Humanizando la tecnología: inteligencia artificial y experiencias inmersivas al servicio de las personas

Carlos Flavián Blanco
Universidad de Zaragoza

Agradecimientos:
Es para mí un inmenso honor impartir la alocución con motivo del día del patrón de la Universidad de Zaragoza, en la festividad de San Braulio. Asumo esta responsabilidad con profunda gratitud y con plena conciencia del privilegio que representa dirigirme a esta comunidad universitaria en un acto de tan alto significado académico e institucional. Quiero pensar, además, que esta distinción trasciende mi persona y constituye también un reconocimiento a la disciplina del marketing, a su madurez académica y a su creciente aportación al conocimiento, a la sociedad y al progreso de nuestra Universidad.
Deseo expresar, en primer lugar, mi más sincero y respetuoso agradecimiento a la ***Rectora de la Universidad de Zaragoza, Rosa María Bolea Bailo****, por su confianza y por el honor que me concede al encomendarme esta intervención. Extiendo igualmente este agradecimiento a su* ***equipo de dirección****, cuyo criterio, compromiso y visión merecen nuestro respeto y en cuya capacidad confiamos para seguir llevando a la Universidad de Zaragoza a las más altas cotas de excelencia, prestigio e influencia.*
Quiero agradecer también, de manera muy especial, a mis padrinos, los profesores ***Luis V. Casaló*** *y* ***Daniel Belanche****, por su generosidad, su apoyo y su acompañamiento en este recorrido académico. Ambos forman parte del grupo de investigación METODO y lo representan con excelencia, compromiso y rigor académico. Su presencia en este acto tiene para mí un valor profundamente simbólico y personal, y constituye también una representación del propio grupo.*
Mi reconocimiento se dirige asimismo a todos y cada uno de los miembros del grupo de investigación ***METODO*** *de la Universidad de Zaragoza. Con el paso del tiempo, un grupo de investigación deja de ser únicamente un espacio de trabajo compartido para convertirse en algo mucho más valioso: una comunidad de confianza, de colaboración, de amistad y de pertenencia. Para mí, METODO representa precisamente eso. Es el grupo con el que me identifico siempre, dentro y fuera de nuestra Universidad, y muy especialmente cada vez que participo en cualquier congreso o conferencia internacional, donde llevo con orgullo su nombre y el de esta institución.*
Finalmente, deseo expresar mi gratitud más íntima y sentida a mi familia. A mi mujer, ***Menchu****, con quien he compartido no solo la vida, sino también una trayectoria investigadora paralela, desarrollada en ámbitos completamente distintos, pero estrechamente relacionados. A lo largo de los años hemos recorrido caminos académicos propios, enriquecidos por el diálogo constante, el intercambio de ideas y el respeto mutuo por nuestras respectivas disciplinas. Su apoyo, su inteligencia y su cercanía han sido fundamentales en mi vida personal y profesional. Y a mis hijos,* ***Marta****,* ***Cristina*** *y* ***Alejandro****, de quienes siento un inmenso orgullo y que representan, sin duda, una de las mayores alegrías y motivos de gratitud de mi vida. Estoy plenamente convencido de que llegarán tan lejos como se propongan, porque poseen el talento, la determinación y los valores necesarios para alcanzar cualquier meta que decidan perseguir.*

A todos ellos, y a cuantos me han acompañado en este camino, mi más profundo agradecimiento.

1. Introducción: la humanización como desafío y oportunidad

1.1. Del "progreso tecnológico" al "progreso humano"

Tradicionalmente, el discurso dominante sobre la introducción de nuevas tecnologías y la automatización de servicios ha girado en torno a una promesa sencilla: más eficiencia, más rapidez, más consistencia y menos costes. Ese marco ayuda a explicar por qué las organizaciones invierten en inteligencia artificial (IA), robots, chatbots o tecnologías inmersivas como la realidad aumentada (RA) o la realidad virtual (RV), pero deja abierta una cuestión más incómoda: ¿qué sucede cuando esas tecnologías dejan de ser meras herramientas y pasan a ocupar una posición central en la prestación del servicio, influyendo en cómo las personas atribuyen intención, responsabilidad y cuidado? El giro contemporáneo no reside solo en que la tecnología desarrolle tareas, sino en que se sitúe en el centro del encuentro de servicio como una presencia capaz de generar nuevas expectativas y orientar la experiencia (van Doorn et al., 2017).

En el contexto de las tecnologías inmersivas, este giro se intensifica porque la interacción se vive como una experiencia inmersiva: el usuario entra en un entorno mediado (a menudo híbrido), percibe presencia y construye significado a partir de la sensación de estar en otro entorno (Flavián et al., 2019b; Orús et al., 2021). Eso reconfigura el *customer journey* (viaje del consumidor): la tecnología actúa como escenario que organiza la atención, la memoria y la emoción, y por ello, humanizar en el contexto de las tecnologías inmersivas exige mejorar el control percibido, la comprensión y la seguridad subjetiva, no solo "maquillar" la interfaz (Femenia-Serra et al., 2024; Lombard & Ditton, 1997).

Este cambio obliga a reformular lo que se entiende por humanización. Si "humanizar" se interpreta como añadir rasgos antropomórficos superficiales (un nombre propio, un rostro, una voz cálida), el término se vuelve estrecho y, en ocasiones, engañoso. La teoría psicológica del antropomorfismo indica que atribuimos cualidades humanas a entidades no humanas mediante procesos automáticos en los que intervienen nuestras percepciones, motivaciones y las señales contextuales (Epley et al., 2007). En otras palabras, una apariencia "humana" puede activar inferencias sociales, pero esa activación no garantiza una experiencia más respetuosa, más segura o más justa (Nass & Moon, 2000). En entornos inmersivos, por ejemplo, la "presencia" (sentir que estoy ahí) puede incrementarse notablemente según el tipo de contenido y el dispositivo, sin que ello implique necesariamente comprensión, control o bienestar (Flavián et al., 2019b; Orús et al., 2021).

Por eso, resulta más razonable concebir la humanización como un criterio tanto normativo como experiencial. Normativo, porque permite ordenar decisiones de diseño y de gestión (qué se automatiza, con qué límites, con qué finalidad); experiencial, porque se evalúa en la interacción concreta: cómo la persona vive el control percibido, la comprensión de lo que ocurre, la dignidad, la seguridad psicológica y la posibilidad de reparación. Desde esta perspectiva, humanizar no equivale a simular humanidad, sino a preservar condiciones humanas básicas en un entorno donde la autonomía del usuario podría erosionarse notablemente (Floridi & Sanders, 2004). En tecnologías inmersivas, además, la propia integración cuerpo-dispositivo ("embodiment") convierte la experiencia en algo más íntimo y susceptible: cuando el dispositivo se integra con los sentidos, aumentan la estimulación sensorial, la implicación y, por tanto, la importancia y las implicaciones del diseño (Flavián et al., 2019a; 2019b).

Esta distinción es clave porque el servicio automatizado incrementa el peso de lo que resulta "invisible", pero que percibimos de forma inconsciente. La inteligencia artificial opera ofreciendo recomendaciones, personalizando opciones, filtrando información, calculando riesgos y priorizando opciones, con efectos profundos incluso cuando el usuario no percibe la intervención algorítmica. En ese terreno, la experiencia humana depende menos del "aspecto externo" del sistema y más de su "funcionamiento interno": quién responde cuando falla, qué límites existen, qué prácticas de transparencia operativa se aplican y qué rutas de derivación hacia la intervención humana están disponibles (Belanche et al., 2024). En paralelo, las evidencias en el contexto de la realidad aumentada sugieren que buena parte del valor para el consumidor se genera a través de mecanismos psicológicos intermedios que mejoran y simplifican el proceso de decisión: reducción del riesgo percibido, aumento

del confort y de la confianza en la decisión tomada, y mayor satisfacción, lo que termina convirtiéndose en un mayor *engagement* con la plataforma o el servicio (Barta et al., 2023a).
Las tecnologías inmersivas añaden un matiz adicional. Si la inteligencia artificial transforma la experiencia desde lo invisible, la inmersión transforma la experiencia desde lo sensible. La promesa no es solo reducir la fricción o personalizar, sino intensificar la presencia, la implicación y la memoria a lo largo del *customer journey* (Femenia-Serra et al., 2024; Flavián et al., 2019b). Sin embargo, la intensificación no es neutral: puede enriquecer el significado, pero también aumentar la vulnerabilidad psicológica, la sobrecarga cognitiva o la captura atencional cuando el diseño prioriza el *engagement* por encima del control percibido (Lombard y Ditton, 1997). En experiencias inmersivas vinculadas a eventos, por ejemplo, se observa la dificultad de los usuarios para focalizar la atención en los elementos principales del evento y percibir la autenticidad; además, la gamificación puede operar como un mecanismo que mitiga parte del problema, pero no elimina el reto de fondo (Flavián et al., 2023). Por consiguiente, la humanización de la tecnología debe abarcar simultáneamente dos frentes: la arquitectura algorítmica (que reorganiza las decisiones tomadas de forma inconsciente) y la arquitectura inmersiva (que reorganiza la atención, la emoción y la experiencia del usuario).

1.2. La Industria 4.0 en servicios y la economía de la experiencia

La llamada Industria 4.0, entendida como la convergencia de tecnologías inteligentes y conectadas, altera la lógica del servicio a través de tres vías principales. En primer lugar, estas tecnologías superan las capacidades humanas específicas en tareas mecánicas y analíticas, lo que desplaza el control de la ejecución al diseño de sistemas y a la supervisión de los resultados (Belk et al., 2023). En segundo lugar, los clientes no solo "usan" tecnología; habitan entornos tecnológicamente mediados, donde la frontera entre canal y experiencia se difumina. Y, en tercer lugar, el límite entre lo humano y lo tecnológico se vuelve más borroso, no como metáfora futurista, sino como hechos cotidianos a través de las interacciones con asistentes conversacionales (Belk et al., 2023; van Doorn et al., 2017).
Desde la economía de la experiencia, este contexto implica que la competitividad no se decide únicamente por la eficiencia operativa, sino por la calidad integral de la interacción con el cliente. El *customer journey* deja de ser una secuencia lineal y se convierte en un tejido de puntos de contacto que combinan espacios físicos, interfaces digitales y, cada vez más, escenarios inmersivos. La evidencia reciente en turismo subraya que las tecnologías inmersivas pueden impactar las respuestas cognitivas, emocionales, sensoriales, conductuales y sociales a lo largo del journey, creando experiencias híbridas y nuevas formas de interacción (Lemon & Verhoef, 2016; Flavián et al., 2019b; Femenia-Serra et al., 2024).
En este contexto, conviene destacar que lo inmersivo no puede considerarse un bloque homogéneo. En el Cuadro 1 (continuo entre realidad–virtualidad), se ordenan las principales tecnologías inmersivas (de la realidad física a la realidad virtual) y se identifican, para cada una, las oportunidades y los riesgos de deshumanización (pérdida de control percibido, saturación, límites difusos o persuasión opaca). Este mapa permite conectar una decisión de diseño (qué grado de inmersión se utiliza y en qué momento del journey) con un criterio de humanización (autonomía, comprensión y seguridad subjetiva).
Además del continuo realidad–virtualidad, conviene disponer de un "mapa" equivalente para la inteligencia artificial en la Industria 4.0. La inteligencia artificial no es un bloque homogéneo: según el nivel (mecánico, analítico o emocional), cambia el tipo de valor que aporta al *customer journey* (eficiencia vs relación) y también el tipo de riesgo de deshumanización (pérdida de control, opacidad, intrusividad o expectativas infladas). El Cuadro 2 resume estas diferencias para conectar la decisión tecnológica con el criterio de humanización (autonomía, comprensión, seguridad y justicia) en cada etapa del journey**.**

Cuadro 1. Continuo realidad–virtualidad y condiciones de humanización de las tecnologías inmersivas (XR)

Tecnología (ejemplo)	Definición breve	Humanización vs. Deshumanización (+) Oportunidades y (-) riesgos
Realidad física *(ej.: compra en tienda física/atención sin capas XR)*	Experiencia sin mediación inmersiva, basada en percepción directa del entorno.	**+** Máxima legibilidad del contexto y del vínculo interpersonal. **–** Menos apoyo para anticipar resultados o comparar opciones.
RA – Realidad aumentada *(ej.: virtual try-on en móvil. Pokémon Go Hola)*	Capa digital superpuesta al mundo real en tiempo real, manteniendo anclaje físico.	**+** Hace visible lo relevante "en contexto", reduce incertidumbre y facilita decisión informada. **–** Puede aumentar carga cognitiva y persuasión opaca.
Realidad pura mixta *(ej.: espacio físico con integración perfecta real–virtual)*	Integración fluida y coherente entre mundo real y virtual, sin costuras perceptibles para el usuario.	**+** Continuidad experiencial: comprensión y acción sin ruptura de atención. **–** Límites difusos real/virtual; exige señalización.
Virtualidad aumentada *(ej.: entorno virtual principal con "ventanas" o elementos del mundo real integrados. Clase en Microsoft Teams en un aula virtual con caras reales)*	Entorno mayoritariamente virtual que incorpora componentes reales (capturas, vídeo, objetos físicos) como parte de la escena.	**+** Aumenta telepresencia y participación; mantiene conexión con lo real. **–** Riesgo de confusión de señales y expectativas (autenticidad); fatiga si la mezcla no es coherente.
RM/MR – Realidad mixta *(ej.: objeto 3D anclado e interactivo en la sala)*	Elementos virtuales anclados al entorno físico con interacción consistente (gestos, proximidad, manipulación).	**+** Interacción más natural (*embodiment*) y aprendizaje experiencial; aumenta recuerdo. **–** Expectativas infladas o confusión si "promete" más de lo que entrega.
RV – Realidad virtual *(ej.: visita 360º con visor HMD)*	Sustituye el entorno físico por uno virtual inmersivo. Se requiere un visor.	**+** Alta presencia para explorar, ensayar y comprender escenarios; útil para pre-experiencias y aprendizaje. **–** Fatiga o malestar (mareo), saturación emocional y pérdida de control si no se regula intensidad, tiempo y salidas.

Fuente: Milgram y Kishino (1994); Flavián et al., (2019b).

Cuadro 2. Tipos de inteligencia artificial en la Industria 4.0 de servicios: contribución al customer journey, oportunidades y riesgos de humanización

Tipo de IA en servicios ***Papel desempeñado***	**Qué aporta al *customer journey* (economía de la experiencia)**	**Oportunidades (+) y riesgos (-) en la humanización (si se diseña bien vs mal)**
IA mecánica (automatización) *Ejecuta tareas repetitivas y estandarizadas*	Reduce fricción en *touchpoints* operativos (confirmaciones, check-in/out, tareas administrativas), mejora consistencia	**+** Libera a empleados para momentos más críticos. - Trato frío, rigidez y escalado humano tardío en incidencias.
IA analítica (predicción, personalización) *Analiza datos para recomendar, priorizar, decidir o anticipar*	Incrementa relevancia y conveniencia (menos esfuerzo de búsqueda, mejores decisiones) y optimiza el recorrido	**+** Competencia percibida y "servicio que entiende" cuando explica y deja control. **-** Opacidad ("caja negra"), sensación de vigilancia y clasificación.
IA emocional (señales sociales/afectivas) *Detecta estados y adapta la interacción social*	Reduce ansiedad/fricción en momentos sensibles; sostiene una interacción más relacional (acompañamiento, contención)	**+** Calidez y apoyo percibidos cuando es útil, proporcional y con límites. - Falsa intimidad, dependencia, expectativas infladas.
IA generativa conversacional (*GenAI*) *Genera contenido dinámico y diálogo natural*	Aumenta fluidez y continuidad narrativa del *journey* (explica, guía, co-crea, resuelve dudas)	**+** Cercanía **si** hay límites, verificación y derivación a humano. **-** Alucinaciones/falsa autoridad, confusión de autoría y daño reputacional.

Fuente: Belanche et al. (2021a; 2021b); Casaló et al. (2025); Huang y Rust (2018) y Huang et al. (2019).

Estudios experimentales indican que el tipo de contenido (p. ej., realismo factual en vídeos 360º) y el dispositivo empleado (p. ej., cascos de RV frente a smartphones) modulan la presencia, la facilidad de imaginación y el atractivo visual; ello, a su vez, condiciona intenciones como la de reservar un hotel. La "calidad humana" de la experiencia depende más de la integración entre el contenido y el dispositivo que del mero uso de tecnologías inmersivas (Orús et al., 2021).

A partir de este punto, la humanización trasciende la mera retórica y se erige como un principio central en el diseño tecnológico: ¿qué momentos del *customer journey* resultan críticos desde una perspectiva humana y qué modalidades de mediación tecnológica son tolerables y legítimas en cada caso? La automatización puede ser altamente eficaz en tareas repetitivas, pero en situaciones de vulnerabilidad, conflicto o reparación del servicio, la ausencia de intervención humana o la opacidad del sistema deteriora la confianza con rapidez, especialmente cuando el usuario ya percibe que "la organización" es responsable del comportamiento del agente tecnológico (van Doorn et al., 2017). En experiencias inmersivas, además, la intensificación emocional puede amplificar tanto los aciertos como los fallos. Por ejemplo, un mayor *embodiment* en una experiencia de realidad virtual puede aumentar las emociones positivas y el *engagement*, pero también eleva la exigencia de coherencia y cuidado en el diseño (Flavián et al., 2019b).

1.3. Enfoque desde marketing y comportamiento del consumidor: percepciones, intención, adopción y conducta real

En el proceso de humanización, el punto de vista del marketing y del comportamiento del consumidor resulta especialmente útil, porque desplaza el foco de lo que la tecnología "es" a lo que la tecnología

“hace” en la mente del usuario. En servicios, la adopción no depende solo de la utilidad percibida ni de la facilidad de uso. Depende también de las percepciones, las normas y los juicios sociales (p. ej., la calidez y la competencia), de las experiencias emocionales, de las atribuciones morales y de la capacidad de la persona para mantener el control percibido en contextos cada vez más automatizados. El servicio, en última instancia, es una relación mediada por percepciones y expectativas, y cuando el mediador es un agente tecnológico con presencia social automatizada, la relación se reescribe por completo.

Además, la literatura reciente permite constatar que el salto entre la intención y la conducta no puede darse por supuesto. La conducta real se consolida cuando el usuario siente que comprende el sistema y puede modular su exposición, sus datos y el grado de automatización que desea experimentar (Flavián et al., 2022). Esto es especialmente relevante en tecnologías inmersivas, donde el “impacto” inicial puede generar intención, pero la continuidad depende de si el diseño evita la saturación, la fatiga y las tensiones psicológicas asociadas a la intensificación sensorial (Lombard & Ditton, 1997).

En el contexto del turismo, por ejemplo, el aumento de la presencia puede impulsar intenciones, pero su efecto suele estar mediado por procesos como la facilidad de imaginación y el atractivo visual, y dichos mediadores son sensibles a la combinación entre el contenido y el dispositivo (Orús et al., 2021). A su vez, enriquecer la experiencia con estímulos multisensoriales (p. ej., aromas agradables y congruentes) puede incrementar la estimulación sensorial y facilitar la imaginación, reforzando reacciones afectivas y conductuales; pero esa misma potencia sensorial hace que la congruencia y el “cuidado” del diseño se conviertan en aspectos críticos (Flavián et al., 2021).

En paralelo, en el ámbito de la inteligencia artificial y la realidad aumentada, emerge otro factor: la ambivalencia. El usuario puede valorar positivamente la personalización y la rapidez, pero también puede sentir intrusión o pérdida de autonomía cuando la personalización se basa en la captura de datos y en una clasificación opaca. El valor, por tanto, no se explica adecuadamente como suma de beneficios; se parece más a un equilibrio dinámico entre beneficios y tensiones. La investigación en realidad aumentada (p. ej., filtros de marca) ilustra bien esa doble cara: la creatividad, la originalidad y el disfrute pueden mejorar la imagen de marca, pero la intrusividad percibida y el reconocimiento publicitario pueden modular (debilitar o reforzar) esos efectos (Orús et al., 2026).

En definitiva, la humanización debe desarrollarse tomando como referencia el comportamiento real del usuario y siendo conscientes de que lo que hace y lo que mantiene en el tiempo están profundamente condicionados por sus percepciones, por las intenciones que se forman durante la experiencia y por cómo interpreta la ambivalencia entre beneficios (p. ej., personalización, conveniencia) y costes percibidos (p. ej., intrusividad, pérdida de autonomía). Por ello, la humanización debe plasmarse a través de variables observables que sustenten esa conducta: control percibido, comprensión, confianza, confort con la decisión tomada, menor carga cognitiva e intrusividad, autenticidad, atención y bienestar (Puntoni et al., 2021). Es en ese conjunto (no en los cambios cosméticos antropomórficos) donde se decide si la tecnología amplifica el valor humano o, por el contrario, lo erosiona.

2. Marco conceptual para entender la tecnología “centrada en personas”

2.1. Tecnología como actor social: presencia social, control percibido y “calidez” relacional

¿Por qué la tecnología activa inferencias sociales incluso cuando sabemos que “no hay nadie” detrás? Parte de la respuesta radica en que ciertos diseños producen señales suficientes para activar expectativas de reciprocidad, intencionalidad y orientación hacia la relación con el interlocutor. La noción de presencia social automatizada captura este fenómeno al describir el grado en que un agente tecnológico hace sentir al cliente que está en compañía de otra entidad social durante la experiencia (van Doorn et al., 2017).

En las interacciones con robots o chatbots, esa presencia emerge mediante señales conversacionales, corporales y de “comprensión”, que funcionan como pistas sociales que el usuario interpreta automáticamente. Para clarificar estas señales de humanización (especialmente las tres más decisivas en los servicios: calidez, competencia y presencia social) y, a la vez, explicitar sus límites éticos (evitar

el engaño, la dependencia o la falsa empatía), el Cuadro 3 sintetiza lo qué significa cada señal, cómo se implementa en la práctica, los efectos produce y las “líneas rojas” que conviene respetar.

Entre las señales conversacionales, no solo importa el ritmo, sino también su cualidad: la rapidez con la que el sistema responde (la latencia mínima sugiere interacción y atención), la forma en que gestiona turnos (no interrumpe, no se adelanta al usuario) y los indicativos de escucha (“entiendo”, “de acuerdo”, “un momento”). También son relevantes la adecuación al contexto o la capacidad de “seguir el hilo” y corregir malentendidos (Nass & Moon, 2000). Con esta finalidad, el sistema parafrasea para confirmar (“si te he entendido bien, quieres cambiar la fecha de entrada, no cancelar”), cuando verifica datos clave antes de actuar (“confirmo: ¿la reserva es para el día 12?”), al reconocer un error explícitamente (“tienes razón; corrijo”). La personalización lingüística puede reforzar la sensación de interacción social si se usa con prudencia, por ejemplo, recordar preferencias explícitas (“prefieres atención por chat”) o ajustar el registro (“si lo prefieres, te lo explico paso a paso”), mientras que el exceso de familiaridad también podría resultar invasivo.

Las señales corporales cumplen un papel análogo cuando el agente tiene una forma física o una representación visual. La sincronía de microgestos (asentir al confirmar el usuario) y la regulación de la proximidad (acercarse para ayudar y alejarse para no invadir) pueden producir una sensación de atención y respeto. En interfaces no físicas, estas señales se traducen en equivalentes visuales, como un avatar que orienta el cuerpo hacia el usuario o expresiones faciales moderadas (Belanche et al., 2021a).

Por último, las señales de “comprensión” (referencias al contexto, memoria de preferencias, consistencia en el seguimiento del caso) suelen ser las más determinantes para que se perciba la presencia social. Recordar el motivo de contacto o evitar que el usuario repita información refuerza el valor de la interacción. Sin embargo, la presencia por sí sola no garantiza la humanización. Puede suavizar la interacción, pero también generar expectativas de apoyo y responsabilidad que el sistema no está en condiciones de cumplir cuando surge un problema relevante. En este sentido, conviene destacar que humanizar no es maximizar la cercanía, sino diseñarla sin erosionar el control percibido.

En tecnologías inmersivas, esta dinámica se intensifica porque la presencia no es solo social, sino también espacial y corporal. La literatura sobre presencia definió tempranamente el concepto como la sensación de “estar allí”, un estado en el que la mediación tecnológica se vuelve psicológicamente "transparente" (Lombard & Ditton, 1997). En ese estado, el juicio del usuario deja de basarse únicamente en la información y pasa a depender de la coherencia sensorial, de la facilidad de imaginación y de la congruencia con las expectativas. Para aclarar estos mecanismos “experienciales” que median la humanización en tecnologías inmersivas (y explican por qué una experiencia puede resultar más comprensible y valiosa o, por el contrario, más intrusiva y exigente), el Cuadro 4 sintetiza los principales procesos psicológicos activados (presencia, *embodiment*, telepresencia, facilidad de imaginación e inmersión multisensorial), indicando cómo se incrementan y qué riesgos conviene prevenir. El resultado es que las experiencias inmersivas facilitan el aprendizaje, la empatía y la memoria; pero también elevan el coste del fallo o de la incoherencia, porque lo vivido “parece” más real.

Cuadro 3. Señales de humanización (calidez, competencia, presencia social) y límites éticos

Señal ***(Definición)***	**Ejemplos prácticos (diseño)**	**Efectos y límites éticos**
Calidez *(Intención benevolente: "me trata bien")*	Tono respetuoso; valida emociones ("entiendo…"); disculpa y repara; ofrece opciones	**Efecto:** sube confianza y valor emocional. **Límite:** Evitar falsa empatía.
Competencia *(Capacidad percibida: "resuelve bien")*	Respuestas correctas; pasos claros; confirma datos; resume; seguimiento del caso	**Efecto:** sube valor funcional y credibilidad. **Límite:** declarar incertidumbre; no inventar.
Presencia social (ASP) *(Sensación de "estar con alguien")*	Turnos naturales; señales de escucha ("déjame revisar"); memoria útil; continuidad del hilo	**Efecto:** reduce fricción/ansiedad y aumenta *engagement*. **Límite:** No simular conciencia ni relación personal.
Autonomía del agente *(Percepción de iniciativa/acción propia)*	Proactividad moderada; recordatorios; ejecutar solo con confirmación; opción "solo sugerencias"	**Efecto:** mejora fluidez del *journey*. **Límite:** consentimiento, control y reversibilidad.
Apariencia humana/ antropomorfismo *(Parecido a humanos: cara/voz/nombre)*	Voz natural; avatar; nombre propio; microexpresiones; "personalidad" consistente	**Efecto:** puede aumentar cercanía y expectativas. **Límite:** evitar "falsa humanidad".

Fuente: Belanche et al. (2020a; 2020b; 2021a) y Puntoni et al. (2021).

Cuadro 4. Mecanismos psicológicos en tecnologías inmersivas que activan o desactivan la humanización

Mecanismo *(definición)*	**Cómo se incrementa en XR *(beneficios)***	**Riesgos asociados**
Presencia *(sensación de "estar ahí" en el entorno mediado)*	Inmersión alta, coherencia audiovisual, baja latencia, interacción fluida *(más comprensión, vividez, aprendizaje experiencial)*	Manipulación emocional si no hay control; malestar/mareo; expectativas irreales
Embodiment / Encarnación *(sentir el cuerpo/acción "dentro" del entorno o avatar)*	Seguimiento de manos/cuerpo, perspectiva 1ª persona *(más control, engagement, confianza en la interacción)*	Fatiga física/mental; presión social a exponerse; efectos de identidad/autoimagen
Telepresencia *(sensación de "estar con otros" o "allí" a distancia)*	Avatares expresivos, audio espacial, interacción en tiempo real, espacios compartidos *(más conexión social)*	Engaño/impersonación; aislamiento si sustituye vínculos; frustración por señales sociales pobres
Facilidad de imaginación *(como de fácil es imaginarse usando/visitando algo)*	Visualización realista (3D/360º), pruebas virtuales *(menos incertidumbre, mejores decisiones, menos arrepentimiento)*	"Efecto demo": decepción si la realidad no coincide; compras impulsivas por exceso de confianza
Inmersión multisensorial *(activación de varios sentidos: sonido/olor/háptica)*	Audio espacial, háptica, estímulos congruentes con el contenido *(más memoria, emoción, sensación de realismo)*	Sobrecarga sensorial; intrusividad; rechazo si hay incongruencia o falta de consentimiento

Fuente: Barta et al. (2023a; 2024b); Orús et al. (2017; 2021); Ruusunen et al., (2023) y Witmer et al (1998).

2.2. Humanización tecnológica: dignidad, autonomía, control, seguridad y justicia percibida

Si se pretende situar a las personas en el centro, no basta con registrar si utilizan la tecnología; es necesario un marco que valore cómo esta afecta a la dignidad, al control y al bienestar de quienes la usan. La satisfacción o la intención de uso aportan información valiosa, pero no basta para juzgar si la experiencia protege condiciones humanas básicas, como la dignidad, la autonomía y la seguridad. La dignidad se aprecia en detalles concretos: si el usuario se siente tratado como un fin y no como un medio; si percibe que puede comprender lo que ocurre y cuestionarlo; y si el sistema reconoce su vulnerabilidad en lugar de explotarla. La autonomía depende del margen real de elección, de la posibilidad de desactivar la personalización y de la existencia de una salida clara cuando el sistema se equivoca (Shneiderman, 2020).

El control percibido es más que una preferencia: es un mecanismo de protección. Cuando el usuario siente control, tolera mejor la incertidumbre y disminuye la probabilidad de atribuir intención manipuladora a la organización. Cuando el control se pierde, la experiencia puede deteriorarse incluso si el funcionamiento "objetivo" del sistema es adecuado. Esa es una de las paradojas del servicio automatizado: el rendimiento no basta si no es comprensible para el usuario y puede ajustarse a sus necesidades (Shneiderman, 2020).

La justicia percibida resulta crítica si la tecnología decide o recomienda. En sistemas de inteligencia artificial, el usuario puede experimentar injusticia no solo en el resultado, sino también por la opacidad del proceso. Y la opacidad no es un problema menor en estos casos. Por ello, la explicabilidad práctica (no necesariamente técnica) y la transparencia operativa se convierten en condiciones fundamentales

para que el usuario mantenga una relación sostenible con el sistema (Guidotti et al., 2019). La clave está en lo práctico: explicar lo suficiente para generar confianza sin generar complejidad.
En las tecnologías inmersivas, la dimensión de seguridad también adquiere un componente corporal y atencional. Un entorno que intensifica los estímulos sin permitir modular su intensidad puede provocar malestar, fatiga o saturación. Si la experiencia está diseñada para captar la atención, la autonomía se erosiona, aunque el usuario "elija" entrar. En este punto, la humanización también requiere establecer límites, pausas, ritmos y opciones de salida comprensibles (Guy, Normand, Jeunet-Kelway, & Moreau, 2023).

2.3. Confianza y riesgo: cómo se forman (y rompen) en interacciones con IA, robots y tecnologías inmersivas

La confianza en servicios tecnológicamente mediados no se genera solo con promesas, sino también con consistencia, transparencia útil y responsabilidad visible. En sistemas automatizados, el usuario tiende a concentrar la responsabilidad en la empresa, incluso cuando el agente es un robot o un chatbot, lo que implica que el fallo se interpreta como un fallo institucional, no meramente técnico (van Doorn et al., 2017). Esta concentración de responsabilidad hace que la reparación sea determinante. En estos casos, no basta con compensar; hay que demostrar aprendizaje.
El riesgo, en consecuencia, se vuelve multidimensional. Incluye riesgo funcional (que el sistema falle), pero también riesgo moral (que la empresa use datos de manera intrusiva), riesgo social (ser observado o evaluado) y riesgo psicológico (sentirse reemplazado, alienado o instrumentalizado). La literatura sobre tensiones psicológicas en experiencias con inteligencia artificial propone que parte del valor radica en aliviar las contradicciones vividas por el consumidor, como sentirse comprendido en lugar de malinterpretado, empoderado en lugar de reemplazado, conectado en lugar de alienado (Belanche et al., 2024; Flavián et al., 2024a). Para operacionalizar estas tensiones, y conectarlas con decisiones concretas de diseño y gestión del servicio, el Cuadro 5 resume cómo las percibe el usuario, qué mejoras suele prometer la inteligencia artificial y cómo podrían humanizarse.
Esta formulación es útil porque no interpreta el rechazo o la incomodidad como una simple fobia a la tecnología. En lugar de atribuirlo a una actitud irracional del usuario, sitúa el foco en cómo se diseña la relación: qué señales de respeto y cuidado transmite el sistema, cuánto control percibido deja a la persona y en qué medida la interacción reduce o intensifica tensiones como sentirse sustituido, vigilado o malinterpretado.
En este contexto, conviene destacar la importancia de un fenómeno adicional: la aversión algorítmica. Tras varios errores, las personas pueden rechazar el uso de algoritmos incluso si su rendimiento es superior al humano, ya que el error algorítmico se percibe como menos tolerable o menos perdonable (Dietvorst et al., 2015). En servicios, esta dinámica sugiere que la confianza es más frágil cuando la tecnología aparece como autoridad decisional sin posibilidad de apelación o sin una explicación suficiente.
En tecnologías inmersivas, el riesgo se intensifica cuando la presencia incrementa el realismo psicológico. Un fallo o una incoherencia puede generar una frustración más intensa, precisamente porque el usuario se sentía "dentro" de la experiencia. Y si el entorno incorpora estrategias persuasivas difíciles de reconocer, el usuario puede sentirse manipulado, lo que acentúa la pérdida de confianza (Madary & Metzinger, 2016).

Cuadro 5. Tensiones psicológicas en experiencias con IA: mejoras del servicio y claves para la humanización

Tensión psicológica *(percepción del usuario)*	**Mejora del servicio**	**Claves para la humanización**
Captura *("me empuja")*	Personalización proactiva; recomendaciones en tiempo real; optimización de tiempos y conversiones	Diseñar **autonomía**: opción "solo sugerencias", límites de frecuencia, pausa/salir visible.
Clasificación *("me están vigilando")*	Predicción de preferencias; segmentación; detección de necesidades; prevención de fraude o riesgos	Diseñar **transparencia y justicia**: "por qué esto", datos mínimos, controles de privacidad comprensibles.
Delegación *(decide por mí")*	Automatización de tareas y decisiones; resolución rápida; ejecución de acciones (cambios, reembolsos, reservas)	Diseñar **control y reversibilidad**: confirmación antes de ejecutar, "deshacer", historial, escalado humano claro.
Social *("no me entiende)*	Interfaces conversacionales; memoria de contexto; lenguaje natural; continuidad del caso	Diseñar **presencia sin engaño**: señales breves de escucha, continuidad útil; disclosure de IA; evitar falsa empatía.
Incertidumbre *("¿será correcto?")*	Respuestas rápidas; síntesis; asistencia 24/7; reducción del esfuerzo cognitivo	Diseñar **humildad epistemológica**: reconocer límites, pedir verificación en decisiones críticas, alertas ante baja confianza.

Fuente: Dietvorst et al. (2015); Flavián et al. (2024a) y Puntoni et al. (2021).

2.4. *Customer journey* híbrido: integración de canales físicos, digitales e inmersivos

Si el recorrido del consumidor es el espacio donde se materializa el valor, entonces su hibridación plantea un reto: cómo asegurar la continuidad sin convertir la transición entre canales en una pérdida de control. La lógica híbrida no consiste solo en sumar canales, sino en coordinar el recorrido con criterio: definir qué conviene automatizar, en qué puntos es preferible la interacción humana, cuándo la inmersión puede funcionar como preexperiencia y cuándo resulta mejor mantener una comunicación sobria y directa (Flavián et al., 2024a).

La literatura sobre la experiencia del cliente enfatiza que el valor emerge a lo largo del recorrido, no en un punto aislado (Lemon & Verhoef, 2016). En el contexto actual, esa afirmación implica que una organización puede diseñar un excelente *touchpoint* inmersivo y, aun así, fracasar si el resto del *journey* es opaco o la reparación de fallos está mal diseñada. Del mismo modo, puede tener un chatbot competente y perder confianza si el usuario no distingue cuándo interactúa con inteligencia artificial y cuándo con humanos, o si no existe un canal claro de reclamación (Tax, Brown, & Chandrashekaran, 1998).

2.5. Valor experiencial ampliado: utilidad, emoción, significado, pertenencia y bienestar

El valor, en servicios, se ha entendido tradicionalmente como equilibrio entre beneficios funcionales y sacrificios. Sin embargo, en experiencias mediadas por inteligencia artificial y tecnologías inmersivas, esa definición resulta insuficiente. El usuario no solo evalúa la utilidad; también evalúa la pertenencia, el significado, la autenticidad y el bienestar. En el metaverso y en la virtualización de eventos, se ha constatado que la gamificación y el diseño de la experiencia pueden generar valor, pero también abrir una "cara oscura" vinculada a expectativas infladas, frustración o efectos no deseados en el bienestar psicológico (Flavián et al., 2023; Puntoni et al., 2021).

Esa ampliación del valor obliga a reformular la pregunta central: ¿cuándo la tecnología mejora la vida del usuario y cuándo solo intensifica la interacción sin aportar significado? Si el objetivo es humanizar, la respuesta debería privilegiar experiencias que permitan a la persona mantener el control, comprender lo que sucede y extraer un valor que no se reduzca a minimizar la fricción.

3. Inteligencia artificial en servicios: de la eficiencia a la relación

3.1. La inteligencia artificial como infraestructura invisible del servicio: recomendación, personalización y toma de decisiones

La inteligencia artificial opera a menudo como infraestructura "invisible" del servicio. No se presenta como una "opción" ni como un "canal", sino como una capa que ordena opciones, prioriza información y sugiere decisiones. Esa invisibilidad tiene una ventaja evidente: reduce la fricción. Pero también introduce un dilema de humanización: cuanto menos visible es la mediación algorítmica, más difícil resulta para el usuario entender cómo se toman las decisiones y qué capacidad tiene para influir, corregir o rechazar los resultados.

En el contexto de los servicios, este dilema se ha intensificado debido a dos tendencias convergentes. Por un lado, el uso cada vez mayor de sistemas que personalizan y recomiendan en tiempo real, con una capacidad creciente para inferir preferencias. Por otro lado, la normalización de interfaces conversacionales que convierten la recomendación en diálogo aumenta la presencia social automatizada y, en consecuencia, la probabilidad de que el usuario atribuya intención al sistema. Este proceso puede enriquecer la experiencia, pero también elevar el riesgo de intrusividad cuando la personalización se percibe como una vigilancia o una clasificación opaca (Belanche et al., 2024).

Desde la perspectiva de la humanización, la cuestión no es solo si la inteligencia artificial mejora los indicadores de servicio (velocidad, consistencia, precisión), sino también si está diseñada e implementada de modo que preserve el control percibido y la capacidad de decisión del usuario. En esa línea, resulta relevante el marco de responsabilidad digital corporativa como conjunto de prácticas para gestionar de manera responsable las tecnologías digitales, integrando privacidad, equidad, seguridad y rendición de cuentas en las decisiones de diseño e implementación (Wirtz et al., 2023). La conclusión es clara: la confianza no se construye solo con mejores resultados; se construye cuando la organización demuestra, mediante mecanismos verificables, que utiliza la inteligencia artificial de forma responsable.

3.2. La inteligencia artificial generativa como "motor creativo" de experiencias: contenido, lenguaje y adaptación contextual

La inteligencia artificial generativa introduce un cambio cualitativo: no solo predice o clasifica, sino que también produce contenido nuevo. El lenguaje, el tono y la capacidad de sostener una conversación natural incrementan la sensación de interacción social y, con ello, el potencial de vínculo, de persuasión y de guía a lo largo del *journey*. Sin embargo, cuanto más "natural" es la generación, más se difuminan la autoría y la rendición de cuentas: crece el dilema de la autenticidad y la transparencia (Dwivedi et al., 2023). Cuando el sistema genera contenido, la atribución de la autoría se vuelve central (resulta clave saber si el mensaje proviene de una persona, de la organización o de un modelo generativo), porque esa diferencia afecta expectativas, responsabilidad y confianza. Y si el usuario no puede distinguir lo sintético de lo humano cuando esa distinción importa, la experiencia puede resultar persuasiva, pero sin ofrecer la transparencia necesaria.

En la lógica de la humanización, por tanto, el criterio no debería ser maximizar la fluidez, sino asegurar fluidez con límites mediante tres mecanismos prácticos: (1) etiquetado cuando corresponda (p. ej., indicar de forma visible que el mensaje o la voz son generados de forma sintética, y señalar cuándo se están usando datos para personalizar); (2) comunicación de capacidades y límites (p. ej., qué puede y qué no puede hacer el sistema, en qué condiciones puede equivocarse y qué fuentes utiliza o no utiliza); y (3) rutas claras de derivación a humano en momentos críticos (p. ej., botones o comandos para "hablar con un agente", escalado automático ante quejas, pagos, decisiones sensibles o señales de frustración). Estas exigencias se conectan con la idea de transparencia operativa como condición

para el control percibido y la confianza. También se alinean con estándares emergentes de gestión responsable. Por ejemplo, la norma ISO/IEC 42001:2023 propone un sistema de gestión de la inteligencia artificial orientado a la gestión responsable (International Organization for Standardization, 2023). El riesgo, por supuesto, es convertir la gestión responsable en burocracia. Pero la alternativa (innovación sin una gestión responsable) suele conllevar un coste reputacional y social creciente.

3.3. La paradoja de la personalización: precisión frente a intrusividad

La personalización suele presentarse como una mejora obvia del servicio: reduce el esfuerzo cognitivo, acorta el tiempo de búsqueda y convierte la abundancia de opciones en un itinerario manejable. Sin embargo, cuanto más precisa es la personalización, más probable es que active una segunda lectura, menos cómoda: "¿Cómo saben esto de mí?". En ese punto, el beneficio funcional puede coexistir con una tensión psicológica difícil de compensar con una mayor eficiencia. Esta tensión puede formularse como una paradoja: la personalización incrementa la relevancia y la satisfacción, pero también puede convertirse en una señal de vigilancia o manipulación, especialmente cuando el sistema no ofrece un marco claro respecto al tipo de datos utilizado, la finalidad buscada y los límites establecidos.

El problema se agrava cuando la personalización deja de limitarse a la oferta y alcanza la dimensión relacional y expresiva del encuentro (tono, lenguaje, gestos y señales de intención). La inteligencia artificial generativa puede adaptar el tono, la estructura del mensaje, el estilo persuasivo o incluso la "personalidad" del agente. Esa adaptación puede considerarse parte de una atención más personalizada, pero también podría interpretarse como una intrusión, porque parece ajustarse a la psicología del cliente en lugar de a sus preferencias declaradas. La teoría del antropomorfismo ayuda a explicar por qué esta frontera es tan delicada: en situaciones de incertidumbre o vulnerabilidad, la atribución de mente a un agente no humano se vuelve más probable, y la interacción parece "más humana" precisamente cuando el usuario necesita más protección, no más persuasión (Epley et al., 2007).

En el contexto de las tecnologías inmersivas, la paradoja adquiere una intensidad particularmente notable. No se trata solo de datos de navegación o de clics: pueden capturarse movimientos, orientación corporal, atención, patrones de interacción espacial y señales conductuales finas que, combinadas, generan una huella íntima incluso cuando el usuario no la percibe como tal (Dwivedi et al., 2022). Desde esta perspectiva, el dilema deja de ser abstracto. La personalización puede ser útil, pero también puede convertirse en una forma de sobreexposición. El ejemplo de los filtros de realidad aumentada es sugerente: incluso cuando la creatividad mejora los resultados de marca, la intrusividad y el reconocimiento publicitario pueden moderar los efectos y activar la resistencia (Orús et al., 2026). La lección trasladable es que el éxito experiencial no depende solo de "acertar más", sino de acertar sin invadir.

Desde una perspectiva de humanización del servicio, el objetivo no es eliminar la personalización, sino reformularla de manera que no sustituya la decisión del usuario, sino que la apoye: el usuario debería poder decidir el nivel de personalización, entender qué se está ajustando y con qué datos, y corregir o desactivar esas adaptaciones cuando lo considere oportuno. En algunos casos, ofrecer una opción explícita de menor personalización, o incluso una experiencia deliberadamente genérica, puede funcionar como señal de respeto. En otros casos, puede bastar con permitir que el usuario ajuste la dimensión personalizada (contenido, ritmo, canal, nivel de proactividad), manteniendo el control percibido y reduciendo la resistencia psicológica o la reactancia. En última instancia, el equilibrio "precisión–intrusividad" es menos un problema técnico que un problema relacional: se resuelve cuando el usuario comprende, elige y puede corregir.

3.4. Transparencia operativa: explicabilidad práctica, comunicación y expectativas realistas

Con frecuencia, la transparencia en inteligencia artificial se aborda como un problema de interpretabilidad técnica, es decir, como el esfuerzo por abrir la "caja negra" del modelo y explicar, al menos de forma aproximada, cómo llega a una recomendación o a una decisión. Sin embargo, en el

ámbito de los servicios, la dimensión más relevante de la transparencia es la que permite al usuario ajustar su nivel de confianza y mantener el control percibido. El usuario no necesita saber qué hace el sistema, qué no hace, qué datos utiliza, qué consecuencias tiene su interacción y qué vías de corrección y reclamación existen si el resultado es inadecuado. Dicho de otro modo, la explicabilidad práctica se parece más a una conversación honesta que a un informe de ingeniería (Parasuraman et al., 2000).

La literatura ha propuesto avanzar desde "cajas negras" hacia "cajas de cristal", es decir, sistemas que ofrezcan una visibilidad razonable sobre cómo recomiendan y deciden, con implicaciones directas para la confianza y la adopción (Rai, 2020). Asimismo, revisiones de orientación más computacional subrayan que la falta de explicaciones no es solo una limitación funcional; también plantea un problema ético, pues dificulta que el usuario evalúe, cuestione y discuta las decisiones automatizadas (Guidotti et al., 2019). En un marco de humanización, ambos argumentos convergen en una misma idea: la transparencia no es un "extra" informativo, sino el requisito que permite al usuario entender lo que hace el sistema, decidir hasta qué punto aceptarlo y, si es necesario, corregir o descartar sus resultados.

En tecnologías inmersivas, la transparencia debe diseñarse con especial cuidado porque la experiencia puede ser tan envolvente que el usuario pierda conciencia de qué elementos son personalizados, qué partes son sintéticas y qué señales están orientadas a influir. Si la transparencia se reduce a avisos legales iniciales, la relación queda desprotegida. De ahí la importancia de una transparencia verdaderamente "operativa", integrada en el flujo de la experiencia, con un lenguaje comprensible y mecanismos sencillos para ajustar preferencias, rechazar la personalización o salir del entorno sin penalización.

Un riesgo adicional es la inflación de expectativas, es decir, que el usuario atribuya al sistema más comprensión, competencia y "criterio" de los que realmente tiene. En inteligencia artificial generativa, la fluidez conversacional (respuestas rápidas, coherentes y con tono humano) puede hacer que el sistema "parezca" entender el contexto de forma profunda o razonar con solidez, cuando en realidad puede estar generando formulaciones plausibles con lagunas, ambigüedades o errores. Esta sobreestimación aumenta el impacto de los fallos por dos vías: (1) el usuario tiende a confiar más y a verificar menos (p. ej., acepta recomendaciones, condiciones o explicaciones sin contrastarlas); y (2) cuando el error aparece, suele hacerlo en momentos de mayor dependencia (decisiones sensibles, pagos, quejas, información crítica), lo que incrementa el coste percibido, la frustración y el daño reputacional. Por consiguiente, humanizar implica también gestionar expectativas: comunicar límites, reconocer incertidumbres y evitar la ficción de conocimiento absoluto. La transparencia, en este sentido, refuerza el enfoque ético de la interacción: no garantiza aciertos constantes, pero sí ofrece información veraz y mecanismos de corrección.

3.5. Impacto en la reputación y la confianza institucional: la "responsabilidad digital corporativa" como ventaja competitiva

Cuando la tecnología actúa como intermediaria del servicio, el usuario rara vez diferencia "el sistema" de "la organización": tiende a interpretar la interacción tecnológica como una extensión de la organización y, por tanto, a evaluar el resultado, la cortesía o el fallo de la interfaz como parte del comportamiento de la organización. En términos de atribución, el sistema no se percibe como un tercero neutral, sino como un componente del *service delivery system;* por ello, las respuestas emocionales (satisfacción, frustración), los juicios de justicia y la intención de recompra se dirigen principalmente al proveedor, incluso cuando la ejecución inmediata proviene de un agente automatizado. Tiende a atribuir la responsabilidad al proveedor, incluso cuando la interacción se realiza a través de un agente automatizado (van Doorn et al., 2017). Por eso, la reputación ya no depende solo de los resultados del servicio, sino también de la asignación de responsabilidades que se percibe detrás: quién responde cuando falla, cómo se repara, qué aprende la organización y qué señales ofrece sobre gestión responsable.

En este escenario, la responsabilidad digital corporativa constituye un marco adecuado para construir confianza institucional, no como un simple gesto reputacional (una acción simbólica o comunicativa orientada a mejorar la imagen), sino como un conjunto de compromisos y prácticas que se reflejan en las interacciones de servicio, los procedimientos y la rendición de cuentas. La lógica es sencilla y clara: cuanto mayores sean la opacidad técnica y la sensibilidad de los datos, más dependerá la confianza en las señales organizativas verificables (políticas concretas y auditables que permiten a terceros verificar que la organización tiene controles, rutinas y resultados medibles detrás de sus promesas). Esta idea se refuerza con el desarrollo de marcos de gestión responsable. Por un lado, aparecen guías de gestión de riesgos como el AI RMF 1.0 de NIST (National Institute of Standards and Technology, 2023). Por otro lado, estándares de sistemas de gestión como ISO/IEC 23894 e ISO/IEC 42001, orientados a integrar la gestión del riesgo y la mejora continua en organizaciones que utilizan la inteligencia artificial de forma intensiva (International Organization for Standardization, 2023).
En la Unión Europea, el Reglamento General de Protección de Datos establece obligaciones sobre el tratamiento de datos personales, especialmente relevantes en el ámbito de la personalización, la biometría y las trazas comportamentales (Reglamento (UE) 2016/679, 2016). Más recientemente, la Ley de Inteligencia Artificial adopta un enfoque basado en el riesgo y fija requisitos para determinados usos, incluyendo obligaciones de transparencia y de gestión responsable (Regulation (EU) 2024/1689, 2024). En última instancia, la competitividad deja de medirse solo por la velocidad con que se lanzan las innovaciones y se ve condicionada por la capacidad de innovar preservando la confianza. Es decir, por su capacidad de introducir nuevas soluciones respaldadas por controles, transparencia y mecanismos de responsabilidad.

4. Robots y agentes antropomórficos: diseñar presencia social sin deshumanizar

4.1. Antropomorfismo como lenguaje de interacción: por qué atribuimos mente e intención a las máquinas

Las interacciones con robots y agentes conversacionales a menudo se interpretan como relaciones sociales porque el antropomorfismo (atribución de propiedades humanas a lo no humano) es una heurística cognitiva habitual. Nos facilita la comprensión e interacción con agentes ambiguos y nos permite inferir intenciones y estados cuando, en realidad, solo hay procesos algorítmicos. Según la teoría de los tres factores (Epley, Waytz y Cacioppo, 2007), la propensión a antropomorfizar se incrementa cuando concurren tres condiciones interrelacionadas: (a) una motivación social para establecer afiliación; (b) la disponibilidad o accesibilidad de un marco humano plausible como categoría interpretativa; y (c) una necesidad instrumental de eficacia o de control sobre el entorno que hace funcional atribuir control humano. En términos prácticos, la teoría postula que, ante agentes ambiguos o poco familiares, las personas recurren con mayor frecuencia a explicaciones antropomórficas (p. ej., atribuyendo intenciones, emociones o creencias) cuando buscan conexión social, cuando los rasgos del agente evocan lo humano o cuando dicha atribución facilita la predicción y el control de su comportamiento. Esto explica por qué la atribución de mente puede intensificarse en situaciones de soledad, incertidumbre o vulnerabilidad, precisamente las condiciones en las que un servicio debería actuar con mayor prudencia.
En el contexto de los servicios, el antropomorfismo funciona como un atajo interpretativo. Si un robot “mira”, “responde”, “recuerda”, “se disculpa” o “se anticipa”, el usuario puede inferir intención, cuidado o incluso responsabilidad moral. Esa inferencia puede aumentar la comodidad del usuario y reducir la fricción afectiva (p. ej., disminuyendo la tensión emocional o facilitando la aceptación del servicio). No obstante, también puede generar expectativas sobre comprensión, empatía o capacidad de reparación que el agente automatizado no está diseñado ni en condiciones de satisfacer. Por tanto, el proceso de humanización no debe limitarse a atribuir rasgos humanos a la máquina; debe centrarse en rediseñar las señales sociales que faciliten la interacción sin inducir a interpretaciones erróneas ni a expectativas de comprensión o de capacidad de decisión para las cuales el sistema no ha sido diseñado. En la práctica, las organizaciones deben evitar transmitir, de forma implícita o ambigua, que el sistema posee comprensión, empatía o agencia moral; en su lugar, conviene declarar explícitamente

las limitaciones del agente, ofrecer mecanismos claros de intervención o de revisión humana y diseñar señales sociales que faciliten la interacción sin generar expectativas de comprensión, responsabilidad o reparación que excedan las capacidades reales del sistema.

4.2. Calidez y competencia: ejes perceptuales que guían aceptación, confort y calidad percibida

La investigación sobre la interacción con agentes tecnológicos indica que la evaluación del usuario se articula en torno a dos cuestiones clave: la competencia del agente (capacidad para realizar la tarea con eficacia) y la calidez o benevolencia percibida (desarrollo de actuaciones con buenas intenciones y respeto hacia el usuario). En robots y asistentes, la competencia suele asociarse con rendimiento, precisión y rapidez; la calidez, con estilo comunicativo, señales de escucha y trato respetuoso. El desafío surge cuando la mejora en la competencia no va acompañada de condiciones de transparencia y de control percibido. En ese caso, el sistema puede parecer "demasiado autónomo", y esa autonomía percibida puede activar emociones negativas o intentos de recuperar el control, especialmente en situaciones de fallo o de conflicto (Schepers et al., 2022).

Para situar y clarificar el análisis, adoptamos una tipología operativa de inteligencia artificial en servicios que distingue tres funciones principales: (1) mecánica, encargada de automatizar tareas repetitivas y estandarizadas; (2) analítica, orientada a la clasificación y predicción a partir de datos; y (3) emocional, destinada a la detección o simulación de respuestas afectivas. Esta distinción práctica facilita la identificación de las expectativas de los usuarios, así como de las responsabilidades, los controles y las medidas de supervisión más adecuados para cada función (Huang & Rust, 2018; Huang et al., 2019). La dificultad no es solo técnica. Es experiencial. A medida que el sistema se desplaza desde lo mecánico hacia lo emocional, la interacción se vuelve más "social" y, por tanto, más sensible a las expectativas morales. Una misma señal (p. ej., una disculpa automatizada) puede resultar útil si se entiende como un procedimiento, pero puede percibirse como simulación si el usuario interpreta que se está "representando" empatía sin experiencia.

4.3. Empatía simulada: cuándo aporta valor y cuándo exige límites y protección al usuario.

La empatía simulada actúa a menudo como una moneda de dos caras: por un lado, facilita el acompañamiento y la contención; por otro, puede inducir a engaño, a la dependencia o a expectativas de comprensión que el sistema no está en condiciones de satisfacer. En algunos contextos, el apoyo emocional de un asistente digital puede reducir el estrés, aumentar el compromiso y la satisfacción, especialmente cuando el usuario percibe que se le escucha y se le orienta (Gelbrich et al., 2020). Sin embargo, el mismo mecanismo puede derivar en engaño implícito si la relación depende de que el usuario crea que hay comprensión humana detrás. En ese caso, el sistema se apoya en una ilusión frágil.

Por tanto, desde un enfoque centrado en las personas, la recomendación es clara: proporcionar contención y apoyo sin dar la impresión de que la máquina es humana. El usuario puede beneficiarse de un estilo cálido y de respuestas cuidadosas, pero necesita señales honestas sobre la artificialidad, los límites y los mecanismos de derivación. En particular, cuando la vulnerabilidad es alta, el riesgo de dependencia o de sustitución afectiva aumenta y la organización debería velar por la protección del usuario y evitar explotar su compromiso. Esta idea se hace patente en el contexto de las tecnologías inmersivas cuando se advierte que la inteligencia artificial generativa puede parecer más humana precisamente en momentos de vulnerabilidad, y que humanizar exige límites y opciones de salida para preservar la autonomía. En la práctica, esto implica diseñar interacciones que combinen calidez con claridad: avisos sencillos sobre la naturaleza automática del agente, controles de la intensidad de la interacción (p. ej., límites de frecuencia o duración) y mecanismos claros que permitan interrumpir o pausar la relación cuando el usuario lo desee.

4.4. Robots en primera línea de servicio: cuándo aportan valor y cuándo deben ceder el protagonismo al humano

Los robots de primera línea pueden aportar valor al reducir la fricción en tareas estandarizadas, ofrecer consistencia y liberar recursos humanos en momentos críticos. Sin embargo, introducir robots y agentes automatizados implica decisiones concretas que transforman la experiencia del cliente y las relaciones entre la organización y el usuario: determinar dónde, cuándo y cómo automatizar cambia quién toma las decisiones, qué se espera del servicio y qué responsabilidades y desequilibrios de poder pueden surgir. La literatura ha insistido en la necesidad de comprender en qué etapas del *customer journey* los clientes prefieren la interacción con la inteligencia artificial y en qué momentos prefieren el contacto humano, especialmente en posventa y en la recuperación del servicio (Flavián & Casaló, 2021).

No es una decisión binaria entre "robot" y "humano"; lo que importa es coordinar qué tareas debe asumir cada actor y en qué momentos, de modo que la automatización mejore la eficiencia sin eliminar el reconocimiento, la empatía y la capacidad de respuesta humana en fases de alta carga emocional o de ambigüedad. El robot puede centrarse en ofrecer información o en ejecutar tareas, mientras que el humano debería mantener el protagonismo en los momentos de alta carga emocional, de ambigüedad o de conflicto. Si se automatiza lo sensible, aumenta el riesgo de deshumanización, no porque la máquina sea "mala", sino porque el usuario siente que su situación no encuentra reconocimiento real. Por consiguiente, la cuestión clave no consiste en optar de manera categórica entre la presencia de agentes automatizados y la intervención humana; se trata, más bien, de identificar qué tareas son susceptibles de automatización de forma segura y eficiente, y cuáles deben permanecer bajo supervisión, juicio y responsabilidad humanos. Esta lógica prepara el terreno para el modelo híbrido de colaboración humano–máquina.

4.5. Diseño responsable de señales sociales: tono, expresividad, proximidad, interactividad y transparencia

Diseñar señales sociales es diseñar inferencias. Un tono conversacional, una sonrisa mecánica, una proximidad corporal o una proactividad excesiva puede activar lecturas de calidez o de invasión, según el contexto. La literatura sobre la implementación de robots de servicio subraya elementos como la proactividad, los tiempos de respuesta, el estilo de interacción y la capacidad de expresar afecto como dimensiones críticas del diseño (Wirtz et al., 2020).

En el contexto de la humanización, estas señales deberían seguir una regla simple: maximizar la claridad y el respeto antes que el antropomorfismo. Esto implica, por ejemplo, evitar estrategias que induzcan a creer que el agente "siente" cuando, en realidad, ejecuta patrones. Implica también diseñar la interactividad con límites, de modo que el usuario pueda regular la intensidad, las pausas y la derivación a un humano. Y, en paralelo, implica transparencia operativa: que el usuario pueda distinguir cuándo interactúa con un sistema automatizado, qué capacidad real tiene y qué responsabilidad asume la organización si el sistema falla.

5. Del reemplazo a la complementariedad: control y rendición de cuentas en servicios automatizados

5.1. La tecnología potencia el servicio; las personas mantienen control y responsabilidad

El debate sobre la integración de la inteligencia artificial en los servicios ha evolucionado progresivamente: se ha pasado de una visión basada en el reemplazo de trabajadores por sistemas automatizados a otra que enfatiza la colaboración y la complementariedad entre humanos y máquinas. Esta transición no es solo parte de una preferencia ética; también es una constatación funcional: muchos sistemas necesitan supervisión humana, y muchos momentos del servicio requieren la capacidad de negociación, la contención emocional, el juicio y el tacto de una persona (habilidades que no pueden reducirse a un mero cálculo). En la literatura reciente se consolida la idea de que el objetivo no es sustituir, sino organizar un reparto de tareas y responsabilidades en el que la

complementariedad genere mejores resultados y, al mismo tiempo, preserve los principios de gestión responsable (Kunz et al., 2025).

El término "control" de este principio rector describe la capacidad del usuario y de la organización para supervisar, comprender y corregir el comportamiento del sistema, así como para asumir la responsabilidad por las decisiones que este toma; se articula en dos dimensiones complementarias: (a) control percibido por el usuario (la posibilidad real de entender, modular y corregir) y (b) control organizativo (designación de responsables, mecanismos de auditoría y procesos de aprendizaje institucional), especialmente cuando las decisiones automatizadas tienen impacto significativo. Por ello, la colaboración entre humanos–máquinas exige una supervisión sólida: sin una gestión responsable, los modelos híbridos corren el riesgo de convertirse en una delegación opaca.

5.2. Reparto de responsabilidades en el servicio: automatizar lo repetitivo, humanizar lo crítico

El reparto de responsabilidades se vuelve claro cuando se observa el servicio como una secuencia de microdecisiones: algunas rutinarias y otras irreversibles.

- En tareas de inteligencia artificial mecánica (automatización de procesos, clasificación básica), el beneficio de eficiencia suele ser alto y el coste relacional relativamente bajo. Esto conlleva efectos claros: mayor productividad, menos errores y la posibilidad de redistribuir personal hacia tareas de mayor valor, pero exige controles de calidad, trazabilidad y vías claras de apelación.
- En tareas de inteligencia artificial analítica (predicción, personalización avanzada), el beneficio puede ser muy alto, pero el riesgo de intrusividad y de injusticia percibida aumenta si no existe transparencia práctica (explicaciones claras sobre las decisiones tomadas, indicadores de confianza, controles para ajustar la personalización y vías claras de apelación y corrección).
- En tareas de inteligencia artificial emocional (detección y respuesta afectiva), el beneficio potencial es la contención, pero el riesgo es prometer una relación que el sistema no puede sostener de forma auténtica (Huang y Rust, 2018; Huang et al., 2019).

Desde esta perspectiva, “humanizar lo crítico” significa reservar la intervención humana para momentos en los que la persona necesita reconocimiento y negociación, por ejemplo, en conflictos, reclamaciones, decisiones sensibles o situaciones de vulnerabilidad (p. ej., emergencias médicas, crisis emocionales, discapacidad, personas mayores o usuarios en situación de precariedad económica). En estos casos suelen ser necesarias la intervención humana, explicaciones claras y vías de apelación accesibles. También significa diseñar la inteligencia artificial para apoyar a las personas, no para sustituir su juicio. Si el humano aparece solo como recurso de emergencia, tarde y mal, el usuario interpreta la organización como distante. En cambio, cuando las personas forman parte visible del sistema (es decir, cuando el usuario identifica y puede contactar a un agente humano dentro del proceso), la tecnología potencia y amplifica sus capacidades (mejorando la rapidez, el alcance y el acceso a la información) en lugar de servir como una "pantalla" que oculta decisiones y diluye responsabilidades.

5.3. Continuidad del customer journey: transición fluida entre máquinas y humanos en momentos sensibles

Si la colaboración híbrida tuviera que resumirse en una sola idea operativa, sería esta: el servicio se gana o se pierde en las transiciones. No porque el resto del recorrido sea irrelevante, sino porque en esos momentos el cliente busca una respuesta implícita: “¿Quién se hace cargo de mí ahora?”. Si la transición entre máquinas y humanos se realiza sin el contexto adecuado, resulta necesario repetir la historia o no se ofrecen las opciones adecuadas; la automatización deja de ser una ayuda y se convierte en una barrera. Por el contrario, una transición bien diseñada conserva el contexto, adapta el tono al caso y proporciona una salida humana visible. Entonces, la tecnología refuerza y facilita la atención relacional, sin sustituir la relación humana. En la práctica, esto implica: (a) transferencias con un resumen del caso y metadatos visibles para el agente receptor; (b) indicadores que señalen el grado

de urgencia o de vulnerabilidad; y (c) un punto de contacto humano claramente señalado y accesible (Kunz et al., 2025).

En experiencias inmersivas, la cuestión añade otra dimensión: la transición no solo cambia de agente, sino también de intensidad sensorial. Cuando la telepresencia o la vividez elevan las expectativas, retornar a un canal menos rico puede percibirse como una pérdida, aunque la información sea correcta. Por eso, debe diseñarse la salida como parte de la experiencia para evitar que lo inmersivo se perciba como un espectáculo desconectado de la realidad.

Desde esta perspectiva, la continuidad del *customer journey* exige dos capas coherentes y simultáneas. La continuidad operativa obliga a que el humano que recibe al cliente disponga de la información procesada por el sistema (decisiones previas, preferencias). La continuidad psicológica exige que el cliente perciba la misma intención de cuidado a lo largo del recorrido (tono, reconocimiento y opciones claras). Cuando esa continuidad se rompe, el usuario no culpa a la interfaz; culpa a la organización.

5.4. Fallos, reparación y justicia percibida: cómo recuperar la confianza si el sistema falla

La automatización se asocia con la consistencia, pero en los servicios la consistencia siempre es relativa. El contexto cambia, el cliente cambia y el sistema puede fallar. Lo decisivo, entonces, es cómo se interpreta el fallo y qué tipo de respuesta produce. En este contexto, la humanización deja de ser una aspiración abstracta y se convierte en una arquitectura de reparación: un conjunto de procesos (reconocimiento inmediato del fallo, explicación práctica, corrección efectiva, escalado humano visible y aprendizaje organizativo) que buscan restaurar la confianza y la justicia cuando los sistemas fallan.

La literatura sobre la recuperación del servicio sintetiza hallazgos bien establecidos: la evaluación de una reparación abarca las dimensiones distributiva, procedimental e interaccional, y cada una influye en la satisfacción y la continuidad de la relación cliente–organización (Tax, Brown & Chandrashekaran, 1998). Estudios longitudinales confirman que las percepciones de justicia se acumulan con el tiempo y que recuperaciones repetidamente deficientes erosionan la relación, con efectos duraderos (Maxham & Netemeyer, 2002). Revisiones y trabajos empíricos posteriores subrayan, además, que la dimensión interaccional (trato respetuoso y reconocimiento) mantiene un papel central (Aguilar-Rojas et al. 2015) incluso en contextos altamente tecnificados.

En los servicios automatizados, la justicia procedimental suele depender del diseño del circuito de resolución. Si el usuario queda atrapado en respuestas genéricas, la experiencia se vive como una pérdida de control. En cambio, cuando el sistema ofrece vías de reclamación comprensibles y un escalado real a humanos, el procedimiento transmite responsabilidad institucional: el usuario percibe que existe un canal claro para presentar quejas, respuestas comprensibles y plazos razonables, y acceso a una revisión humana con poder de decisión; esos elementos reducen la sensación de abandono y ayudan a restaurar el control y la dignidad del cliente. En la literatura especializada, esta idea se formula con precisión: el traspaso a un humano no es opcional, sino un mecanismo central de la justicia procedimental. Su calidad se evalúa mediante criterios operativos concretos: que el usuario no tenga que repetir su historia, que pueda escoger o aceptar el canal de atención, y que el traspaso preserve la continuidad informativa y el respeto y la autonomía del cliente.

Además, existe un matiz social que complica la reparación: cuando el sistema ha mostrado señales de calidez o empatía, la brecha entre esa promesa relacional y la incapacidad para resolver un fallo intensifica la frustración del usuario; deja de ser solo un error técnico y se percibe como una forma de engaño. Para reparar eficazmente hay que evitar dos extremos contrapuestos: la deshumanización defensiva (ocultar decisiones y dejar al usuario a su suerte) y la empatía simulada (expresar afecto sin capacidad real de resolución). Una reparación humanizada combina el reconocimiento inmediato del fallo, una explicación práctica y comprensible y un camino claro y efectivo hacia la resolución, sin teatralizar emociones que no se pueden sostener.

Además, existe una asimetría psicológica relevante: cuando los sistemas emiten señales sociales (p. ej., calidez o empatía), un fallo puede percibirse como engaño y agravar la reacción negativa.

Asimismo, estudios sobre aversión algorítmica muestran que las personas tienden a penalizar de forma desproporcionada los errores de los algoritmos (Dietvorst, Simmons & Massey, 2015), aunque investigaciones posteriores matizan este efecto según el contexto y la confianza previa. Por ello, automatizar los puntos críticos del recorrido incrementa el coste de los fallos. En definitiva, la excelencia en sistemas híbridos requiere integrar el diseño de errores y sus reparaciones en el propio servicio: planificar, medir y ensayar los mecanismos de recuperación en lugar de considerarlos anomalías puntuales (Tax et al., 1998; Maxham III & Netemeyer, 2002).

5.5. Trabajo aumentado: bienestar, competencias, aceptación interna y cultura de servicio

La colaboración humano–máquina depende tanto del diseño como del trabajo humano. Detrás de chatbots, robots o experiencias inmersivas hay empleados que supervisan, corrigen excepciones y reparan fallos; ignorar esa dimensión laboral empobrece la humanización y afecta tanto a la experiencia del cliente como a la del trabajador (Kellogg, Valentine & Christin, 2020).

Automatizar lo costoso para la empresa no equivale a automatizar lo apropiado para la relación con el cliente. Cuando la automatización afecta puntos de contacto que los usuarios consideran centrales, puede interpretarse como desinversión relacional, aun cuando la función se cumpla; del mismo modo, si deja a los trabajadores como "parches", sin autoridad ni recursos, deteriora la cultura de servicio. La solución exige un reparto explícito de roles, formación y objetivos compartidos, no solo la implementación tecnológica aislada (Belanche et al., 2024).

La distribución de la responsabilidad es decisiva: la colaboración debe amplificar capacidades sin expropiar la autonomía y la capacidad de decisión de las personas. Si la responsabilidad se difumina, se erosiona la confianza de clientes y empleados y aumentan las ineficiencias y los costes operativos (Rahwan et al., 2019).

El trabajo aumentado es una oportunidad para revalorizar lo humano (evaluación in situ, atención centrada en la persona y capacidad de reparación), pero solo si la automatización libera tiempo de tareas repetitivas y no intensifica la carga ni precariza el encuentro. Sin bienestar laboral y una cultura de servicio, la tecnología más sofisticada puede generar experiencias frías y desconectadas (Kellogg, Valentine y Christin, 2020).

6. Experiencias inmersivas: diseñar presencia que mejora decisiones

6.1. Tecnologías inmersivas como nuevo lenguaje experiencial: presencia, participación y memoria

Las experiencias inmersivas se diferencian por la combinación de tres elementos clave (Flavián et al., 2019b): *embodiment* (integración corporal y sensorial), presencia psicológica (la sensación real de «estar allí») e interactividad significativa (posibilidad de actuar y afectar el entorno). No son los gráficos los que la definen, sino el grado en que estos tres componentes se articulan. Esta triple articulación tiene implicaciones directas para generar un mayor valor percibido. La presencia, entendida como la sensación de estar en el entorno, tiende a incrementar la implicación, la atención y la memoria episódica, lo que convierte a las tecnologías inmersivas en herramientas especialmente atractivas para industrias en las que la intangibilidad genera incertidumbre (Slater y Sanchez-Vives, 2016). En el contexto del turismo, por ejemplo, la inmersión funciona como una preexperiencia que reduce las dudas y hace la decisión más "vivida" antes de que sea real. Sin embargo, si la inmersión se diseña únicamente como un impacto, se agota rápido. El valor sostenible depende de la capacidad de sostener significados como pertenencia, autenticidad y autoexpresión sin desbordar la atención del usuario ni erosionar su control percibido.

Surge aquí un paralelismo interesante: la presencia social automatizada (p. ej., agentes que emiten señales de empatía) puede reducir tensiones psicológicas y aumentar el valor percibido (Flavián et al. 2024a); de forma análoga, la inmersión puede generar efectos similares por vías sensoriales y espaciales. No obstante, humanizar no equivale a intensificar el realismo técnico: implica garantizar que el usuario conserve la libertad de entrar, de modular su grado de implicación y de disponer de controles claros sobre la experiencia.

6.2. Del entorno físico al inmersivo: cómo las tecnologías inmersivas moldean las decisiones, potencian el aprendizaje y los vínculos con marcas y organizaciones

La transformación clave no radica en que el entorno sea "virtual", sino en el concepto de *embodiment* (encarnamiento). Este concepto permite describir cómo la experiencia integra la acción corporal, la percepción sensorial y la presencia corporal; es decir, la toma de decisiones integra la percepción y la acción corporales (sensaciones, movimientos y control), no solo un procesamiento abstracto de información (Guy et al., 1991). En entornos inmersivos, el *embodiment* explica por qué recorrer un espacio o manipular objetos afecta las emociones, la memoria y las valoraciones (la decisión se anticipa y se siente) y, por tanto, por qué la coherencia entre lo mostrado y lo real es crítica para evitar desajustes de expectativas. Recorrer un espacio, manipular objetos o explorar un servicio en primera persona convierte la evaluación en una experiencia anticipada. Esta vivencia puede reducir la percepción de riesgo, aumentar la confianza en la decisión y reforzar el vínculo con la oferta (p. ej., una visita virtual a un hotel que permite valorar la habitación, las vistas y los servicios antes de reservar), pero también puede inflar expectativas si lo mostrado no se corresponde con la realidad (Orús et al., 2021). Por ello, la humanización exige coherencia entre la promesa inmersiva y la entrega posterior, y el diseño de salidas que traduzcan lo vivido en información práctica y opciones de control para el usuario.

6.3. Multisensorialidad y coherencia: intensificar sin saturar: equilibrio emoción–carga cognitiva

La multisensorialidad es, a la vez, la potencia y el riesgo de las tecnologías inmersivas. Aumentar los estímulos puede enriquecer la emoción y la presencia, pero también puede saturar, fatigar o generar malestar. De hecho, hay límites humanos duros y uno de los más relevantes es la cibernausia o *cybersickness*. La literatura destaca que el malestar puede presentarse incluso en usuarios motivados, reduciendo el tiempo tolerable y afectando la evaluación global de la experiencia. Además, la relación entre presencia y *cybersickness* es compleja, lo que refuerza una implicación crítica: maximizar la presencia como objetivo único puede ser una forma inadvertida de deshumanización, porque obliga al usuario a elegir entre "aguantar" y abandonar. En consecuencia, la ergonomía, la estabilidad visual, las opciones de pausa y el control de la intensidad no son detalles técnicos; son condiciones de cuidado (Guy et al., 2023).

A pesar de lo anterior, la saturación no es solo fisiológica. También es cognitiva. Cuando el entorno compite por la atención con demasiados estímulos, el usuario puede perder el foco en los elementos centrales y disminuir la autenticidad percibida. La gamificación puede ayudar a mantener la atención en ciertos contextos, pero no sustituye la necesidad de orientación y claridad. Humanizar, desde esta perspectiva, significa diseñar un recorrido atencional que respete límites humanos y que no convierta la intensidad en obligación.

Aquí no basta con afirmar que la inmersión puede deshumanizar; hay que identificar qué formas adopta el riesgo y qué salvaguardas lo previenen. Por ello, el Cuadro 6 sintetiza los dilemas más frecuentes cuando las tecnologías inmersivas se combinan con optimización algorítmica (captura de atención, promesas irreales, publicidad camuflada, decisiones aceleradas e identidad confusa) y los traduce en medidas simples de diseño: transparencia, pausas, control del usuario, etiquetado comprensible y fricción mínima para salir o pedir ayuda humana. Así, la intensidad deja de ser un fin en sí mismo y se convierte en un medio proporcional al objetivo: enriquecer la experiencia sin capturar al usuario, sin inflar expectativas ni desplazar su autonomía.

Cuadro 6: Riesgos de persuasión, confianza y bienestar en asociados a la combinación de las tecnologías inmersivas y la inteligencia artificial (y cómo mitigarlos)

Riesgo *(tipo):* explicación	**Intensificación *(impacto)***	**Estrategias de mitigación**
Captura de atención *(sobrecarga)*: la experiencia "tira" del usuario con estímulos y tareas.	XR añade novedad + estímulos 360; IA ajusta el "gancho" al momento. *(cansancio, desconexión, estrés)*	Diseño con foco (1 tarea); pausas; límite de sesión; modo "baja estimulación".
Promesa irreal *(expectativas)*: lo vivido parece mejor que lo real y luego decepciona.	3D/360 eleva realismo; IA "embellece" o suaviza defectos. *(frustración, desconfianza)*	Aviso "recreación"; mostrar condiciones reales; comparador (real vs simulado); pruebas con usuarios.
Publicidad camuflada *(intrusión)*: la marca entra como juego sin quedar claro.	Filtros RA mezclan ocio y venta; IA personaliza mensaje y timing. *(rechazo, sensación de invasión)*	Etiqueta "patrocinado"; control de frecuencia; opt-out visible; separar capa lúdica/comercial.
Compra sin entender *(decisión rápida)*: reduce dudas pero no asegura comprensión.	RA "prueba" y baja riesgo; IA recomienda para cerrar la compra. *(impulso, arrepentimiento)*	"Por qué se sugiere"; confirmación antes de pagar; info clave simple; devolución fácil.
Identidad confusa *(confianza)*: no queda claro quién es humano, IA o avatar.	Avatares + voz IA parecen "personas"; fácil imitar estilos. *(inseguridad, engaño)*	Indicador "IA/humano"; verificación en servicios sensibles; botón "hablar con persona"; reportar abuso.

Fuente: Barta et al. (2023b; 2024a); Flavián et al. (2023); Madary y Metzinger (2016) y Oliver (1980).

6.4. Tecnologías inmersivas para bienestar y apoyo emocional: usos legítimos, límites y responsabilidades

Las tecnologías inmersivas pueden contribuir al bienestar y al apoyo emocional si son concebidas con un propósito específico (p. ej., favorecer la calma), se ofrecen en sesiones breves y cerradas (con inicio, duración prevista y cierre) y garantizan el control del usuario (p. ej., ajustar la intensidad, pausar y salir de inmediato sin fricción). La evidencia sobre la presencia indica que la VR amplifica la experiencia subjetiva; por ello, las recomendaciones éticas insisten en establecer salvaguardas concretas para minimizar los riesgos y el malestar (Madary y Metzinger, 2016; Sánchez-Vives y Slater, 2005).

En este contexto, resulta clave reforzar las salvaguardas: proporcionalidad entre la intensidad y el objetivo, reversibilidad (salidas inequívocas, pausas y recuperación) y prevención del malestar, incluida la mitigación de la *cybersickness*, que afecta el uso y la evaluación incluso en usuarios motivados (Madary y Metzinger, 2016). De igual forma, debe preservarse la autonomía: guiar sin capturar y proteger la libertad del usuario para entrar, permanecer o salir (Sánchez-Vives y Slater, 2005).

6.5. Tecnologías inmersivas para garantizar inclusión y accesibilidad

La inclusión debe incorporarse desde el principio. Para ello, conviene ofrecer formas alternativas de interacción (p. ej., control por mirada o mando además de gestos), permitir ajustar la intensidad (p. ej., volumen, brillo, velocidad de movimiento), asegurar la compatibilidad con tecnologías asistivas (p. ej., lectores de pantalla en menús y subtítulos) y prever recorridos equivalentes cuando la inmersión no sea posible (p. ej., versión 2D del recorrido o audio-guía).

Las recomendaciones del W3C ofrecen criterios para hacer los contenidos web más accesibles y son relevantes para las interfaces web y móviles que rodean experiencias inmersivas (p. ej., en un servicio VR para visitar un museo: configurar subtítulos y volumen; orientarse con un tutorial y un mapa simple) (W3C, 2023). Además, la inclusión no puede separarse del acceso real: de poco sirve diseñar

una experiencia “accesible” si, en la práctica, una parte de los usuarios no puede participar debido a barreras de hardware, conectividad, coste o a la falta de apoyos disponibles.
Por último, la gestión de los datos es particularmente sensible en entornos inmersivos, ya sea por los datos biométricos capturados o por las huellas de uso que pueden revelar problemas de salud. Proteger a los usuarios implica recoger lo mínimo, explicarlo con claridad y darles control sobre el consentimiento y el borrado (Sánchez-Vives y Slater, 2005).

7. Industria 4.0 y sociedad: retos, oportunidades y consecuencias

7.1. Promesa de la automatización inteligente: eficiencia, conveniencia y nuevos modelos de servicio

La automatización inteligente ofrece una gran promesa: mayor eficiencia operativa, mayor conveniencia para el usuario y el desarrollo de nuevos modelos de servicio, lo que potencia la personalización, la atención proactiva y los servicios bajo demanda. No obstante, desde una perspectiva centrada en las personas, esa promesa debe matizarse: la digitalización no solo crea canales y capacidades técnicas. También redistribuye quién puede participar, en qué condiciones y con qué beneficios. De hecho, la automatización, además de potenciar capacidades, también puede acentuar las desigualdades. Los colectivos más vulnerables se enfrentan a riesgos adicionales como consecuencia de su menor alfabetización digital, menor capacidad para interpretar señales persuasivas o mayor dificultad para reclamar (Wirtz et al., 2018).
Como consecuencia, si consideramos la humanización un objetivo del sistema, no podemos limitarnos a desarrollar interfaces más amables. Hay que revisar las reglas, las responsabilidades y los recursos. Debe establecerse quién decide (algoritmo o persona) y crear un proceso de apelación atendido por personas. Mantener vías equivalentes de atención (teléfono, chat humano, oficinas) para usuarios con baja alfabetización digital. Contar con registros de las decisiones relevantes (fecha, aspectos clave, responsable humano) para auditoría y revisión o para analizar quién recibe qué beneficios y establecer mecanismos de compensación si algunos grupos se vieran desfavorecidos.
Estas medidas de "innovación responsable" reducen la asimetría informativa (los usuarios entienden las razones), preservan la autonomía (pueden corregir o elegir) y mantienen la confianza pública (hay responsables y remedios claros).

7.2. Riesgos de deshumanización: fricción emocional, pérdida de confianza, alienación y dependencia

En servicios mediados por inteligencia artificial y en entornos inmersivos, la deshumanización surge cuando la persona no puede corregir errores, no entiende por qué el sistema actúa así, no encuentra un responsable visible o percibe que su vulnerabilidad no fue considerada en el diseño (Flavián et al., 2024a).
Este conjunto de tensiones psicológicas organiza la vivencia del consumidor (Puntoni et al. 2021): sentirse servido o explotado; comprendido o malinterpretado; empoderado o reemplazado; conectado o alienado. La tensión se vuelve crónica cuando los diseños optimizan exclusivamente métricas de eficiencia o de *engagement* y no integran el bienestar, la autonomía y la justicia percibida como criterios explícitos de éxito (Puntoni et al., 2021).
Un riesgo afín es la dependencia: los sistemas inmersivos y los procesos de personalización pueden diseñarse para intensificar la permanencia, la repetición y la exposición (Flavián et al., 2024b). Si la retención se logra mediante estímulos adictivos o mecánicas de gamificación sin propósito, el *engagement* compite con el bienestar y la experiencia se reduce a una mera retención. Por ello, resulta recomendable establecer algunas contramedidas, como incorporar puntos de salida claros y fáciles de usar en experiencias inmersivas, establecer protocolos de escalado humano o definir límites éticos para técnicas de retención y evitar mecánicas que exploten sesgos cognitivos.

7.3. Desigualdades y brechas tecnológicas: acceso, alfabetización digital y distribución del valor

La Industria 4.0 puede democratizar el acceso, pero también puede amplificar las desigualdades. El acceso al servicio depende de los dispositivos, de la conectividad, de la alfabetización digital y, en contextos inmersivos, de la tolerancia física y cognitiva ante experiencias intensas. A ello se suma una desigualdad menos visible: la distribución del valor en sistemas basados en datos. Por ejemplo, los usuarios con amplios historiales podrían recibir ofertas personalizadas y atención prioritaria; quienes no tienen datos, en cambio, podrían recibir respuestas genéricas o ser ignorados. La brecha no es solo material: también es relacional y simbólica (Puntoni et al., 2021).

En tecnologías inmersivas, el problema se manifiesta con claridad si la inmersión se convierte en la vía principal de acceso a ciertos servicios, pero el hardware y la conectividad quedan restringidos a segmentos privilegiados; la innovación no amplía la participación, sino que la segmenta.

Estos desajustes podrían corregirse mediante las medidas propuestas por el World Wide Web Consortium (2023) (W3C): ofrecer alternativas equivalentes no inmersivas; asegurar la accesibilidad en el registro y la configuración; permitir múltiples modos de interacción (teclado, voz, controles adaptativos); incluir ajustes de accesibilidad (subtítulos, descripciones de audio, opción de reducir el movimiento); e implementar modos de baja fidelidad y requisitos de conectividad mínimos.

En este contexto, humanizar no significa uniformizar, sino asegurar la equivalencia del valor recibido: ofrecer alternativas no inmersivas que no sean de segunda categoría, rutas de ayuda humana accesibles o potenciar la transparencia para quienes tienen menos recursos para interpretar señales persuasivas o procedimientos automatizados.

8. Ética, inclusión y gestión responsable: condiciones para la confianza pública

8.1. Privacidad en inteligencia artificial y tecnologías inmersivas: datos sensibles, biometría, comportamiento y "huella" digital e inmersiva

La privacidad se ha convertido en uno de los aspectos más sensibles y polémicos de la digitalización avanzada, no solo por la cantidad de datos, sino también por su naturaleza y su capacidad inferencial. En los servicios mediados por inteligencia artificial, los riesgos principales se asocian al desarrollo de perfiles, a las inferencias (p. ej., sobre rasgos de personalidad, estados emocionales o condiciones de salud) y a las decisiones automatizadas.

En entornos inmersivos, el riesgo se desplaza hacia datos de carácter corporal (movimientos de cabeza y de manos, trayectorias de mirada, patrones de navegación espacial y tiempos de atención). Todo ello compone una huella inmersiva que puede revelar preferencias, emociones, hábitos y, en ciertos casos, incluso enfermedades o vulnerabilidades. Junto a la huella digital tradicional (registros de uso, historiales y metadatos), la huella inmersiva (señales cinestésicas, oculares y temporales) amplía la capacidad de inferencia y reidentificación, elevando los riesgos para la privacidad, la autonomía y la toma de decisiones informada (Kosinski, Stillwell y Graepel, 2013; Sánchez-Vives y Slater, 2005).

En este contexto, el protocolo de "consentimiento" puede convertirse en un mero ritual. La comprensión del alcance real es difícil incluso para usuarios expertos, y la combinación de huellas inmersivas con otras fuentes permite construir perfiles predictivos que transforman la experiencia psicológica de "estar en un entorno" en "estar siendo medido" (Sánchez-Vives y Slater, 2005; Madary y Metzinger, 2016).

Ante esta situación, resulta necesario pasar de los avisos legales a ofrecer un control efectivo sobre el usuario. Esto implica ofrecer opciones comprensibles para limitar la recogida, modular la personalización, borrar huellas y evitar usos secundarios. Implica también reconocer que ciertos datos (biométricos e inmersivos) son cualitativamente más sensibles y que tratarlos como equivalentes al rastreo convencional es un error e insuficiente.

En el marco europeo, estas exigencias no son solo normativas; también son estratégicas. El RGPD establece principios de licitud, transparencia, minimización y limitación de la finalidad, con implicaciones directas para entornos que recolectan trazas conductuales y datos biométricos (Regulation (EU) 2016/679, 2016).

8.2. Sesgos y discriminación: riesgos en modelos, contenidos y decisiones automatizadas

Los sesgos algorítmicos suelen tratarse como fallos técnicos, pero en el ámbito de los servicios operan como fallos morales porque afectan el acceso, el trato y la igualdad de oportunidades. Un sistema de recomendación que invisibiliza a ciertos perfiles, un *scoring* que penaliza sistemáticamente a determinados colectivos o un generador de contenido que reproduce estereotipos pueden deteriorar la confianza y la credibilidad incluso cuando la experiencia individual parece "correcta". La discriminación algorítmica, además, puede ser invisible para quien no la sufre y, por tanto, carecer de retroalimentación espontánea que permita corregirla (Yang et al., 2024).

Los sesgos y la discriminación no se plasman solo en decisiones explícitas de inclusión o exclusión; también pueden apreciarse en el tono del lenguaje, en la priorización de la atención y en el propio proceso de construcción de perfil, donde historiales desiguales producen experiencias desiguales (Akter et al., 2023). Por otra parte, la opacidad hace que el usuario intuya la existencia de un trato injusto sin poder demostrarlo, lo que erosiona la confianza, incluso aunque no se presente una queja formal (Dalenberg et al., 2018). La consecuencia es clara: la humanización exige revisión humana, transparencia y supervisión del lenguaje, especialmente en inteligencia artificial generativa, y exige mecanismos de explicación suficientes para garantizar la confianza y la credibilidad.

En ese punto, la lógica de "gestión de riesgos" resulta más útil que la de "buenas intenciones". Conviene adoptar un marco operativo que traduzca principios como la equidad, la transparencia y la rendición de cuentas en prácticas concretas: evaluaciones de impacto y pruebas de sesgo; auditorías independientes y registros de la toma de decisiones; y supervisión humana con mecanismos de reparación a lo largo del ciclo de vida del sistema (Akter et al., 2023).

9. Conclusión: Humanizando la tecnología para mejorar nuestras vidas.

¿En qué consiste realmente humanizar cuando la tecnología deja de ser solo una herramienta y pasa a ser presencia? No es maquillar interfaces con voces amables o avatares simpáticos para hacer aceptable una lógica de extracción de datos o de reducción de costes. Tampoco es maximizar el realismo ni el *engagement*. Humanizar significa diseñar condiciones en las que la persona conserve el control, comprenda el sistema y se sienta segura al interactuar con sistemas capaces de intensificar las emociones, evitando que la tecnología determine por ella lo que debe sentir o decidir. Esta idea se despliega en cuatro ejes aparentemente simples, pero que entrañan cierta complejidad:

1. La tecnología actúa como actor social porque la interpretamos así: atribuimos intenciones, evaluamos la calidez y la competencia, y construimos confianza a partir de señales que antes se dirigían a personas. Por eso no basta con el antropomorfismo; hacen falta señales sociales responsables que hagan legible la interacción sin inducir engaño ni dependencia (van Doorn et al., 2017; Belanche et al., 2020b).
2. El valor experiencial supera la utilidad y la conveniencia; incluye emoción, significado, pertenencia y bienestar. En los servicios automatizados, una parte decisiva del valor se manifiesta cuando se alivian tensiones psicológicas propias de la automatización (el sistema debe funcionar, pero el usuario debe percibir que tiene el control y que no está siendo dirigido) (Puntoni et al., 2021).
3. La responsabilidad percibida no desaparece con la automatización, sino que se concentra. Si el sistema falla, la organización debe asumir la responsabilidad y la reparación. Es su responsabilidad etiquetar el contenido sintético, explicar las decisiones en un lenguaje práctico y diseñar transiciones claras para los humanos en momentos críticos (Belanche et al., 2020a).
4. La idea de reemplazar a los humanos debe ser sustituida por la colaboración humano–máquina. El futuro es híbrido. La clave es identificar las tareas a automatizar, las reservadas al juicio humano y cómo diseñar las transiciones que preserven la dignidad del usuario, la justicia percibida y el trato responsable (Kunz et al., 2025).

Para facilitar su aplicación, el Cuadro 7 presenta un *marco unificado de humanización en inteligencia artificial y tecnologías inmersivas* que integra criterios de experiencia y de prestación responsable del

servicio. El cuadro traduce la humanización en indicadores auditables: control y salidas, transparencia práctica, bienestar y seguridad, justicia e inclusión y cuidado relacional, e incorpora, además, tres condiciones clave para que el enfoque sea viable en la práctica: responsabilidad y reparación ante fallos, intención percibida de la automatización (mejora del servicio vs recorte de costes) y colaboración humano–máquina (reparto de tareas y transiciones).

Si entendemos el servicio como un ecosistema, la humanización deja de ser un atributo opcional y se convierte en un criterio de calidad. Un sistema es más humano cuando permite decidir sin sentirse conducido, reduce la fricción sin aumentar la opacidad, sorprende sin capturar y protege la vulnerabilidad sin convertirla en una variable de optimización. La inmersión amplifica el valor, pero también los riesgos si no se incorporan límites, pausas y salidas comprensibles. Muchos fracasos tecnológicos no son fallos técnicos, sino fracasos de percepción o de relación: recomendaciones precisas que generan intrusividad, entornos inmersivos que saturan, robots que aparentan atención pero no resuelven conflictos y obligan a repetir el problema de forma reiterada. El daño reputacional a menudo supera al daño funcional cuando la organización parece priorizar la eficiencia por encima de la dignidad.

Por eso conviene adoptar un principio rector: la tecnología amplifica; las personas conservan el control percibido y la responsabilidad. Ese principio evita el falso dilema entre la automatización y la humanidad y orienta el debate hacia el reparto de tareas y la responsabilidad visible. La inteligencia artificial mecánica puede liberar recursos de tareas repetitivas; la inteligencia artificial analítica puede respaldar la toma de decisiones y la personalización; y la inteligencia artificial emocional puede ofrecer contención limitada en contextos bien definidos. Pero el criterio no debe ser “hacer más”, sino “hacer con límites”: preservar la justicia percibida, la transparencia operativa y las rutas humanas cuando la situación exija el juicio, la capacidad de negociación o la atención prestada por otra persona.

En definitiva, la tecnología se vuelve humana cuando la organización y el ecosistema que la rodea toman decisiones coherentes con un ideal de respeto práctico. Cuando integra la privacidad como un control efectivo, no como un ritual. Cuando trata la accesibilidad como criterio de calidad y no como una adaptación tardía. Cuando mide la permanencia, pero también la fatiga. Cuando la reparación forma parte del servicio y no es una excepción. Cuando se potencia la presencia social que acompaña frente a la que captura. En última instancia, la humanización es una forma de gestionar la innovación (Ostrom, Parasuraman, Bowen, Patrício y Voss, 2015). Si se acepta esa premisa, la tecnología pasa a ser un instrumento de cuidado colectivo y la innovación podrá utilizarse para ampliar las libertades, reducir las vulnerabilidades y potenciar oportunidades reales para todos.

Cuadro 7. Marco unificado de humanización en inteligencia artificial y tecnologías inmersivas

Pilar de humanización (pregunta guía)	**Implicaciones en IA**	**Importancia / implicaciones en XR**	**Indicaciones para auditoría (*check* rápido)**
Control *(¿La persona puede salir fácilmente?)*	La IA puede empujar decisiones y automatizar acciones sin límites claros.	La inmersión puede "arrastrar" a seguir; sin controles, el usuario se siente sobrecargado.	¿Hay pausa/salir siempre visible? ¿Ajustes simples de intensidad/sonido?
Comprensión y transparencia *(¿Entiende qué está viendo y por qué?)*	No se entiende por qué recomienda, qué datos usa o el objetivo.	En XR se confunde fácil lo real con lo recreado	¿"Por qué esto" en recomendaciones? ¿Lenguaje simple, sin jerga?
Seguridad y bienestar *(¿Evita saturación o malestar?)*	Puede optimizar "enganche" a costa del bienestar.	Riesgo de mareo, fatiga o sobreestimulación	¿Límites de sesión y descansos? ¿Avisos de contenido intenso?
Justicia e inclusión *(¿Funciona para distintas capacidades y contextos?)*	Puede introducir sesgos (reconocer peor acentos/rostros, recomendar de forma desigual).	XR puede excluir por hardware, mareo, visión/oído, movilidad o espacio físico.	¿Revisión de sesgos? ¿Pruebas con perfiles diversos? ¿Subtítulos/controles accesibles?
Cuidado relacional *(¿Acompaña sin engañar ni sustituir lo humano cuando hace falta?)*	Agentes IA pueden parecer humanos; riesgo de engaño o dependencia.	Avatares y "presencia" pueden crear falsa cercanía o presión social.	¿Se indica si es IA o humano? ¿Botón "hablar con persona"? ¿Registro de incidencias?
Responsabilidad y reparación *(¿Quién responde y cómo se repara si falla?)*	En fallos, la responsabilidad percibida suele recaer en la organización.	En XR, el realismo puede amplificar el daño percibido del fallo (desorientación, engaño, pérdida económica).	¿Protocolo de recuperación del servicio? ¿Explicación post-fallo y trazabilidad? ¿Canal humano y compensación?
Intención percibida de automatización *(¿Mejora servicio o reduce costes?)*	La reducción de costes reduce la aceptación; si se percibe mejora, sube valor.	En XR, la "magia" puede ocultar objetivos (publicidad, manipulación).	¿Se comunica beneficio al usuario? ¿Evidencia de mejora (tiempo, calidad)?
Colaboración humano-máquina *(¿Complementa y dignifica roles?)*	El valor aumenta si IA y humanos colaboran: IA para escala; humano para juicio, empatía y excepciones.	En XR, la transferencia debe ser fluida; supervisión humana en momentos críticos.	¿Supervisión y escalado? ¿Roles claros? ¿Formación del personal?

Fuente: Belanche et al. (2020b; 2021a; 2024); Flavián et al. (2019a; 2019b); Ibáñez-Sánchez et al. (2022); Kunz et al. (2025) y Madary y Metzinger (2016); Sanchez-Vives y Slater (2005).

Nota: Este texto fue revisado y mejorado con la asistencia de herramientas de edición y corrección (función Canvas y Grammarly). Estas herramientas se emplearon únicamente con fines de revisión y mejora estilística; la autoría intelectual y la responsabilidad del contenido final corresponden en su totalidad al autor.

Bibliografía:

Aguilar-Rojas, O., Fandos-Herrera, C., y Flavián-Blanco, C. (2015). What may lead you to recommend and revisit a hotel after a service failure instead of complaining?. *International Journal of Contemporary Hospitality Management*, *27*(2), 214-235. https://doi.org/10.1108/IJCHM-06-2013-0265

Akter, S., Sultana, S., Mariani, M., Wamba, S. F., Spanaki, K., y Dwivedi, Y. K. (2023). *Advancing algorithmic bias management capabilities in AI-driven marketing analytics research*. Industrial Marketing Management, 114, 243–261. https://doi.org/10.1016/j.indmarman.2023.08.013

Barta, S., Gurrea, R., y Flavián, C. (2023a). *How augmented reality increases engagement through its impact on risk and the decision process*. *Cyberpsychology, Behavior, and Social Networking*. *26*(3), 177-187. https://doi.org/10.1089/cyber.2022.0087

Barta, S., Gurrea, R., y Flavián, C. (2023b). *Telepresence in live-stream shopping: An experimental study comparing Instagram and the metaverse*. *Electronic Markets, 33*, 29. https://doi.org/10.1007/s12525-023-00643-6

Barta, S., Gurrea, R., y Flavián, C. (2024a). *Augmented reality experiences: Consumer-centered augmented reality framework and research agenda*. *Psychology & Marketing*, 42(2), 634-650. https://doi.org/10.1002/mar.22143

Barta, S., Ibáñez-Sánchez, S., Orús, C., y Flavián, C. (2024b). *Avatar creation in the metaverse: A focus on event expectations*. *Computers in Human Behavior, 156*, 108192. https://doi.org/10.1016/j.chb.2024.108192

Belanche, D., Belk, R. W., Casaló, L. V., y Flavián, C. (2024). *The dark side of artificial intelligence in services*. *The Service Industries Journal, 44*(3–4), 149–172. https://doi.org/10.1080/02642069.2024.2305451

Belanche, D., Casaló, L. V., Flavián, C., y Schepers, J. (2020a). *Robots or frontline employees? Exploring customers' attributions of responsibility and stability after service failure or success*. *Journal of Service Management, 31*(2), 267–289. https://doi.org/10.1108/JOSM-05-2019-0156

Belanche, D., Casaló, L. V., Flavián, C., y Schepers, J. (2020b). *Service robot implementation: A theoretical framework and research agenda*. *The Service Industries Journal, 40*(3–4), 203–225. https://doi.org/10.1080/02642069.2019.1672666

Belanche, D., Casaló, L. V., Schepers, J. J. L., y Flavián, C. (2021a). *Examining the effects of robots' physical appearance, warmth, and competence in frontline services: The Humanness-Value-Loyalty model*. *Psychology & Marketing, 38*(12), 2357–2376. https://doi.org/10.1002/mar.21532

Belanche, D., Casaló, L. V., y Flavián, C. (2019). *Artificial intelligence in FinTech: Understanding robo-advisors adoption among customers*. *Industrial Management & Data Systems, 119*(7), 1411–1430. https://doi.org/10.1108/IMDS-08-2018-0368

Belanche, D., Casaló, L. V., y Flavián, C. (2021). Frontline robots in tourism and hospitality: Service enhancement or cost reduction? *Electronic Markets, 31*, 477–492. https://doi.org/10.1007/s12525-020-00432-5

Belanche, D., Casaló, L. V., y Flavián, C. (2021b). *Frontline robots in tourism and hospitality: Service enhancement or cost reduction? Electronic Markets, 31*, 477–492. https://doi.org/10.1007/s12525-020-00432-5

Belk, R. W., Belanche, D., & Flavián, C. (2023). Key concepts in artificial intelligence and technologies 4.0 in services: RW Belk et al. *Service Business*, *17*(1), 1-9. https://doi.org/10.1007/s11628-023-00528-w

Casaló, L. V., Millastre-Valencia, P., Belanche, D., y Flavián, C. (2025). *Intelligence and humanness as key drivers of service value in Generative AI chatbots*. International Journal of Hospitality Management, 128, 104130. https://doi.org/10.1016/j.ijhm.2025.104130

Dalenberg, D. J. (2018). Preventing discrimination in the automated targeting of job advertisements. *Computer Law & Security Review*, 34(3), 615–627. https://doi.org/10.1016/j.clsr.2017.11.009

Dietvorst, B. J., Simmons, J. P., y Massey, C. (2015). *Algorithm aversion: People erroneously avoid algorithms after seeing them err*. *Journal of Experimental Psychology: General, 144*(1), 114–126. https://doi.org/10.1037/xge0000033

Dwivedi, Y. K., Hughes, L., Wang, Y., Alalwan, A. A., Ahn, S. J., Balakrishnan, J., ... & Wirtz, J. (2023). Metaverse marketing: How the metaverse will shape the future of consumer research and practice. Psychology & Marketing, 40(4), 750-776. DOI: 10.1002/mar.21767

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). Opinion Paper:"So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International journal of information management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Epley, N., Waytz, A., y Cacioppo, J. T. (2007). *On seeing human: A three-factor theory of anthropomorphism*. *Psychological Review, 114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Flavián, C., Belk, R. W., Belanche, D., y Casaló, L. V. (2024a). *Automated social presence in AI: Avoiding consumer psychological tensions to improve service value*. *Journal of Business Research, 175*, 114545. https://doi.org/10.1016/j.jbusres.2024.114545

Flavián, C., Ibáñez-Sánchez, S., Orús, C., y Barta, S. (2023). *The dark side of the metaverse: The role of gamification in event virtualization*. *International Journal of Information Management*, 102726. https://doi.org/10.1016/j.ijinfomgt.2023.102726

Flavián, C., Ibáñez-Sánchez, S., y Orús, C. (2019). The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research, 100*, 547–560. https://doi.org/10.1016/j.jbusres.2018.10.050

Flavián, C., Ibáñez-Sánchez, S., y Orús, C. (2019a). *Integrating virtual reality devices into the body: Effects of technological embodiment on customer engagement and behavioral intentions toward the destination*. *Journal of Travel & Tourism Marketing, 36*(7), 847–863. https://doi.org/10.1080/10548408.2019.1618781

Flavián, C., Ibáñez-Sánchez, S., y Orús, C. (2019b). *The impact of virtual, augmented and mixed reality technologies on the customer experience*. *Journal of Business Research, 100*, 547–560. https://doi.org/10.1016/j.jbusres.2018.10.050

Flavián, C., Pérez-Rueda, A., Belanche, D., & Casaló, L. V. (2022). Intention to use analytical artificial intelligence (AI) in services–the effect of technology readiness and awareness. *Journal of Service Management*, *33*(2), 293-320. https://doi.org/10.1108/JOSM-10-2020-0378

Flavián, C., Tussyadiah, I. P., y Orús, C. (2024b). *How immersive technologies are redefining the customer journey*. *International Journal of Contemporary Hospitality Management, 36*(11), 3557–3565. https://doi.org/10.1108/IJCHM-11-2024-215

Flavián, C., y Casaló, L. V. (2021). *Artificial intelligence in services: Current trends, benefits and challenges*. *The Service Industries Journal, 41*(13–14), 853–859. https://doi.org/10.1080/02642069.2021.1989177

Floridi, L., y Sanders, J. W. (2004). *On the morality of artificial agents.* Minds and Machines, 14(3), 349–379. https://doi.org/10.1023/B:MIND.0000035461.63578.9d

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., y Pedreschi, D. (2019). *A survey of methods for explaining black box models*. *ACM Computing Surveys, 51*(5), Article 93, 1–42. https://doi.org/10.1145/3236009

Guy, M. G., Normand, J.-M., Jeunet-Kelway, C., y Moreau, G. (2023). The sense of embodiment in Virtual Reality and its assessment methods. *Frontiers in Virtual Reality, 4*, 1141683. https://doi.org/10.3389/frvir.2023.1141683

Huang, M.-H., y Rust, R. T. (2018). *Artificial intelligence in service*. *Journal of Service Research, 21*(2), 155–172. https://doi.org/10.1177/1094670517752459

Ibáñez-Sánchez, S., Orús, C., y Flavián, C. (2022). Augmented reality filters on social media: Analyzing the drivers of playability based on uses and gratifications theory. *Psychology & Marketing*. https://doi.org/10.1002/mar.21639

International Organization for Standardization. (2023). *ISO/IEC 23894:2023 Information technology: Artificial intelligence: Guidance on risk management*. ISO.

Kellogg, K. C., Valentine, M. A., y Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals, 14*(1), 366–410. https://doi.org/10.5465/annals.2018.0174

Kosinski, M., Stillwell, D., y Graepel, T. (2013). Private traits and attributes are predictable from digital records. *Proceedings of the National Academy of Sciences*, 110(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Kunz, W. H., Sajtos, L., y Flavián, C. (2025). Beyond replacement: Human–machine collaboration in the age of AI. *Journal of Service Management. 36 (4): 477–494.* https://doi.org/10.1108/JOSM-04-2025-0194

Lemon, K. N., y Verhoef, P. C. (2016). *Understanding customer experience throughout the customer journey*. *Journal of Marketing, 80*(6), 69–96. https://doi.org/10.1509/jm.15.0420

Lombard, M., & Ditton, T. (1997). *At the heart of it all: The concept of presence. Journal of Computer-Mediated Communication, 3(2).*
DOI: 10.1111/j.1083-6101.1997.tb00072.x

Madary, M., y Metzinger, T. K. (2016). *Real virtuality: A code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology*. *Frontiers in Robotics and AI, 3*, 3. https://doi.org/10.3389/frobt.2016.00003

Maxham III, J. G., y Netemeyer, R. G. (2002). A longitudinal study of complaining customers' evaluations of multiple service failures and recovery efforts. *Journal of Marketing, 66*(4), 57–71. https://doi.org/10.1509/jmkg.66.4.57.18512

Milgram, P., & Kishino, F. (1994). *A taxonomy of mixed reality visual displays. IEICE Transactions on Information and Systems, 77(12), 1321–1329.*

Ming-Hui Huang, Roland T. Rust, & Vojislav Maksimovic (2019). *The Feeling Economy: Managing in the Next Gel Intelligence (AI).* California Management Review, 61(4), 43–65. https://doi.org/10.1177/0008125619863436

Nass, C., y Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1)*. U.S. Department of Commerce.

OECD. (2019). *Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449)*. OECD Legal Instruments.

Oliver, R. L. (1980). *A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions.* Journal of Marketing Research, 17(4), 460–469. https://doi.org/10.1177/002224378001700405

Orús, C., Gurrea, R., y Flavián, C. (2017). Facilitating imaginations through online product presentation videos: Effects on imagery fluency, product attitude and purchase intention. *Electronic Commerce Research, 17*, 661–700. https://doi.org/10.1007/s10660-016-9250-7

Orús, C., Ibáñez-Sánchez, S., y Flavián, C. (2021). Enhancing the customer experience with virtual and augmented reality: The impact of content and device type. *International Journal of Hospitality Management, 98*, 103019. https://doi.org/10.1016/j.ijhm.2021.103019

Orús, C., Ibáñez-Sánchez, S., y Flavián, C. (2026). *The impact of creativity in social AR filters on brand awareness, image, and behavioral intentions: The role of intrusiveness and ad recognition*. *Technological Forecasting & Social Change, 225*, 124521. https://doi.org/10.1016/j.techfore.2025.124521

Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L., y Voss, C. A. (2015). *Service research priorities in a rapidly changing context*. *Journal of Service Research, 18*(2), 127–159. https://doi.org/10.1177/1094670515576315

Parasuraman, R., Sheridan, T. B., y Wickens, C. D. (2000). *A model for types and levels of human interaction with automation*. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, 30*(3), 286–297. https://doi.org/10.1109/3468.844354

Puntoni, S., Reczek Walker, R., Giesler, M., y Botti, S. (2021). *Consumers and Artificial Intelligence: An Experiential Perspective.* Journal of Marketing, 85(1), 131–151. DOI: 10.1177/0022242920953847

Rai, A. (2020). *Explainable AI: From black box to glass box*. *Journal of the Academy of Marketing Science, 48*(1), 137–141. https://doi.org/10.1007/s11747-019-00710-5

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). (2016). *Official Journal of the European Union*.

Ruusunen, N., Hallikainen, H., y Laukkanen, T. (2023). Does imagination compensate for the need for touch in 360-virtual shopping? *International Journal of Information Management, 70*, 102622. https://doi.org/10.1016/j.ijinfomgt.2023.102622

Sanchez-Vives, M. V., y Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience, 6*(4), 332–339. https://doi.org/10.1038/nrn1651

Schepers, J., Belanche, D., Casaló, L. V., y Flavián, C. (2022). *How smart should a service robot be? Journal of Service Research, 25*(4), 565–582. https://doi.org/10.1177/10946705221107704

Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human–Computer Interaction, 36*(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human-Computer Interaction, 36*(6), 495-504. https://doi.org/10.1080/10447318.2020.1741118

Slater, M., y Sanchez-Vives, M. V. (2016). *Enhancing our lives with immersive virtual reality*. *Frontiers in Robotics and AI, 3*, 74. https://doi.org/10.3389/frobt.2016.00074

Tax, S. S., Brown, S. W., y Chandrashekaran, M. (1998). *Customer evaluations of service complaint experiences: Implications for relationship marketing*. *Journal of Marketing, 62*(2), 60–76. https://doi.org/10.1177/002224299806200205

van Doorn, J., Mende, M., Noble, S. M., et al. (2017). *Domo arigato Mr. Roboto: Emergence of automated social presence in organizational frontlines and customers' service experiences.* Journal of Service Research, 20(1), 43–58. DOI: 10.1177/1094670516679272

van Doorn, J., Mende, M., Noble, S. M., Hulland, J., Ostrom, A. L., Grewal, D., y Petersen, J. A. (2017). *Domo arigato Mr. Roboto: Emergence of automated social presence in organizational frontlines and customers' service experiences*. *Journal of Service Research, 20*(1), 43–58. https://doi.org/10.1177/1094670516679272

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., y Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management, 29*(5), 907–931. https://doi.org/10.1108/JOSM-04-2018-0119

Witmer, B. G., y Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225–240. https://doi.org/10.1162/105474698565686

World Wide Web Consortium (W3C). (2023). *Web Content Accessibility Guidelines (WCAG) 2.2.* W3C Recommendation. https://www.w3.org/TR/WCAG22/

Yang, M., Guo, J., Zhu, L., Zhu, H., Song, X., Zhang, H., y Xu, T. (2024). Fairness evaluation of marketing algorithms: A framework for equity distribution. *Journal of Electronic Business & Digital Economics*, 3(3), 251–274. https://doi.org/10.1108/JEBDE-10-2023-0024