internos no expertos

Modelado Basado en Datos

AutoML

Las herramientas AutoML facilitan el desarrollo y optimización de modelos de aprendizaje automático a usuarios no expertos en ML

A

Preprocesamiento y Limpieza de Datos

Eliminación de valores atípicos, manejo de valores faltantes y normalización de datos.

U

Ingeniería de Características

Generación y selección automática de características relevantes en los datos (transformación de variables, combinación de características, etc.), optimizando las entradas para cada modelo.

T

Selección de Modelos

Comparación de diferentes algoritmos para identificar el mejor modelo para el problema específico.

O

Optimización de hiperparámetros

Ajuste de hiperparámetros mediante técnicas de optimización (búsqueda aleatoria, la búsqueda en cuadrícula y el optimizador bayesiano, buscando los valores óptimos para maximizar la precisión del modelo.

M

Evaluación y Validación del Modelo

Validación cruzada y métricas específicas para evaluar la precisión y rendimiento de cada modelo. Una vez identificado el mejor modelo, se realiza una validación exhaustiva para asegurar que los resultados no estén sobreajustados.

L

Implementación y despliegue Automático

Monitorización de modelos en producción en un sistema existente o en una infraestructura en la nube. Esto facilita la escalabilidad y permite realizar ajustes o actualizaciones automáticas del modelo.

Auto-ML Ventajas

Reducción del Tiempo de Desarrollo

Automatizar el proceso de selección, ajuste y evaluación de modelos reduce drásticamente el tiempo necesario para un modelo predictivo efectivo.

Accesibilidad

Facilita que usuarios sin experiencia en machine learning utilicen técnicas avanzadas de modelado, permitiendo a empresas y equipos con recursos limitados aprovechar el machine learning.

Optimización Continua

Los modelos pueden ajustarse automáticamente a medida que cambian los datos, lo que permite adaptarse rápidamente a nuevas condiciones sin requerir una reingeniería completa del modelo.

Escalabilidad

AutoML es escalable y puede gestionar grandes volúmenes de datos en infraestructuras en la nube, lo que facilita su uso en sistemas de producción y aplicaciones empresariales



Abstract words

2242-2243

1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 26

Introduction

0000-0000-0000-0000 TEL: 77777777

temperature

Years	2011	2012	2013
2014	2015	2016	2017
2018	2019	2020	2021

Quantification of Temperature

Vibrations

Weight	1,227	210 lb (95 kg)	9.1%
WHR	0.93	99 cm	1.27
Cholesterol	2.96	30.8 mmol/L	0

Spoken Part 1 (15 minutes)

Precautions

Height	1.70m	Weight	74kg
Age	2000	Age	21.00
Starters	0	Enders	0

audrey@blackstudies.org

Failure

Age	0.00	100 days	0.13
Sex	0.00	Male	1.00
Condition	1.43	Full record	0

Abstract Find out whether:

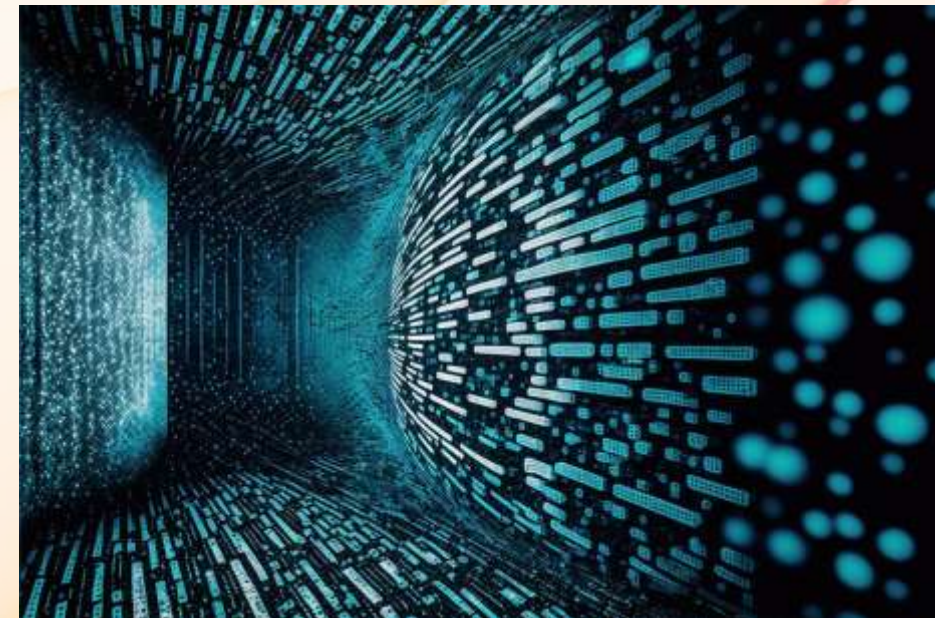


Gemelos digitales y AI Act

Un sistema de IA es “un sistema basado en técnicas computacionales que, para un conjunto de objetivos definidos por humanos, genera salidas como predicciones, recomendaciones o decisiones, que influyen en entornos físicos o virtuales.”



Utiliza algoritmos de **aprendizaje automático**, **sistemas expertos**, **modelado predictivo** o **toma de decisiones automatizada**.



Procesa datos (de sensores, históricos o simulaciones) para generar recomendaciones o actuar sobre un sistema físico.

Alto riesgo

Sistemas de IA que pueden afectar significativamente derechos o seguridad.

Infraestructura crítica

Supervisión o control de redes eléctricas, agua, transporte.

- IA que ajusta automáticamente el flujo en redes eléctricas
- Gemelo digital que gestiona presión en sistemas hidráulicos urbanos

Industria

Optimización que influye en seguridad de maquinaria o procesos.

- Gemelo digital que ajusta parámetros de una línea de producción automatizada
- IA que prevé fallos en turbinas y actúa en tiempo real

Educación

Evaluación automatizada de rendimiento académico o decisiones de acceso.

- Sistema que califica ensayos mediante PLN (procesamiento de lenguaje natural)
- IA que recomienda niveles educativos según rendimiento histórico

Recursos Humanos

Selección, evaluación o gestión de personal mediante IA.

- Plataforma que filtra currículums mediante clasificación automática
- IA que predice rendimiento o rotación del personal a partir de datos internos

Sanidad

Diagnóstico o tratamiento automatizado.

- IA que detecta tumores en imágenes de TAC.
- Algoritmo que recomienda dosis de insulina

Riesgos clave para plataformas no-code



“Caja negra” (poca explicabilidad)

Usar modelos interpretables y documentación automática



Automatización no controlada

Restringir ajustes automáticos sin validación humana



Reentrenamientos opacos

Registrar y validar cada nueva versión del modelo



Sesgos no detectados por usuarios no técnicos

Incorporar módulos automáticos de análisis de sesgos

Requisitos para sistemas de alto riesgo



Gestión de riesgos

Debe permitir definir riesgos, registrar evaluaciones, y activar controles manuales o automáticos.

01
10

Gobernanza de datos

Debe ofrecer mecanismos para validar calidad de datos, equilibrio de clases, y detección de sesgos.



Documentación técnica

Debe facilitar documentación clara de estructura, lógica, entradas, salidas y parámetros.



Logging

Debe mantener registros auditables de cada entrenamiento, predicción y actualización.



Transparencia

Debe permitir mostrar de forma clara cómo funciona el sistema (lógica, límites, recomendaciones).



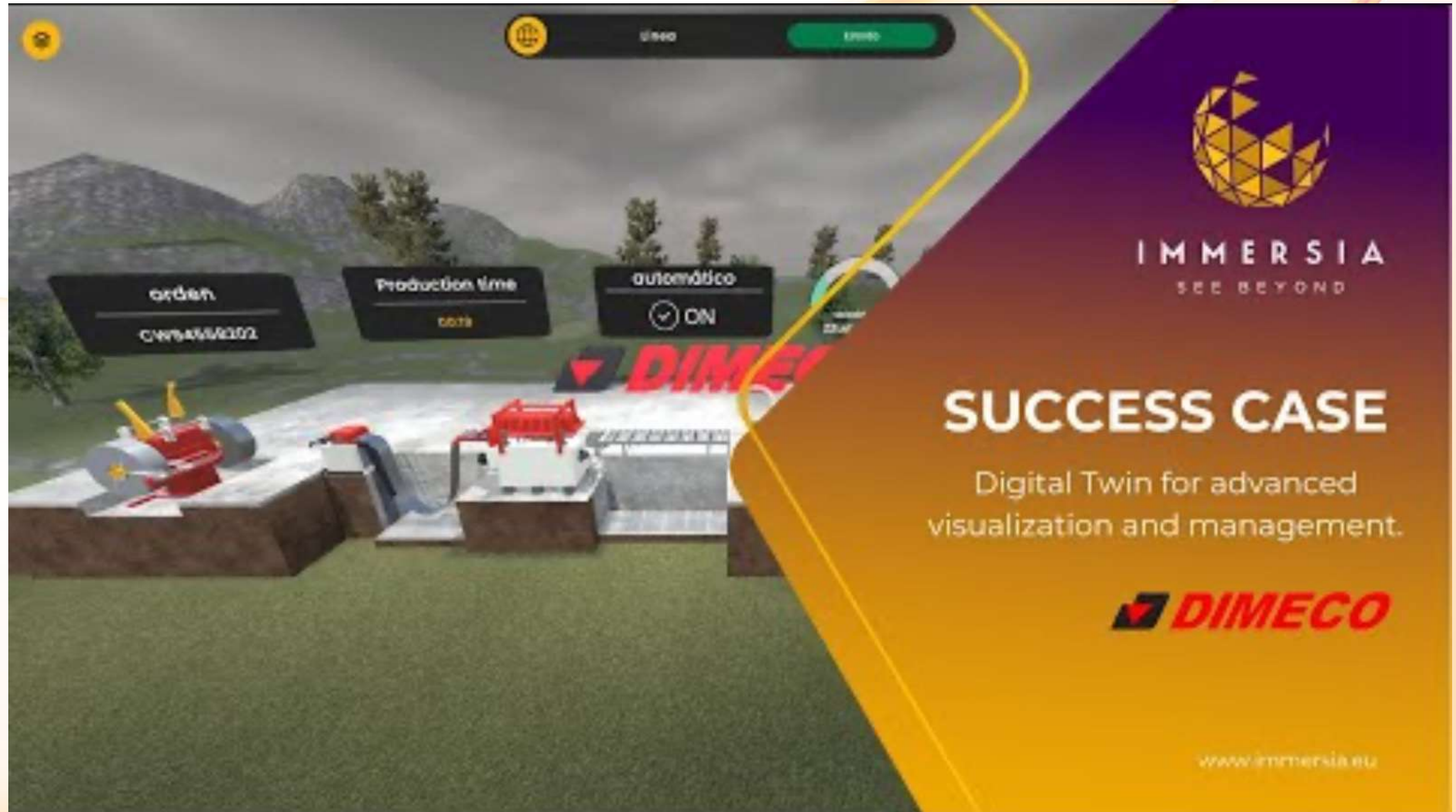
Supervisión humana

Debe incluir funciones de control humano (validación manual, revisión de salidas, override).

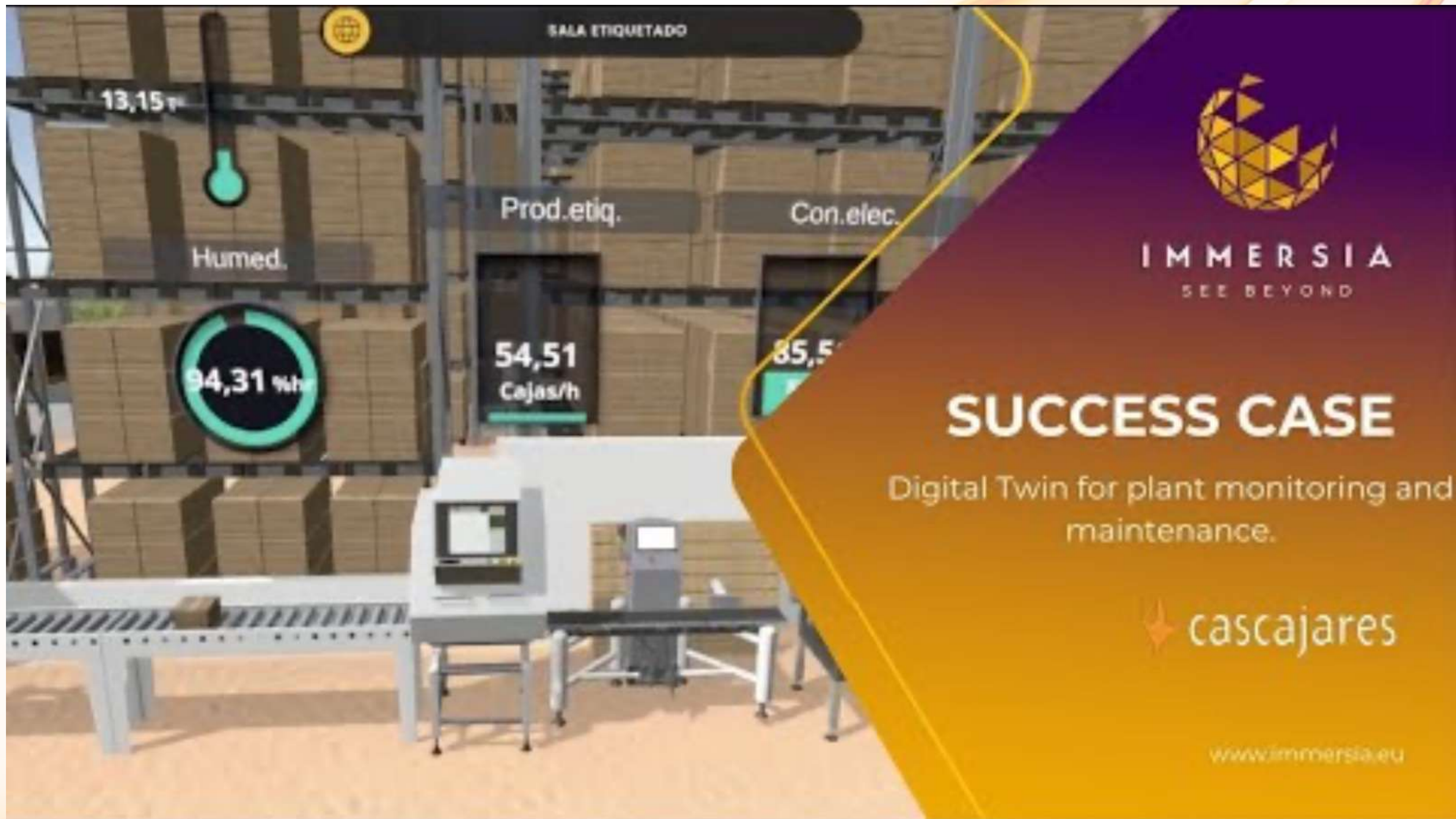


Robustez y ciberseguridad

Debe ofrecer pruebas de estrés, control de versiones, validación cruzada, y defensa frente a manipulaciones.







3D rendering of a warehouse environment with digital overlays showing production metrics. The scene includes a conveyor belt, shelving units, and a control station. The digital overlays display the following data:

- 13,15 T
- Humed. 94,31 %hr
- Prod.etiq. 54,51 Cajas/h
- Con.elec. 85,5

SALA ETIQUETADO

IMMERSIA
SEE BEYOND

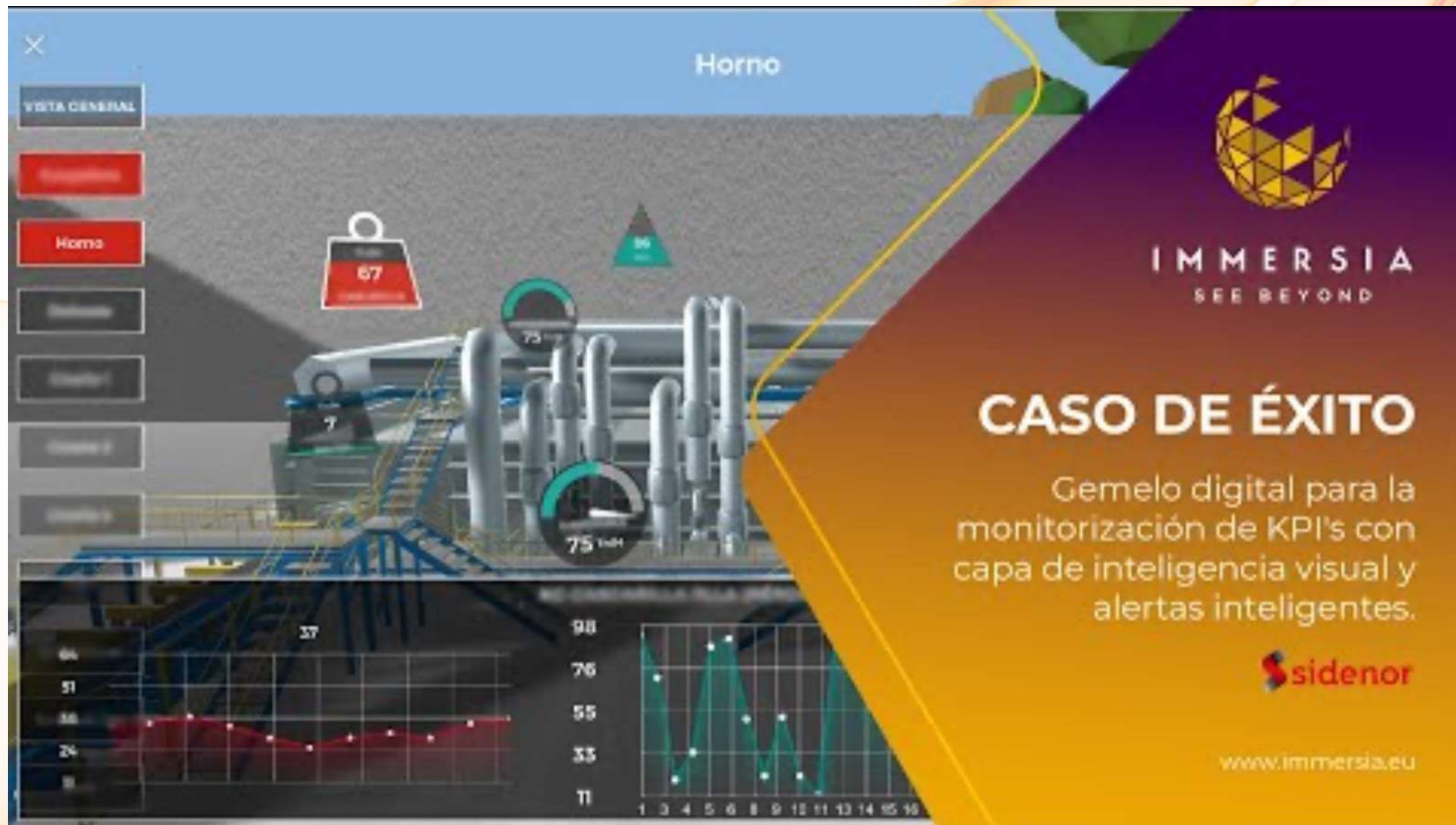
SUCCESS CASE

Digital Twin for plant monitoring and maintenance.

 cascajares

www.immersia.eu





ESKERRIK ASKO

Samara Ruiz
samara.ruiz@immersia.eu
