



Bibliothèque nationale de France

Préconisations pour l'implémentation d'ARK par les sous-autorités nommantes et autorités d'adressage BnF

**Bibliothèque nationale
de France**

direction des Services et des réseaux
département des Métadonnées
service Ingénierie des métadonnées

version 1 du 12 mars 2018
émetteur : Sébastien PEYRARD
affaire suivie par : Bertrand CARON
référence : BnF-ADM-2018-022324-01



Table des mises à jour du document

Version	Auteur	Date	Objet de la mise à jour
v. 1	Bertrand Caron	12/03/2018	Version initiale validée.



TABLE DES MATIERES

1. INTRODUCTION	4
1.1. OBJET	4
1.2. DOMAINE D'APPLICATION.....	4
1.3. GESTION ET EVOLUTION.....	4
2. DOCUMENTS APPLICABLES ET DE REFERENCE	4
3. TERMINOLOGIE	5
3.1. GLOSSAIRE	5
3.2. ABREVIATIONS.....	6
4. UTILISATION DU NAAN BNF	7
5. CONSTRUCTION DU NOM ARK	7
5.1. DEFINITION DU PREFIXE	7
5.2. CARACTERES SPECIFIQUES A LA RESSOURCE.....	7
5.3. CARACTERE DE CONTROLE	7
6. EVOLUTION DES RESSOURCES	8
6.1. SUPPRESSION.....	8
6.2. DEPUBLICATION OU CHANGEMENT DE PERIMETRE DE DIFFUSION	8
6.3. MODIFICATION SUBSTANTIELLE DE L'IDENTITE.....	8
6.4. SCISSION	9
6.5. REMPLACEMENT D'UNE RESSOURCE	9
7. DEFINITION DES QUALIFICATIFS	9
7.1. QUALIFICATIFS DE GRANULARITE	9
7.2. QUALIFICATIFS DE SERVICE	9
7.3. QUALIFICATIFS COMMUNS A TOUTES LES AUTORITES D'ADRESSAGE BnF	10
7.3.1. <i>Qualificatif de service « description »</i>	10
7.3.2. <i>Qualificatif de service « policy »</i>	10
8. RESOLUTION DES ARK	11
8.1. IDENTIFICATION DU NAAN.....	12
8.2. IDENTIFICATION DU PREFIXE	12
8.3. CALCUL DU CARACTERE DE CONTROLE.....	12
8.4. RESOLUTION DES QUALIFICATIFS.....	13
8.4.1. <i>Résolution des qualificatifs de granularité</i>	13
8.4.2. <i>Résolution des qualificatifs de service</i>	13
8.4.3. <i>Redirection vers une URL d'accès</i>	13
9. ANNEXE INFORMATIVE : LISTE DES CODES HTTP RECOMMANDES PAR CAS D'USAGE	15



1. Introduction

1.1. **Objet**

Ce document détaille les préconisations d'implémentation de l'identifiant ARK au sein de la BnF, et plus particulièrement des principes d'attribution par les sous-autorités nommantes et de résolution par les autorités d'adressage. Il réunit des recommandations mentionnées par la spécification ARK et d'autres spécifiques à la BnF et considérées comme des bonnes pratiques.

1.2. **Domaine d'application**

Ce document s'applique à toutes les sous-autorités nommantes et autorités d'adressage BnF, c'est-à-dire à toutes les applications publiant ou manipulant des ARK dont le NAAN est celui de la BnF (12148). Si plusieurs d'entre elles ne suivent pas encore l'ensemble de ces préconisations, ces dernières sont néanmoins destinées à être implémentées de manière homogène, à moins d'une spécificité justifiant l'exception.

Ce document est également destiné

- aux autorités nommantes extérieures à la BnF qui désirent connaître la politique générale de la BnF concernant les ARK afin de s'en inspirer pour leur propre implémentation (toutes sections)
- aux autorités d'adressage extérieures à la BnF qui désirent réutiliser des ARK BnF pour leur propre usage (section 8 ci-dessous « Résolution des ARK »).

1.3. **Gestion et évolution**

Ce document définit un cadre général à l'attribution et à la résolution des ARK à la BnF. Certaines considérations d'implémentations décrites dans les sections 7.3, 8 et 9 peuvent être revues dans le cadre de futures instructions menées par la BnF.

2. Documents applicables et de référence

Intitulé	Description	Référence
<i>Spécification ARK</i>	<i>Spécification ARK soumise le 5 avril 2013 pour normalisation à l'IETF</i>	https://tools.ietf.org/pdf/draft-kunze-ark-18.pdf
<i>Vade-mecum sur les identifiants pérennes</i>	<i>Identifiants pérennes pour les ressources culturelles : vade-mecum pour les producteurs de données, version 1.0, Ministère de la Culture et de la communication : décembre 2015</i>	www.bnf.fr/documents/identifiants_perennes_vademecum.pdf
<i>Registre ARK</i>	<i>Liste des autorités nommantes</i>	http://www.cdlib.org/uc3/naan_registry.txt http://ark.bnf.fr/services/naan_registry.txt



3. Terminologie

3.1. Glossaire

Terme	Définition
Autorité d'adressage	Organisation capable de rendre un ARK actionnable en fournissant un accès à la ressource. Selon la spécification ARK, une autorité d'adressage se définit par sa capacité à fournir, pour un ARK donné, la ressource demandée ou un substitut approprié, des métadonnées et une politique de permanence appliquée à la ressource. Voir https://tools.ietf.org/pdf/draft-kunze-ark-18.pdf#page=26
Autorité nominante	Organisation créant des associations pérennes entre des identifiants et des ressources informationnelles. L'autorité nominante peut déléguer cette création à une autre organisation.
Inflexion	Méthode d'appel d'un ARK en vue d'obtenir des métadonnées sur la ressource ou la politique de permanence appliquée à la ressource. Selon la spécification ARK, la méthode d'appel consiste à ajouter respectivement « ? » (point d'interrogation) ou « ?? » à la fin de l'ARK. Voir https://tools.ietf.org/pdf/draft-kunze-ark-18.pdf#page=31
Nom ARK	Chaîne de caractères suivant immédiatement le numéro d'autorité nominante et éventuellement suivie de qualificatifs.
Numéro d'autorité nominante	Numéro sur cinq chiffres assigné par l'agence de maintenance de l'identifiant ARK (la CDL) à une organisation.
Politique de permanence	Politique appliquée par une autorité d'adressage aux ressources auxquelles elle donne accès. Elle résulte de pratiques du ressort de l'autorité nominante (le comportement en cas d'évolution de la ressource, la méthode de nommage) et du ressort de l'autorité d'adressage (les variantes proposées de la ressource sous forme de qualificatifs de service). Voir https://tools.ietf.org/pdf/draft-kunze-ark-18.pdf#page=29
Préfixe	Élément initial du nom ARK. Composé d'un nombre défini de caractères, il est attribué par l'autorité nominante à une sous-autorité nominante.
Qualificatif de granularité	Qualificatif précédé du caractère « / » (barre oblique) indiquant que la ressource identifiée par le qualificatif est contenue dans celle identifiée par l'ARK non qualifié.
Qualificatif de service	Qualificatif précédé du caractère « . » (point) indiquant que la ressource identifiée par le qualificatif est une variante de celle identifiée par l'ARK non qualifié.
Registre ARK	Fichier géré par la CDL listant les organisations déclarées comme autorités nominantes et leur attribuant un NAAN. Le registre ARK est répliqué dans deux institutions partenaires de la CDL : la BnF et la <i>National Library of Medicine</i> .
Ressource	On désignera par le terme de « ressource » l'objet informationnel identifié par un ARK.



Terme	Définition
Schème d'identifiant	Partie liminaire de l'identifiant définissant son type. Pour un identifiant ARK, il s'agit de l'élément « ark: ».
Sous-autorité nommante	Organisation ou système à laquelle l'autorité nommante a délégué la responsabilité de créer des noms ARK. La sous-autorité nommante est identifiée par un préfixe.

3.2. Abréviations

Abréviation	Terme développé
ARK	<i>Archival Resource Key</i>
CDL	<i>California Digital Library</i>
NAA	<i>Name Assigning Authority</i> (autorité nommante)
NAAN	<i>Name Assigning Authority Number</i> (numéro d'autorité nommante)
NMA	<i>Name Mapping Authority</i> (autorité d'adressage)
NMAH	<i>Name Mapping Authority Hostport</i> (serveur d'accès d'une autorité d'adressage)



4. Utilisation du NAAN BnF

Les ARK attribués par les sous-autorités nommantes BnF utilisent le NAAN 12148 (BnF). **La responsabilité du maintien des ARK produits dans ce NAAN incombe donc à la BnF. À terme, toute sous-autorité nommante BnF doit appliquer les recommandations ci-dessous.**

Il est fondamental de noter que **la publication d'un identifiant ARK, qu'il soit utilisé comme URL ou simplement mentionné comme permalien, engage la BnF** à maintenir l'association entre l'ARK en question et la ressource concernée, dans son identité d'origine.

5. Construction du nom ARK

Le nom ARK doit être opaque et ne révéler aucune information signifiante de la ressource.

De ce fait, les règles suivantes sont appliquées :

- **Utilisation exclusive des 29 caractères « 0123456789bcdfghjkmnpqrstvwxyz » (lettres en minuscules). Les voyelles sont exclues** afin d'éviter de créer fortuitement une suite de caractères signifiants. **La lettre « l » l'est également** en raison de la confusion possible avec le caractère « I » (i majuscule).
- **Pas plus de deux lettres de suite**, afin d'éviter de créer fortuitement un acronyme.

5.1. Définition du préfixe

La BnF définit des préfixes correspondant à différentes sous-autorités nommantes à qui elle délègue l'attribution de noms ARK. Le préfixe correspond donc à un **contexte d'attribution** visant à éviter que deux sous-autorités nommantes attribuent le même ARK au sein du NAAN BnF, **et non à une information descriptive sur le type de la ressource.**

Note : les autorités nommantes sont incitées à prévoir dès l'origine un préfixe afin de gérer d'éventuels nouveaux contextes d'attribution.

5.2. Caractères spécifiques à la ressource

À la suite des caractères du préfixe, on trouvera une série de caractères suivant les règles mentionnées dans la section 5 ci-dessus. On veillera, pour le confort de l'utilisateur souhaitant citer une ressource par son ARK, à générer des noms ARK les plus courts possible¹.

5.3. Caractère de contrôle

Le dernier caractère d'un ARK BnF est un caractère de contrôle. Il a pour but d'éviter de renvoyer une erreur « 404 » lorsqu'un utilisateur fait une faute de frappe en appelant un ARK. En ce cas, l'application résolvant un ARK est en mesure de signaler à l'utilisateur que l'ARK est erroné et qu'il doit vérifier sa forme.

¹ Ainsi, on pourra générer 811 valeurs distinctes pour un nom ARK comprenant un préfixe sur deux caractères, deux caractères spécifiques à la ressource parmi les 29 caractères recommandés et un caractère de contrôle. Si le nom ARK comprend trois caractères spécifiques à la ressource, ce chiffre est porté à 24 389.

6. Evolution des ressources

Une implémentation d'identifiants ARK doit répondre à deux nécessités contradictoires : la conservation à long terme de l'association entre un identifiant et une ressource, et, bien souvent, une évolutivité plus ou moins importante des ressources identifiées. **Pour chaque implémentation ARK, il est donc nécessaire d'envisager les cas d'évolution suivants et, le cas échéant, de prévoir une réponse technique appropriée.**

6.1. Suppression

La suppression ne peut se contenter de supprimer l'enregistrement / la ressource dans la base : il est nécessaire qu'un ARK publié continue à pointer, si ce n'est sur la ressource elle-même, au moins vers ses métadonnées et une information sur l'événement qui a donné lieu à la fin de l'accès à la ressource telle que diffusée originellement. En outre, une ressource de substitution peut être le cas échéant signalée à l'utilisateur.

6.2. Dépédition ou changement de périmètre de diffusion

La dépédition ou le changement de périmètre de diffusion, qui aboutit à l'inaccessibilité d'une ressource dans les conditions de consultation d'origine², **ne peut se contenter de supprimer l'enregistrement / la ressource dans la base d'origine** : il est nécessaire qu'un ARK publié continue à pointer, si ce n'est sur la ressource elle-même, au moins vers ses métadonnées et une information sur l'événement qui a donné lieu à la fin de l'accès à la ressource telle que diffusée originellement. En outre, des modalités de consultation alternatives peuvent être le cas échéant signalées à l'utilisateur.

6.3. Modification substantielle de l'identité

Il est fondamental pour les utilisateurs de ressources nommées par des ARK de connaître le degré de variabilité de la ressource³. En particulier :

- si la ressource est strictement invariante ;
- ou si elle peut être soumise à corrections ou enrichissement ;
- ou enfin si elle s'accroît.

Cependant, en cas de modification substantielle de l'identité d'une ressource identifiée par un ARK on prévoira

- soit la création d'une nouvelle ressource et la redirection vers la nouvelle ;
- soit, à défaut, si une telle opération n'est pas envisageable, une information dans les métadonnées de la ressource modifiée signalant la teneur de la modification.

La nature de la modification devant faire l'objet d'une telle procédure sera à définir au cas par cas, selon des principes édictés par l'autorité nominante⁴.

Exemple : ark:/12148/cb41242894n

Dans certains cas bien précis, le changement de titre d'un périodique fait l'objet de la création d'une nouvelle ressource, avec mention en zone 785 [en Intermarc](#) de la ressource nouvelle. Cette pratique catalographique peut être considérée comme une réponse adaptée à une modification substantielle d'une ressource identifiée par un ARK.

² Ainsi, une ressource jusque-là proposée sans limitation d'accès devenant accessible sur authentification ou exclusivement dans l'enceinte de l'institution serait dans ce cas.

³ PEYRARD, Sébastien; KUNZE, John A.; TRAMONI, Jean-Philippe, « The ARK Identifier Scheme: Lessons Learnt at the BnF and Questions Yet Unanswered ». dans *International Conference on Dublin Core and Metadata Applications*, oct. 2014. ISSN 1939-1366, p. 91 (« Machine-readable commitments »). <<http://dcpapers.dublincore.org/pubs/article/view/3704/1927>>. Consulté le 19/02/2018.

⁴ La spécification ARK, p. 31 (<https://tools.ietf.org/pdf/draft-kunze-ark-18.pdf#page=31>), cite parmi les critères à étudier la similarité (la limite à partir de laquelle une chose diffère suffisamment d'une autre pour être considérée comme une ressource à part entière) et la granularité (la limite à partir de laquelle une composante ou une variante d'une chose est considérée comme une ressource à part entière).

6.4. Scission

Une scission (répartition du périmètre d'une ressource sur plusieurs nouvelles ressources) doit donner lieu à la redirection vers les différentes ressources qui correspondent au périmètre de la ressource originelle.

6.5. Remplacement d'une ressource

Le remplacement d'une ressource par une autre, qu'il soit opéré pour proposer une version plus récente d'une ressource ou pour fusionner deux ressources dont la similarité apparaît ultérieurement, **ne peut se contenter de supprimer l'enregistrement / la ressource obsolète dans la base** : il est nécessaire que l'ARK publié à l'origine redirige vers la ressource nouvelle.

Exemple (partiel) : la ressource originellement identifiée par [ark:/12148/cb33348652z](https://nbn-resolving.org/urn:nbn:fr:bnf-12148-cb33348652z) renvoie le même résultat que la ressource par laquelle elle a été remplacée ([ark:/12148/cb32757566w](https://nbn-resolving.org/urn:nbn:fr:bnf-12148-cb32757566w))⁵.

7. Définition des qualificatifs

La spécification ARK permet d'étendre le nom ARK proprement dit par des qualificatifs. Ces derniers identifient soit des composantes (qualificatif de granularité) ou des variantes (qualificatif de service) d'une ressource afin de les rendre citables.

La définition de qualificatifs pertinents peut être réalisée par l'autorité nommante et/ou par l'autorité d'adressage. À la BnF, les sous-autorités nommantes et les autorités d'adressage les définissent en concertation avec les coordinateurs ARK afin de garantir l'homogénéité des pratiques de nommage.

Deux qualificatifs de service sont obligatoires pour toute autorité d'adressage utilisant des ARK BnF ; le fonctionnement de ceux-ci est décrit dans la section 7.3 ci-dessous.

Les contraintes spécifiques à la BnF sur la forme des qualificatifs sont les suivantes :

- **ils doivent être composés de caractères alphanumériques** (la contrainte de l'opacité ne portant pas nécessairement sur les qualificatifs, les voyelles et le caractère « l » peuvent être employés) ;
- **les lettres doivent être en minuscules** ;
- **le premier caractère du qualificatif de service doit être une lettre.**

7.1. Qualificatifs de granularité

Un qualificatif de granularité suit immédiatement le nom ARK ; il est séparé de celui-ci par le caractère « / » (barre oblique). Un ARK ainsi qualifié identifie une ressource contenue dans une autre identifiée par l'ARK non qualifié. Ainsi, la ressource [ark:/12148/btv1b525049362/f108](https://nbn-resolving.org/urn:nbn:fr:bnf-12148-btv1b525049362-f108) identifie la vue n° 108 du document numérique identifié par l'ARK [ark:/12148/btv1b525049362](https://nbn-resolving.org/urn:nbn:fr:bnf-12148-btv1b525049362).

Formellement, publier un ARK comportant un qualificatif de granularité équivaut à publier deux ARK : l'ARK qualifié et l'ARK non qualifié. **Une autorité d'adressage doit donc fournir une ressource ou à défaut des métadonnées sur cette ressource en cas d'appel d'un ARK non qualifié si elle a publié un ARK qualifié correspondant.**

Il est à noter que **la politique de permanence appliquée à une ressource identifiée par un ARK qualifié peut différer de celle appliquée à la ressource identifiée par le même ARK non qualifié**. Ainsi, la ressource [ark:/12148/cc12415m/ca3158](https://nbn-resolving.org/urn:nbn:fr:bnf-12148-cc12415m-ca3158) a une permanence limitée par rapport à [ark:/12148/cc12415m](https://nbn-resolving.org/urn:nbn:fr:bnf-12148-cc12415m) : si l'accès à la première est compromis en raison d'une suppression ou d'un déplacement, l'utilisateur a la garantie qu'il pourra accéder à la seconde. On notera également que les exigences d'opacité pesant sur le nom ARK peuvent être allégées.

7.2. Qualificatifs de service

Un qualificatif de service suit immédiatement le nom ARK ou, en cas d'utilisation d'un qualificatif de granularité, le dernier d'entre eux ; il en est séparé par le caractère « . » (point). Un ARK ainsi qualifié identifie une variante de la ressource identifiée par l'ARK non qualifié. Ainsi, la ressource

⁵ Mais via une réponse « 200 » (« OK »), ce qui ne correspond pas aux préconisations.

<ark:/12148/btv1b525049362/f108.pdf> identifie une variante de la vue 108 du document numérique, identifiée par <ark:/12148/btv1b525049362/f108>, sous la forme d'un fichier PDF.

Il est à noter que la politique de permanence appliquée à une ressource identifiée par un ARK qualifié peut différer de celle appliquée à la ressource identifiée par le même ARK non qualifié. Ainsi, la ressource <ark:/12148/bpt6k134019r.epub> a une permanence limitée par rapport à <ark:/12148/bpt6k134019r> : si l'accès à la première est compromis en raison d'une évolution technique, l'utilisateur a la garantie qu'il pourra accéder à la seconde.

7.3. Qualificatifs communs à toutes les autorités d'adressage BnF

Les qualificatifs suivants sont obligatoires pour toute autorité d'adressage résolvant des ARK BnF.

7.3.1. Qualificatif de service « description »

Le qualificatif « description » permet aux autorités d'adressage BnF de répondre à l'une des exigences de la spécification ARK : fournir des métadonnées minimales associées aux ressources identifiées par un ARK. Ce qualificatif remplace la solution technique proposée par la spécification ARK (à savoir l'ajout du caractère « ? » à la suite de l'ARK), jugée problématique par la BnF pour des raisons techniques.

Exemple hors BnF : si l'on ajoute le caractère « ? » à l'ARK <ark:/47881/m6dz06h9> identifiant un mémoire, on obtient des métadonnées minimales décrivant la ressource (<ark:/47881/m6dz06h9?>).

Un des usages essentiels de ce qualificatif est, **en cas de remplacement, suppression ou dépublication de la ressource, de conserver une information sur la ressource**. On fournira donc les éléments suivants⁶ :

- Métadonnées descriptives minimales (identifiant ARK, titre, auteur, date, correspondant aux éléments « where », « what », « who », « when » de la spécification Dublin Core Kernel⁷ ainsi que le type de la ressource – notice descriptive, document numérique, ressource pédagogique, etc.).
- Type de l'événement ayant entraîné l'inaccessibilité de la ressource (suppression ou dépublication).
- Date de l'événement.
- Cause de l'événement.
- Optionnellement, agent ayant effectué l'opération.
- Optionnellement, modalités alternatives de communication de la ressource.

Note : une instruction ultérieure définira la forme recommandée de l'information fournie à l'utilisateur en cas d'utilisation du qualificatif « description ». Entre temps, cette forme est laissée au choix de l'autorité d'adressage.

7.3.2. Qualificatif de service « policy »

Le qualificatif « policy » permet aux autorités d'adressage BnF de répondre à l'une des exigences de la spécification ARK : fournir une politique de permanence associée aux ressources identifiées par un ARK⁸. Ce qualificatif remplace la solution technique proposée par la spécification ARK (à savoir l'ajout des caractères « ?? » à la suite de l'ARK), jugée problématique par la BnF pour des raisons techniques.

Exemple : <ark:/12148/bpt6k1320017q.policy> renvoie la politique générale de la BnF vis-à-vis des identifiants ARK qu'elle attribue.

On propose le canevas suivant pour la définition d'une politique de permanence spécifique à une autorité d'adressage :

« Considérations générales sur les ARK BnF

⁶ Voir « Identifiants pérennes pour les ressources culturelles : vade-mecum pour les producteurs de données, version 1.0 », p. 10, sur www.bnf.fr/documents/identifiants_perennes_vademecum.pdf#page=10.

⁷ Voir <http://dublincore.org/groups/kernel/spec/>.

⁸ Voir « Identifiants pérennes pour les ressources culturelles : vade-mecum pour les producteurs de données, version 1.0 », p. 11, sur www.bnf.fr/documents/identifiants_perennes_vademecum.pdf#page=11.

La BnF attribue des identifiants dans le domaine ARK 12148 (Bibliothèque nationale de France) selon les principes suivants.

- Aucun identifiant ARK ne sera attribué à nouveau ; c'est-à-dire qu'une fois qu'un lien entre un identifiant ARK et un objet aura été publié, ce lien devra être considéré comme unique, et ce pour une durée indéfinie.
- Les identifiants ARK attribués par la BnF ne contiennent, dans la mesure du possible, aucune information sémantique facilement reconnaissable ; cela contribue à faciliter leur utilisation indépendamment d'un contexte de temps ou de lieu.
- Les identifiants ARK attribués par la BnF contiennent un caractère de contrôle qui les garantit contre des erreurs de caractère isolées et des erreurs de transposition. Un utilisateur ayant fait une faute de frappe en tapant un ARK obtiendra une réponse HTTP 400 et un message l'informant que l'ARK fourni est erroné.

Mutabilité des ressources

La mutabilité des ressources [définition des ressources concernées par la politique de permanence] est définie comme suit.

- [description de l'éventuelle évolution de l'identité des ressources et mécanismes permettant à l'utilisateur de l'identifier]
- [éventuels cas de scission, remplacement, suppression ou dépublication et mécanismes permettant à l'utilisateur de les identifier]

Autorité d'adressage

L'autorité d'adressage [nom de l'application BnF concernée] résout les ARK des ressources décrites ci-dessus.

L'autorité d'adressage [serveur d'accès de l'autorité d'adressage] gère les qualificatifs de service génériques suivants :

- « description » : métadonnées minimales sur la ressource,
 - o [exemple]
- « policy » : politique de permanence de la ressource.
 - o [exemple]

En outre, elle gère les qualificatifs de granularité et de service suivants :

- « [nom du qualificatif] » : [description du niveau de granularité ou de service],
 - o [exemple]

Disponibilité

[Décrire les conditions de disponibilité et de continuité de service de la ressource]

[Mentionner une information de contact en cas de constat par l'utilisateur de l'inaccessibilité de la ressource.]

8. Résolution des ARK

Afin de résoudre les ARK de son périmètre, toute autorité d'adressage BnF résoudra les ARK de ressources dont elle a la charge **en définissant un nom de domaine permettant de résoudre l'ensemble des ARK de son périmètre.**

L'objectif des recommandations ci-dessous est d'**éviter, autant que possible, que l'utilisateur fournissant à une autorité d'adressage BnF un ARK, quelles que soient ses caractéristiques, obtienne une erreur 404.** De ce fait, ce document définit une série de cas d'erreur et exploite la richesse des codes HTTP (voir notamment l'annexe « Annexe informative : liste des codes HTTP recommandés par cas d'usage » ci-dessous) afin de fournir au consommateur des ARK BnF le maximum d'informations pour lui permettre d'adapter son comportement et ainsi de synchroniser efficacement ses données avec celles de la BnF.



Le traitement du nom ARK est décrit dans la spécification ARK à la section 2.7 « Normalization and lexical equivalence »⁹. On n'en donne ici que les étapes essentielles, pertinentes pour le processus de résolution d'un ARK BnF.

Il est fortement recommandé que les autorités d'adressage intègrent, lors de l'évolution de leur application, des tests de non-régression afin de vérifier que les différents mécanismes de résolution détaillés dans les sections ci-dessus continuent de fonctionner normalement.

8.1. Identification du NAAN

Dans le cas de l'appel à l'autorité d'adressage avec un identifiant commençant par « ark:/ », l'application devra identifier le numéro d'autorité nommante qui suit immédiatement la chaîne citée.

Cas 1 : le NAAN est « 12148 » => passage à l'étape « identification du préfixe ».

Cas 2 : le NAAN n'est pas « 12148 » => **redirection avec code HTTP « 302 » (redirection temporaire) de l'ARK concerné vers le résolveur ark.bnf.fr.**

8.2. Identification du préfixe

Dans le cas de l'appel à l'autorité d'adressage avec un identifiant commençant par « ark:/12148/ », l'application devra identifier le préfixe qui suit immédiatement la chaîne citée.

Cas 1 : le préfixe identifie un ARK géré par l'application => passage à l'étape « calcul du caractère de contrôle ».

Cas 2 : le préfixe n'identifie pas un ARK géré par l'application => **redirection avec code HTTP « 302 » (redirection temporaire) de l'ARK concerné vers le résolveur ark.bnf.fr.**

Exemple de cas 2 : <http://gallica.bnf.fr/ark:/12148/cb329111107> renvoie vers <http://ark.bnf.fr/ark:/12148/cb329111107>, lui-même renvoyant vers <http://catalogue.bnf.fr/ark:/12148/cb329111107>.

8.3. Calcul du caractère de contrôle

Le calcul du caractère de contrôle relève de la responsabilité de chaque autorité d'adressage pour les ARK qu'elle est en mesure de résoudre. **Il est fortement recommandé que chacune d'entre elles implémente le calcul du caractère de contrôle** tel que décrit ci-dessous **lorsqu'un ARK de son périmètre lui sera fourni.**

Le calcul du caractère de contrôle porte sur le nom ARK (ARK non qualifié).

Table de correspondance base10 / base29 :

xdigit:	0	1	2	3	4	5	6	7	8	9	b	c	d	f	g
value:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

xdigit:	h	j	k	m	n	p	q	r	s	t	v	w	x	z
value:	15	16	17	18	19	20	21	22	23	24	25	26	27	28

Algorithme :

1. Vérifier que la chaîne correspond au motif « [préfixe][0-9bcdghjkmnpqrstvwzx]* »
2. Pour chaque caractère, multiplier sa valeur en base 10 par sa position dans la chaîne, puis faire la somme.
3. Calculer le modulo base 29 de la valeur précédemment obtenue. Le caractère de contrôle correspond à ce modulo exprimé en base 29.

Cas 1 : le dernier caractère du nom ARK fourni en entrée à l'autorité d'adressage correspond au résultat de l'algorithme appliqué aux précédents caractères du nom ARK => passage à l'étape « traitement du nom ARK ».

⁹ Voir <https://tools.ietf.org/pdf/draft-kunze-ark-18.pdf#page=18>.

Cas 2 : le dernier caractère du nom ARK ne correspond pas au résultat de l'algorithme => **réponse 400** (« **Bad request** ») avec le **texte explicatif** « **ARK erroné : l'ARK que vous avez entré ([ARK fourni]) ne correspond pas à un ARK valide, merci de vérifier sa structure.** »¹⁰.

Exemples : <ark:/12148/cb34533084g>, <ark:/12148/bpt6k3411272d>.

8.4. Résolution des qualificatifs

Il est à noter que **l'ordre des qualificatifs de granularité est significatif** : l'ARK <ark:/12345/ab/cd/ef> identifie une ressource contenue dans la ressource <ark:/12345/ab/cd>, elle-même contenue dans la ressource <ark:/12345/ab>. En revanche, **l'ordre des qualificatifs de service est non significatif** : l'ARK <ark:/12145/ab.pdf.version2> est strictement équivalent à <ark:/12145/ab.version2.pdf>.

En raison de la diversité des formes que peuvent prendre des ARK qualifiés, **il est de la responsabilité de l'autorité d'adressage de définir des comportements par défaut si les qualificatifs fournis par l'utilisateur ne peuvent être traités**. Les sections 8.4.1 et 8.4.2 ci-dessous décrivent les principes de base de la résolution des qualificatifs et des exemples de comportements par défaut adaptés à différentes autorités d'adressage.

Les cas de comportements par défaut abordés sont les suivants :

- Qualificatif attendu mais non présent ;
- Qualificatif pas ou plus géré pour la ressource spécifiée ;
- Qualificatif inconnu.

8.4.1. Résolution des qualificatifs de granularité

Les qualificatifs de granularité sont à résoudre en premier lieu, du plus spécifique au plus général. Ainsi, si l'ARK <ark:/12345/ab/cd/ef> ne peut être servi à l'utilisateur, on tentera de lui fournir <ark:/12345/ab/cd> et, à défaut, <ark:/12345/ab>.

Exemple BnF Archives et manuscrits : <ark:/12148/cc87367c/cc87293v> ne peut être résolu ; l'autorité d'adressage fournit donc à défaut <ark:/12148/cc87367c>.

Exemple Gallica : un utilisateur demande la vue n° 690 d'un document numérique via l'ARK <ark:/12148/bpt6k204254b/f690>. L'autorité d'adressage, constatant qu'il n'existe pas de vue avec ce numéro, retourne la plus proche, c'est-à-dire la dernière vue du document.

8.4.2. Résolution des qualificatifs de service

Un ARK comportant plusieurs qualificatifs de service doit pouvoir être résolu quel que soit l'ordre dans lequel ces qualificatifs se présentent.

Les qualificatifs de service, s'il s'en trouve plusieurs, doivent donc être classés dans l'ordre alphabétique. On considère généralement, à moins qu'un comportement par défaut plus pertinent soit défini par l'autorité d'adressage, qu'**un qualificatif non géré ou inconnu doit être ignoré**.

Exemple : le qualificatif « epub » ne peut être résolu pour la ressource <ark:/12148/cc87367c>. L'ARK <ark:/12148/cc87367c.epub> renvoie donc le même contenu que l'ARK <ark:/12148/cc87367c>.

8.4.3. Redirection vers une URL d'accès

La spécification ARK fournit des recommandations pour la construction d'URI HTTP déréférencables construites de manière à donner les garanties maximales de pérennité. Néanmoins, certaines autorités d'adressage peuvent ne pas souhaiter remplacer leurs URL d'accès existantes par des URL incluant un ARK. Bien que cette pratique ne soit pas recommandée¹¹, elle n'est pas exclue par la spécification. Dans ce cas, la mise en place de tables de redirection de l'URL contenant l'ARK vers l'URL d'accès est nécessaire.

¹⁰ On notera que cette formule diffère de celle affichée actuellement par Gallica et le Catalogue général (« ARK malformé, caractère de contrôle incorrect ») car elle est plus explicite : l'ARK que l'utilisateur a entré est globalement incorrect ; l'erreur ne concerne pas forcément le caractère de contrôle seul.

¹¹ Notamment parce que les internautes ont pour habitude de copier l'URL présente dans la barre d'adresse de leur navigateur plutôt que de chercher dans la page une mention de permalien.



Exemple : <http://data.bnf.fr/ark:/12148/cb16459435n> redirige vers http://data.bnf.fr/16459435/brutti_sporchi_e_cattivi_film/ par le biais d'une réponse HTTP « 303 ».

Il est dans ce cas souhaitable de proposer à l'utilisateur un moyen de copier l'identifiant ARK afin de le citer.

Exemples : bouton « Adresse permanente de la notice » en haut à droite des notices du catalogue Médailles et Antiques (<http://medaillesetantiques.bnf.fr/ark:/12148/c33gbf0zz>), champ « Permalien » en bas des pages de data.bnf.fr (<http://data.bnf.fr/ark:/12148/cb16459435n>).



9. Annexe informative : liste des codes HTTP recommandés par cas d'usage

Code HTTP	Intitulé	Cas d'usage
200	« OK »	La ressource est servie à l'utilisateur telle qu'elle est attendue.
300	« Multiple choices »	La ressource a été scindée et plusieurs ressources alternatives sont disponibles.
301	« Moved Permanently »	La ressource a été remplacée et son équivalent est fourni à l'utilisateur.
302	« Found »	L'autorité d'adressage ne peut résoudre cet ARK et le transmet à un résolveur.
303	« See other »	La ressource identifiée par un ARK est accessible via une autre URL d'accès.
400	« Bad request »	L'ARK entré est erroné.
403	« Forbidden »	La ressource n'est pas accessible pour des raisons légales.
410	« Gone »	La ressource a été supprimée.