



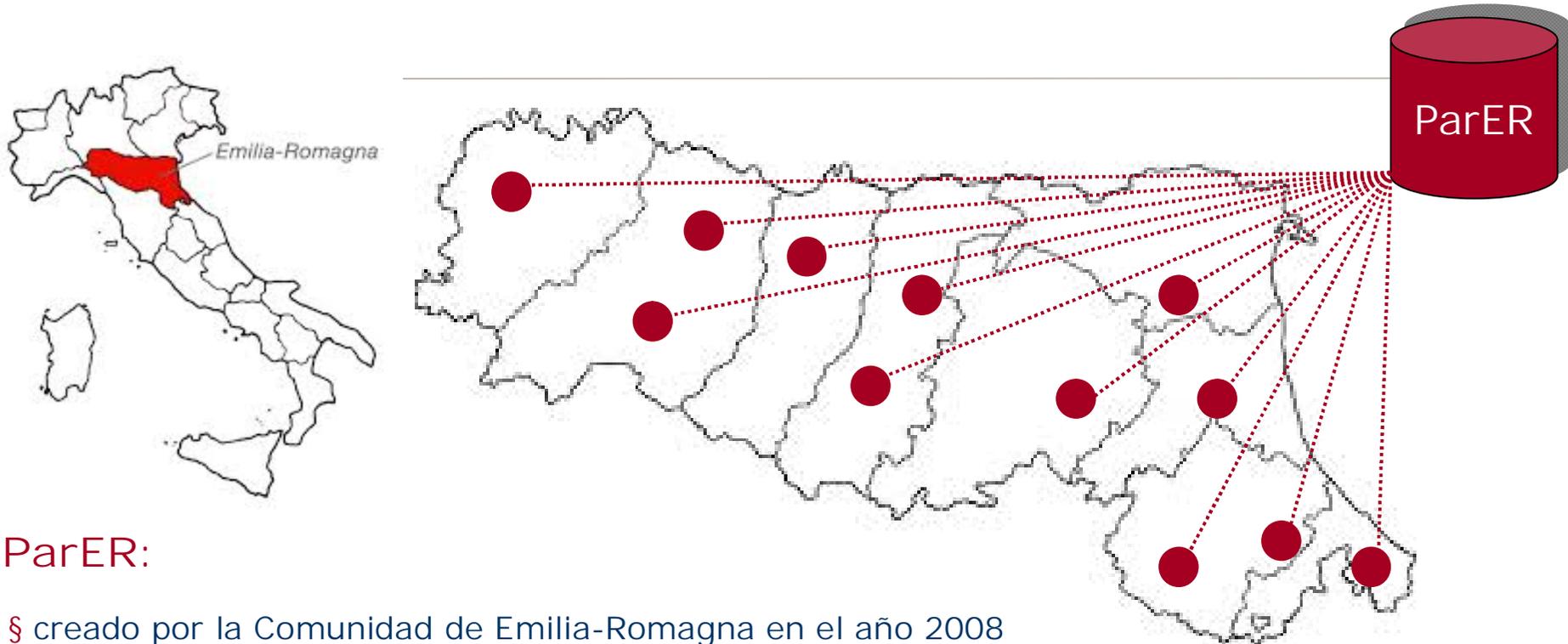
La experiencia en la preservación digital de un custodio público regional:

- § modelos de referencia utilizados
- § políticas desarrolladas
- § cuestiones críticas emergentes

Alessandro Alfier,

Bogotá, Archivo General de la Nación Colombia, 4 de diciembre de 2014

¿Qué es el "Polo archivístico dell'Emilia-Romagna"?



ParER:

§ creado por la Comunidad de Emilia-Romagna en el año 2008

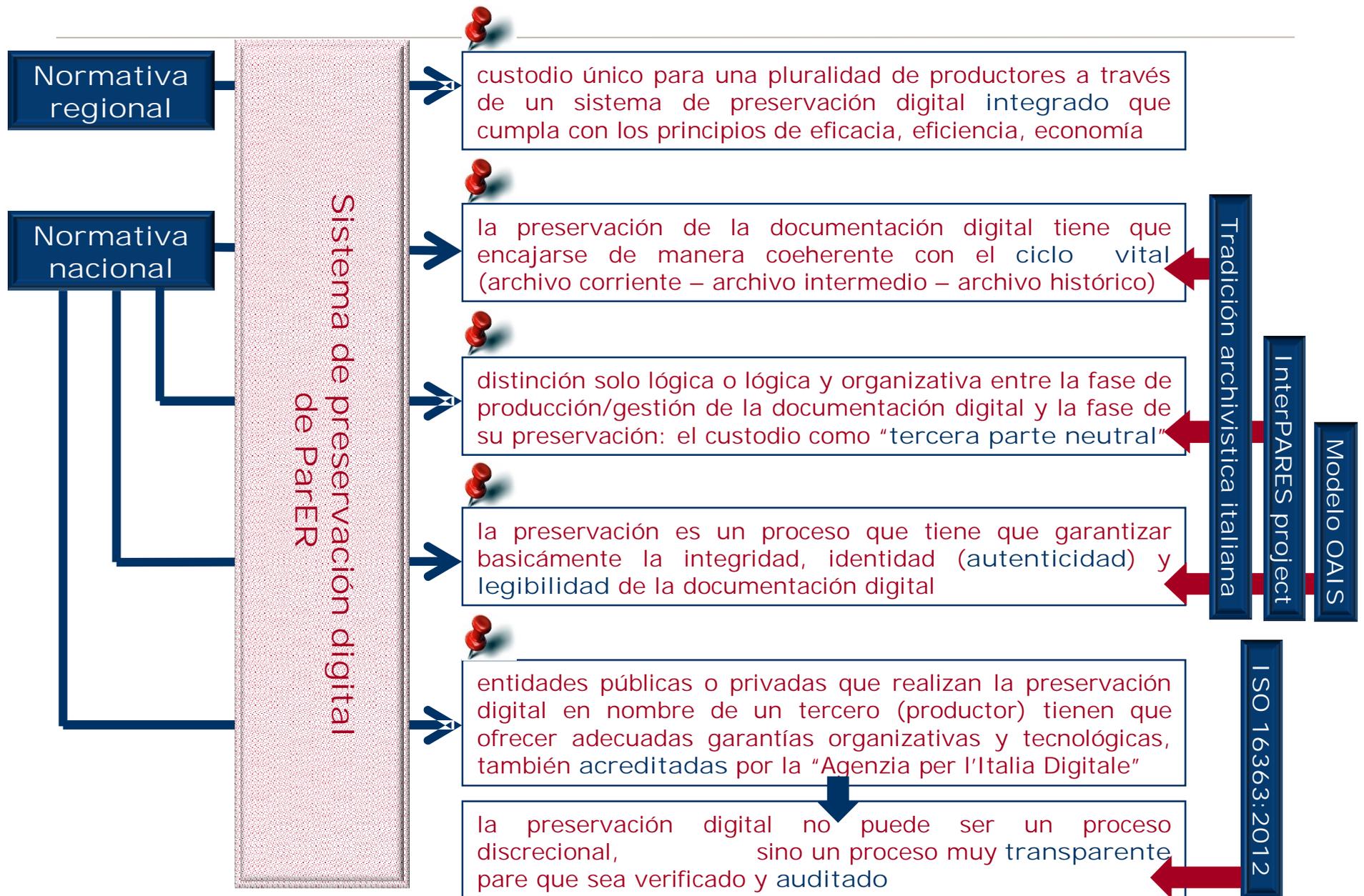
§ como archivo único de concentración al servicio de múltiples administraciones públicas

§ para la preservación a largo plazo de la documentación digital

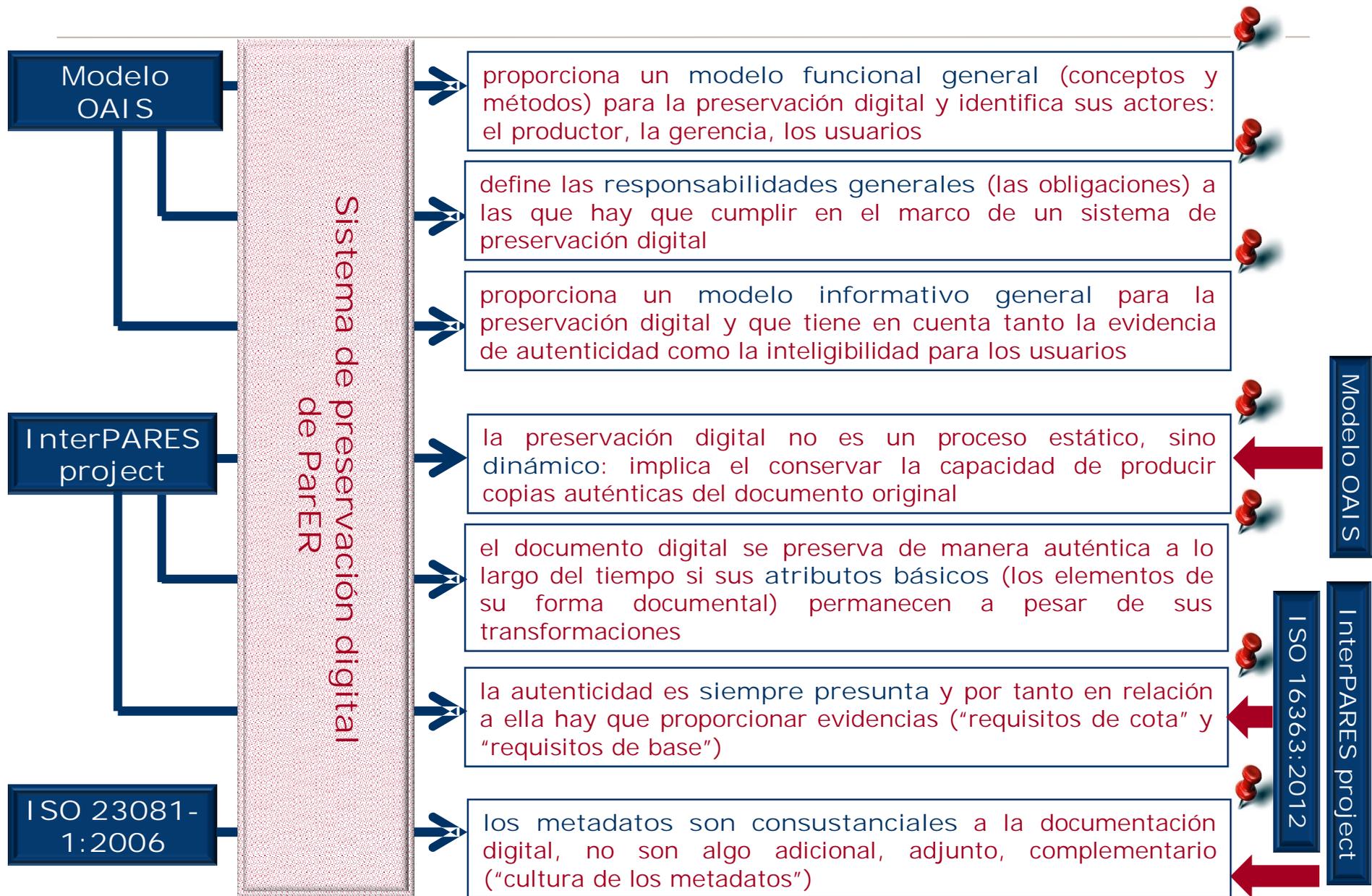
se conservan más de
26 millones
de documentos digitales

- la Comunidad de Emilia-Romagna y sus agencias
- provincias
- ayuntamientos
- universidades
- instituciones sanitarias

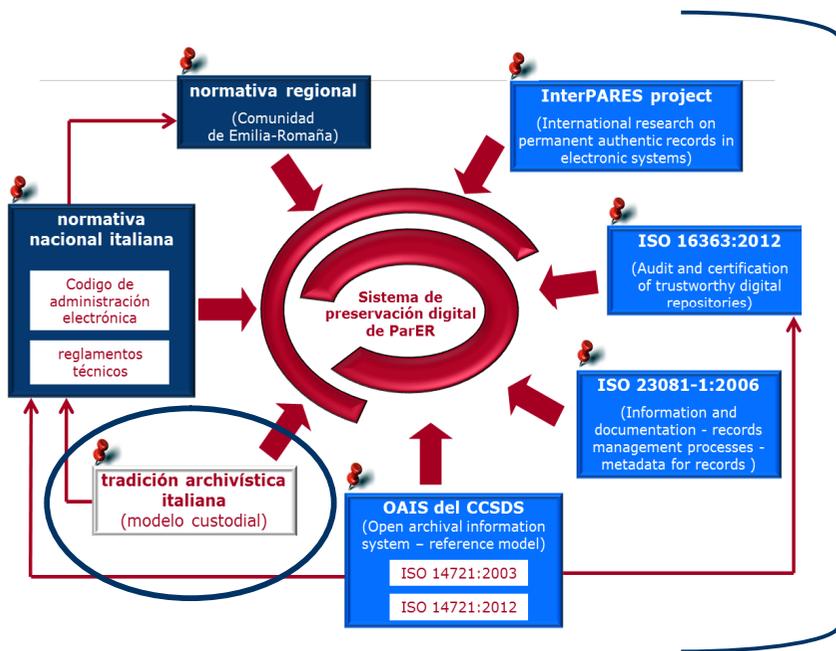
Las distintas aportaciones de los modelos de referencia utilizados por ParER (1)



Las distintas aportaciones de los modelos de referencia utilizados por ParER (2)



Los modelos de referencia utilizados por ParER: ¿en búsqueda de un paradigma archivístico renovado?



¿Los modelos de referencia utilizados son todos coherentes o compatibles con el **paradigma archivístico custodial** (tradición archivística italiana)?



Primera serie de cuestiones críticas:

¿El paradigma custodial sigue siendo exactamente lo mismo en la transición de los archivos analógicos a los archivos digitales?

es decir

¿Las diferencias entre la documentación analógica y la digital no afectan de alguna manera al funcionamiento concreto del paradigma custodial?



El ParER dentro el paradigma custodial: la visión retrospectiva



FINALIDAD:

preservar la autenticidad presunta de la documentación digital que se le confía. Por lo tanto tutelar los documentos en su naturaleza de evidencia y de instrumento de accountability



NATURALEZA:

sujeto de naturaleza pública dotado de plena autonomía, capaz de obrar como tercera parte neutral respecto a los productores. Como tal, puede garantizar un proceso de conservación transparente, exento de cualquier posible conflicto de interés



METODOLOGÍA:

preserva la autenticidad presunta de los documentos digitales dentro de dos horizontes metodológicos: el de la seguridad, para asegurar la integridad de la documentación conservada, y el de la estabilidad, para garantizar su completa identidad

se concreta cuando el documento sobrepasa el

UMBRAL ARCHIVÍSTICO

y se puede definir para ello un **CONTEXTO COMPLETO** (definido e inmutable)

CICLO VITAL

- archivo corriente
- archivo intermedio
- archivo histórico



PARADIGMA CUSTODIAL
(InterPARES)

mirada al documento sólo cuando ha sido ya producido, gestionado, usado:

visión retrospectiva

El ParER

entre la visión retrospectiva y la visión prospectiva (1)

relación de ParER con su territorio circunstante

DEPÓSITO "ANTICIPADO"

VISIÓN DE LOS RECORDS MANAGERS

Depósito "anticipado": ¿qué es?

al ParER se le confían documentos digitales aún activos, es decir mucho antes de que hayan perdido su estricta utilidad para las actividades del productor



ParER comunica al productor los episodios que invalidan ab origine la naturaleza de evidencia del documento, episodios no evidenciados por los sistemas de gestión documental, por ejemplo:

Depósito "anticipado": ¿para qué?

permite al productor disfrutar, desde un principio, de los recursos del ParER en tema de seguridad de la integridad de las cadenas de bits y de los certificados de firma digital que presentan una caducidad temporal natural



§ la alteración criptográfica de la cadena de bits de los ficheros

§ el recurso a formatos de ficheros que no garantizan la inmodificabilidad del contenido

Depósito "anticipado": ¿a qué sirve?

ofrece a la documentación producida desde hace poco un umbral mínimo de protección, aligerando las responsabilidades que de otra forma recaerían por entero en el marco del productor



§ la aplicación de firmas digitales sucesivamente a la caducidad natural o revocación de los certificados de firma

El ParER

entre la visión retrospectiva y la visión prospectiva (2)

Relación de ParER con su territorio circunstante

DEPÓSITO "ANTICIPADO"

VISIÓN DE LOS RECORDS MANAGERS

los productores implican al ParER con continuidad sobre aspectos decisivos de la producción y gestión de la documentación digital, por ejemplo:

§ definición correcta de las formas documentales

§ correcta generación y codificación de los metadatos de contexto

§ correcta gestión documental en los sistemas de recordkeeping

← dificultades para distinguir entre sistemas informativos y sistemas de gestión documental

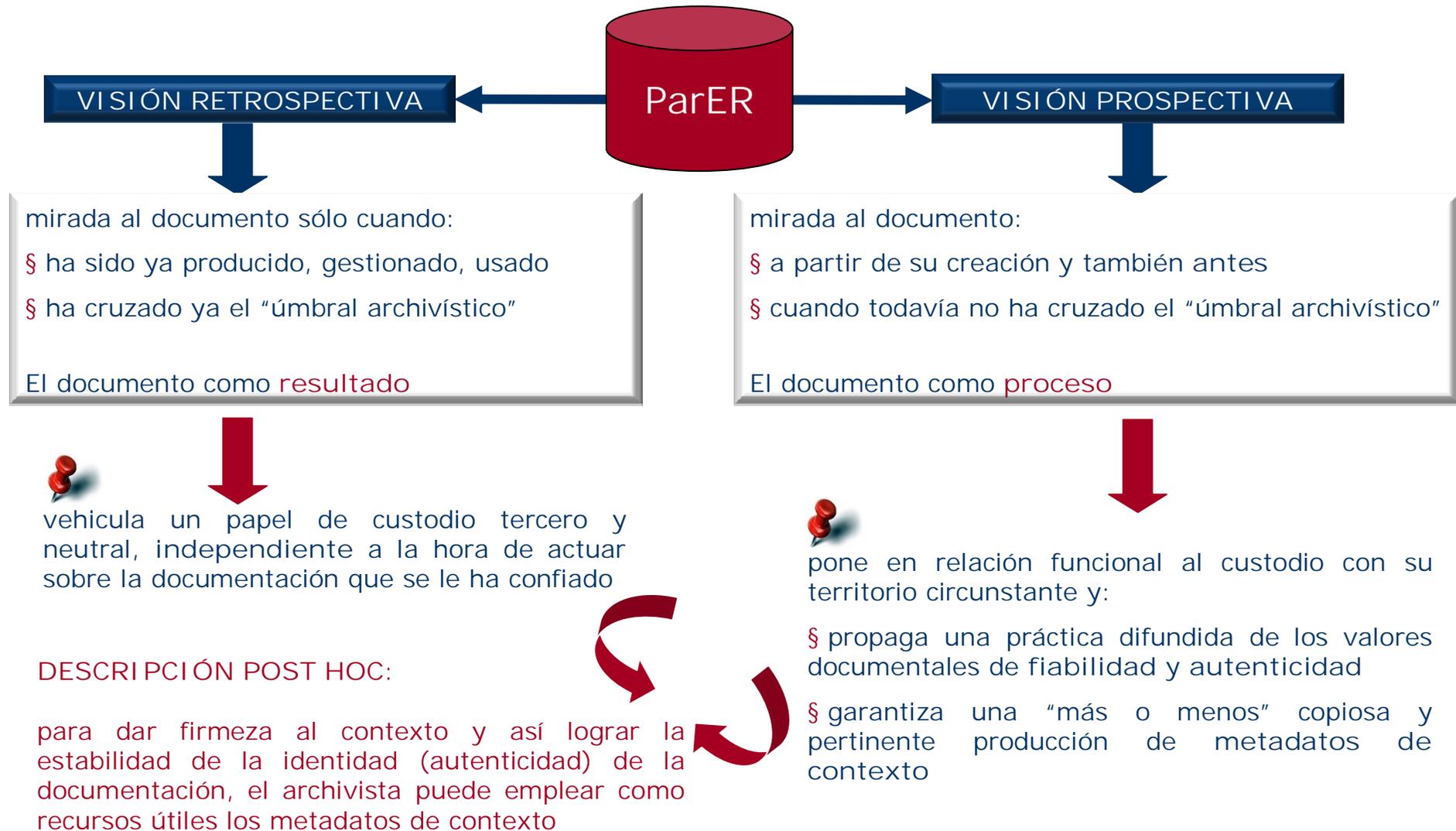
← dificultades para distinguir entre sets de datos y documentos

← dificultades para comprender las diferencias entre forma documental, contexto, contenido

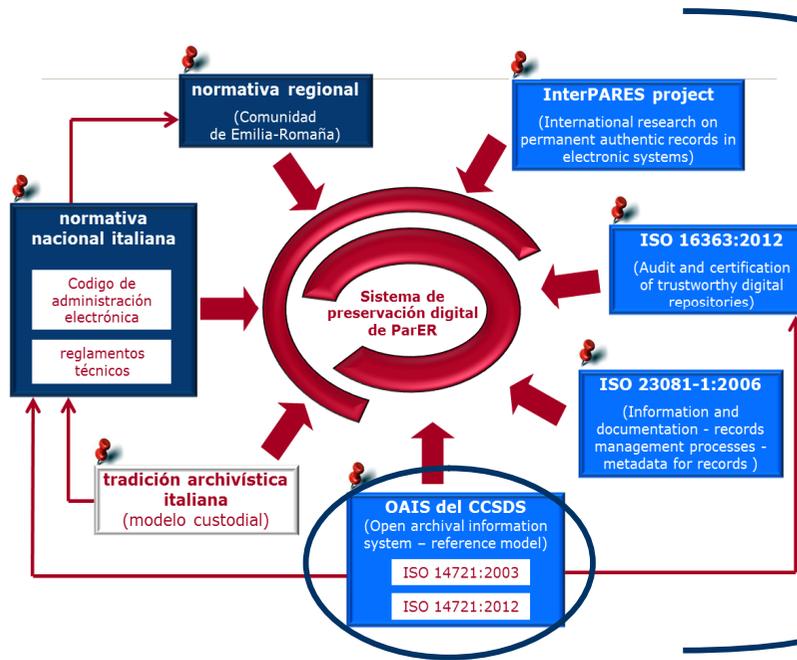
← presunción de la "libertad" de estructura del documento digital, es decir, de su forma documental

El ParER

entre la visión retrospectiva y la visión prospectiva (3)



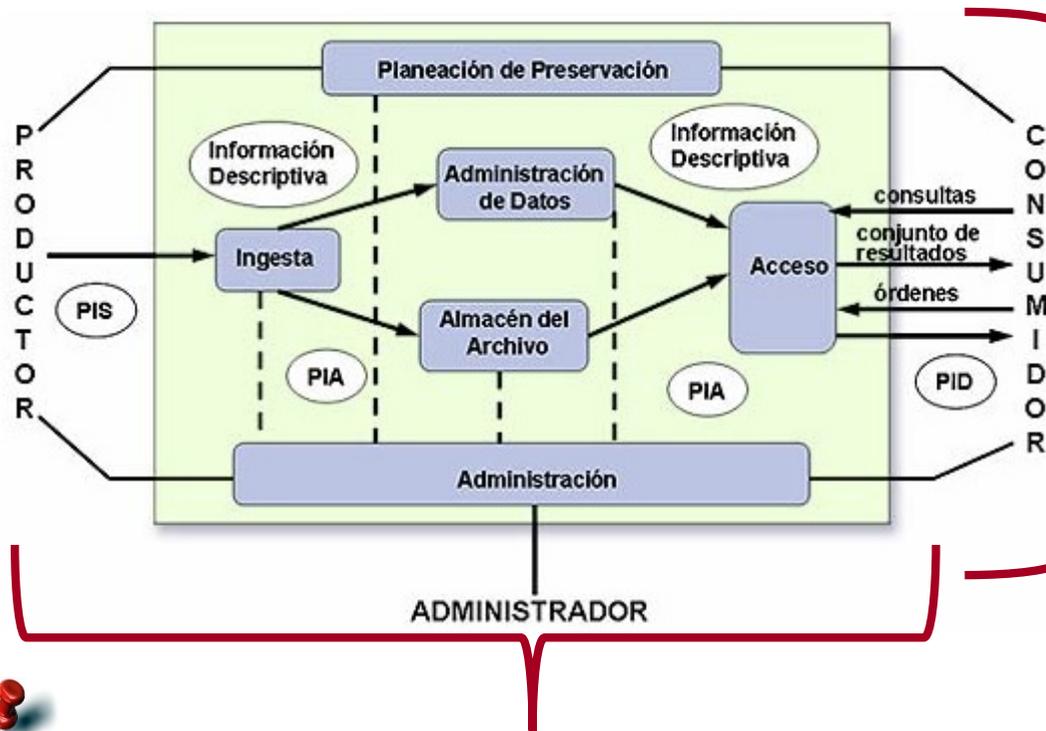
El ParER y el modelo OAIS: ¿una relación omnicomprendensiva?



¿Hacer referencia al modelo OAIS permite cumplir con todos los requisitos y exigencias relacionadas con la implementación del sistema de preservación digital del ParER?

El modelo OAI S:

¿ una panacea para todos los problemas ? (1)



OAIS es un "reference model" entonces:

§NO incluye el diseño aplicativo para un sistema de preservación digital

§NO contiene especificaciones y soluciones técnicas para la implantación de una plataforma de preservación digital

(Mariella Guercio)

OAIS se aplica:

§a cualquier tipo de objeto digital (pero el documento archivístico es una particular entidad digital)

§a cualquier tipo de organización (pública, privada, cultural, científica etc.)

§a cualquier tipo de marco archivístico (modelo custodial versus modelo post-custodial)

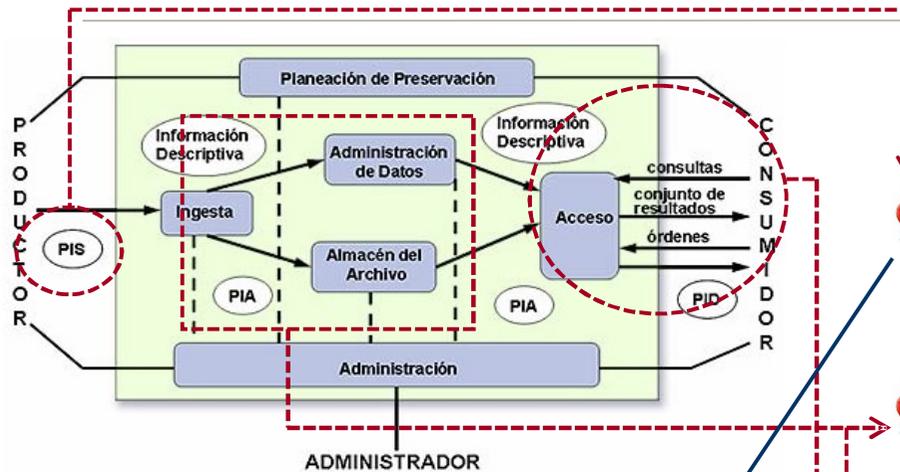
Un sistema de preservación digital requiere, para su implementación, que

el modelo OAI S sea integrado por una serie de recursos/políticas adicionales que permitan al propio modelo ser declinado en un marco real concreto

El modelo OAI S requiere ser contextualizado

El modelo OAIS:

¿ una panacea para todos los problemas ? (2)



El modelo OAIS requiere ser integrado por (lista de ejemplos no exhaustiva):

un estándar para la metodología de transferencia de la documentación digital del sistema de gestión documental del productor hacia el archivo del custodio

un estándar para la producción de los metadatos relacionados con los procesos de preservación digital llevados a cabo por el custodio en el archivo (por ejemplo PREMIS)

un estándar para los procesos de migración que el custodio tiene que llevar a cabo sobre la documentación digital, a través de distintos soportes y formatos de ficheros

un estándar para la búsqueda, recuperación, representación de la documentación digital preservada en el archivo, en relación a los usuarios de la web

un estándar para la certificación y auditoría del sistema de preservación digital en su conjunto (por ejemplo ISO 16363:2012)

un estándar para el protocolo de comunicación utilizado en las transacciones entre el productor y el custodio (por ejemplo HTTPS)

un estándar que pueda definir estructura y metadatos (sobre todo metadatos de contexto) de la documentación digital transferida del productor hacia el archivo del custodio

¿Estándares de facto versus estándares de iure?

¿Estándares internacionales, nacionales, "locales"?

El modelo OAI S:

¿ una panacea para todos los problemas ? (3)



+ estándares (recursos/políticas) adicionales



¿Estándares de facto versus estándares de iure?

¿Estándares internacionales, nacionales, locales?



Segunda serie de cuestiones críticas:

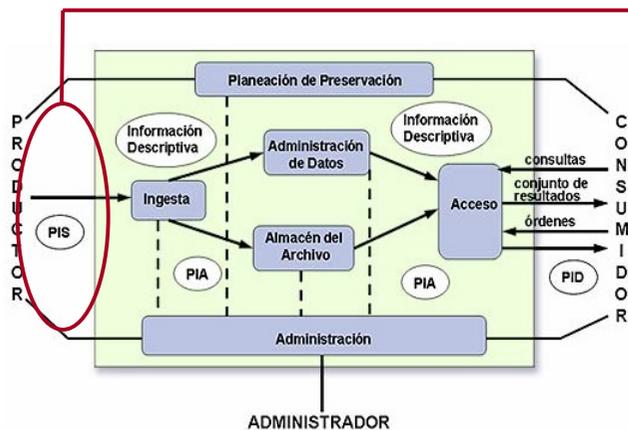
¿Qué estándar elegir?

Si ya no existe un adecuado estándar a nivel internacional o nacional para el tema específico a tratar, el custodio tendrá que elaborar un "estándar local", es decir un recurso que valga dentro de su archivo o en el marco delimitado de los productores y usuarios que interactúan con ello

Si ya existe un adecuado estándar a nivel internacional o nacional para el tema específico a tratar, cuanto más el propio estándar tenga una naturaleza flexible tanto más el custodio tendrá que obrar para adaptarlo a sus necesidades específicas



El proceso de preservación digital en el ParER: recursos y políticas dentro el marco del OAIS (2)



La fase de pre-transferencia de la documentación digital al custodio

¿Cuándo se utiliza?

Fase opcional: está presente cuando, por razones técnicas, no se puede transferir directamente la documentación digital al custodio en forma de un PIT estandarizado según las normas del ParER (por ejemplo documentación compuesta por ficheros de gran tamaño)

¿En qué consiste?

La documentación digital entra en contacto con el sistema de preservación digital del ParER en forma de un PIT no estandarizado. Si supera los controles establecidos, el sistema de preservación digital lo convierte en un PIT perfectamente estandarizado según las normas del ParER

Estándar para el protocolo de comunicación entre el productor y el custodio

ParER ha elegido como protocolo "normal" el HTTPS, que no se puede pero utilizar siempre en todos los contextos: por lo tanto como protocolo "alternativo" se utiliza el FTP (para la fase de pre-transferencia)

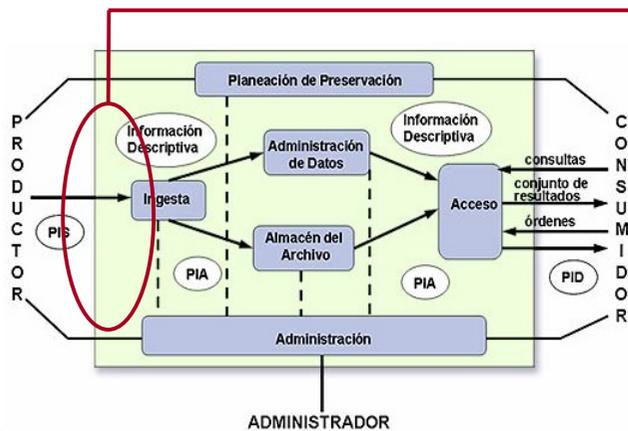


¡MUCHO CUIDADO!

Cuestion crítica:

a veces no hay sólo que elegir un único estándar para el tema específico a tratar, sino que hay que prever más casos de uso y entonces elegir más de un estándar para el mismo marco

El proceso de preservación digital en el ParER: recursos y políticas dentro el marco del OAIS (3)



Transferencia de la documentación digital al custodio

¿En qué consiste?

Fase obligatoria:

§ la transferencia de la documentación digital al custodio se realiza a través de un paquete de información sobre transferencia (PIT)

§ el PIT está formado por ficheros y metadatos (estructurales, contextuales, técnicos y de gestión)

§ el PIT está normalizado según las normas del ParER, que se actualizan continuamente

§ el PIT normalizado varía en función de la tipología (forma) documental y tipo de entidad archivística (unidad documental simple y unidad documental compuesta)

§ el PIT normalizado según las normas del ParER está codificado en el lenguaje XML

Estándar para
definir estructura y metadatos de la
documentación digital transferida

ParER ha elaborado un
"estándar local", en forma de normas
sobre el PIT que valen en el marco
delimitado de los productores que
interactúan con ello

Italia:

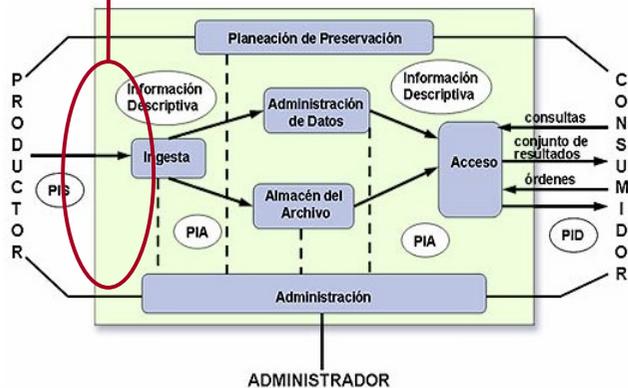
no existe un estándar nacional que define, de manera exhaustiva, estructura y metadatos de la documentación digital enviada/recibida por distintos sistema de gestión documental y preservación digital

España:

"Esquema nacional de metadatos para la gestión de documentos electrónicos" (eEMGDE)

El proceso de preservación digital en el ParER: recursos y políticas dentro el marco del OAIS (4)

Transferencia de la documentación digital al custodio



Estándar para
definir estructura y metadatos de la
documentación digital transferida

ParER ha elaborado un
"estándar local", en forma de normas
sobre el PIT que valen en el marco
delimitado de los productores que
interactúan con ello



Cuestiones críticas:

§ las normas del ParER tienen un valor en un marco local. Por lo tanto los sistemas de gestión documental de los productores no están ya predispuestos para producir el PIT normalizado según las normas del Parer;

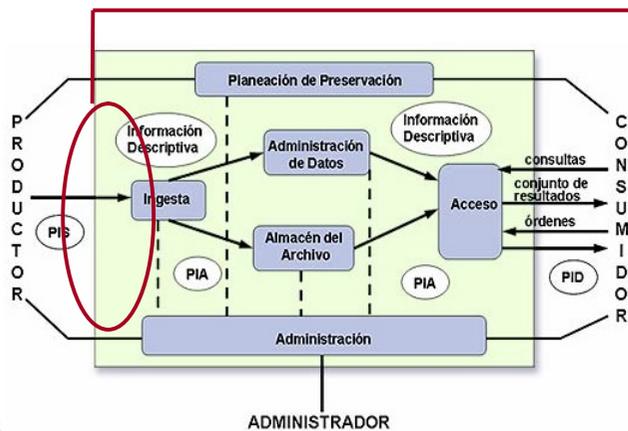
§ cuando un productor quiere interactuar con el ParER, se requiere una actividad más o menos costosa para la construcción de una interfaz entre su sistema de gestión documental y el sistema de preservación digital, de acuerdo con la tecnología del Web Service y la normas del ParER sobre el PIT;

§ como las normas del ParER sobre el PIT se actualizan, también la interfaz del sistema de gestión documental tiene que actualizarse, en la medida en que el productor quiera utilizar los nuevos servicios del sistema de preservación digital (costos adicionales);

§ la actividad de "poblar" el PIT con los metadatos requeridos por las normas del ParER es a veces muy costosa, porque los metadatos tienen que extraerse de la base de datos propia del sistema de gestión documental del productor;

§ no hay una estandarización muy extensa de los sistemas de gestión documental, con respecto a los metadatos producidos y procedimientos de gestión documental que afectan a las características de la documentación digital

El proceso de preservación digital en el ParER: recursos y políticas dentro el marco del OAIS (5)



La fase de comprobación del paquete de información sobre transferencia (PIT)



¿En qué consiste?

el PIT estandarizado está sometido a unos controles automáticos por el sistema de preservación digital, para poner de relieve cualquier anomalía antes de aceptar la entrega del propio PIT. Los principales controles se actúan sobre:

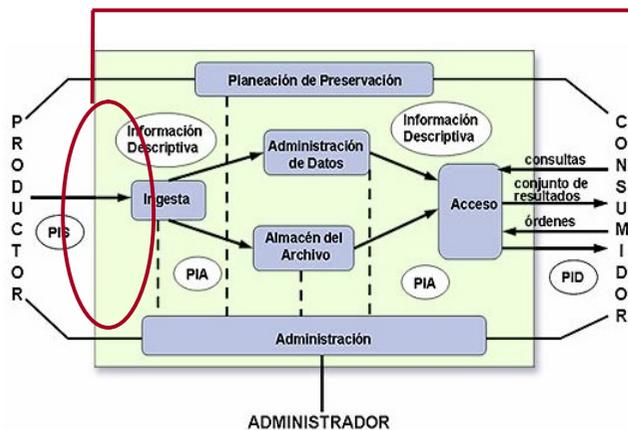
- § la identificación del productor que pone en acción la transferencia de la documentación digital;
- § el cumplimiento, por parte del PIT que se está intentando transferir, del modelo estandarizado que el ParER ha definido para los PIT;
- § la consistencia de los metadatos (presencia y valores atribuidos), en particular de los metadatos de contexto;
- § la comprobación de la validez de las eventuales firmas digitales asociadas a la documentación digital;
- § admisibilidad de los formatos de ficheros que componen la documentación digital que se está intentando transferir

Reglas de individuación de las fechas útiles para comprobar la validez de las firmas digitales

La legislación italiana ofrece al ParER unas fechas ciertas que son "oponibles a terceros", es decir pueden ser exhibidas como prueba en los tribunales (por ejemplo la fecha en la que un documento digital queda registrado en el registro de salida/entrada de documentos)

● Sello de tiempo

El proceso de preservación digital en el ParER: recursos y políticas dentro el marco del OAIS (6)



La fase de aceptación o rechazo del paquete de información sobre transferencia (PIT)

¿En qué consiste?

si al menos uno de los controles de la fase anterior tiene un resultado negativo, se rechaza el PIT. En este caso:

§ el sistema de preservación digital devuelve al productor un documento XML, "Resultado de la transferencia", que contiene la información sobre los controles efectuados y sus resultados;

§ el PIT rechazado se almacena en un área temporal del sistema de preservación digital, lógicamente exterior al archivo del productor, de modo que el productor y el ParER puedan analizar las anomalías.

Si todos los controles tienen éxito, el PIT entra en el sistema de preservación digital para su incorporación lógica en el archivo del productor. Para cada PIT aprobado se genera automáticamente el "Informe sobre la transferencia"



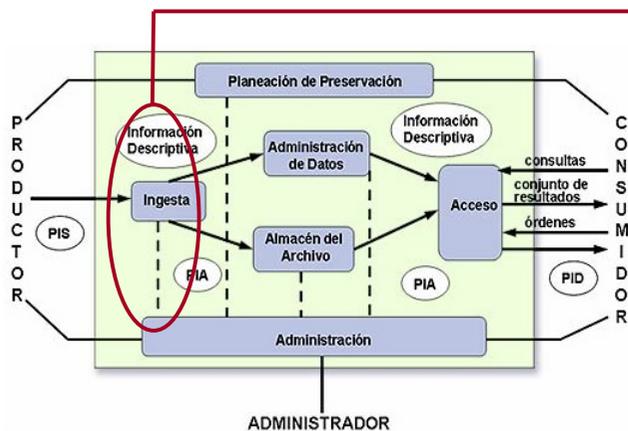
Question crítica:

si los productores no tienen una "cultura" y políticas para una adecuada gestión documental, el número de anomalías en la fase de transferencia resulta ser muy relevante y así las propias anomalías "caen al vacío" sin tener resolución.

No se activa ningún feedback entre el custodio y el productor (visión prospectiva)

Recursos que el ParER invierte para profundizar la "cultura" de gestión documental de los productores que obran en su territorio de competencia

El proceso de preservación digital en el ParER: recursos y políticas dentro el marco del OAIS (7)



La fase de toma en custodia y generación del "Informe sobre la transferencia"

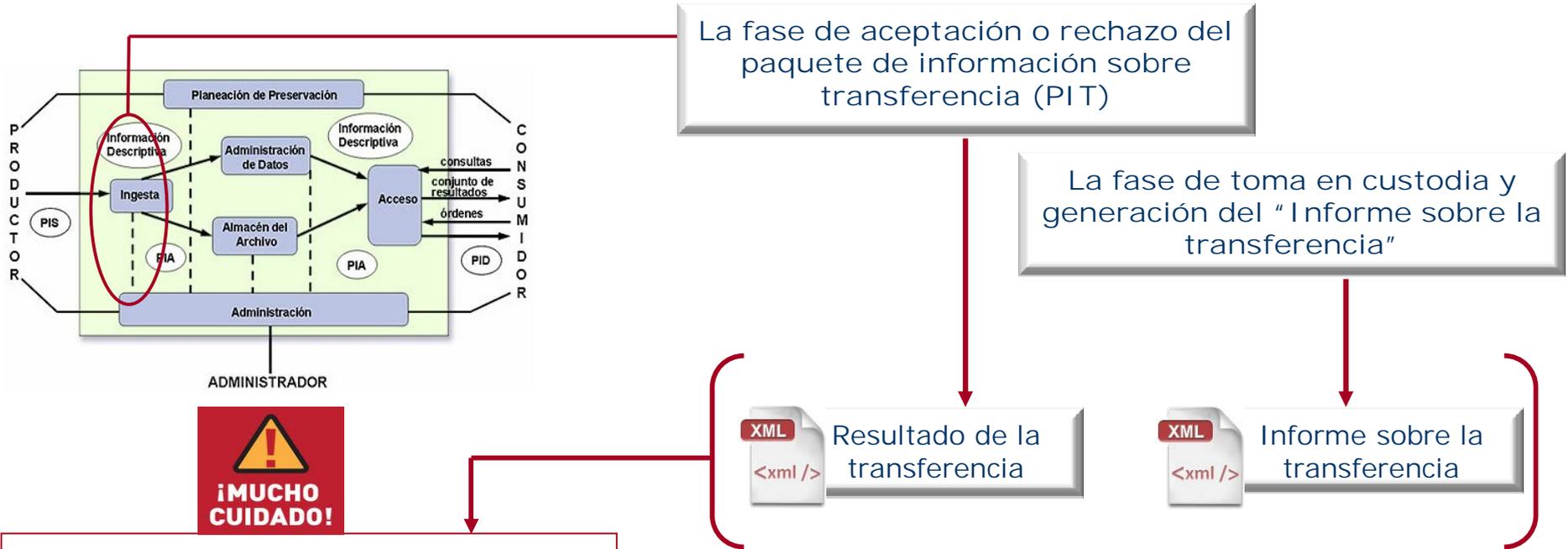
¿En qué consiste?

Para atestiguar la exitosa toma en custodia del PIT, se genera automáticamente el "Informe sobre la transferencia" que se almacena en el sistema de preservación digital y se asocia con el relacionado PIT. Este informe contiene entre otras cosas:

- § la fecha de su creación, producida de manera que valga como fecha cierta, "oponible a terceros" de acuerdo con la normativa italiana;
- § la huella del PIT en su conjunto, producida con la función Hash;
- § las huellas producidas con la función Hash para cada uno de los ficheros que componen el PIT;
- § una breve descripción del propio PIT.

Este informe se pone directamente a disposición del productor, a través de un documento en formato XML, "Resultado de la transferencia".

El proceso de preservación digital en el ParER: recursos y políticas dentro el marco del OAIS (8)



Cuestión crítica:

no sólo el sistema de preservación digital recibe en custodia documentos digitales de los productores, sino ello mismo produce documentación digital.

§ ¿Qué hacer con ella?

§ ¿Simplemente almacenarla?

§ **¿Preservarla?**



"la espiral de la preservación digital"

El proceso de preservación digital en el ParER: recursos y políticas dentro el marco del OAI S (9)



La fase de producción del paquete de información archivística (PIA)

¿Cuándo se realiza?

El paquete de información archivística (PIA) se crea:

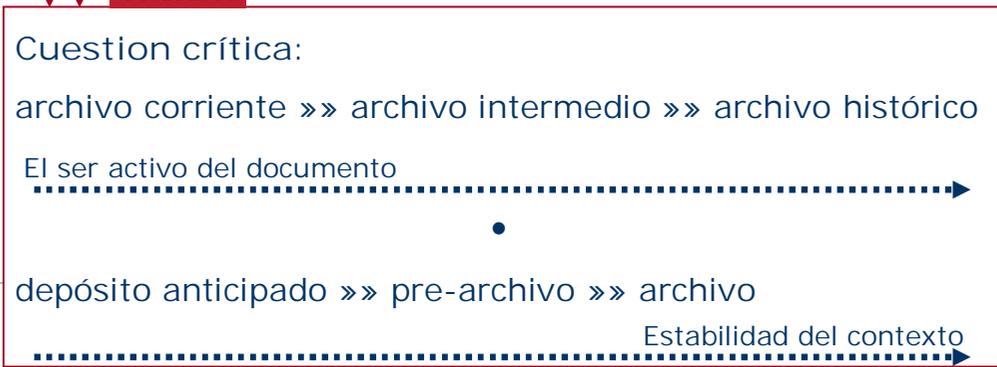
- § a nivel de la unidad documental simple transferida, en la fase del depósito "anticipado";
- § a nivel de unidad documental compuesta y serie archivística, en la fase del "pre-archivo".



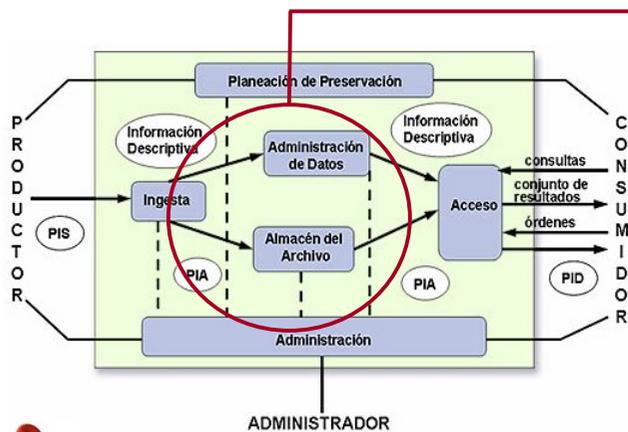
Por lo tanto:

- ü cualquier entidad tomada en custodia en forma de PIT, para ser gestionada en el sistema de preservación digital tiene que convertirse en un PIA;
- ü a la altura del "pre-archivo", en el sistema de preservación digital hay una red de relaciones entre diferentes PIA que pertenecen a distintos niveles;
- ü dentro de la fase del "pre-archivo" el sistema de preservación valora los PIA, para establecer su retención y entrada en el archivo o destrucción según plazos establecidos

en la fase del "pre-archivo" el contexto documental más próximo perteneciente a la unidad documental simple (ya transferida) se aproxima a un primer grado de estabilidad (por ejemplo a la conclusión de un expediente)



El proceso de preservación digital en el ParER: recursos y políticas dentro el marco del OAI S (10)



La fase de producción del paquete de información archivística (PIA)

¿En qué consiste?

El PIT tomado en custodia está sometido a un proceso para que sirva de base para la generación del PIA. Este proceso consiste en:

- § la normalización de los metadatos contenidos en el PIT según las normas establecidas por el ParER sobre el PIA;
- § la eventual normalización de la estructura del documento digital transferido a través del PIT según las normas establecidas por el ParER sobre el PIA;
- § la migración de los formatos de ficheros a través de un procedimiento que garantiza la producción de copias auténticas del documento digital

Estándar para definir metadatos y estructura de la documentación digital custodiada

ParER ha elaborado unas normas sobre el PIA, utilizando como referencia el estándar nacional UNI 11386:2010, requerido por la normativa italiana



Question crítica:

¿El estándar nacional cumple perfectamente con todas las necesidades de la preservación digital a largo plazo?

El PIA que supera la fase de valoración y entra en el archivo, tiene que proporcionar un contexto completo (definido e inmutable) ¿Para proporcionar este contexto son suficientes los metadatos o sirve también la description post hoc del archivero?

La relación entre metadatos, descripción archivística y contexto: el ejemplo de la documentación sanitaria digital (1)

Las tecnologías de la información están extendidas en el campo sanitario (e-Health) con un grado de madurez que es desconocido en otros sectores de la producción documental

DINAMICIDAD de la documentación sanitaria digital

Principio de
"continuidad asistencial":

al paciente se le debe garantizar una actividad de asistencia que se despliegue como un continuum, independientemente de todas las posibles "fracturas"

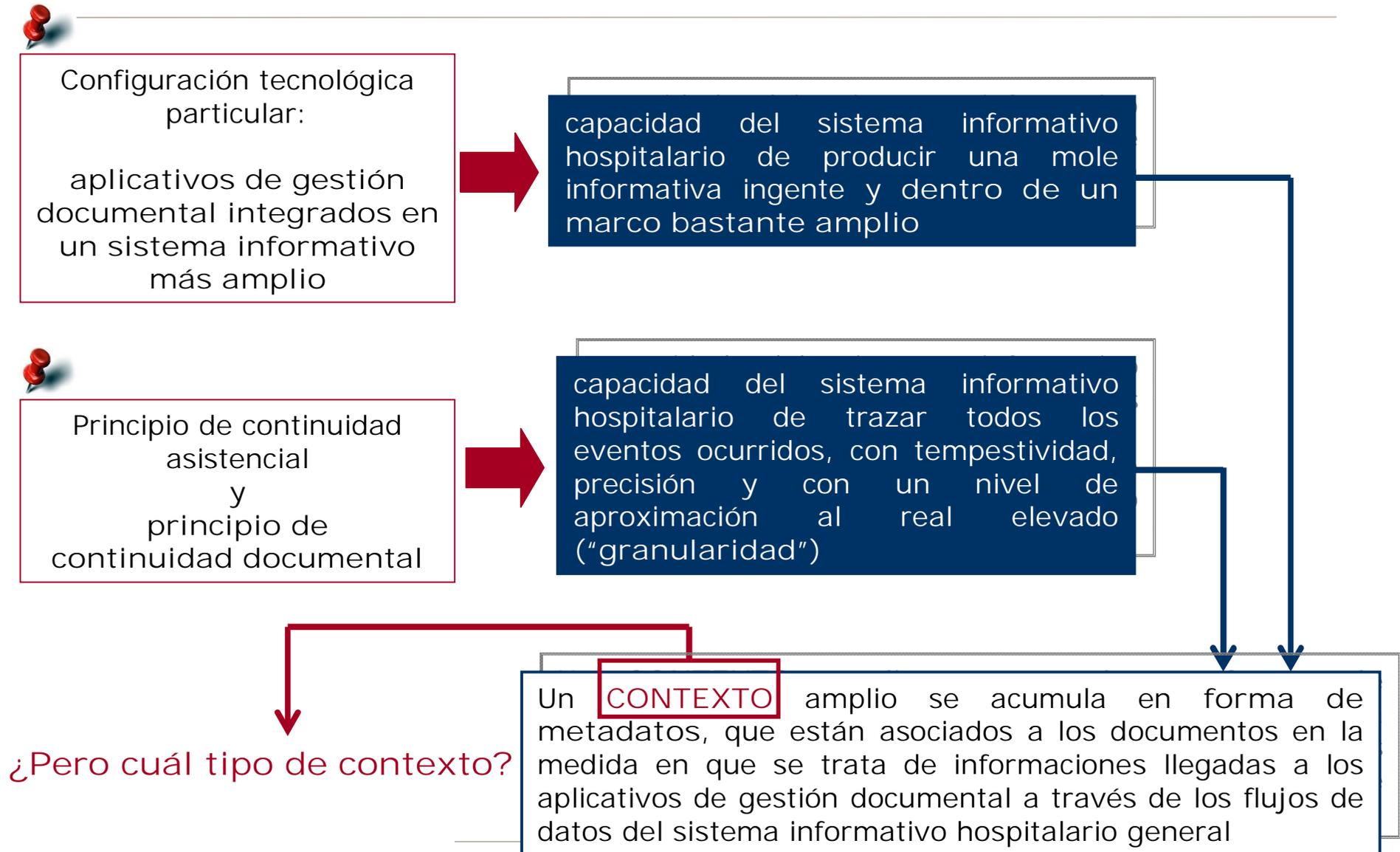
Principio de
"continuidad documental"

la documentación sanitaria digital no es el producto de un simple sistema de gestión documental

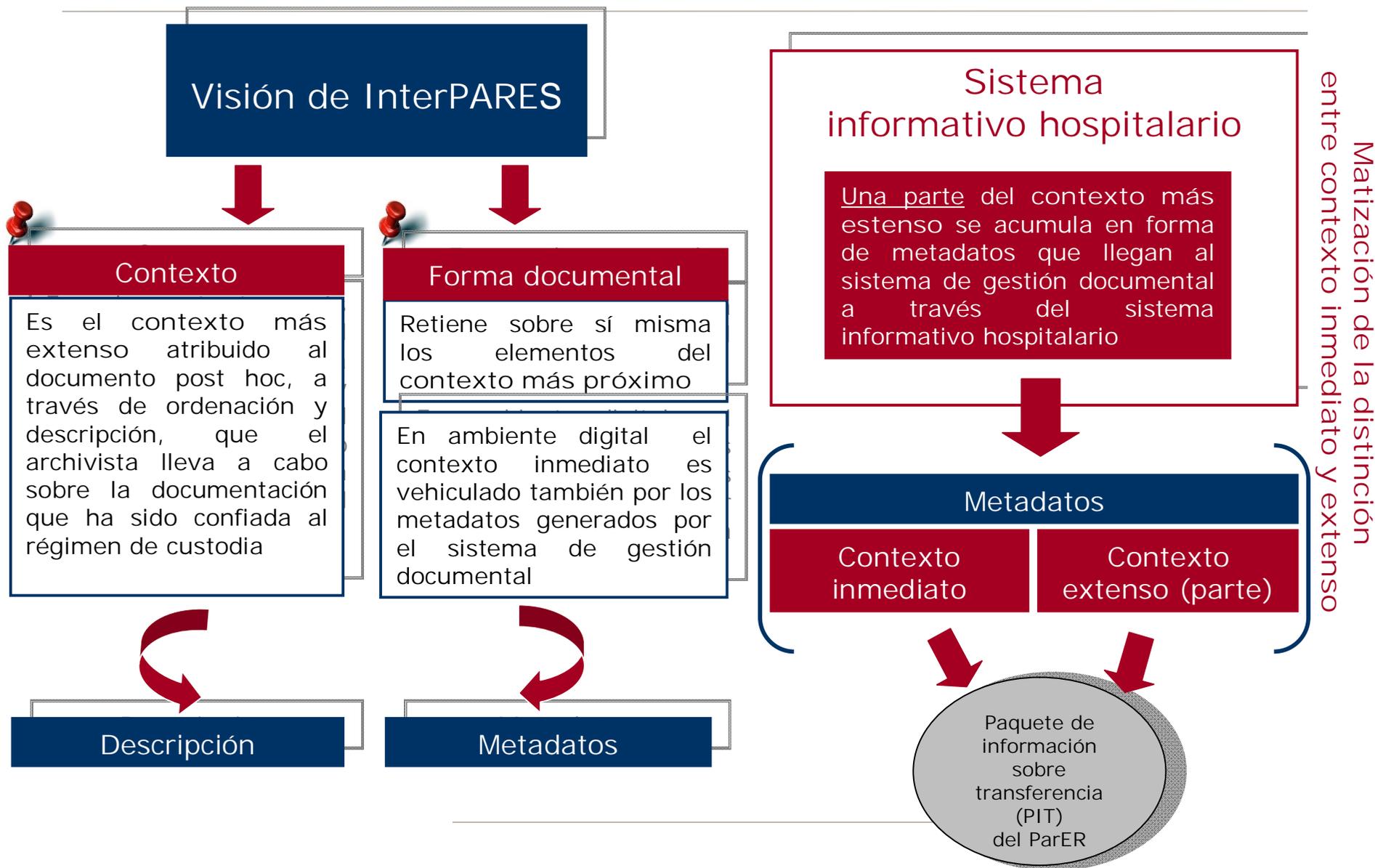
sino el resultado de un sistema informativo articulado: el "sistema informativo hospitalario", un conjunto de aplicaciones interconectadas que interactúan con un elevado grado de dinamicidad y **entre las cuales se encuentran también los módulos de gestión documental**

su lógica funcional es la de la actualización granular, puntual y tempestiva de la mole informativa, que se considera imprescindible para que las estructuras sanitarias puedan proporcionar cuidados adecuados a los pacientes

La relación entre metadatos, descripción archivística y contexto: el ejemplo de la documentación sanitaria digital (2)



La relación entre metadatos, descripción archivística y contexto: el ejemplo de la documentación sanitaria digital (3)



La relación entre metadatos, descripción archivística y contexto: el ejemplo de la documentación sanitaria digital (4)

Documentación
sanitaria digital



¿evolución de la relación entre forma documental,
contexto inmediato y contexto extenso?

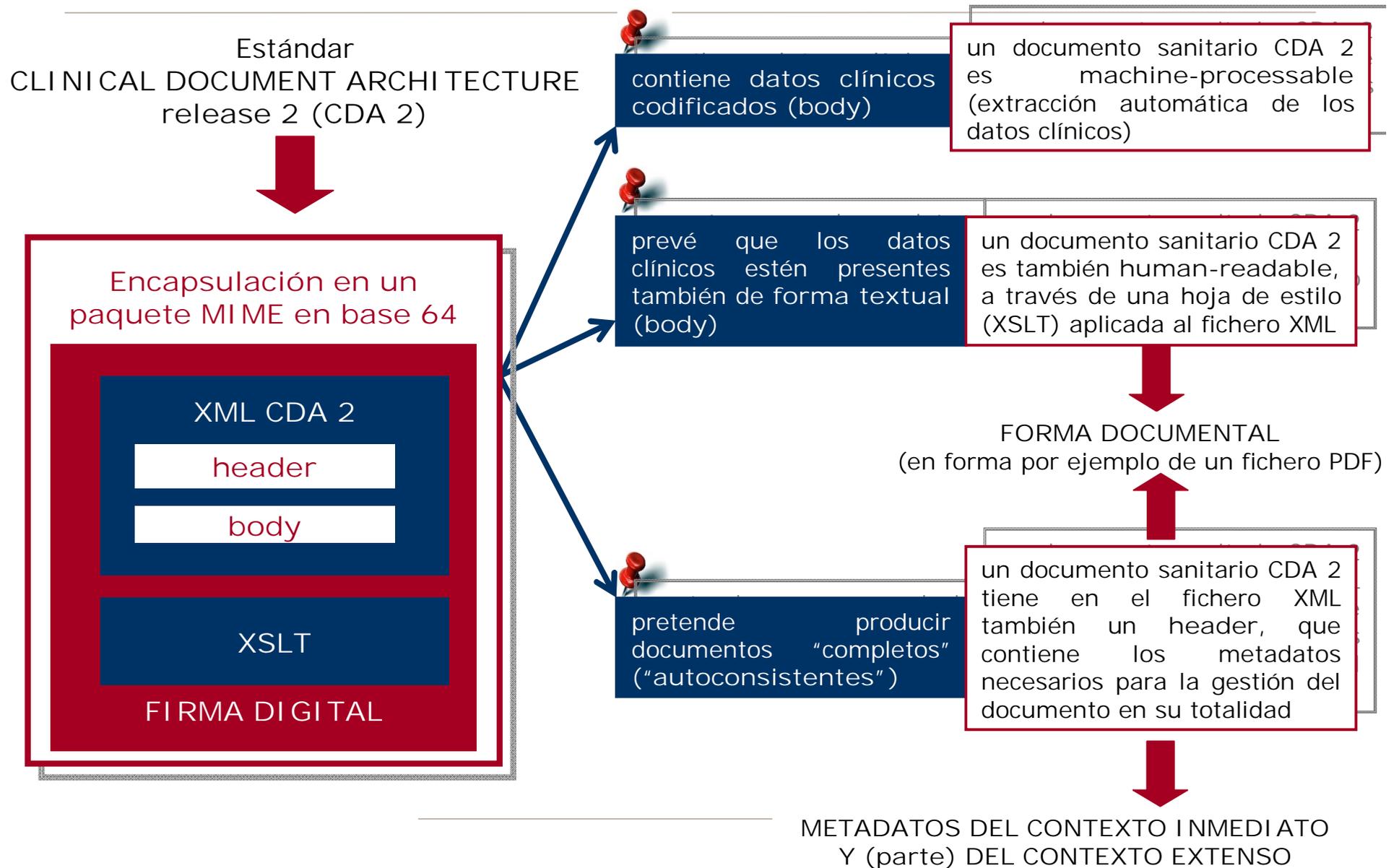
Estándar
CLINICAL DOCUMENT ARCHITECTURE
release 2 (CDA 2)
(ISO/HL7 27932:2009)

Mantenido por la organización sin
animo de lucro High Level Seven
(HL7)

Creado para permitir a los documentos
sanitarios digitales ser
intercambiados entre diferentes
aplicativos

Por lo tanto es implementado en el
lenguaje de codificación XML

La relación entre metadatos, descripción archivística y contexto: el ejemplo de la documentación sanitaria digital (5)



La relación entre metadatos, descripción archivística y contexto: el ejemplo de la documentación sanitaria digital (6)

Estándar
CLINICAL DOCUMENT ARCHITECTURE
release 2 (CDA 2)



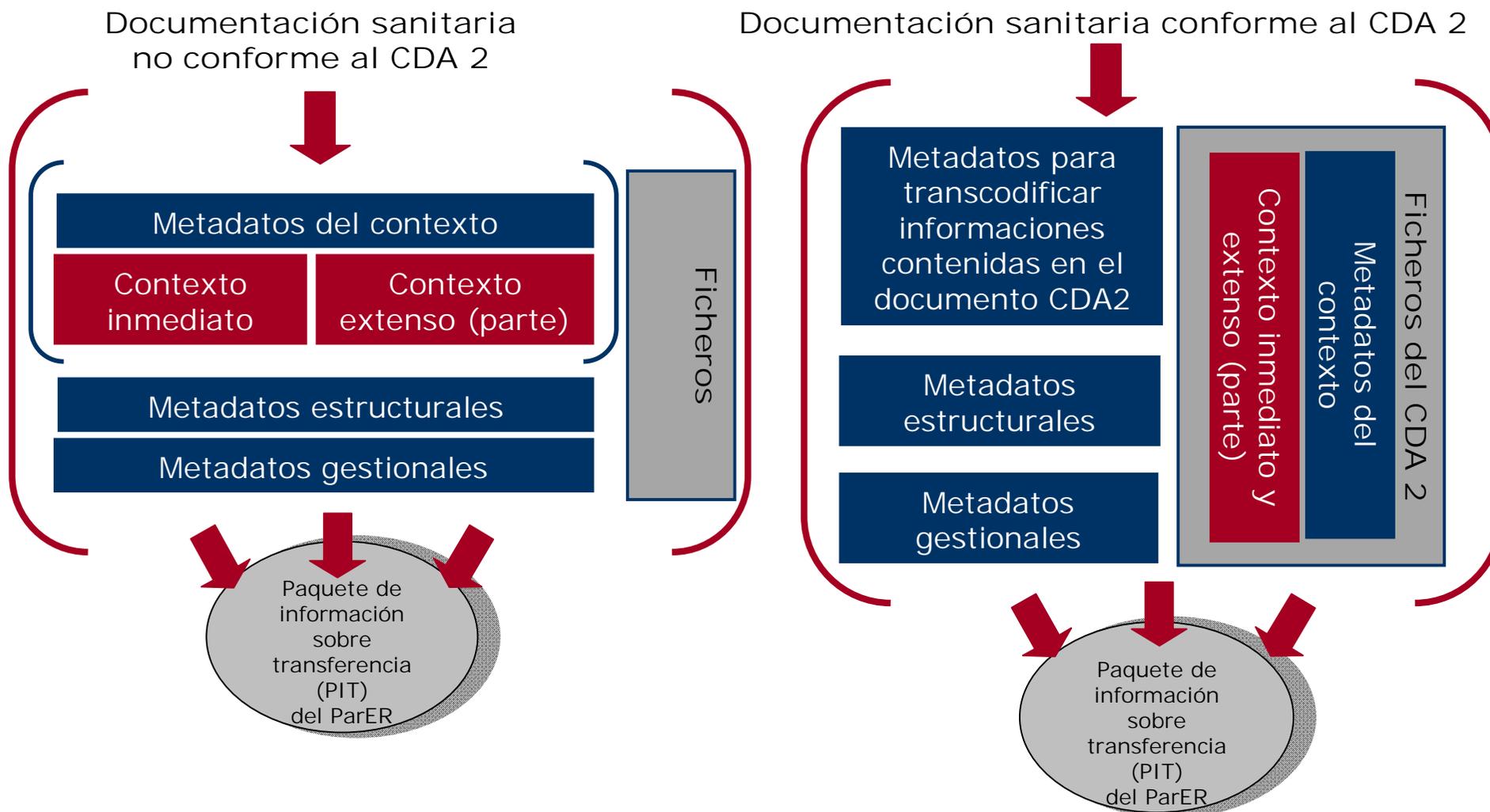
el contenido (los datos clínicos del body) y el contexto (los metadatos del header) están "desvinculados" de la forma documental tradicional

contenido y contexto viajan en el tiempo y en el espacio sin servirse de la forma documental tradicional, que no es un continuum coexistente a los mismos (es "invocada" solo cuando el humano procesa contenido y contexto a través de la hoja de estilo original)

la forma documental tradicional se eclipsa – por supuesto no desaparece – y parece reducirse a metadatos

la distinción entre metadatos del contexto inmediato y metadatos del contexto extenso aparece fuertemente debilitada

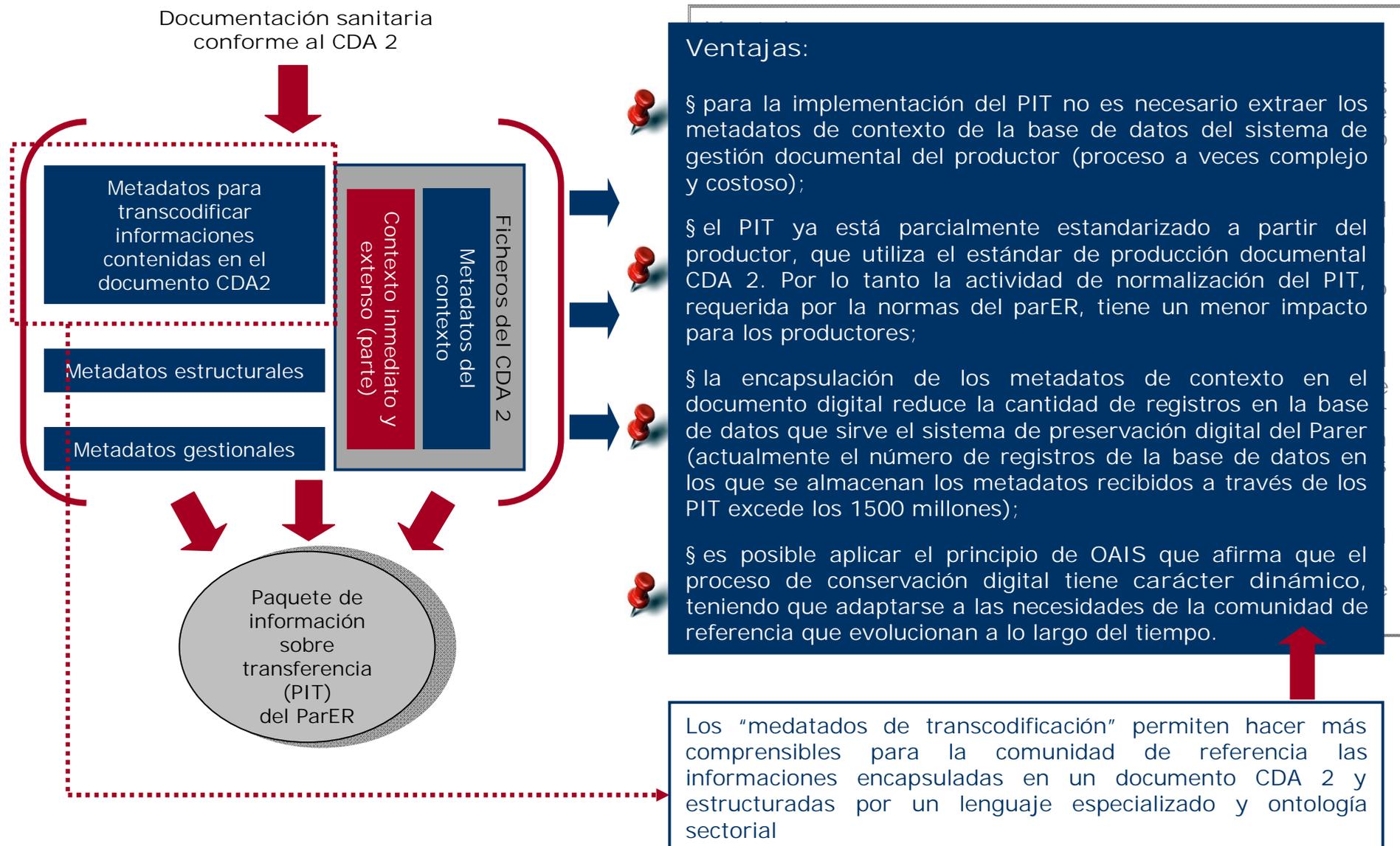
La relación entre metadatos, descripción archivística y contexto: el ejemplo de la documentación sanitaria digital (7)



Encapsulación →

- § MEOS'S (Metadata Encapsulated Objects) de David Bearman
- § HERO (Hurley's Enduring Recordkeeping Object) de Chris Hurley

La relación entre metadatos, descripción archivística y contexto: el ejemplo de la documentación sanitaria digital (8)





Muchas gracias a todos

<http://parer.ibc.regione.emilia-romagna.it>