FAQ sobre identificadores ARK

Preguntas frecuentes y respuestas sobre ARK

- · Lo esencial
 - o ¿Cómo puedo dar comentarios sobre este documento?
 - O ¿Qué son los ARK?
 - o ¿Qué es un identificador?
 - o ¿Qué es un identificador persistente?
 - o ¿Qué es un resolvedor?
 - o ¿A qué tipo de cosas se asignan los ARK?
 - ¿Quién está usando ARKs?
- Empezando
 - ¿Qué necesito para crear ARK?
 - ¿Cómo empiezo a crear las cadenas de caracteres que se convierten en ARK?
 - Qué son los identificadores opacos?
 - ° ¿Cómo hago que el contenido del servidor sea direccionable con ARK?
 - ¿Cómo cito o publicito un ARK?
 - o ¿Existen herramientas y servicios para ayudar con los ARK?
- Mas allá de lo básico
 - o ¿Qué es el N2T?
 - Si la mayoría de los ARK se ejecutan en sus propios resolvedores, ¿por qué existe también un resolvedor global para ARK?
 Mi organización tiene su propio resolvedor ARK. ¿Debería preocuparme por N2T.net?

 - ° ¿Por qué el resolvedor ARK global (n2t.net) no tiene la palabra "ARK"?
 - O ¿Qué significa "paso de sufijo"?
 - ¿Cuáles son las partes de un ARK?
 - ¿Puedo asignar ARK a cosas dentro de algo que ya tiene un ARK?
 - o ¿Cuál es el propósito de la NAAN, y puedo hacer cambios a ella?
 - ¿Existen restricciones en el uso de NAAN?
- ARK y otros identificadores
 - ° ¿Por qué usaría ARK en comparación con, por ejemplo, DOI?
 - o ¿Qué tienen en común ARK, DOI, Handle, PURL y URN?
 - Espera, ¿estás diciendo que ARK, DOI, Handle, PURL y URN son inútiles?
 - ° ¿En qué se diferencian los ARK de los identificadores como DOI, Handles, PURL y URN?
 - La respuesta corta
 - Más diferencias entre ARKs, DOIs, Handles, PURLs y URNs
 - O Pero si se pueden eliminar los ARK, ¿cómo se puede confiar en ellos?
 - ¿Puede un objeto tener un ARK y un DOI?
 - o ¿Cuándo debo usar ARK en comparación con DOI, Handles, PURL o URN?
- De la cuna a la tumba
 - o ¿Qué se entiende por ARKs que apoyan el desarrollo temprano de objetos?
 - Si los ARK no lo requieren, ¿por qué molestarse en crear metadatos?
 - ¿Qué metadatos se recomiendan para los ARK?
 - O ¿Por qué veo metadatos ARK con las etiquetas de quién, qué, cuándo y dónde ?
 - ° ¿Qué es una "inflexión" de ARK y en qué se diferencia de la "negociación de contenido"?
 - ¿Qué quieres decir con depósitos?

Lo esencial

¿Cómo puedo dar comentarios sobre este documento?

Se puede insertar los comentarios en esta versión del documento.

¿Qué son los ARK?

Los ARK ("Archival Resource Key," claves de recursos de archivo) son identificadores de alto funcionamiento que lo conducen a cosas y a descripciones de esas cosas. Por ejemplo, este ARK,

https://n2t.net/ark:/67531/metadc107835/

te lleva a una disertación. Al agregar dos '??' al final del ARK se debería llegar a la descripción del objeto:

https://n2t.net/ark:/67531/metadc107835/??

¿Qué es un identificador?

En Internet, un identificador es una URL, o parte de una URL. Por ejemplo, este identificador ARK central,

ark:/12148/btv1b8449691v/f29

aparece dentro de dos URL diferentes (Localizadores uniformes de recursos, también conocidos como enlaces web o direcciones web):

http://ark.bnf.fr/ark:/12148/btv1b8449691v/f29

Los ARK son especialmente buenos para ser identificadores persistentes.

¿Qué es un identificador persistente?

Una vez se dijo que la vida media de una URL era de 44 días. Al final de su vida útil, se *rompe* un enlace de URL, lo que significa que le da el temido error "404 No encontrado" que la mayoría de nosotros hemos visto. Por irritante que sea, es políticamente incómodo cuando se busca investigación financiada con fondos públicos, y es un desastre cultural para bibliotecas, archivos, museos y otras organizaciones de memoria.

Entre los muchos enlaces que pueden o podrían conducirlo a cosas, un *identificador persistente* es un enlace que, en principio, sigue funcionando en el futuro. Los servicios que proporcionan descubrimiento e interconexión (como entre artículos de investigación, autores, datos de respaldo e investigaciones relacionadas) prefieren identificadores persistentes debido a esa estabilidad.

Los identificadores persistentes deberían seguir funcionando incluso cuando las cosas se mueven entre sitios web. Normalmente, cuando las cosas se mueven, todos los que alguna vez registraron los enlaces antiguos necesitarían saber cuáles son los nuevos enlaces, lo que es casi imposible. Ahí es donde entran los resolvedores de identificadores.

¿Qué es un resolvedor?

Un resolvedor es un sitio web que se especializa en reenviar identificadores entrantes (los que se anunciaron originalmente a los usuarios) a los sitios web que estén mejor capacitados para manejarlos. En general, el reenvío se llama resolución; Un paso en un proceso de resolución se llama redirección.

Para que un resolvedor funcione, su nombre de host (n2t.net o ark.bnf.fr) en los identificadores anteriores) debe elegirse cuidadosamente para que nunca sea necesario cambiarlo. Las organizaciones de memoria, algunas de ellas con siglos de antigüedad, tienden a tener nombres de host adecuados para ser resolvedor. Algunos resolvedores más jóvenes y conocidos son n2t.net (el resolvedor ARK), identifiers.org, doi.org, handle.net y purl.org.

¿A qué tipo de cosas se asignan los ARK?

Para cualquier cosa digital, física o abstracta. Eso puede incluir cosas que aún no existen, pero a las que debe hacer referencia desde los objetos que está en proceso de crear o planificar, como un enlace de un borrador de artículo a un conjunto de datos en preparación, o un enlace de un carta digital archivada a una ayuda de búsqueda planificada. Una advertencia es que generalmente debe asignar ARK a cosas que posee, controla o administra. Se desaconseja asignar ARK a cosas que no controlas porque tales identificadores tienden a ser frágiles.

Los tipos de cosas que tienen ARK incluyen los que se enumeran a continuación. Los números son aproximados, actuales a partir de septiembre de 2019, y son autoinformados por las organizaciones vinculadas.



¿Quién está usando ARKs?

Eso es un poco difícil de decir porque las ARK están muy descentralizadas, pero más de 600 organizaciones registradas han creado, entre ellas, aproximadamente 3.200 millones de ARK. Puede encontrar ARKs utilizados como enlaces permanentes en

- el Data Citation Index (vinculado a la Web of Science),
- · Artículos de Wikipedia,
- Registros de Wikidata,
- Colecciones de archivo de Internet,
- · Perfiles de investigadores ORCID, etc.

Aquí está la distribución global de las organizaciones registradas para crear ARK a partir de abril de 2020. Al hacer clic en la imagen estática a continuación, debería acceder a un mapa actualizado y ampliable.



Empezando

¿Qué necesito para crear ARK?

Primero necesita un NAAN ("Name assigning authority number," Número de autoridad de asignación de nombre), que es un número reservado exclusivamente para su organización. Debe aparecer en cada ARK que su organización asigne, justo después de la etiqueta "ark:/". La NAAN en todas estas ARKs

ark:/12148/btv1b8449691v/f29

http://ark.bnf.fr/ark:/12148/btv1b8449691v/f29

https://n2t.net/ark:/12148/btv1b8449691v/f29

es 12148, e identifica de forma exclusiva la Biblioteca Nacional de Francia. Cada NAAN está asociado con la URL de un resolvedor para sus ARK, por ejemplo, para resolver los 12148 ARK, agréguelos a http://ark.bnf.fr/ como se muestra arriba. El resolvedor N2T.net es inusual en el sentido de que dirige cualquier ARK al resolvedor registrado bajo su NAAN.

Es gratis obtener o usar un NAAN, y se lo puede solicitar a través de este formulario. Más de 600 organizaciones tienen un NAAN (bibliotecas, archivos, museos, departamentos universitarios, agencias gubernamentales, editoriales académicas y educativas, proyectos, etc.), todos incluidos en el registro público de NAAN.

¿Cómo empiezo a crear las cadenas de caracteres que se convierten en ARK?

Puede crear cadenas ARK como lo desee, siempre que use solo dígitos, letras (ASCII, sin signos diacríticos) y los siguientes caracteres:

```
= ~ * + @ _ $ . /
```

Los dos últimos caracteres están reservados en el caso de que desee revelar relaciones ARK.

Otra característica única de los ARK es que pueden aparecer guiones ('-') pero son *inertes de identidad*, lo que significa que las cadenas que difieren solo por guiones se consideran idénticas; por ejemplo, estas cadenas

```
ark:/12345/141e86dc-d396-4e59-bbc2-4c3bf5326152
```

ark:/12345/141e86dcd3964e59bbc24c3bf5326152

Identificar lo mismo. La razón de esta característica es que los procesos de formateo de texto en el mundo introducen rutinariamente guiones adicionales en los identificadores, rompiendo enlaces a cualquier servidor que trate los guiones como significativos.

Los ARK distinguen entre letras minúsculas y mayúsculas, lo que hace posibles identificadores más cortos (52 vs 26 letras por posición de carácter). Sin embargo, la "forma ARK" es usar minúsculas a menos que necesite ARK más cortos. La restricción hace que sea más fácil para los resolvedores admitir sus ARK en caso de que lleguen del mundo con letras mayúsculas o mixtas, lo que sucede lamentablemente a menudo debido a la suposición persistente de 50 años de que los identificadores no distinguen entre mayúsculas y minúsculas. También puede considerar el uso del repertorio de caracteres de la herramienta Noid, que crea cadenas seguras para la transcripción utilizando el algoritmo más fuerte de dígitos de verificación del identificador principal; util iza solo dígitos y consonantes menos "l" (letra ell, a menudo confundida con el dígito 1):

Con respecto a la asignación, una estrategia común es aprovechar los identificadores heredados. Por ejemplo, un número de muestra de polilla de museo cd456f9_87 podría anunciarse debajo del ark:/12345/cd456f9_87. Es posible que sea necesario modificar algunos identificadores heredados en vista de las restricciones de caracteres ARK. La segunda estrategia común es crear cadenas completamente nuevas para sus ARK. En este caso, es importante considerar si hacerlos *opacos* o no opacos (o un poco de ambos).

¿Qué son los identificadores opacos?

Las cadenas de identificadores persistentes son típicamente *opacas*, revelando deliberadamente poco sobre a qué están asignadas, porque los identificadores no opacos no envejecen ni viajan bien. Los nombres de las organizaciones son notoriamente transitorios, razón por la cual los NAAN son números opacos. A medida que se corrigen los títulos y las fechas, los significados de las palabras evolucionan (p. Ej., Los acrónimos más inocentes pueden volverse ofensivos o infractores), las cadenas destinadas a ser persistentes pueden volverse confusas o políticamente desafiantes. La generación y asignación de cadenas completamente opacas también conlleva un riesgo, por ejemplo, los números asignados secuencialmente revelan información de tiempo y las cadenas que contienen letras pueden deletrear palabras involuntariamente (razón por la cual faltan vocales en el repertorio de caracteres recomendado).

| Ejemplos de cadenas con un rango de opacidad | | | |
|--|--------------------------------------|---------------------------|-----------------------------|
| no opaco | Archivo permanente de Netscape | Gay_Divorcee_1934_April_1 | Resolvedor de nombre a cosa |
| opaco-ish | x0001, x0002,, x9998 | GD/1934/04/01 | n2t.net |
| opaquer | 141e86dc-d396-4e59-bbc2-4c3bf5326152 | 19340401 | n2t |
| opaquest | 141e86dcd3964e59bbc24c3bf5326152 | h8k74926g | 12148 |

No se requiere que los ARK sean opacos, pero se recomienda que el nombre del objeto base se haga opaco, ya que tiende a nombrar el foco principal de persistencia. Si alguna cadena calificadora sigue ese nombre, es menos importante que sea opaca. Para ayudar a elegir su enfoque de opacidad, es posible que desee considerar la compatibilidad con identificadores heredados y la facilidad de generación y transcripción de cadenas (por ejemplo, brevedad, dígitos de verificación). Se pueden crear nuevas cadenas (minted) con fecha/hora, UUID y generadores de números, así como minters Noid (Nice Opaque Identifiers).

Las cadenas opacas son "mudas" y, por lo tanto, difíciles de manejar, por eso los ARK fueron diseñados para ser identificadores "parlantes". Esto significa que si hay ARK Identifiers FAQ # metadata, un ARK que llega a su servidor con el '?' la inflexión debería poder hablar de sí misma.

¿Cómo hago que el contenido del servidor sea direccionable con ARK?

Primero, decida cuál será la experiencia del usuario al acceder a sus ARK, por ejemplo, un archivo de hoja de cálculo, un PDF, una imagen, una página de destino llena de metadatos formateados y un rango de opciones, etc. Cualquiera que elija, planifique su servidor para poder responder con metadatos si su ARK debería llegar con un '?' inflexión después de eso.

De lo contrario, servir ARK es como servir URL. Normalmente, las cadenas de URL entrantes *direccionan* (se asignan) al contenido que devuelve su servidor web. Si su servidor reconoce ARK, los ARK entrantes (expresados como URL) deben asignarse al mismo contenido. Un enfoque común es asignar el ARK a la URL utilizando una tabla de software que actualiza cada vez que cambia la URL. En este caso, su servidor está actuando como un *res olvedor local*. Si no desea implementar esto usted mismo, existen herramientas y servicios de software ARK que pueden ayudarlo.

Otro enfoque es ejecutar su servidor web sin cambios, pero en lugar de actualizar las tablas locales, actualizaría las tablas de mapeo de ARK a URL que residen en un resolvedor no local. Se pueden encontrar ejemplos de esto entre los proveedores y en cualquier organización que actualice las tablas a trav és de EZID.cdlib.org (que, debido a una relación especial, actualiza las tablas de resolución en n2t.net).

¿Cómo cito o publicito un ARK?

Se prefiere la forma de URL (https o http) del ARK, por ejemplo,

```
https://n2t.net/ark:/99166/w66d60p2
```

Un ARK destinado para uso externo generalmente se publicita (libera, publica, difunde) de esta manera para que sea un *identificador accionable*. Si se necesita una visualización visual más compacta de un ARK, debe estar hipervinculado; por ejemplo, se puede lograr una visualización compacta de un hipervínculo HTML con

```
<a href=" https://n2t.net/ark:/99166/w66d60p2 "> ark:/99166/w66d60p2 </a>
```

Una decisión importante es si sus ARK basados en URL utilizarán el nombre de host de su resolvedor local o el resolvedor N2T.net. Si el control local o el desarrollo de la marca es lo suficientemente importante, anunciaría ARK basados en su resolvedor local (consulte la publicación de contenido con ARK). Si le preocupa la estabilidad de su nombre de host local, anunciaría sus ARK basados en n2t.net (vea ejemplos de ambos).

Resolver sus ARK a través de N2T siempre es posible para los usuarios, independientemente de cómo los anuncie.

¿Existen herramientas y servicios para ayudar con los ARK?

Aquí hay una lista parcial de herramientas de software para identificación persistente que incluye

- · Noid (Nice Opaque Identifiers), software de código abierto para acuñar y resolver ARK por su cuenta
- · ArchivesSpace, aplicación de código abierto para administrar y proporcionar acceso web a archivos, manuscritos y objetos digitales
- Complemento ARK para Omeka, que crea y gestiona ARK para la plataforma de publicación web de código abierto Omeka
- Módulo ARK para Drupal, que permite que su sitio Drupal actúe como una Autoridad de asignación de nombres (NMA)

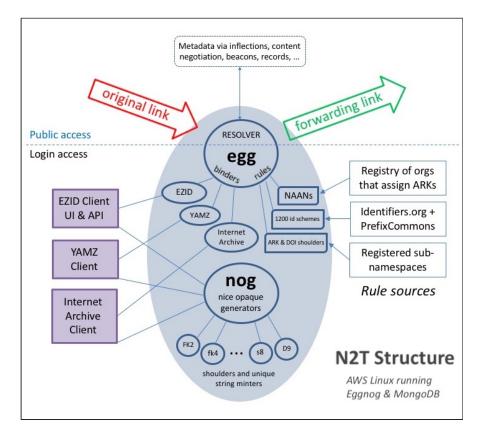
Mas allá de lo básico

¿Qué es el N2T?

N2T.net es un resolvedor ARK global. N2T, que si gnifica "Name-to-Thing" (nombre-a-cosa), es en realidad un resolvedor generalizado para asignar nombres a cosas, por lo que sabe dónde enrutar más de 600 otros tipos de identificadores: ARK, DOI, PMID, Taxon, PDB, ISSN, etc. estás interesado, el diagrama y el resto de esta respuesta dan un poco más de detalle.

Una solicitud llega del público en general como una URL que consiste en "https://n2t.net/" seguido de un identificador (nombre). N2T busca ese identificador y redirige el enlace original a un enlace de reenvío. Para hacer esto, utiliza dos patrones de resolución diferentes. Primero N2T intenta resolver de acuerdo con la información encontrada en un identificador almacenado individual. De lo contrario, N2T intenta resolver de acuerdo con las reglas de clase almacenadas, en función del tipo de identificador. También hay una API N2T que requiere credenciales de inicio de sesión que permite operaciones por lotes y la generación de identificadores únicos.

N2T usa dos tipos de datos almacenados. Primero , almacena registros individuales de más de 20 millones de identificadores de objetos (por ejemplo, ARK, DOI) que obtiene de tres fuentes: E ZID.cdlib.org, Internet Archive y YAMZ.net. Cuand o dichos registros incluyen una URL de redireccionamiento (destino) y metadatos descripti vos de preguntas frecuentes sobre los identificadores ARK, N2T puede actuar sobre las inflexiones de preguntas frecuentes sobre los identificadores ARK, así como realizar el paso de sufijos y la "negociación de contenido".



En segundo lugar, N2T almacena más de 3500 registros de "reglas" para identificadores de enrutamiento que no se encuentran individualmente en N2T, pero para los cuales tiene información de redireccionamiento vinculada al tipo de identificador que se está resolviendo. Obtiene registros de reglas de varias fuentes, incluido el registro NAAN, una base de datos de *hombros* ARK y DOI y una asociación formal sobre identificadores compactos con identifier s.org.

Si la mayoría de los ARK se ejecutan en sus propios resolvedores, ¿por qué existe también un resolvedor global para ARK?

La mayoría de los ARK son creados por organizaciones que los anuncian ("publican") en función de sus propios resolvedores. Por ejemplo, este ARK se publicó en función del resolvedor ark.bnf.fr :

http://ark.bnf.fr/ark:/12148/btv1b8449691v/f29

Tener que ejecutar y mantener su propio resolvedor es el costo de una autonomía completa. El uso de su propio resolvedor también le permite crear marcas a través del nombre de host, la desventaja es que las marcas son transitorias y tienden a hacer que los identificadores sean frágiles. Las presiones políticas e incluso legales (p. Ej., Marcas comerciales) pueden dificultar el soporte de nombres de host de marca más antiguos, por lo tanto, sus identificadores.

Esa es otra razón para tener el resolvedor global ARK. Las personas que se encuentren con un identificador roto en el futuro pueden encontrar que su no mbre de host ya no existe, y si se trata de un ARK, pueden extraer la identidad central (comenzando con "ark:") y presentarla al resolvedor global n2t.net, como en

https://n2t.net/ark:/12148/btv1b8449691v/f29

Mi organización tiene su propio resolvedor ARK. ¿Debería preocuparme por N2T.net?

Sí, por dos razones principales. Primero, si sus ARKs "en estado salvaje" se muestran sin su nombre de host de resolución (lo que significa que comienzan con "ark:...", que no es raro ver), la persona que quiera usarlos no necesitará saber el hostname siempre que puedan recordar agregar "n2 t.net" delante de ellos. Esto funciona porque N2T conoce el nombre de host de resolución correcto.

En segundo lugar, aunque algunas organizaciones y sus nombres de host de resolución son de larga duración, la mayoría no lo son. Una persona que intente utilizar un ARK que contenga un nombre de host de resolución que no funcione puede reemplazar la parte que no funciona con "n2t.net". Si las circunstancias alguna vez lo obligan a cambiar su resolvedor, este paso de reemplazo le da a las ARK que publicó antes del cambio una mejor oportunidad de trabajar.

Para evitar futuros inconvenientes, algunas organizaciones que ejecutan sus propios resolvedores pueden elegir desde el principio suprimir sus nombres de resolvedor y simplemente anunciar ("publicar") sus ARK basados en n2t.net.

¿Por qué el resolvedor ARK global (n2t.net) no tiene la palabra "ARK"?

Cuando surgió la demanda de un resolvedor ARK global, los principios básicos de apertura y generalidad impidieron que los diseñadores crearan otro silo en el molde DOI/Handle/PURL. En cambio, el resolvedor ARK fue construido para ser un resolvedor genérico, independiente del esquema, llamado N2T (Name-to-Thing), que ahora resuelve más de 600 tipos de identificadores, incluidos ARK, DOI, Handles, PURL, URN, ORCID, ISSN, La resolución es esencialmente buscar en una tabla una cadena de identificación, independientemente del tipo, y redirigirla al lugar correcto.

Los mismos principios básicos guiaron el diseño de una herramienta anterior llamada noid, que fue construida para ARK pero también es utilizada regularmente por organizaciones que manejan Handles.

¿Qué significa "paso de sufijo"?

Brevemente, el paso de sufijo es una característica de N2T. Supongamos que solo tiene un ARK registrado, https://n2t.net/ark:/12345/6789, y lo redirige a la página del servidor web,

```
https://a.example.org/dataset542
```

Y supongamos que el mismo servidor también sirve estas páginas:

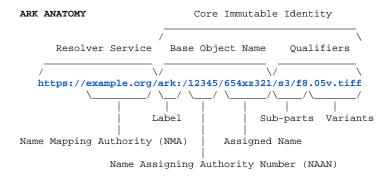
```
https://a.example.org/dataset542/volume3
https://a.example.org/dataset542/volume3/part2
https://a.example.org/dataset542/volume3/part2.pdf
```

Lo que hace el sufijo es dejar que su ARK registrado actúe como si también hubiera registrado estos tres ARK a continuación, lo que resolvería las páginas anteriores, respectivamente:

```
https://n2t.net/ark:/12345/6789/volume3
https://n2t.net/ark:/12345/6789/volume3/part2
https://n2t.net/ark:/12345/6789/volume3/part2.pdf
```

En este caso, el paso de sufijo le ahorró tener que mantener registros para tres páginas más. De hecho, funciona para un número ilimitado de páginas.

¿Cuáles son las partes de un ARK?



¿Puedo asignar ARK a cosas dentro de algo que ya tiene un ARK?

Sí, los ARK se pueden asignar a cualquier nivel de *granularidad*, como un manuscrito, capítulos dentro de él, secciones de capítulos, subsecciones, etc. Un ARK también se puede asignar a una cosa que encierra otras cosas. En los ARK, el carácter '/' está reservado para ayudar al destinatario a comprender la contención, por ejemplo, el primer objeto a continuación contiene el segundo:

```
ark:/12148/btv1b8449691v
ark:/12148/btv1b8449691v/f29
```

Ese es el calificador de contención. Solo hay otro calificador ARK, e indica formas variantes de una cosa usando el carácter reservado ". delante de un sufijo Por ejemplo, si estos ARK identifican documentos,

```
ark:/12148/btv1b8449691v/f29.pdf
ark:/12148/btv1b8449691v/f29.html
```

debido a que difieren solo por el sufijo .pdf o .html, se puede inferir que identifican dos formas diferentes del mismo documento.

¿Cuál es el propósito de la NAAN, y puedo hacer cambios a ella?

Los NAAN subdividen el conjunto de todos los ARK posibles (el espacio de nombres ARK). El subconjunto de ARK bajo una NAAN dada se puede subdividir en hombros (por ejemplo, 12345/x2, 98765/b4), lo que puede facilitar la delegación de la asignación de ARK autónoma a los departamentos de una organización grande. La resolución ARK se basa libremente en los NAAN, pero debido a que las organizaciones se dividen, los ARK se acomodan al problema de división del espacio de nombres al apoyar la gestión de un espacio de nombres por más de una organización. Si hace la transición dentro o fuera de una relación de proveedor, no hay impedimento para llevar su NAAN con usted.

Puede cambiar una NAAN completando el mismo formulario en línea utilizado para solicitar una nueva NAAN. Ejemplos de razones para un cambio pueden incluir

- notificar a N2T que la persona de contacto de su organización o la URL de resolución cambiarán,
- actualizar la política de asignación de nombre de su organización (política de muestra),
- solicitando una NAAN adicional para un nuevo cuerpo significativo de ARK o una nueva división organizacional, y
- haciendo la transición de su NAAN a otra organización que continuará con su trabajo y se hará cargo de su NAAN.

¿Existen restricciones en el uso de NAAN?

Sí, es importante nunca inventar o usar una NAAN que no esté incluida en el registro público. Sin embargo, hay dos NAAN especiales que cualquiera puede usar:

- 99999, para "prueba", "desarrollo" o ARK experimentales, y
- 12345, para ARK no funcionales que aparecen en la documentación.

Para las personas con suficiente capacitación, es fácil reconocer y eliminar los ARK con estos NAAN de los informes de enlaces rotos que deben tratarse. A pesar de los mejores esfuerzos de los proveedores, estos ARK con frecuencia "escapan a la naturaleza", donde terminan confundiendo a los usuarios y los verificadores de enlaces.

ARK y otros identificadores

¿Por qué usaría ARK en comparación con, por ejemplo, DOI?

- Para mantener bajos los costos (detalles).
- Para trabajar exactamente con los metadatos que desea.
- Para poder crear identificadores sin metadatos.
- Para poder crear un identificador incluso antes de que exista su objeto.
- Tener un identificador tan pronto como cree el primer borrador de sus datos.
- Mantener ese identificador privado mientras los datos y metadatos evolucionan, y decidir (quizás años) más tarde, publicarlo o descartarlo.
- Para retener ese identificador después de la publicación, quizás luego asigne un identificador adicional, como un DOI.
- Debido a que los ARK, creados para aplicaciones genéricas y no específicamente para contenido publicado, encajan naturalmente con objetos físicos como muestras o estaciones de campo.
- Debido a que los resolvedores de ARK pueden lidiar con identificadores dañados rutinariamente en el mundo mediante procesos de formateo de texto que introducen guiones.
- Debido a que la mayoría de los ARK llevan un dígito de verificación de Noid que se puede usar para detectar todos los errores de transcripción comunes en lugar de solo algunos de ellos.
- Para poder crear identificadores más cortos, ya que las mayúsculas y minúsculas permiten cadenas más densas (un mayor número de cadenas de una longitud dada).
- Para poder cambiar el proveedor y/o la infraestructura sin tener que coordinar las transferencias de bases de datos con una autoridad central.
- Para poder lidiar con el problema de división del espacio de nombres sin perder el control de sus identificadores.
- Para vincular identificadores a diferentes tipos de compromisos de persistencia matizados.
 Para poder agregar consultas (por ejemblo 2 Lang en) al resolver sus identificadores.
- Para poder agregar consultas (por ejemplo,? Lang = en) al resolver sus identificadores.
- Utilizar una infraestructura abierta coherente con los valores de su organización.
 Para vincular directamente a los objetos que valora en lugar de a las páginas de destino.
- Para crear un identificador que permita millones (paso de sufijo).
- Para acceder a metadatos convenientes y con todas las funciones a través de las identificaciones de preguntas frecuentes # de identificadores ARK.
- Para integrarse fácilmente con API IIIF utilizando calificadores ARK.

¿Qué tienen en común ARK, DOI, Handle, PURL y URN?

Estos son los principales tipos de identificadores persistentes (o esquemas).

- Todos han existido desde 2001 o antes.
- Todos se encuentran en lugares como los perfiles de Data Citation Index , Wikipedia y ORCID.org.
- Todos dan acceso a casi cualquier tipo de cosa, ya sea digital, física, abstracta, persona, grupo, etc.

También tienen una estructura muy similar, como se ve en los ejemplos a continuación, que consta de cuatro partes:

```
https://n2t.net/ark:/99999/
12345
https://doi.org/10.99999/1
2345
https://handle.net/10.99999
/12345
https://purl.org/99999/123
https://purl.org/99999/123
https://cvarious>/urn:99999
:12345
```

Y todos tienen poco efecto sobre la persistencia. Vea 10 mitos persistentes sobre identificadores persistentes.

Espera, ¿estás diciendo que ARK, DOI, Handle, PURL y URN son inútiles?

No, esa es una declaración demasiado fuerte. Pero mantengamos estos esquemas de identificación (tipos) en perspectiva.

- Todos ellos no logran detener las causas principales de los enlaces rotos: pérdida de fondos, desastres naturales, agitación social, guerra, remoción deliberada, error humano y negligencia del proveedor.
- Todos requieren que usted, el proveedor final, actualice las tablas de reenvío a medida que cambian las URL.
- Todos identifican contenido que está sujeto a cambios o eliminación en futuras visitas.
- · Todos tienen identificadores que se rompen regularmente y en grandes cantidades, muchos miles y más.
- Todos confían en la redirección ordinaria integrada en los servidores web desde 1994 y proporcionada de forma gratuita por cientos de servicios de acortamiento de URL.

Dado lo poco que hacen los esquemas por usted, al elegir uno, es probable que desee considerar factores como el costo, el riesgo y la apertura.

¿En qué se diferencian los ARK de los identificadores como DOI, Handles, PURL y URN?

La respuesta corta

Los ARK son los únicos identificadores convencionales, sin depósito y sin pagos que puede registrarse para usar en aproximadamente 48 horas. Los DOI, Handle y PURL requieren resolución y otros servicios que provienen de sus respectivos sistemas centralizados (depósitos).

Eso no quiere decir que la persistencia sea gratuita. Hacer que cualquier identificador sea persistente lo carga a usted, el proveedor, con los costos de administración de contenido, alojamiento, monitoreo y reenvío. Puede hacer esas cosas usted mismo o con la ayuda de un proveedor. Pero con las ARK, al igual que con las URL, no se le cobrará por separado por sus identificadores y no se lo bloqueará en un depósito de resolución de propósito especial que también bloquea otros identificadores.

Los ARK son inusuales en ser descentralizados. Si bien uno *puede* obtener servicios de resolución de un resolvedor ARK global llamado n2t.net, más del 90% de los ARK en el mundo se publican sin usar n2t.net en el URL hostname. Más de 600 organizaciones registradas en todo el mundo han creado, entre ellas, un estimado de 3.200 millones de ARK y, al igual que con las URL, nadie ha pagado una tarifa de identificación para crearlas. Por supuesto que *mantenerlos* no es gratis. Nunca es gratuito mantener el acceso al contenido persistente a largo plazo, independientemente del tipo de identificador.

Más diferencias entre ARKs, DOIs, Handles, PURLs y URNs

- Páginas de destino: los DOI Crossref y DataCite enlazan con páginas de destino de editores construidas alrededor, pero no directamente, de los
 objetos que le interesan, pero los ARK pueden conectarse libremente directamente a los objetos que le interesan, lo que es amigable para las
 máquinas y los humanos, ya que no requiere un extra paso de navegación humana para tareas comunes como
 - o abrir el archivo PDF de un artículo para leerlo,
 - º haciendo referencia a un archivo de imagen destinado a incorporarse automáticamente en línea en un documento, y
 - o citando una hoja de cálculo que se utilizará para el análisis directo de datos por software.
- Los DOI, los otros identificadores, etc. no admiten las inflexiones de preguntas frecuentes sobre los identificadores ARK de estilo ARK que
 permiten el acceso a los metadatos, independientemente de si un identificador apunta a un objeto o su página de destino.
- A diferencia de los DOI y los controladores, los ARK no tienen requisitos de metadatos. Los ARK que no se han lanzado al mundo son fáciles de eliminar.
- Todas las cosas finalmente pasan, incluidos los nombres de host y la propia web y el protocolo " https://". Cuando esa primera parte del identificador deja de tener significado, solo los ARK y los URN incluirán la etiqueta (por ejemplo, "ark:") que indica el tipo de identificador que queda.
- · Para DOI, Handle y PURL, debe usar sus respectivos resolvedores. ARKs y URNs, le permiten usar su propio resolvedor.
- Para crear DOI y Handle, se le exige que pague una tarifa de membresía y, para los DOI, las agencias de asignación imponen los cargos por DOI de varias maneras. No hay tarifas para ARK, PURL y URN.
- Para crear Handles, debe instalar y mantener un servidor Handle local, que le brinda otro sistema para monitorear, parchear y solucionar problemas.
- Aunque puede usar un resolvedor local o de proveedor para sus ARK y URN, los ARK pueden resolverse a través del resolvedor global n2t.net.
- La infraestructura de resolución de URN prevista nunca se construyó, por lo que los URN se resuelven actualmente como URL y no hay un resolvedor global de URN como URL designado. Para registrarse para crear URN, debe solicitar un espacio de nombres URN.
- Los ARK tienen algunas características únicas que admiten el desarrollo temprano de objetos: los ARK se pueden eliminar, pueden nacer sin metadatos y pueden existir con cualquier metadato que desee almacenar.

Pero si se pueden eliminar los ARK, ¿cómo se puede confiar en ellos?

Realmente hace que los ARK sean más confiables. La capacidad de eliminar es una parte vital de una gestión de colección conveniente que se niega a aquellos tipos de identificadores que no son ARK que prohíben la eliminación bajo la presunción de que las personas, una vez que se les pide que se comprometan, no cometerán errores. Las personas armadas con software de gestión de identificadores convierten regularmente errores humanos simples en errores a gran escala, incluso en el umbral del compromiso. Al dificultar su limpieza, obligamos a los sistemas a arrastrar esos problemas a perpetuidad.

Si bien no son inmunes a tales errores, los ARK tienen la gran ventaja de que pueden crearse y eliminarse en las sombras, independientemente de su lanzamiento, publicación o compromiso de archivo.

¿Puede un objeto tener un ARK y un DOI?

Si. A veces es útil tener dos identificadores, aunque puede volverse confuso cuando sucede con frecuencia. Muchas personas comienzan asignando ARK a cada cosa que crean para tener una referencia estable desde el principio, incluso antes de saber si quieren publicarla, y mucho menos conservarla.

El objeto y sus metadatos se desarrollan juntos, y para el subconjunto de cosas que finalmente desea publicar en lugares que requieren DOI, puede asignar DOI en el momento de la publicación. Si su ARK es estable y tiene metadatos básicos, ya está haciendo todo lo que necesita para admitir un DOI adecuado. Esta es una forma en que los ARK admiten el desarrollo temprano de objetos.

Para admitir dos identificadores de manera eficiente, se recomienda que cree el DOI de modo que redirija al ARK original. Esto no solo elimina la necesidad de actualizar la redirección de DOI, sino que también mantiene el ARK persistente para cualquiera que lo haya grabado o marcado previamente.

¿Cuándo debo usar ARK en comparación con DOI, Handles, PURL o URN?

No hay respuestas simples. Los identificadores (no las cosas, sino sus nombres) son difíciles de describir, por lo que si escucha respuestas simples en otro lugar, tenga cuidado con las falacias comunes.

Nada inherente en ARK, DOI, Handles, PURL o URN los hace más o menos adecuados para un campo, dominio o sector en particular. Con un identificador de resolución y un servicio de administración administrativa, todos brindan el servicio central de resolución (y también lo hacen las URL administradas adecuadamente).

Las generalizaciones sobre los tipos de identificadores a veces se aplican cuando la resolución y la administración de ese tipo están bloqueadas en un proveedor o proveedor en particular. Por ejemplo, muchas características y restricciones de PURL y Handle están bien definidas por sus respectivos silos de administración, al igual que los de DOI, que se construyen sobre los Handles. Pero los DOI tienen prácticas de metadatos que son diversas y evolucionan en diferentes agencias de registro de DOI.

Las diferencias concretas que experimentamos, como los *metadatos*, las páginas de destino y la integración de herramientas (por Ej., Herramientas de publicación), no son propiedades de los esquemas de identificadores per se, sino propiedades de resolución, administración y servicios de citas que varios proveedores extienden o retienen. de diferentes tipos de identificadores. Esos servicios están conformados a su vez por las comunidades de práctica y los mercados. Los servicios básicos se basan en una base de datos confiable que almacena cada identificador junto con elementos de metadatos (creador, título, fecha, URL de redireccionamiento, etc.) que describen el objeto identificado. Los servicios adicionales incluyen verificación de enlaces, detección de duplicados, generación de informes y búsqueda.

De la cuna a la tumba

¿Cuándo en mi flujo de trabajo debo crear ARK?

Al nacer el objeto, o incluso antes. A veces nombramos a nuestros bebés antes de que nazcan, y nombramos y hacemos referencia a objetos en las etapas de concepción, a veces mucho antes de que den fruto. Dependiendo de cuán elaborada sea la planificación, sus objetos no nacidos podrían tener ARK de función completa que se resuelven en un sustituto apropiado y devuelven metadatos enriquecidos, incluidas las declaraciones de persistencia.

La única advertencia es tener cuidado al liberar ARK (publicidad) que tengan perspectivas inciertas a largo plazo. Algunos sistemas de administración de identificadores tienen características para ayudar a administrar y resolver identificadores inéditos (por ejemplo, EZID tiene un estado "reservado"). Cuantas más personas conozcan un ARK, más difícil será eliminarlo.

¿Cómo es que los ARK pueden ser fáciles de eliminar?

Si nadie conoce un identificador que no sea usted, no hay ningún daño en eliminarlo o retirarlo. Retrocediendo, un identificador es en realidad una afirmación de que una cadena de caracteres dada está asociada con algo específico. A cuantas menos personas le digas, más fácil será desechar esa afirmación. Si crea una URL y la comparte solo con sus colegas más cercanos, es mucho más fácil de retirar que si la URL apareciera durante un mes en un sitio web público, del cual fue obtenida por los motores de búsqueda de Internet. Por el contrario, es difícil eliminar DOI y Handle porque una vez registrados y resueltos, se lanzan efectivamente al mundo.

Los ARK se comportan como URL a este respecto. Los proveedores son libres de crear y compartir ARKs por poco, en cuyo caso son fáciles de eliminar.

Quizás sorprendentemente, incluso si se comparten de manera más amplia, los ARK deberían venir con declaraciones de persistencia que le digan cuánto o cuán poco compromiso se les hace. Los ARK fueron diseñados para articular una variedad de declaraciones de persistencia, pero ciertamente no están solos entre los identificadores y objetos que exhiben una variedad de "sabores" de compromiso. Esta es la razón por la cual los ARK se conocen como identificadores de alto funcionamiento que son buenos para la persistencia en lugar de "identificadores persistentes".

Finalmente, la gente comete errores. Los ARK, DOI, Handles, PURL y URN a veces se transmiten por error y deben retirarse. Cuando eso sucede, la mejor práctica del proveedor es hacer que el identificador retirado se resuelva en una página de "lápida" que explica y tal vez se disculpe por las molestias. A pesar de los rumores, los identificadores persistentes nunca están garantizados.

¿Qué se entiende por ARKs que apoyan el desarrollo temprano de objetos?

Las personas necesitan identificadores antes de saber exactamente a qué objeto se refieren, o si se refieren a algo que valga la pena conservar. No se puede crear un identificador que requiera metadatos consolidados durante el desarrollo temprano ya que se sabe poco sobre el objeto. Por lo tanto, los creadores de objetos casi siempre asignan inicialmente identificadores que no tienen requisitos de metadatos, como URL o ARK.

Si comienza con un ARK, se beneficia de poder mantener el identificador original hasta su lanzamiento público a medida que los metadatos maduran. Muc hos objetos pasan por fases intensivas de desarrollo y revisión, que a veces duran años, durante los cuales son demasiado inmaduros para cumplir con la mayoría de los requisitos de metadatos. Sin embargo, cada objeto necesita algún tipo de identificador desde la concepción hasta la madurez, donde la madurez podría parecer una publicación pública y una mejora adicional o abandono. Es fácil abandonar los ARK que no se han lanzado al mundo.

Al igual que el objeto en sí, los elementos de metadatos necesitan un lugar flexible para crecer y madurar con el tiempo:

- comenzando en la fase de planificación, cuando solo necesita un identificador,
- en el momento del nacimiento, cuando su primera representación digital necesita una URL de destino de redireccionamiento,
- después del primer análisis, cuando emerge su significado y un título provisional,
- al crear docenas de elementos de metadatos específicos de la disciplina que violan la mayoría de los estándares de metadatos, excepto los suyos,
- durante el procesamiento posterior por un colega cuyo nombre agregará como creador adicional,
- cuando la retroalimentación temprana basada en el identificador tuiteado revela una idea clave y un nuevo contribuyente,
- y así sucesivamente, hasta el archivo, abandono, publicación pública, corrección, revisión, mejora, etc.

A diferencia de los DOI Crossref y DataCite, que requieren metadatos específicos (por ejemplo, ver el esquema DataCite), los ARK no limitan ninguna de estas actividades. Además, el resolvedor N2T.net realmente los admite a todos.

Si los ARK no lo requieren, ¿por qué molestarse en crear metadatos?

Crear metadatos (información adicional asociada o que describe un objeto) tiene varios beneficios clave. En primer lugar, no importa lo que los vuelve a dirigir a ARK -ya sea una página de destino o un archivo- metadatos ofrece a los usuarios información vital sobre el objeto, como referencias a las versiones más recientes, fecha de creación, procedencia, etc. Por lo general ARKs metadatos se accede a través ARK Identificadores Preguntas frecuentes # inflexiones.

Los metadatos realmente alivian la dificultad de trabajar con identificadores opacos, que no revelan pistas sobre lo que identifican. En ausencia de metadatos, se ve obligado a acceder al objeto en sí para recordar qué es y también a confiar en que está accediendo al objeto correcto. Además, las discrepancias entre los metadatos devueltos y el objeto al que se accede ayudan a todos a detectar cambios de identificador y errores.

Los metadatos son para adultos y son mucho menos importantes para los objetos inmaduros y sus identificadores que para los que han sido liberados. Te ner metadatos demuestra credibilidad básica del proveedor y compromiso con identificadores de alto funcionamiento. No todos los proveedores están a la altura de esta tarea.

No tiene por qué ser costoso. Construir metadatos desde cero puede ser costoso, pero generalmente es creado y administrado por proveedores de objetos, en cuyo caso se puede aprovechar de manera eficiente para los identificadores. Idealmente, para una fuerte persistencia, los metadatos maestros (mantenidos por proveedores de objetos) deberían reflejarse en sistemas independientes, de modo que sea difícil para alguien manipular sin detección las asociaciones de identificadores. Por ejemplo, los repositorios de objetos digitales que obtienen ARK y DOI del servicio EZID almacenan una copia de sus metadatos con EZID.cdlib.org, que a su vez almacena otra copia con el resolvedor N2T.net.

¿Qué metadatos se recomiendan para los ARK?

Los metadatos son datos desordenados para todos los identificadores, no solo para ARK. En todos los dominios y tipos de objetos hay miles de estándares, muchos de ellos superpuestos pero conflictivos, y cada uno se aplica de acuerdo con las costumbres organizacionales locales y con diferentes niveles de cumplimiento. Elegir o crear una especificación para sus metadatos depende de factores como

- si actualmente está administrando metadatos (sugerencia: quédese con él a menos que tenga una buena razón para cambiar),
- si desea publicar objetos oficialmente (sugerencia: prepárese para poder proporcionar autor, título, fecha, editor/archivo y tipo de objeto),
- los requisitos y capacidades de su resolvedor (sugerencia: su personal de TI o proveedor podría tener sus propios requisitos), y
- si desea almacenar elementos no estándar (sugerencia: N2T lo permite, pero la mayoría de los estándares y proveedores no).

La interoperatibilidad confiable entre dominios puede permanecer fuera del alcance, pero Dublin Core, DataCite, Schema.org y Dublin Kernel son especificaciones de metadatos comunes a tener en cuenta para su uso con ARK.

¿Por qué veo metadatos ARK con las etiquetas de quién, qué, cuándo y dónde?

Los ARK fueron diseñados para identificar cualquier cosa, no solo cosas que son, por ejemplo, publicables o que se pueden comprar. No es natural modelar un fósil, una muestra de tejido, un término de vocabulario o Marie Curie como si cada uno tuviera un Autor, Título, Editor, Copyright y Precio. En cambio, desde 2001, un ARK generalmente tiene un núcleo de cuatro elementos de metadatos altamente genéricos (Dublin Kernel, inspirado en Dublin Core (DC)), seguido de cualquier otro elemento de metadatos (pares de nombre/valor) que el proveedor desee proporcionar.

Los metadatos del kernel están estructurados como si respondieran a las preguntas, quién, qué, cuándo y dónde con respecto a la expresión o "revelación" de un objeto:

- quién "lo contó" (similar a DC Creator, contributor y Publisher, pero también Inventor, Descubridor, Conductor, etc.),
- qué se llamaba "tell" (similar a DC Title, pero también TissueSampleNumber, ArtifactBarcode, etc.),
- cuando fue "dicho" (Fecha DC similar, pero incluye rangos de fechas, fechas aproximadas y BCE),
- · dónde se puede encontrar el "relato" (desde DC Identifier, pero generalmente no es necesario porque este es el ARK)

Hay mucho más que decir sobre los metadatos ARK, por ejemplo, aplicar quién, qué, cuándo y dónde al contenido de una biografía, o cómo un archivo planea soportar un conjunto de datos. Más pautas de metadatos ARK estarán disponibles en arks.org. Otros elementos son clave, como

• cómo se "contó" (similar a ResourceType), que puede dictar asignaciones a especificaciones de metadatos externos y elementos adicionales

- URL de destino de redirección, que generalmente se almacena como un elemento distinguido de metadatos
- elementos que contienen declaraciones de persistencia, para expresar la fuerza o debilidad de un compromiso de archivo

¿Qué es una "inflexión" de ARK y en qué se diferencia de la "negociación de contenido"?

Una *inflexión* es un cambio al final de una palabra para expresar un cambio en el significado. Nos permite definir una palabra como "ir" sin definir también "va" y "va". A un ARK que conduce a un objeto, simplemente agregando un '?' hasta el final (el '?' es un ejemplo de una inflexión ARK) nos permite solicitar metadatos sin tener que definir un identificador separado para los metadatos del objeto. Esta simple técnica puede ser utilizada por un humano con un navegador web. El resolvedor N2T admite inflexiones y negociación de contenido.

La negociación de contenido para metadatos es una técnica de software para solicitar formatos alternativos de un objeto, como el formato PDF o RTF de un archivo HTML. Aunque no fue diseñado para ello, la "negociación de contenido" histórica fue criticada (rebuscada) en ciertos contextos para solicitar metadatos bajo el supuesto sorprendente de que los formatos utilizados a menudo para contener metadatos son de hecho metadatos y nunca serán objetos por derecho propio. A diferencia de las inflexiones, la "negociación de contenido para metadatos" no funciona en absoluto para los *objetos* represe ntados en esos formatos (cuya lista está creciendo y se conoce solo por acuerdo privado), ni es lo suficientemente fácil como para ser utilizada directamente por la mayoría de los usuarios humanos.

Aunque las inflexiones se asocian comúnmente con ARK, no son "propiedad" de ARK. Contrariamente a la creencia popular, los identificadores no *hacen* nada, es su resolución de que *hacen* o *no* soportan tales características. Así, por ejemplo, inflexiones y el sufijo de paso son compatibles con n2t.net para todos los tipos de identificadores, pero no por doi.org o handle.net (que tiene una funcionalidad relacionada llamada handle de plantilla) para cualquier tipo de identificador.

¿Qué quieres decir con depósitos?

Por lo general, los servicios basados en esquemas están diseñados como depósitos o repositorio o plataformas cerradas, que sirven a un tipo de identificador particular, como Handle, DOI o PURL. Cada depósito o repositorio realiza las mismas funciones principales: asignar nombres (cadenas de identificadores) a cosas (objetos o metadatos). Excluir todos menos un tipo de cadena de identificación puede ayudar a capturar mercados, pero es un desperdicio y no incluye. Requiere construir el mismo conjunto de servicios una y otra vez para cada tipo y viola los principios básicos de apertura.

En contraste, el resolvedor N2T (Name-to-Thing) y la interfaz de administración EZID (identificadores fáciles) fueron diseñados para funcionar con todos los identificadores. El esfuerzo puesto en cualquier nueva característica se puede aprovechar de manera eficiente en todos los tipos, lo que a veces crea una flexibilidad sorprendente. Por ejemplo, los ARK a menudo se almacenan en EZID con "metadatos DOI", y cada DOI almacenado en N2T puede beneficiarse de las "características de resolución ARK", como las inflexiones y el paso de sufijos, que no están disponibles a través del resolvedor principal DOI (doi.org).