



**HAL**  
open science

## ARK is in the Air: ARKs Trending in the French-speaking Area and the BnF's Role in the ARK Story

Bertrand Caron, Sébastien Peyrard, Jordan de La Houssaye, Stéphane Reecht, Thomas Ledoux, Jean-Philippe Tramoni

### ► To cite this version:

Bertrand Caron, Sébastien Peyrard, Jordan de La Houssaye, Stéphane Reecht, Thomas Ledoux, et al.. ARK is in the Air: ARKs Trending in the French-speaking Area and the BnF's Role in the ARK Story. iPRES - 15th International Conference on Digital Preservation, Sep 2018, Boston, MA, United States. 10.17605/OSF.IO/R6ZJY . hal-02143045

**HAL Id: hal-02143045**

**<https://bnf.hal.science/hal-02143045>**

Submitted on 29 May 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ARK is in the Air: ARKs Trending in the French-speaking Area and the BnF's Role in the ARK Story

Bertrand Caron  
Bibliothèque nationale de France \*  
bertrand.caron@bnf.fr

Jordan de La Houssaye  
Bibliothèque nationale de France \*  
jordan.de-la-houssaye@bnf.fr

Thomas Ledoux  
Bibliothèque nationale de France \*  
thomas.ledoux@bnf.fr

Sébastien Peyrard  
Bibliothèque nationale de France \*  
sebastien.peyrard@bnf.fr

Stéphane Reeceht  
Bibliothèque nationale de France \*  
stephane.reeceht@bnf.fr

Jean-Philippe Trameni  
Bibliothèque nationale de France \*  
jean-philippe.trameni@bnf.fr

\* Paris, France

## ABSTRACT

In recent years we have seen a growing adoption of Archival Resource Key (ARK) identifiers in France and in French-speaking countries, a growing reliance on National Library of France (BnF) ARKs for data dissemination, and a growing demand for data permanence from the BnF community – the French libraries. The BnF has acknowledged these changes through renewed efforts in education, communication, and development of good practices. This progress report describes the recent upsurge of interest in ARKs, the main difficulties posed by identifier management for the BnF and its community, and finally the “ARK Summit 2018”, a major event in the construction of a francophone community of interest around ARK identifiers.

## KEYWORDS

ARK, Persistent identifier, Resolver, Community building

## INTRODUCTION

In 2014 the National Library of France (BnF) and the California Digital Library (CDL) published a paper entitled “*The ARK Identifier Scheme: Lessons Learnt at the BnF and Questions Yet Unanswered*” [11]. Looking back over eight years of using Archival Resource Key (ARK) identifiers [9], the paper described how the BnF had gained some solid experience in dealing with functional, technical, and organizational issues in identifier curation. Since 2014, the BnF has made consistent efforts to inform cultural institutions of the advantages of the ARK identifier scheme, and adoption has progressed significantly in France – the number of registered French-speaking organizations saw a 230% increase over the last four years. During the same period, the BnF extended its use of ARKs to new areas.

Interestingly, these two trends led the BnF, as well as other organizations, to face several new questions. ARKs began to be seen as the identifier of choice for internal BnF resources and for bridging internal applications. Meanwhile, users began citing BnF ARKs extensively and started building cultural services on top of them. This forced the BnF to review its practices and technical framework for assigning and resolving identifiers, to evaluate its capacity to actually maintain access to resources, and to clarify its public promises regarding stewardship. On their side, libraries, archives, museums, publishers, and other private sector organizations began creating

ARKs with little knowledge of the cost and commitment that the maintenance of long-term identifiers implied. Often the cost and commitment for these institutions became more muddled because control of their identifiers was the responsibility of third-party data dissemination system providers.

In order to better understand the community and to share the benefit of its own experiences, the BnF, with support from the CDL, organized an international ARK Summit<sup>1</sup> in March 2018, targeted at the French-speaking community. Its goals were to get a sense of the impact and adoption of ARKs in France, to establish more consistent and robust guidelines for assignment and resolution mechanisms, to help new ARK name assigning authorities get up to speed, and to provide a place for all French-speaking institutions to begin gathering into a community interested in long-term identifiers. Discussions with participants highlighted the lack of specific guidance on ARK implementation. They also raised awareness of general, persistent access issues and of the role that the BnF, as a leader of the professional community of French libraries, could play as a national-level provider of expertise.

## 1 2014-2018: ARKS GAINING GROUND

### 1.1 Increasing Francophone Adoption

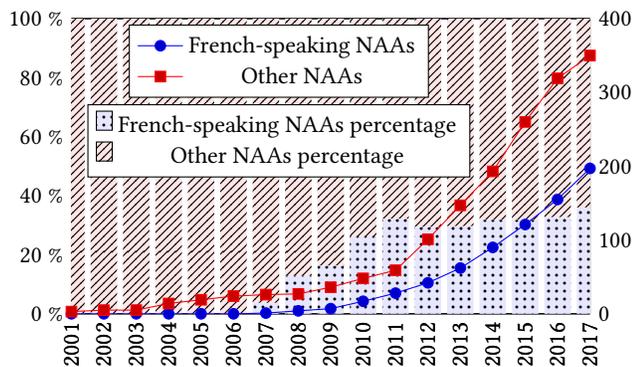
In March 2018 there were more than 210 French-speaking organizations (95% of which were French) registered in the ARK registry [12], among a total of 550 registered organizations. The percentage of French-speaking institutions has been growing steadily, from 13% in 2008, to 29% in 2013 and 36% in 2017. Figure 1 shows this progression.

Indeed, in 2014 a task force on persistent identifiers was established by the French Ministry of Culture, with representatives from various cultural heritage organizations (libraries, archives, museums), which resulted in a “*Vademecum on persistent identifier implementation for data producers*” [6] published in April 2015. At the same time, the 2015 revision of the “*General Reference Framework for Interoperability*” for the public sector [2] mentioned the ARK system as one of the three recommended standards for identifiers, along with URIs and ISNIs.

Meanwhile, in December 2014 the BnF was invited to present the ARK system at a workshop<sup>2</sup> of the European DASISH project, dedicated to persistent identifiers. The subject of community adoption and advocacy for ARKs was explicitly raised: the lack of a fully structured community of potential adopters could turn to seem to

<sup>1</sup>[http://www.bnf.fr/en/m\\_anx\\_articles/a.jp\\_180321\\_ark.html](http://www.bnf.fr/en/m_anx_articles/a.jp_180321_ark.html)

<sup>2</sup><http://dasish.eu/dasishevents/pidworkshop/presentations/>



**Figure 1: Percentage of French-speaking Name Assigning Authorities (NAAs) in the ARK registry.**

have hindered ARK adoption, in spite of the high quality of the specification and overall system. This highlighted the need for ARK advocacy, an activity for which the BnF, as a major implementer in Europe, was seen to be naturally suited. The BnF took on this task in France more energetically from that date on, and the ARK system was thus introduced to various organizations representing different sectors in 2016 and 2017. The BnF also introduced ARKs to the Inist (French Institute for Scientific and Technical Information), which built an on-line infrastructure for e-publications from the research sector for the ISTE<sup>3</sup> project, a French program of acquisitions of scientific resources (21 million in early 2018). The BnF also presented the ARK system to the French Ministry of Education, which runs Eduthèque<sup>4</sup>, a portal aggregating educational resources from public and private stakeholders for the benefit of teachers and students. In the education and research sectors, these outreach efforts have been seminal in advocating for persistent identifiers in general and ARKs in particular. The BnF also conducted training sessions with its digital library partners in 2016. Last but not least, it gave a presentation to the archivist community at the triennial French Archivists Forum<sup>5</sup>. The session for the archivist community was slightly different, since archival institutions in France have been using ARKs for a long time. This is related to early adoption of ARKs by vendors of records management systems and archival portals, which supported ARK as the default identifier. This meant that each new institution signing up with a vendor was registered as a distinct new organization in the Name Assigning Authority Number (NAAN) registry<sup>6</sup> [12]. These sessions aimed at raising awareness of the role of the institution in identifier management and persistence policy definition [7, 8].

## 1.2 ARKs at the BnF: a Complex Landscape

At the end of the 2000s, the BnF as a Name Assigning Authority (NAA) had defined two major “subnaming authorities” that assigned ARKs independently: the catalog of bibliographic records and Gallica, the digital library of the BnF. These sub-NAAs effectively assigned ARKs within two separate namespaces. Rather than implementing a single resolver having full knowledge of all the

<sup>3</sup><http://inist-registry-3.ark.inist.fr/>

<sup>4</sup><http://www.edutheque.fr/accueil.html>

<sup>5</sup><http://forum2016.archivistes.org/>

<sup>6</sup>The registry reserves a unique 5-digit number for each assigning authority to begin its ARKs, e.g., BnF’s look like [ark:/12148/...](http://ark:/12148/...)

Subnaming authority	Name Mapping Authority (NMA)											
	<a href="http://mandragore.bnf.fr">mandragore.bnf.fr</a>	<a href="http://archivesetmanuscrits.bnf.fr">archivesetmanuscrits.bnf.fr</a>	<a href="http://data.bnf.fr">data.bnf.fr</a>	<a href="http://catalogue.bnf.fr">catalogue.bnf.fr</a>	<a href="http://medaillesetantiques.bnf.fr">medaillesetantiques.bnf.fr</a>	<a href="http://reliures.bnf.fr">reliures.bnf.fr</a>	<a href="http://presselocaleancienne.bnf.fr">presselocaleancienne.bnf.fr</a>	<a href="http://bpl6.bnf.fr">bpl6.bnf.fr</a>	<a href="http://gallica.bnf.fr">gallica.bnf.fr</a>	<a href="http://gallica.intramuros.bnf.fr">gallica.intramuros.bnf.fr</a>	<a href="http://images.bnf.fr">images.bnf.fr</a>	<a href="http://classes.bnf.fr">classes.bnf.fr</a>
Mandragore	X											
Archives et manuscrits		X	X									
Catalogue Général			X	X								
reliures.bnf.fr						X						
Médailles et antiques					X							
Digitized documents entry chain									X	X	X	
SPAR									X	X	X	
Classes												X

**Figure 2: BnF subnaming authorities and their corresponding Name Mapping Authorities (NMAs).**

ARKs created and their corresponding access Uniform Resource Locators (URLs), the BnF developed an ARK “routing” application, which simply associates a given ARK, according to its prefix, with a Name Mapping Authority (NMA), or “resolver”<sup>7</sup>. The prefix corresponding to each sub-NAA is a short extension to the BnF NAAN, sometimes called a “shoulder”<sup>8</sup>, for example,

- [ark:/12148/cb...](http://ark:/12148/cb...)
- [ark:/12148/b...](http://ark:/12148/b...)

In this way, each BnF subnaming authority was independent, and the access URL could be easily deduced from the syntax of each ARK. Upon access, redirection occurred only if an NMA could not resolve an ARK; in this case, it used the routing application to automatically forward the user to the correct NMA. This technical choice, which was intended to simplify the resolution process, worked well with these first two subnaming authorities and their NMAs. Since then, new subnaming authorities, each of which assigns ARKs independently and without centralized control, have continued to emerge. This includes the two currently most important sources of assignment, the preservation repository (SPAR) and the BnF Manuscripts and Archives catalog<sup>9</sup>.

The BnF now manages more than 50 different websites; out of which 16 act as ARK NMA resolvers. Figure 2 shows a simplified view of the publicly available resolvers.

It is worth noting that several NMAs now routinely resolve ARKs that their respective sub-NAA did not assign. The routing mechanism is quite neutral, so everything works fine despite significant differences in the access conditions, functionalities, and nature of the identified entities.

Thus, [ark:/12148/cb32800864g](http://ark:/12148/cb32800864g) identifies a periodical record that can be displayed differently depending on whether one accesses it via the BnF catalog<sup>10</sup>, the linked open data portal [data.bnf.fr](http://data.bnf.fr)<sup>11</sup> or the new site “Presse locale ancienne”<sup>12</sup>, which for each newspaper title identifies physical and digitized copies in different French or foreign libraries. Effectively, one uniquely identified archival object supports a variety of context-dependent access experiences.

One main incentive to assign ARKs came from BnF’s linked open data policy to assign ARKs to datasets. Data providers are

<sup>7</sup>An NMA is a web server or local resolver that can take an incoming identifier and deliver (“map” it to) appropriate content.

<sup>8</sup><https://ezid.cdlib.org/learn/id-concepts>

<sup>9</sup><http://archivesetmanuscrits.bnf.fr>

<sup>10</sup><http://catalogue.bnf.fr/ark:/12148/cb32800864g>

<sup>11</sup><http://data.bnf.fr/ark:/12148/cb32800864g>

<sup>12</sup><http://presselocaleancienne.bnf.fr/ark:/12148/cb32800864g>

offered a trade: if they want to benefit from the data.bnf.fr search engine optimization features, they have to put some thought into a persistence policy (elaborated below).

## 2 ARK IMPLEMENTATION CHALLENGES

### 2.1 Users Working with ARKs in the Wild

When in 2014 the BnF decided to open up access to its metadata using the “Licence Ouverte (Open License)” [10], ARK identifiers seemed the natural choice. As a consequence, ARKs are being used to reference a wide diversity of recently exposed data. For example, the Work<sup>13</sup>, *Le Mystère de la chambre jaune* has the actionable ARK <http://data.bnf.fr/ark:/12148/cb119626305>; the use of fragments (such as <http://data.bnf.fr/ark:/12148/cb119626305#about>) allows to differentiate the record from the actual work in order to attach different statements on those distinct URIs – *Le Mystère de la chambre jaune* was created in 1907, but the BnF record was created in 1984. A similar distinction plays out for many other kinds of entities described by records, including persons, subjects, places, etc. All these entities have BnF ARKs. Yet another application of ARKs, independent of semantic-web-related questions, came from an open license requirement that we explicitly mention the source whenever data was to be reused outside the BnF.

Due to public exposure of these identifiers, external users are able to link to them in a persistent and secure way. As an example, Wikidata<sup>14</sup> now recognizes “BnF ID”<sup>15</sup> and “Gallica ID”<sup>16</sup> properties for registering authorities and digital material identifiers that are linked to their ARKs. Continuing the previous example, the Q2563691 entity in wikidata is linked to [ark:/12148/cb119626305](http://ark:/12148/cb119626305), supporting reference and information extraction. This, combined with clustering and alignments done by “Virtual International Authority File (VIAF)”<sup>17</sup>, has led to a significant boost in usage. A simple query shows that, in April 2018, among the 4 226 232 human beings described in Wikidata, 334 069 were linked to their descriptions in BnF (about 7%), and if we limit this to the 158 321 French people, we have a 43% alignment rate (68 930). More interestingly, these ARKs are also used to source the information.

The wide visibility of these identifiers exerts pressure on institutions to fulfill their commitments to them. Depending on the use case, the object, and the institution, there is a variety of legitimate ways to support persistence, immutability, and citability via stable URLs. To avoid surprising users, persistent identifiers need to have clearly defined scope and persistence promises. For example, in some portals that provide access to digitized civil registries, ARK identifiers are generated on demand to point to user-defined fragments in digitized pages; UUIDs are generated as a basis for the ARKs and a user is usually unaware that the archival institution has no knowledge of or commitment to them. While this situation is not desirable and ARKs are not the only kind of identifier to suffer from this kind of practice, ARKs do provide an “inflection” feature by which an end-user should be able to retrieve explicit persistence policies. ARKs need not necessarily provide access to

objects forever, or even for the next year, but they need to be “talking identifiers” that tell their end users what should be expected about their persistence.

### 2.2 A BnF Concern: Reconciling Access Persistence with Resource Mutability

As a leading institution among French libraries and a pioneering user of ARK identifiers, the BnF strove to set a good example by building implementations that conformed to the spirit as well as to the letter of the ARK specification. The proliferation of BnF applications that manipulated ARKs, whether creating them or adapting them to new services over existing resources, was starting to pose problems. For example, the absence of a centralized knowledge base of all identifiers, their statuses, and resolution methods, was forcing each NMA that consumed identifiers to query the NMA APIs of the subnaming authorities in order to resolve the ARKs that they were simply reusing.

In addition, many cases of mutability complicated the situation. Unlike published documents, for example, descriptive records produced by the BnF are subject to many changes in their lifecycles. Even though metadata producers are strongly discouraged from reusing a record for a document other than the one it originally described, record deletion remains a residual practice, with no tombstone mechanism (surrogate object indicating basic metadata, the reason and the date when the resource became unavailable). Whenever possible, metadata producers are encouraged to delete with redirection, allowing the user to be automatically redirected to the alternate resource. Another unaddressed problem is when a record needs to be split, namely in the following cases:

- when cataloging is done with a progressive degree of precision (a resource describing a set of documents is produced, then replaced by a series of resources describing individual documents from the set);
- when an identifier assigned to a subject term later becomes obsolete (e.g., “Women in Mythology” becomes split into “Women” and “Mythology”).

In the case of digital documents, the most common case of mutability at the BnF happens when resources are withdrawn (“unpublished”) by a collection manager. The dissemination of the document becomes restricted to the institution’s premises, sometimes for intellectual property reasons, or because the original document was digitized again and the older copy is considered obsolete. These situations, coupled with the reuse of ARKs by NMAs that do not control the identifiers created in another database, make a strong case for a centralized BnF knowledge base, and the publication of honest persistence statements based on the institution’s real ability to keep the resources accessible over the long term.

For the BnF preservation system, Scalable Preservation and Archiving Repository (SPAR), the mutability of catalog resources is a particularly crucial issue. The reference information in SPAR is based on a small set of descriptive metadata, mainly identifiers (ISBN, ISSN, ISRC, etc.). Among these, the ARK identifier constitutes a bridge to the complete object metadata in the catalogs. Based on cold storage and a robotic tape library, SPAR is not updated when the intellectual content description of its information packages evolves in the catalogs. Thus the mutability of descriptive

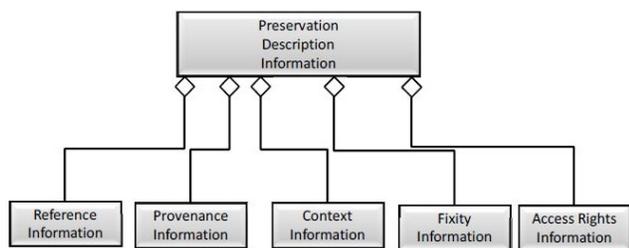
<sup>13</sup>“Work” as defined in the Functional Requirements for Bibliographic Records (FRBR) model

<sup>14</sup><https://www.wikidata.org>

<sup>15</sup><https://www.wikidata.org/wiki/Property:P268>

<sup>16</sup><https://www.wikidata.org/wiki/Property:P4258>

<sup>17</sup><https://viaf.org/>



**Figure 3: Preservation Description Information (source: Figure 4-16 of the OAIS standard)**

records leads to desynchronization that only the identifier can mitigate. Operations on descriptive records that alter the association between the identifier and its resource (deletion or relocation of digital copies) would otherwise compromise the ability of the digital repository to access representation and Preservation Description Information (PDI) [1] (Figure 3).

### 2.3 A Community Concern: Maintaining Control over Identifier Management

In many discussions we initiated with internal users at the BnF and the French-speaking ARK community generally, it became clear that basic understanding of persistent identifiers is limited. For example, people often mistakenly believe that ARKs (and other persistent identifiers) automatically, by technical magic, guarantee reliable access forever. The consequence is that archival organizations request identifier assignment from providers of access services, but such services might only commit to persistence for the duration of their organizational contracts. In addition, mistrust of international identifier management services may also encourage creation of homemade solutions. For example, OCLC's failure in 2015 to maintain PURL.org as a library community effort led some service providers to abandon PURLs in favor of URLs that they can control themselves, even though the Internet Archive now maintains PURL.org<sup>18</sup>.

In discussions with stakeholders in the archival sector, it became clear that archival institutions should have control over their identifiers but often cannot or do not want to assert it. Most identifier service design concerned minting and resolving rather than long-term stewardship, for which the following questions should be addressed by NAAs and NMAs:

- At what level of granularity should I mint an ARK? At the level of an archival fonds, at the level of a document, or at the level of a single page?
- What happens if my resource changes over time? For instance, if I discover that a page is missing in a digitized document that has been published, and I want to insert a page, if the URL for the page contains an order number, an insertion results in shifting URLs for all the pages after the inserted page, which can cause citation problems.

Other questions to address concern overall interoperability at a particular time and through time:

- At a given time, how will a vendor achieve persistent access for its ARK-based URLs if it does not also manage the identified assets themselves?
- Through time, how will the ARKs outlive the software and organizations that minted and/or resolved them? Other systems may implement ARKs with different algorithm and policy choices. Moreover, one organization may need to import ARKs previously supported by a predecessor and export them in turn to a successor.

In a nutshell, we found a strong need for organizations to take on the task of identifier policy definition: what should be identified, how might objects be expected to change, and what service can be promised at a given time and through time? Institutions should play a bigger role in controlling their identifiers, and should realize that this task is a core part of their digital asset management activities. Unfortunately, dissemination and discovery tools are currently concerned primarily with object access, in other words, making ARKs actionable and building services on top of them. The vendors and heads of archival and library networks are likely the best facilitators of the necessary changes, but the first step is to raise awareness so that individual organizations begin to demand guidance on best practices. These findings were all borne out at the ARK Summit.

## 3 GOING TO THE NEXT STAGE

The BnF is a seasoned practitioner, a center of expertise, and a promoter of best practices and services for the francophone community. Its role in helping French-speaking institutions overcome technical and language barriers, still an important obstacle for many French institutions, is crucial. Efforts to organize and strengthen ARK implementations are split in three complementary directions:

- producing naming recommendations with respect to sub-naming authorities and NMAs [5]<sup>19</sup>, as well as supporting organizations, within or outside the BnF that need to apply them;
- raising awareness of good practices among data producers, as persistence is not a mere technical question but requires that agents who manipulate and produce data be conscious of the impact of their actions on long-term citability;
- stressing the commitment, costs, skills, and infrastructure required for identifier management, while promoting and expanding existing services for assignment, description, and resolution of ARK identifiers.

On March 21st 2018, these goals led the BnF to organize an 'ARK Summit', a French-speaking gathering which aimed to present the principles ARK is based on, share best practices, take a glimpse of the impact and adoption of ARK in France and get people from the same professional sector to talk about persistent identification.

### 3.1 The ARK Summit, Effect 1: from Individual Users to a Community

The ARK Summit was instrumental in raising community awareness among a wide audience that included heritage institutions and vendors from various sectors serving the French-speaking community. It was crucial to clarify that identifiers can be easy to

<sup>18</sup><http://blog.archive.org/2016/09/27/persistent-url-service-purl-org-now-run-by-the-internet-archive/>

<sup>19</sup>There are similar efforts in other domains such as the one initiated by the Committee on Earth Observation Satellites [3]

understand, and that technical HTTP / URL considerations are just a means to ensure persistent access to institutional objects. Service design and identifier management are squarely the institution's responsibility. This is the first step before defining the technical means of fulfilling that responsibility, whether by internal efforts or by delegating to an external operator (private vendor or public operator).

The "free kitten" metaphor, borrowed from the open source community, was highly retweeted at the ARK Summit. Minting ARKs is as easy as adopting an abandoned kitten, but once you accept responsibility, the young cat becomes a serious commitment, involving mental load, time, and money. This metaphor really helped surface the often hidden responsibilities of practical persistent identifier management.

To follow up the rich discussions and channel the energy engaged during the day, an ARK French-speaking email discussion list<sup>20</sup> was put in place by the BnF and its partners to complement the English-speaking list<sup>21</sup>.

### 3.2 The ARK Summit, Effect 2: the Community, our Best Allies

The summit also had an unforeseen benefit: it attracted internal attention to BnF's own ARK implementation. As mentioned, the BnF manages 12 subnaming authorities and almost 16 NMAs under the `bnf.fr` domain name, all of which have independent implementation policies and autonomous road maps. Matters of persistence are often forgotten in the incremental evolution of each road map, and organizing the summit with more than 270 participants drew internal attention to ARKs and persistent access. Not surprisingly, in the weeks just before and after the event, the ARK team received welcome feedback on implementation issues and broken links for deleted, replaced or unpublished resources (both issues described in section 2.2). This momentum will be leveraged to improve the BnF implementation and make it more consistent across its NMAs.

### 3.3 Beyond the Summit: Towards Shared Governance of ARKs

The ARK Summit also happened at a time when the CDL was considering governance changes around ARKs. After having served as the incubator for global ARK infrastructure since 2001, the CDL considered it had gathered enough experience to initiate a joint project with Duraspace called "ARKs in the open" [4]. The primary task of this project is to begin transitioning maintenance of that infrastructure to a partnership with multiple global organizations and community participants.

Notably, this means assembling an advisory group and working groups to help guide its future, developing shared maintenance procedures for the ARK NAAN registry, consulting with the community on requirements for a shared infrastructure and management, and shepherding the ARK specification through the IETF RFC process to make it an Internet Standard. At the same time, an experts group met on March 22nd at the BnF, the day after the summit, and started considering some of those questions [13]. A joint BnF-CDL online survey is being prepared to understand the

community better: how do providers use ARKs? For which objects? What features do they implement? What challenges did they face or are they currently facing? We believe these activities can usefully align with the "ARKs in the Open" project.

## CONCLUSION

Persistent identification is like digital preservation: it is a path, not a destination. In both pursuits, it is easier to know when one has failed than when one has succeeded. In both pursuits, the experience of persistence through time will evolve along with technologies and user expectations, and it will be improved by continuous discussion among the concerned community of users and providers. ARKs have been around since 2001 and adopted by over 550 organizations. To follow the "free kitten" metaphor, with well over a hundred million ARK kittens born since 2001, the community at large is now ever more mindful of the care and feeding required for a robust experience of long-term access. The ARK Summit at BnF, the interest it raised in community building, and the lively discussions that took place demonstrated the demand for this open and flexible approach to stewarding heritage resources on the Web.

## ACKNOWLEDGMENTS

The authors would like to thank John Kunze from CDL, without whom this paper would not have been the same, all the members of the SPAR team for their commitment, and Claire Sibille from the SIAF for the fruitful discussions about ARKs.

## REFERENCES

- [1] 2012. *Reference Model for an Open Archival Information System (OAIS)*. Recommended practice. CCSDS. 135 pages. <https://public.ccsds.org/Pubs/650x0m2.pdf>
- [2] 2015. *Référentiel Général d'Interopérabilité (RGI)* [General Reference Framework for Interoperability]. (December 2015).
- [3] 2017. *Persistent Identifiers Best Practices*. Technical Report CEOS/WGISS/DSIG/PIDBP. CEOS-WGISS. [http://ceos.org/document\\_management/Working\\_Groups/WGISS/Documents/WGISS%20Best%20Practices/CEOS%20Persistent%20Identifier%20Best%20Practices\\_v1.2.pdf](http://ceos.org/document_management/Working_Groups/WGISS/Documents/WGISS%20Best%20Practices/CEOS%20Persistent%20Identifier%20Best%20Practices_v1.2.pdf)
- [4] 2018. ARKs in the Open project, an initiative from the California Digital Library and Duraspace. (March 2018). <https://wiki.duraspace.org/display/DSP/ARKs+in+the+Open+Project>
- [5] 2018. *Préconisations pour l'implémentation d'ARK par les sous-autorités nommées et autorités d'adressage BnF*. (March 2018).
- [6] Ministère de la Culture et de la Communication. 2015. *Vade-mecum pour les producteurs de données : identifiants pérennes pour les ressources culturelles*. Technical Report.
- [7] John Kunze. 2003. Towards electronic persistence using ARK identifiers. In *Proceedings of the 3rd ECDL Workshop on Web Archives*. <https://confluence.ucop.edu/download/attachments/16744455/arkcdl.pdf>
- [8] John Kunze, Scout Calvert, Jeremy D DeBarry, Matthew Hanlon, Greg Janée, and Sandra Sweat. 2017. Persistence statements: describing digital stickiness. *Data Science Journal* 16 (2017). <http://doi.org/10.5334/dsj-2017-039>
- [9] John A. Kunze and R. P. C. Rodgers. 2013. *The ARK Identifier Scheme*. Internet-Draft draft-kunze-ark-18. Internet Engineering Task Force. Retrieved 2018-04-13 from <https://tools.ietf.org/html/draft-kunze-ark-18> Work in Progress.
- [10] French Prime Minister/Etalab. 2017. *Licence ouverte* [Open licence]. (April 2017). <https://www.etalab.gouv.fr/wp-content/uploads/2017/04/ETALAB-Licence-Ouverte-v2.0.pdf> version 2.0.
- [11] Sébastien Peyrard, Jean-Philippe Tramon, and John Kunze. 2014. The ARK Identifier Scheme: Lessons Learnt at the BnF and Questions Yet Unanswered. In *Proceedings of the 2014 International Conference on Dublin Core and Metadata Applications (DCMI'14)*. Dublin Core Metadata Initiative, 83-94. <http://dl.acm.org/citation.cfm?id=2771234.2771244>
- [12] California Digital Library (CDL) University of California. 2001. ARK registry. (08 March 2001). Retrieved 2018-04-13 from [https://www.cdlib.org/uc3/naan\\_registry.txt](https://www.cdlib.org/uc3/naan_registry.txt)
- [13] Collective work. 2018. ARK experts day @ National Library of France (BnF). (22 March 2018). Retrieved 2018-06-29 from [https://www.cdlib.org/services/infrastructure/docs/ARK\\_experts\\_day.pdf](https://www.cdlib.org/services/infrastructure/docs/ARK_experts_day.pdf)

<sup>20</sup>arks-forum-fr@framalistes.org

<sup>21</sup>arks-forum@googlegroups.com