

Journée Systèmes et Données - 4 décembre 2020

Titre : **Le projet FNE**

Anila Angjeli (BnF) et Benjamin Bober (Abes), chef(fe)s de projet FNE

Toutes les questions ont été posées via le chat lors de la journée. Certaines ont été posées en direct par des modératrices. Celles qui n'ont pu être traitées en direct ont été transmises aux intervenants pour une réponse par écrit.

Source de référence pour l'information sur le projet FNE : <https://www.transition-bibliographique.fr/fne/fichier-national-entites/>

Réponses apportées par écrit après la journée :

- A-t-on un calendrier pour la phase 2 (ouverture aux producteurs tiers) ? Et si c'est post 2023, que faire en attendant ? Continuer avec les autorités BnF, IdRef, etc. comme actuellement ?

Réponse : Tant que la réalisation du FNE est en phase de démarrage (mise en place de l'infrastructure, alimentation avec de vraies données, etc.) il est difficile de se projeter au-delà du premier déploiement BnF/Abes.

Entretemps, oui, il faudra continuer à travailler avec les autorités de la BnF et de l'Abes, comme actuellement. Les services en charge de la diffusion des métadonnées dans les deux agences, en collaboration avec le groupe Systèmes et données de la TB, veillent à ce que ces questions soient prises en compte dans les instructions du FNE.

- Comment le FNE pourra-t-il prendre en compte et gérer les besoins divergents des différents producteurs en cas d'élargissement ? Notamment aux archivistes et aux données relatives aux collectivités administratives.

Réponse : Dans la gouvernance du FNE l'instance Conseil scientifique comprend des représentants de communautés de producteurs de données susceptibles de rejoindre le FNE. Une de ses missions est de remonter les différents besoins de ces communautés afin qu'elles soient prises en compte et que des instructions soient programmées.

Par ailleurs, le FNE s'appuie sur la norme RDA-Fr, dont les groupes de travail sont ouverts au-delà de la communauté des bibliothèques stricto sensu (ex.: des archivistes siègent au sein du groupe Collectivités) et dont la publication est soumise à enquête publique dans le cadre de la Commission nationale 46 de l'AFNOR.

- Peut-on trouver quelque part la composition du Conseil scientifique ?

Réponse : Elle sera diffusée prochainement sur les pages FNE du site TB, de même que la composition des autres instances de la gouvernance du FNE. La consolidation de la gouvernance a demandé pas mal de travail jusqu'à cet automne. Ce qui explique que l'équipe FNE n'était pas encore en mesure de donner publiquement des détails de cette gouvernance.

- Comment expliquer aux non-spécialistes le pourquoi du FNE ?

Réponse en deux points :

1. Les acteurs du FNE travaillent avec le groupe Formation de la TB pour faciliter la compréhension du FNE, au travers des formations dispensées aux professionnels et grâce au site web transition-bibliographique qui se veut accessible à toute personne intéressée.
 2. Tentative d'explication en quelques lignes :
 - les données d'autorité ont prouvé leur utilité, mais aujourd'hui chacun en crée dans son environnement - duplication du travail, des transactions dans tous les sens, de l'expertise dispersée...
 - il est temps de passer à un autre stade - mutualiser les efforts, identifier et décrire une entité une fois pour toutes dans une seule base centrale - l'utiliser autant de fois que nécessaire. Bénéfices attendus :
 - Intellectuels / Réseau de connaissances / Échanges entre professionnels
 - Economiques
 - Flux de données
 - L'initiative FNE réunit plusieurs conditions pour mettre en œuvre cette vision :
 - Un cadre conceptuel (modèle IFLA LRM) avec le corps des règles le RDA-FR
 - Une volonté de la part des acteurs majeurs de la TB - les deux agences bibliographique
 - Le cadre légal le permet - licence Etalab
 - Les technologies le permettent
-
- Les autorités toutes construites Concept-Lieu-Temps dans Rameau, pourront-elles exister dans le FNE sans être déconstruites ?
 - Il y a des autorités toutes construites concept-lieux-temps. Pourront-elles être dans le FNE?
 - Pour rebondir sur la dernière question de la matinée, les autorités toutes construites Concept-lieu-temps Rameau pourront-elles exister dans le FNE sans être déconstruites ?
 - Comment seront traitées les entités concepts, lieux, temps dans le FNE, qui sont plus éloignées des données traditionnelles du catalogage que les entités œuvres, expressions, ou mêmes agents ?

Réponse : Le FNE ne fera que permettre la mise en œuvre de la TB. Les décisions sur les contours des entités relevant du langage Rameau sont prises par le sous-groupe Concept-Lieu-Temps (CLT) du groupe Normalisation de la TB et seront mises en œuvre par le Centre National Rameau (CNR) qui est l'instance de gestion intellectuelle du langage Rameau.

Le FNE, quant à lui, se doit de créer les conditions pour gérer ces entités (identification univoque de chacune des entités, possibilité de les décrire de manière structurée et granulaire, établissement des liens entre les entités, etc.) : les instructions sur les entités CLT seront menées en étroite collaboration avec le CNR et le sous-groupe CLT de la TB.

Pour les Lieux, il y aura transfert des notices Noms géographiques qui seront fusionnées avec les notices Rameau Noms géographiques.

- Comment vont se faire les fusions BnF/IDRef sans perte d'information (+ choix de la forme autorisée si différente, etc.) ?

Réponse : La question des fusions est délicate et complexe. Un groupe spécifique Alignements-Fusions réunissant des agents de l'ABES et de la BnF a été mis en place pour trouver des solutions à ces questions. Tout cela est encore à l'étude et devrait se préciser en 2021. Les instructions combinent plusieurs axes :

1. Alignements : définir les critères, les méthodes, expérimenter avec des outils, évaluer la qualité des alignements possibles, etc. Pour simplifier, beaucoup de travail en amont pour s'assurer que de part et d'autre il s'agit de la même entité pour qu'ensuite on puisse acter la décision de les fusionner. Des alignements existent déjà entre certaines données de la BnF et de l'ABES. Ils ont été élaborés soit dans data.bnf.fr ou dans ID-Ref, avec récupération des identifiants, soit dans le cadre de projets internationaux (par exemple VIAF). Cet état de l'existant fournit déjà une importante base de travail.
2. La fusion sans perte des données est une des exigences majeures pour le FNE. Pour gérer les éventuelles divergences, plusieurs pistes peuvent être envisagées : priorité des données venant d'une institution ou de l'autre en fonction du corpus auquel elles appartiennent, redondance de certaines zones pour permettre l'affichage de toutes les données, en signalant la provenance de la donnée... Les règles sont à définir. La BnF et l'Abes élaboreront conjointement une politique de fusion fondée sur la connaissance des caractéristiques des données de part et d'autre et des enseignements issus des instructions.
3. Sur le choix des points d'accès autorisés, c'est le niveau d'application du code RDA-FR qui constitue la référence.

Réponses apportées au cours de l'intervention :

- Comment pourrions-nous tous concrètement nous aligner ? Y aura-t-il des sortes de notices d'autorité, œuvres, expressions, que l'on ira récupérer dans le FNE ? Car le but n'est pas que nous fassions chacun dans notre coin une notice d'œuvre ou d'expression - à la limite seul l'item serait à décrire, le reste est à récupérer.

Réponse : La réponse est contenue dans la question ! Bien qu'on n'emploie plus le terme "notices", oui, de fait il y aura des rassemblements d'informations concernant chacune des entités, et ce dans le respect du modèle IFLA LRM et du code RDA-FR. Les données seront exposées et, de ce fait, pourront être récupérées.

Concernant la fin de la question "- à la limite seul l'item serait à décrire, le reste est à récupérer" - ce serait l'idéal, mais pour des raisons liées aux missions de chacune des agences et autres institutions, les circuits et les processus de travail en place, la décision a été prise de ne pas inclure les Manifestations dans le FNE.

- Est-ce que ce projet sera étendu aux bibliothèques spécialisées hors Sudoc ou BnF, par exemple les grandes bibliothèques musicales nationales ?

Réponse : page 5 du PDF de présentation. Oui bien sûr, ces bibliothèques seront la cible de la deuxième phase de déploiement.

- Est-ce que cette plateforme sera également consultable par les institutions étrangères (Belgique, Luxembourg etc. ?)

Réponse : Le FNE, tout au moins dans un premier temps, est le projet d'une base de production. Il n'y aura pas d'interface publique de consultation. Par contre les données seront exposées dans le respect des standards, via des API, sur le web de données, au moyen des requêtes SPARQL, etc. Donc elles pourront être "vues" par n'importe qui et récupérées. Rappelons que les données du FNE seront des données ouvertes, sous licence Etalab.

Mais encore une fois, ne pas confondre ici données ouvertes et interface publique de consultation.

- Peut-on déjà récupérer des entités du FNE ?

Réponse : non, le FNE n'existe pas pour le moment.

(A noter qu'une partie des entités que l'on retrouvera à terme dans le FNE seront constituées à partir des données d'autorité existantes et d'ores et déjà récupérables via Idref : <https://abes.fr/reseaux-idref-orcid/reutiliser-les-donnees-idref/formats-de-donnees-disponibles/> et les outils de la BnF : <https://api.bnf.fr/>)

- Et pourra-t-on contribuer, par exemple, avec des spécificités locales (personnes, lieux...) ?

Réponse : Oui, c'est le but. La valeur du FNE se situera également dans la Longue Traîne, elle permettra la mise en lumière et la mise à disposition de tous de données spécifiques à des fonds locaux. Il est, toutefois, trop tôt pour dire avec précision comment la contribution va pouvoir se faire.

- Quel est le lien entre le FNE et VIAF ?

Réponse : VIAF agrège des données (notices d'autorité) de sources diverses. Donc il ne fait qu'utiliser, consommer, les données des autres. On ne peut pas produire (créer / modifier / gérer) dans VIAF. Il n'y a que les algorithmes qui travaillent avec la matière première qui sont les données d'autorité qui lui sont fournies par ses membres. Ces algorithmes "clustérisent", donc alignent et rassemblent en des "noeuds" les notices identifiées comme portant sur une même entité. Les données sont juxtaposées, et pas fusionnées (c'est pour cela qu'on parle de clusters). Ensuite ces rassemblements sont publiés via l'interface publique de consultation et exposés dans le web de données.

L'objectif du FNE, quant à lui, est de créer un environnement de co-production. On y créera des données. Ces données pourront y être modifiées, fusionnées, et subir toutes les opérations qu'on mène dans une base de données. Il y aura de l'activité humaine, bien qu'on mettra en œuvre des outils automatiques également. Il y aura une politique de qualité des données, des règles de travail, en somme, de la gouvernance des données.

Autre différence : VIAF est un outil international, alors que le FNE est un outil national ou à terme francophone, avec comme objectif de mobiliser l'expertise, le savoir-faire local partout où il est.

Les données du FNE auront en tout cas vocation à alimenter à terme VIAF au moment de la mise en production du FNE.

- Pourquoi ne pas faire cataloguer tous les catalogueurs dans le FNE au lieu d'avoir plusieurs outils (BnF avec Noemi/ Abes avec WinIBW et IdRef)

Réponse : Les spécificités de production des deux réseaux, l'écosystème des logiciels de bibliothèques et de leur implémentation constituent un existant complexe dont il faut tenir compte. Les deux agences et leurs tutelles ont fait un choix pragmatique de mettre en place, dans un calendrier maîtrisable, un outil de production commun qui permettra à l'ensemble des catalogueurs de produire conjointement une part importante des entités du modèle et rendra ainsi effective l'adoption de nouvelles règles communes de catalogage par des communautés professionnelles diverses, construites sur RDA-FR. : c'est déjà un projet ambitieux en soi.

- Comment seront récupérées les données via WinIBW ? Synchronisation immédiate ou différée (tous les 15 jours comme les notices Bnf actuellement) ? Récupération comme une base externe ?

Réponse : A préciser au cours du projet. C'est un axe à instruire par le groupe de coordination technique. Les évolutions techniques permettent d'envisager une fréquence beaucoup plus élevée, si ce n'est même du temps réel.

- Est-ce que dans mon catalogue public je pourrai afficher les notices du FNE ?

Réponse : la version 1 du FNE devrait contenir la majeure partie des données d'autorité actuelles de la BnF et de l'Abes (les Oeuvres et les Expressions viendront dans un deuxième temps). Si votre catalogue gère et affiche déjà ces données d'autorité, il paraît logique qu'il puisse afficher les données du FNE dont la diffusion sera assurée par les agences via leurs circuits de diffusion ou que vous pourrez récupérer depuis les API du FNE.

- Les fonctions des Correspondants Autorités de l'Abes vont-elles donc être fortement modifiées ?

Réponse : Probablement. La question est traitée par le groupe "Cas d'usages" du FNE.

- Quel sera le format de production ?

Réponse : L'instruction est en cours.

- Travailler dans le FNE permettrait que les catalogueurs de la Bnf et de l'Abes ne produisent pas des doublons d'entités ? Ou alors détection à l'entrée du FNE quand synchronisation des 2 flux (Abes et BnF) de doublons possibles ?

Réponse : Ne pas produire des doublons est une des exigences de la qualité des données dans le FNE. Mais les moyens pour y arriver sont à l'étude. L'élimination des doublons est déjà une priorité dans les processus qualité de la BnF et de l'ABES dans leurs fichiers respectifs.

- Comment envisage-t-on que la qualité des données soit vérifiée ? Y aura-t-il un groupe d'experts propre au FNE qui s'y consacrera ?

Réponse : C'est une des questions majeures, et complexes, du FNE. Le traitement est en cours au niveau de la gouvernance du FNE. Soulignons que le FNE se dotera d'une charte qualité, avec comme premier principe la mise en commun des politiques et processus qualité existants. Pour ce faire, le souhait est de réinvestir dans le FNE les compétences et l'expertise actuelles, ainsi que d'utiliser les outils actuels de contrôle qualité en œuvre à l'Abes et la BnF.

- Qui s'occupe de la partie logiciels et programmation : service en interne ou externalisé ?

Réponse : Dans la gouvernance du FNE le groupe de Coordination technique, constitué de spécialistes Abes et BnF traite ces questions. Manifestement on devra avoir recours aussi à des prestataires pour certains développements spécifiques (sur-couches à wikibase ou la structure de wikibase même)

- Qu'appellez-vous le modèle IFLA LRM ?

Réponse : on préfère vous renvoyer vers le site Transition bibliographique https://www.transition-bibliographique.fr/17-06-12_publication_ifla_lrm/

- Qu'est-ce que l'ESR ?

Réponse : Sigle pour l'Enseignement Supérieur et la Recherche

- Quels seront les types d'identifiants ? Vont-ils corrélés avec ceux existants ?

Réponse : La politique des identifiants constitue un axe à part d'instruction dans le FNE. Le principe même d'une base centrale veut que les entités qui y sont stockées et gérées soient équipées d'identifiants propres à la base. Mais il y aura forcément gestion des liens avec les identifiants BnF et Abes, Ark et IdRef/PPN. Les identifiants internationaux actuellement utilisés par l'Abes et la BnF (ex. ISNI, ORCID, ARK, Q Wikidata, ...) seront également maintenus au sein du FNE pour assurer une interopérabilité internationale maximale des entités.

- Est-ce pertinent de travailler avec IdRef pour nos autorités (Suisse romande) dans un futur très proche ? =>

Réponse : L'ensemble des producteurs qui participent aujourd'hui aux réseaux existants, ou qui les intégreront avant la mise en place du FNE, ont vocation à produire dans le FNE.