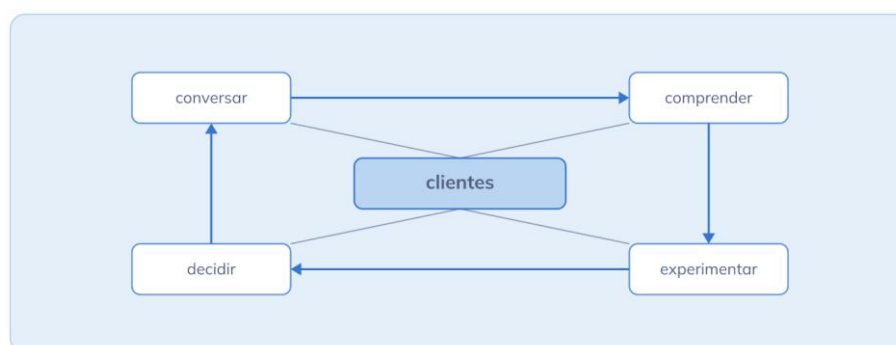


State of the art

Descubrimiento de producto y validación con clientes

Mapa de referencia del conocimiento profesional vigente



Actualizado a septiembre de 2026

Sobre este documento

Este documento es un mapa curado de los conocimientos, prácticas y modelos empleados profesionalmente a la fecha de su publicación para practicar el descubrimiento de producto (*product discovery*) y validar con clientes qué merece construirse antes de construirlo, reduciendo la incertidumbre con evidencia útil de clientes y mercado. Para cada uno, este mapa lo sitúa, indica su grado de adopción en la práctica profesional actual y orienta sobre dónde encontrar información de referencia.

Úsalo como observatorio y punto de referencia para contrastar si el conocimiento profesional que empleas está alineado con la práctica actual y en vanguardia.

Se complementa con un documento de desarrollo que profundiza en cada concepto y con la plataforma de entrenamiento y evaluación en Skill Arena. En el área [Descubrimiento de producto y validación con clientes](#) puedes contrastar tu nivel de conocimiento y, si lo superas, obtener un diploma que acredita curricularmente la solvencia y vanguardia profesional en esta área.



Estado del conocimiento

El conocimiento sobre el descubrimiento de producto evoluciona de forma progresiva: su núcleo lleva años probado y cambia más por refinamiento de la práctica que por rupturas. Lo que se mueve deprisa es su instrumental. La inteligencia artificial ya transcribe, sintetiza y hasta entrevista, y con ella han llegado prácticas nuevas, evidencia todavía escasa y obligaciones legales que conviene conocer antes de usarla. En los distintos apartados del documento, las etiquetas (ESTABLECIDO, EN CONSOLIDACIÓN, EMERGENTE) ayudan a identificar la madurez de cada concepto:

ESTABLECIDO consenso asentado; conocimiento que se da por necesario.

EN CONSOLIDACIÓN gana adopción con rapidez; aún no universal pero ya relevante.

EMERGENTE frontera reciente; alta relevancia y alta volatilidad.

Bloque A — Conversar y capturar señal

El punto de partida del descubrimiento: separar el problema de la solución, escuchar a las personas adecuadas para aprender y no para confirmar, y dejar un registro fiel de lo que dijeron, con el cuidado que exigen sus datos. Sin buena señal de entrada, todo lo que viene después se contamina.

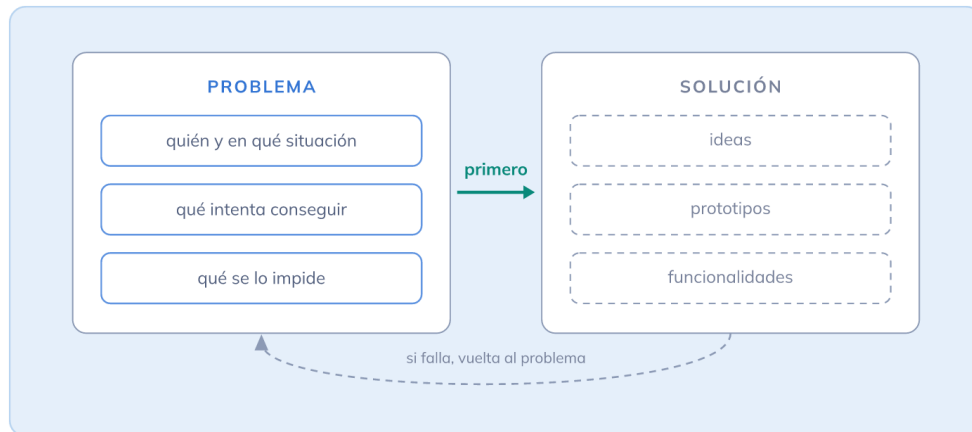


Figura 1. Dos espacios que no se mezclan: el problema se entiende antes de comprometerse con una solución, y se vuelve a él cada vez que una solución falla.

Espacio del problema y espacio de la solución · ESTABLECIDO · nuevo desde 2026

Descubrir exige mantener separados dos espacios. El del problema recoge a quién le pasa algo, en qué situación, qué intenta conseguir y qué se lo impide; el de la solución, las ideas, los prototipos y las funcionalidades con que podríamos responder. Mezclarlos pronto hace que la conversación gire alrededor de una idea propia y que la evidencia se busque para defenderla. La disciplina consiste en entender el problema con precisión antes de comprometerse con una solución, y en volver a él cada vez que una solución falla.

Por qué está aquí ahora. Es la primera capacidad que acredita esta *skill* y el origen de casi todos los errores posteriores, porque una petición como «necesitan un panel de control» llega ya con la solución dentro. Dan Olsen hizo explícita la distinción en *The Lean Product Playbook*, y el descubrimiento continuo de Teresa Torres la recorre de principio a fin. Es un principio establecido que ninguna práctica del área discute; lo que falla es sostenerlo cuando el equipo ya se ha encariñado con una idea.

Dónde mirar. [Scrum Manager BoK — Product discovery](#) · [Dan Olsen — The Lean Product Playbook](#)

La entrevista que aprende, no la que confirma · ESTABLECIDO

Una entrevista de descubrimiento explora el comportamiento pasado y el contexto real de la persona; no es una encuesta oral ni una venta encubierta. Antes de redactar el guion conviene poder completar la frase «queremos aprender X y lo sabremos si observamos Y». Las preguntas útiles piden historias concretas del pasado («cuéntame la

última vez que...»), que es la base de la entrevista basada en historias (*story-based interviewing*) de Teresa Torres; las inútiles piden predicciones sobre el futuro u opiniones sobre nuestra idea.

Por qué está aquí ahora. Es el cimiento de la disciplina y su lección mejor documentada: el sesgo de confirmación, diseñar preguntas que validen una solución ya imaginada, sigue siendo el fallo más común. La técnica no ha cambiado, pero su contexto sí. La entrevista moderada por IA ya se vende como producto, y el Nielsen Norman Group concluyó en enero de 2026 que sirve para recoger opiniones sobre un producto o para cribar participantes, pero no para explorar un problema desconocido, que es precisamente lo que pide esta entrevista (bloque F).

Dónde mirar. [Rob Fitzpatrick — The Mom Test](#) · [Teresa Torres — Customer interviews](#)

A quién escuchar: segmento, muestra y reclutamiento · ESTABLECIDO

Lo que se aprende depende de a quién se escucha. Antes de reclutar se fija el segmento que interesa, con criterios de inclusión y de exclusión que se comprueban con un cuestionario de selección (*screener*) redactado sin delatar la respuesta que se busca. Los participantes deben ser externos y representativos: la red propia, los compañeros y los clientes más entusiastas tienen incentivos para agradar. Y la muestra se dimensiona por el propósito, porque unas pocas entrevistas bastan para descubrir patrones, pero no para medir cuántos clientes los comparten.

Por qué está aquí ahora. Reclutar mal invalida la señal de raíz y ninguna técnica posterior lo corrige: el sesgo de selección aparece igual en la entrevista, en el test de prototipo y en la validación de la demanda. Es una decisión previa a cualquier conversación, y cada vez la toman más personas sin formación específica en investigación. Según Maze, en 2026 el 39 % de los encuestados afirma que en su organización los responsables de producto ya investigan con usuarios.

Dónde mirar. [Nielsen Norman Group — Screening questions](#) · [Nielsen Norman Group — Interviewing users](#)

Notas de calidad: separar hecho, cita e interpretación · ESTABLECIDO

Durante la entrevista y justo después se separa lo que la persona dijo (la cita), lo que hizo o contó que hizo (el hecho, con su contexto de quién, cuándo y en qué situación), lo que interpreta quien entrevista, marcado como tal, y lo que queda pendiente de explorar. Cuando la conversación se graba y se transcribe, la cita literal deja de ser escasa y el valor de las notas pasa al contexto y a la interpretación marcada. Si las notas suenan a adjetivos («le encantó», «está contento»), falta evidencia concreta.

Por qué está aquí ahora. La síntesis solo es tan buena como el material que la alimenta. Teresa Torres dio forma a la práctica de cerrar las notas en caliente con el resumen de entrevista (*interview snapshot*): una página con una cita memorable, las oportunidades, los hallazgos y un mapa de la experiencia, que se rellena en los 15 a 25 minutos que siguen a la conversación. La transcripción automática lo facilita, siempre que el resumen

que genera la herramienta se trate como una interpretación más y se contraste con la transcripción.

Dónde mirar. [Teresa Torres — Interview snapshots](#) · [Scrum Manager BoK — Interview snapshot](#)

Consentimiento, datos y transparencia al conversar con clientes · EN **CONSOLIDACIÓN** · nuevo desde 2026

Grabar, transcribir y analizar una entrevista es tratar datos personales, porque la voz y lo que se cuenta pueden identificar a la persona. Hay que informar antes de grabar y mientras se graba, pedir un consentimiento específico para cada grabación y comprobar qué hace la herramienta con los datos: si los usa para reentrenar su modelo, eso es otro tratamiento con su propia base jurídica. Las notas con datos identificables no se pegan en un asistente de IA sin anonimizarlas. Y si quien entrevista es una IA, la persona debe saberlo.

Por qué está aquí ahora. Lo que era buena costumbre tiene desde 2026 criterios de autoridad. La Agencia Española de Protección de Datos publicó en enero y en abril sus criterios sobre la transcripción de voz con IA: consentimiento por grabación, indicador visible mientras se graba, exactitud y rectificación de la transcripción, y el aviso de que la detección de emociones puede caer en usos que el Reglamento europeo de IA restringe o prohíbe. Desde el 2 de agosto de 2026, ese Reglamento exige además que los sistemas que interactúan con personas les informen de que son una IA. Está en consolidación porque la práctica profesional alrededor de estos criterios aún se está formando.

Dónde mirar. [AEPD — Transcripción de voz con IA \(I\)](#) · [AEPD — Transcripción de voz con IA \(II\)](#) · [Reglamento de IA — artículo 50](#)

Bloque B — De los datos a la comprensión

Convertir conversaciones sueltas en aprendizaje accionable: detectar patrones que sostengan un insight y enmarcar el problema sin esconder dentro la solución. Aquí está buena parte del valor diferencial del profesional.

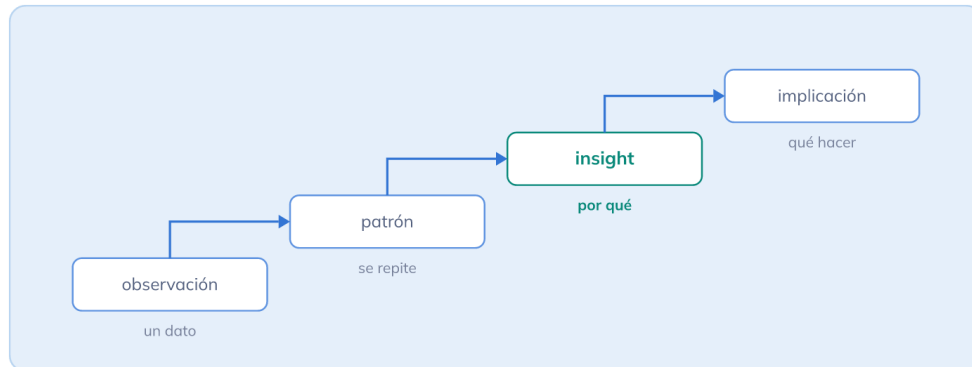


Figura 2. De la observación a la implicación: la síntesis asciende por niveles y el insight es el que conecta el patrón con una causa.

Síntesis: de la observación al insight accionable · ESTABLECIDO

El análisis cualitativo asciende por cuatro niveles: la observación (un dato de una fuente), el patrón (lo que se repite en varias fuentes), el insight (la comprensión que conecta el patrón con una causa) y la implicación (lo que conviene hacer en consecuencia). Un insight responde a qué hemos aprendido que no sabíamos y qué significa para el producto, y se apoya en la recurrencia entre varias fuentes, nunca en una cita aislada. Las excepciones se tratan como posible señal de otro segmento, y no como ruido.

Por qué está aquí ahora. Confundir una cita interesante con un insight es el error que más decisiones desvía, y la IA le ha dado una forma nueva. Teresa Torres advierte contra la síntesis perezosa con IA (*lazy AI interview synthesis*), que consiste en volcar las transcripciones en un asistente y aceptar su resumen, y propone sintetizar primero, pedir después a la IA la misma síntesis y comparar las dos (bloque F). El ritual de síntesis en grupo tiene skill propia en Scrum Manager, Design Thinking; aquí interesa su vertiente analítica.

Dónde mirar. [Nielsen Norman Group — Affinity diagramming](#) · [Skill Arena — Design Thinking](#)

Enmarcar el problema: el trabajo que el cliente quiere hacer · ESTABLECIDO

Definir el problema como «falta la funcionalidad X» mete la solución dentro del problema y cierra la exploración. El marco de los *jobs to be done* (JTBD) describe en su lugar el progreso que busca la persona en una situación concreta, con independencia de la solución. Su forma más usada en producto es la *job story*, que nació en Intercom con Paul Adams y bautizó Alan Klement: «cuando [situación], quiero [motivación], para [resultado]». Conviene completarla con la alternativa con que la persona lo resuelve hoy, que marca el listón que hay que superar.

Por qué está aquí ahora. Es el puente entre la evidencia y las decisiones de producto, y la defensa más eficaz contra la fábrica de funcionalidades (*feature factory*). El marco tiene dos escuelas que conviene distinguir: la de Clayton Christensen y Bob Moesta, que entiende el trabajo como el progreso que la persona intenta hacer en una circunstancia, y la de Tony Ulwick, cuya innovación guiada por resultados (*Outcome-Driven Innovation*, ODI) formula el trabajo como un verbo con su objeto y mide los resultados que el cliente desea. La plantilla de la job story pertenece a la primera tradición.

Dónde mirar. [Intercom — Job stories](#) · [Christensen Institute — Jobs to Be Done](#) · [Strategyn — Outcome-Driven Innovation](#)

Bloque C — Diseñar experimentos para aprender

Hacer explícitos los supuestos arriesgados y diseñar la prueba mínima que los ponga a prueba: del mapeo de supuestos al test de prototipo y al producto mínimo viable (minimum viable product, MVP) entendido como experimento. Es la zona donde más se confunde actividad con aprendizaje.

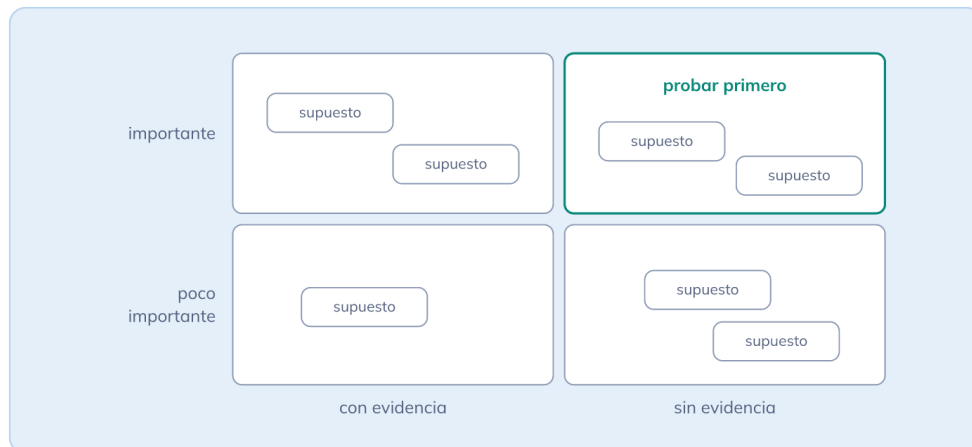


Figura 3. El mapa de supuestos: se prueba primero lo que es importante y todavía no tiene evidencia.

Mapeo de supuestos: los riesgos que hay que probar primero

ESTABLECIDO

Toda idea descansa sobre supuestos que pueden resultar falsos. Los tres tipos clásicos son la deseabilidad (¿lo quieren?), la viabilidad (¿es un negocio?) y la factibilidad (¿podemos construirlo?), y las fuentes de referencia añaden otros, como la usabilidad y la ética en Teresa Torres o la adaptabilidad en David Bland. Se priorizan en la matriz de importancia y evidencia del mapeo de supuestos (*Assumptions Mapping*) de Bland, que manda probar primero lo que es importante y todavía carece de evidencia, aunque sea incómodo de probar.

Por qué está aquí ahora. El patrón de fallo habitual es probar lo fácil, casi siempre lo técnico, y dejar sin explorar lo crítico. Bland atribuye el ejercicio a su trabajo con Jeff Gothelf y Josh Seiden, autores de *Lean UX*, y lo difundió en *Testing Business Ideas*, con Alexander Osterwalder. Empezar por la deseabilidad es el orden habitual, pero manda el cuadrante: en un producto cuyo valor depende de un modelo de IA, la factibilidad puede ser el supuesto que lo decide todo (bloque F).

Dónde mirar. [David Bland — Assumptions Mapping](#) · [Teresa Torres — Assumption testing](#) · [Marty Cagan — The four big risks](#)

Tests de prototipo: tareas como objetivos, no instrucciones

CONSOLIDACIÓN

Un test de prototipo solo da señal de usabilidad si las tareas se plantean como objetivos («crea un proyecto llamado Prueba Q1») y no como instrucciones paso a paso, si modera alguien distinto de quien diseñó el prototipo y si los participantes son externos y representativos. Lo que cuenta es lo que la persona hace (dónde pulsa, dónde se atasca,

qué no encuentra), no lo que opina. La fidelidad se ajusta a la pregunta de investigación, y basta la mínima que la responda.

Por qué está aquí ahora. La teoría es estable pero su ejecución disciplinada no es universal, y 2026 añade dos tentaciones. Las herramientas de IA generan prototipos de alta fidelidad en minutos, y un prototipo que parece terminado invita a opinar sobre el acabado y a encariñarse con él. Además llegan agentes y moderadores automáticos que prometen evaluar la usabilidad; la investigación que los entrena avanza, pero hoy no sustituyen la observación de personas reales del segmento.

Dónde mirar. [Steve Krug — Rocket Surgery Made Easy](#) · [Nielsen Norman Group — Usability testing 101](#) · [Scrum Manager BoK — Think-aloud protocol](#)

El MVP como experimento, no como producto reducido · EN CONSOLIDACIÓN

Un MVP es lo mínimo necesario para aprender algo concreto, no una versión pequeña de todo. Antes de construir se consideran alternativas más ligeras, elegidas según la hipótesis: el servicio manual a la vista del cliente (*concierge*), la operación humana oculta tras una interfaz aparente (Mago de Oz), la página de captación (*landing page*) o una sola funcionalidad (*single feature MVP*). Que el servicio manual no escale es irrelevante al principio, porque si nadie lo quiere no hay nada que escalar.

Por qué está aquí ahora. El malentendido del MVP como lista de componentes sigue extendido, y la IA lo refuerza. Marty Cagan distingue construir para aprender de construir para ganar (abril de 2026) y observa que, con las herramientas actuales, diez o veinte prototipos a la semana están al alcance de casi cualquiera, así que el cuello de botella ya no es construir. En julio citaba a Hilary Gridley: como nunca ha sido tan rápido construir, nunca ha sido tan fácil correr diez veces más deprisa en la dirección equivocada. Construir barato no exime de preguntar qué es lo mínimo para aprender.

Dónde mirar. [Marty Cagan — Build to learn vs build to earn](#) · [Bland y Osterwalder — Testing Business Ideas](#) · [Scrum Manager BoK — Mínimo Producto Viable](#)

Bloque D — Decidir con evidencia

Cerrar el bucle: fijar el criterio de éxito antes de mirar los datos, ponderar la fuerza de la evidencia frente a los sesgos y decidir entre perseverar, iterar, pivotar o parar. Es donde el descubrimiento se traduce en decisiones que el equipo defiende.

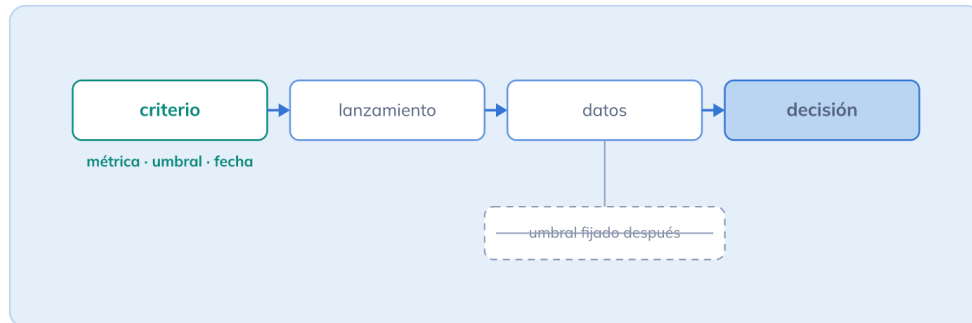


Figura 4. El criterio se fija antes de lanzar: el umbral de éxito queda escrito antes de que lleguen los datos.

Criterios de éxito y señales definidos de antemano · ESTABLECIDO

El criterio de éxito se fija antes de lanzar el experimento: una sola métrica primaria con su umbral, métricas secundarias de contexto y límites que no deben cruzarse aunque la primaria salga bien (*guardrail metrics*). Decidir después de ver los datos qué resultado parece bueno, o formular la hipótesis a la vista de los resultados, es lo que la estadística llama *p-hacking* y *HARKing*. Y hay que reconocer lo que el tamaño de la muestra permite concluir, porque sin grupo de control la evidencia es un indicio y no una relación causal.

Por qué está aquí ahora. Sin criterio previo, casi cualquier resultado se racionaliza como éxito. Teresa Torres lo resume en trazar una línea en la arena antes de probar un supuesto. El matiz que conviene no perder es de escala: un test A/B con grupo de control fiable exige un tráfico que la mayoría de los descubrimientos tempranos no tiene, como explican Ron Kohavi y sus coautores en *Trustworthy Online Controlled Experiments*, y en ese caso lo honesto es declarar la evidencia como direccional.

Dónde mirar. [Teresa Torres — Assumption testing](#) · [Kohavi, Tang y Xu — Trustworthy Online Controlled Experiments](#)

La fuerza de la evidencia y el control de sesgos · ESTABLECIDO

La evidencia es más fuerte cuanto más se parece lo observado a una decisión real con coste para quien la toma: una opinión pesa menos que un registro, un registro menos que una reserva y una reserva menos que un pago. Cuatro sesgos frecuentes la erosionan (confirmación, supervivencia, selección y deseabilidad social), y las señales de bajo coste, como los me gusta, los comentarios de la red propia o las aperturas de un correo, se confunden a menudo con validación.

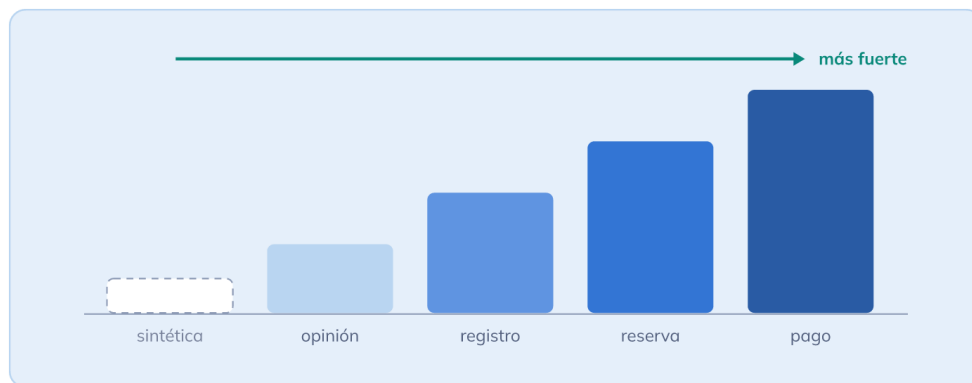


Figura 5. La fuerza de la evidencia: la evidencia pesa más cuanto más se parece lo observado a una decisión real con coste para quien la toma.

Por qué está aquí ahora. Es el filtro que separa el descubrimiento riguroso del teatro de validación. Alexander Osterwalder lo formuló como un gradiente: la evidencia se fortalece a medida que se acerca a una compra real, porque lo que las personas dicen y lo que hacen son cosas distintas. En 2026 ha aparecido una fuente nueva de falso positivo, las respuestas de usuarios simulados por IA, que ocupan el peldaño más bajo de esta escala (bloque F).

Dónde mirar. [Alexander Osterwalder — How strong is your innovation evidence?](#) · [Rob Fitzpatrick — The Mom Test](#) · [Scrum Manager BoK — Sesgo cognitivo](#)

Perseverar, iterar, pivotar o parar · EN CONSOLIDACIÓN

Ante la evidencia, las opciones son cuatro: perseverar en la dirección actual, iterar ajustando elementos concretos, pivotar cambiando un aspecto fundamental (el segmento, la propuesta de valor, el canal) sin perder lo aprendido, o parar. Cambiar de segmento es uno de los tipos de pivote que describió Eric Ries en *The Lean Startup*. La falacia del coste hundido, que justifica continuar por lo ya invertido, es el sesgo más peligroso; lo que cuenta es la mejor inversión de los próximos recursos.

Por qué está aquí ahora. Conecta el descubrimiento con la decisión real de producto y con el descubrimiento continuo, donde decidir es un acto recurrente y no un hito final. Está en consolidación porque el espectro de opciones está aceptado, pero su uso explícito todavía madura: documentar lo aprendido antes de cambiar de rumbo y ponderar la evidencia mixta por su calidad, su magnitud, el coste de oportunidad y la reversibilidad de la decisión.

Dónde mirar. [Eric Ries — The Lean Startup](#) · [Teresa Torres — Product discovery basics](#)

Bloque E — El descubrimiento como hábito del equipo

Descubrir funciona mejor como hábito semanal del equipo que decide qué construir que como fase previa al desarrollo. Este bloque recoge cómo se organiza ese hábito y la herramienta que conecta lo que se aprende con el resultado que se persigue.

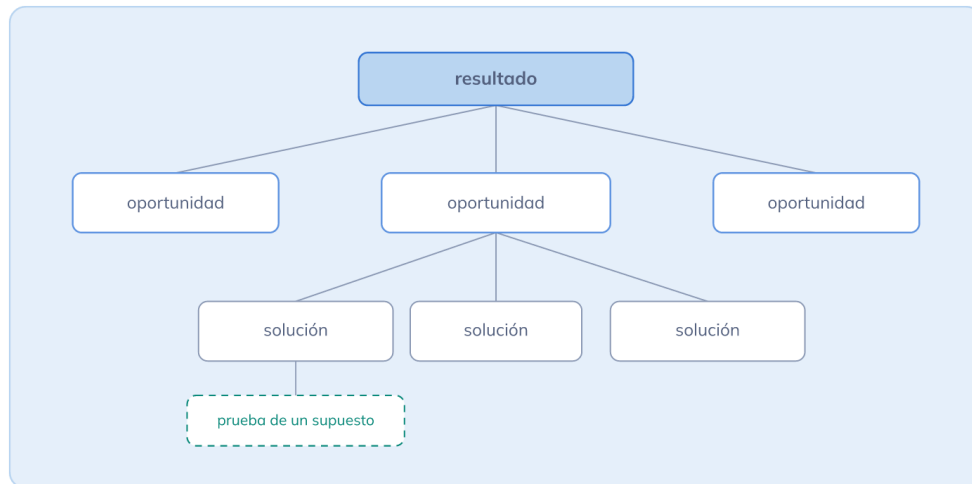


Figura 6. El árbol de oportunidades: el resultado se descompone en oportunidades, cada oportunidad en soluciones y cada solución en pruebas de sus supuestos.

Descubrir cada semana: el trío de producto y el contacto continuo · EN CONSOLIDACIÓN · nuevo desde 2026

El descubrimiento continuo (*continuous discovery*) es, en la definición de Teresa Torres, el contacto al menos semanal con clientes del equipo que construye el producto, mediante actividades pequeñas de investigación orientadas a un resultado (*outcome*). Lo sostiene el trío de producto (*product trio*): quien lleva el producto, quien diseña y quien desarrolla, que descubren juntos en lugar de pasarse un informe. Descubrir y entregar corren en paralelo, lo que el campo llamó doble vía (*dual-track*), y lo aprendido se guarda en un repositorio al que todo el equipo tiene acceso.

Por qué está aquí ahora. Es la práctica canónica del área desde *Continuous Discovery Habits* (2021) y sigue siendo minoritaria: en la encuesta de ProductPlan de finales de 2025, solo el 34 % de los equipos recoge conocimiento de clientes con regularidad y lo usa para priorizar, frente a casi un 20 % que vive de peticiones sueltas y un 12 % sin ningún proceso. El qué del descubrimiento continuo lo enseña la certificación Product Owner de Scrum Manager; aquí interesa cómo se sostiene la cadencia, quién entrevista, cómo se recluta cada semana y cómo se conserva la evidencia.

Dónde mirar. [Teresa Torres — Continuous Discovery Habits](#) · [Scrum Manager BoK — Dual-track discovery & delivery](#) · [Skill Arena — Product Owner](#)

El árbol de oportunidades: del resultado al experimento · EN CONSOLIDACIÓN

· nuevo desde 2026

El árbol de oportunidades y soluciones (*opportunity solution tree*) de Teresa Torres ordena el descubrimiento en cuatro niveles: arriba, el resultado que el equipo persigue; debajo, las oportunidades (necesidades, fricciones y deseos que salen de las entrevistas); después, las soluciones posibles para cada oportunidad, y al final, las pruebas de los supuestos de cada solución. Obliga a elegir una oportunidad antes de idear, a comparar varias soluciones para ella y a probar supuestos en lugar de funcionalidades enteras.

Por qué está aquí ahora. Es la herramienta que conecta lo que se aprende en las entrevistas con el resultado de negocio e impide que cada conversación dispare una funcionalidad. En 2026 han aparecido borradores del árbol generados con IA a partir de entrevistas reales. Torres los defiende como punto de partida que el equipo revisa y advierte contra los árboles de un clic hechos con datos inventados, porque el sentido del árbol está en sintetizar lo que se aprende de clientes reales. El resultado de la raíz se define con objetivos y resultados clave, que tienen skill propia.

Dónde mirar. [Teresa Torres — Opportunity solution trees](#) · [Scrum Manager BoK — Árbol de soluciones de oportunidades](#) · [Skill Arena — OKRs y estrategia de producto](#)

Bloque F — Descubrir con inteligencia artificial

La inteligencia artificial transcribe, resume, entrevista y hasta simula clientes. Este bloque la trata como fuente y como filtro de evidencia: dónde acelera el descubrimiento, dónde lo erosiona y cómo se valida un producto cuyo valor depende de un modelo.

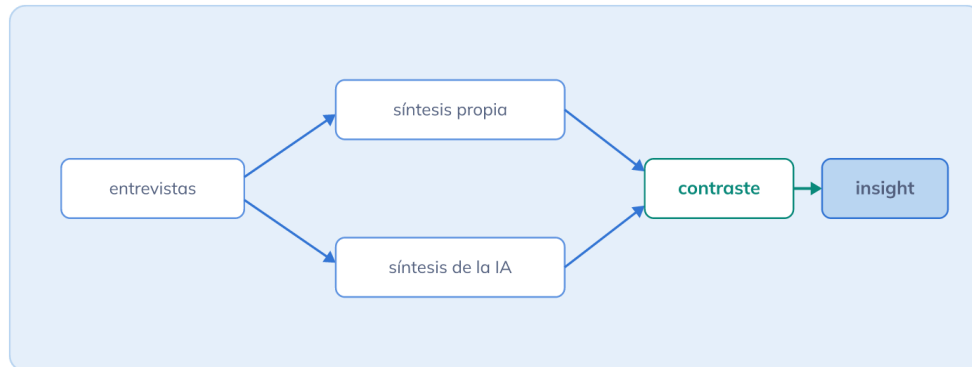


Figura 7. Comparar, no delegar: la síntesis propia y la de la IA se hacen por separado y se contrastan antes de afirmar un insight.

Entrevistar y sintetizar con IA: dónde ayuda y dónde erosiona · EN CONSOLIDACIÓN · nuevo desde 2026

En el trabajo de descubrimiento, la IA ya transcribe, resume, agrupa temas y propone borradores de síntesis, y hay plataformas que moderan entrevistas sin entrevistador humano. Su uso sensato sigue dos reglas. En la síntesis se compara y no se delega: primero la síntesis propia, después la de la IA y al final el contraste entre ambas. En la entrevista, la moderación automática sirve para opiniones sobre un producto, cribado y estudios en varios idiomas, pero no para explorar un problema desconocido.

Por qué está aquí ahora. La adopción está medida. Según Maze, en 2026 el 69 % de los profesionales de la investigación usa IA en al menos parte de sus estudios, y el 82 % sostiene que interpretar el matiz y la emoción sigue siendo humano. ProductPlan encuentra el mismo patrón: la IA se usa sobre todo para resumir el *feedback* de clientes (47 %) y muy poco para priorizar problemas (14 %). El Nielsen Norman Group probó dos plataformas de entrevista con IA y solo 3 de 10 participantes encontraron natural la conversación. Está en consolidación porque el criterio de uso ya tiene fuentes, aunque las herramientas cambian cada pocos meses.

Dónde mirar. [Nielsen Norman Group — AI-moderated interviews](#) · [Teresa Torres en airfocus — AI product management](#) · [Maze — Future of User Research Report 2026](#)

Lo sintético en la escala de la evidencia · EMERGENTE · nuevo desde 2026

Los usuarios sintéticos son perfiles generados por IA que responden como lo haría un segmento de clientes. Sirven para preparar el trabajo real (investigación documental, hipótesis, borradores de guion), pero sus respuestas no son evidencia sobre clientes: son opinión sin coste, complaciente y demasiado coherente, que capta las tendencias

generales y falla en su magnitud y su variabilidad. En la escala de la fuerza de la evidencia ocupan el peldaño más bajo.

Por qué está aquí ahora. La herramienta se ofrece con insistencia y la evidencia que se acumula es negativa. Un estudio publicado en arXiv en julio de 2026 encontró que los modelos inflaban entre dos y cuatro veces las diferencias entre segmentos y llevaban a elegir el segmento equivocado en la mitad de los casos estadounidenses; otro del mismo mes, con 37 modelos, midió un sesgo de aquiescencia de 0,84 desviaciones típicas. El 79 % de los investigadores teme que las partes interesadas (*stakeholders*) confíen de más en estos hallazgos. Es emergente porque la práctica está en disputa y su permanencia no está garantizada.

Dónde mirar. [Nielsen Norman Group — Synthetic users](#) · [Chen, Zhu y Zheng — When synthetic users fail \(arXiv, 2026\)](#) · [User Interviews — State of Synthetic Users](#)

Validar un producto cuyo núcleo es un modelo · EMERGENTE · nuevo desde 2026

Cuando el valor del producto depende de un modelo de IA, a las preguntas habituales se suma otra: si aporta valor con una fiabilidad aceptable para quien lo usa. El descubrimiento se amplía con el análisis de errores, que consiste en observar con clientes reales dónde falla el sistema, convertir esos fallos en criterios de evaluación (*evals*) y aceptar que los criterios cambian a medida que se aprende. La factibilidad deja de ser una cuestión técnica que se resuelve al final y pasa a ser un supuesto que se prueba pronto.

Por qué está aquí ahora. Teresa Torres y Petra Wille lo formularon en septiembre de 2025: el análisis de errores funciona como actividad de descubrimiento, y Torres ha llegado a llamar a los evals un nuevo hábito del descubrimiento. Es emergente porque la práctica está todavía en formación. Aquí interesa solo el tramo de validar el valor con una fiabilidad aceptable; la operación de los evals antes y después del lanzamiento la tratan la certificación Product Owner y la skill Harness Engineering.

Dónde mirar. [Torres y Wille — AI evals & discovery](#) · [Skill Arena — Harness Engineering](#) · [Skill Arena — Product Owner](#)

Nota sobre las referencias

*Los enlaces incluidos estaban disponibles y verificados en la fecha de actualización.
Dado el ritmo de cambio del área, algunos pueden modificarse; cada revisión del mapa actualiza también sus referencias.*

© 2026 Scrum Manager®. Esta obra se publica bajo licencia Creative Commons Atribución – No Comercial 4.0 Internacional (CC BY-NC 4.0). Los formadores y centros oficiales de Scrum Manager quedan licenciados bajo los términos CC BY 4.0 para su actividad formativa.