Informe “El Reto” enerTIC

Comité de Expertos de Innovación para la eficiencia en Infraestructuras IT

El regreso del centro de datos corporativo: IA soberana y control del dato



Septiembre 2026

Sumario

Resumen ejecutivo..... 3

Reconocimientos..... 4

Identificación del reto por el Comité de Expertos..... 5

DECALOGO

1. Soberanía del dato, del modelo y del conocimiento..... 6

2. Una infraestructura para cada carga: CPD propio, colocation, edge y nube híbrida 9

3. IA privada y factorías corporativas de conocimiento..... 13

4. Alta densidad y refrigeración híbrida: rediseñar la infraestructura física.... 16

5. Energía, flexibilidad y resiliencia: el CPD como aliado del sistema eléctrico.. 20

6. Nuevas arquitecturas de computación, red y almacenamiento..... 24

7. IA eficiente y sostenible: utilizar el modelo adecuado para cada necesidad..... 28

8. Talento, organización y capacitación para una transformación permanente..... 32

9. Del CPD monitorizado al CPD autónomo y gobernado..... 36

10.Transformar el parque instalado sin detener el negocio..... 40

Resultados esperados..... 46

Anexo: Marcos normativos, certificaciones y *frameworks* de referencia..... 50

Informes El Reto 2026..... 51

Acerca de enerTIC..... 54

Resumen ejecutivo

La expansión de la inteligencia artificial, el crecimiento del dato y las nuevas exigencias de soberanía, eficiencia y resiliencia están devolviendo al centro de datos corporativo una función estratégica. Este cambio no implica renunciar a la nube ni concentrar todas las cargas en infraestructura propia, sino recuperar la capacidad de decidir **dónde se procesan los datos, qué modelos se utilizan, cómo se protege el conocimiento y de qué proveedores depende la organización**.

El informe identifica una prioridad central: evolucionar desde una gestión fragmentada de infraestructuras y servicios hacia un **modelo híbrido, gobernado y permanentemente adaptable**. Esto exige coordinar arquitectura tecnológica, capacidad eléctrica y térmica, operación, talento e inversión, evitando tanto la dependencia excesiva de terceros como la sustitución constante de activos que todavía aportan valor.

El decálogo plantea cinco prioridades de actuación:

- **Recuperar soberanía y capacidad de decisión**. Las organizaciones deben conocer dónde residen sus datos, modelos y conocimientos críticos, controlar sus dependencias y asignar cada carga al entorno más adecuado entre CPD propio, *colocation*, *edge* y nube híbrida.
- **Construir capacidades corporativas de inteligencia artificial**. La IA privada y las factorías de conocimiento permiten aprovechar información sensible y experiencia interna con mayor control, siempre que se apoyen en datos fiables, arquitecturas gobernadas y casos de uso vinculados a resultados concretos.
- **Rediseñar la infraestructura para las nuevas cargas**. La alta densidad, los aceleradores y el crecimiento del tráfico interno obligan a coordinar computación, red, almacenamiento, energía y refrigeración. Las soluciones híbridas y modulares permiten adaptar capacidad sin transformar toda la instalación al mismo tiempo.
- **Integrar eficiencia, flexibilidad y resiliencia energética**. El CPD debe gestionar la energía como una restricción estratégica y, cuando el contexto lo permita, como una capacidad flexible al servicio del sistema eléctrico. La eficiencia debe medirse por trabajo útil y extenderse también a la selección, entrenamiento y operación de los modelos de IA.
- **Gobernar una transformación permanente**. El talento, la automatización supervisada, la observabilidad y la modernización continua del parque instalado son capacidades inseparables. La autonomía debe avanzar de forma gradual, con límites, trazabilidad, reversión y autoridad humana sobre las decisiones críticas.

Estas prioridades deben aplicarse según la madurez, la criticidad y las necesidades reales de cada organización. No todas requieren una factoría de IA propia, refrigeración líquida o automatización avanzada. En algunos casos, la prioridad será completar inventarios, clasificar cargas, medir consumos o documentar dependencias; en otros, desplegar zonas de alta densidad, incorporar mantenimiento predictivo o automatizar actuaciones reversibles.

El criterio para priorizar es claro: **comenzar por el servicio o la carga, definir sus requisitos de soberanía, rendimiento, continuidad y eficiencia, y seleccionar después la infraestructura y el grado de automatización adecuados**. Cada actuación debe contar con una línea de base, un responsable, límites operativos, métricas de resultado, capacidad de reversión y una visión completa de su coste durante el ciclo de vida.

Este enfoque permite convertir el CPD en una plataforma evolutiva, capaz de combinar activos propios, servicios externos y nuevas capacidades sin comprometer la operación. La finalidad no es disponer de la infraestructura más avanzada ni automatizar el mayor número de decisiones, sino construir una capacidad digital que refuerce la **soberanía, la eficiencia, la competitividad, la resiliencia y la sostenibilidad de la organización**.

Reconocimientos

Este informe ha sido posible gracias a la colaboración de los miembros del Comité y de su Presidente, cuya experiencia y compromiso han enriquecido el análisis y las conclusiones aquí presentadas.

Comité de Expertos de Innovación para la eficiencia en Infraestructuras IT

Presidente



Antonio Ruiz-Falcó
PUE Control

Miembros titulares



Pedro Muñoz
Global Switch



Jorge Valenzuela
Nutanix



Jose Luis Casaus
Aquads Technologies



Víctor Mojarrieta
Kynegos



Carolina Grases
Repsol Fundación



Miguel Ángel Solana
Komtes



Antonio Cortés
Inetum

Otros expertos participantes

Fernando Inurrigarro
Nutanix

Identificación del Reto

Las organizaciones necesitan hoy más capacidad de computación, almacenamiento y conectividad para responder al crecimiento del dato, la inteligencia artificial y los servicios digitales. Sin embargo, desde el Comité de Expertos identificamos una brecha entre esa demanda y la capacidad de muchas infraestructuras para evolucionar con la rapidez necesaria sin comprometer la continuidad, la eficiencia o el control. El centro de datos vuelve a ocupar una posición estratégica, pero debe responder a necesidades muy diferentes de aquellas para las que fue diseñado.

Esta brecha se aprecia cuando las cargas se mantienen en un entorno por inercia, sin valorar sus requisitos de soberanía, latencia o coste; cuando la organización desconoce dónde se procesan sus datos o qué dependencias introduce un modelo; o cuando la capacidad nominal de una instalación oculta restricciones eléctricas, térmicas, estructurales o de conectividad. También aparece al incorporar aceleradores o soluciones de inteligencia artificial sin prever su impacto sobre la red, el almacenamiento, la refrigeración y la operación.

El reto tiene una dimensión organizativa y económica. Las decisiones sobre infraestructura suelen distribuirse entre tecnología, instalaciones, seguridad, energía, compras y áreas de negocio, sin una visión común de las cargas y sus necesidades. A ello se suman modelos de inversión centrados en la adquisición inicial, que no siempre consideran integración, consumo, mantenimiento, actualización, dependencia del proveedor o retirada. Una solución técnicamente avanzada puede convertirse así en una nueva fuente de coste, complejidad y riesgo.

La expansión de la inteligencia artificial acentúa estas tensiones. Aprovechar datos y conocimiento corporativo exige preservar su trazabilidad, proteger la información sensible y mantener control sobre los modelos y sus resultados. Al mismo tiempo, las nuevas cargas elevan la densidad energética, generan mayores exigencias térmicas y obligan a revisar arquitecturas de computación, red y almacenamiento. Incorporar más capacidad no basta si la infraestructura no puede alimentarla, refrigerarla y operarla de forma segura y eficiente.

Tampoco existe un punto de partida único. Conviven instalaciones recientes con CPD diseñados para cargas convencionales, infraestructura propia con *colocation*, *edge* y nube, y organizaciones con capacidades muy distintas de operación y gobierno. La respuesta no puede consistir en trasladar todo a un único entorno ni en sustituir indiscriminadamente el parque instalado. Debe identificar qué activos siguen aportando valor, qué cargas requieren mayor control y dónde es necesario adaptar, ampliar, externalizar o retirar capacidad.

El Comité sitúa aquí el reto central del informe: **transformar el centro de datos en una plataforma evolutiva capaz de soportar las nuevas cargas y recuperar soberanía sobre el dato, los modelos y el conocimiento, sin detener el negocio ni introducir dependencias innecesarias**. El decálogo desarrolla las capacidades arquitectónicas, energéticas, operativas, humanas y de gobierno necesarias para avanzar en esa dirección, reforzando la eficiencia, la competitividad, la resiliencia y la sostenibilidad de las organizaciones.

1. Soberanía del dato, del modelo y del conocimiento

Definir qué activos digitales requieren un control reforzado y establecer las condiciones necesarias para conservar capacidad de decisión sobre ellos

La incorporación de la inteligencia artificial a procesos críticos amplía el alcance de la soberanía digital. Ya no basta con conocer dónde se almacenan los datos: también es necesario saber **quién puede utilizarlos, con qué finalidad, cómo se procesan, qué conocimiento generan y si esa capacidad podría trasladarse a otro entorno** sin perder funcionalidad, trazabilidad o autonomía.

Esta cuestión resulta especialmente relevante para compañías energéticas, *utilities*, empresas industriales y responsables de infraestructuras públicas. Sus sistemas digitales intervienen en la operación de activos críticos, la prestación de servicios esenciales y la toma de decisiones con impacto económico, ambiental y social. La pérdida de control puede afectar tanto a la competitividad como a la resiliencia.

La soberanía debe entenderse, por tanto, como la capacidad verificable para **decidir, supervisar, proteger, trasladar e interrumpir**, cuando sea necesario, el uso de una tecnología. La ubicación física importa, pero no determina por sí sola el grado de control.

De la soberanía del dato a la soberanía del conocimiento

Los datos corporativos abarcan desde información pública hasta diseños industriales, telemetría operacional, información personal, patrones de consumo, documentación técnica o procedimientos vinculados a infraestructuras críticas. Cuando se emplean para entrenar, ajustar o contextualizar modelos, generan una nueva capa de conocimiento cuyo valor puede igualar o superar el de la información original.

ACTIVO QUE DEBE GOBERNARSE	VALOR Y RIESGO PRINCIPAL
Datos de origen	Reflejan clientes, activos y operaciones; pueden sufrir acceso, pérdida, alteración o reutilización indebida.
Documentación y contexto	Incorporan experiencia y propiedad intelectual que puede quedar expuesta.
Modelos y adaptaciones	Contienen aprendizaje y criterios específicos; pueden crear dependencia tecnológica.
Prompts, reglas y mecanismos de recuperación	Determinan cómo se contextualiza el modelo y pueden ser manipulados.
Conocimiento derivado	Incluye predicciones, recomendaciones y patrones cuyo uso debe ser trazable.
Registros de operación	Permiten reconstruir consultas, respuestas y actuaciones, pero también pueden exponer información.

También deben gobernarse las relaciones entre estos elementos: qué datos alimentaron una versión, qué transformaciones se aplicaron, qué parámetros cambiaron y qué fuentes sustentaron una respuesta. Sin esa cadena de evidencias, una organización puede conservar físicamente sus datos y haber perdido la capacidad de comprender o justificar el conocimiento producido.

Control efectivo, no necesariamente propiedad física

Un CPD propio puede facilitar la supervisión y la autonomía, pero no garantiza soberanía si depende de hardware propietario, licencias difíciles de sustituir, modelos cerrados o servicios externos que la organización no controla. Del mismo modo, acudir a *colocation*, una nube privada gestionada o un servicio especializado no implica renunciar al control si existen garantías suficientes de transparencia, segregación, seguridad, auditoría y portabilidad.

La evaluación debe considerar conjuntamente **localización y jurisdicción; accesos, cifrado y control de claves; capacidad de operación y auditoría; continuidad, portabilidad y reversibilidad**. Controlar no significa poseer toda la infraestructura, sino conservar capacidad efectiva de decisión sobre los activos críticos y las condiciones en las que se explotan.

Clasificar antes de decidir

No todos los activos necesitan el mismo nivel de protección. Aplicar el máximo control a cualquier carga eleva costes y dificulta la innovación; utilizar criterios laxos puede exponer información o procesos esenciales. Antes de elegir proveedor, plataforma o modelo de despliegue, responsables de datos, infraestructura, seguridad, operación, cumplimiento y negocio deben valorar:

- **Sensibilidad y criticidad**: naturaleza de la información y consecuencias de su pérdida, alteración o indisponibilidad.
- **Contexto operativo**: intervención del sistema en decisiones o actuaciones sobre activos críticos.
- **Autonomía y continuidad**: capacidad mínima para operar si falla un proveedor, componente o enlace.
- **Trazabilidad y jurisdicción**: necesidad de reconstruir resultados y marcos legales aplicables.
- **Reversibilidad**: posibilidad de sustituir la solución sin perder información o capacidades esenciales.
- **Impacto sectorial**: efectos sobre servicios esenciales, activos críticos o el interés público.

Esta clasificación evita tanto concentrar todos los activos en un único perímetro como asumir que cualquier servicio comercial sirve para cualquier finalidad. Una prueba con datos anonimizados no exige el mismo control que un sistema conectado con una red eléctrica, una planta industrial o una infraestructura urbana.

Gobernar el ciclo de vida y preparar la reversibilidad

La soberanía debe mantenerse durante la preparación de los datos, el desarrollo, el entrenamiento, la puesta en producción, las actualizaciones y la retirada. El alcance puede cambiar al combinar fuentes, incorporar información más sensible o añadir conectores, agentes y proveedores.

La organización necesita capacidades permanentes para:

- inventariar datos, modelos, versiones, integraciones, dependencias y responsables;
- documentar origen, transformaciones, usos autorizados y criterios de conservación;
- segregar entornos y controlar los contenidos destinados a entrenamiento, ajuste o inferencia;

- registrar accesos, cambios y resultados para reconstruir decisiones;
- revisar permisos, proveedores y riesgos, y suspender o revertir componentes cuando sea necesario.

Esta gobernanza debe compartirse entre tecnología, seguridad, operación, cumplimiento y negocio. El conocimiento está distribuido por toda la organización y no puede protegerse mediante una decisión exclusiva del departamento IT.

La dependencia, además, ya no se limita a la extracción de datos. Puede residir en APIs, formatos, modelos, identidades, catálogos, herramientas de orquestación y observabilidad, aceleradores o competencias vinculadas a un proveedor. Por ello, la reversibilidad debe diseñarse desde el inicio mediante formatos exportables, documentación, separación entre datos y servicios propietarios, alternativas para componentes críticos, pruebas de traslado y condiciones contractuales de salida.

El *Data Act* impulsa el cambio entre proveedores y la portabilidad, pero la regulación no sustituye el diseño técnico: una interfaz de exportación no garantiza que otro entorno pueda reproducir el servicio.

El marco europeo como criterio de gobierno

El marco europeo relaciona protección de datos, resiliencia, gobernanza algorítmica y portabilidad. El *RGPD* mantiene la responsabilidad de la organización cuando intervienen terceros; el *Reglamento de IA* refuerza, según el nivel de riesgo, la gobernanza de datos, la documentación, la supervisión humana, la robustez y la ciberseguridad; y *NIS2* amplía la atención hacia la continuidad, los incidentes y la cadena de suministro.

Estas normas no obligan con carácter general a ejecutar las cargas en un CPD propio. Exigen conocer los riesgos, asignar responsabilidades, mantener evidencias y demostrar un control proporcionado, con independencia de que se utilice infraestructura propia, gestionada o compartida.

Prioridades sectoriales e indicadores

La aplicación concreta varía por sector. Energía y *utilities* priorizan integridad, continuidad y segregación IT-OT; la industria añade la protección de diseños, procesos y propiedad intelectual; y las infraestructuras públicas deben combinar seguridad, interoperabilidad, continuidad y responsabilidad ante la ciudadanía.

En todos los casos, la organización debe poder identificar sus datos, modelos y conocimientos críticos; quién accede a ellos; qué dependencias condicionan su uso; cómo mantendría el servicio ante un fallo; y si podría sustituir la solución sin perder capacidades esenciales.

INDICADOR	EVIDENCIA QUE DEBE APORTAR
Activos críticos clasificados	Inventario actualizado, criticidad y responsable.
Trazabilidad de datos y modelos	Linaje, versiones y documentación.
Servicios con estrategia de salida	Planes, dependencias y condiciones contractuales.
Portabilidad probada	Resultados, incidencias y tiempos de migración.
Dependencias de terceros identificadas	Proveedores, accesos, ubicaciones y responsabilidades.
Control sobre claves y registros	Políticas criptográficas, gestión de claves y acceso a logs.

Los indicadores deben ajustarse a la criticidad y revisarse periódicamente. No todas las cargas precisan el mismo nivel de trazabilidad o una migración inmediata, pero **las excepciones deben ser conscientes, proporcionadas y documentadas**.

De la soberanía declarativa a la soberanía operativa

La soberanía no es una etiqueta asociada a una ubicación o proveedor, sino un conjunto de decisiones, controles y evidencias. Requiere clasificar antes de elegir, aplicar medidas proporcionales, mantener trazabilidad, diseñar la portabilidad, limitar dependencias y medir de forma periódica la autonomía alcanzada.

El objetivo no es aislarse de servicios externos, sino adoptar nuevas capacidades sin renunciar al gobierno de los activos que sostienen la competitividad, la continuidad y el servicio público. **La cuestión decisiva no es solo dónde está el servidor, sino si la organización puede comprender, proteger, auditar, trasladar y gobernar el conocimiento que depende de él**. Sobre esa base podrá decidirse qué combinación de CPD propio, *colocation*, *edge*, *cloud* o infraestructura especializada resulta adecuada para cada carga.

2. Una infraestructura para cada carga: CPD propio, *colocation*, *edge* y nube híbrida

Seleccionar el entorno más adecuado para cada necesidad según criterios de control, criticidad, latencia, escalabilidad, eficiencia y capacidad operativa

La evolución de las infraestructuras digitales ha dejado obsoleta la idea de alojar todas las cargas en un único lugar. CPD propio, *colocation*, nube privada, *cloud* público, *edge computing* y proveedores especializados ofrecen capacidades diferentes que pueden combinarse dentro de una misma estrategia.

La ubicación condiciona el rendimiento, el coste, la seguridad, la huella energética, la escalabilidad, la dependencia tecnológica y la continuidad. Una solución adecuada para una aplicación corporativa puede no serlo para un sistema conectado con una planta industrial, una red energética o una infraestructura pública.

La cuestión no es qué modelo resulta mejor en términos generales, sino **qué combinación permite atender cada necesidad, conservar el control definido en el capítulo anterior y mantener opciones de evolución**. La infraestructura debe concebirse como un ecosistema distribuido y gobernado con criterios comunes.

Del centro de datos como lugar al continuo de computación

La tradicional elección entre infraestructura interna y nube ya no refleja la realidad. Una organización puede desplegar una nube privada en su CPD o en *colocation*, coordinar nodos *edge* desde una plataforma *cloud* y consumir capacidad acelerada para necesidades puntuales. Incluso las distintas fases de una carga pueden ejecutarse en entornos diferentes.

La Comisión Europea describe esta evolución como un continuo *cloud-edge-IoT*, en el que datos y computación se distribuyen entre infraestructuras centrales y recursos próximos a su lugar de uso. El objetivo es combinar seguridad, sostenibilidad e interoperabilidad sin convertir la descentralización en fragmentación.

MODELO	APORTACIÓN PRINCIPAL	CONDICIONANTES
CPD propio	Control físico y operativo para cargas críticas, estables o integradas con sistemas internos	Inversión, energía, actualización y talento
Colocation	Infraestructura dedicada sin operar el edificio y sus servicios físicos	Dependencia del operador, conectividad, acceso y contrato
Nube privada	Automatización y recursos compartidos dentro de un perímetro dedicado	Complejidad, coste operativo y capacidad interna
Nube pública	Elasticidad, rapidez y acceso a servicios gestionados	Coste variable, portabilidad y dependencia de servicios propietarios
Edge computing	Procesamiento próximo a activos, sensores y usuarios	Gestión distribuida, seguridad física y recursos limitados
Neocloud o GPU as a Service	Capacidad acelerada sin inversión permanente	Disponibilidad, movimiento de datos y continuidad

Estas categorías no son compartimentos estancos. Lo relevante es definir qué función cumple cada entorno, qué responsabilidades conserva la organización y cómo se conectan y gobiernan todos ellos.

La carga debe preceder a la plataforma

Seleccionar primero una plataforma y adaptar después todas las cargas puede generar sobrecostes, restricciones o dependencias difíciles de corregir. La decisión debe comenzar por el comportamiento de la carga y considerar que desarrollo, validación, entrenamiento e inferencia quizá necesiten ubicaciones distintas.

Además de la soberanía, deben evaluarse:

- **Criticidad, latencia y determinismo**: consecuencias de una interrupción y tiempos de respuesta exigidos.
- **Autonomía e integración**: necesidad de operar sin conectividad externa o junto a sistemas locales, legacy u OT.
- **Variabilidad y horizonte temporal**: estabilidad, estacionalidad y duración prevista de la demanda.
- **Gravedad del dato**: volumen, frecuencia y coste de trasladar la información.
- Especialización: necesidad de aceleradores o servicios difíciles de mantener internamente.
- **Capacidad operativa**: disponibilidad de energía, espacio, conectividad y profesionales.
- **Coste total y reversibilidad**: inversión, consumo, dependencias y posibilidad de cambiar de entorno.

Ningún criterio debe aislarse. Una carga puede necesitar baja latencia y, al mismo tiempo, una capacidad computacional inviable localmente; otra puede manejar información sensible y requerir una elasticidad difícil de ofrecer en un CPD propio. La elección debe documentar tanto la solución como los compromisos aceptados.

Qué aporta cada modelo

El **CPD propio** conserva valor para cargas previsibles, críticas y estrechamente integradas con sistemas internos u OT. Facilita procedimientos específicos de acceso, recuperación y supervisión, pero solo resulta adecuado si puede responder a las nuevas necesidades de potencia, refrigeración, conectividad, seguridad y actualización. No debería sobredimensionarse para picos ocasionales ni mantenerse únicamente porque ya existe.

El ***colocation*** permite conservar el control sobre servidores, almacenamiento, red y configuración mientras un tercero gestiona edificio, energía, climatización y seguridad física. Puede aportar capacidad, recuperación o un emplazamiento especializado sin ampliar el CPD corporativo. Su evaluación debe incluir potencia realmente disponible, conectividad, refrigeración, acceso, servicios operativos, escalabilidad y condiciones de salida, no solo el precio por rack.

La **nube privada** aplica automatización, provisión bajo demanda y recursos compartidos dentro de un perímetro dedicado, ya sea propio, en *colocation* o gestionado. Aporta valor cuando estandariza y agiliza la provisión; pierde sentido si mantiene procesos manuales, baja utilización o tecnologías difíciles de sustituir.

El ***cloud* público** resulta especialmente útil para experimentación, entornos temporales, cargas variables y servicios gestionados. Reduce la inversión inicial, pero exige gobernar consumo, identidades, cifrado, registros, datos autorizados y dependencia tecnológica. Cuando la carga se estabiliza, aumenta el volumen de datos o se intensifica el uso de servicios propietarios, el coste y la dificultad de salida pueden crecer. El marco europeo favorece la interoperabilidad y el cambio entre servicios de procesamiento, pero no sustituye una arquitectura preparada para migrar.

El ***edge computing*** procesa información cerca del activo o usuario cuando la latencia, la autonomía o el tráfico hacen ineficiente una solución exclusivamente centralizada. Es especialmente relevante en energía, industria, ciudades y movilidad. Según *ETSI*, combina proximidad, baja latencia y acceso al contexto local. Su despliegue requiere inventario, configuración homogénea, actualización remota, protección física y lógica, continuidad sin conexión y coordinación con las plataformas centrales.

Los servicios ***neocloud* o *GPU as a Service*** permiten acceder a aceleradores para pruebas, entrenamientos puntuales o picos de demanda sin una inversión permanente. Deben valorarse según frecuencia de uso, reserva y disponibilidad, sensibilidad y movimiento de datos, continuidad y posibilidad de reproducir el entorno con otro proveedor.

Diseñar la arquitectura híbrida

Muchas organizaciones ya operan en entornos híbridos como resultado de proyectos, adquisiciones y decisiones departamentales. La hibridación solo aporta valor cuando cada entorno tiene una función definida y existe una capa común de gobierno. De lo contrario, multiplica plataformas, costes, identidades y puntos de fallo.

El modelo debe establecer:

- qué cargas y datos pueden operar en cada entorno;
- cómo se gestionan identidades, permisos y políticas;
- qué conectividad y niveles de servicio unen las ubicaciones;
- cómo se supervisan coste, rendimiento, seguridad y capacidad;
- cómo se recupera o traslada una carga.

Una misma iniciativa puede usar *cloud* para el desarrollo, capacidad especializada para el entrenamiento e infraestructura controlada o *edge* para la inferencia. Esta distribución debe planificarse para evitar movimientos innecesarios y dependencias imprevistas.

Una política gobernada y revisable

La asignación de cargas no puede depender de decisiones aisladas. Debe seguir un proceso común: clasificar carga y datos; definir requisitos; comparar alternativas; evaluar coste total, riesgos y salida; validar la decisión entre infraestructura, seguridad, operación y negocio; registrar excepciones; y revisar periódicamente la ubicación.

Las prioridades variarán. Energía y *utilities* pondrán el foco en continuidad, segregación IT-OT y autonomía de activos distribuidos; la industria, en latencia e integración con la producción; y las infraestructuras públicas, en protección, interoperabilidad, continuidad y reversibilidad contractual.

INDICADOR	QUÉ PERMITE EVALUAR
Cargas con ubicación justificada	Aplicación efectiva del marco de decisión
Latencia y disponibilidad	Adecuación a las necesidades operativas
Utilización y coste total por carga	Riesgo de sobredimensionamiento e impacto económico
Tiempo de provisión	Agilidad para responder a nuevas demandas
Transferencia de datos	Efecto de la distribución y la gravedad del dato
Alternativas de recuperación y portabilidad probada	Resiliencia y capacidad real de cambio
Incidentes de integración	Complejidad introducida por la hibridación
Nodos *edge* inventariados y actualizados	Control de la infraestructura distribuida

Los indicadores deben analizarse conjuntamente: reducir costes puede aumentar la latencia o la dependencia; elevar el control puede limitar la elasticidad; y acelerar la provisión puede fragmentar la arquitectura.

El valor reside en conservar opciones

La infraestructura adecuada no es la que concentra más activos internamente ni la que consume más servicios externos. Es la que responde a las necesidades reales, mantiene el control sobre los elementos críticos y permite revisar las decisiones cuando cambia el contexto.

Cada carga debe ejecutarse en el entorno que ofrezca el mejor equilibrio entre control, rendimiento, resiliencia, eficiencia y capacidad de evolución. La hibridación no es una solución por sí misma: solo adquiere valor cuando está gobernada, documentada y preparada para cambiar. Definida la ubicación de cada carga, el siguiente capítulo abordará la capacidad necesaria para explotar el conocimiento corporativo mediante IA privada y factorías de conocimiento.

3. IA privada y factorías corporativas de conocimiento

Convertir los datos y el conocimiento de la organización en una capacidad controlada, reutilizable y orientada a generar valor

Después de determinar qué activos requieren un control reforzado y dónde puede ejecutarse cada carga, el reto consiste en utilizar ese conocimiento de forma segura, repetible y escalable. La dificultad no reside solo en acceder a un modelo avanzado, sino en conectarlo con información fiable, procesos reales y responsabilidades claras.

Muchas organizaciones han comenzado con pruebas departamentales y asistentes aislados sobre plataformas diferentes. Estas iniciativas aceleran el aprendizaje, pero pueden crear datos duplicados, conexiones no supervisadas, evaluaciones incompatibles y aplicaciones que nunca llegan a integrarse en la operación.

La **IA privada** permite superar esta fragmentación. No es un producto ni exige ejecutar todas las cargas en un CPD propio: es un modelo de control mediante el que la organización decide qué información y modelos pueden utilizarse, cómo se validan los resultados y qué evidencias deben conservarse. Su objetivo es convertir el conocimiento corporativo en una capacidad común, evitando que cada caso de uso reconstruya su propia infraestructura de datos, seguridad y gobierno.

IA privada no significa infraestructura aislada

El carácter privado depende del control ejercido sobre datos, modelos, configuraciones, accesos y resultados. Puede desplegarse en infraestructura propia, *colocation*, nube dedicada o varios entornos, siempre que respete los criterios de soberanía y ubicación definidos anteriormente. A la inversa, instalar un modelo *on-premise* no garantiza el control si los permisos son excesivos, las versiones no se gestionan o no existe trazabilidad.

CAPA	QUÉ DEBE GOBERNARSE	RIESGO PRINCIPAL
Datos y documentos	Fuentes, calidad, permisos, conservación y finalidad	Exposición o uso fuera de contexto
Modelos	Procedencia, versión, licencia y limitaciones	Dependencia o resultados imprevisibles
Contexto e instrucciones	*Prompts*, reglas y restricciones	Manipulación o pérdida de criterios corporativos
Herramientas y conectores	Sistemas accesibles y acciones permitidas	Escalada de privilegios o impacto operativo
Aplicaciones y resultados	Usuarios, validaciones, trazabilidad y uso posterior	Soluciones sin responsable o decisiones no justificables
Operación	Monitorización, cambios, incidencias y retirada	Degradación silenciosa u obsolescencia

La posición estratégica es clara: **la IA privada constituye un modelo de gobierno y explotación del conocimiento, no una categoría de hardware ni una localización concreta**.

Del proyecto aislado a la factoría corporativa

Cuando cada área elige por separado modelos, fuentes y herramientas, aumentan los costes y resulta difícil reutilizar conectores, evaluaciones y aprendizajes. La alternativa es una plataforma común que acompañe a los casos de uso desde su propuesta hasta su retirada.

El concepto de *AI Factory* describe ecosistemas que combinan computación, datos y talento. Adaptado a la organización, representa una factoría corporativa capaz de:

- priorizar casos de uso con necesidad, valor y responsable definidos;
- proporcionar acceso controlado a fuentes corporativas;
- seleccionar y sustituir modelos sin rehacer toda la solución;
- reutilizar conectores, *prompts*, evaluaciones y controles;
- aplicar criterios comunes de seguridad, despliegue y supervisión;
- actualizar o retirar aplicaciones cuando cambien las necesidades.

Su madurez no debe medirse por las GPU instaladas ni por el número de modelos disponibles. **Industrializar no significa multiplicar aplicaciones**, sino reducir el esfuerzo necesario para construirlas y elevar el nivel mínimo de calidad, seguridad y trazabilidad.

El conocimiento corporativo como materia prima

Los modelos generalistas desconocen los procedimientos, activos, configuraciones, contratos y decisiones de cada organización. El valor diferencial aparece al conectarlos con información corporativa fiable. Sin embargo, un gran repositorio no equivale a una base de conocimiento: puede contener duplicados, versiones contradictorias, documentos caducados o permisos heredados.

Antes de incorporar una fuente deben comprobarse su autoridad, vigencia, calidad, propietario, permisos y finalidad. Los especialistas de cada dominio deben validar qué contenido resulta fiable; infraestructura puede proporcionar la plataforma, pero no decidir qué procedimiento industrial está vigente o qué documentación representa el estado real de un activo.

Las arquitecturas RAG permiten recuperar información de fuentes autorizadas antes de generar una respuesta. Así pueden actualizarse contenidos sin reentrenar el modelo, respetarse permisos, mostrar evidencias y reutilizarse repositorios en distintas aplicaciones.

RAG no garantiza respuestas correctas. Documentos irrelevantes, desactualizados o mal fragmentados pueden producir resultados convincentes pero erróneos. La evaluación debe abarcar la consulta, la recuperación, la selección de fuentes, la respuesta y las evidencias, verificando también que se respetan los permisos. La plataforma debe permitir combinar o sustituir modelos sin perder la base de conocimiento ni sus controles.

Capacidades compartidas durante todo el ciclo de vida

Una factoría no es una herramienta monolítica, sino un conjunto coordinado de capacidades: gestión de demanda; catálogos de conocimiento y modelos; entornos de experimentación; evaluación; MLOps y LLMOps; seguridad y cumplimiento; observabilidad; reutilización; y soporte a los usuarios.

El objetivo es aplicar procesos estables sin bloquear la evolución tecnológica. El *AI Risk Management Framework de NIST* ofrece una referencia basada en gobernar, contextualizar, medir y gestionar los riesgos durante todo el ciclo de vida.

La distancia entre una demostración y un sistema operativo es considerable. Antes de avanzar desde la experimentación al piloto y a producción, cada iniciativa debe acreditar:

- el problema que resuelve, el proceso que mejora y el responsable del resultado;
- la calidad de las fuentes y la mejora frente al procedimiento actual;
- los errores posibles, sus consecuencias y la supervisión humana necesaria;

- los sistemas que consulta, las acciones permitidas y los controles aplicados;
- el coste al escalar, el mantenimiento y el procedimiento de retirada;
- la posibilidad de cambiar el modelo sin reconstruir la aplicación.

La factoría debe establecer puertas de decisión con exigencias crecientes de validación, documentación, seguridad e integración. Detener un proyecto que no aporta valor, aunque funcione técnicamente, es una señal de madurez.

Seguridad y responsabilidad integradas

Los modelos forman parte de una cadena que incluye datos, librerías, APIs, conectores, repositorios, usuarios y proveedores. Deben contemplarse el envenenamiento de datos, documentos o instrucciones manipulados, extracción de información, componentes de procedencia dudosa, permisos excesivos y resultados utilizados fuera de contexto. ENISA recomienda abordar estos sistemas como activos multicapa.

Los controles deben ser proporcionales al riesgo. Cuando un agente puede ejecutar acciones, deben limitarse sus herramientas y privilegios, validarse los parámetros y mantenerse confirmación humana ante posibles efectos sobre activos, personas o servicios críticos.

Cada aplicación necesita un responsable de negocio, un propietario de las fuentes, un responsable técnico, especialistas que validen los resultados y participación de seguridad y cumplimiento según el riesgo. El equipo de plataforma aporta componentes comunes, pero no puede asumir el significado de los resultados ni la responsabilidad de las decisiones de cada área.

Aplicaciones sectoriales e indicadores

En energéticas y *utilities*, la factoría puede facilitar el acceso controlado a procedimientos, normativa, históricos y conocimiento sobre activos. En industria, puede conectar manuales, ingeniería, mantenimiento y experiencia acumulada. En infraestructuras públicas, puede mejorar el acceso a expedientes, procedimientos y catálogos, manteniendo permisos y evitando confundir un resultado generado con una decisión administrativa validada.

INDICADOR	QUÉ PERMITE EVALUAR
Casos con responsable y objetivo medible	Calidad de la selección y rendición de cuentas
Pilotos que alcanzan producción	Capacidad de industrialización
Fuentes con propietario, versión y permisos	Preparación del conocimiento
Respuestas fundamentadas y trazables	Calidad, evidencia y capacidad de auditoría
Componentes reutilizados	Eficiencia de la plataforma común
Incidentes, excepciones y sistemas retirados	Riesgo y control del ciclo de vida
Uso efectivo y mejora frente a la línea base	Integración y valor real

Las métricas deben definirse antes del despliegue. La cantidad de proyectos iniciados no demuestra capacidad corporativa: el valor aparece cuando los casos alcanzan producción, reutilizan activos anteriores y mejoran de forma comprobable los procesos.

Convertir el conocimiento en una capacidad sostenible

La ventaja no residirá únicamente en disponer del modelo más avanzado, sino en **conectarlo con conocimiento fiable, evaluarlo en el contexto real y sustituirlo sin perder los activos construidos a su alrededor**. La IA privada convierte este objetivo en una política operativa y la factoría corporativa lo lleva a la práctica mediante un núcleo común de datos, modelos, seguridad, evaluación, despliegue y supervisión.

Este enfoque evita tanto la proliferación descontrolada de herramientas como una plataforma excesivamente rígida. Una vez definida esta capacidad, el siguiente reto será adaptar la infraestructura física a las concentraciones de potencia y calor que determinadas cargas exigirán.

4. Alta densidad y refrigeración híbrida: rediseñar la infraestructura física

Adaptar la distribución eléctrica y la gestión térmica a cargas que concentran grandes cantidades de potencia sin sobredimensionar el conjunto del centro de datos

La computación acelerada está alterando la relación entre espacio, potencia y capacidad de procesamiento. Algunos clústeres concentran en un rack una demanda antes distribuida por una sala completa, con efectos sobre alimentación eléctrica, evacuación del calor, estructura, mantenimiento y crecimiento.

La densidad extrema todavía no representa la situación habitual. Según la *encuesta global de Uptime Institute de 2025*, los racks por encima de 100 kW siguen siendo excepcionales. La transformación se caracteriza por la convivencia entre cargas convencionales y zonas concentradas.

El reto es soportar esos nuevos perfiles sin diseñar toda la instalación para el escenario más extremo. **Potencia y refrigeración deben planificarse conjuntamente**, porque la energía consumida termina convirtiéndose en calor que debe extraerse de forma segura y continua.

No existe una única densidad de referencia

La densidad media puede ocultar restricciones locales. Un CPD puede disponer de potencia agregada suficiente y no ser capaz de transportarla hasta un rack determinado o evacuar el calor en esa posición. La planificación debe considerar la carga real, su duración, utilización prevista y posible evolución.

PERFIL ORIENTATIVO	CARGAS HABITUALES	IMPLICACIÓN FÍSICA
Hasta 10-15 kW por rack	Aplicaciones y almacenamiento convencionales	Compatible con diseños tradicionales bien gestionados
Entre 15 y 30 kW	Infraestructura consolidada y servidores potentes	Revisión de distribución de aire y capacidad eléctrica local
Entre 30 y 80 kW	HPC y configuraciones aceleradas	Refrigeración próxima o soluciones líquidas parciales
Más de 80-100 kW	Clústeres densos de computación acelerada	Diseño específico de alimentación, refrigeración y estructura
Cientos de kW	Sistemas altamente integrados de nueva generación	Arquitecturas eléctricas y térmicas diferentes

Estas bandas son orientativas. La viabilidad depende del caudal, la contención, la temperatura y el diseño del servidor. Preparar toda una sala para densidades que solo afectarán a pocos racks puede reducir la eficiencia y desviar inversiones.

Del diseño de sala al diseño por rack, zona y pod

Las nuevas cargas obligan a sustituir la visión homogénea de la sala por unidades más granulares. Hay que conocer la potencia entregable en cada posición, el calor evacuable, el peso admisible y las condiciones de mantenimiento.

La revisión debe abarcar:

- acometidas, transformadores, cuadros, UPS, *busway*, líneas y PDU;
- coordinación de protecciones, secciones de cable y calidad del suministro;
- picos y variaciones rápidas de las cargas aceleradas;
- redundancia real en cada circuito y tramo de distribución;
- acceso seguro y capacidad de crecer sin reconstruir el conjunto.

Una arquitectura global N+1 o 2N no evita que aparezcan limitaciones en un circuito o una zona. Al concentrarse gran capacidad computacional en pocos racks, cualquier fallo localizado adquiere más impacto.

Evolución de la distribución eléctrica

La infraestructura convencional transporta corriente alterna hasta el rack y realiza sucesivas conversiones a corriente continua. Al aumentar la densidad, distribuir energía a baja tensión exige intensidades mayores, más cobre, más espacio y genera más pérdidas y calor.

Los *power shelves* compartidos y los *power racks* o *sidecars* permiten centralizar la alimentación y liberar espacio en el rack IT. Open Compute Project trabaja en alimentación desagregada que distribuye ±400 V DC —800 V entre polos— desde un *power rack* lateral; su especificación Diablo plantea escalar desde unos 100 kW hacia 1 MW.

NVIDIA sitúa su adopción de 800 V DC a partir de 2027 para racks de 1 MW o más. Es una hoja de ruta, no una solución generalizada para los racks actuales de 140 o 150 kW. Tampoco elimina todas las conversiones: reorganiza la cadena y acerca la energía a mayor tensión al consumo.

El mensaje para las organizaciones no es adoptar inmediatamente 800 V DC, sino **evitar decisiones que cierren la puerta a futuras arquitecturas**, especialmente en ampliaciones con horizontes largos.

La refrigeración será híbrida

El aire continuará siendo válido para buena parte de las cargas. La contención y una gestión adecuada de caudales mejoran su rendimiento. El límite aparece cuando el calor concentrado exige más ventilación, presión y espacio y aumenta los puntos calientes.

El líquido transporta más calor con menor volumen y puede acercarse al punto de generación. Aun así, no sustituirá de inmediato al aire: incluso los servidores con *cold plates* mantienen calor residual en memoria, almacenamiento o alimentación. La convivencia de ambas tecnologías será estructural.

TECNOLOGÍA	APLICACIÓN PRINCIPAL	CONSIDERACIÓN CLAVE
Aire con contención	Densidades moderadas	Madurez, pero menor eficacia al crecer la concentración
In-row o close-coupled	Filas o zonas más densas	Proximidad a la carga y ocupación adicional
Rear Door Heat Exchanger	Racks con alta expulsión de calor	Circuito líquido y espacio de mantenimiento
Direct-to-Chip	CPU, GPU y aceleradores	Alta extracción y necesidad de tratar calor residual
Inmersión	Configuraciones especializadas	Cambios en formato, materiales y mantenimiento
Aire-líquido	Instalaciones heterogéneas	Adaptación por zonas con mayor complejidad operativa

ASHRAE recomienda alinear el diseño con la densidad prevista y contemplar desde el principio opciones *liquid-ready*. La elección debe valorar compatibilidad, temperaturas, vida útil, mantenimiento, repuestos, riesgos y capacidad de evolución, no solo la extracción máxima.

El circuito líquido como infraestructura crítica

Introducir líquido en la sala implica diseñar una cadena completa entre los componentes y el rechazo de calor del edificio. Las CDU regulan y separan circuitos, caudal, presión y temperatura, y deben coordinarse con bombas, intercambiadores, tuberías, colectores y monitorización.

La operación debe supervisar temperaturas, caudal y presión, calidad del fluido, bombas y filtros, fugas, capacidad térmica y autonomía. También debe conocer qué ocurre si falla una CDU, una bomba o un tramo y cuánto tiempo puede funcionar el equipo antes de reducir rendimiento o detenerse.

Los procedimientos cambian: conectar y purgar circuitos, mantener el fluido o responder a una fuga exige coordinación entre IT, ingeniería térmica, mantenimiento y fabricantes.

Agua, estructura y espacio

Refrigeración líquida no equivale necesariamente a mayor consumo de agua. El fluido interno puede circular en un circuito cerrado; el consumo depende principalmente del rechazo final mediante *dry coolers*, sistemas adiabáticos, torres evaporativas u otras soluciones. La evaluación debe distinguir agua recirculada y evaporada, calidad y tratamiento, estrés hídrico local, energía necesaria para evitar evaporación, fuentes alternativas y posible reutilización térmica.

La densidad también aumenta el peso. ASHRAE advierte de racks líquidos que pueden superar los 1.800 kg. Deben verificarse estructura, transporte, dimensiones, espacio para *power racks* y CDU, accesos, compatibilidad de materiales y seguridad ante fugas o tensiones elevadas.

Adaptar zonas sin sobredimensionar el CPD

La coexistencia de densidades aconseja crear filas, *pods*, islas o módulos con capacidades eléctricas y térmicas diferenciadas. Así pueden mantenerse soluciones de aire para cargas convencionales y reservar refrigeración avanzada para clústeres concretos.

Estas zonas deben concebirse como unidades completas: rack, PDU, circuito, refrigeración, estructura y mantenimiento deben evolucionar conjuntamente. Preparar canalizaciones y espacios *liquid-ready* permite acompasar inversión y demanda sin asumir desde el inicio todo el coste y la complejidad.

Indicadores para gobernar potencia y refrigeración

El PUE sigue siendo útil, pero no muestra si un rack dispone de capacidad, si las GPU se utilizan o si la arquitectura térmica funciona correctamente. Se necesitan métricas más granulares:

INDICADOR	QUÉ PERMITE EVALUAR
Potencia real y capacidad por rack o zona	Utilización, margen y nuevas cargas posibles
Redundancia por circuito	Resiliencia local
Capacidad térmica utilizada	Ocupación de aire, CDU e intercambiadores
Temperatura, caudal y presión	Comportamiento y capacidad de extracción
Energía de ventiladores y bombas	Consumo auxiliar
Calor capturado por líquido	Reparto entre líquido y aire residual
Incidencias térmicas	Puntos calientes, alarmas y reducción de rendimiento
WUE y consumo hídrico	Uso de agua en el rechazo de calor
Peso, ocupación y capacidad *liquid-ready*	Compatibilidad y preparación futura

Estos indicadores deben relacionarse con la utilización computacional: reservar potencia y refrigeración para equipos infrautilizados bloquea capacidad y reduce el valor de la inversión.

Prepararse para la densidad sin convertirla en objetivo

La densidad es una consecuencia de determinadas arquitecturas, no un indicador de éxito. Para muchas organizaciones, la solución adecuada será una zona limitada de computación acelerada dentro de una infraestructura mayoritariamente convencional.

El CPD debe incorporar alta densidad donde aporte valor, sin trasladar su coste y complejidad al resto de las cargas. Esto exige segmentación, modularidad y un diseño coordinado de potencia, refrigeración, espacio y operación. Resuelta esta capacidad física, el siguiente capítulo abordará cómo garantizar una energía suficiente y resiliente y convertir el CPD en un recurso flexible para el sistema eléctrico.

5. Energía, flexibilidad y resiliencia: el CPD como aliado del sistema eléctrico

Garantizar potencia suficiente para el crecimiento digital y convertir una demanda intensiva en un recurso más previsible, gestionable e integrado con la red

La energía ya no es solo una partida del coste operativo. La potencia, los plazos de conexión y la continuidad condicionan dónde pueden desplegarse nuevas infraestructuras y a qué velocidad pueden crecer.

El desafío también alcanza a energéticas, *utilities*, industrias y administraciones que amplían sus CPD, incorporan computación acelerada o despliegan capacidad distribuida. Un proyecto viable puede quedar bloqueado por falta de potencia o plazos incompatibles con el negocio.

La Comisión Europea estima que los centros de datos representan alrededor del 2,5 % del consumo eléctrico de la UE y que su capacidad podría pasar de 12 GW en 2025 a 28 GW en 2030. Concentrar solicitudes puede aumentar la congestión y competir con otros procesos de electrificación.

La respuesta no consiste únicamente en construir red y contratar más potencia. Los CPD deben mejorar su eficiencia, ofrecer previsiones fiables y coordinar almacenamiento, generación local y gestión de la demanda. **El objetivo es pasar de una relación pasiva con el sistema eléctrico a una integración planificada y, cuando sea posible, flexible**.

De optimizar el consumo a garantizar la potencia

Reducir el consumo anual sigue siendo importante, pero un CPD eficiente puede no disponer de capacidad eléctrica para incorporar nuevas cargas. La variable crítica es el ***time-to-power***: el tiempo necesario para convertir una necesidad en potencia realmente utilizable, incluidos estudios, permisos, refuerzos, subestaciones y puesta en servicio.

También deben diferenciarse potencia solicitada, concedida, contratada, disponible y utilizada. Reservar capacidad que tardará años en ocuparse bloquea recursos y eleva costes; infravalorar la demanda obliga a replantear la infraestructura cuando las cargas ya están comprometidas.

VARIABLE	PREGUNTA DE PLANIFICACIÓN	RIESGO PRINCIPAL
Potencia disponible	¿Cuánta capacidad puede utilizarse en cada ubicación?	Confundir disponibilidad global y local
Crecimiento comprometido	¿Qué cargas llegarán y en qué fases?	Solicitar capacidad demasiado pronto o tarde
Perfil y factor de carga	¿Cómo varía la demanda y cuánto se utilizará?	Dimensionar solo por el máximo
Tiempo de conexión	¿Qué actuaciones externas condicionan el suministro?	Desalinear calendarios energéticos y tecnológicos
Calidad eléctrica	¿Qué perturbaciones toleran las cargas?	Fallos pese a disponer de energía
Flexibilidad	¿Qué consumo puede desplazarse o modularse?	Tratar toda la demanda como rígida

La planificación debe apoyarse en escenarios y despliegues modulares, sin ejecutar toda la ampliación ni solicitar toda la potencia desde el inicio.

Una hoja de ruta tecnológica y energética común

Los equipos IT trabajan con ciclos más cortos que la infraestructura eléctrica: un clúster puede adquirirse en meses, mientras una ampliación de red puede requerir años. Por ello, infraestructura digital, energía, instalaciones, sostenibilidad, finanzas y negocio deben compartir una hoja de ruta.

Esta coordinación permite asignar un presupuesto de potencia a:

- cargas actuales y crecimiento ya comprometido;
- proyectos en validación y ampliaciones sujetas a hitos;
- capacidad flexible o interrumpible;
- almacenamiento y generación local;
- márgenes de contingencia.

Así se evita que varias áreas comprometan simultáneamente una capacidad inexistente y se facilita decidir si una necesidad debe atenderse en el CPD actual, en otra ubicación o mediante recursos externos.

Resiliencia más allá de la redundancia

La continuidad depende de una cadena que comienza en la red y termina en cada equipo. Una doble acometida o una configuración 2N no garantiza que todos los elementos sean independientes ni que soporten un incidente prolongado.

La estrategia debe contemplar fallos de red, calidad eléctrica, averías, combustible, mantenimiento, errores humanos y eventos ambientales, diferenciando qué cargas deben mantenerse y cuáles pueden degradarse.

Cada horizonte requiere una respuesta. Las UPS cubren milisegundos o minutos; baterías y almacenamiento térmico sostienen transiciones o gestionan picos; y la generación de emergencia responde ante interrupciones largas. La resiliencia exige probar la secuencia completa con cargas y sistemas térmicos, no cada componente aisladamente.

Hibridar recursos sin sobredimensionar

La combinación de red, autoconsumo, PPA, baterías UPS, BESS, generación de respaldo, almacenamiento térmico y gestión de cargas puede reducir picos, mejorar la continuidad y aprovechar mejor la infraestructura.

Ningún recurso resuelve por sí solo todas las necesidades. La solar reduce demanda en determinadas horas, pero no garantiza continuidad nocturna; las baterías desplazan consumo o cubren picos, pero tienen autonomía limitada; y la generación de emergencia aporta duración a cambio de condicionantes ambientales, logísticos y regulatorios.

La coordinación debe optimizar **continuidad, coste, emisiones y utilización de red**. Si los recursos se dimensionan por separado, pueden duplicarse capacidades o surgir conflictos sobre cuál debe actuar en cada situación.

Flexibilidad verificable y segura

En un CPD conviven cargas con comportamientos diferentes. Los sistemas transaccionales, de control y comunicaciones esenciales requieren continuidad; determinados entrenamientos, simulaciones, procesos *batch*, informes o copias pueden desplazarse dentro de una ventana.

La flexibilidad puede ser:

- **temporal**, adelantando o retrasando procesos;
- **modular**, reduciendo recursos asignados a tareas no prioritarias;

- **espacial**, trasladando cargas entre ubicaciones;
- **térmica**, aprovechando inercia o almacenamiento de frío;
- **eléctrica**, cargando o descargando baterías;
- **operativa**, priorizando servicios esenciales durante una contingencia.

La capacidad teórica debe probarse. Desplazar una carga exige conservar su estado, cumplir plazos y evitar que el movimiento de datos consuma más recursos de los ahorrados. La continuidad, la seguridad y la vida útil de los equipos prevalecen sobre cualquier ingreso energético.

Almacenamiento y participación en el sistema

Las baterías UPS se destinan principalmente a la continuidad, pero parte de su capacidad puede utilizarse para limitar picos o ajustar la demanda. También pueden complementarse con BESS independientes para no comprometer la reserva de respaldo.

El uso dual exige definir estado de carga mínimo, duración, degradación, garantías, coordinación con la generación, seguridad y prioridad automática de la continuidad. **El almacenamiento no es capacidad gratuita**: cada ciclo afecta a su vida útil y al riesgo de la instalación.

Un CPD puede contribuir al sistema mediante previsiones precisas, reducción de cargas no críticas, desplazamiento hacia periodos con más renovables, almacenamiento y acuerdos de conexión flexible. La Comisión Europea identifica estas infraestructuras como posibles participantes en balance, servicios auxiliares o gestión de congestiones. En España, el Servicio de Respuesta Activa de la Demanda permite reducir potencia a consumidores habilitados que cumplen requisitos técnicos y operativos.

No obstante, la primera contribución consiste en informar con precisión sobre el crecimiento, evitar solicitudes especulativas, mejorar el factor de carga y desplegar por fases.

Energía renovable y oportunidad para España

La contratación anual de energía renovable no garantiza correspondencia horaria. Un PPA puede favorecer nueva generación y estabilizar precios, pero un CPD consume continuamente mientras la producción varía. La estrategia debe considerar cobertura, adicionalidad, coincidencia horaria, proximidad, respaldo y almacenamiento.

España parte de una posición favorable. En 2025, las renovables generaron el 55,5 % de la electricidad nacional, se añadieron 10 GW de potencia y el almacenamiento alcanzó 3.427 MW, según Red Eléctrica. Sin embargo, la abundancia nacional no garantiza capacidad local inmediata y la red debe atender otros procesos de electrificación.

Convertir esta ventaja en capacidad digital exige red, tramitación previsible, despliegues escalonados, almacenamiento, flexibilidad, conectividad, talento e integración territorial. **La energía renovable solo se convierte en ventaja competitiva cuando puede entregarse con calidad, continuidad y dentro del plazo necesario**.

Aragón ilustra la convergencia entre renovables, suelo y atracción de centros de datos. Su Plan Energético 2024-2030 destaca el refuerzo de redes, almacenamiento y generación próxima. La lección es coordinar demanda digital, red, agua, generación limpia y beneficios locales.

Gobernar la energía como capacidad corporativa

La energía debe integrarse en la gobernanza digital. La organización necesita revisar periódicamente potencia disponible y reservada, evolución de cargas, flexibilidad, autonomía, cobertura renovable, ampliaciones de red, inversiones aplazables y riesgos de continuidad.

INDICADOR	QUÉ PERMITE EVALUAR
Potencia disponible, contratada y utilizada	Margen de crecimiento e infrautilización
Factor de carga y precisión de previsiones	Calidad de la planificación
Time-to-power	Alineación energética y tecnológica
Potencia flexible identificada y probada	Capacidad real de respuesta
Potencia, duración y reserva del almacenamiento	Respaldo y gestión de picos
Incidentes y autonomía por nivel de carga	Resiliencia del suministro
Correspondencia horaria renovable	Coincidencia entre consumo y generación limpia
Energía desplazada o participación en servicios	Uso efectivo de la flexibilidad
Capacidad bloqueada por proyectos no ejecutados	Eficiencia en la asignación de potencia

Ningún indicador debe maximizarse aisladamente: más flexibilidad puede reducir el respaldo, aumentar la autonomía puede elevar costes y emisiones, y reservar demasiada potencia puede impedir otros proyectos.

De cuello de botella a activo estratégico

La energía seguirá condicionando el crecimiento digital, pero la respuesta no puede limitarse a solicitar más potencia. El CPD debe evolucionar hacia un consumidor previsible, eficiente y flexible, preservando una distinción esencial: **la continuidad de las cargas críticas no es negociable, pero parte del consumo sí puede gestionarse**.

Un centro de datos no debe medirse solo por la energía que consume, sino por la inteligencia con la que planifica, utiliza y adapta esa demanda. Garantizada la potencia, el siguiente capítulo abordará las arquitecturas de computación, red y almacenamiento necesarias para transformarla en capacidad útil.

6. Nuevas arquitecturas de computación, red y almacenamiento

Diseñar plataformas equilibradas capaces de alimentar, interconectar y gobernar cargas con necesidades muy diferentes sin crear nuevos cuellos de botella

La computación acelerada está transformando los centros de datos, pero no todas las organizaciones ni todos los casos de uso necesitan la misma plataforma. Entrenamiento distribuido, inferencia en tiempo real, aplicaciones RAG, analítica y procesamiento *edge* generan patrones muy distintos.

Reducir esta evolución a comprar GPU produce arquitecturas desequilibradas. Un acelerador permanecerá infrautilizado si la red no coordina los nodos, el almacenamiento no entrega datos con suficiente rapidez o la plataforma no asigna correctamente los recursos. **La arquitectura debe integrar computación, memoria, red, almacenamiento, software, seguridad y operación**, dimensionados según el trabajo útil esperado.

Para energéticas, *utilities*, industrias y administraciones, diseñar toda la infraestructura como si fuera a entrenar modelos fundacionales puede generar inversiones desproporcionadas. La mayoría combinará analítica, inferencia, RAG, automatización y capacidad distribuida.

No existe una única carga de IA

Las necesidades dependen del procesamiento, los datos y el nivel de servicio. Incluso una iniciativa puede utilizar perfiles distintos durante desarrollo, validación y producción.

PERFIL	PRIORIDAD	EXIGENCIA ARQUITECTÓNICA PRINCIPAL
Entrenamiento distribuido	Reducir el tiempo total	Comunicación intensiva entre aceleradores, lectura sostenida y *checkpoints*
Ajuste y experimentación	Flexibilidad	Recursos compartidos, datos versionados y asignaciones temporales
Inferencia *batch*	Maximizar volumen	Alimentación continua y trabajos en cola
Inferencia en tiempo real	Latencia y disponibilidad	Réplicas, acceso rápido a modelos y respuesta previsible
RAG	Calidad y trazabilidad	Conectividad con repositorios, búsqueda e índices actualizados
Analítica sobre activos	Integración con datos operativos	Ingesta, series temporales e históricos
Procesamiento *edge*	Proximidad y autonomía	Recursos locales y sincronización selectiva

Una arquitectura idónea para entrenamiento distribuido puede ser excesiva para inferencia documental; una plataforma corporativa convencional quizá no soporte el tráfico este-oeste de un clúster. Antes de ampliar capacidad debe identificarse si el rendimiento depende de escalar dentro del nodo, entre servidores, en red o en almacenamiento.

El acelerador no elimina el resto del sistema

Las CPU siguen preparando datos, coordinando tareas y ejecutando lógica general. Los aceleradores realizan operaciones paralelas, pero dependen de memoria, interconexión, red y almacenamiento. El equilibrio incluye:

- CPU, GPU y otros aceleradores adecuados a cada trabajo;
- memoria con capacidad y ancho de banda suficientes;
- interconexión entre aceleradores y adaptadores de red;
- DPU o *SmartNIC* para descargar funciones cuando aporte valor;
- almacenamiento local para caché, temporales y *checkpoints*;
- sistemas compartidos para datos, modelos y resultados.

Añadir aceleradores sin revisar estas relaciones puede elevar el coste sin reducir proporcionalmente el tiempo de ejecución. **El objetivo no es maximizar el número de GPU, sino su utilización útil y sostenida dentro del sistema completo**.

Scale-up, *scale-out* y redes de clúster

El s***cale-up*** aumenta la capacidad que opera como un dominio estrechamente acoplado dentro de un nodo, chasis o rack. Tecnologías como NVLink permiten comunicación directa entre GPU, aunque introducen dependencias específicas y límites físicos.

El ***scale-out*** distribuye el trabajo entre servidores, racks o pods. Su rendimiento depende de topología, transporte y congestión. No todas las organizaciones necesitan grandes dominios: para inferencia, ajuste o analítica puede resultar más eficiente una plataforma contenida y fácil de operar.

La red se convierte así en parte del sistema de cómputo. Cuando un nodo espera parámetros, gradientes o estados de otro, los aceleradores permanecen ociosos. Latencia, pérdida de paquetes y congestión afectan al clúster completo.

RDMA permite transferir información entre memorias con menor intervención de CPU. *InfiniBand* integra este mecanismo en una arquitectura de baja latencia, mientras RoCE lo aplica sobre Ethernet. La *especificación 1.0 de Ultra Ethernet Consortium* incorpora transporte y control de congestión orientados a IA y HPC.

La elección no debe plantearse como una oposición genérica entre Ethernet e *InfiniBand*. Debe considerar tamaño del clúster, patrones de comunicación, latencia media y de cola, operación *lossless*, compatibilidad, herramientas, talento, coste y dependencia. Topologías *spine-leaf* o Clos aportan múltiples caminos, pero aumentar la velocidad nominal no corrige un diseño desequilibrado.

Conviene separar el tráfico de computación, almacenamiento, gestión y acceso externo mediante redes físicas o segmentación reforzada. Una transferencia masiva no debería afectar a la coordinación de aceleradores ni al control del clúster.

Alimentar la computación con datos

Una plataforma acelerada solo rinde si recibe información de forma sostenida. El ciclo incluye ingesta, preparación, versionado, distribución, *checkpoints*, registro de modelos, publicación para inferencia y archivado.

La arquitectura puede combinar objetos, ficheros paralelos, NVMe local, repositorios de modelos, bases de datos e índices vectoriales. La jerarquización mantiene los conjuntos activos cerca de la computación y traslada históricos a capas menos costosas.

Medir solo *terabytes* o *petabytes* resulta insuficiente. Deben analizarse tiempo de espera de entrada/salida, *throughput* sostenido, latencia de metadatos y duración de

checkpoints. Estos datos muestran si el almacenamiento alimenta los aceleradores al ritmo necesario.

Una arquitectura técnica para la factoría corporativa

La factoría definida en el capítulo anterior necesita capas coordinadas de fuentes e ingesta, preparación y catálogo, almacenamiento, computación, interconexión, plataforma, servicio, observabilidad y seguridad. No tiene por qué residir en un único clúster: puede utilizar recursos y proveedores distintos si mantiene arquitectura, trazabilidad y control comunes.

El diseño debe permitir sustituir componentes. Si datos, modelos y aplicaciones quedan ligados inseparablemente a una plataforma, la organización reproducirá las dependencias que pretendía evitar.

Orquestar recursos escasos y costosos

Los aceleradores deben tratarse como capacidad compartida, no como equipos reservados permanentemente por cada proyecto. La orquestación debe gestionar colas, prioridades, cuotas, reservas para producción, asignaciones temporales, particionamiento, trabajos distribuidos, recuperación y consumo por proyecto.

Los *schedulers* de HPC están orientados a trabajos *batch* coordinados; las plataformas de contenedores facilitan servicios, API y aplicaciones persistentes. Una arquitectura corporativa puede necesitar ambos modelos. La inferencia, además, debe absorber variaciones y mantener latencia sin reservar siempre el máximo: *batching*, caché y asignación dinámica pueden mejorar la utilización.

Seguridad y resiliencia desde la infraestructura

Los clústeres incorporan firmware, drivers, librerías, imágenes, schedulers, redes e interfaces de gestión de múltiples proveedores. La seguridad debe incluir:

- redes de gestión separadas y protección de interfaces BMC;
- identidades individuales, privilegios mínimos y secretos centralizados;
- control de procedencia para firmware, drivers, imágenes y modelos;
- inventario de dependencias y segmentación entre entornos;
- cifrado proporcionado al riesgo y registro de cambios administrativos.

En entornos IT-OT, las aplicaciones analíticas deben acceder a la información mediante zonas intermedias, API controladas y flujos autorizados, sin conexión directa innecesaria con sistemas operativos críticos.

Los fallos son inevitables en clústeres grandes. *Checkpoints*, reinicio consistente, sustitución de nodos, aislamiento de componentes, réplica de inferencia, redundancia de gestión y conservación externa de modelos permiten recuperar el trabajo. La frecuencia del checkpoint debe equilibrar rendimiento y pérdida potencial. Para servicios críticos también deben existir modos degradados, modelos alternativos o una capacidad básica de continuidad.

Heterogeneidad y modularidad

Durante años convivirán distintas generaciones de CPU, GPU, red y almacenamiento. La arquitectura debe asumirlo mediante interfaces documentadas, entornos reproducibles, catálogos de configuraciones, automatización, observabilidad común, políticas de compatibilidad y módulos ampliables.

Los *diseños abiertos de clústeres de Open Compute Project* buscan reducir configuraciones únicas y facilitar bases reutilizables. Estas referencias reducen riesgo, pero deben adaptarse a las cargas reales: un diseño para gran entrenamiento puede ser desproporcionado para una plataforma de RAG e inferencia.

Las prioridades también varían. Una *utility* puede combinar inferencia y series temporales manteniendo la operación segregada; una industria, *edge*, visión artificial y análisis centralizado; y una infraestructura pública puede valorar más la interoperabilidad y los servicios compartidos que un clúster especializado de baja utilización.

Indicadores para evaluar la arquitectura

INDICADOR	QUÉ PERMITE EVALUAR
Utilización de aceleradores y recursos reservados	Aprovechamiento y sobredimensionamiento
Espera en cola y finalización de trabajos	Capacidad y rendimiento integral
Throughput y latencia de inferencia	Calidad del servicio en producción
Espera por comunicación y latencia de red	Congestión y efecto sobre la computación
Throughput de almacenamiento y checkpoints	Alimentación de datos y resiliencia
Trabajos fallidos o reiniciados	Estabilidad de la plataforma
Componentes reutilizados	Estandarización
Tiempo para incorporar nuevo hardware	Modularidad y evolución
Cobertura de observabilidad	Visibilidad de dependencias y cuellos de botella
Configuraciones fuera de catálogo	Deriva arquitectónica y riesgo

Las métricas deben interpretarse juntas. Una utilización elevada puede indicar eficiencia o falta de capacidad; añadir hardware para reducir colas puede empeorar su aprovechamiento; y maximizar *throughput* puede perjudicar la latencia interactiva.

De la potencia teórica al trabajo útil

Las técnicas de supercomputación deben adoptarse de forma selectiva, en función de problemas reales. Red, almacenamiento, CPU, software y operación determinan cuánto valor puede extraerse de cada acelerador.

La arquitectura no debe dimensionarse por el número de GPU ni por su rendimiento máximo, sino por su capacidad para transformar datos en resultados útiles de forma previsible, resiliente y gobernable. Una vez construida la plataforma, el siguiente capítulo abordará cómo seleccionar modelos, reducir computación innecesaria y medir los recursos consumidos por cada resultado.

7. IA eficiente y sostenible: utilizar el modelo adecuado para cada necesidad

Evolucionar desde la eficiencia del centro de datos hacia una gestión integral de los recursos consumidos por modelos, aplicaciones y servicios

La sostenibilidad de la inteligencia artificial no depende solo de la energía del CPD. Modelo, datos, contexto, frecuencia de uso, software y utilización del hardware condicionan los recursos necesarios.

Dos organizaciones pueden resolver una tarea con consumos muy diferentes: una mediante un gran modelo generalista; otra combinando reglas, búsqueda, modelos especializados y RAG.

La primera pregunta es **qué computación resulta proporcionada al valor y la complejidad de la tarea**. La eficiencia es responsabilidad de negocio, datos, desarrollo, arquitectura, infraestructura y sostenibilidad.

Del PUE al trabajo útil

El PUE continúa siendo relevante para medir la relación entre energía total y consumo IT, pero no indica si la computación genera valor. Un CPD puede mejorar su PUE y aumentar su consumo total, alojar equipos eficientes infrautilizados o ejecutar modelos sobredimensionados.

La evaluación debe recorrer varios niveles:

- **instalación**: energía, PUE, agua, calor recuperado y suministro;
- **plataforma**: utilización de CPU, GPU, memoria, red y almacenamiento;
- **modelo**: energía de entrenamiento, ajuste e inferencia;
- **aplicación**: consultas, latencia, contexto y frecuencia de uso;
- **proceso**: tiempo, calidad y decisiones mejoradas;
- **cartera**: consumo agregado, duplicidades y efecto rebote.

El *marco europeo de reporte para centros de datos* incorpora métricas de energía, potencia, temperatura, agua, renovables y calor reutilizado. El siguiente paso es conectarlas con una unidad funcional —inferencia, documento, predicción o tarea validada— para conocer el consumo por resultado útil.

El modelo más grande no siempre es el adecuado

Los LLM generalistas aportan flexibilidad, pero muchos casos empresariales están acotados. Consultar documentación, clasificar incidencias, extraer datos o aplicar procedimientos puede no requerir máxima escala.

Los *Small Language Models* reducen normalmente memoria y computación por inferencia cuando la tarea está bien definida, aunque su eficiencia real depende también de arquitectura, hardware, contexto y utilización. Combinados con RAG pueden resultar especialmente adecuados para dominios corporativos controlados.

ENFOQUE	APORTACIÓN PRINCIPAL	USO ESPECIALMENTE ADECUADO
LLM generalista	Flexibilidad y tareas abiertas	Razonamiento general y uso multidominio
SLM especializado	Menores requisitos y mayor control	Clasificación, extracción y dominios acotados
Modelo con RAG	Conocimiento actualizable y trazable	Consulta de documentación corporativa
SLM con RAG	Especialización y menor demanda	Asistentes técnicos, normativos o procedimentales
Cascada de modelos	Reserva el modelo mayor para casos complejos	Consultas de dificultad desigual
Reglas o búsqueda	Bajo coste y previsibilidad	Procesos deterministas y consultas estructuradas

La elección debe basarse en pruebas de calidad, latencia y robustez. Si un modelo pequeño cumple los objetivos, otro mayor añade consumo injustificado; si no los alcanza, escalar debe documentarse. La privacidad depende de la arquitectura y los controles, no del tamaño.

Una jerarquía de menor a mayor intensidad

Antes de recurrir a generación avanzada conviene comprobar si el problema puede resolverse con reglas, consultas estructuradas, búsqueda o recuperación sin generación. Después pueden evaluarse un modelo existente, RAG, un SLM especializado, ajuste eficiente, entrenamiento intensivo y, solo cuando exista una necesidad diferencial, un modelo desde cero.

Ascender en esta secuencia debe responder a una limitación demostrada. Si las reglas resuelven el proceso, añadir generación aumenta coste y variabilidad; si RAG aporta el conocimiento necesario, incorporarlo mediante entrenamiento dificulta su actualización y trazabilidad. Mantener conocimiento y lógica fuera del modelo también facilita sustituirlo.

Optimizar antes de ampliar infraestructura

Una vez seleccionado el enfoque, pueden reducirse recursos mediante:

- cuantización, destilación, poda o adaptadores eficientes;
- *batching* y escalado dinámico para mejorar la utilización;
- caché para evitar cálculos repetidos;
- reducción de contexto y límites de salida;
- enrutamiento según la complejidad de la consulta;
- apagado de entornos de desarrollo y prueba sin uso.

Estas técnicas deben evaluarse junto con la calidad. Ahorrar energía carece de valor si aumenta los errores, obliga a repetir consultas o requiere revisión manual constante. Tampoco debe optimizarse solo la latencia: aceptar unos segundos adicionales puede permitir agrupar solicitudes y evitar recursos sobredimensionados.

La inferencia y el software determinan el consumo operativo

El entrenamiento concentra energía durante un periodo limitado, pero una aplicación continua puede acumular mayor demanda de inferencia. El impacto depende de consultas, volumen, respuestas, reintentos, modelo, utilización y vida del servicio.

También deben medirse las respuestas descartadas, duplicadas o mal fundamentadas. Una aplicación aparentemente pequeña puede consumir mucho si se integra automáticamente en numerosos procesos sin comprobar la utilidad de sus resultados.

El software influye mediante llamadas redundantes, contextos extensos, transferencias repetidas, modelos duplicados y recursos ociosos. La especificación *Software Carbon Intensity*, reconocida como ISO/IEC 21031:2024, relaciona energía, intensidad de carbono y emisiones incorporadas con una unidad funcional. Los equipos deben observar consumo, latencia, errores y coste.

Medir energía, rendimiento y calidad

Comparar vatios de forma aislada puede ser engañoso: un sistema más potente puede completar la tarea antes y consumir menos energía total. La medición debe relacionar energía, tiempo, volumen, calidad, utilización, recursos compartidos y repeticiones.

MLPerf Power aplica este principio a entrenamiento e inferencia. Para contratar o comparar plataformas deben utilizarse datos, modelos y niveles de servicio representativos, no solo rendimientos máximos del fabricante.

Flexibilidad energética sin comprometer la continuidad

Las cargas *batch*, simulaciones y ajustes con ventanas amplias pueden programarse cuando exista mayor disponibilidad renovable o menor intensidad de carbono. Si hay ubicaciones equivalentes, puede considerarse también el desplazamiento espacial, siempre que el movimiento de datos, la soberanía y la continuidad no anulen el beneficio.

La programación *carbon-aware* debe combinar previsiones de carbono, renovables, precio y congestión, estado del almacenamiento, plazo, prioridad y coste del traslado. Los servicios en tiempo real y críticos mantienen su prioridad. Además, el ahorro debe verificarse frente a una línea base: una previsión favorable no demuestra por sí sola que se hayan evitado emisiones.

Agua, carbono incorporado y ciclo de vida

La fabricación de aceleradores, servidores, baterías, redes y refrigeración también consume materiales, agua y energía. Sustituir hardware con demasiada frecuencia puede trasladar gran parte de la huella a la cadena de suministro.

La renovación debe considerar fabricación, transporte, materiales críticos, vida útil, ampliación, reutilización, reparabilidad, reciclaje, energía ahorrada y compensación del carbono incorporado. El agua debe diferenciar consumo directo, recirculación, generación eléctrica, semiconductores y cadena de suministro.

Evitar la carrera tecnológica y el efecto rebote

Una nueva generación solo se justifica si mejora el consumo por tarea, la calidad, la consolidación o el valor y si compensa los cambios exigidos en otras capas. La novedad no debe desplazar al análisis del equipamiento sustituido, la dependencia y el retorno energético y económico.

Además, reducir el coste por inferencia puede aumentar el uso total: más asistentes, llamadas automáticas, alternativas nunca revisadas o herramientas equivalentes. **La eficiencia unitaria debe analizarse junto con el consumo agregado y la proporción de resultados utilizados**. Una capacidad más barata necesita mayor gobierno, no menos.

Evaluar el beneficio sistémico

La IA puede ahorrar energía, materiales y emisiones en industria, redes, mantenimiento, logística, movilidad o servicios públicos, pero ese beneficio debe demostrarse. Es

necesario comparar el proceso anterior con el nuevo, descontar el consumo digital y comprobar que la mejora persiste.

Una aplicación intensiva puede generar un gran ahorro neto; otra eficiente, un beneficio marginal. La sostenibilidad depende de recursos, mejora, escala, duración, efecto rebote y alternativas.

Un cuadro de mando orientado al resultado

INDICADOR	NIVEL	QUÉ PERMITE EVALUAR
PUE y WUE	Instalación	Eficiencia energética e hídrica
Energía renovable y calor reutilizado	Suministro	Origen y aprovechamiento energético
Utilización de aceleradores	Plataforma	Uso del hardware desplegado
Energía por trabajo	Plataforma/modelo	Coste energético del procesamiento
Joules por inferencia	Modelo	Consumo unitario
Contexto, caché y resultados utilizados	Aplicación	Sobrecarga y utilidad efectiva
Energía por tarea validada	Proceso	Consumo asociado a un resultado útil
Emisiones por unidad funcional	Servicio	Intensidad climática normalizada
Carbono incorporado amortizado	Ciclo de vida	Impacto de fabricación asignado al uso
Ahorro generado y consumo de cartera	Impacto/gobierno	Beneficio neto y efecto rebote

Las métricas necesitan contexto: modelo, hardware, longitud, latencia y calidad. Cada caso de uso debe llegar a producción con una estimación de demanda, una alternativa menos intensiva, un objetivo de calidad, una métrica de valor y criterios para optimizarlo o retirarlo.

De la eficiencia declarada a la eficiencia gobernada

La IA sostenible no es la que siempre utiliza menos energía, sino la que emplea los recursos estrictamente necesarios para producir un resultado valioso, medible y proporcionado. Este criterio desplaza el foco desde la potencia disponible hacia la disciplina con la que se utiliza.

La transformación también requiere profesionales capaces de elegir modelos, interpretar métricas, cuestionar casos de uso y formar a los usuarios. Esas competencias constituyen el objeto del siguiente capítulo.

8. Talento, organización y capacitación para una transformación permanente

Construir las capacidades necesarias para diseñar, operar y aprovechar infraestructuras cada vez más complejas

Energía, suelo, computación y refrigeración no generan valor por sí solos. La capacidad real depende de quienes diseñan, integran y operan la infraestructura y utilizan sus herramientas.

El talento forma parte de la infraestructura crítica. La tecnología puede adquirirse con rapidez; el conocimiento operativo y del negocio necesita años. Esto es especialmente relevante en energía, industria y servicios públicos, donde los sistemas digitales interactúan con activos físicos y servicios esenciales.

El reto no se resuelve con contrataciones o cursos puntuales. La evolución tecnológica exige aprender, transferir conocimiento y reorganizar responsabilidades de forma continua.

Una brecha de capacidades, no solo de especialistas

La falta de especialistas es solo la parte visible. Una organización puede fracasar si sus equipos trabajan en silos, la dirección desconoce los límites o los usuarios carecen de criterios de seguridad.

NIVEL	NECESIDAD PRINCIPAL	RIESGO SI FALTA
Especialización técnica	Diseñar y operar computación, red, almacenamiento, energía, refrigeración y seguridad	Dependencia de proveedores y dificultades de diagnóstico
Integración multidisciplinar	Coordinar IT, OT, energía, sostenibilidad, datos, ciberseguridad y negocio	Decisiones incompatibles o alejadas de la operación
Dirección y gobierno	Priorizar inversiones, responsabilidades y riesgos	Proyectos inconexos y ausencia de rendición de cuentas
Capacitación del usuario	Utilizar herramientas de forma eficaz y responsable	Baja adopción, exposición de datos y resultados no verificados

Actuar solo sobre uno de estos niveles traslada el cuello de botella a los demás. Formar masivamente a los usuarios servirá de poco sin plataformas fiables y reglas comprensibles; contratar especialistas no resolverá la integración, el gobierno o la adopción.

Especialización técnica con visión de conjunto

La evolución del CPD combina disciplinas antes separadas. Ganan relevancia:

- arquitectura y orquestación de computación acelerada;
- redes y almacenamiento para cargas intensivas;
- energía y refrigeración de alta densidad;
- operación coordinada de IT, OT e instalaciones;
- plataformas de datos y modelos, RAG, MLOps y LLMOps;

- ciberseguridad y resiliencia de infraestructura y cadena de suministro;
- gestión energética y sostenibilidad;
- gobierno del dato y del conocimiento.

El *European Cybersecurity Skills Framework de ENISA*, por ejemplo, ayuda a traducir las necesidades de seguridad en perfiles y competencias.

No todas las organizaciones necesitan internalizar cada especialidad. Pueden combinar equipos propios, proveedores y socios tecnológicos según la criticidad y el control requerido. Lo que debe conservarse internamente es la **capacidad de arquitectura, supervisión y decisión**: externalizar una tarea no equivale a delegar la responsabilidad.

Los especialistas deben comprender las disciplinas adyacentes. El área de infraestructura debe saber cómo entrenamiento e inferencia afectan a capacidad, red y energía; los equipos de datos, cómo soberanía, latencia, seguridad y coste condicionan la ubicación.

Además, el conocimiento concentrado en una sola persona constituye un punto único de fallo. Documentación, revisión entre pares, rotación de responsabilidades y sucesión permiten convertir experiencia individual en capacidad institucional.

Equipos multidisciplinares y organización federada

La colaboración entre infraestructura, energía, datos, seguridad y negocio debe mantenerse durante todo el ciclo de vida. Un modelo federado combina un núcleo común de plataformas y controles con equipos próximos a los procesos sectoriales.

El equipo de plataforma proporciona cómputo, datos, seguridad y despliegue; la función de gobierno mantiene principios reutilizables; los dominios definen necesidades; operación garantiza continuidad; y seguridad y cumplimiento integran salvaguardas.

El centro de excelencia no debe aprobar cada decisión ni acumular proyectos. Su función es proporcionar criterios, componentes y acompañamiento para que los equipos avancen dentro de límites conocidos. Cada servicio necesita un propietario responsable de objetivos, datos, riesgos, costes y resultados.

Recualificar antes de sustituir

Muchas capacidades pueden desarrollarse en los equipos existentes. Quienes conocen el CPD, los activos o los servicios aportan un contexto difícil de reemplazar. La recualificación combina esa experiencia con nuevas competencias.

El proceso debe incluir:

- **mapa de capacidades**, brechas y dependencias críticas;
- **itinerarios por rol** para arquitectos, operadores, seguridad, energía, datos, dirección y usuarios;
- **aprendizaje práctico** mediante laboratorios, simulaciones y proyectos supervisados;
- **mentoría y rotación** entre infraestructura, operación, seguridad y negocio;
- **actualización recurrente** conforme cambien plataformas, riesgos y casos de uso.

Las certificaciones estructuran una base común, pero no sustituyen la experiencia. La evaluación debe comprobar si la persona sabe diagnosticar, interpretar métricas, aplicar controles y detener un proceso cuando supera los límites.

Capacitar al usuario: de la curiosidad al uso responsable

Los empleados necesitan saber cuándo utilizar las herramientas, qué información introducir, cómo valorar las respuestas y cuándo solicitar revisión. De lo contrario, aparecerán baja adopción o usos indiscriminados.

La formación debe adaptarse al rol y cubrir usos autorizados, información permitida, consultas, incertidumbre, verificación, supervisión humana, errores y valor obtenido.

Esta capacitación también es una obligación de gobierno. El artículo 4 del Reglamento europeo de IA exige medidas suficientes de alfabetización para quienes operan o utilizan estos sistemas. La Comisión Europea señala que deben adaptarse al conocimiento previo, el contexto y el riesgo: leer unas instrucciones puede ser insuficiente. **Nadie puede supervisar eficazmente un sistema que no comprende**.

Aprender con IA sin perder criterio

Los asistentes pueden acelerar la consulta de documentación, la preparación de informes, el código y la incorporación de profesionales. También permiten convertir procedimientos, incidencias y lecciones aprendidas en una base de conocimiento accesible.

Su uso formativo exige validar contenidos y preservar la comprensión de los fundamentos. Una respuesta plausible no equivale a conocimiento correcto, y automatizar el aprendizaje no debe impedir que un operador entienda por qué actúa, especialmente ante situaciones imprevistas. La herramienta debe difundir conocimiento institucional, no sustituir el juicio experto.

Colaboración entre educación, tecnología y empresa

El reto requiere colaboración entre empresas, universidades, formación profesional, centros tecnológicos, administraciones y proveedores, trasladada a currículos, laboratorios, prácticas y proyectos reales.

El programa de *Data Center Operations de Northern Virginia Community College* vincula formación práctica y empresas del territorio. Para España, el aprendizaje es asociar los CPD con formación profesional, estudios multidisciplinares, microcredenciales, laboratorios, formación dual e investigación aplicada.

La Comisión Europea mantiene para 2030 el objetivo de alcanzar veinte millones de especialistas TIC. Los contenidos deben evitar una dependencia excesiva de un fabricante para preservar la transferibilidad del conocimiento.

Retener conocimiento para reforzar la soberanía

La soberanía también exige personas capaces de evaluar tecnologías, negociar, operar y recuperar servicios sin depender completamente de terceros. Poseer infraestructura sin saber gobernarla ofrece una autonomía limitada.

La retención requiere carreras técnicas, movilidad interna, participación, tiempo para aprender y comunidades de práctica. La diversidad y la incorporación de nuevas generaciones amplían la base de talento. En territorios con CPD, las capacidades locales convierten inversión en empleo cualificado, investigación y actividad empresarial.

Medir capacidad, no horas de formación

INDICADOR	QUÉ PERMITE EVALUAR
Roles críticos con competencias y suplencias	Resiliencia frente a ausencias y rotación
Cobertura interna de capacidades estratégicas	Dependencia de terceros
Tiempo para habilitar a un profesional	Eficacia del aprendizaje
Participación en prácticas y simulaciones	Exposición a situaciones reales
Incidentes por falta de capacitación	Impacto operativo de la brecha
Adopción y supervisión correcta por colectivo	Uso real y responsable
Movilidad y retención de perfiles críticos	Conservación de experiencia
Participación multidisciplinar	Integración entre tecnología, operación y negocio
Conocimiento documentado y revisado	Conversión de experiencia en activo corporativo

Los indicadores deben interpretarse conjuntamente. Una adopción elevada no es positiva si aumenta los errores; una dependencia externa puede ser eficiente si la organización conserva capacidad para especificar, supervisar y sustituir el servicio.

Una capacidad que debe renovarse permanentemente

La transformación exige especialización, integración multidisciplinar, dirección informada y usuarios capacitados. El objetivo no es que todos dominen la tecnología, sino que **cada persona disponga del conocimiento necesario para ejercer correctamente su responsabilidad**.

El talento debe planificarse con el mismo rigor que capacidad, energía, seguridad y continuidad. Contratar, formar, documentar, transferir y retener conocimiento es un proceso permanente. La automatización no elimina el papel humano: lo desplaza hacia la supervisión, interpretación y decisión, cuestiones que desarrollará el siguiente capítulo.

9. Del CPD monitorizado al CPD autónomo y gobernado

Utilizar analítica y automatización para gestionar infraestructuras más complejas sin perder control, resiliencia ni criterio operativo

Las nuevas cargas aumentan la densidad, el consumo y la complejidad del centro de datos, pero también ofrecen más información para anticipar incidencias y automatizar determinadas decisiones.

El reto es convertir medidas dispersas en una visión coherente del CPD. Un centro autónomo no es una instalación sin personas: es una infraestructura capaz de **observar, interpretar, anticipar y actuar dentro de límites definidos**.

La autonomía debe subordinarse a la gobernanza, especialmente en servicios críticos. Su finalidad es mejorar eficiencia y resiliencia sin introducir más coste, complejidad o riesgo del que resuelve.

Una autonomía gradual y específica para cada función

La autonomía combina instrumentación, datos fiables, integración, reglas y supervisión. No debe asignarse un único nivel al CPD: puede automatizar ajustes térmicos, exigir aprobación para reasignar cargas y reservar las maniobras eléctricas al personal autorizado.

NIVEL	CAPACIDAD	PAPEL DE LAS PERSONAS	CONDICIÓN PARA AVANZAR
0.Manual	Detección y respuesta mediante inspección y procedimientos	Interpretan y ejecutan	Inventario y procedimientos fiables
1.Monitorizado	Métricas, alarmas y cuadros de mando	Analizan varias herramientas	Sensores, datos y umbrales revisados
2.Correlacionado	Relación de eventos y posibles causas	Validan y deciden la respuesta	Modelo común de activos y dependencias
3.Predictivo	Anticipación de degradaciones, incidencias y capacidad	Evalúan recomendaciones	Modelos validados y precisión conocida
4.Supervisado	Acciones predefinidas dentro de límites	Aprueban políticas y excepciones	Trazabilidad, reversión y parada
5.Acotado	Optimización integral dentro de una política	Conservan autoridad y revisión	Evidencia sostenida de seguridad y eficacia

El nivel superior no siempre es el más adecuado. Si una decisión es infrecuente, el proceso no está estabilizado o el error resulta difícil de revertir, mantener intervención humana puede ser la opción más madura. **Saber qué no debe automatizarse también forma parte de la autonomía**.

Una base operativa común antes de incorporar inteligencia

DCIM, BMS, SCADA, ITSM, controladores y orquestadores suelen utilizar identificadores y escalas temporales diferentes. Una alarma térmica solo adquiere contexto al relacionarse con el rack, la carga, la refrigeración, el consumo y el servicio afectado.

La operación avanzada necesita:

- un inventario actualizado de activos físicos y digitales;
- relaciones y dependencias conocidas entre componentes;
- identificadores coherentes y sincronización temporal;
- criterios sobre calidad, frecuencia y conservación del dato;
- registros de cambios, incidencias y actuaciones;
- interfaces controladas entre análisis y ejecución;
- responsables definidos para cada fuente.

No es imprescindible concentrarlo todo en una plataforma, pero sí disponer de un **modelo operativo común** que explique qué representa cada medida, su origen, precisión y alcance. Automatizar sobre información incompleta no elimina el error: lo propaga con mayor rapidez.

De la alarma a la decisión contextualizada

La monitorización indica que una variable ha superado un umbral; la operación avanzada busca explicar la causa y proponer una respuesta. Correlación, detección de anomalías, predicción y análisis posterior permiten reducir ruido, anticipar restricciones y mejorar procedimientos.

Estas capacidades deben mostrar evidencias, incertidumbre y alternativas. El resultado se mide mediante una mejor detección, diagnóstico y recuperación, no por el número de alarmas. Una recomendación crítica que no pueda justificarse debe elevarse para revisión.

Gemelos digitales para ensayar antes de intervenir

Un gemelo digital es una representación dinámica, actualizada con datos operativos. Permite simular nuevos racks, redistribuciones, fallos, configuraciones térmicas o contingencias antes de intervenir. El Departamento de Energía estadounidense ha aplicado este enfoque a la refrigeración eficiente. *U.S. Department of Energy — COOLERCHIPS*

Su utilidad depende de la calibración con datos reales y de la declaración de sus límites. Antes de decidir deben conocerse el ámbito representado, la fecha de actualización, las hipótesis, los escenarios de validación, las variables excluidas y la desviación frente al comportamiento real. **El gemelo no sustituye las pruebas controladas ni las protecciones físicas**; ayuda a seleccionar mejor qué merece la pena probar.

Mantenimiento basado en condición

El análisis de vibraciones, temperaturas, presiones, consumos y calidad eléctrica permite detectar degradaciones y priorizar intervenciones. Puede anticipar fallos, evitar sustituciones prematuras y planificar repuestos.

No sustituye de forma general al mantenimiento preventivo. Las obligaciones regulatorias, las recomendaciones del fabricante y los fallos poco observables siguen exigiendo revisiones programadas. La estrategia más sólida combina mantenimiento preventivo, predictivo y correctivo según criticidad, observabilidad y experiencia, evaluando siempre falsos positivos y falsos negativos.

Automatización de circuito cerrado con límites verificables

Cuando el sistema deja de recomendar y actúa sobre la infraestructura, aumentan tanto el valor potencial como las consecuencias de un error, una mala configuración o un ataque. El régimen de control debe ajustarse al impacto y a la reversibilidad de cada actuación.

ACTUACIÓN	EJEMPLO	CONTROL RECOMENDADO
Observación	Correlacionar eventos o prever capacidad	Ejecución automática y evaluación periódica
Recomendación	Proponer una intervención	Validación humana
Reversible y acotada	Ajustar un parámetro autorizado	Límites, duración máxima y reversión
Con impacto en el servicio	Reasignar cargas o aislar un componente	Aprobación o política explícita para casos conocidos
Sobre sistemas críticos	Maniobras eléctricas o de emergencia	Controles independientes, segregación y autoridad definida

Toda automatización de circuito cerrado debe incorporar:

- condiciones inequívocas de entrada y salida;
- límites de alcance, duración y velocidad de cambio;
- comprobaciones previas y confirmación posterior del resultado;
- reversión automática y parada accesible al operador;
- registro íntegro de datos, decisiones y acciones;
- autorizaciones revisables, modo manual probado y protecciones independientes.

La validación debe avanzar desde el **modo sombra**, sin intervención, hacia recomendaciones aprobadas y, solo con evidencia suficiente, acciones limitadas. La guía NIST SP 800-82 recuerda que la seguridad de los entornos operacionales debe preservar también rendimiento, fiabilidad y protección física. *NIST — Guide to Operational Technology Security*

Proporcionalidad tecnológica y optimización global

La decisión debe partir del problema operativo. Con frecuencia conviene corregir primero la instrumentación o el procedimiento, aplicar después reglas deterministas o análisis estadístico e incorporar modelos predictivos y automatización únicamente cuando aporten una mejora demostrable.

La evaluación debe incluir integración, consumo, datos, validación, dependencia, errores, reversión y valor obtenido. Una regla transparente puede ser preferible a un modelo avanzado, y una recomendación predictiva a la ejecución automática.

Además, la optimización debe ser global. Reducir la temperatura puede elevar el consumo; concentrar cargas puede crear restricciones térmicas; desplazar trabajos puede perjudicar la latencia o los acuerdos de servicio. Las políticas deben priorizar entre continuidad, rendimiento, energía, agua, capacidad térmica, desgaste, coste, emisiones y margen de contingencia, definiendo **límites que ninguna optimización local pueda superar**.

Gobernar los modelos que gobiernan la infraestructura

Cada modelo o automatización debe tener propietario, finalidad y régimen de control. Su expediente operativo debe identificar:

- el problema, los activos afectados y los datos utilizados;
- las condiciones en las que no debe emplearse;

- los indicadores de precisión y utilidad;
- las versiones, pruebas y límites de actuación;
- quién puede modificarlo, detenerlo o autorizar su escalado;
- los procedimientos de revisión, retirada y respuesta a incidentes.

El NIST AI Risk Management Framework articula la gestión del riesgo mediante cuatro funciones: gobernar, contextualizar, medir y gestionar. Es una referencia útil para integrar los modelos en la supervisión general de la organización. *NIST — AI Risk Management Framework*

También debe vigilarse la degradación. Renovar equipos, modificar la refrigeración o incorporar nuevas cargas puede invalidar un modelo antes preciso; la ausencia de errores visibles no demuestra que siga funcionando correctamente.

Ciberseguridad, autoridad humana y contexto crítico

La conexión entre análisis y control amplía la superficie de ataque. Deben separarse observación, decisión y ejecución, aplicarse identidades y privilegios mínimos, segmentarse IT y OT, controlarse versiones y validarse las entradas.

Las pruebas deben cubrir pérdida de comunicaciones, datos contradictorios, sensores comprometidos e indisponibilidad del modelo. Ante información insuficiente, el sistema debe degradarse hacia un estado conocido y controlable.

Las personas reducen tareas repetitivas, pero definen políticas, validan modelos y resuelven excepciones. La supervisión solo es real si el operador comprende la decisión, dispone de tiempo y puede detenerla. El modelo interpreta; la automatización ejecuta; operación fija límites; seguridad controla la exposición; y la dirección acepta el riesgo residual.

Los límites deben adaptarse al contexto. Energía e industria exigen separación respecto a OT y respeto a la autoridad de control; los servicios públicos requieren disponibilidad, trazabilidad y capacidad de revisión; y el *edge* necesita operar durante pérdidas de conectividad. **La criticidad no impide automatizar: exige controles más rigurosos e independientes**.

Indicadores para evaluar la autonomía

DIMENSIÓN	INDICADORES POSIBLES
Detección y respuesta	Tiempos de detección, diagnóstico y recuperación; incidencias anticipadas
Calidad predictiva	Precisión, falsos positivos y negativos; error de previsión
Automatización y fiabilidad	Acciones correctas, excepciones, reversiones, disponibilidad y cambios fallidos
Eficiencia	Mejora normalizada por unidad de servicio o trabajo útil
Datos y gobernanza	Cobertura, retrasos y modelos con propietario, límites y revisión vigente
Seguridad	Intentos bloqueados, privilegios excesivos y tiempo de revocación
Supervisión humana	Intervenciones, anulaciones justificadas y tiempo para decidir
Evolución	Desviación frente al comportamiento real y cambios de contexto

Los indicadores necesitan una línea base anterior a la automatización. También deben registrarse los sistemas detenidos o no desplegados: renunciar a una automatización que no demuestra valor o seguridad puede ser una señal de buena gobernanza.

Autonomía con propósito y bajo control

El CPD autónomo evoluciona desde la visibilidad hacia la anticipación y la actuación. Su valor reside en gestionar una complejidad creciente y coordinar decisiones sobre capacidad, energía, refrigeración y continuidad.

La posición estratégica es inequívoca: **la autonomía es un medio para operar mejor, no un objetivo tecnológico**. Cada función debe recibir el grado adecuado de automatización según su riesgo, utilidad y reversibilidad, preservando responsabilidad, trazabilidad, protección y control humano.

Las organizaciones más avanzadas no serán necesariamente las que automaticen más, sino las que combinen reglas, analítica y actuación automática con criterio. Construir esta capacidad sobre infraestructuras en servicio conduce al siguiente reto: transformar el parque instalado sin comprometer la operación ni desaprovechar inversiones útiles.

10. Transformar el parque instalado sin detener el negocio

Evolucionar los activos existentes mediante modularidad, coexistencia tecnológica y una inversión continua orientada al valor

Buena parte de la infraestructura que soportará las nuevas cargas ya existe. Centros corporativos, salas técnicas y nodos distribuidos deberán seguir operando mientras incorporan mayor densidad, refrigeración avanzada y computación acelerada.

La transformación no consiste en sustituir todo el parque instalado, sino en decidir qué conservar, optimizar, adaptar, reubicar, sustituir o retirar según su utilidad, riesgo y coste futuro. **Preservar la inversión no significa mantener indefinidamente cada activo**, sino aprovechar su valor mientras siga siendo útil.

Tampoco existe un estado final estable. La evolución de las cargas y los modelos de servicio exige pasar de grandes renovaciones aisladas a un proceso permanente de diagnóstico, inversión, validación y ajuste.

El parque instalado como activo, no como obstáculo

Identificar como *legacy* toda infraestructura anterior a la última generación confunde antigüedad con falta de valor. Un CPD existente puede aportar conectividad, permisos, conocimiento, procedimientos y capacidad amortizada.

También puede presentar restricciones eléctricas, estructurales o hidráulicas, tecnologías sin soporte o configuraciones difíciles de intervenir. Deben evitarse tanto el descarte prematuro como la defensa automática de activos inadecuados. La cuestión es **qué función debe desempeñar el CPD en el modelo futuro** y con qué coste, riesgo y flexibilidad.

Comenzar por los servicios y diagnosticar el conjunto

Una modernización centrada en equipos puede ampliar una instalación antes de conocer su función. Cada servicio debe clasificarse por criticidad, continuidad, sensibilidad del dato, latencia, crecimiento, relación con IT y OT y tolerancia a la migración.

Así se evita dimensionar todo el CPD para el requisito extremo: la alta densidad puede concentrarse en zonas específicas y otras cargas permanecer en plataformas convencionales o externas.

El diagnóstico debe integrar infraestructura física, plataformas digitales, operación y economía:

DIMENSIÓN	EVIDENCIA NECESARIA
Servicios y cargas	Catálogo, criticidad, SLA, utilización y escenarios de crecimiento
Energía y refrigeración	Medidas por circuito y zona, redundancia, calidad eléctrica y mapas térmicos
Espacio y estructura	Planos, accesos, resistencia y rutas de instalación
Red y almacenamiento	Latencia, ancho de banda, rendimiento, topología y crecimiento
Operación y seguridad	Procedimientos, incidencias, pruebas, arquitectura de seguridad y cumplimiento
Ciclo de vida	Soporte, repuestos, contratos, dependencias y conocimiento disponible
Economía	CAPEX, OPEX, consumo, mantenimiento, riesgo y coste de transición

La evaluación debe basarse en datos medidos: la potencia total puede ocultar restricciones intermedias y la capacidad térmica no estar disponible donde se necesita. Los escenarios conservador, central y acelerado permiten distinguir decisiones aplazables de las que requieren anticipación.

Conservar, optimizar, adaptar, sustituir o retirar

Cada activo o conjunto funcional debe recibir un tratamiento explícito. Mantener por inercia genera costes ocultos; sustituir por antigüedad puede destruir valor aprovechable.

DECISIÓN	CUÁNDO RESULTA ADECUADA	RESULTADO ESPERADO
Conservar	Cumple su función, mantiene soporte y no limita el servicio	Prolongar su vida útil
Optimizar	Existe margen mediante configuración, operación o consolidación	Liberar capacidad y reducir consumo
Adaptar	La base es válida, pero requiere refuerzos localizados	Incorporar nuevas cargas sin reconstruir
Sustituir	Limita capacidad, fiabilidad, seguridad o eficiencia	Renovar manteniendo el servicio
Reubicar	La carga encaja mejor en otro modelo de despliegue	Evitar una inversión desproporcionada
Retirar	El servicio está duplicado, ha desaparecido o cuesta más de lo que aporta	Eliminar consumo, riesgo y deuda operativa

La decisión debe contemplar dependencias: retirar una aplicación puede eliminar servidores, licencias y redes, mientras un equipo sustituible puede sostener sistemas que aún no pueden migrar. Toda transformación necesita un **plan de desmantelamiento**; modernizar no es añadir capas.

Coexistencia tecnológica y adaptación localizada

Servidores convencionales y acelerados, refrigeración por aire y líquida, redes de distintas generaciones y herramientas heterogéneas convivirán durante años. No es una transición breve, sino una condición estructural que exige diseñar para la diversidad.

La coexistencia requiere capacidad diferenciada por zonas, separación de fallos, herramientas compatibles, seguridad y criterios de retirada. La estandarización debe reducir complejidad sin bloquear la evolución.

Las **islas de alta densidad** concentran refuerzos eléctricos, térmicos y estructurales donde existe demanda. Facilitan probar, aislar incidencias y ampliar gradualmente, pero deben evaluar su impacto sobre potencia, evacuación de calor y respaldo, y aplicar reglas comunes.

Modularidad para conservar opciones

Las grandes ampliaciones pueden inmovilizar capital o llegar tarde. La modularidad incorpora unidades desplegables y ampliables de forma progresiva: zonas eléctricas, circuitos de refrigeración, bloques de racks, almacenamiento o capacidad externa.

Para que sea efectiva necesita:

- interfaces técnicas y compatibilidad entre generaciones;
- reservas de espacio y rutas de conexión;
- pruebas para cada incremento;
- umbrales que activen la siguiente ampliación;
- proveedores, plazos y límites globales identificados.

La modularidad no elimina la planificación: acometidas, obra civil y permisos requieren anticipación. Combina esa visión futura con inversiones parciales ligadas a demanda real y conserva la posibilidad de cambiar de decisión.

Una transformación organizada en ciclos recurrentes

La hoja de ruta no debe ser un trayecto lineal hacia una arquitectura definitiva. Cada ciclo debe:

1. actualizar la línea base de demanda, capacidad, consumo, riesgo y ciclo de vida;
2. identificar restricciones y dependencias;
3. comparar conservación, adaptación, sustitución o cambio de despliegue;
4. priorizar por valor, urgencia, riesgo y reversibilidad;
5. probar en un entorno acotado;
6. desplegar por fases con capacidad de reversión;
7. medir el resultado frente a la línea base;
8. incorporar lo aprendido a estándares y previsiones.

Así se mantiene la dirección estratégica sin ejecutar inversiones que hayan perdido sentido.

Proteger la continuidad durante la intervención

Modificar una instalación activa puede revelar dependencias no documentadas. Antes de intervenir deben identificarse servicios afectados, configuraciones temporales, puntos sin redundancia y condiciones de cancelación.

Cada actuación necesita revisión independiente, procedimiento, ventana acordada, copias verificadas, coordinación y reversión. La puesta en servicio debe comprobar el conjunto en condiciones normales, degradadas y de contingencia y actualizar la documentación.

El marco de ASHRAE, NEMA y *Pacific Northwest National Laboratory* integra planificación, diseño, adaptación, puesta en servicio y operación de instalaciones nuevas y existentes. Esta visión evita tratar el *retrofit* como una suma de sustituciones independientes. *ASHRAE — AI Data Center Energy Performance Framework*

Un caso de negocio basado en riesgo y flexibilidad

El coste de adquisición no basta. Una adaptación barata puede elevar el gasto operativo; una alternativa más costosa puede reducir plazos, mejorar la resiliencia o conservar opciones.

El caso de negocio debe incorporar:

- inversión, operación, mantenimiento y capacidad temporal;
- riesgo de interrupción y coste de las ventanas;
- ciclo de vida, soporte, repuestos y retirada;
- dependencia de proveedores y capacidad de sustitución;
- soberanía, cumplimiento y portabilidad;
- flexibilidad para ampliar, reducir o reubicar;
- valor liberado y coste de no actuar.

La inversión ya realizada explica el contexto, pero las decisiones deben basarse en costes y beneficios futuros. Deben evitarse tanto la permanencia por coste hundido como la sustitución de infraestructura útil por la mera existencia de una alternativa reciente.

En paralelo, conviene priorizar actuaciones de **bajo arrepentimiento**: actualizar inventarios, mejorar la medición, retirar equipos sin uso, documentar dependencias, normalizar interfaces y corregir riesgos inmediatos. Las inversiones vinculadas a previsiones concretas necesitan umbrales y compromisos de demanda.

Compras, eficiencia y ciclo de vida

La contratación debe prever integración y sustitución. Los requisitos han de incluir interfaces y datos utilizables, identidades, soporte, corrección de vulnerabilidades, interoperabilidad, condiciones de salida y asistencia durante la retirada. También debe conocerse la dependencia del proveedor y el conocimiento transferido.

La modernización permite corregir pérdidas y retirar capacidad inactiva. Sus resultados deben compararse con una línea base y normalizarse por trabajo útil, considerando capacidad, consumo, agua, disponibilidad, emisiones y activos retirados.

El Código de Conducta europeo ofrece buenas prácticas voluntarias de eficiencia y sostenibilidad. *Joint Research Centre — EU Code of Conduct for Data Centres* El marco europeo también refuerza las obligaciones de monitorización y reporte para los centros incluidos en su ámbito. *Comisión Europea — Energy performance of data centres*

Gobierno de una cartera continua

La evolución del CPD debe administrarse como una cartera de capacidades y riesgos. Infraestructura, servicios, energía, seguridad, finanzas y sostenibilidad deben compartir previsiones, resolver prioridades y decidir ampliaciones, pausas o retiradas.

Cada actuación necesita un propietario y un responsable del servicio que acepte la migración y el riesgo residual. La dirección debe conocer los servicios habilitados, riesgos mitigados, plazos y flexibilidad conservada.

El ritmo y los controles varían según el contexto. Energía e industria deben coordinarse con OT y ventanas operativas; las administraciones han de preservar continuidad, protección del dato y auditoría; y el *edge* requiere estandarización, gestión remota y funcionamiento durante pérdidas de conectividad.

Indicadores para gobernar la transformación

DIMENSIÓN	INDICADORES POSIBLES
Capacidad y utilización	Recursos útiles incorporados, liberados o consolidados
Continuidad	Interrupciones, ventanas excedidas y reversiones
Ejecución	Desviaciones de plazo, coste, alcance y dependencias
Flexibilidad	Tiempo y coste para ampliar, reducir o reubicar
Ciclo de vida	Activos sin soporte, retirados o con sucesión definida
Economía y eficiencia	Coste total, ahorro verificado y consumo por trabajo útil
Riesgo y soberanía	Restricciones críticas, dependencias únicas, control y portabilidad
Aprendizaje y retirada	Pilotos convertidos en estándares; equipos, contratos y licencias eliminados

Las métricas de ejecución no deben desplazar a las de resultado. Finalizar en plazo carece de valor si la capacidad no puede utilizarse o crea nuevos riesgos. Pausar o reorientar una inversión también puede demostrar buena gestión.

Errores que deben evitarse

Los principales errores son:

- comprar equipamiento antes de clasificar cargas;
- dimensionar toda la instalación para el requisito extremo;
- utilizar capacidades nominales en lugar de medidas por zona;
- modernizar componentes sin analizar dependencias;
- mantener pilotos que nunca se convierten en estándares;
- incorporar plataformas sin plan de integración o salida;
- ampliar sin retirar activos obsoletos o innecesarios;
- evaluar solo la inversión inicial;
- tratar la hoja de ruta como inmutable;
- considerar concluida la transformación tras una renovación.

Transformar de forma continua y con propósito

La próxima generación de CPD combinará activos existentes, módulos especializados y servicios externos dentro de una arquitectura coherente.

La posición estratégica es clara: el **parque instalado debe gestionarse como una plataforma evolutiva**, no como una herencia que conservar intacta ni como un obstáculo que sustituir por completo. La modernización debe proteger el servicio, aprovechar las inversiones útiles, retirar lo que ha perdido valor y conservar opciones ante una evolución tecnológica incierta.

Esto exige diagnóstico, modularidad, disciplina financiera, puesta en servicio y gobernanza. También exige asumir que la transformación permanente no tiene una fecha de finalización. La ventaja corresponderá a quienes evolucionen al ritmo de sus necesidades, sin confundir velocidad con improvisación ni innovación con sustitución constante.

Resultados esperados

La aplicación del decálogo debe traducirse en cambios observables: mayor control sobre los datos y los modelos, infraestructuras mejor adaptadas a cada carga, uso más eficiente de la energía, operaciones más resilientes y una capacidad sostenida para incorporar nuevas tecnologías. El resultado buscado es una organización capaz de **evolucionar su infraestructura digital sin perder soberanía, continuidad ni control sobre sus decisiones**.

El avance dependerá del punto de partida, la criticidad de los servicios y las características del parque instalado. Aun así, pueden identificarse resultados comunes que deberían hacerse visibles cuando las recomendaciones se incorporen a la estrategia, la arquitectura, la contratación y la operación del CPD.

Mayor soberanía sobre el dato, los modelos y el conocimiento

La organización deberá conocer dónde se almacenan y procesan sus datos, qué modelos intervienen en sus servicios, qué conocimiento corporativo utilizan y de qué proveedores e infraestructuras dependen. Esta visibilidad permitirá decidir con mayor criterio qué cargas deben permanecer bajo control directo y cuáles pueden utilizar servicios externos.

Esto permitirá:

- **Clasificar datos, modelos y cargas según su criticidad**, sensibilidad, latencia, rendimiento y requisitos regulatorios.
- Identificar dependencias tecnológicas, contractuales y operativas que puedan limitar la continuidad o la capacidad de decisión.
- Mantener trazabilidad sobre fuentes, versiones, transformaciones y usos del conocimiento corporativo.
- Establecer condiciones de portabilidad, reversibilidad y salida desde el inicio de cada contratación.
- Evitar que información sensible o conocimiento estratégico se traslade a entornos no autorizados.
- Reducir la dependencia de formatos, plataformas, proveedores o especialistas concretos.

La soberanía no implicará concentrar todo en infraestructura propia. El resultado será un modelo híbrido gobernado en el que el CPD corporativo, el *colocation*, el *edge* y la nube se utilicen según las necesidades reales de cada servicio, manteniendo una visión común sobre costes, riesgos y responsabilidades.

Infraestructuras y capacidades de IA alineadas con cada carga

La arquitectura deberá responder al comportamiento de las cargas y no únicamente a tendencias tecnológicas o previsiones agregadas. La organización podrá asignar recursos de computación, red y almacenamiento según las necesidades de entrenamiento, inferencia, análisis, operación transaccional o conservación de datos.

Entre los resultados esperados se encuentran:

- **Mejor correspondencia entre cada carga y la infraestructura utilizada**, evitando sobredimensionamientos y migraciones realizadas por inercia.
- Mayor utilización de servidores, aceleradores, almacenamiento y conectividad disponibles.

- Reducción de cuellos de botella mediante una planificación conjunta de computación, red y datos.
- Arquitecturas capaces de combinar recursos convencionales y acelerados sin fragmentar la operación.
- Selección de modelos de IA proporcionados al problema, utilizando alternativas más pequeñas cuando ofrezcan el resultado necesario.
- Comparación del rendimiento, el coste y el consumo por tarea o resultado útil.
- Capacidad para escalar, reducir o trasladar cargas cuando cambien la demanda, el riesgo o las condiciones económicas.

Las factorías corporativas de conocimiento deberán facilitar la reutilización segura de datos, modelos, componentes y aprendizajes. Esto reducirá duplicidades, acelerará nuevos casos de uso y permitirá aplicar IA privada sobre información sensible con controles comunes de acceso, calidad, evaluación y supervisión.

Mayor capacidad física, eficiencia energética y resiliencia

La incorporación de alta densidad deberá realizarse con una visión integrada de potencia, refrigeración, estructura, red y operación. La capacidad nominal dejará de ser la única referencia: la organización deberá conocer qué recursos están realmente disponibles, en qué zonas y bajo qué condiciones de redundancia y contingencia.

La evolución deberá reflejarse en:

- Mapas actualizados de capacidad eléctrica, térmica, estructural y de conectividad.
- **Despliegues de alta densidad vinculados a demanda confirmada**, mediante zonas especializadas o módulos ampliables.
- Combinación de refrigeración por aire y líquida según densidad, criticidad y viabilidad.
- Mayor visibilidad sobre consumos, temperaturas, pérdidas y márgenes operativos.
- Reducción de capacidad inactiva, equipos sin uso y consumos no asociados a trabajo útil.
- Mejora de la disponibilidad mediante separación de dominios de fallo, reservas adecuadas y pruebas de contingencia.
- Evaluación conjunta de eficiencia energética, uso de agua, emisiones, rendimiento y continuidad.

La relación con el sistema energético deberá ser más activa. Cuando la naturaleza de las cargas lo permita, el CPD podrá adaptar consumos, desplazar trabajos diferibles, utilizar almacenamiento o coordinarse con generación renovable sin comprometer los acuerdos de servicio. La flexibilidad deberá apoyarse en límites operativos, previsiones fiables y capacidad para volver al funcionamiento normal.

La sostenibilidad se convertirá así en una **variable técnica y económica medible**, relacionada con la capacidad útil y el trabajo realizado, no únicamente con reducciones absolutas de consumo.

Operación más inteligente, gobernada y centrada en las personas

La mejora de la instrumentación, la observabilidad y la calidad del dato deberá trasladarse a una operación capaz de detectar antes las degradaciones, relacionar eventos y responder con mayor rapidez. La automatización aportará valor cuando actúe sobre procesos conocidos, dentro de límites verificables y con autoridad humana efectiva.

Los resultados operativos incluirán:

- **Menores tiempos de detección, diagnóstico y recuperación** ante incidencias.
- Reducción del ruido de alarmas mediante correlación contextual, sin ocultar señales críticas.
- Predicciones de capacidad y mantenimiento evaluadas mediante su precisión, falsos positivos y falsos negativos.
- Uso de gemelos digitales para ensayar cambios, ampliaciones y contingencias antes de intervenir.
- Automatizaciones desplegadas progresivamente desde el modo sombra hasta acciones reversibles y acotadas.
- Registro íntegro de datos, recomendaciones, decisiones, acciones y reversiones.
- Modos manuales, mecanismos de parada y protecciones independientes comprobados periódicamente.

Los equipos deberán asumir funciones más orientadas a definir políticas, validar modelos, interpretar excepciones y coordinar tecnología e infraestructura física. La organización contará con propietarios claros para servicios, activos, datos y automatizaciones, y reducirá la dependencia de personas concretas mediante documentación, formación, procedimientos y transferencia de conocimiento.

La supervisión humana será efectiva cuando el operador comprenda la situación, disponga de tiempo para intervenir y tenga autoridad real para detener una actuación. La automatización no eliminará la responsabilidad: permitirá concentrarla allí donde aporta más valor.

Una transformación continua que preserve el valor de la inversión

El parque instalado deberá gestionarse como una cartera de capacidades, riesgos y opciones. Cada activo tendrá una función definida y una decisión asociada: conservar, optimizar, adaptar, sustituir, reubicar o retirar. Esto permitirá aprovechar infraestructuras útiles sin prolongar dependencias que hayan dejado de ser sostenibles.

Los resultados esperados serán:

- Inventarios y líneas de base actualizados para orientar decisiones de capacidad e inversión.
- **Modernizaciones modulares y por fases**, vinculadas a umbrales de demanda y resultados verificables.
- Coexistencia ordenada entre tecnologías y generaciones diferentes.
- Pilotos con criterios claros de validación, ampliación, integración o retirada.
- Intervenciones realizadas con ventanas, protecciones y procedimientos de reversión.
- Contratos que contemplen integración, soporte, actualización, portabilidad y salida.
- Retirada efectiva de equipos, licencias y servicios que hayan dejado de aportar valor.
- Casos de negocio que incorporen coste total, riesgo, flexibilidad y coste de no actuar.

Las actuaciones de bajo arrepentimiento —mejorar la medición, documentar dependencias, corregir riesgos, liberar capacidad o preparar interfaces— permitirán avanzar incluso cuando las previsiones de crecimiento sean inciertas. Las inversiones de mayor alcance se activarán cuando exista demanda suficiente y se hayan verificado sus condiciones técnicas y económicas.

La gobernanza de la cartera facilitará coordinar prioridades entre infraestructura, tecnología, energía, seguridad, continuidad, finanzas y negocio. También permitirá pausar o reorientar decisiones cuando cambie el contexto, sin interpretar esa adaptación como un fracaso de la planificación.

En conjunto, el avance será visible cuando la organización pueda demostrar que controla mejor sus datos y modelos, asigna cada carga al entorno adecuado, obtiene más rendimiento de la energía y de los activos disponibles, responde con mayor rapidez ante incidencias y adapta su infraestructura sin detener los servicios. El resultado esperado es un CPD **soberano, eficiente, resiliente y evolutivo**, capaz de convertir la transformación tecnológica en competitividad, continuidad y sostenibilidad.

Anexo: Marcos normativos, certificaciones y *frameworks* de referencia

La siguiente selección reúne las referencias con mayor aplicación al diseño, operación, transformación y gobierno de centros de datos, con especial atención al contexto europeo y español.

- **Directiva (UE) 2023/1791 de eficiencia energética y Reglamento Delegado (UE) 2024/1364** (normativa europea). Establecen obligaciones de información para los centros de datos sujetos a su ámbito —a partir de 500 kW de demanda de potencia IT instalada— y una metodología común para medir energía, agua, renovables, calor residual y refrigeración. Directiva 2023/1791 · Reglamento 2024/1364
- **Reglamento Europeo de Inteligencia Artificial — *AI Act* (UE) 2024/1689** (normativa europea). Introduce obligaciones proporcionales al riesgo y al papel de cada organización en materia de gobierno, documentación, transparencia, supervisión humana, modelos de propósito general y sistemas de IA de alto riesgo. Más información
- **Directiva NIS2 (UE) 2022/2555** (normativa europea). Refuerza la gestión del riesgo de ciberseguridad, la continuidad, la seguridad de la cadena de suministro y la notificación de incidentes para sectores esenciales y proveedores de infraestructuras digitales, incluidos los servicios de centros de datos cuando resulten aplicables sus criterios. Más información
- **Reglamento General de Protección de Datos y Reglamento Europeo de Datos** (normativa europea). El RGPD protege los datos personales, mientras el *Data Act* refuerza la portabilidad, la interoperabilidad y el cambio de proveedor en servicios de tratamiento de datos y nube, aspectos esenciales para la soberanía y las arquitecturas híbridas. RGPD · Data Act
- **Esquema Nacional de Seguridad — Real Decreto 311/2022** (normativa española). Define los principios, requisitos mínimos y medidas de seguridad aplicables a los sistemas del sector público español y a los proveedores comprendidos en su ámbito, incluyendo disponibilidad, trazabilidad, continuidad y evaluación de conformidad. Más información
- **Serie EN 50600 / ISO/IEC 22237** (estándares de infraestructura). Proporciona un marco integral para diseñar y operar instalaciones de centros de datos considerando ubicación, construcción, energía, control ambiental, cableado, seguridad, disponibilidad y eficiencia. ISO/IEC 22237-1
- **ISO/IEC 27001 e ISO 22301** (normas certificables). Permiten implantar sistemas de gestión coordinados para proteger la información y asegurar la continuidad y recuperación de los servicios ante interrupciones. ISO/IEC 27001 · ISO 22301
- **ISO 50001:2018** (norma certificable). Establece un sistema de gestión de la energía basado en indicadores, líneas base y mejora continua para controlar consumos, costes y rendimiento energético. Más información
- **ISO/IEC 42001:2023** (norma certificable). Define los requisitos de un sistema de gestión de la inteligencia artificial para gobernar sus riesgos, responsabilidades, controles y mejora continua durante todo el ciclo de vida. Más información
- **Serie ISO/IEC 30134** (estándares de indicadores). Normaliza la medición de la eficiencia de los centros de datos mediante indicadores comparables, entre ellos el PUE, facilitando líneas base, seguimiento y comunicación homogénea de resultados. ISO/IEC 30134-2:2026 — PUE
- ***Uptime Institute Tier Standard* y *Tier Certification*** (certificación de centros de datos). Clasifica y certifica la topología y las capacidades operativas de la infraestructura mediante cuatro niveles, que deben seleccionarse de acuerdo con la criticidad y las necesidades reales del negocio. Más información
- ***EU Code of Conduct for Data Centres Energy Efficiency*** (marco voluntario europeo). Recopila buenas prácticas técnicas y operativas para reducir el consumo energético y mejorar la sostenibilidad de los centros de datos nuevos y existentes. Más información
- ***NIST AI Risk Management Framework — AI RMF*** (framework voluntario). Propone integrar la confianza y la gestión del riesgo de la IA mediante cuatro funciones —gobernar, contextualizar, medir y gestionar— aplicables al diseño, despliegue y supervisión de modelos. Más información
- ***NIST SP* 800-82 Rev. 3 — Guide to Operational Technology Security** (guía de ciberseguridad). Orienta la protección de sistemas OT, BMS, SCADA y otras plataformas de control respetando sus requisitos específicos de seguridad física, disponibilidad, fiabilidad y rendimiento. Más información

- **AI Data Center Energy Performance Framework** (framework técnico). El marco elaborado por ASHRAE, NEMA y *Pacific Northwest National Laboratory* aborda conjuntamente planificación, diseño, construcción, puesta en servicio, operación y adaptación de centros de datos para IA, incluyendo eficiencia energética, agua, flexibilidad y resiliencia. Más información

La aplicabilidad jurídica de cada referencia debe comprobarse según el país, el sector, la potencia instalada, el tamaño de la organización y su función como propietario, operador, proveedor o usuario; los estándares y marcos voluntarios no sustituyen las obligaciones regulatorias.

Informes El Reto 2026

Cinco perspectivas sobre energía, inteligencia y decisiones

En 2026, la expansión de la IA y de la capacidad de cómputo coincide con la electrificación de la industria y la movilidad, la integración de renovables y almacenamiento y la llegada de grandes consumidores a redes que también deben atender la demanda existente. La potencia disponible y los plazos de conexión condicionan proyectos digitales e industriales. La crisis de suministro vinculada al estrecho de Ormuz vuelve a situar la seguridad energética y la exposición geopolítica en primer plano; los riesgos ciberfísicos amplían las exigencias de continuidad en servicios cada vez más digitalizados.

Para enerTIC.org, que cumple quince años, esta coyuntura confirma la necesidad de analizar conjuntamente la transición energética y digital. Una nueva carga de IA exige decisiones sobre arquitectura, potencia, refrigeración y gobierno del dato; su suministro requiere previsiones y capacidad de red; su valor depende después de cómo se utilice para operar activos, mejorar procesos industriales o prestar servicios urbanos. Las compañías energéticas y *utilities* están en el centro de esta relación, junto con los grandes consumidores y los territorios que intervienen en la demanda.

Una relación de doble sentido

La infraestructura digital necesita electricidad y una gestión responsable de su energía y refrigeración. Al mismo tiempo, el sistema energético necesita infraestructura digital para observar activos dispersos, anticipar congestiones y coordinar generación, almacenamiento y demanda. Los informes de CPDs y Energía IT abordan ambos lados de este vínculo; el de Energía OT lo traslada a funciones físicas que deben mantenerse incluso cuando falla una conexión o un proveedor.

El resultado se comprueba fuera de la plataforma

Un caso de IA industrial necesita datos de energía y producción con significado común, acceso seguro a OT y una decisión sobre cuándo consumir electricidad. Una plataforma urbana requiere capacidades semejantes a otra escala, además de responsables, mantenimiento y presupuesto para prestar el servicio. Por eso la colección relaciona consumo por tarea útil, capacidad de red, energía por unidad producida, continuidad de los activos y calidad de los servicios urbanos.

La pregunta que comparten los cinco informes es qué información, capacidades y decisiones permiten mejorar un resultado sin trasladar costes o riesgos a otra parte del sistema.

Infraestructuras IT · El regreso del CPD corporativo

El primer informe parte de una decisión que ya no es puramente tecnológica: dónde procesar cada carga y qué control conservar sobre datos, modelos y conocimiento generado. Propone elegir entre CPD propio, *colocation*, *edge* y nube híbrida según sensibilidad,

latencia, coste y continuidad. La IA privada y las factorías de conocimiento pueden aprovechar información corporativa si existe gobierno de sus usos, trazabilidad y capacidad de trasladar el servicio o cambiar de proveedor.

Las nuevas cargas elevan la densidad de potencia, el tráfico interno y las necesidades térmicas. El decálogo plantea diseñar conjuntamente computación, red, almacenamiento y refrigeración híbrida; medir también el consumo por trabajo útil; y modernizar el parque instalado por fases. Examina además cuándo un CPD puede desplazar cargas, utilizar almacenamiento o participar en la flexibilidad del sistema eléctrico sin poner en riesgo la continuidad.

Energía IT · El dato como infraestructura crítica

El segundo informe estudia la dependencia inversa: generación, redes, almacenamiento y nuevos consumos precisan información para planificar, operar y coordinarse. Plantea identificar los datos esenciales, conocer su origen y sus transformaciones, asignar responsables y preservar la calidad con criterios ajustados a cada decisión. La interoperabilidad IT/OT debe conservar el significado de cada activo y señal, además de proteger los límites de seguridad de la operación.

Sobre esa base, la analítica y la IA pueden anticipar demanda, congestiones e incidencias, evaluar capacidad de conexión y estimar flexibilidad realmente disponible. El informe incorpora observabilidad ampliada, ciberresiliencia, soberanía tecnológica y una distinción necesaria entre datos observados, inferidos y sintéticos para no presentar incertidumbre como certeza operacional.

La conexión es concreta: el operador del CPD necesita potencia y condiciones de red verificables; la *utility* necesita previsiones fiables de la nueva demanda digital y conocer qué cargas pueden ser flexibles, dónde y durante cuánto tiempo.

Energía OT · Inteligencia sobre activos físicos

Este decálogo parte de redes, plantas, refinerías, instalaciones gasistas y otros activos de larga vida donde una recomendación puede afectar a personas, equipos y continuidad. Propone mapear funciones críticas y modernizar sistemas heredados mediante instrumentación, virtualización y cambios reversibles. El mantenimiento predictivo y prescriptivo, la simulación y el *edge* aportan contexto y respuesta cerca del activo; su autonomía debe ajustarse a límites físicos, calidad del dato y capacidad humana para intervenir. Los *smart meters* abren nuevas opciones de gestión de la demanda con beneficios que han de verificarse.

Industria · Energía, productividad y retorno

El informe industrial vincula la IA a casos con retorno y al rediseño de procesos. Un consumo aislado dice poco: debe relacionarse con producto, receta, estado de la línea y unidades válidas para conocer su efecto sobre coste, calidad y EBITDA. La integración IT/OT y la modernización selectiva de planta permiten actuar sin comprometer producción. Autoconsumo, almacenamiento y flexibilidad se dimensionan según perfiles y restricciones reales; la ciberseguridad y el control del know-how protegen el conocimiento que diferencia a la empresa.

Smart Cities · Datos útiles y capacidad de sostener servicios

El informe urbano identifica una brecha entre sensores y plataformas desplegados y decisiones que mejoran la ciudad. Exige calidad y contexto del dato, una Oficina del Dato con responsabilidades, interoperabilidad y espacios de confianza para coordinar energía, agua, movilidad, edificios y otros servicios. La ciudad de energía Net Zero incorpora generación local y flexibilidad; cada proyecto necesita además mantenimiento, presupuesto operativo y una hoja de ruta proporcionada a la madurez municipal.

Industria y territorios forman parte de la demanda que Energía IT debe prever y Energía OT debe poder gestionar. Su flexibilidad solo es útil cuando se conoce su localización, sus límites de proceso y su disponibilidad real; la infraestructura digital debe aportar esos datos sin añadir dependencias difíciles de sostener.

La colección completa

Cinco decálogos independientes y complementarios, elaborados desde los Comités de Expertos de enerTIC.org.



Infraestructuras IT:
Cargas IA · potencia · soberanía



Energía IT:
Dato crítico · red · flexibilidad



Energía OT:
Activos · edge · continuidad



Industria:
Proceso · energía · retorno



Smart Cities:
Dato fiable · servicios · territorio

Acerca de enerTIC

2011–2026 · 15 años impulsando la eficiencia energética desde la tecnología

La Plataforma enerTIC.org reúne a empresas, profesionales e instituciones para desarrollar el potencial de las tecnologías digitales en la mejora de la eficiencia energética. Su misión relaciona esa mejora con la sostenibilidad, la resiliencia y la competitividad: tres objetivos que hoy deben abordarse de forma conjunta en las decisiones sobre infraestructuras, operaciones e inversión.

En 2026, esa misión adquiere una dimensión concreta. La electrificación incrementa las exigencias sobre las redes; la expansión de la inteligencia artificial y los centros de datos añade nuevas demandas de energía y capacidad; y la incertidumbre geopolítica obliga a prestar más atención a la seguridad de suministro y a la continuidad de los servicios. Sensores, analítica, automatización o gemelos digitales pueden ayudar a responder, pero su utilidad depende de la calidad de los datos, de su integración con los sistemas existentes y de resultados que puedan medirse.

enerTIC trabaja en el punto de encuentro entre esas necesidades y las soluciones disponibles. El Consejo Consultivo aporta una visión estratégica; la Red de Colaboración Institucional amplía el diálogo; y cinco Comités de Expertos examinan los retos de los centros de datos, las infraestructuras IT y las operaciones del sector energético, la industria y los territorios inteligentes. Ese trabajo toma forma en informes, casos de uso y encuentros donde se contrastan experiencias y se identifican oportunidades de colaboración. El Smart Energy Congress reúne cada año esas perspectivas en una conversación común.

El 15° aniversario invita a mirar hacia la siguiente etapa. La eficiencia de un centro de datos depende también de su acceso a la energía; una red más inteligente necesita coordinar IT y OT; y una industria o una ciudad capaces de gestionar mejor su demanda pueden contribuir a un sistema más flexible. enerTIC impulsa esa visión conectada para que la digitalización se traduzca en decisiones verificables sobre consumo, emisiones, seguridad operativa y retorno de la inversión.

Empresa asociadas

